L'ECONOMIA DEI DATI

Tendenze di mercato e prospettive di policy

Roma, Gennaio 2018

Con il contributo scientifico di

Università Bocconi

ASK
Art, Science and Knowledge
Laboratorio di Economia
e gestione delle istituzioni
e delle iniziative artistiche
e culturali

Ringraziamenti e attribuzioni

L'economia dei dati: tendenze di mercato e prospettive di policy è uno studio realizzato da ITMedia Consulting con il contributo scientifico del Centro di Ricerca ASK Università Bocconi.

L'obiettivo del lavoro è di fornire spunti di riflessione che possano consentire a soggetti interessati a vario titolo (decision makers, imprese, autorità di regolazione e concorrenza, università centri di ricerca, studiosi e consulenti) di approfondire gli aspetti economici e di policy relativi all'economia digitale e in particolare alla cd. data driven economy, aggiornando modelli interpretativi e sviluppando chiavi di lettura anche in materia di concorrenza, regolazione e privacy, adeguati alla complessità e velocità dei mutamenti in atto.

Il lavoro è strutturato in due parti: la prima, relativa al mercato, si compone di 2 capitoli, la seconda, relativa ai temi di policy, si compone di tre capitoli. Augusto Preta è il responsabile del progetto e coordinatore editoriale della prima parte; Mariateresa Maggiolino ha coordinato editorialmente la seconda parte. Sebbene il lavoro sia frutto di uno sforzo comune, Giulia Filosa e Carlo Salatino hanno redatto la prima parte; nella seconda parte, con riferimento al capitolo 3, Mariateresa Maggiolino ha redatto i paragrafi 3.1, 3.3, 3.4.1, 3.4.2, 3.5, 3.6, 3.7 e 3.8, Giulia Filosa ha redatto il paragrafo 3.2 e Augusto Preta il paragrafo 3.4.3; con riferimento al capitolo 4, Giuseppe Colangelo ha coordinato il capitolo, mentre Augusto Preta ha redatto il paragrafo 4.1, Giuseppe Colangelo e Mariateresa Maggiolino il paragrafo 4.2, Paolo Biedanelli i paragrafi 4.3 e 4.4; Aura Bertoni e Carlo Rossi Chauvenet hanno redatto l'intero capitolo 5. Paolo Biedanelli e Giulia Filosa hanno curato l'editing finale.

I dati contenuti nel rapporto, aggiornati a ottobre 2017, provengono da fonti interne di ITMedia Consulting, dati pubblici, paper, saggi e altre pubblicazioni scientifiche, rapporti di analisti, bilanci e documentazione interna delle società, rapporti e relazioni annuali.

Per la sua realizzazione sono stati consultati e coinvolti, anche in workshop ad hoc, operatori ed esperti del settore: Paola Allamano (Co-founder & CEO WaterView); Luigi Capello (CEO LVenture Group); Andrea Conte (Manager Reply); Stefano Da Empoli (Presidente I-Com); Barbara Vecchi (Founder & CEO Hopenly) e Andrea Stazi di Google Italia, che ha garantito il sostegno economico al progetto. Si ringraziano inoltre Emilio Calvano (Università di Bologna) e Federico Ghezzi (Università Bocconi) per i suggerimenti e l'amichevole e prezioso contributo di idee.

Nell'esprimere a tutti sincero apprezzamento e gratitudine, va peraltro precisato che i contenuti e le opinioni espressi in questo lavoro impegnano esclusivamente ITMedia Consulting, insieme al Centro di Ricerca ASK Università Bocconi e gli autori, per le parti di rispettiva competenza.

Il presente studio (http://www.itmedia-consulting.com/DOCUMENTI/economiadeidati.pdf) viene distribuito sotto licenza creative commons (attribuzione – non commerciale – non opere derivate, CC BY-NC-ND), al fine di consentirne la più ampia diffusione e circolazione.

Indice degli argomenti

Key Findings 9

Executive summary.....12

PARTE I. Tendenze di mercato.....28

Introduzione.....28

Capitolo 1. L'economia digitale 31

1.1 La quarta rivoluzione industriale e la società connessa 31

1.1.1. Cloud computing 38

1.1.2. Internet of Things 45

1.1.3. Intelligent Automation: AI, AugI, AR e VR 52

1.2. Industry 4.0: nuove prospettive economiche, impatti sul mercato e sulle imprese.....61

1.3. Overview sugli sviluppi di politica industriale 71

1.3.1 Lo scenario europeo.....74

1.3.2 Le iniziative italiane: Piano Nazionale Industria 4.0 e lancio delle sperimentazioni 5G ..88

Capitolo 2. L'economia dei dati 93

2.1 Il ruolo dei dati nell'ecosistema digitale.....93

2.1.1 La catena del valore.....98

2.1.2 L'economia dei dati e gli impatti sulle imprese.....100

2.1.3 I dati nel nuovo ecosistema digitale.....105

2.1.4 I mercati per la commercializzazione dei dati digitali 107

2.1.5 *Data driven innovation*.....111

2.1.5.1 La *blockchain*.....118

2.1.5.2 Ulteriori ambiti di applicazione e criticità.....122

2.2 Dove risiede il valore dei dati 130

2.2.1 Come massimizzare il valore dei dati 130

2.2.1.1 La raccolta e la condivisione dei dati all'interno dell'impresa: i data lake 139

2.2.2. L'analisi dei dati tramite algoritmi 144

2.2.2.1 *Data mining* 145

2.2.2.2 *Machine Learning* 149

2.2.2.3 Le applicazioni del *Machine learning* e lo *statistical learning* 156

2.2.2.4 *Reccomender System* 165

2.2.2.5 Visual Analytics & Data Visualization.....175

PARTE II. Prospettive di policy 183

Introduzione 183

Capitolo 3. Le questioni di diritto antitrust 185

3.1 Lo spirito della disciplina antitrust europea e nazionale (cenni) 185

3.2 Lo stato dell'arte 187

3.3 I big data e la necessità di procedere caso per caso 194

3.4 I big data e il potere di mercato 196

3.4.1 Il potere di mercato nei mercati a valle di prodotti e servizi offerti usando big data 197

3.4.2 Barriere all'ingresso nei mercati a valle di prodotti e servizi offerti usando big data 198

3.4.3 I mercati digitali, le piattaforme *multisided* e i prezzi zero 202

3.4.3.1 I mercati a più versanti 204

3.4.3.2 Prezzi zero 208

3.4.3.3 *Data driven digital markets* 210

3.5 I mercati della generazione e raccolta dei dati digitali: l'accesso ai big data 212

3.5.1 Accesso esclusivo ai dati digitali, risorse essenziali e potere di mercato 215

3.5.2 Le barriere all'accesso ai big data 218

3.5.3 Condotte imprenditoriali che accrescono i costi all'accesso: le concentrazioni 219

3.5.4 (Segue) Condotte unilaterali: la raccolta di dati come attività economica 227

3.5.5 (Segue) Condotte unilaterali di sfruttamento 229

3.5.6 (Segue) Condotte unilaterali che escludono i concorrenti dall'accesso ai dati digitali 233

3.6 L'uso anticompetitivo dei big data 234

3.6.1 I big data e la collusione 235

3.6.1.1 La nozione di intesa (cenni) 235

3.6.1.2 Gli algoritmi e le condizioni strutturali che favoriscono la collusione 235

3.6.1.3 Gli algoritmi e i comportamenti delle imprese 237

3.6.1.4 Il caso particolare delle piattaforme della sharing economy 242

3.6.1.5 Possibili soluzioni rispetto alla collusione algoritmica 244

3.6.2 I big data e i prezzi personalizzati 245

3.6.3 I big data e il rifiuto a contrarre 252

3.6.3.1 Le difficili licenze obbligatorie 254

3.7 I big data: tra privacy e antitrust 256

3.8 I big data: tra consumer protection e antitrust (cenni) 262

Capitolo 4. La regolamentazione economica e di settore 264
4.1 Big data e regolazione delle comunicazioni 264
4.1.1 Le comunicazioni elettroniche 264
4.1.2 Il pluralismo informativo 266
4.2 I big data tra monopoli naturali, esternalità positive e asimmetrie informative 271
4.3 Dei consumatori più consapevoli del valore dei dati 275
4.3.1 La disclosure regulation 277
4.4 Una regolazione complementare al diritto antitrust 280
4.4.1 Switching cost e data portability 280
4.4.2 L'apertura imposta dei dati 284
4.4.3 L'apertura dei dati posseduti dalle pubbliche amministrazioni 288
Capitolo 5. Privacy: economia dei dati e protezione dei dati personali 291
5.1 La tutela dei dati personali nel nuovo eco-sistema digitale 291
5.2 Le principali novità introdotte dal GDPR 293
5.2.1 Estensione dell'ambito territoriale 294
5.2.2 Il diritto individuale alla protezione dei dati personali 295
5.2.3 Il trasferimento dei dati verso Paesi terzi o Organizzazioni internazionali 297
5.2.4 Le Autorità di controllo 299
5.2.5 Rimedi e sanzioni 300
5.2.6 Gli obblighi a carico del Titolare (e del Responsabile) del trattamento 301
5.3 I principi che regolano il trattamento dei dati personali 304
5.4 I nuovi strumenti a tutela della sicurezza delle informazioni: Accountability, Data protection by design e by default, DPIA e DPO 306
5.5 I nuovi diritti dell'interessato: il diritto all'oblio e la data portability 314
5.6 La condivisione dei dati e gli accordi di data sharing nella c.d. "Membership economy" 317
5.7 Big Data e Regolamento Privacy 322
Bibliografia 325

Indice dei grafici e delle tabelle

Figura 1. Cronologia delle rivoluzioni industriali 31
Figura 2. Modello della Grande onda 34
Figura 3. La crescita del traffico internet 35
Figura 4. Le stime della crescita digitale 36
Figura 5. La crescita esponenziale dell'universo digitale 36
Figura 6. Impatto relativo di internet su diverse industry 37
Figura 7. Modelli di servizio Cloud 39
Figura 8. Principali player di mercato Cloud 40
Figura 9. Catena del valore dei servizi Cloud 41
Figura 10. Evoluzione del Cloud verso architetture edge/fog computing e servizi abilitati 42
Figura 11. Previsioni sulla spesa globale dei servizi Public Cloud 44
Figura 12. Market Share dei Servizi Cloud (2016) 44
Figura 13. Requisiti 5G 46
Figura 14. Soluzioni di connettività per IoT 47
Figura 15. Previsioni sul numero di dispositivi IoT per il 2020 47
Figura 16. Dispositivi IoT installati per settore dal 2013 al 2020 (in milioni) 48
Figura 17. Il modello connected cars 49
Figura 18. Architettura logica e applicazioni dell'Internet of Things 50
Figura 19. Catena del valore per l'IoT 51
Figura 20. Valore aggiunto complessivo dell'IoT entro il 2020 52
Figura 21. Il ciclo dei processi cognitivi 53
Figura 22. Evoluzioni nell'ambito dell'Intelligenza Artificiale 55
Figura 23. Investimenti in ambito AI 55
Figura 24. Esempio di Rete Neurale Artificiale 56
Figura 25. Abilitatori tecnologici 57
Figura 26. Gartner Hype Cycle for emerging technologies, 2016 59
Figura 27. Crescita di Pokemon Go 60
Figura 28. Crescita AR 60
Figura 29. L'industy 4.0 62
Figura 30. I nuovi modelli di business nella digital economy 63
Figura 31. Applicazione del modello sharing economy 64
Figura 32. Digital Supply Chain Management 65
Figura 33. La crescita delle imprese digitali 66
Figura 34. I vantaggi della digital trasformation 67
Figura 35. Industria 4.0, il quadro delle tecnologie digitali 68
Figura 36. Il passaggio da un approccio product-oriented ad un approccio platform-focused 69
Figura 37. I livelli di implementazione della digital economy a livello mondiale 71
Figura 38. Quota del digitale del PIL paese per paese (2015 and 2020) 72
Figura 39.Dimensioni e tendenze dell'economia dei dati dell'UE 73
Figura 40. Le iniziative della Commissione in chiave di digitalizzazione dell'industria europea 76
Figura 41. I principali programmi industria 4.0 nel mondo 79

Figura 42. Digital Single Market Strategy for Europe 82
Figura 43. Le direttrici del Piano nazionale industria 4.0 e gli obiettivi 88
Figura 44. Cabina di regia industria 4.0 91
Figura 45. Il nuovo ecosistema digitale 93
Figura 46. Le origini del fenomeno "Big Data" 94
Figura 47. Le caratteristiche dei big data: le quattro "V" 95
Figura 48. La "Veracity" 96
Figura 49. Schema del ciclo di vita dei dati e valore aggiunto nella catena del valore dei dati 98
Figura 50. La catena del valore dei dati 99
Figura 51. Il processo di scoperta della conoscenza 100
Figura 52. I risultati dello studio di McKinsey sull'acquisizione del valore dei dati per settore 103
Figura 53. Unified Capabilities Information Management 105
Figura 54. Semplificazione del modello funzionale per l'analisi dei dati 113
Figura 55. La scoperta e lo sfruttamento economico di nuovi dati 114
Figura 56. Due generazione di analisi dei dati a confronto 116
Figura 57. Esempi applicativi nel settore financial services & insurance 117
Figura 58. Le potenziali applicazioni della blockchain per settore 122
Figura 59. L'intelligenza artificiale nella marketing automation. L'esempio di Netflix 123
Figura 60. I vantaggi della "personalizzazione" nei diversi settori 130
Figura 61. Classificazione dei Big Data 132
Figura 62. I 5 elementi alla base del successo dei dati e della loro trasformazione tramite analitycs 133
Figura 63. La creazione di valore all'interno del data ecosystem 134
Figura 64. Il processo di Enterprise Data Management 141
Figura 65. I vantaggi dei data lake 143
Figura 66. Knowledge Discovery in Database (KDD) 145
Figura 67. Il processo di Machine Learning 150
Figura 68. Tipologie di analisi combinabili con il ML 155
Figura 69. Le tre tipologie di problemi risolvibile tramite l'applicazione del ML 156
Figura 70. Applicazioni pratiche del Machine learning 161
Figura 71. Un caso pratico: il servizio Google AdWords 163
Figura 72. Le righe dell'homepage di Netflix 170
Figura 73. Personalized Video Ranker (PVR) 171
Figura 74. Search algorithms 172
Figura 75. Gli effetti della personalizzazione dell'algoritmo PVR 173
Figura 76. Il processo di ricerca e sviluppo di Netflix di tipo iterativo-incrementale 174
Figura 77. L'importanza della Visual Analytics 176
Figura 78. Le aree di ricerca della Visual Analytics 177
Figura 79. Il processo di Visual Analytics 178
Figura 80. The disciplines that contribute towards visual analytics 179
Figura 83. Data driven digital markets 210

Tabella 1. L'evoluzione di internet 32
Tabella 2. Artificial Intelligence e Augmented Intelligence a confronto 58
Tabella 3. SWOT analisi industria 4.0 70
Tabella 4. Abbinamento Business problem e tipologia di dato 131
Tabella 5. Alcuni esempi applicativi 147
Tabella 6. Vantaggi e limiti delle tecniche di data mining 148
Tabella 7. Modelli di filtering 168

Key Findings

1. Siamo all'alba di una nuova era dello sviluppo economico e sociale, frutto di un continuo, inarrestabile processo di innovazione, che ha caratterizzato nel corso degli ultimi due decenni lo sviluppo di internet, attraverso **la diffusione della *digital economy* in tutti i settori, non più soltanto legati all'ICT.**
2. Tecnologie, piattaforme e sistemi innovativi come ***Cloud Computing, Internet of Things, Big Data & Analytics, Blockchain, Artificial Intelligence, Augmented Reality & Virtual Reality, Advanced robotics & 3D printing*** **e** ***5G*** costituiscono i nuovi strumenti abilitanti della digital economy che, grazie alla sua diffusione pervasiva in tutti i settori, promette, dunque, di dar luogo a una nuova era dello sviluppo economico e sociale e più in generale a una nuova e più evoluta fase dell'esistenza umana.
3. Uno dei fattori chiave che guidano lo sviluppo dell'economia digitale è rappresentato dunque dai dati: una risorsa che rende sempre più importante per le imprese dotarsi di strumenti analitici adeguati, ma che richiede soprattutto un importante cambiamento nella struttura economica e culturale delle società digitali.
4. La letteratura prevalente, tende a concentrarsi sull'aspetto quantitativo e dunque concernente il ***volume*** dei dati (da qui il termine big data). Tuttavia, il volume non è l'unica caratteristica importante. La ***velocità*** alla quale i dati vengono generati e resi accessibili è parimenti impressionante. Il fenomeno dei big data è definito anche in ragione della capacità di analizzare una ***varietà*** di insiemi di dati non strutturati provenienti da fonti diverse
5. La digitalizzazione pervasiva pone, dunque, i dati quale elemento catalizzatore dell'innovazione. Si rende necessario quindi che **le organizzazioni si trasformino puntando a diventare sempre più *data driven***, I dati possono incidere positivamente sul miglioramento dei prodotti o servizi di un'impresa; possono consentire alle imprese di sfruttare nuove opportunità di business; infine possono anche essere utilizzati per orientare meglio i potenziali clienti, fornendo loro servizi o prodotti personalizzati. I benefici non sono però circoscritti unicamente alle imprese ma anche ai cittadini a alle amministrazioni pubbliche
6. In ambito **Antitrust i**l fatto che una impresa disponga di big data non implica necessariamente che quell'impresa disponga di **potere nel mercato** o nei mercati della generazione e raccolta di quei dati. Analogamente, non si può stabilire un nesso causale necessariamente vero tra l'esistenza di big data in capo a una impresa e il fatto che essa possa esercitare potere nei mercati che si collocano a valle della generazione e raccolta dei dati.
7. Le **barriere all'accesso** ai dati digitali non rappresentano una particolarità dei mercati della generazione e raccolta dei dati digitali. Con riguardo alle operazioni di **concentrazione**, l'esperienza accumulata sinora ha quasi sempre escluso l'eventualità che imprese, pur depositarie di big data, riuscissero, aggregando i propri patrimoni di informazioni, a impedire che i propri concorrenti avessero accesso a dati sufficienti per produrre beni e servizi in concorrenza con i loro. Con riferimento alla **collusione algoritmica**, a destare maggiore preoccupazione è l'ipotesi che riproduce l'interdipendenza oligopolistica, ossia l'ipotesi per

cui ogni operatore arrivi a praticare un prezzo collusivo, sebbene abbia sviluppato il proprio algoritmo in modo autonomo e indipendente. Con riferimento ai **prezzi personalizzati**, a suscitare maggiori perplessità è il caso dei prezzi discriminatori che non si rivelino né escludenti né anticompetitivi, bensì di sfruttamento. Riguardo al **rifiuto di condividere i dati**, **il soddisfacimento dei requisiti della dottrina dell'essential facilities è problematico** ed ancor più complessi sono i risvolti pratici dell'applicazione della suddetta dottrina.

8. Il rapporto tra il **diritto antitrust e la disciplina a tutela dei dati personali** può essere declinato lungo diverse direttrici: a) le nuove regole in materia di portabilità dei dati aumenteranno la concorrenza tra le imprese che analizzano quei dati personali per produrre conoscenza e valore; b) sarebbe del tutto fisiologico un intervento antitrust che fosse teso a sanzionare eventuali strategie capaci di ridurre la quantità e la qualità di beni e servizi; c) il caso Facebook avviato dal Bundeskartellamt suggerisce che le condotte per effetto delle quali le imprese raccolgono dati personali possano risultare, oltre che lesive della privacy, anche abusive in quanto inique ed ingiustificatamente gravose; d) il diritto antitrust e la disciplina della privacy potrebbero trovare un punto di convergenza laddove una riduzione di quest'ultima venga qualificata come una riduzione della qualità dei beni e servizi offerti.
9. L'Autorità per le Garanzie nelle Comunicazioni, è fra le autorità di **regolazione**, quella che guarda con particolare interesse ai big data dovendo garantire un accesso al mercato secondo criteri di **obiettività, trasparenza, non discriminazione e proporzionalità per i servizi di sua competenza.** Fra i compiti dell'AGCOM vi è anche quello di promuovere **il pluralismo informativo**, un tema che, per le sue connotazioni, legate ad aspetti che vanno oltre gli ambiti strettamente economici, da sempre ha creato problemi applicativi. La dimensione individuale e sociale (diritto all'informazione) richiede una maggiore attenzione all'evoluzione in atto (fake news, post verità), ma sempre allo scopo di migliorare una strumentazione fin qui carente in tema di pluralismo informativo.
10. Essendo l'azione del regolatore più "invasiva", si ritiene necessaria solamente in circostanze particolari, quelle in cui si reputa necessario un suo intervento, al fine di ripristinare le corrette dinamiche del mercato. Ciò si verifica quando si rileva la presenza di **fallimenti di mercato**.
11. In particolare c'è chi ritiene appropriato l'intervento della regolazione in caso di **asimmetria informativa** legata anche all'**ignoranza/inconsapevolezza dei consumatori "fonti di dati" e dei consumatori "destinatari di informazioni e beni"**. Un approccio basato **sull'obbligo di informare** si è rivelato però fin qui inefficace. Rendere **i consumatori più consapevoli** del valore dei propri dati è il presupposto da cui partire per superare le asimmetrie informative e valutare il trade-off tra **disclosure** e protezione.
12. Di segno positivo sono le misure di recente implementate in Europa per favorire la ***data portability***: in assenza di rilevanti costi di spostamento per via della portabilità dei dati, a parità di condizioni economiche (a parità di prezzi nulli) le piattaforme saranno incentivate a far leva sull'altro parametro di competizione (la qualità). **Poco auspicabili risulterebbero, invece, interventi sulle politiche di prezzo delle piattaforme**. La non neutralità della struttura dei prezzi applicati dalla piattaforma è funzionale alla necessità di alimentare

entrambi i versanti, sicché un **eventuale intervento regolatorio teso a proibire i prezzi nulli** rischierebbe di minare non solo il *business model* ma l'esistenza stessa di una piattaforma.

13. Si è detto in precedenza di quanto **sia inapplicabile la dottrina dell'essential facility** al fenomeno dei big data. Imporre a soggetti privati di rendere disponibile la loro banca dati o i loro algoritmi di elaborazione gratuitamente potrebbe essere una misura compromettente gli equilibri del mercato, **che finirebbe per produrre un danno oltre che per le imprese anche per il benessere dei cittadini/consumatori.**
14. **Privacy.** L'immagine virtuale di ognuno di noi è la stratificazione di dati ceduti involontariamente, di dati che siamo costretti a cedere se vogliamo usufruire di un servizio e di dati che mettiamo in rete volontariamente per definire la nostra identità pubblica.
15. **Il nuovo Regolamento 2016/679 (noto come GDPR**, General Data Protection Regulation) riconosce la natura transnazionale della data economy adottando un approccio estensivo al proprio ambito di applicazione territoriale e rafforzando gli strumenti per garantire l'adeguatezza del trasferimento extra-UE dei dati personali.
16. La liceità del trattamento e conservazione dei big data resta un aspetto non definito dal GDPR che rimane aperto all'interpretazione giuridica, alle prese con il **difficile bilanciamento tra principi di trasparenza, *data minimization* ed esigenze di sviluppo di nuovi prodotti e servizi**.

Executive summary

Parte I – Mercato

L'economia digitale

Siamo all'alba di una nuova era dello sviluppo economico e sociale, frutto di un continuo, inarrestabile processo di innovazione, che ha caratterizzato nel corso degli ultimi due decenni lo sviluppo di internet, attraverso **la diffusione della *digital economy* in tutti i settori, non più soltanto legati all'ICT.**

A tal proposito si può parlare di un vero e proprio ecosistema, caratterizzato dai seguenti elementi: una pluralità di soggetti che concorre alla creazione e offerta dei servizi digitali; la trasformazione dei prodotti anche più comuni in prodotti "intelligenti" e, dunque, in servizi; l'ampliamento dei contatti e delle interrelazioni tra i diversi mercati, che innescano un processo di allargamento del perimetro, con l'inclusione di nuove attività e nuovi ambiti merceologici.

Questa trasformazione è il risultato di un processo che ha inizio con la prima diffusione del world wide web (1995-2005) e procede con le reti sociali e il cosiddetto web 2.0 (2005-2015).

Così come queste trasformazioni sono state guidate dall'innovazione tecnologica che ha introdotto nuovi paradigmi di connettività, di elaborazione e di utilizzo delle informazioni, allo stesso modo la prossima trasformazione – la terza generazione di internet - sarà guidata da cambiamenti tecnologici che comporteranno significativi mutamenti nella domanda e nelle abitudini di consumo, così come nell'economia, nella struttura delle imprese e nei modelli di business.

Innanzitutto, cresce esponenzialmente la disponibilità di dati digitali: se, in passato, del mondo si offrivano rappresentazioni analogiche che potevano poi essere trasformate in sequenze di uno e zero, oggi è aumentata in modo significativo la capacità di istituzioni, imprese e individui di "datificare" la realtà, ossia di disporre di rappresentazioni degli eventi e dei comportamenti umani che nascono già digitali.

Inoltre, coerentemente con la legge di Moore, l'innovazione tecnologica ha determinato e continua a determinare, con una velocità sempre maggiore, a costi sempre minori e con una diffusione pervasiva, l'evoluzione esponenziale delle prestazioni nonché la semplicità delle modalità di accesso alle tre operazioni base sulle informazioni e sui dati: elaborazione, in termini di potenza di calcolo dei microprocessori; memorizzazione, in termini di disponibilità e capacità; trasmissione, in termini di capacità e velocità grazie allo sviluppo delle reti broadband e ultrabroadband.

Tecnologie, piattaforme e sistemi innovativi come ***Cloud Computing, Internet of Things, Big Data & Analytics, Blockchain, Artificial Intelligence, Augmented Reality & Virtual Reality, Advanced robotics & 3D printing*** **e** ***5G*** costituiscono i nuovi strumenti abilitanti della digital economy che, grazie alla sua diffusione pervasiva in tutti i settori, promette, dunque, di dar luogo a una nuova era

dello sviluppo economico e sociale e più in generale a una nuova e più evoluta fase dell'esistenza umana.

Proprio attraverso l'introduzione e l'implementazione di tali tecnologie e la loro diffusione generalizzata, l'esperienza umana sta vivendo un profondo cambiamento, caratterizzato da un alto grado di efficienza dovuto all'impiego dei dati, al ritorno a paradigmi produttivi locali, a sistemi di rete *cloud* integrati capaci di mettere in relazione persone e oggetti, luoghi fisici e virtuali.

Questa nuova fase, comunemente definita "**quarta rivoluzione industriale**", è caratterizzata da una serie di innovazioni tecnologiche destinate a modificare totalmente le nostre abitudini, toccando ogni aspetto della nostra vita, migliorandone la qualità.

Uno dei fattori chiave che guidano lo sviluppo dell'economia digitale è rappresentato dunque dai dati: una risorsa che rende sempre più importante per le imprese dotarsi di strumenti analitici adeguati, ma che richiede soprattutto un importante cambiamento nella struttura economica e culturale delle società digitali.

Mentre la terza rivoluzione industriale o Industria 3.0 seguendo l'odierna classificazione, si è incentrata sulla automazione della singola macchina e del singolo processo, l'Industria 4.0 si concentrerà sulla **digitalizzazione *end-to-end*** di tutti i beni materiali e della loro integrazione all'interno dell'ecosistema digitale.

Questo processo passerà attraverso diverse fasi:

(i) Digitalizzazione e integrazione della catena del valore orizzontale e verticale. L'industria 4.0 porterà alla digitalizzazione e all'integrazione dei processi verticali passando attraverso l'intera organizzazione, dallo sviluppo all'acquisto del prodotto, mediante sistemi di produzione, logistica e vendita.

(ii) Digitalizzazione dell'offerta di prodotti/servizi. La digitalizzazione dei prodotti comprenderà l'espansione della linea dei prodotti esistenti, per esempio tramite l'aggiunta di sensori intelligenti o di dispositivi di comunicazione che possono essere utilizzati con gli strumenti d'analisi dei dati, nonché attraverso la creazione di nuovi prodotti digitali che si concentreranno su soluzioni completamente integrate. Attraverso nuovi metodi di raccolta e analisi dei dati, le aziende saranno in grado di generare dati relativi al prodotto atti a perfezionare il prodotto stesso, per soddisfare al meglio le crescenti esigenze dei consumatori finali e quindi assoceranno ai loro prodotti un ventaglio di servizi;

(iii) Modelli di business digitali e "accesso" diretto al consumatore. I principali leader di mercato, stanno già ampliando la loro offerta, essendo in grado di fornire soluzioni digitali importanti come, ad esempio, servizi basati completamente su dati e soluzioni fortemente integrate con le piattaforme. Modelli di business digitali dirompenti saranno spesso incentrati sullo sviluppo di ulteriori ricavi digitali e sull'ottimizzazione delle interrelazioni con il consumatore finale.

L'economia dei dati

In questo scenario, il **dato rappresenta, pertanto, il motore della trasformazione nella *digital economy*** ed è per questo che viene considerato "il lubrificante" della nuova economia.

In particolare, negli ultimi anni, più che parlare di semplici o singoli dati, si è soliti riferirsi a un concetto molto più complesso e ampio, riconducibile al termine "***smart data"*** o più comunemente "***big data***", proprio in virtù delle ingenti quantità di dati disponibili all'interno del nuovo ecosistema digitale, prodotti ad alta velocità e provenienti da una moltitudine di fonti, la cui gestione e analisi richiedono nuovi, più performanti e intelligenti strumenti in termini di processori e algoritmi.

Il primo elemento di diversità dei dati rispetto al passato è rappresentato dalla loro *quantità*: database ordinari non sono più in grado di gestire il numero sempre crescente di dati, conseguentemente si è reso necessario sviluppare nuovi modelli di database capaci di memorizzare, classificare ed elaborare grandi moli di dati a velocità supersoniche.

Il secondo aspetto riguarda le modalità di analisi dei dati che, come detto, sono decisamente cambiate, si parla infatti oggi di tecniche di analisi innovative ed avanzate: "*predictive analytics*", "*data mining*" e "*data science*", tecniche queste alla portata di tutti, grazie anche alla presenza di tecnologie *open source* destinate proprio alla loro applicazione. Attraverso tali modalità di analisi è possibile personalizzare una ricerca, un prodotto/servizio, la pubblicità, rivoluzionando completamente le strategie di marketing e in generale di business.

La letteratura prevalente, tende a concentrarsi sull'aspetto quantitativo e dunque concernente il *volume* dei dati che, per alcuni autori, rappresenta uno dei principali problemi connessi a questo genere di fenomeno, in ragione del fatto che risulta particolarmente complesso stare al passo con la loro costante crescita esponenziale. Secondo le stime, infatti, ogni giorno ne vengono prodotti circa 2,5 exabyte, un numero elevatissimo.

Tuttavia, il volume non è l'unica caratteristica importante. La *velocità* alla quale i dati vengono generati e resi accessibili è parimenti impressionante. Per citare alcuni esempi, secondo alcune stime del 2016, relativamente ai dati generati sui social, Facebook genera circa 10 miliardi di "like" ogni giorno da parte di 1,09 miliardi di utenti, mentre su Istagram vengono condivise più di 95 milioni di immagini al giorno con una media giornaliera di 4,2 miliardi di "like" e infine su YouTube si registrano oltre 400 ore di contenuti caricati dagli utenti ogni minuto di ogni singola giornata.

La velocità di sviluppo dei dati tuttavia desta non poche perplessità soprattutto in merito alla loro gestione, in particolare quando tale velocità supera di gran lunga quella di elaborazione. Infatti, perché i dati riescano a garantire le migliori performance, devono essere analizzati a una velocità pari a quella, fulminea, che caratterizza il flusso di informazioni nei sistemi di Enterprise Information Management. Nel giro di microsecondi occorre decidere se vale la pena acquisire un determinato dato, valutandone la rilevanza dopo averlo confrontato o combinato con le altre informazioni disponibili.

In alcuni casi, il fenomeno dei *big data* è definito anche in ragione della capacità di analizzare una *varietà* di insiemi di dati non strutturati provenienti da fonti diverse come: registri web, social media, smartphone, sensori e transazioni finanziarie. Ciò richiede la capacità di collegare insiemi di dati diversi, oltre alla capacità di estrarre informazioni da un insieme di dati destrutturati. Il carattere della varietà dei dati fa riferimento essenzialmente alla loro struttura: i dati strutturati entrano nei sistemi di Enterprise Information Management già contrassegnati con un tag e sono facili da smistare; oggi, tuttavia, i dati sono per la maggior parte non strutturati, informazioni casuali, difficili da analizzare e gestire.

Queste tre proprietà appena illustrate - volume, velocità e varietà - sono considerate le tre principali caratteristiche dei *big data* e sono comunemente indicate come le tre V. Si tratta di proprietà "tecniche" che sostanzialmente dipendono dalle capacità di memorizzazione dei dati e dalle tecnologie di elaborazione. Lo sviluppo dell'analisi del fenomeno ha portato a considerare oltre le tre V sopracitate, una quarta V rappresentata dalla caratteristica del *valore*, che dipende fortemente dal crescente potenziale economico e dalla valenza sociale che a loro volta attribuiscono ai dati la qualità di "nuovi fattori di produzione".

Quanto alle modalità di classificazione esistenti, è possibile distinguere diversi approcci di catalogazione dei dati, con riferimento alla:

(i) *Tipologia di informazione*: i dati possono essere classificati in base alle diverse tipologie di informazioni che sono in grado di fornire e che, a loro volta, possono avere un diverso grado di utilità per una data impresa. I dati possono, ad esempio, fornire informazioni su individui, sui loro comportamenti, le loro preferenze e la loro posizione geografica; su entità economiche rispetto al fatturato realizzato e al numero delle transazioni commerciali effettuate; ovvero su oggetti, rispetto ad esempio alla posizione in tempo reale o alla velocità di una macchina.

(ii) *Struttura del dato*: i dati possono essere differenziati a seconda se siano strutturati o meno. Da questa peculiarità dipende la possibilità di estrarne valore economico e la scelta della modalità necessaria per farlo. Per dati strutturati, ossia dati numerici o di natura finanziaria (dunque tabelle, record, documentazioni di *office automation*, dati M2M, ecc.) è necessario un modello di estrazione che definisca una serie di parametri quali: il settore di riferimento; la tipologia di dati che interessano quel dato settore; e infine la relazione esistente tra i dati stessi. Un esempio di dati strutturati potrebbe essere rappresentato da una base di dati di un consumatore che contiene informazioni relative al nome, cognome, indirizzo, età, recapito telefonico, ecc. I dati strutturati possono essere più facilmente elaborati e utilizzati per scopi commerciali rispetto a dati destrutturati, soprattutto con riferimento ai metodi tradizionali di elaborazione.

(iii) *Modalità di raccolta*: i dati possono essere raccolti attraverso diverse modalità e a seconda della modalità adottata si registra un diverso impatto sul loro valore economico. I dati sono perlopiù forniti volontariamente ed attivamente dagli utenti. I dati possono essere anche raccolti semplicemente attraverso fonti liberamente disponibili su internet (*open source*) o

attraverso lo studio del comportamento dell'utente, pur in mancanza di alcuna informazione personale sull'utente stesso. Esempi del primo tipo sono i cd. motori di ricerca generici che si basano sulla raccolta sistematica e l'elaborazione di ogni pagina del web a loro disposizione - una tecnica nota anche come "*crawling*". La seconda modalità è anche quella più diffusa. Infatti, un gran numero di imprese tracciano il percorso fatto dall'utente da una pagina web all'altra - e talvolta anche all'interno di una singola pagina (tecnicamente è possibile monitorare quale parte di una pagina web viene realmente vista dall'utente) - attraverso tecniche diverse. Ad oggi, queste tecniche sono state ulteriormente sviluppate al fine di incrementare lo studio sul comportamento dell'utente su più dispositivi mobili e non solo sul web. Alcune delle tecniche utilizzate per questo scopo possono però essere facilmente superate dall'utente senza particolari sforzi, altre invece sono difficilmente evitabili. I dati così generati possono anche (tecnicamente) essere facilmente combinati con dati rilasciati volontariamente ad una società, allo scopo di creare profili utente altamente differenziati. Tuttavia, le leggi sulla privacy possono concretamente limitare il campo di applicazione di questo genere di approccio. Da ultimo, i dati possono essere generati deducendo nuove informazioni utilizzando dati già esistenti.

La digitalizzazione pervasiva pone, dunque, i dati quale elemento catalizzatore dell'innovazione. Si rende necessario quindi che **le organizzazioni si trasformino puntando a diventare sempre più *data driven***, adottando una strategia olistica di gestione delle informazioni che includa e integri nuove tipologie di dati e di tecniche di gestione e di analisi.

I dati possono incidere positivamente sul miglioramento dei prodotti o servizi di un'impresa; possono consentire alle imprese di sfruttare nuove opportunità di business; infine possono anche essere utilizzati per orientare meglio i potenziali clienti, fornendo loro servizi o prodotti personalizzati.

Non è ancora del tutto chiaro quali saranno, complessivamente, gli effetti economici generati dal fenomeno *data-driven*, tuttavia dagli studi ad oggi disponibili emerge chiaramente l'esplicarsi di un effetto positivo sulla produttività delle imprese e delle pubbliche amministrazioni.

Complessivamente, questi risultati prevedono un aumento della produttività pari a circa 5-10% per quei soggetti economici che decidono di adottare un approccio *data-driven*.

I benefici non sono però circoscritti unicamente alle imprese: un uso efficace del Big Data Analytics da parte dei governi e delle amministrazioni pubbliche a livello europeo, potrebbe ridurre del 15%-20% i costi amministrativi grazie a una maggiore efficienza derivante dalla maggiore trasparenza, maggiore gettito fiscale (grazie, ad esempio, all'erogazione di servizi personalizzati) e un minor rischio di frodi o errori, legati più direttamente alla capacità di analisi dei dati.

Tuttavia secondo il McKinsey Global Institute, è necessario investire fortemente sulla formazione di figure professionali adatte e più in generale far aumentare la consapevolezza sul valore dei dati nelle organizzazioni. Uno dei più grandi ostacoli alla diffusione dei big data è rappresentato, infatti, dalla scarsità di figure professionali adeguate per cui una delle principali sfide dei big data è di natura umana e consiste da una parte nel diffondere la "cultura del dato" all'interno di tutta

l'organizzazione e dall'altra nello sviluppare sempre più la capacità di porre le giuste domande, riconoscere i pattern descrittivi di determinati fenomeni, predire i comportamenti e effettuare assunzioni basate su fatti ed evidenze supportate dai dati.

Nel contesto in esame, un ruolo particolarmente rilevante, soprattutto per il loro ampio potenziale di applicazione di tipo ***disruptive***, è rivestito dalla *Blockchain* e dall'*Artificial Intelligence*.

La tecnologia **Blockchain**, inizialmente utilizzata per la gestione della moneta virtuale Bitcoin, è un database distribuito che sfrutta la tecnologia *peer-to-peer,* consentendo lo scambio di valute, la loro tracciabilità e soprattutto di conoscere l'identità dei soggetti che effettuano le relative operazioni. In altre parole è il libro contabile in cui sono registrate tutte le transazioni fatte in Bitcoin dal 2009 ad oggi, e in cui si "registra" l'identità dei titolari di valuta.

Nello specifico, il protocollo *blockchain* è caratterizzato da una catena distribuita di informazioni tra più nodi di una rete (pubblica o privata) costituita da macchine di varia potenza e connesse tra di loro. Ogni nodo detiene una copia del registro, costituito da una serie di blocchi concatenati, all'interno dei quali sono contenute le informazioni degli scambi avvenuti tra coloro che partecipano alla rete. Il fatto che tali informazioni siano distribuite contemporaneamente, in maniera sincrona su tutti i nodi della rete, rende impossibile l'eliminazione dei dati. Ogni transazione effettuata in questo ambiente è soggetta ad una 'conferma', o 'consenso', espresso attraverso una validazione effettuata risolvendo un problema matematico. Tale conferma viene, dunque, eseguita da un algoritmo condiviso ed accettato nell'ecosistema della rete stessa.

Diversamente dalle piattaforme in capo alle banche, che sono di proprietà delle banche stesse e utilizzabili a livello centralizzato, la *blockchain* è una tecnologia "open" e "free" utilizzabile online da qualsiasi privato e organizzata come un maxi database distribuito. Chiunque può scaricarne una "copia", ossia accedere al sistema: le identità sono protette da crittografia intelligente. La *blockchain* è in grado di assicurare il corretto scambio di titoli e azioni, può sostituire un atto notarile e può garantire la bontà delle votazioni, ridisegnano il concetto di seggio elettorale, proprio perché ogni transazione viene sorvegliata da una rete di nodi che ne garantiscono la correttezza e ne possono mantenere l'anonimato.

Estrapolata dal suo contesto, questa innovativa tecnologia può essere utilizzata in tutti gli ambiti in cui è necessaria una relazione tra più persone o gruppi. Infatti, la *blockchain* non è solo bitcoin, anzi, è una tecnologia capace di avere un impatto rivoluzionario su diversi settori: sanità, tecnologia, finanza, *retail*, media, energia, fino alla pubblica amministrazione. Ciò che rende la *blockchain* uno strumento dalle molteplici applicazioni sono: affidabilità, trasparenza, solidità, irrevocabilità, digitale.

L'**Intelligenza Artificiale** si riferisce all'abilità di un computer di svolgere funzioni e ragionamenti tipici della mente umana. Questo approccio tecnologico sta emergendo come il principale catalizzatore che accelererà lo sviluppo di settori e applicazioni innovative come ***connected car*** **e** ***smart home***, connotandosi tra i diversi abilitatori tecnologici, come una sorta di super abilitatore dei servizi digitali. Nel suo aspetto puramente tecnologico, essa comprende la teoria e le tecniche per lo sviluppo di algoritmi che consentono alle macchine di mostrare un'abilità e/o attività intelligente,

almeno in domini specifici. Tra gli attuali e principali campi di applicazione vi sono gli **assistenti virtuali** (come ad esempio Alexa di Amazon, Cortana di Microsoft, Siri di Apple, Google Home, Google Assistant e Google Now nonché il più recente Bixby di Samsung) che offrono assistenza su richieste generiche, richieste contestualizzate per ottenere previsioni su viaggi, pianificazione attività, etc.

Un altro ambito di applicazione in crescita è quello relativo alle **Conversational Interfaces**, e in particolar modo ai **Chatbot**, considerati alla base di una potenziale rivoluzione del mercato enterprise grazie alle interessanti prospettive di risparmio di tempi e costi offerte dalle opportunità di automazione. Si tratta di applicazioni incorporate nei servizi di messaggistica, come Messenger di Facebook, capaci di guidare autonomamente gli utenti attraverso una serie di domande nello svolgimento di diversi compiti. Nello specifico, si prevede che queste applicazioni produrranno nuovo valore condiviso tra le complesse operazioni all'interno delle aziende e all'esterno in ambito B2B e B2C, spaziando su tre fonti:

(i) ***Automazione**, tramite la sostituzione tra macchine e persone nelle attività ripetitive e ridondanti;*

(ii) ***Semplificazione**, per generare livelli maggiori di adozione e usabilità dei sistemi e dei processi aziendali;*

(iii) ***Innovazione**, abilitando nuovi servizi basati sull'analisi automatica degli insights aziendali in modalità non sostenibili dalle persone.*

Il crescente interesse verso l'Intelligenza Artificiale è legato alla prospettiva di sostituzione degli esseri umani in vari ruoli e compiti. Tale prospettiva si basa sulla convinzione che man mano che saranno utilizzate nella risoluzione di problemi complessi, tali applicazioni evolveranno sempre più grazie all'apprendimento automatico dai dati e dagli input coinvolti.

Tuttavia, numerosi sono i dubbi circa le reali possibilità che queste applicazioni, orientate in tale ottica alla creazione della cd. "conoscenza perfetta", possano davvero sostituirsi alle persone. Per tale motivo, in letteratura si parla di un nuovo approccio, definito **Augmentend Intelligence (AugI)** orientato ad aumentare l'intelligenza e la conoscenza di un individuo al fine di supportarlo nel processo decisionale e nell'acquisizione di ulteriore conoscenza, senza che un sistema artificiale lo sostituisca.

Le applicazioni di AugI funzionano come un partner personalizzato che offre supporto presentando risposte, spunti e suggerimenti rilevanti basati sull'accesso a dati e informazioni esterne e soprattutto operando sotto la guida dell'utente che potrà dinamicamente valutare la rilevanza di quanto suggerito attraverso modalità di interazione sofisticate e al tempo stesso semplici finalizzate anche a "educare" la macchina/applicazione. Le risposte e i suggerimenti vengono determinati sulla base di algoritmi che calcolano e visualizzano la probabilità associata con differenti risposte potenziali aiutando a comprendere perché una data risposta è stata suggerita o meno. Inoltre, grazie all'apprendimento supervisionato dagli utenti, al crescere del loro utilizzo queste applicazioni diventeranno maggiormente precise e affidabili tendendo al comportamento ideale atteso per questo tipo di sistemi.

Parte 2 - Policy

Antitrust

La funzione industriale dei big data, ossia la loro capacità di produrre valore grazie alla formazione di nuova conoscenza aiuta chi svolge **l'analisi antitrust** a fissare un punto: i big data sono una risorsa che si colloca "a monte" delle diverse attività che un'impresa può svolgere servendosene. I big data hanno cioè una natura strumentale e la loro ragion d'essere deve essere quindi colta in termini relazionali, ossia guardando alle soluzioni che grazie ai big data vengono ideate. Pertanto, non solo **i big data sono input** tanto nell'ipotesi in cui un'impresa li impieghi per migliorare i propri processi, quanto nel caso in cui li usi per creare nuovi beni capaci di meglio soddisfare i consumatori, quanto ancora nella circostanza che li utilizzi per elaborare raccomandazioni, recensioni o altro. Di più, **le implicazioni e gli effetti competitivi dei big data devono analizzati caso per caso** in rapporto all'uso che le singole imprese ne fanno.

La necessità di procedere guardando alle specifiche circostanze di ogni scenario di mercato trova ad esempio conferma anche con riguardo al binomio big data-potere di mercato. In particolare, in relazione a detta questione ci si pone due interrogativi: per un verso, ci si chiede se i big data conferiscano potere di mercato o se, per lo meno, contribuiscano a schermarlo. Per altro verso, si discute del caso particolare delle *multi-sided media platforms*, perché ci si interroga sulle modalità da seguire per apprezzare il loro potere di mercato.

Sotto il primo profilo, **il fatto che una impresa disponga di big data non implica necessariamente che quell'impresa disponga di potere nel mercato o nei mercati della generazione e raccolta di quei dati**. Analogamente, non sembra che si possa stabilire un nesso causale necessariamente vero tra l'esistenza di big data in capo a una impresa e il fatto che essa possa esercitare potere nei mercati che si collocano a valle della generazione e raccolta dei dati, che si tratti di quelli della commercializzazione dei dati, o che si tratti di quelli dei beni e dei servizi ideati muovendo dalle informazioni estratte dai big data.

Con riferimento poi all'eventualità che i big data costituiscano delle barriere all'ingresso nei mercati a valle dei beni e servizi grazie ad essi migliorati, la dottrina si è divisa, dando luogo ad un dibattito assai vivace, ancorché spesso intrecciato a valutazioni non immediatamente legate ai soli dati, quanto più genericamente connesse alle tecnologie dell'informazione e, in particolare, alle piattaforme digitali multi-versante, anche quando queste non offrono prodotti e servizi a prezzo nullo. In particolare, anche sulla scorta di alcuni casi empirici, si è negata l'esistenza di un *first mover advantage*, esaltando il fenomeno del *multi-homing*, evidenziando come gli effetti di rete siano unidirezionali e ricordando come anche l'accumulo dei dati digitali soffra di rendimenti decrescenti. Così si è arrivati a negare l'ipotesi secondo la quale i mercati digitali si concentrerebbero, portando ad esempio casi di imprese che oggi hanno acquistato importanti quote di mercato senza essere state le *prime* ad operare nei relativi mercati digitali.

Il secondo profilo rilevante per il binomio big data-potere di mercato riguarda il caso delle ***multisided media platforms***, vale a dire imprese che raccolgono big data e offrono i loro beni non solo a prezzi

nulli, ma altresì impiegando modelli di business a più versanti. Questo è un caso "particolarissimo" non perché si voglia disconoscere il ruolo che queste imprese giocano nell'economia digitale, ma perché i problemi che attengono all'accertamento del loro potere di mercato non sono connessi alla detenzione dei big data, ma alla scelta di queste imprese di praticare prezzi nulli all'interno di quelli che, per giunta, sono modelli di business che aggregano due o più domande. Il punto di convergenza tra i fenomeni dei "big data", degli "zero-price goods" e delle "multi-sided platforms" è cioè di natura puramente empirica: capita che imprese come Google o Facebook offrano i loro servizi di searching e di social networking a prezzi nulli, finanziando questa loro offerta con i dati dei loro utenti e i ricavi provenienti dagli inserzionisti pubblicitari interessati all'attenzione di quegli stessi utenti. Nondimeno, nella realtà non solo si danno imprese che detengono sì big data, ma che offrono i loro prodotti a prezzi positivi, attraverso modelli di business che si rivolgono ad una sola domanda. Di più, nella realtà esistono imprese che praticano prezzi nulli senza raccogliere dati o senza ricorrere a modelli multi-versante, così come esistono delle imprese multi-versante che applicano prezzi positivi in ogni mercato in cui operano. A titolo di esempio, si consideri che i prezzi nulli caratterizzano strategie di pricing come il cross-selling e il bundling che non prevedono né la raccolta di dati, né l'esistenza di una seconda domanda interdipendente. Allo stesso modo, i sistemi di carte di credito sono sì multi-mercato, ma non applicano prezzi nulli. La possibilità delle multi-sided platforms di offrire servizi gratuiti deriva infatti dalla loro capacità di internalizzare gli effetti di rete diretti e indiretti e le divergenti elasticità della domanda dei loro due /o più gruppi di utenti.

Sotto il profilo antitrust, il prezzo zero introduce un elemento di ulteriore complessità rispetto alla natura multi-sided del mercato, dal momento che il metodo convenzionale di definizione dei mercati – il test del monopolista ipotetico o SSNIP test – si basa sui prezzi positivi e su un determinato aumento degli stessi (5-10%) per individuare i vincoli concorrenziali in un determinato mercato. Inoltre, il prezzo nullo fa sì che la qualità divenga un elemento cruciale nell'apprezzamento degli effetti dei comportamenti delle imprese, ancorché gli effetti sulla qualità non possano essere analizzati come gli effetti sui prezzi.

Ciò non toglie però che nei mercati finanziati dalla pubblicità e caratterizzati da una concorrenza per attenzione/base utenti, a parità di prezzo (nullo), la concorrenza sulla qualità mostra dinamiche simili a quelle della tradizionale concorrenza di prezzo: un aumento del potere di mercato (attenzione) potrebbe incentivare una riduzione della qualità (più pubblicità), sfruttando gli effetti di rete. Alla luce di ciò, ci si domanda se si può applicare la stessa logica dei "tradizionali" mercati (multiversante) a prezzo zero basati sull'attenzione ai nuovi mercati digitali basati sui dati.

Con riferimento al tema del**l'accessibilità dei dati digitali**, se i dati corrispondono alle registrazioni che un individuo fa dei fatti del mondo, allora pare possibile che quell'individuo (eventualmente un'impresa) sia l'unico ad avere accesso a dei dati quando egli risulti anche l'unico soggetto capace di assistere all'evento o di disporre della tecnologia necessaria a percepirlo. Nondimeno, come spiega la teoria della preclusione anticompetitiva, **l'esistenza di una esclusiva di fatto** – ma si supponga anche l'esistenza di una, per quanto remota, esclusiva di diritto – su dei dati digitali non dovrebbe portare a concludere né che quei dati siano, per ciò stesso, anche delle risorse importanti per la produzione dei prodotti/servizi a valle, né che le imprese che vantano quell'esclusiva siano *naturaliter* in posizione

dominante. Insomma, il fatto che un'impresa detenga il controllo esclusivo su alcuni dati non significa – o, *rectius*, non prova di per sé – che i suoi rivali non possano accedere ad altri input ugualmente capaci di soddisfare la loro domanda.

Circa poi le **barriere all'accesso ai dati digitali**, queste possono certo esistere, ma non rappresentano una particolarità dei mercati della generazione e raccolta dei dati digitali. Possiamo ad esempio verificare l'esistenza di barriere naturali, quando notiamo che la raccolta dei dati sulle precipitazioni di acqua piovana in una data regione dipende soltanto dall'installazione in quella stessa regione di una centralina di registrazione. Oppure, possiamo constatare l'esistenza di barriere strutturali, quando ad esempio ricordiamo che le imprese interessate a raccogliere dati digitali dovranno dotarsi delle necessarie infrastrutture e competenze e, in particolare, se vorranno agire tramite delle piattaforme online, dovranno considerare i network effects di cui godono le imprese loro rivali e già incumbent. Ancora, se un'unica impresa avesse il diritto esclusivo ad esercitare un'attività per effetto della quale raccoglie alcuni tipi di dato, allora il mercato di quei dati sarebbe difeso da barriere legali. Nondimeno il diritto antitrust non può agire contro le barriere naturali e strutturali che proteggono un particolare mercato o che generalmente proteggono certi tipi di mercato. L'indagine circa queste barriere serve al diritto antitrust a fini descrittivi, ossia per individuare i mercati rilevanti e apprezzare il potere di mercato che le imprese in essi già presenti detengono. Con riguardo poi ai comportamenti delle imprese che potrebbero rendere alcuni dati meno accessibili, l'attenzione viene innanzitutto attratta dalle concentrazioni. Al riguardo, l'esperienza accumulata sinora – esperienza che potrebbe modificarsi in altri contesti di mercato e avere ad esempio ad oggetto dati che anziché fungere da input, fungono da beni finali – ha quasi sempre escluso l'eventualità che imprese, pur depositarie di big data, riuscissero, aggregando i propri patrimoni di informazioni, a impedire che i propri concorrenti avessero accesso a dati sufficienti per produrre beni e servizi in concorrenza con i loro. Ad esempio, in *Facebook/WhatsApp* la Commissione ha riconosciuto la liceità dell'acquisizione, sottolineando come la presenza di numerosi e importanti operatori capaci di collezionare dati circa i consumatori avrebbe impedito a Facebook di avere il controllo del prezzo delle inserzioni pubblicitarie, anche nell'ipotesi in cui Facebook avesse utilizzato i dati di WhatsApp a fini pubblicitari. La Commissione, cioè, ha escluso che la concentrazione comportasse un danno al benessere del consumatore perché «there will continue to be a large amount of internet user data that are valuable for advertising purposes and that are not within Facebook's exclusive control».

Con riguardo poi all'uso dei big data, si discute di collusione algoritmica, prezzi personalizzati e rifiuto a condividerli. Sotto il primo profilo, posto che non pare possibile stabilire se l'uso di **algoritmi** chiamati a lavorare sui big data renderà i mercati più proni alla **collusione**, in virtù del fatto che gli algoritmi consistono in una serie di comandi necessari al raggiungimento di determinati fini le imprese possono congegnare gli stessi per perseguire i più disparati obiettivi o svolgere le più differenti funzioni. Di conseguenza, non è possibile – e non sarebbe neanche corretto – qualificare in termini antitrust la classe degli algoritmi in sé considerata. Occorre, piuttosto, operare dei distinguo. In particolare, a destare maggiore preoccupazione è l'ipotesi per cui ogni operatore arriva a praticare un prezzo collusivo, sebbene abbia sviluppato il proprio algoritmo in modo autonomo e indipendente.

Con riferimento ai **prezzi personalizzati**, a suscitare maggiori perplessità è il caso dei prezzi discriminatori che non si rivelino né escludenti né anticompetitivi, bensì di sfruttamento. Qualora questi prezzi fossero praticati da un'impresa in posizione dominante, infatti, si produrrebbe uno scenario in cui il benessere del consumatore è nullo sebbene l'efficienza sia massima.

In ultimo, riguardo al rifiuto di condivisione, è necessario valutare quanto il *framework* definito dalla casistica europea in tema di ***essential facilities*** (casi *Magill, IMS, Microsoft* e *Huawei*) sia applicabile al nuovo scenario rappresentato dai big data. Per giunta, anche laddove la dottrina dell'*essential facility* fosse applicabile, l'imposizione di un obbligo a contrarre porrebbe non pochi problemi in termini, ad esempio, di tipi di dati condivisi, tempi della condivisione e profili di privacy.

A quest'ultimo riguardo, **il rapporto tra il diritto antitrust e la disciplina a tutela dei dati personali** può essere declinato lungo diverse direttrici. In primo luogo, le nuove regole in materia di portabilità dei dati, le quali riguarderanno i dati personali volontariamente trasferiti e quelli tracciati, aumenteranno la concorrenza tra le imprese che analizzano quei dati personali per produrre conoscenza e valore. In secondo luogo, sarebbe del tutto fisiologico un intervento antitrust che, nei mercati per la commercializzazione di prodotti e servizi "privacy-friendly" o di tecnologie "privacy-enhancing", fosse teso a sanzionare eventuali strategie capaci di ridurre la quantità e la qualità di quei beni e di quelle tecnologie. Del resto nelle decisioni *Facebook/WhatsApp* e *Microsoft/LinkedIn* la Commissione ha osservato per la prima volta che i consumatori potrebbero vedere la privacy come un fattore significativo che incide sulla qualità dei beni e dei servizi disponibili su internet. In terzo luogo, come il caso *Facebook* avviato dal Bundeskartellamt suggerisce, è possibile che le condotte per effetto delle quali le imprese raccolgono dati personali risultino, oltre che lesive della privacy, anche inique ed ingiustificatamente gravose ai sensi dell'art. 102 (a) del Trattato. In quarto luogo, è altresì possibile che il diritto antitrust e la disciplina della privacy possano trovare un punto di convergenza laddove una riduzione di quest'ultima venga qualificata come una riduzione della qualità dei beni e servizi offerti e, in tal modo, come una riduzione del benessere dei consumatori.

Regolamentazione

Vi sono determinate circostanze in cui, per correggere situazioni potenzialmente dannose, viene considerato più opportuno un intervento ex-ante da parte del regolatore, rispetto a un'azione ex-post da parte dell'antitrust. In questi casi interviene la **regolazione economica** che essendo comunque, più "invasiva", si ritiene necessaria solamente in circostanze particolari, quelle in cui si reputa necessario un suo intervento, al fine di ripristinare le corrette dinamiche del mercato. Ciò si verifica quando si rileva la presenza di **fallimenti di mercato**, ovvero in tutti quei casi in cui il sistema economico non è in grado di raggiungere la piena efficienza in quanto i fattori impiegati non sono utilizzati in modo ottimale.

In generale l'azione del regolatore si esercita su un intero settore economico. In particolare in quei settori strategici, caratterizzati dalla presenza di reti (telecomunicazioni, trasporti, energia), la regolazione ha assunto un ruolo preponderante nel definire una serie di regole che garantissero

l'accesso a condizioni eque e non discriminatorie nei mercati a valle. In questo senso l'interrelazione sempre più ampia tra i settori economici spinge nell'economia digitale, per le caratteristiche approfondite nella prima parte del lavoro, **al superamento della tradizionale regolazione verticale, sector specific**, in favore di una orizzontale e sempre più convergente tra le diverse industries coinvolte, collegate dal tema della connettività e caratterizzate da sempre più ampie interrelazioni (smart cities, connected cars, IoT, AI, ecc..).

L'Autorità per le Garanzie nelle Comunicazioni, è fra le autorità di regolazione, quella che guarda con particolare interesse ai big data e ai loro effetti sulla concorrenza nei mercati. Fra le sue prioritarie funzioni infatti, ai sensi del Codice delle Comunicazioni elettroniche (d.lgs 1 agosto 2003, n. 259), c'è quella di **garantire un accesso al mercato secondo criteri di obiettività, trasparenza, non discriminazione e proporzionalità per i servizi di sua competenza**. Fra di essi rientrano tutti i servizi di comunicazioni fissi e mobili, i servizi media audiovisivi e i servizi postali. Ognuno di essi si inserisce all'interno di settori interessati allo sviluppo dei big data, in quanto caratterizzati da un elevato numero di consumatori che, essendo interconnessi tramite la piattaforma IP alle reti digitali, costituiscono non solo grandi utilizzatori dei servizi in rete, ma anche tra i principali fornitori di dati.

Fra i compiti dell'AGCOM vi è anche quello di promuovere **il pluralismo informativo**, un tema che, anche per le sue connotazioni, legate ad aspetti che vanno oltre gli ambiti strettamente economici, da sempre ha creato problemi applicativi. Oggi, in ragione dello sviluppo di nuove modalità di accesso all'informazione (contenuti online, social media) ci si chiede se non sia opportuno regolare il pluralismo informativo, non solo dal punto di vista dell'offerta, ma anche dal punto di vista della domanda. Se la dimensione economica (concorrenza e posizioni dominanti) pare più efficacemente affrontabile attraverso il modello antitrust caso per caso, quella legata alla dimensione individuale e sociale (diritto all'informazione) richiede certamente una maggiore attenzione all'evoluzione in atto (fake news, post verità), ma sempre allo scopo di migliorare una strumentazione carente in tema di pluralismo informativo, attraverso l'uso più efficiente delle tecnologie e dei big data.

In una chiave più direttamente correlata alla dimensione economica della regolazione, la scarsa competenza dei consumatori "fonti di dati" e dei consumatori "destinatari di informazioni e beni" determinerebbe **asimmetrie informative** a vantaggio delle imprese. Dinanzi alle circostanze ed ai fattori appena illustrati che consentirebbero alle imprese attive nei mercati per la generazione e raccolta dei big data di acquisire potere di mercato, il ruolo e **gli strumenti della regolazione differiscono da quelli del diritto della concorrenza**. Al contempo però, anche sul versante regolatorio che a differenza dell'antitrust permetterebbe un intervento sui mercati per scongiurarne i fallimenti, i margini di intervento (quantomeno quelli auspicabili) appaiono limitati. E' possibile certamente ipotizzare interventi tesi a mitigare le asimmetrie informative attraverso una puntuale definizione degli adempimenti richiesti per l'utilizzo dei dati personali degli utenti. In questo caso ci muoveremmo nel solco dell'approccio tradizionalmente seguito per le questioni relative alla privacy e, più in generale, al deficit informativo dei consumatori-utenti.

Tale approccio, tuttavia, negli anni si è dimostrato non particolarmente efficace: gli interventi con i quali si è aggravato l'obbligo di informare per le imprese hanno finito per tradursi in adempimenti prevalentemente formali, inadeguati a fronteggiare il tema sostanziale di fondo dell'adeguata

informazione del consumatore e delle scelte consapevoli compiute dagli stessi. Nell'ambito digitale, il quadro complessivo è ancor più desolante. La convinzione che il consenso possa rappresentare il principale elemento di legittimazione anche nel contesto digitale espone ogni intervento normativo all'esito tragico dello sforzo di Sisifo, specie se si considera che non si intravedono modelli di business alternativi rispetto a quelli che ricavano profitti dalla **profilazione degli utenti** e dalla **pubblicità.**

Di segno diverso sono le misure di recente implementate in Europa per favorire la mobilità degli utenti tra le piattaforme riducendo gli *switching costs* attraverso la garanzia della ***data portability* .** Una siffatta misura è tesa ad alimentare la competizione tra le piattaforme e potrebbe risultare funzionale a favorire una concorrenza sulla qualità dei servizi: in assenza di rilevanti costi di spostamento per via della portabilità dei dati, si deve, infatti, ipotizzare che, a parità di condizioni economiche (a parità di prezzi nulli) le piattaforme saranno incentivate a far leva sull'altro parametro di competizione (la qualità).

Poco auspicabili risulterebbero, invece, interventi sulle politiche di prezzo delle piattaforme. E' opportuno, infatti, ricordare che a caratterizzare i mercati in questione sono innanzitutto le interrelazioni e l'interdipendenza tra i versanti delle piattaforme, le quali devono portare «both sides on board», ossia intercettare un numero sufficiente di agenti economici su ciascuno dei versanti coinvolti in modo da assicurarsi una massa critica adeguata per alimentare gli effetti di rete indiretti. Nel caso, ad esempio, delle piattaforme che forniscono servizi di ricerca o di *social networking*, nel facilitare l'incontro tra utenti ed inserzionisti pubblicitari, seguendo il tradizionale modello degli *advertising-supported media,* mentre ai primi i servizi vengono offerti gratuitamente, ai secondi viene applicata una tariffa per gli spazi pubblicitari messi a disposizione per il *behavioural targeting.* Questa funzione obiettivo impone sovente alle piattaforme di praticare prezzi asimmetrici ai gruppi presenti sui diversi lati (*skewed pricing*), che non riflettono il costo effettivo del servizio offerto ad un particolare gruppo di utenti ma tengono conto delle differenze nell'elasticità della domanda. La non neutralità della struttura dei prezzi applicati dalla piattaforma è funzionale alla necessità di alimentare entrambi i versanti, sicché un **eventuale intervento regolatorio teso a proibire i prezzi nulli rischierebbe di minare non solo il *business model* ma l'esistenza stessa di una piattaforma**.

Rendere **i consumatori più consapevoli** del valore dei propri dati è il presupposto da cui partire per superare le asimmetrie informative e valutare il trade-off tra **disclosure** e protezione. Vari studi si sono interrogati sul tema, con conclusioni tendenzialmente problematiche. L'educazione del consumatore viene per lo più vista come un'attività che non è di per sé sufficiente a creare un bilanciamento ideale fra l'esigenza di proteggere i dati e la condivisione di informazioni. Complessivamente, tuttavia, sebbene riequilibrare la conoscenza che individui e imprese hanno circa il valore dei dati possa implicare costi e rischi, questo rimane uno dei principali obiettivi della regolazione. Programmi educativi a tale proposito potrebbero quindi aiutare in tal senso, anche se dovrebbero essere sempre gestiti con attenzione.

Si è detto in precedenza di quanto **sia inapplicabile la dottrina dell'essential facility** al fenomeno dei big data. Laddove, dunque, si volesse ovviare a questo limite si potrebbe configurare un'altra forma di intervento dei pubblici poteri nell'economia, ossia si potrebbero prevedere **degli obblighi di apertura dei dati.** Il tema dell'obbligo generalizzato a condividere i dati è stato più volte affrontato

dagli studiosi che hanno rimarcato come bisognerebbe sempre tenere presente **il trade-off fra i benefici** di una maggiore apertura dei dati, derivanti da una maggiore facilità nello svolgere analisi o acquisire notizie, e **i danni** che si rischiano di provocare rinunciando alla loro segretezza. Imporre a soggetti privati di rendere disponibile la loro banca dati o i loro algoritmi di elaborazione gratuitamente potrebbe essere una misura compromettente gli equilibri del mercato, **che finirebbe per produrre un danno oltre che per le imprese anche per il benesseri dei cittadini/consumatori.**

Privacy

In questa realtà sempre più digitale, nella quale i dati rappresentano la linfa della cosiddetta ***knowledge economy,*** occorre poi chiedersi quale ruolo tenda ad assumere la specifica categoria dei **dati personali**. Le attività di raccolta ed elaborazione dei dati inerenti agli utenti dei servizi online è infatti impiegata in un numero crescente di settori, sia privati che pubblici. Dalle assicurazioni all'ambito sanitario, dal comparto bancario alla pubblicità, per citarne solamente alcuni, l'attività di trattamento dei dati personali appare sempre più consistente.

D'altro canto, questa stessa realtà digitale ha profondi effetti anche sul piano sociale, definibili in termini di *knowledge society,* in cui la rappresentazione di ogni individuo è legata sempre più alla sua persona virtuale, vale a dire all'immagine derivante dalle informazioni che lo riguardano e che sono reperibili in rete. L'immagine virtuale di ognuno di noi è la stratificazione di dati ceduti involontariamente, senza rendercene conto, di dati che siamo costretti a cedere se vogliamo usufruire di un servizio, e di dati che mettiamo in rete volontariamente per definire la nostra identità pubblica.

Il Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016 (noto come GDPR, *General Data Protection Regulation*), relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati, abroga la precedente direttiva 95/46/CE (cosiddetta Direttiva madre), presenta sia elementi di continuità che elementi di discontinuità rispetto alla Direttiva madre. La principale differenza è di tipo giuridico e sistemico: la Direttiva definiva un mero sistema di armonizzazione attraverso un obbligo di risultato a carico degli Stati membri, lasciando loro spazio quanto ai mezzi all'interno delle singole discipline nazionali, mentre il nuovo Regolamento, già in vigore e che sarà applicato dal 25 maggio 2018, è immediatamente obbligatorio in tutti i suoi elementi e per tutti i paesi membri. Questa scelta trova la sua ragion d'essere nel carattere transnazionale sia delle trasformazioni economiche della *knowledge economy,* sia dell'interesse pubblico alla tutela dei dati personali che riguarda la *knowledge society*.

Dalla diversa natura giuridica delle due normative derivano anche diversi criteri per stabilire la legge applicabile al trattamento dei dati personali. Laddove la Direttiva doveva individuare la legge applicabile anche all'interno dell'Unione, oltre che al di fuori dei suoi confini, avendo una semplice funzione di armonizzare le legislazioni nazionali, il Regolamento pone quali uniche questioni l'applicabilità della normativa europea al caso di specie e, in caso affermativo, l'Autorità di controllo tra quelle dei diversi Stati membri chiamata a garantirne l'applicazione. A questo proposito, è bene

sottolineare la portata estensiva dell'ambito di applicazione territoriale del nuovo Regolamento: esso si applica per il solo fatto che il trattamento dei dati avvenga "nel contesto delle attività di uno stabilimento" nell'Unione, e, dunque, a prescindere dal fatto, come attualmente previsto, che vengano impiegati strumenti situati nel territorio dell'Unione, ad esempio server, siano situati in ambito extra-UE. Del tutto innovativa, rispetto alla della Direttiva madre, è anche la disposizione del paragrafo n. 2 dell'articolo 3, che, nel caso di Titolari o Responsabili del trattamento non stabiliti nel territorio dell'Unione, fonda l'applicabilità delle nuove norme sulla base di due differenti criteri: **1) l'offerta di beni o la prestazione di servizi a persone interessate nell'Unione; e 2) il monitoraggio del comportamento di tali soggetti all'interno dell'Unione, a fini di 'profilazione' e dunque per adottare decisioni che li riguardano, analizzarne o prevederne le preferenze ed i comportamenti.**

Il delicato tema del trasferimento dei dati fuori dal territorio dell'Unione è una parte fondamentale del nuovo Regolamento. In primo luogo, viene meno il requisito dell'autorizzazione nazionale: il trasferimento verso un Paese terzo ritenuto "adeguato" dalla Commissione Europea, oppure sulla base di clausole contrattuali modello, o di norme vincolanti d'impresa (Binding Corporate Rules) approvate da un'Autorità di controllo, potrà avere inizio senza attendere l'autorizzazione nazionale del Garante — a differenza di quanto attualmente previsto. Tuttavia, l'autorizzazione del Garante sarà ancora necessaria se un titolare desidera utilizzare clausole contrattuali *ad-hoc* (cioè non riconosciute come adeguate tramite decisione della Commissione Europea) oppure accordi amministrativi stipulati tra autorità pubbliche — una delle altre novità introdotte dal Regolamento.

Il Legislatore europeo, a tutela del soggetto Interessato, ha riaffermato, oltre alla necessità del consenso, ulteriori diritti conoscitivi (informativa ed accesso) e di controllo (limitazione al trattamento, diritto di revoca, opposizione, rettifica, cancellazione e portabilità dei dati).

Il titolare, ai sensi delle nuove norme europee, **può trattare i dati anche al di fuori di un preventivo consenso dell'interessato**, purché il trattamento sia effettuato nell'ambito della tutela di propri interessi legittimi e questi non prevalgano sugli interessi e le libertà dell'interessato (c.47), ovvero nel caso in cui l'utilizzo dei dati per finalità diverse da quelle previste possa considerarsi compatibile con le finalità originarie (secondo un giudizio di compatibilità effettuato dal titolare sulla scorta di una serie definita di condizioni art. 6.(4)).

Il rapporto tra consenso, prestazione contrattuale e finalità rimane quindi con il nuovo regolamento il paradigma sul quale **misurare la liceità dei trattamenti effettuati dal titolare**, ma si apre anche all'utilizzo dei dati anche per altre finalità, ove sia possibile individuare una base giuridica adeguata all'ulteriore trattamento e non vi siano rischi per la tutela dei diritti dell'interessato.

In questo senso, uno dei problemi che rimangono irrisolti anche a seguito dell'applicazione del GDPR è il difficile **bilanciamento tra le esigenze di trasparenza (c. 39), di specificità del consenso e di minimizzazione del trattamento**, di cui agli artt. 4 e 5 del GDPR, con le oggettive difficoltà dei titolari del trattamento nella redazione delle proprie privacy policy a fronte di finalità di trattamento non sempre prevedibili e *data retention* difficili da stimare.

La continua evoluzione dei modelli di business e la necessità di accumulare informazioni dei propri utenti per rispondere, o meglio prevedere, le loro esigenze rendono assai complesso l'adeguamento dei titolari alle previsioni contenute nel GDPR, sia per i dati di cui già dispongono, sia nel design di nuovi prodotti o servizi che, sempre più spesso, si fondano sull'utilizzo massivo di Big Data.

Profili di tensione con la normativa privacy si possono infatti riscontrare e permangono in tutte le fasi del trattamento di questa tipologia di dati. Che tipo di informativa, ed in rapporto a quale specifica finalità, l'interessato dovrà essere trasparentemente avvertito della raccolta di dati apparentemente non-personali? La semplice menzione della finalità del trattamento normalmente utilizza di"attività di analisi e statistiche" può considerarsi troppo generica alla luce delle nuove norme? Ancora una volta agli interpreti, in primis ai giuristi e poi alle corti superiori nazionali ed europee, viene affidato il difficile compito di bilanciamento dei diritti inviolabili dell' individuo con le esigenze di trattare volumi di dati sempre più grandi e per finalità che non possono essere sempre preventivamente definite.

* * *

In conclusione, il contesto in esame qui analizzato in tutte le sue complesse sfaccettature, legato al mondo dei dati e dell'economia digitale, si presenta mutevole, affascinante e aperto a ogni tipo di sfide.

In questo senso il lavoro è diviso in due parti: la prima, incentrata sul mercato, si compone di due capitoli: il primo capitolo descrive caratteristiche e modelli di sviluppo dell'economia digitale; il secondo capitolo è dedicato all'economia dei dati e alla loro centralità nella *digital economy*, con particolare riferimento ai nuovi strumenti di analisi dei dati. La seconda parte, che si focalizza sulle policy, si compone di 3 capitoli: antitrust, regolazione e privacy.

PARTE I. Tendenze di mercato

Introduzione

Siamo al centro di una nuova era dello sviluppo economico e sociale. Ciò è frutto di un continuo, inarrestabile processo di innovazione, che ha caratterizzato nel corso degli ultimi due decenni lo sviluppo di internet, attraverso la diffusione della *digital economy* in tutti i settori, non più soltanto legati all'ICT.

A tal proposito, si può parlare di un vero e proprio ecosistema, caratterizzato dai seguenti elementi:

(i) una pluralità di soggetti che concorre alla creazione e offerta dei servizi digitali;

(ii) la trasformazione dei prodotti, anche comuni, in prodotti "intelligenti" e, dunque, in servizi;

(iii) l'ampliamento dei contatti e delle interrelazioni tra i diversi mercati, che innescano un processo di allargamento del perimetro, con l'inclusione di nuove attività e nuovi ambiti merceologici.

Questa trasformazione è il risultato di un processo che, secondo la letteratuta prevalente[1], ha inizio con la diffusione di internet, attraverso lo sviluppo del World Wide Web (1995-2005) e procede con le reti sociali e la diffusione *multi-device* e multipiattaforma, attraverso il cosiddetto web 2.0 (2005-2015).

Così come queste trasformazioni sono state guidate dall'innovazione tecnologica che ha introdotto nuovi paradigmi di connettività, di elaborazione e di utilizzo delle informazioni, allo stesso modo la prossima trasformazione - la terza generazione di internet - sarà guidata da cambiamenti tecnologici che comporteranno significativi mutamenti nella domanda e nelle abitudini di consumo, così come nell'economia, nella struttura delle imprese e nei modelli di business.

Innanzitutto, cresce esponenzialmente la disponibilità di dati digitali: se, in passato, del mondo si offrivano rappresentazioni analogiche che potevano poi essere trasformate in sequenze di uno e zero, oggi è aumentata in modo significativo la capacità di istituzioni, imprese e individui di "datificare" la realtà, ossia di disporre di rappresentazioni degli eventi e dei comportamenti umani che nascono già digitali. Inoltre, coerentemente con la legge di Moore, l'innovazione tecnologica ha determinato e continua a determinare, con una velocità sempre maggiore, a costi sempre minori e con una diffusione pervasiva, l'evoluzione esponenziale delle prestazioni delle tre operazioni base sulle informazioni e sui dati:

(i) elaborazione, in termini di potenza di calcolo dei microprocessori;

[1] Questa prima parte dello studio sui trend di mercato dà conto e utilizza le analisi e i risultati delle principali pubblicazioni e studi strategici in materia di economia digitale e dei dati, segnalate come fonti nelle figure e riportate in dettaglio in bibliografia.

(ii) memorizzazione, in termini di disponibilità e capacità;

(iii) trasmissione, in termini di capacità e velocità grazie allo sviluppo delle reti broadband e ultrabroadband.

Proprio attraverso l'introduzione e l'implementazione di tali tecnologie e la loro diffusione generalizzata, l'esperienza umana sta vivendo un profondo cambiamento, caratterizzato da un alto grado di efficienza dovuto all'impiego dei dati, al ritorno a paradigmi produttivi locali, a sistemi di rete *cloud* integrati capaci di mettere in relazione persone e oggetti, luoghi fisici e virtuali[2].

Siamo dunque agli inizi di una nuova fase dello sviluppo di quella che è comunemente definita "la quarta rivoluzione industriale", caratterizzata da una serie di innovazioni tecnologiche, quali:

1. **Cloud.** Le tradizionali reti di comunicazioni sono oggetto di rilevanti cambiamenti a livello di architettura, basandosi sui paradigmi del Cloud Computing *Cloud-integrated Network* (CIN) e sulle tecniche di virtualizzazione per fornire elevate prestazioni in termini di flessibilità, scalabilità, prontezza di risposta, ottimizzazione delle risorse e personalizzazione dei servizi. Attraverso il *Cloud* si è in grado di gestire simultaneamente e in maniera ottimale, anche dal punto di vista dei costi di esercizio, sia i servizi di rete primari per l'impresa che il traffico dati;

2. **Internet delle cose**. L'aumento esponenziale di macchine e dispositivi connessi a internet - Internet of Things e Machine-to-Machine - digitalizzeranno significativamente la quasi totalità degli oggetti e dei processi del mondo contemporaneo, generando grandi quantità di informazioni che necessiteranno di nuove modalità di accesso, archiviazione dei dati, potenza di calcolo, ampia larghezza di banda e connettività a bassa latenza;

3. **Big Data e Analytics.** Nell'ecosistema digitale l'accesso, la raccolta, l'elaborazione e l'utilizzo dei dati assumono un valore primario in ogni settore di attività. Infatti, la trasformazione in corso è strettamente legata alla pervasività dei dati, fattori comuni a ciascuna infrastruttura abilitante propria di questa fase evolutiva, che soprattutto grazie alle crescenti abilità computazionali, i c.d. *Analytics*, capaci di estrarre valore dall'esame di un set di dati grezzi, nonché al conseguente abbassamento dei costi degli strumenti di raccolta, elaborazione e analisi, sono sempre più in grado di trasformarsi in informazioni destinate a diventare motore di innovazione e fonte di crescita per l'economia e la società. La raccolta e l'analisi dei dati rappresenta dunque il centro di questa quarta rivoluzione, in termini di maggiore velocità/flessibilità/efficienza dei processi di produzione, aumento degli standard qualitativi e rilevante riduzione dei costi;

4. **Intelligenza artificiale/aumentata**. Nuovi sistemi di *intelligence augmentation* / *artificial intelligence* che saranno in grado di supportare l'intelligenza umana, sostituendosi in alcune attività. È interessante notare come il concetto di *intelligence augmentation* si afferma per

[2] Tutto ciò richiederà anche sistemi di sicurezza particolarmente innovativi e affidabili rispetto agli attuali basati su crittografia, connessioni *secure application-level*, nonché su altre forme di sicurezza separate dalla rete.

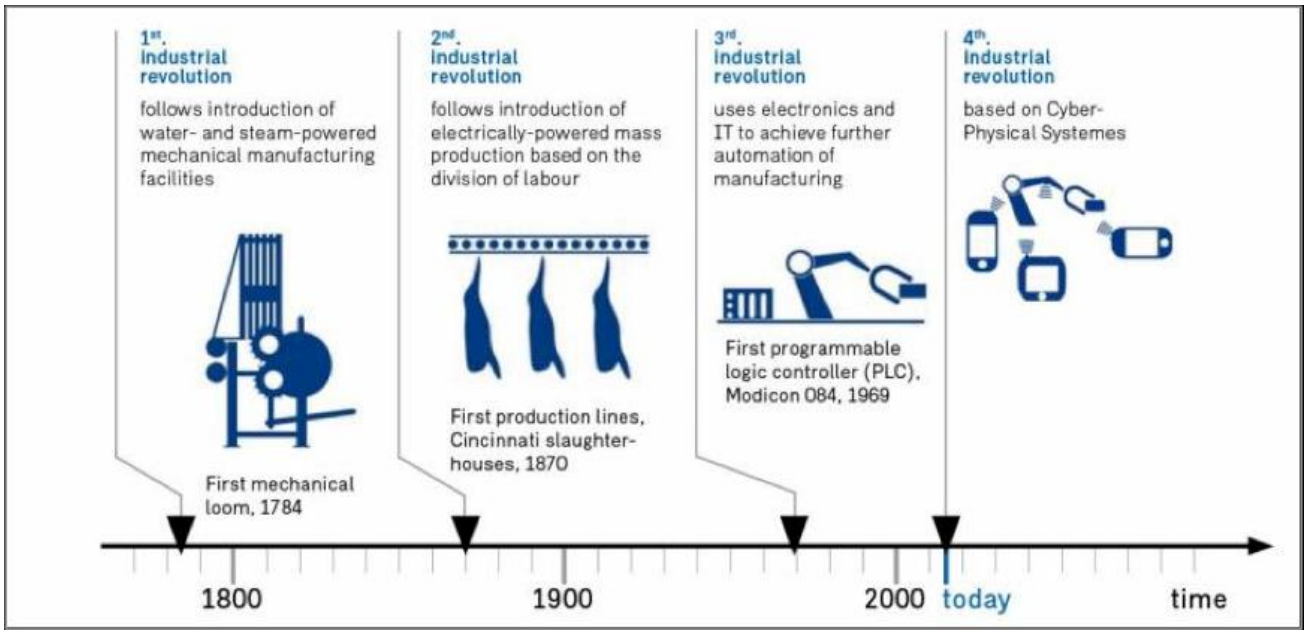
certi versi in alternativa a quello di *artificial intelligence,* che è indirizzato alla creazione della c.d. "conoscenza perfetta" nei sistemi artificiali.

Capitolo 1. L'economia digitale

1.1 La quarta rivoluzione industriale e la società connessa

Per definizione, ciascuna rivoluzione tecnologica comporta, anzitutto, la creazione di qualcosa di nuovo, che a sua volta determina l'introduzione di una nuova infrastruttura necessaria affinché l'innovazione di grande portata, che caratterizza quella particolare rivoluzione industriale, possa essere diffusa, e conseguentemente possa condurre ad una vera e propria rivoluzione economica e ad avere un forte impatto sui comportamenti individuali e sui rapporti sociali.

Figura 1. Cronologia delle rivoluzioni industriali



Fonte: DFKI

Nel corso del tempo, attraversando le varie fasi, le imprese hanno acquisito una dimensione sempre più globale ampliando la propria offerta di prodotti e servizi, nonché la rispettiva base di utenti. Prima della rivoluzione industriale, la maggior parte delle imprese assumevano una dimensione per lo più locale e conseguentemente godevano di limitate possibilità di crescita.

Dalla metà del 1700 alla metà del 1800, la rivoluzione industriale ebbe un effetto dirompente sulle strutture economiche esistenti, attraverso l'introduzione di una vasta gamma di tecnologie e macchinari. Tutto quello che, prima della rivoluzione, veniva rigorosamente realizzato a mano, successivamente venne prodotto attraverso le macchine: una nuova produzione basata su economie di risultato. Nel 1840, il telegrafo e il codice Morse, resi disponibili sul mercato, facilitarono la comunicazione, velocizzando i processi decisionali e dunque la produzione stessa.

Con l'avvento del telefono nel 1870, ebbe inizio una lunga ondata di trasformazioni economiche, che videro le imprese estendere la propria dimensione geografica. Tra la fine del XIX e gli inizi del XX

secolo le imprese nazionali assunsero grandi dimensioni con apparati e linee di assemblaggio particolarmente articolate e organizzate.

Il 1960 fu l'anno dello sviluppo delle prime multinazionali, di un sistema delle comunicazioni molto più efficiente, dello sviluppo del trasporto aereo e delle relative infrastrutture associate.

Nel 1982, fu introdotto il protocollo TCP/IP come protocollo standard per ARPANET – la prima rete IP al mondo. L'età di internet ebbe inizio nel 1990, dopo la diffusione dei personal computer (1970), e l'invenzione del World Wide Web (1989) e dei primi servizi internet basati su IP tramite modem e linee telefoniche. Le tecnologie Ethernet e TCP/IP furono introdotte per la prima volta nella realtà imprenditoriale alla fine del 1980 esclusivamente per uso interno. Il grande successo di queste tecnologie determinò negli anni novanta la loro adozione, facendo sì che le imprese maggiormente efficienti potessero collegare le loro reti LAN aziendali ai siti, consentendo operazioni in grado di aumentare considerevolmente l'efficienza.

Tabella 1. L'evoluzione di internet

Web 1.0 (Web of content) (1990-2000)	Web 2.0 (Web of communication) (2000-2010)	Web 3.0 (Web of context) (2010-2020)	Web 4.0 (Ubiquitous Webs) (2020-2030)
• Company generated content • Taxonomy • Navigazione: menu • Interazione utente/sito • Online quando serve • "Pesantezza" (funzioni client side) • Banda stretta • Servizi "chiusi" • E-commerce "si paga" • Release successive	• User generated content • Folksonomy spontaneismo della rete • Navigazione: search & peer • Rete come spazio sociale: Social network • Sempre online • "Leggerezza" • Banda larga • Servizi "aperti" • "Freemium"	• Web semantico • Applicazioni "intelligenti" • Mobile Internet/Mobile App • Cloud computing • Machine-to-machine communications • Big data • Deep e machine learning • Banda ultra larga • IoT	• Geospatial web • Ubiquitous/seamless connettivity (seamless) • IoE • Smart City • Smart grid • Connected space • Articifial Intelligence (Robots, personal assistant, smart agents) • Wearable technology • Augmented/Virtual reality • Data Analytics • Valuable Small Data

Fonte: ITMedia Consulting / LUISS DREAM, Il video e la terza rivoluzione di internet

Originariamente, il web, nella sua prima fase di sviluppo (Web 1.0), era caratterizzato prevalentemente da siti statici, senza alcuna possibilità di interazione con l'utente, eccetto la normale navigazione tra le pagine, l'uso della posta elettronica e delle funzioni di ricerca.

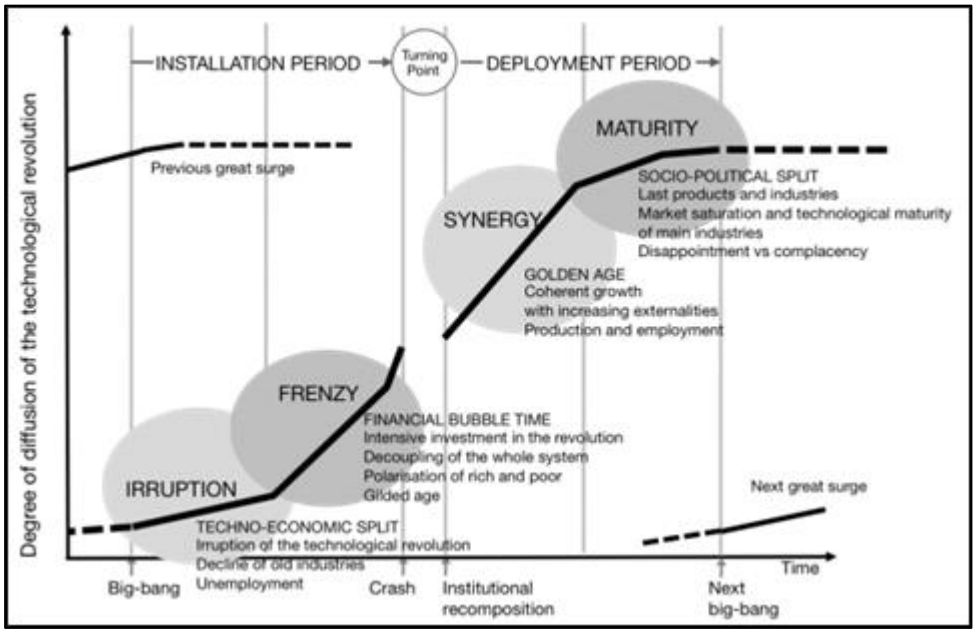
La fase successiva, quella del Web 2.0, è sinonimo di condivisione, in cui internet assurge a un insieme di molteplici applicazioni online che permettono uno spiccato livello di interazione tra il sito web e l'utente, ottenuto tipicamente attraverso opportune tecniche di programmazione web, afferenti al paradigma del web dinamico in contrapposizione al così detto web statico proprio del Web 1.0. Siamo agli inizi della prima *commercial digital era* caratterizzata dalla disponibilità di contenuti multimediali e diritti digitali associati; dall'avvento del commercio elettronico e quindi dalla possibilità di acquistare beni fisici su internet; dalla capacità di condividere immagini, informazioni, video, pensieri sui social media, e contenuti personali in genere, mediante soprattutto dispositivi mobili.

Si arriva così all'attuale Web 3.0, o terza generazione di internet, costituita da servizi che utilizzano web semantico, *natural language search, data mining, machine learning* e tecnologia di intelligenza artificiale. Il web 3.0 è caratterizzato dalle presenza di internet in qualsiasi cosa ci circonda (*connected society* e *internet of things*), anche in chiave di offerta di servizi che evidenziano un processo di comprensione delle informazioni facilitato dalle macchine, al fine di fornire un'esperienza più produttiva e intuitiva all'utente.

La successiva fase di sviluppo del web sarà verosimilmente definita Web 4.0 e dovrebbe integrare pienamente le due fasi evolutive precedenti (Web 2.0 e Web 3.0) ai fini della realizzazione di un web per così dire "Ubiquitous": sostanzialmente le applicazioni presenti in rete avrebbero lo scopo di connettere in modo automatico le persone (così come il web semantico mette in connessione in modo automatico i contenuti), sulla base delle attività che stanno svolgendo, per facilitare la collaborazione e il raggiungimento di scopi condivisi unendo le rispettive risorse e competenze. Si tratterebbe quindi di un internet pienamente integrato con la realtà fisica, al servizio delle relazioni, per moltiplicarle e potenziarle.

Ne discende che l'innovazione è stata al centro di tutte le rivoluzioni industriali che risultano accomunate quindi, dalla sostituzione della tecnologia principale della rivoluzione precedente, ad esempio il vapore al posto dell'acqua, oppure l'elettricità al posto del vapore; secondo il modello dell'onda lunga, o **"**Grande onda**",** come illustrato nel grafico che segue.

Figura 2. Modello della Grande onda



Fonte: Perez C., Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages

Nelle fasi iniziali si assiste ad una lotta tra il nuovo paradigma tecnico-economico, legato all'irruzione di nuove innovazioni, e quello precedente radicato in strutture di produzione ormai consolidate e integrate nell'ambiente socio-culturale e nel quadro istituzionale. Una volta vinta tale battaglia, il nuovo paradigma inizia a diffondersi in tutta l'economia e in tutta la società propagandosi in due periodi distinti, Installation Period e Deployment Period, della durata di circa 20/30 anni ciascuno. Il turning point tra i due periodi è di solito un periodo di recessione che prevede una ricomposizione dell'intero sistema, in particolare del contesto normativo che consente la ripresa della crescita e il pieno sfruttamento della rivoluzione tecnologica. Nel periodo dell'Installazione, il nuovo paradigma tecnico-economico dell'ICT basato sui dati e sulle informazioni, ha coinvolto i settori basati largamente sull'utilizzo delle tecnologie ICT, come ad esempio il settore finanziario o dei contenuti multimediali (audio/video). Ora ci troviamo nel pieno della fase di Deployment, ovvero nel periodo in cui le innovazioni ICT si stanno diffondendo in modo pervasivo in tutta l'economia e nella società in generale. Questa è la fase in cui si realizzano i maggiori benefici, la c.d. *Golden Age* caratterizzata dalla ripresa della crescita abilitata dalla nuova economia dei dati. Un presupposto essenziale è dato dalla disponibilità, con adeguati livelli di qualità e affidabilità, di tecnologie e infrastrutture abilitanti sia in termini di reti a banda larga e ultra-larga per la trasmissione delle informazioni che in termini di diffusione di dispositivi e strumenti di elaborazione e memorizzazione delle informazioni, quali Cloud Computing, smartphone, tablet, sensori, smart objects, applicazioni software, ecc.

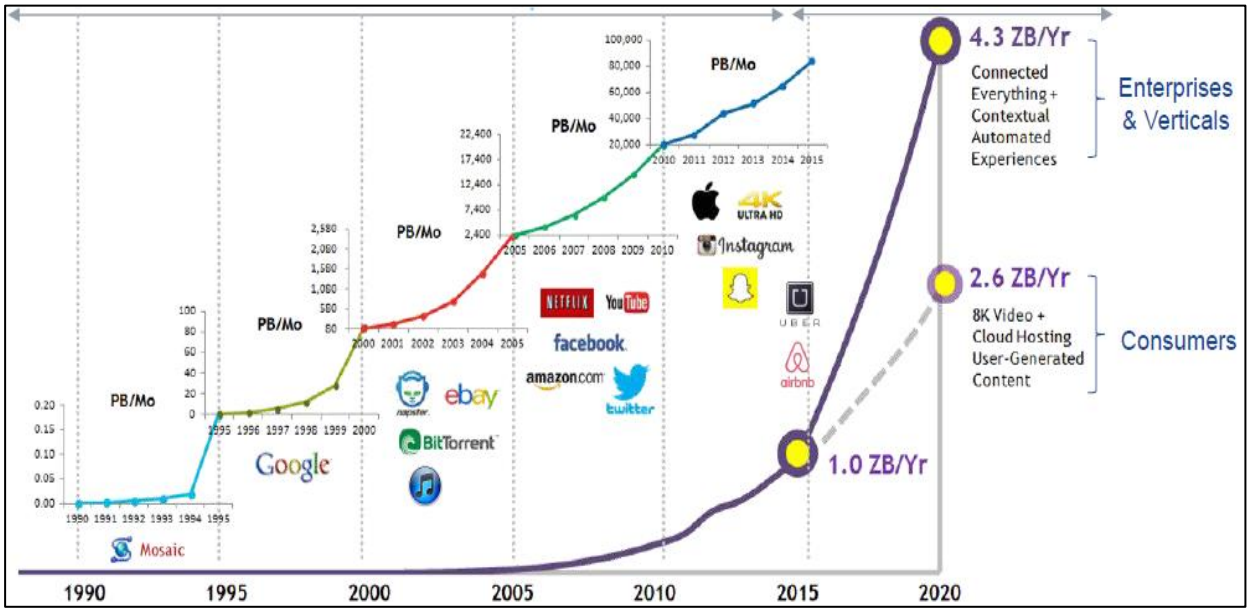
L'attuale espansione dell'ICT consente miglioramenti nella produttività delle aziende di qualsiasi settore industriale, delle organizzazioni pubbliche e della vita dei cittadini, dagli scenari *di Smart City* con applicazioni *water management, waste management, traffic management, energy management* alle nuove opportunità di formazione, tramite *e-learning* e MOOC, di condivisione di contenuti, grazie

ai social networks, ma anche di risorse e beni, tramite le applicazioni di *sharing economy*, di intrattenimento, ad esempio grazie ai servizi VOD in streaming.

All'interno dell'ecosistema digitale, sotto la spinta impressa dalla diffusione del paradigma internet che agisce dunque come potente forza di cambiamento, si afferma il nuovo ecosistema "connesso" di persone e cose/oggetti. In tale scenario, la crescita dei dati è sempre più esponenziale.

Il grafico sottostante mostra la crescita del traffico di rete (curva grigia) sin dagli albori dell'era di internet, attraverso la mappatura delle tre fasi del web in piccoli incrementi quinquennali, che mostrano i trend relativi agli attori che hanno caratterizzato ciascun intervallo analizzato. L'ultimo ciclo quinquennale corrisponde ad una nuova era di internet, caratterizzata da una crescita senza precedenti del traffico dati legata alla connessione di qualsiasi cosa in termini di oggetti e persone.

Figura 3. La crescita del traffico internet



Legenda: PB: Petabytes= 10^{15} bytes; ZB: Zettabytes = 10^{21} bytes

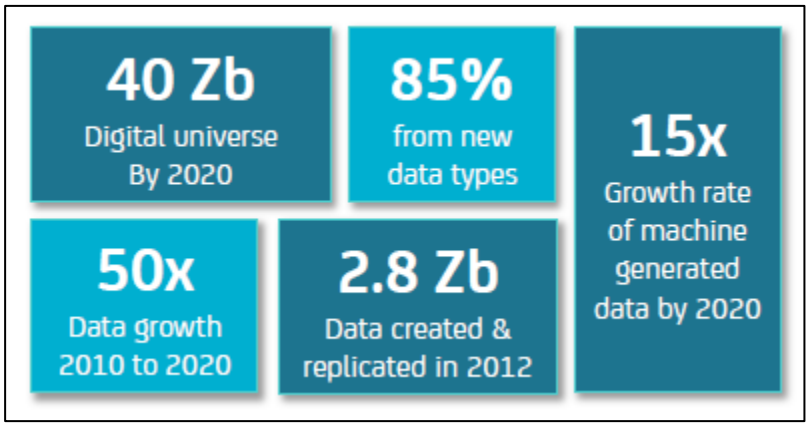
Fonte: The Future X Network: A Bell Labs Perspective

L'azione sinergica di digitalizzazione e connettività pone il campo delle comunicazioni al centro dei processi di trasformazione.

Lo sviluppo delle reti internet a banda ultra larga determinerà a sua volta il raggiungimento di un livello di automazione senza precedenti, attraverso il collegamento alla rete di persone e oggetti digitalizzati. Tutto ciò trasformerà le imprese attorno a un nuovo paradigma globale-locale che inciderà positivamente sul fattore tempo attraverso l'ottimizzazione di attività e processi ripetitivi, ridisegnando le architetture e i modelli di business delle imprese.

Si stima infatti che nel 2020 il processo di digitalizzazione in atto porterà l'universo digitale a raggiungere addirittura una quota pari a 40 Zb, con una crescita digitale 50 volte superiore su base decennale (2010-2020).

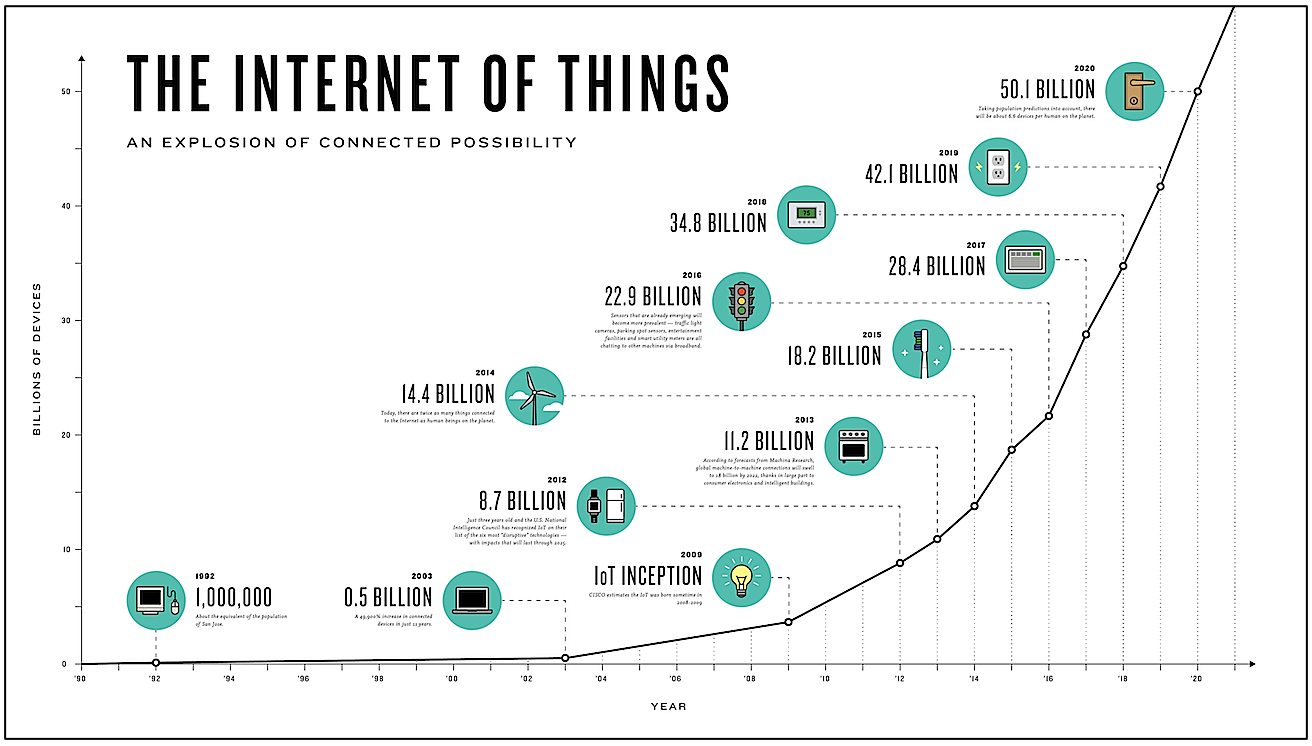
Figura 4. Le stime della crescita digitale



Fonte: Unicredit

Inoltre, il tasso di crescita delle macchine e dei dispositivi capaci di generare dati entro il 2020 aumenterà esponenzialmente di circa 15 volte come mostrato dal grafico seguente e a un ritmo sempre crescente.

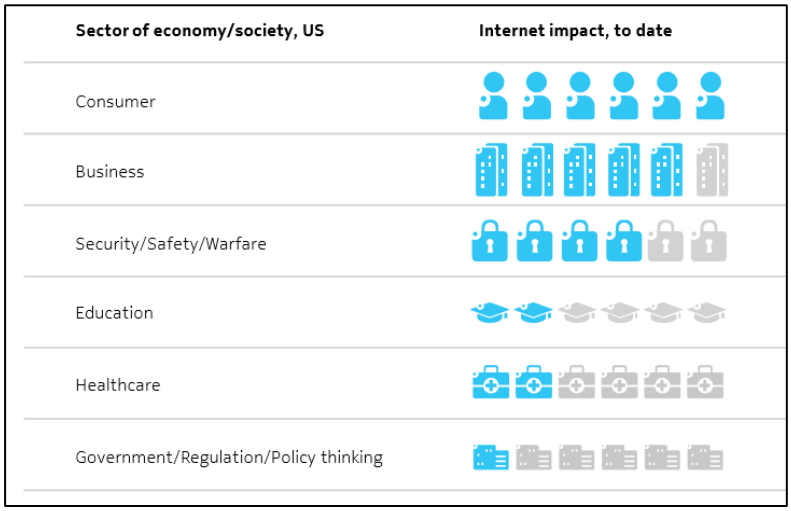
Figura 5. La crescita esponenziale dell'universo digitale



Fonte: Cisco

La figura seguente mostra il livello di maturità dell'impatto di internet su diverse *industry*. Appare evidente che ad oggi il livello più elevato sia stato raggiungo nei servizi B2C (si pensi alla produzione e distribuzione di contenuti audio/video) e nei servizi B2B (si pensi al mondo della finanza), mentre molto può essere ancora realizzato in ambiti B2A quali istruzione, salute e cura della persona, servizi pubblici, etc.

Figura 6. Impatto relativo di internet su diverse industry



Fonte: Meeker, M., "Internet Trends 2015"

Come già anticipato, la quarta rivoluzione industriale o industria 4.0, presuppone l'utilizzo di macchine intelligenti, interconnesse e collegate a internet e, ai fini del suo sviluppo, necessita di tutta una serie di tecnologie abilitanti, definite anche fattori di cambiamento (o *disruptive technologies*) che la caratterizzano e che porteranno al sistema economico-industriale nuovi e importanti vantaggi quali:

- *Maggiore flessibilità*, dal punto di vista dello sviluppo di nuovi prodotti, data dalla possibilità di produrre nella stessa struttura produttiva beni diversi, grazie all'automazione e alla trasmissione continua di dati lungo l'intera filiera. Inoltre la capacità di adattamento dei macchinari a determinate specifiche grazie a opportune e immediate programmazioni, consente una personalizzazione di massa capace di spingersi, in certi casi, fino alla realizzazione di pezzi unici in tempi molto limitati;
- *Maggiore velocità*, dal prototipo alla produzione in serie attraverso tecnologie innovative;
- *Maggiore produttività*, attraverso ridotti tempi di set-up, diminuzione degli errori e fermi macchina;
- *Migliore qualità* e minori scarti mediante sensori che monitorano la produzione in tempo reale;

- *Maggiore competitività* del prodotto, grazie alle maggiori funzionalità derivanti dall'implementazione delle tecnologie abilitanti, a cominciare dall'Internet of things.

Per tecnologie abilitanti si intendono in particolare: *Cloud computing*, *Internet of Things* (IoT), *Intelligent Automation* (IA), *Analytics* e *Big data*, *Advanced robotics* e *3D printing*. Le prime tre, che vengono approfondite di seguito, si trovano attualmente in uno stadio di sviluppo più avanzato, mentre alla quarta, che è abilitatore ma al contempo input essenziale di tutte le altre e costituisce l'oggetto primario del presente studio, è dedicato l'intero secondo capitolo.

1.1.1. Cloud computing

La letteratura prevalente fornisce diverse definizioni del Cloud. Si riportano di seguito le più affermate:

- "Il Cloud Computing è un modello per consentire un accesso tramite rete ad un insieme di risorse di calcolo configurabili (reti, server, storage, applicazioni e servizi) che possono essere istanziate e rilasciate con un minimo sforzo di gestione ed interazione con il service provider", NIST - United States National Institute of Standards and Technology;

- "Il Cloud Computing è uno stile di computazione secondo cui le risorse IT sono erogate come servizio, elasticamente ed in maniera scalabile" – Gartner.

La rivoluzione abilitata dal cloud nell'ambito delle modalità di creazione ed erogazione di servizi e applicazioni digitali per le imprese, nonché di contenuti, può essere fatta risalire al lancio di Salesforce.com avvenuto nell'autunno del 1999. Sebbene internet stesse entrando sempre più nelle vite dei consumatori (Netscape era stata fondata nel 1994 mentre Amazon e Ebay nel 1995), il concetto di accedere a servizi e applicazioni di business per le imprese attraverso il browser era praticamente sconosciuto. Grazie a Salesforce.com le imprese non avevano più bisogno di acquistare prodotti hardware e software costosi per le proprie applicazioni in quanto qualsiasi esigenza in tal senso poteva essere soddisfatta su internet da server localizzati in data center di proprietà di aziende terze.

La prima era del Cloud nacque così decretando di fatto la trasformazione di internet da grande "contenitore" di informazioni eterogenee a piattaforma attraverso cui servizi e applicazioni informatiche a livello d'impresa potevano essere facilmente creati ed erogati in maniera scalabile su base globale. In questa prima fase la creazione dell'infrastruttura sottostante, ovvero i data center composti da numerosi server e dispositivi di rete, costituì il principale ostacolo di sviluppo di questo nuovo paradigma.

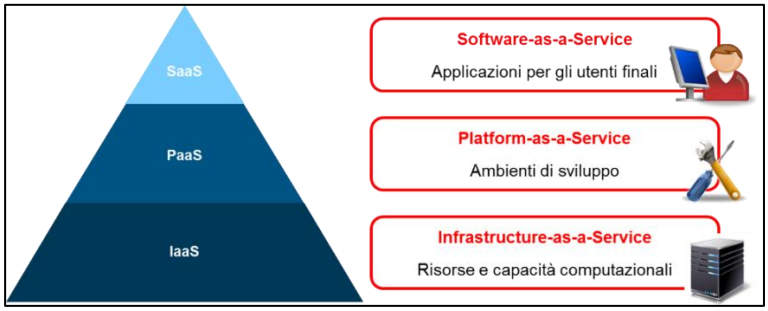
Nel 2006 il lancio di Amazon Web Services – AWS segnò un'ulteriore e decisiva evoluzione rendendo di fatto di massa l'infrastruttura di Cloud Computing. Potendo utilizzare in modo economicamente conveniente, grazie alla modalità *pay-as-you-go*, le risorse hardware messe a disposizione da Amazon, le barriere all'entrata per lo sviluppo di applicazioni e servizi si ridussero enormemente. In questo modo chiunque, dai fornitori di servizi ai piccoli gruppi di sviluppatori di applicazioni,

potevano realizzare offerte di servizi per milioni di utenti senza doversi preoccupare di investimenti e problematiche relative alle sottostanti infrastrutture di elaborazione, memorizzazione e connettività.

Da allora il Cloud e le tecnologie di virtualizzazione ad esso strettamente connesse, si sono diffuse rapidamente con tassi di crescita a doppia cifra in tutto il mondo e nella quasi totalità dei settori economici, rivoluzionando di fatto le modalità con cui le imprese implementano le applicazioni aziendali interne e gestiscono le risorse, così come i servizi rivolti ai clienti e le funzioni di business come ad esempio il *Customer Relationship Management* (CRM). Tutto ciò consente un più efficiente utilizzo delle risorse, scalabilità rapida e flessibile, e un abbassamento dei rischi attraverso il bilanciamento tra spese operative e spese in conto capitale per i nuovi servizi e funzionalità. Infatti, in questo modo le aziende possono esternalizzare con maggiore efficacia gran parte del loro operato.

La figura seguente mostra le tre tipologie di modello di servizio per il Cloud.

Figura 7. Modelli di servizio Cloud



Fonte: elaborazioni ITMedia Consulting su fonti varie

In ogni modello è presente l'espressione **"***as a service***"** per indicare la possibilità che ha l'utente di richiedere l'utilizzo delle varie soluzioni IT «come servizio». Ciò significa che l'utente può decidere quando utilizzare le risorse, per quanto tempo e definendone le caratteristiche sulla base delle proprie necessità

In un servizio di tipo IaaS (*Infrastructure as a Service*) – Infrastruttura come Servizio – il fornitore di soluzioni Cloud mette a disposizione l'utilizzo di risorse hardware, accessibili da remoto; il servizio è quindi rappresentato dalle capacità di calcolo, dalle componenti della struttura di rete, o dalle capacità di archiviazione richieste dall'utente finale in base alle proprie esigenze.

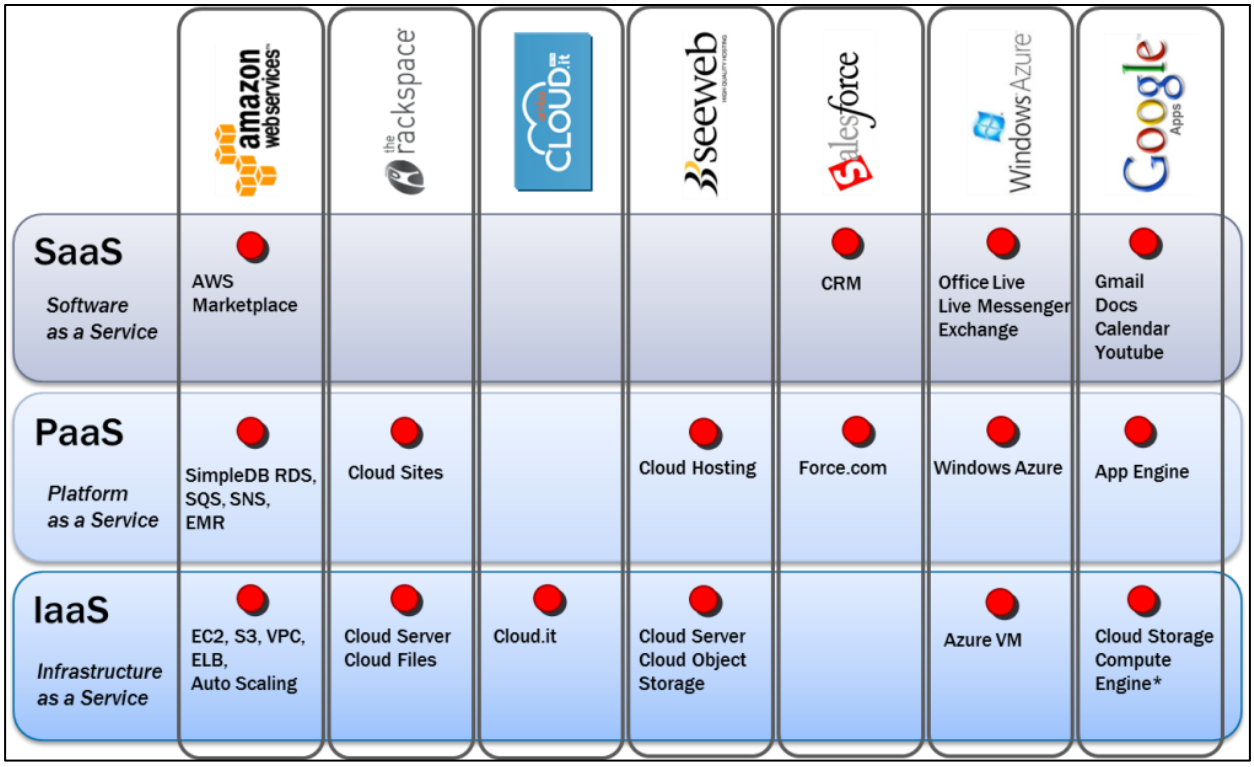
Nei servizi di tipo PaaS (*Platform as a Service*) – Piattaforma come Servizio – il fornitore di soluzioni Cloud mette a disposizione una piattaforma, ospitata all'interno della propria infrastruttura hardware, che comprende gli strumenti necessari allo sviluppo di applicazioni, oltre allo spazio per ospitarle ed eseguirle. Gli utenti di tali servizi possono quindi sviluppare le proprie applicazioni attraverso l'utilizzo delle API (Application Programming Interface) fornite, oppure tramite interfacce grafiche, senza la necessità di installare alcun software sulle proprie macchine. Esempi tipici di soluzioni PaaS sono gli ambienti di sviluppo software Google App Engine, Microsoft Windows Azure, Force.com di Salesforce.com.

Per i servizi erogati in modalità SaaS (*Software as a Service*) – Software come Servizio – il provider di servizi Cloud fornisce agli utenti finali delle applicazioni software, ospitate ed eseguite in ambiente Cloud. Gli utenti finali interagiscono con tali applicazioni tramite il web, solitamente senza la necessità di installare ed eseguire le applicazioni sul proprio computer. È infatti il provider il responsabile della gestione di tutte le risorse, hardware e software, necessarie per l'erogazione del servizio. Esempi di soluzioni SaaS sono: Spotify, Facebook e Youtube. Vi sono poi soluzioni CRM quali Sales Cloud (Salesforce.com), servizi di posta elettronica come Gmail, e servizi di storage online quali Dropbox.

Infine, un'ulteriore classificazione riguarda i modelli di distribuzione dei servizi Cloud. Quando si parla di *Public Cloud*, i servizi della nuvola sono messi a disposizione di chiunque e a disposizione di grandi industrie tramite un'organizzazione Service Provider. Nel caso del *Private Cloud*, il Cloud è realizzato esclusivamente per una specifica organizzazione e può essere gestito dall'organizzazione stessa all'interno di un proprio data center o da una terza parte in aree dedicate. Si parla, invece, di *Hybrid Cloud* quando un'organizzazione utente fornisce ed utilizza alcune risorse *in house* mentre per altre risorse si affida ad una fornitura esterna. Infine si ha il *Community Cloud* quando l'infrastruttura Cloud è condivisa tra diverse organizzazioni.

Nella figura seguente sono presentati i principali player del mercato dei servizi Cloud.

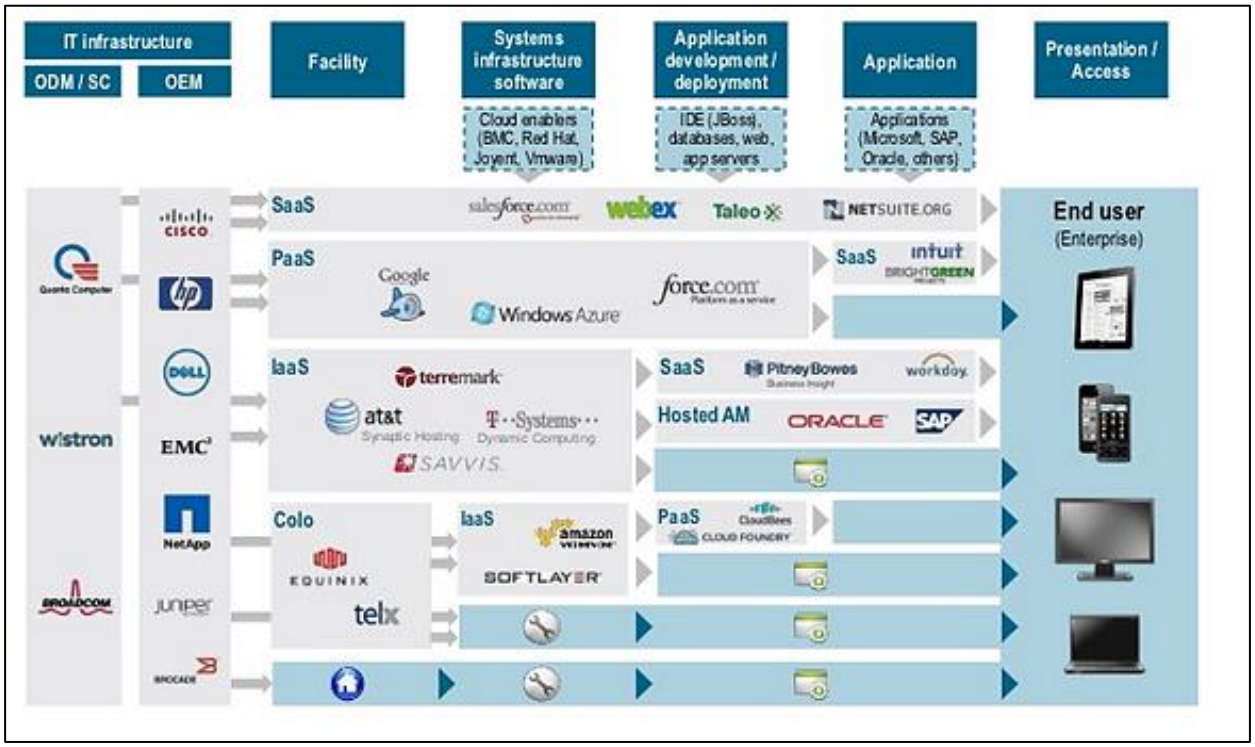
Figura 8. Principali player di mercato Cloud



Fonte: elaborazioni ITMedia Consulting su fonti varie

La figura seguente illustra invece la catena del valore dei servizi Cloud con i principali player di mercato.

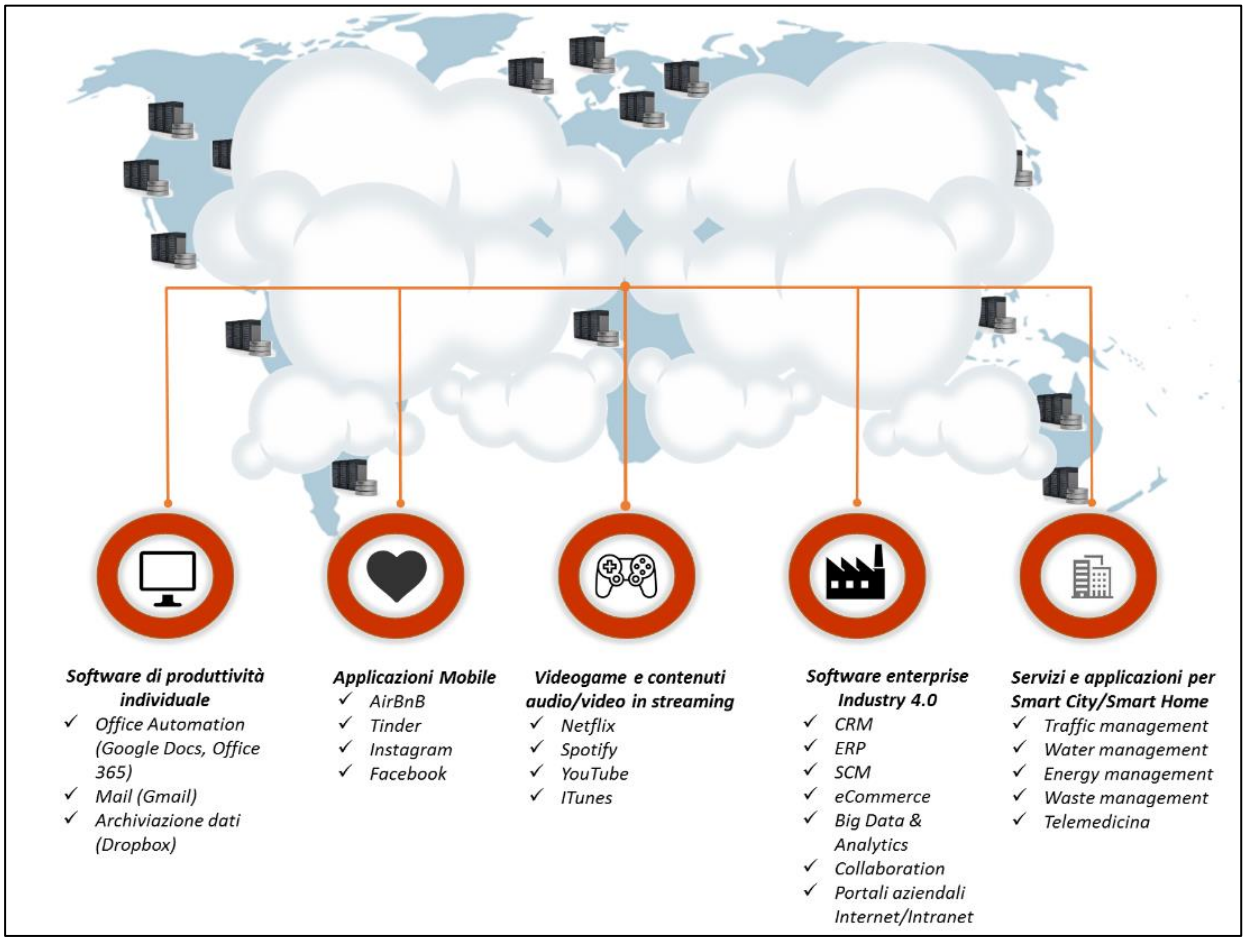
Figura 9. Catena del valore dei servizi Cloud



Fonte: Cisco, The Cloud Value Chain Exposed

Ad oggi i servizi offerti su Cloud sono maturi e largamente utilizzati in numerosi processi di business e da svariate applicazioni come ad esempio CRM, ERP, SCM, Office Automation, Posta Elettronica, HR Management, Enterprise Social Network & Collaboration, eCommerce, sviluppo Portali Internet/Intranet, piattaforme di calcolo in ambito Big Data & Analytics, Videogame, Streaming di contenuti audio e video (quest'ultimi in particolare tramite CDN - Content Delivery Networks), ecc. L'immagine seguente offre un quadro sintetico dei diversi settori in cui il Cloud rappresenta di fatto una delle principali tecnologie abilitanti e illustra l'attuale trend evolutivo su scala globale.

Figura 10. Evoluzione del Cloud verso architetture edge/fog computing e servizi abilitati



Fonte: elaborazioni ITMedia Consulting su fonti varie

Oggi è in corso una seconda evoluzione rispetto al modello iniziale basato su architetture centralizzate che ha consentito di ottenere rilevanti economie di scala.

Per soddisfare la duplice esigenza di una crescente capacità di banda e di tempi di latenza sempre più contenuti, due fattori, banda e latenza, strettamente necessari per lo sviluppo di nuovi servizi innovativi, nonché per sostenere l'aumento esponenziale del numero di dispositivi, città, imprese connessi, è necessario un approccio distribuito che porti il cloud in prossimità degli utenti e di fatto ai bordi della rete. Si parla infatti di *edge computing/edge cloud* anche se recentemente è stato introdotto il termine *fog computing* proprio per sottolineare che più di una nuvola si tratta di una vera e propria nebbia distribuita un po' dappertutto.

Il *cloud computing*, nonostante la sua estrema importanza, presenta infatti un limite: il grado di latenza che si può raggiungere nelle fasi di invio dati dal dispositivo al server cloud, di elaborazione e di inoltro della risposta dal server cloud al dispositivo. In alcune applicazioni, come nel *gaming* e nelle sessioni di videoconferenza, queste latenze sono tali da diventare insostenibili e rendere così il cloud una strada impraticabile. Tuttavia, quest'ostacolo può essere arginato attraverso l'implementazione di tecnologie di tipo *edge/fog computing*.

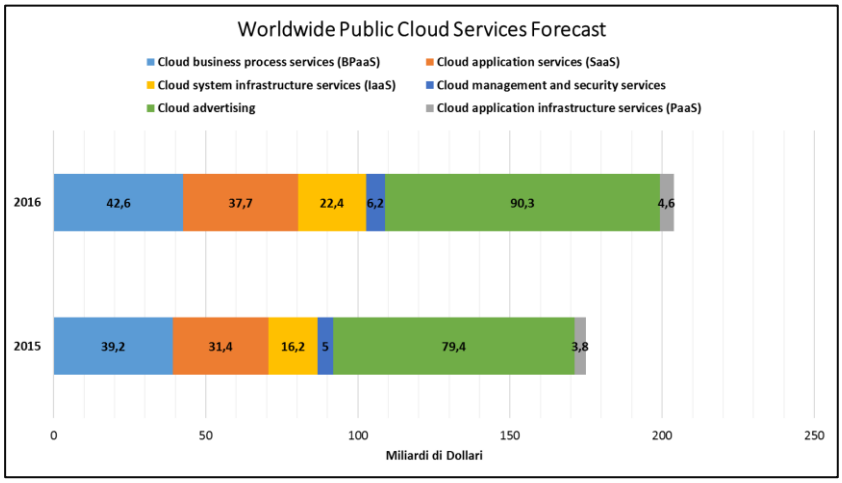
Le enormi latenze a cui può giungere un'elaborazione cloud dipendono principalmente da un fattore limitativo: la disponibilità di banda. Nonostante lo sviluppo di tecnologie più all'avanguardia come le reti 4G, 5G e la fibra, che comunque presentano delle limitazioni dovute al costo del traffico e della banda disponibile, trasferire i dati di elaborazione verso il cloud, per poi elaborarli e interpretarli può diventare piuttosto dispendioso in termini di banda, di tempo e, conseguentemente, di costi. Un esempio su tutti: il backup nel cloud è un'ottima soluzione per assicurarsi un archivio dati geo-delocalizzato e sicuro, ma la povertà di banda a cui si è sottoposti rende il processo lento e capace di saturare tutta la banda a disposizione, con conseguenti innalzamenti dei costi infrastrutturali. Il problema diventa ancora più importante se si guarda all'IoT, considerando che ogni oggetto connesso ha un indirizzo IP e comunica in rete tutta una serie di dati da elaborare, a volte anche di notevole entità dal punto di vista delle dimensioni.

L' *edge/fog computing* risolve questa difficoltà di trasferimento dati, permettendo ad esempio ai router o ai dispositivi network locali di processare in locale la maggior parte dei dati provenienti dagli oggetti o dai *device* emittenti, almeno fin tanto che non sia richiesta una potenza computazionale o un'allerta tale da dover far intervenire i server presenti nel cloud. Per semplicità, si pensi al caso del motore di un jet, questo dispositivo è in grado di generare ogni 30 minuti circa 10 TB di dati relativi alle sue performance e alle condizioni di funzionamento. Una mole di dati che deve essere trasferita, elaborata e validata con conseguenti problematiche a livello di latenze. Con l'*edge/fog computing*, invece che inviare ai cloud server tutte le comunicazioni in cui il motore da conto del corretto funzionamento di tutte le componenti, questi dati vengono passati al router locale o all'infrastruttura network locale, che non invia al cloud nulla fin tanto che non si generano dati che possano far presagire un malfunzionamento o una richiesta di intervento tecnico. Dunque, l'elaborazione dei dati non avviene sulla nuvola, ma su strutture intelligenti locali, capaci di mettere a disposizione la propria potenza computazionale limitata per svolgere task semplici, ideati però per permettere l'invio e l'elaborazione nel cloud dei soli dati critici e utili (come rappresentato in figura 10). In questo modo, vengono notevolmente ridotte se non eliminate del tutto le latenze, registrando anche un sostanziale risparmio di banda.

Tutti i maggiori player di mercato stanno investendo in tal senso: Google ad esempio sta costruendo Data Center in numerosi Paesi dall'Australia a Mumbai, Amazon sta per inaugurarne di nuovi a Parigi e nel Regno Unito mentre IBM lo scorso anno ne ha aperto uno a Milano da 50 milioni di euro. Tutto ciò, inoltre, consentirà di superare alcuni vincoli normativi legati a ragioni di privacy dei dati.

In estrema sintesi tale evoluzione sta determinando l'industrializzazione delle infrastrutture informatiche diventate sempre più integrate e fondamentali per tutti i processi e attività della società odierna. A ulteriore conferma di quanto sopra, gli investimenti nel settore continuano a crescere significativamente con un valore di mercato globale pari a 204 miliardi di dollari a fine 2016, nel solo segmento Public Cloud, con una crescita del 16.5% rispetto ai 175 miliardi del 2015.

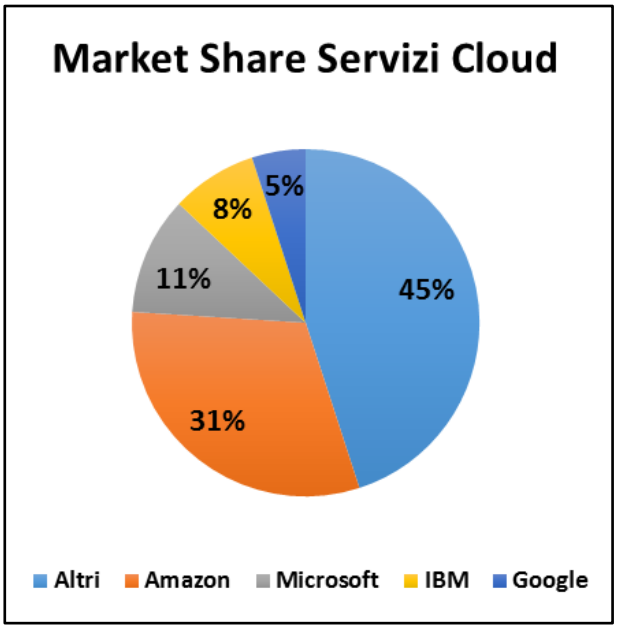
Figura 11. Previsioni sulla spesa globale dei servizi Public Cloud



Fonte: Gartner, Forecast: Public Cloud Services, Worldwide, 2015-2021

Per quanto concerne le quote di mercato tra i principali player il grafico seguente mostra la situazione attuale. Data l'esiguità del numero di player a livello globale alcuni sostengono che vi sia un rischio di oligopolio legato soprattutto al *vendor lock-in* per le aziende utilizzatrici dei servizi Cloud. Secondo l'ultimo report IDC[3], il cloud diventerà sempre più concentrato con i 5 principali *vendo*r IaaS/PaaS che controlleranno il 75% del mercato entro il 2020.

Figura 12. Market Share dei Servizi Cloud (2016)



Fonte: elaborazioni ITMedia Consulting su dati Gartner, Cisco

[3] *IDC, FutureScape: Worldwide IT Industry 2017 Predictions*, 2016.

Sempre entro tale anno il 67% dell'intera spesa IT, a livello di infrastrutture e software, sarà guidata dallo sviluppo di offerte di servizi basati sul cloud. Già oggi la strategia *cloud-first* è sempre più *cloud-only*, sulla base del trend secondo cui quasi ogni azienda diventerà un *cloud service provider* di servizi innovativi su un proprio *marketplace*, trasformando le *capability* cloud da un mero aspetto IT ad un asset fondamentale a livello di *operations* del core business.

1.1.2. Internet of Things

Gli enti internazionali di standardizzazione forniscono varie definizioni dell'Internet of Things:

- "una rete di elementi (ciascuno con sensori integrati) connessi a internet", IEEE – Institute of Electrical and Electronics Engineers;
- "rete mondiale di oggetti interconnessi indirizzabili in modo univoco, basata su protocolli di comunicazione standard", IETF – Internet Engineering Task Force;
- "infrastruttura globale (*ubiquitous network* disponibile ovunque, in qualsiasi momento, da qualsiasi cosa e da chiunque*)* per la società dell'informazione, che abilita servizi avanzati grazie all'interconnessione degli «oggetti» fisici e virtuali sulla base di soluzioni ICT (esistenti ed evolutive)", ITU – International Telecommunications Union;
- "un'infrastruttura di rete globale e dinamica con capacità di autoconfigurazione basata su protocolli di comunicazione standard e interoperabili nei quali le «cose» fisiche e virtuali hanno identità, attributi fisici e una «personalità» virtuale, utilizzano interfacce intelligenti e sono perfettamente integrati nella Rete", IERC – IoT European Research Cluster.

Indipendentemente dalla sua definizione, l'Internet of Things in realtà, consiste in una moltitudine di oggetti fisici dotati di un'interfaccia digitale per connettersi a internet, diventando di fatto degli *smart objects* e facendo sì che il mondo digitale tracci una mappa di quello reale, dando un'identità elettronica alle cose e ai luoghi dell'ambiente fisico. Le interfacce digitali possono essere le più diverse, basandosi su tecnologie quali identificazione a radio frequenza (RFID), codici QR, bluetooth, Zigbee, Wi-Fi, Reti cellulari, Reti satellitari, etc.

Tutto ciò è sostanzialmente reso possibile dalla trasformazione dei dati e delle informazioni in conoscenza, attraverso metodologie e processi di raccolta ed elaborazione, nonché di scambio informativo, in virtù di particolari caratteristiche e capacità. Gli oggetti creano un sistema pervasivo ed interconnesso avvalendosi di molteplici tecnologie di comunicazione (tipicamente a corto raggio). Queste tecnologie di più basso livello, quando integrate in architetture basate sul protocollo IP, possono dar concretamente vita alla visione dell'Internet delle cose, essendo in grado di dialogare con i nodi della rete internet.

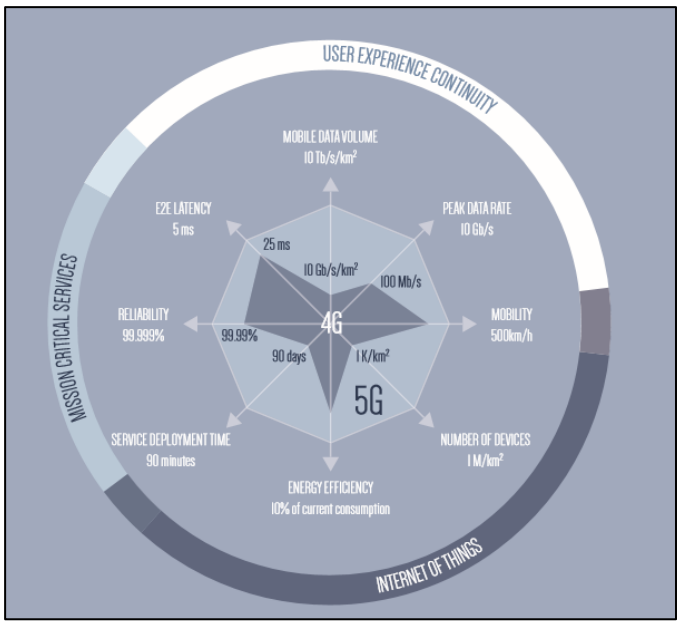
L'Internet delle cose (IoT) è finalizzata a migliorare la qualità della nostra vita attraverso l'introduzione di servizi innovativi pensati per una vasta gamma di applicazioni: dall'automazione industriale agli elettrodomestici, dalla sanità alla elettronica di consumo, e molti settori che quotidianamente sono impiegati in diversi contesti della vita umana. La visione di un "mondo

intelligente" dove tutto è direttamente connesso alla rete non è più un miraggio. Inoltre, la crescita degli oggetti è sostenuta dal costante aumento del numero di dispositivi in grado di monitorare ed elaborare informazioni dal mondo fisico, nonché dall'abbassamento dei rispettivi costi. La maggior parte delle soluzioni basate sul sistema dell'IoT, infatti, richiede servizi *cloud*, facendo leva sulle loro capacità virtualmente illimitata a sfruttare efficacemente le potenzialità di enormi minuscoli sensori e attuatori verso una cosiddetta nube di cose.

Nonostante tutte le condizioni sembrano essere particolarmente favorevoli, rimane ancora molto da fare prima di raggiungere un ecosistema di oggetti che sia al contempo funzionante, affidabile ed efficiente. Per esempio, l'eterogeneità dei dispositivi, tecnologie e modalità di interazione (*Machine-to-machine, machine-to-human e machine-to-cloud*) pone importanti sfide circa il processo di comunicazione. In questa prospettiva, una vasta gamma di tecnologie wireless a bassa potenza e a corto raggio, sono state progettate per fornire una connettività efficiente tra i diversi dispositivi IoT e la rete internet. Recentemente, anche le reti a lungo raggio, proprie dei cellulari, vengono prese in considerazione al fine di garantire l'*internetworking* desiderato con i dispositivi dell'IoT, soprattutto in virtù dei vantaggi offerti in termini di una maggiore copertura di rete, alta velocità di trasmissione dei dati, bassa latenza, basso costo per bit, alta efficienza dello spettro, etc.

In questo contesto, sono state introdotte importanti innovazioni nel settore delle comunicazioni proprio al fine di facilitare e ottimizzare la connessione dei dispositivi IoT alla rete: si tratta in particolare del wireless di quinta generazione (5G). Il 5G non sarà solo una pura e semplice evoluzione delle correnti generazioni di rete ma, più significativamente, una rivoluzione tecnologica nel settore dell'informazione e della comunicazione. In particolare, come evidenziato in figura 13, il 5G abiliterà la connessione di milioni di dispositivi per km^2 assicurando un'elevatissima efficienza energetica tale da consentire che la batteria degli oggetti IoT duri per periodi lunghi fino a 10 anni.

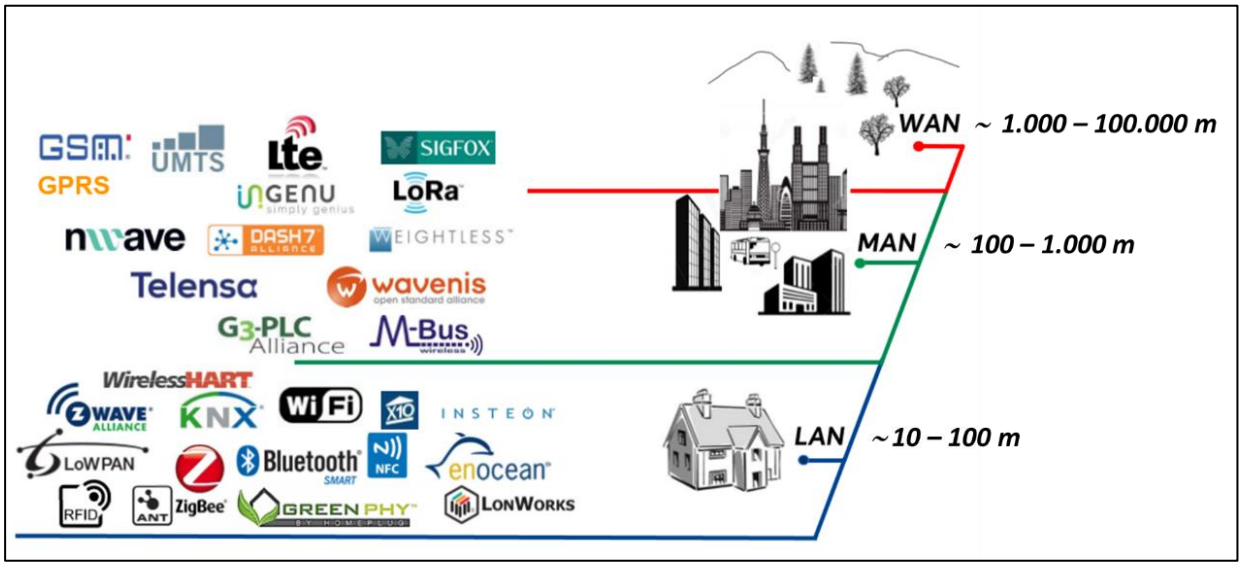
Figura 13. Requisiti 5G



Fonte: 5G Vision, 5GPP

L'immagine seguente offre un quadro sintetico delle diverse soluzioni di connettività per l'IoT per diversi *range* di copertura.

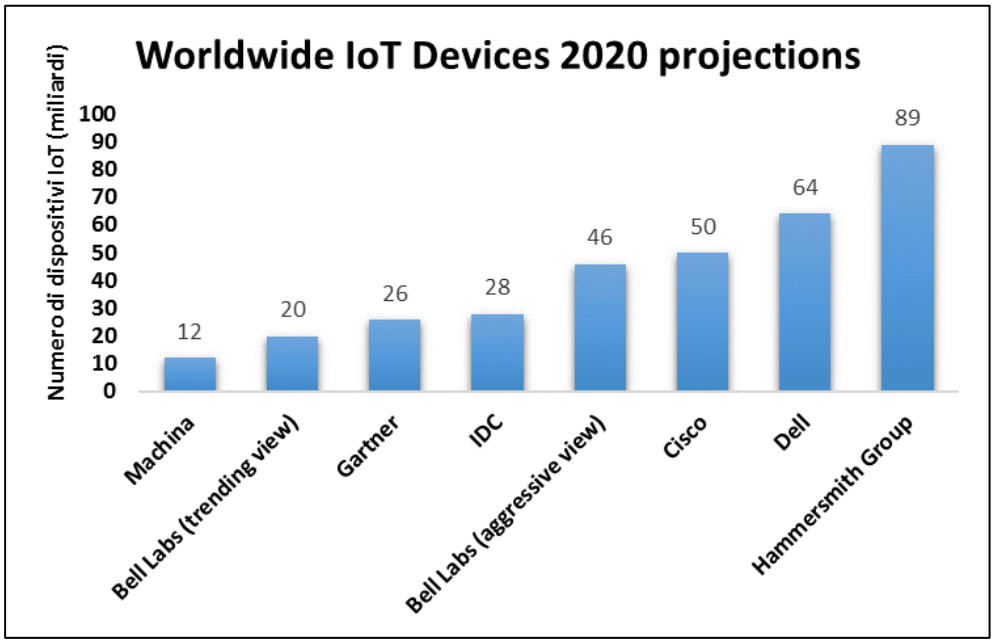
Figura 14. Soluzioni di connettività per IoT



Fonte: elaborazioni ITMedia Consulting su fonti varie

La crescita dell'IoT è oggetto di numerosi studi e previsioni in relazione al numero di *smart objects* che saranno connessi in rete nei prossimi anni. Come mostrato in figura 15 i numeri previsti variano in modo significativo ma tutti indicano una direzione comune: una vera e propria esplosione del numero di dispositivi connessi che si attesta su multipli di decine di miliardi a partire dal 2020.

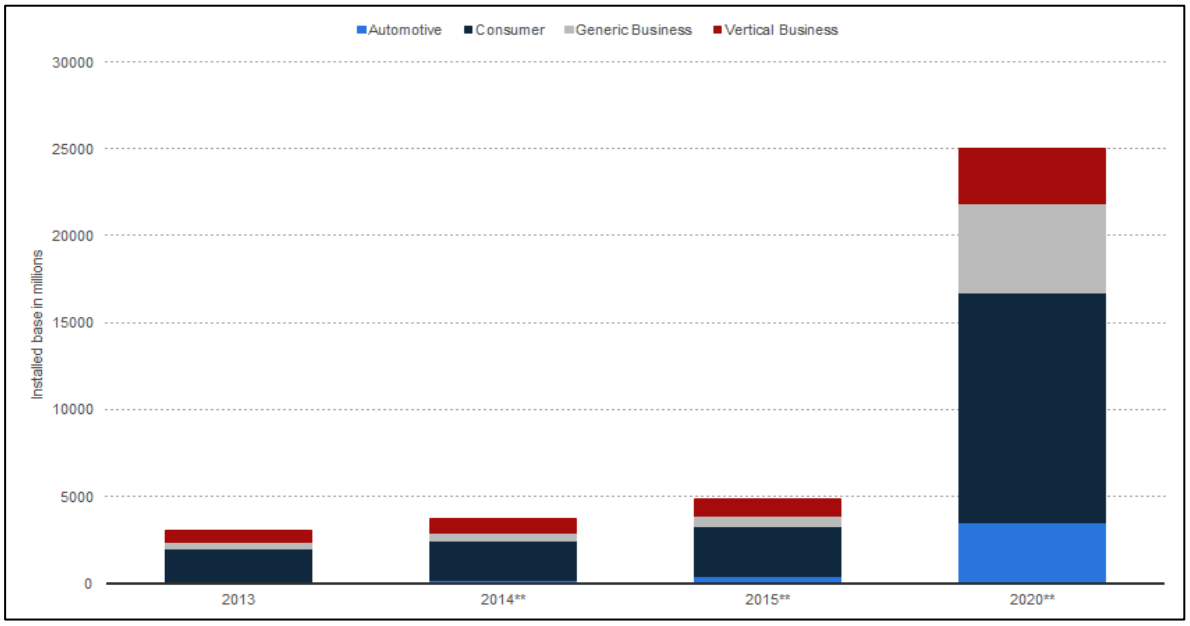
Figura 15. Previsioni sul numero di dispositivi IoT per il 2020



Fonte: elaborazioni ITMedia Consulting su fonti varie

In particolare, la figura 16 mostra la crescita, dal 2013 al 2020, dell'impiego dei dispositivi IoT rispetto a 4 settori (Automotive, Consumer, Generic business e Vertical business).

Figura 16. Dispositivi IoT installati per settore dal 2013 al 2020 (in milioni)



Fonte: Gartner

Come si evince dal grafico, il mercato dell'*automotive* sta rapidamente evolvendo e, secondo queste stime è destinato ad aumentare sempre più, verso la produzione di veicoli smart. L'IoT consente alle case automobilistiche di sviluppare mezzi di trasporto sempre più intelligenti e grazie alla collaborazione intersettoriale tra imprese di telecomunicazioni, aziende tecnologiche e case automobilistiche, tutto questo sta diventando realtà. Sarà infatti possibile sviluppare sistemi avanzati di assistenza alla guida, in modo tale da avere auto senza conducente, laddove l'infrastruttura stradale lo consenta, o di mettere a punto veicoli con un livello di sicurezza di gran lunga superiore all'attuale. Secondo l'Organizzazione mondiale della sanità, le auto autonome, dotate di sistema di sicurezza avanzata, potrebbero ridurre del 90% il numero di incidenti mortali su strada, oggi stimati nell'ordine di una cifra prossima a 1,2 milioni a livello globale. Per ridurre al minimo il rischio di incidenti le case automobilistiche dovranno garantire che le auto siano in grado di comunicare tra di loro, riconoscere l'ambiente circostante, adattarsi alle condizioni meteo, ricevere gli aggiornamenti del sistema di mappatura e respingere gli attacchi informatici. Alla base dell'auto intelligente ci sono ancora una volta i dati, raccolti in funzione del servizio richiesto. Quest'ultimo deve poi essere codificato da un processo che implica l'acquisizione e l'elaborazione dei dati, nonché una costante interazione tra tutti gli elementi che lo definiscono. In tal senso, risulta necessario disporre di un'infrastruttura di comunicazione con qualità trasmissiva adeguata e tempi di latenza residuali per poter sostenere aggiornamenti in tempo reale. Tutto ciò richiederà un'enorme potenza di elaborazione e genererà un volume di dati senza precedenti. Gestire questa immensa mole di dati - si stimano circa 4.000 gigabyte per auto ogni giorno - richiederà la connettività 5G e una robusta infrastruttura di data center. La tipologia di dati che vorranno essere acquisiti in funzione del servizio

erogato comporterà anche la definizione di un sistema di comunicazione che dovrà avvalersi di infrastrutture *over the cloud* e distribuite. In questo contesto, diventerà sempre più importante la comunicazione di prossimità, ovvero un'infrastruttura distribuita che consenta di ottimizzare al massimo lo scambio di dati. Ecco, quindi, le due diverse dimensioni di computing che sovrintendono un possibile processo *connected car*: il *cloud computing* e l'*edge/fog computing* ovvero l'elaborazione centralizzata e/o l'elaborazione di prossimità.

Figura 17. Il modello connected cars

Fonte: elaborazioni ITMedia Consulting su fonti web

I casi d'uso dell'IoT possono riguardare ogni settore, con prospettive notevoli di sviluppo di nuovi servizi innovativi, ottimizzazione di servizi esistenti e generazione di valore aggiunto per l'intera economia: ad esempio in ambito sanitario tramite applicazioni di gestione automatica delle malattie croniche, basate su sensori di parametri connessi in rete, si potrebbero realizzare cospicui risparmi, limitando il numero di visite in ospedale dei pazienti oltre a garantire un monitoraggio costante e affidabile; in ambito *smart city* si potrebbe rendere particolarmente efficiente la gestione delle risorse idriche attraverso applicazioni di monitoraggio e controllo di tutti i punti di raccolta, distribuzione e consumo di acqua implementando anche soluzioni di analisi dei trend della domanda basata sui dati storici raccolti.

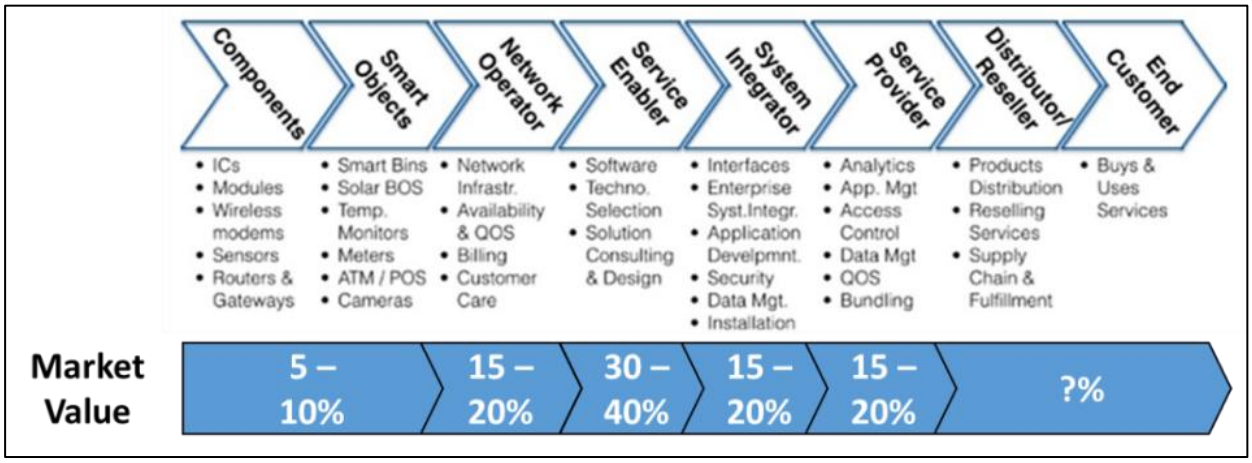
La figura seguente offre una panoramica dei principali settori coinvolti, nonché un'architettura logica di riferimento in cui si sottolinea il ruolo strategico della gestione dei dati (in particolare dei big data prodotti dagli *smart objects*) mediante soluzioni innovative di *analytics*.

Figura 18. Architettura logica e applicazioni dell'Internet of Things

Fonte: Italtel Group S.p.A.

Data l'eterogeneità di tutti gli elementi presenti in una tipica architettura IoT, la relativa catena del valore presenta numerosi attori che si contendono il valore di mercato. Nello specifico la componente più grande del valore di mercato è data dai "*System Integrator*" e dai "*Service enablers*", che rappresentano le tecnologie che danno accesso ai servizi e che insieme rappresentano dal 45% al 60% del valore totale. Il beneficio della maggior parte dei prodotti e dei servizi dell'IoT va quindi alle aziende che sviluppano e supportano soluzioni complete per i fornitori di servizi. Questo conferma l'intuizione sul fatto che l'industria dell'IoT è guidata principalmente dall'innovazione e dalle capacità fornite dai "*system integrators*" e, in misura minore, dalle aziende che sviluppano componenti di base per l'IoT e per gli oggetti intelligenti. I margini di profitto e la velocità media richiesta per convertire gli investimenti in R&S in ricavi variano considerevolmente da un segmento d'applicazione all'altro, soprattutto nella misura in cui il prodotto o servizio finale si debba conformare a regolamentazioni governative o a rigorosi e complessi standard di qualità e sicurezza.

Figura 19. Catena del valore per l'IoT



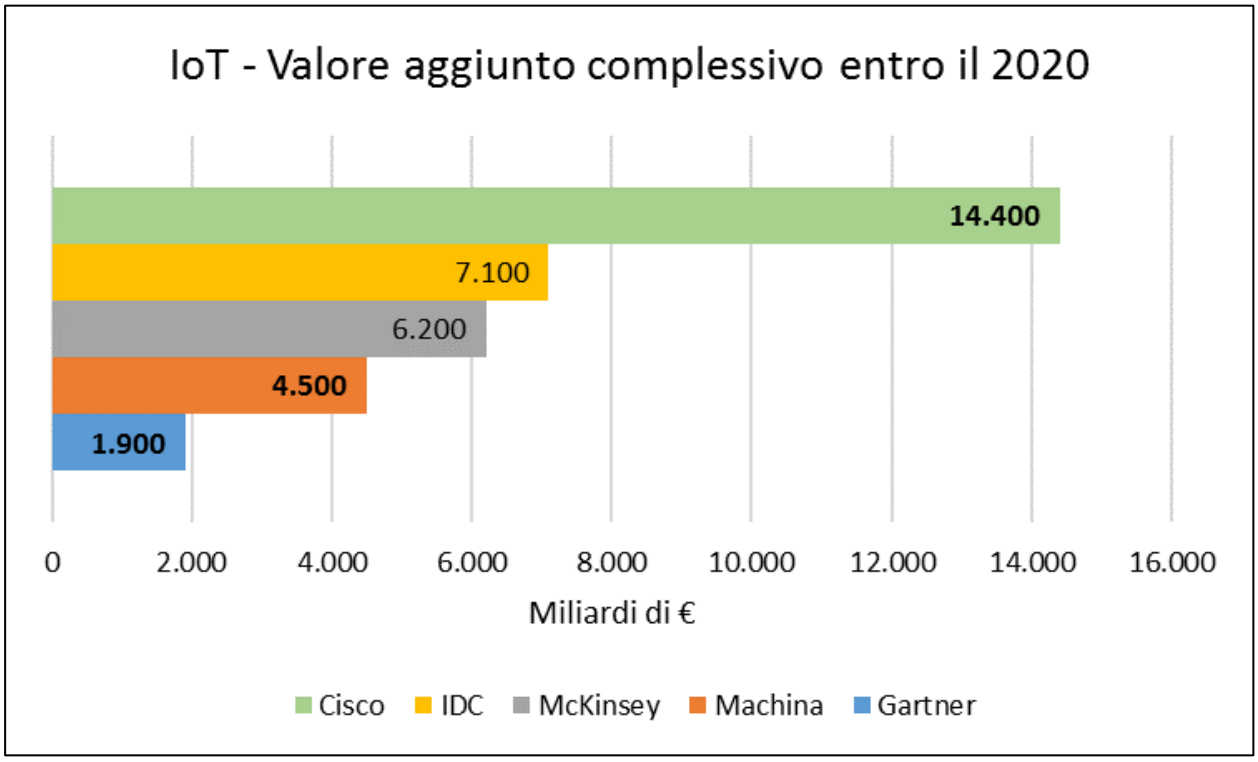
Fonte: Arthur D.Little

Le enormi potenzialità offerte dall'IoT potranno essere adeguatamente sfruttate in scenari caratterizzati da un elevato livello di sviluppo di:

- soluzioni cloud per la memorizzazione ed elaborazione dei dati (anche in ottica *edge cloud* per le applicazioni *mission critical*);
- soluzioni big data e *analytics* per estrapolare le informazioni essenziali per l'esecuzione ottimale delle applicazioni e dei servizi;
- soluzioni di *cybersecurity* per un uso appropriato e sicuro di dati e informazioni sensibili e particolarmente rilevanti per il corretto funzionamento di infrastrutture e/o interi settori, soprattutto nel segmento business e nell'ambito del settore pubblico. Miliardi di nuovi dispositivi connessi daranno luogo a nuove sfide amplificando le attuali problematiche di sicurezza (si pensi ad esempio a cyber attacchi ad auto connesse, impianti di produzione industriale connessi, dispositivi per il monitoraggio e la regolazione del traffico stradale, etc.)

Infine, anche in relazione al valore aggiunto complessivo previsto entro il 2020 per l'economia globale le stime, pur variando largamente, a seconda delle fonti (Cisco, IDC, McKinsey, Machina e Gartner), indicando inequivocabilmente un potenziale enorme dell'IoT per tutti i settori economici.

Figura 20. Valore aggiunto complessivo dell'IoT entro il 2020



Fonte: elaborazioni ITMedia Consulting su fonti varie

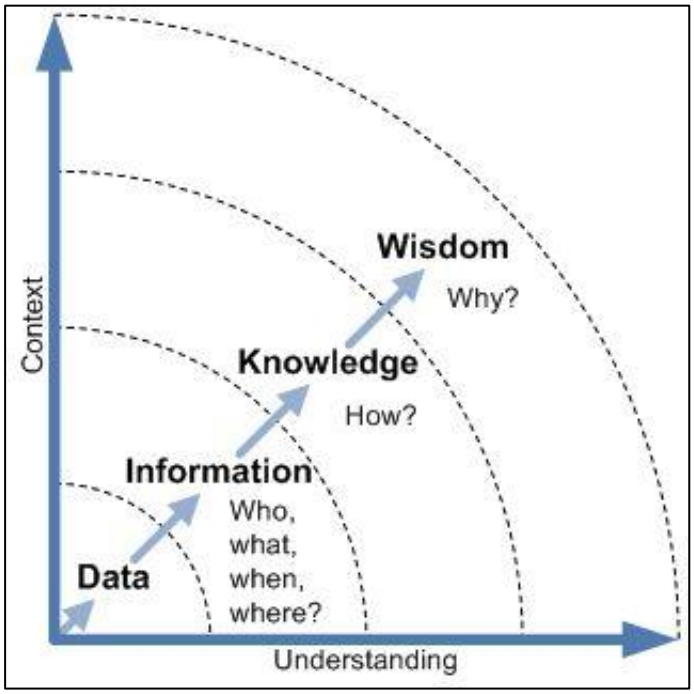
1.1.3. Intelligent Automation: AI, AugI, AR e VR

Le sorprendenti e inarrestabili evoluzioni precedentemente descritte sono alla base di una vera e propria rivoluzione in merito al futuro dell'informazione. Secondo il ciclo DIKW (Data – Information – Knowledge – Wisdom) che plasma l'impostazione seguita da coloro che studiano knowledge management e computer science, l'informazione è alla base dei processi cognitivi e decisionali di qualsiasi organizzazione e/o individuo.

In tale ciclo si distinguono:

- *Data:* tutto ciò che viene direttamente osservato e rilevato (misure, dati grezzi, etc.)
- *Information:* l'individuazione di relazioni, ad esempio strutturali e/o funzionali, tra i dati;
- *Knowledge:* l'individuazione di modelli che descrivono il significato delle informazioni;
- *Wisdom:* abilità di aumentare l'efficacia della conoscenza aggiungendovi valore tramite la capacità di giudizio.

Figura 21. Il ciclo dei processi cognitivi



Fonte: IT Infrastructure Library (ITIL) v.3, Office of Government Commerce (OGC)

L'acquisizione dell'informazione richiede l'abilità di catturare i dati, computarli ovvero effettuare elaborazioni più o meno complesse su un insieme parziale o totale di dati e infine comunicare i risultati. Il costo dell'informazione può essere quindi attribuito al costo di queste "3 C". Grazie allo sviluppo di internet il costo della trasmissione dei dati si è ridotto drasticamente. Inoltre, la diffusione su scala globale del paradigma del *Cloud Computing*, e in particolare della sua integrazione con le reti in ottica *Cloud Integrated Network*, ha ridotto in maniera analoga il costo della computazione. Dunque, complessivamente il costo dell'informazione si è drasticamente ridotto.

La raccolta dei dati, che fino a poco tempo fa sembrava essere rimasta l'ultima barriera da superare, oggi viene progressivamente eliminata grazie all'avvento dell'IoT e alla digitalizzazione di ogni oggetto. Come affermato nei paragrafi precedenti, stiamo assistendo infatti alla crescita esponenziale del traffico dati, gran parte dei quali sono disponibili gratuitamente (si pensi ad esempio ai dati pubblicati dalle Pubbliche Amministrazioni, i c.d. *Open Data*, relativi a una moltitudine di aspetti geografici, anagrafici, demografici, infrastrutturali, etc. particolarmente rilevanti per lo sviluppo di servizi, così come ai dati resi disponibili dagli utenti i c.d. *User Generated Content*, in prima istanza video/foto ma anche condivisione di posizioni, commenti, recensioni etc., parimenti importanti per la realizzazione di applicazioni e servizi).

L'effetto combinato della riduzione dei costi e dell'aumento travolgente dei dati e delle risorse accessibili porta allo sviluppo di una quarta "C" relativa alla "Cognizione" ovvero alla possibilità che un'organizzazione o un individuo possa avere sempre più informazioni pertinenti e rilevanti per un dato contesto, applicazione o compito ricevendo supporto anche da agenti intelligenti (sia sotto forma di applicazioni software che di veri e propri automi/robot) che possono aiutare o addirittura

sostituire l'uomo nello svolgimento di compiti non solo materiali ma anche cognitivi richiedenti capacità di ragionamento e giudizio.

Si tratta della c.d. Intelligenza Artificiale (*Artificial Intelligence* - AI) ovvero dell'abilità di un computer di svolgere funzioni e ragionamenti tipici della mente umana. Nel suo aspetto puramente tecnologico, essa comprende la teoria e le tecniche per lo sviluppo di algoritmi che consentono alle macchine di mostrare un'abilità e/o attività intelligente, almeno in domini specifici. Tra gli attuali e principali campi di applicazione vi sono gli assistenti virtuali (come ad esempio Alexa di Amazon, Cortana di Microsoft, Siri di Apple, Google Home, Google Assistant e Google Now nonché il più recente Bixby di Samsung) che offrono assistenza su richieste generiche, richieste contestualizzate per ottenere previsioni su viaggi, pianificazione attività, etc.

Questo mercato già affollato appare particolarmente vivace in termini di investimenti, partnership e operazioni di M&A. Significativo in tal senso è l'accordo siglato a fine agosto da Amazon e Microsoft per integrare i rispettivi assistenti virtuali. In tal modo, si potrà utilizzare l'assistente Alexa tramite Cortana, ad esempio su un computer Windows, e Cortana tramite Alexa, ad esempio utilizzando gli speaker Echo di Amazon. Dal punto di vista strategico, l'integrazione tecnologica verte sul fatto che i due assistenti virtuali presentano specificità diverse. Così, mentre gli utenti di Amazon potranno usare Cortana per controllare l'agenda di lavoro o farsi leggere un'email, viceversa, con Alexa si potrà ascoltare musica e fare shopping online. Se Microsoft ha tutto l'interesse a spingere Cortana fuori dal raggio d'azione dei pc, verso le auto che si guidano da sole o verso voci come quelle di Alexa, quest'ultima interagirà meglio nell'Outlook di Microsoft, che consente di leggere le email o creare agende e calendari di lavoro. In questo modo, Amazon avrà accesso ai 500 milioni di utilizzatori di Windows 10, allargando così la platea dei suoi utenti.

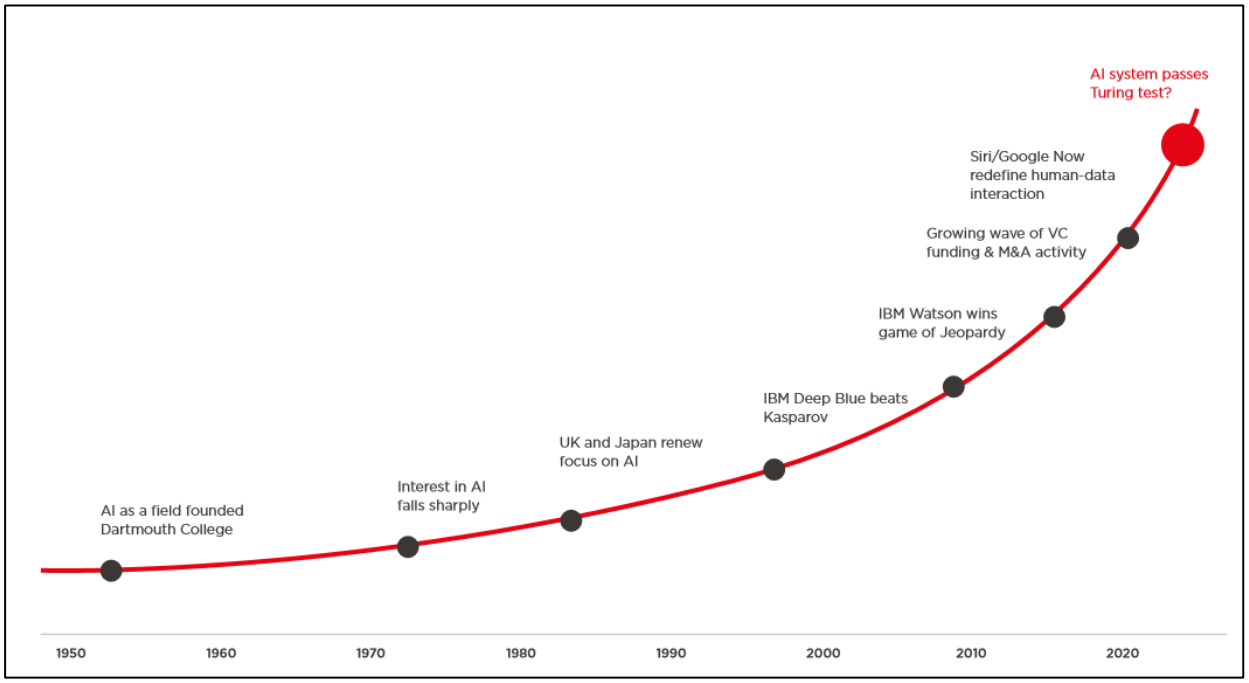
Un altro ambito di applicazione in crescita è quello relativo alle Conversational Interfaces, e in particolar modo ai Chatbot, considerati alla base di una potenziale rivoluzione del mercato enterprise grazie alle interessanti prospettive di risparmio di tempi e costi offerte dalle opportunità di automazione. Si tratta di applicazioni incorporate nei servizi di messaggistica, come Messenger di Facebook, capaci di guidare autonomamente gli utenti attraverso una serie di domande nello svolgimento di diversi compiti. Nello specifico, si prevede che queste applicazioni produrranno nuovo valore condiviso tra le complesse operations all'interno delle aziende e all'esterno in ambito B2B e B2C, spaziando su tre fonti:

- *Automazione*, tramite la sostituzione tra macchine e persone nelle attività ripetitive e ridondanti;
- *Semplificazione*, per generare livelli maggiori di adozione e usabilità dei sistemi e dei processi aziendali;
- *Innovazione*, abilitando nuovi servizi basati sull'analisi automatica degli insights aziendali in modalità non sostenibili dalle persone.

Sul versante B2C, attualmente la maggior parte dei chatbot è focalizzata sulle operazioni di customer service; tuttavia aziende leader come Budweiser, General Electric, Nissan, Pizza Hut, e Vodafone stanno portando avanti numerose sperimentazioni relative all'automazione delle campagne di

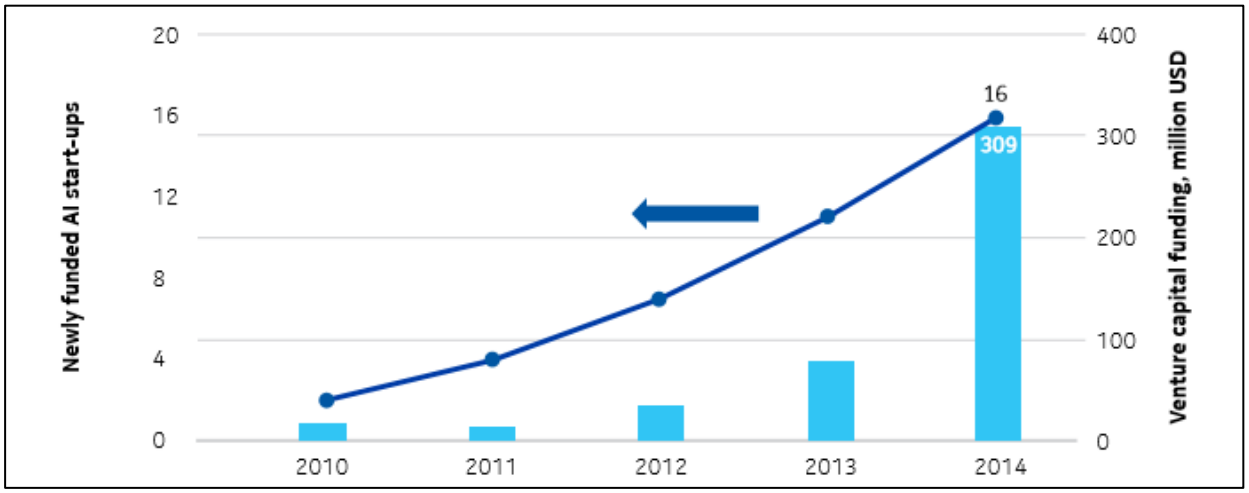
marketing finalizzate a identificare nuovi clienti e valorizzare il proprio brand. In ambito B2B, tali prodotti abiliteranno l'automazione delle attività ripetitive dei dipendenti includendo anche alcune operazioni effettuate con i sistemi ERP. Tale automazione accelera l'adozione dei sistemi IT generando risparmi fino al 18% del costo totale dell'implementazione pari ad alcune centinaia di migliaia di dollari. Le immagini seguenti presentano un quadro dell'evoluzioni nel campo dell'Intelligenza Artificiale anche in relazione al trend degli investimenti.

Figura 22. Evoluzioni nell'ambito dell'Intelligenza Artificiale



Fonte: GSMA Intelligence, Global mobile trends

Figura 23. Investimenti in ambito AI

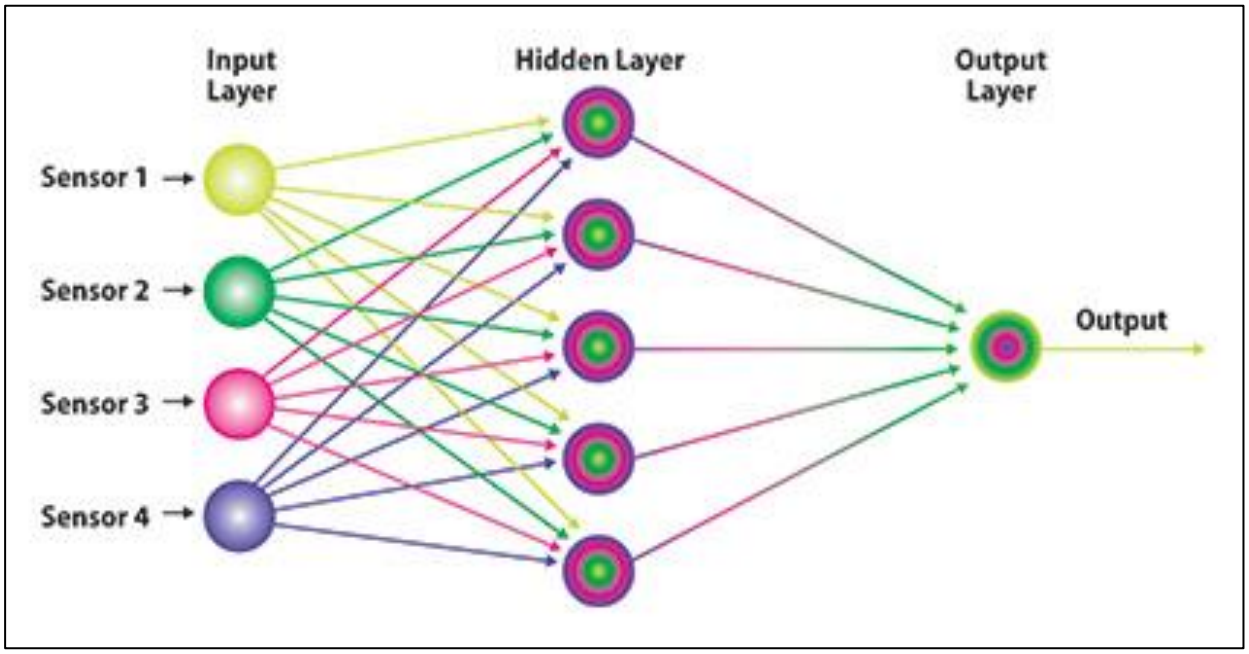


Fonte: The Future X Network: A Bell Labs Perspective

Tra le principali tecniche di AI vi è il *Machine Learning*. Si tratta di tecniche di apprendimento automatico che si occupano della realizzazione di sistemi e algoritmi che si basano su osservazioni come dati per la rappresentazione di nuovi contenuti informativi. L'apprendimento può avvenire catturando caratteristiche di interesse provenienti da esempi, strutture dati o sensori, per analizzare e valutare le relazioni tra le variabili osservate.

Le tecniche di apprendimento automatico sono implementate per mezzo di reti neurali artificiali che emulano il comportamento del cervello umano per definire le connessioni tra i vari neuroni artificiali e la loro attivazione consentendo di ottenere prestazioni elevate in applicazioni quali riconoscimento vocale, riconoscimento di immagini, elaborazione del linguaggio naturale, etc.

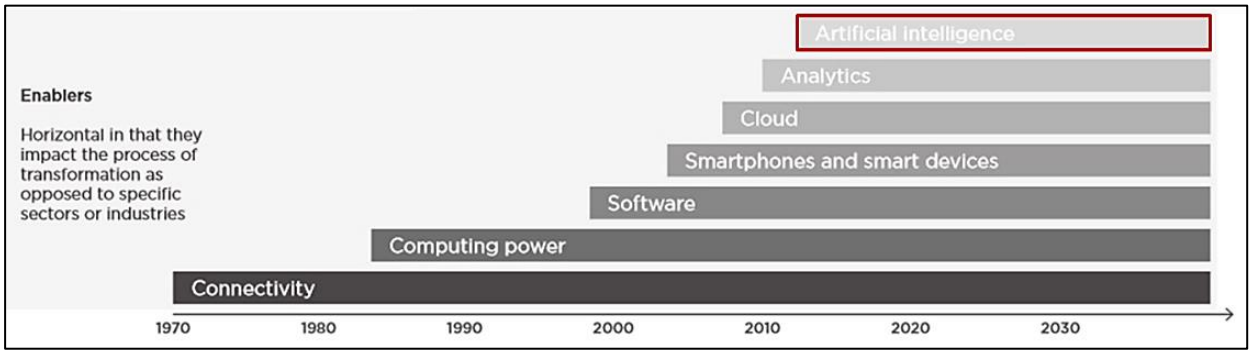
Figura 24. Esempio di Rete Neurale Artificiale



Fonte: elaborazioni ITMedia Consulting su fonti varie

L'Intelligenza Artificiale sta emergendo come il principale catalizzatore che accelererà lo sviluppo di settori e applicazioni innovative come *connected car* e *smart home*. Infatti, tra i diversi abilitatori tecnologici, essa è considerata come una sorta di super abilitatore di numerose applicazioni come riportato nella figura seguente.

Figura 25. Abilitatori tecnologici



Fonte: GSMA Intelligence, Global mobile trends

Il crescente interesse verso l'Intelligenza Artificiale è legato alla prospettiva di sostituzione degli esseri umani in vari ruoli e compiti. Tale prospettiva si basa sulla convinzione che man mano che saranno utilizzate nella risoluzione di problemi complessi, tali applicazioni evolveranno sempre più grazie all'apprendimento automatico dai dati e dagli input coinvolti.

Tuttavia, numerosi sono i dubbi circa le reali possibilità che queste applicazioni, orientate in tale ottica alla creazione della cd. "conoscenza perfetta", possano davvero sostituirsi alle persone.

Per tale motivo, in letteratura si parla di un nuovo approccio, definito *Augmentend Intelligence* (AugI) orientato ad aumentare l'intelligenza e la conoscenza di un individuo al fine di supportarlo nel processo decisionale e nell'acquisizione di ulteriore conoscenza, senza che un sistema artificiale lo sostituisca.

L'obiettivo di queste tecniche è quello di fornire potenziali risposte o possibili percorsi di approfondimento di un problema piuttosto che mostrare un elenco di risposte ordinate per rilevanza che tuttavia non consentono di risolvere la problematica posta.

Le applicazioni di AugI funzionano come un partner personalizzato che offre supporto presentando risposte, spunti e suggerimenti rilevanti basati sull'accesso a dati e informazioni esterne e soprattutto operando sotto la guida dell'utente che potrà dinamicamente valutare la rilevanza di quanto suggerito attraverso modalità di interazione sofisticate e al tempo stesso semplici finalizzate anche a "educare" la macchina/applicazione. Le risposte e i suggerimenti vengono determinati sulla base di algoritmi che calcolano e visualizzano la probabilità associata con differenti risposte potenziali aiutando a comprendere perché una data risposta è stata suggerita o meno. Inoltre, grazie all'apprendimento supervisionato dagli utenti, al crescere del loro utilizzo queste applicazioni diventeranno maggiormente precise e affidabili tendendo al comportamento ideale atteso per questo tipo di sistemi.

La tabella seguente offre un quadro sinottico comparativo delle principali caratteristiche dell'*Artificial Intelligence* e dell'*Augmented Intelligence*.

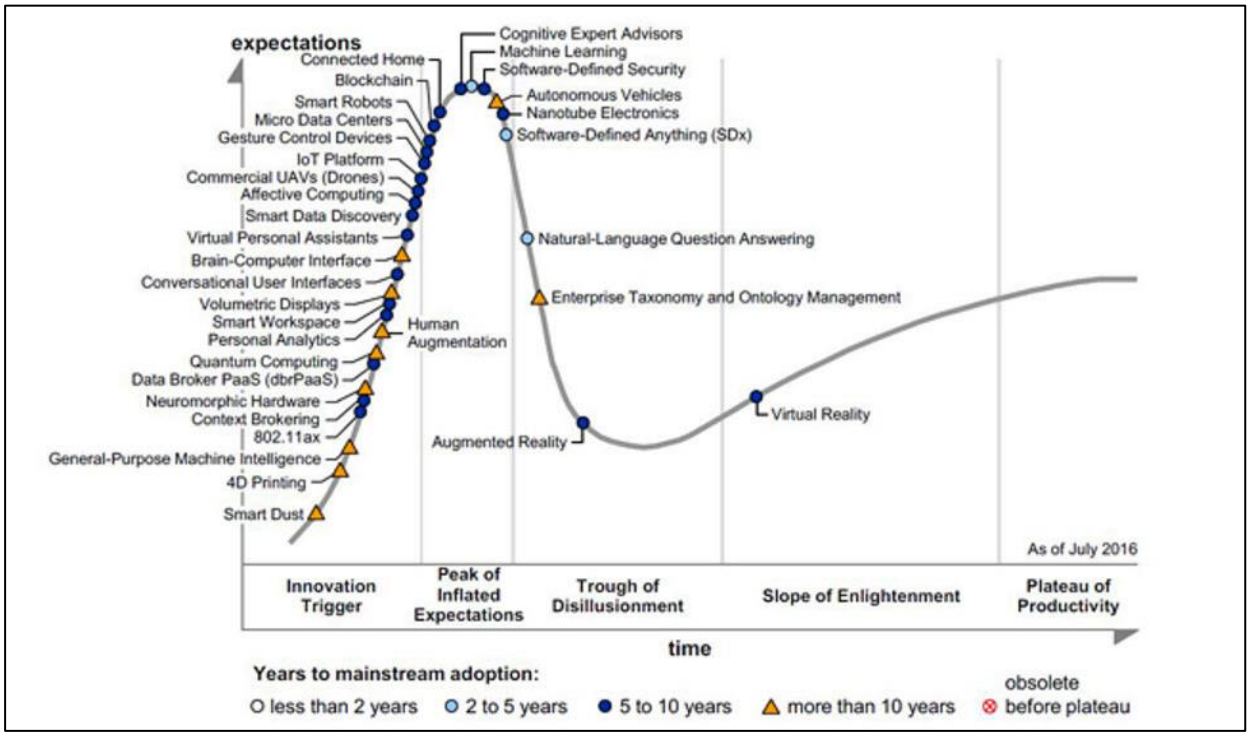
Tabella 2. Artificial Intelligence e Augmented Intelligence a confronto

	Artificial Intelligence (AI)	Augmented Intelligence (AugI)
Focus	Macchine/sistemi pensanti che possono percepire la realtà e la situazione in cui sono interagendovi	Supportare e potenziare la capacità di pensiero delle persone (processi di decision-making, etc.)
Metodo	Reti neurali e logica matematica	Algoritmi di computazione di risposte rilevanti su base probabilistica
Aspetti chiave	Autonomia	Aumento dell'intelligenza, personalizzazione e interazione
Sfide	Comprensione del modo in cui gli esseri umani pensano	Algoritmi e metodi di interazione
Output	Sistemi e agenti automatici intelligenti	Agenti/sistemi che partecipano come partner nel processo cognitivo dell'utente
Applicazioni	Robotica, automazione industriale, domestica, etc.	"Assistenti" personali

Fonte: ITMedia Consulting

Nell'ambito delle tecnologie relative ai sistemi/agenti intelligenti, tra i campi di applicazione più interessanti e attuali a livello di trend (si veda l'immagine relativa all'Hype Cycle di Gartner), vi sono le applicazioni di realtà virtuale (*Virtual Reality* – VR) e realtà aumentata (*Augmented Reality* – AR).

Figura 26. Gartner Hype Cycle for emerging technologies, 2016

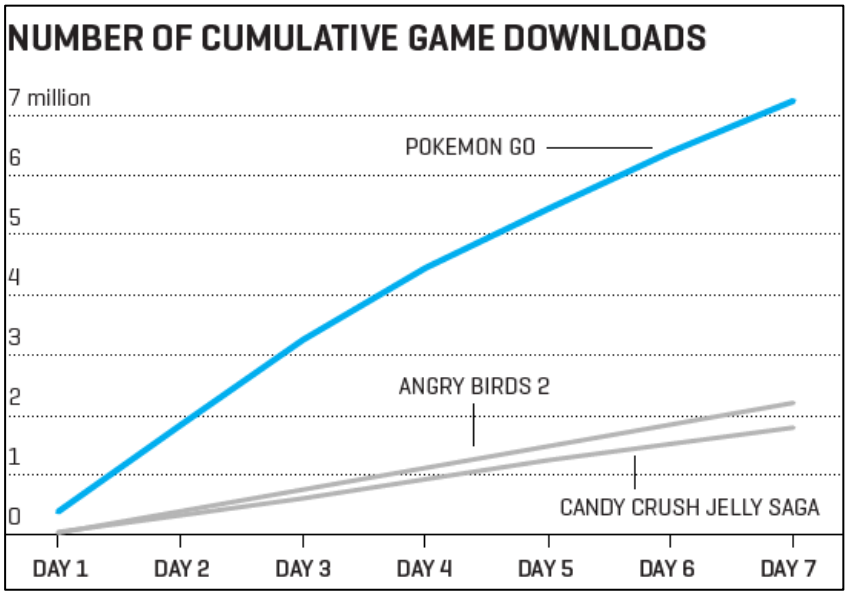


Fonte: Gartner

In tali applicazioni confluiranno anche altre evoluzioni tecnologiche relative a: nuove modalità di UX – User Experience, piattaforme di contenuti, aspetti di sicurezza e robustezza applicativa, connettività veloce, ubiqua e con latenza minima grazie al 5G e all'*edge computing*, sensoristica e IoT, dispositivi sempre più performanti e meno costosi, etc.

Sono previste applicazioni sia consumer che business con un focus iniziale sul B2C, specialmente in ambito marketing, e sul *gaming*. Attualmente il caso maggiormente emblematico del potenziale di tali applicazioni è rappresentato dalla mobile Game Pokemon Go, videogioco di tipo free-to-play basato sulla realtà aumentata geolocalizzata con GPS. Il suo imponente successo in termini di numero e caratteristiche di utenti e rapidità di crescita testimonia che l'AR è semplice e può essere a portata di tutti, non solo delle generazioni più digitalizzate come i *millennials* o i nativi digitali.

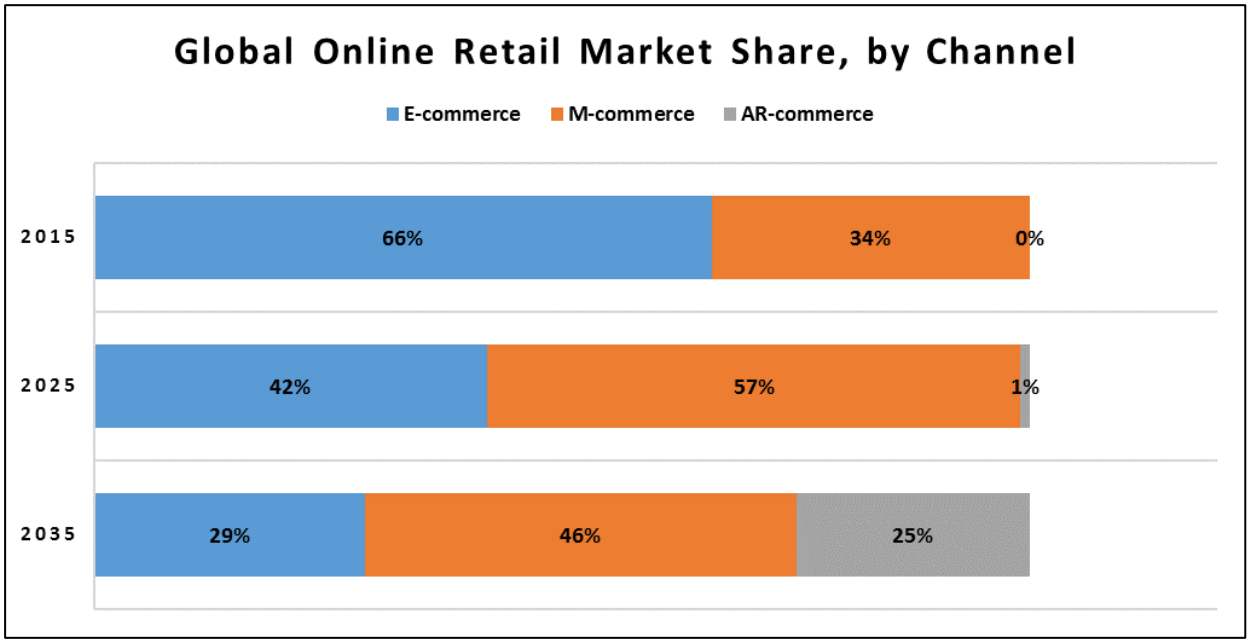
Figura 27. Crescita di Pokemon Go



Fonte: Sensortower estimates, Data that Drive App Growth

Diverse sono le previsioni sul mercato dell'AR. Secondo Juniper Research tale mercato raggiungerà oltre 2,3 miliardi di App scaricate entro il 2021 specificando che queste non riguarderanno solo i videogame. Anche nell'ambito delle vendite online, l'AR avrà un'influenza sempre maggiore passando dal 1% sul totale del 2025 al 25% del 2035.

Figura 28. Crescita AR



Fonte: ITMedia Consulting su fonti varie

Inoltre, secondo IDC, nel 2017, il 30% delle aziende Fortune Global 2000 sperimenteranno soluzioni di AR/VR nelle loro campagne di marketing. Tali soluzioni abiliteranno nuove e più potenti interfacce per il *customer engagement* che saranno adeguatamente capitalizzate a partire dal 2021, termine entro il quale si prevede l'adozione di massa delle tecnologie AR/VR.

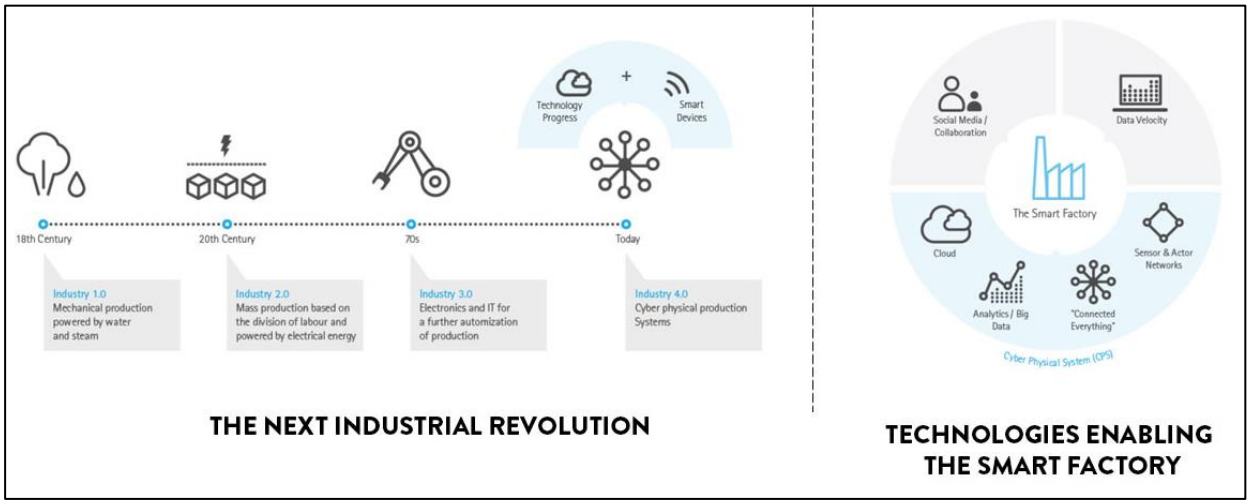
Molti grandi player internazionali stanno guardando a questo settore con particolare interesse. Microsoft sta investendo nello sviluppo degli HoloLens un Head-mounted display in grado di arricchire la realtà, completandola con nuovi dettagli attraverso le sue lenti. L'aggiunta di due telecamere posteriori all'iPhone 7, insieme all'acquisizione avvenuta nel 2013 della azienda israeliana PrimeSense che ha sviluppato il Kinect per la Xbox, indica la direzione di Apple di penetrare la tecnologia AR facendo leva sull'ecosistema dei propri smartphone. Inoltre, anche Google ha pubblicato da poco sul suo Play Store l'applicazione Tango, che sempre attraverso gli smartphone porta l'utente in un mondo alieno e misterioso mescolato con la nostra realtà. Infine, Sony insieme a Nolan ha annunciato un casco dedicato ai motociclisti dotato del sistema N-Com ARX, in grado di proiettare un ologramma sull'head-up display nel campo visivo del motociclista, che riesce a fornire utili informazioni per il supporto alla sicurezza e al comfort di guida.

1.2. Industry 4.0: nuove prospettive economiche, impatti sul mercato e sulle imprese

Le tecnologie digitali sopra analizzate avranno un impatto profondo nell'ambito di quattro direttrici di sviluppo: la prima, riguarda l'utilizzo dei dati, la potenza di calcolo e la connettività, e si declina in *Big data, Internet of Things e cloud computing* per la centralizzazione delle informazioni e la loro conservazione. La seconda è quella degli *Analytics*: una volta raccolti i dati, è necessario ricavarne valore. La terza direttrice di sviluppo è l'interazione tra uomo e macchina, che coinvolge le interfacce "*touch*", sempre più diffuse, e la realtà aumentata. Infine c'è tutto il settore che si occupa del passaggio dal digitale al "reale", e che comprende la manifattura additiva, la stampa 3D, la robotica, le comunicazioni, le interazioni *machine-to-machine* e le nuove tecnologie per immagazzinare e utilizzare l'energia in modo mirato, razionalizzando i costi e ottimizzando le prestazioni. All'interno di questo scenario, si sviluppa dunque un nuovo paradigma industriale, che viene comunemente definito Industria 4.0.

In particolare*, Industry 4.0* o Industria 4.0. è la definizione basata sull'evoluzione di internet, dal web 1.0, fino ai nostri giorni. In termini più operativi, l'industria 4.0 comporta l'organizzazione di processi di produzione basati sulla tecnologia e su *device* in grado di comunicare autonomamente tra di loro lungo la catena del valore: ossia un modello di *smart factory* dove i sistemi controllati dai computer gestiscono processi fisici, creando un mondo virtuale e parallelo a quello fisico.

Figura 29. L'industy 4.0



Fonte: Accenture

Sostanzialmente, le fabbriche dovrebbero diventare luoghi cyber-fisici, costantemente connessi, in cui il mondo reale e il mondo digitale si integrano: la produzione, così come è stata a lungo immaginata, comprenderà un'interazione diretta *machine to machine* (M2M) oltre che "uomo-macchina" e le nuove tecnologie integreranno gli oggetti, trasformandoli in sistemi intercomunicanti e dotati di "intelligenza". La disponibilità enorme di dati ed informazioni, peraltro, sarà in grado di influenzare indistintamente il processo produttivo, i modelli di business e la creazione di prodotti, consentendo inoltre la possibilità di prevedere il fallimento degli eventi e di abilitare l'impresa ad auto-configurarsi al fine di adattarsi ai continui cambiamenti.

In definitiva, dal quadro qui delineato emerge come l'innovazione sia contemporaneamente sinonimo di opportunità e di "distruzione". Infatti, se da un lato i vantaggi legati alle nuove trasformazioni sono senza dubbio evidenti, dall'altro non mancano elementi di criticità, legati essenzialmente al fatto che la quasi totalità dei tradizionali business sono nati e si sono sviluppati nella fase pre-digitale. Conseguentemente risulta particolarmente complesso adattare strutture economiche tradizionali alla nuova realtà digitale, riuscendo a connettere *asset* fisici ad internet o al *cloud*, ovvero utilizzare interfacce e infrastrutture proprie della rete digitale.

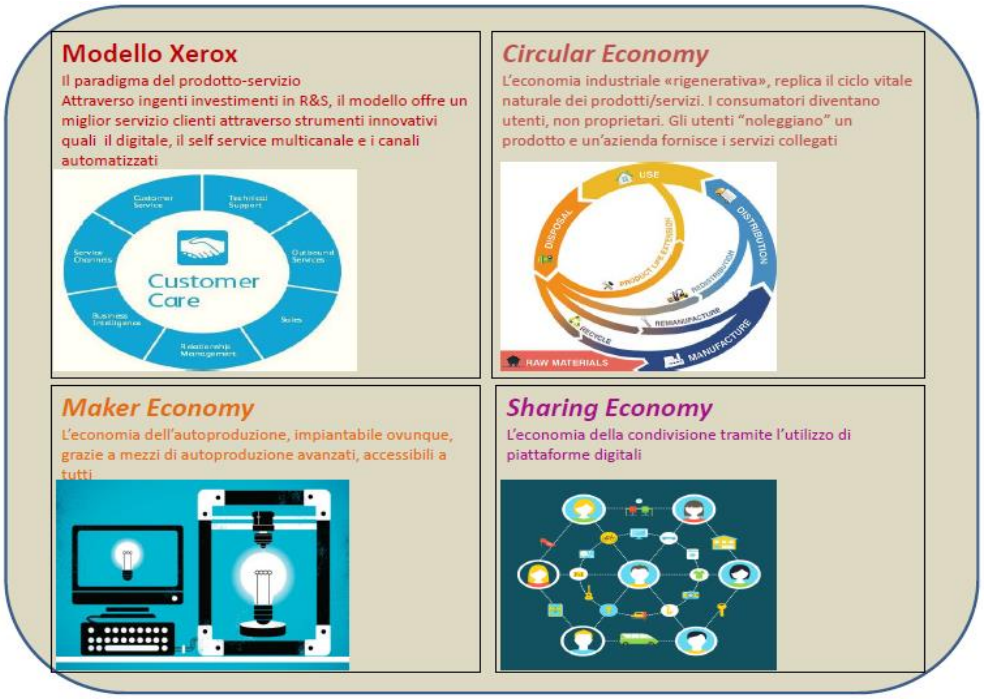
Molti di questi business sono rimasti completamente esclusi dal mondo digitale, o comunque sono destinati a diventare business secondari, privi di una intelligenza/conoscenza completa. La natura del prodotto di un'impresa – modalità di produzione, consegna e consumo – è essenzialmente influenzata dalla rete che a sua volta determina quanto velocemente l'impresa può espandersi: il passaggio dall'analogico al digitale, dall'hardware al software, dal materiale al virtuale, ha portato a prodotti che possono essere realizzati in massa o semplicemente copiati, che possono essere spediti automaticamente da magazzini localizzati o più semplicemente scaricati o fruibili in modalità streaming, e utilizzati ovunque tramite applicazioni istallate su dispositivi mobili. Queste trasformazioni digitali hanno modificato il modo in cui le imprese sono destinate a operare.

Ne consegue che l'attuale trasformazione è certamente dipendente dall'infrastruttura, ma determina soprattutto una trasformazione economica della realtà imprenditoriale. Il mondo delle imprese, può costituire un elemento in grado di rallentare lo sviluppo della "*digital economy*", in quanto ancora strettamente legato alla realtà fisico/analogica, anche se la situazione è destinata a cambiare radicalmente.

Secondo uno studio Nokia[4], si stima che le imprese digitali, diventeranno vere e proprie imprese "virtuali", intendendo con questo termine le possibili realtà imprenditoriali che potrebbero svilupparsi nei prossimi decenni e che saranno caratterizzate da efficienza, velocità di adattamento agli scenari di mercato in continua evoluzione e del tutto indipendenti da qualsiasi limite di carattere fisico e spaziale. Le imprese saranno in grado si estendersi oltre la propria infrastruttura fisica, trasformando e digitalizzando la propria offerta e il proprio patrimonio aziendale, dunque perdendo qualsiasi connotazione fisica o materiale che sia.

Le imprese "virtuali" saranno dotate di auto-apprendimento, capacità intuitive e abilità cognitive, tali da meglio organizzare la propria forza lavoro. Le aziende che adotteranno questa nuova visione implementeranno nuovi modelli di business "out of the site", aumenteranno la produttività della forza lavoro, coinvolgeranno sempre più il consumatore nella propria attività e diminuiranno i costi operativi. Tali fattori innovativi creeranno nuove forme di valore e saranno in grado di ridefinire la competitività globale e locale dell'era digitale.

Figura 30. I nuovi modelli di business nella digital economy

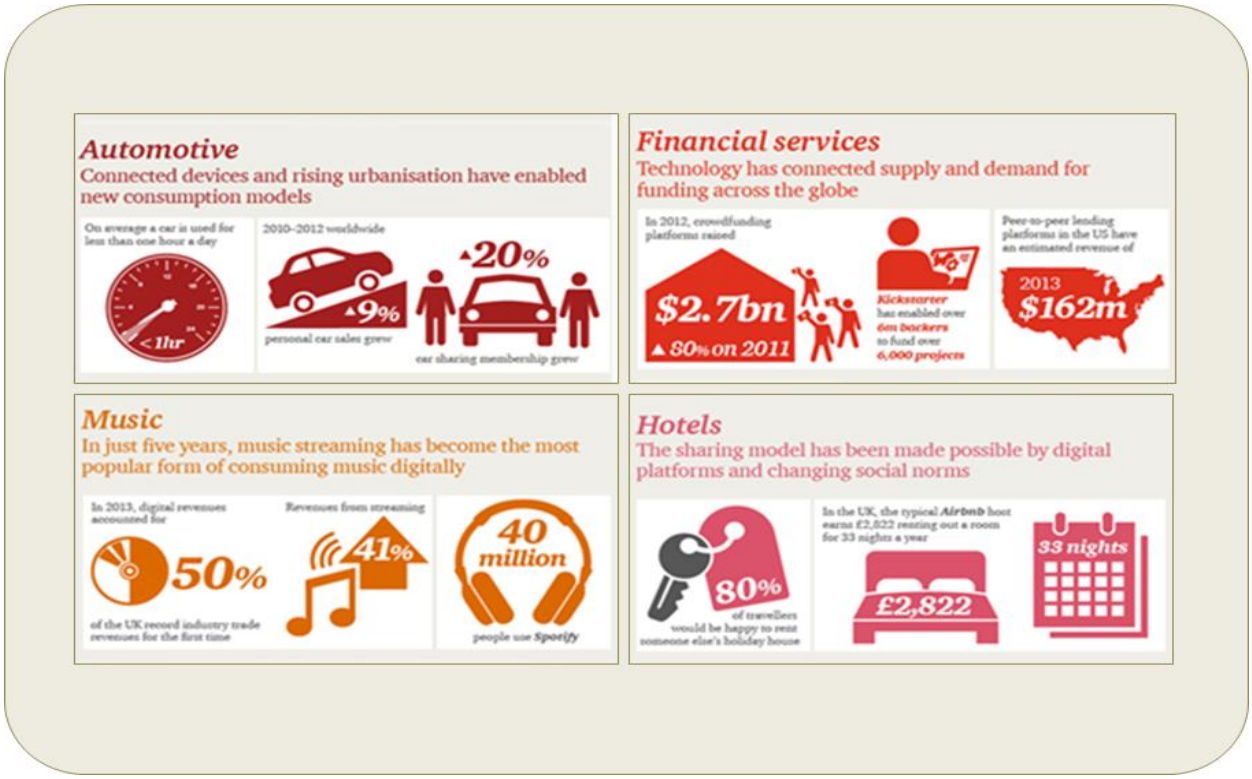


Fonte: elaborazioni ITMedia Consulting su fonti varie

[4] Nokia, *The Future X Network: A Bell Labs Perspective*, 2016.

Questo profondo cambiamento si riflette nella crescente consapevolezza da parte delle imprese di doversi necessariamente trasformare. Realtà quali Airbnb, Spotify e Uber hanno già "sconvolto" i settori in cui si sono insediati, il turismo, l'industria della musica e i trasporti, in pochissimi anni. Il loro vantaggio deriva dalla capacità di raggiungere la domanda su scala globale e di collegarsi costantemente con il consumatore finale attraverso applicazioni istallate sui rispettivi dispositivi mobili. Il loro successo e, in generale, il successo di gran parte di queste nuove realtà aziendali esercita una forte pressione sui business tradizionali.

Figura 31. Applicazione del modello sharing economy

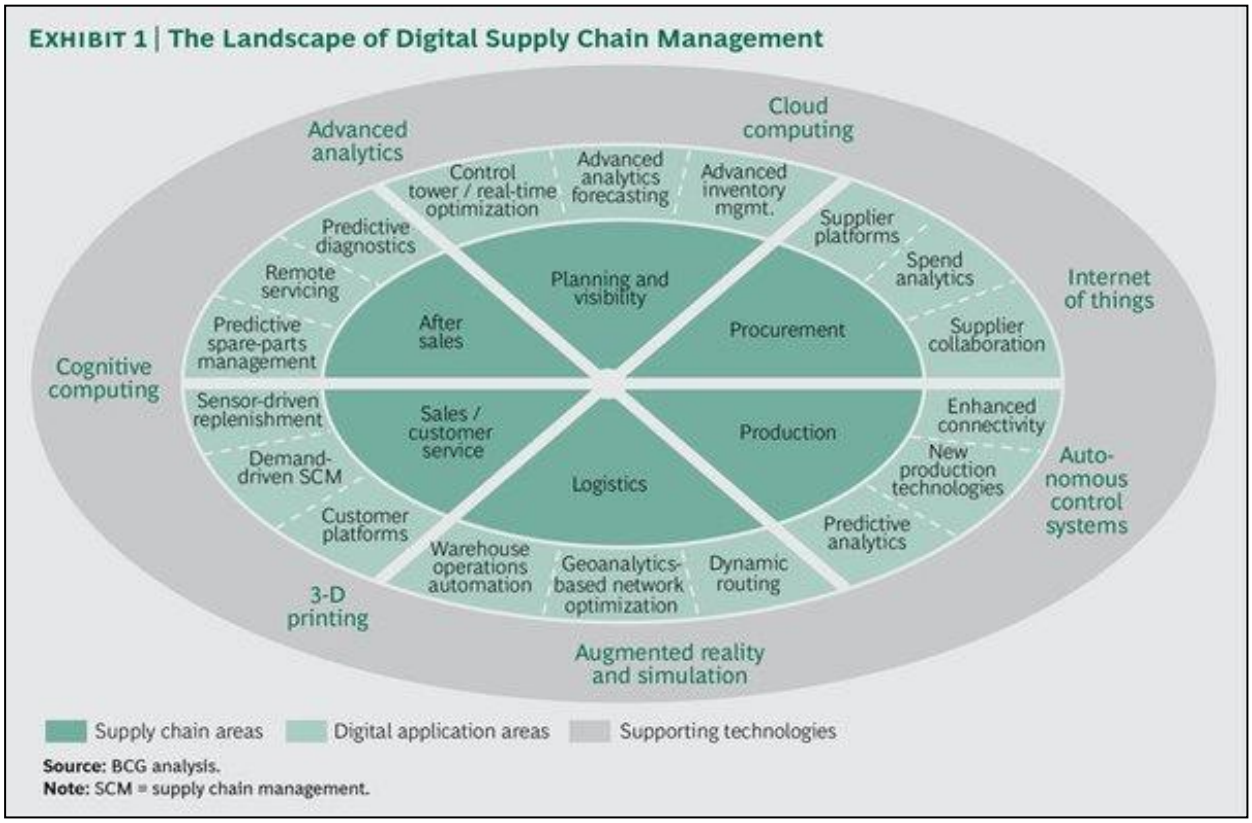


Fonte: PWC

Le aziende, dunque, se da un lato riconoscono il valore strategico dell'Industria 4.0, dall'altro sono consapevoli delle importanti sfide che dovranno affrontare per integrare le nuove tecnologie. Per costruire e mantenere un vantaggio competitivo, le aziende hanno bisogno di ampliare e approfondire le loro conoscenze pratiche sulle tecnologie digitali e le relative modalità d'uso. Proprio questo nuovo sapere fornisce le basi per lo sviluppo e l'attuazione delle strategie di produzione digitale.

La gestione "digitale" della *supply chain* ha maturato e sta generando un valore significativo. Il grafico seguente sulla *digital revolution,* mostra schematicamente lo scenario della *digital supply chain management.*

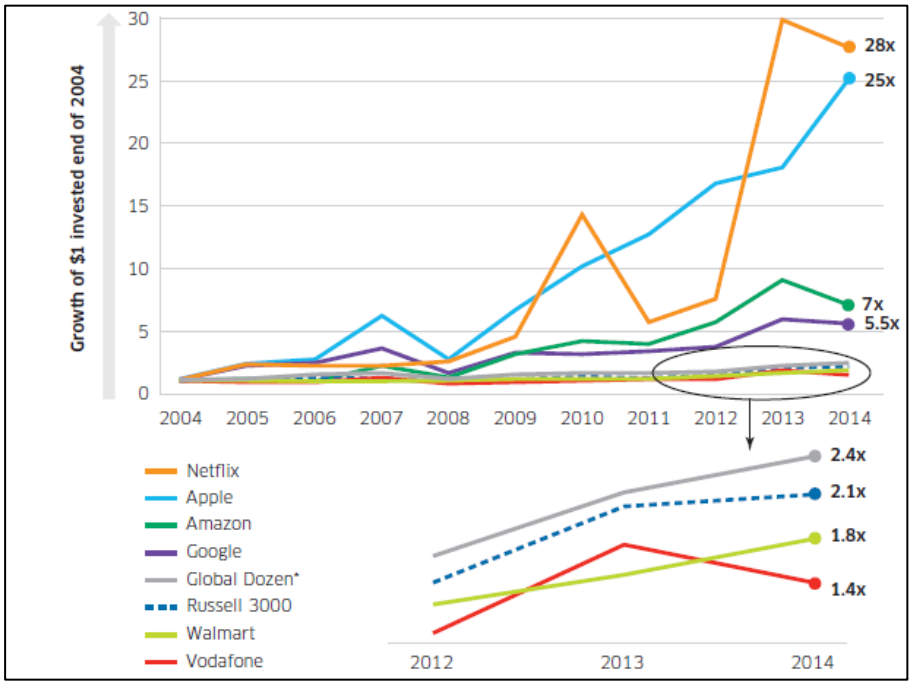
Figura 32. Digital Supply Chain Management



Fonte: BCG

Le principali aziende leader sono le prime a promuovere ed attuare il processo di digitalizzazione. In primo luogo rispetto alle proprie funzioni interne, sia relativamente ai processi verticali che orizzontali lungo la catena del valore. Stanno inoltre arricchendo la propria offerta con funzioni digitali, introducendo servizi innovativi basati sull'utilizzo di dati. I vantaggi, in termini economici, delle imprese che investono nella digitalizzazione è piuttosto evidente, soprattutto confrontando la loro crescita rispetto a quella delle imprese tradizionali.

Figura 33. La crescita delle imprese digitali



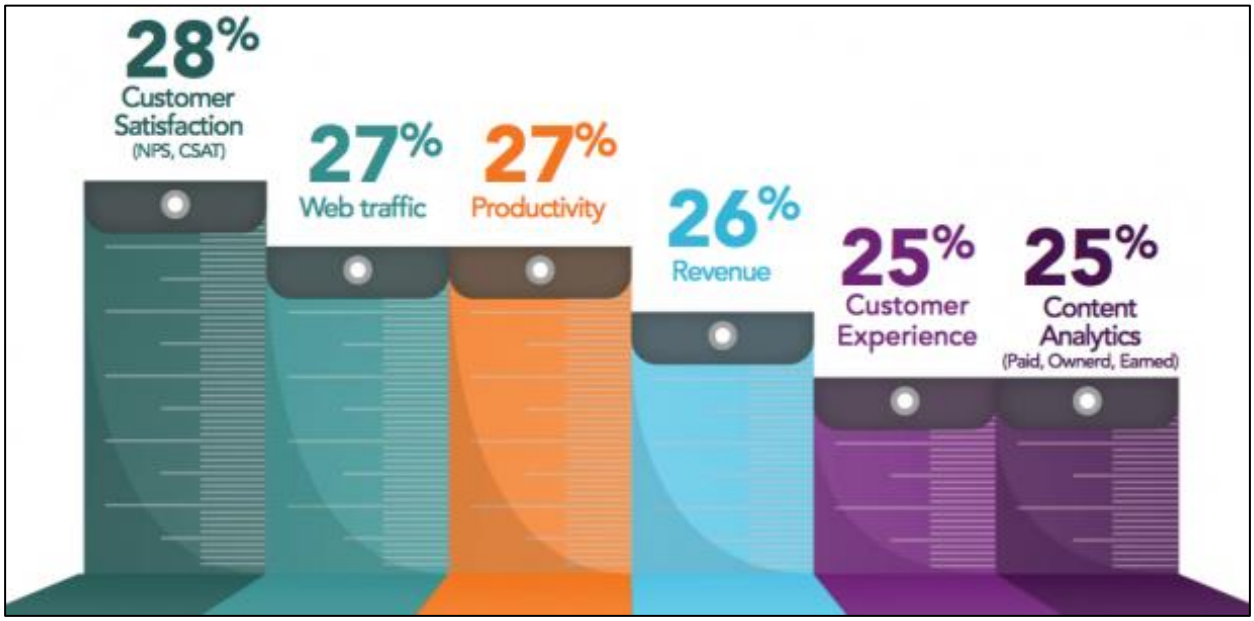
Fonte: The Future X Network: A Bell Labs Perspective

In definitiva, siamo alle porte di una rivoluzione di dimensione globale che stravolgerà le singole realtà aziendali e trasformerà al contempo le dinamiche di mercato in tutta una serie di settori. Si stima che i principali benefici per le imprese si avranno in termini di:

- aumento annuale dei ricavi digitali, in media fino al 2,9%;
- riduzione dei costi, in media del 3,6% l'anno;
- forte aumento degli investimenti (in tecnologie digitali come sensori o dispositivi di connettività, in software e applicazioni come il *Manufacturing Execution Systems*, in formazione per i dipendenti e infine in sistemi di sicurezza)[5].

[5] Le percentuali riportate rappresentano elaborazioni su dati PWC del 2016.

Figura 34. I vantaggi della digital trasformation



Fonte: Altimeter Group, State of digital transformation

Mentre la terza rivoluzione industriale o Industria 3.0 seguendo l'odierna classificazione, si è incentrata sulla automazione della singola macchina e del singolo processo, l'Industria 4.0 si concentrerà sulla digitalizzazione *end-to-end* di tutti i beni materiali e della loro integrazione all'interno dell'ecosistema digitale.

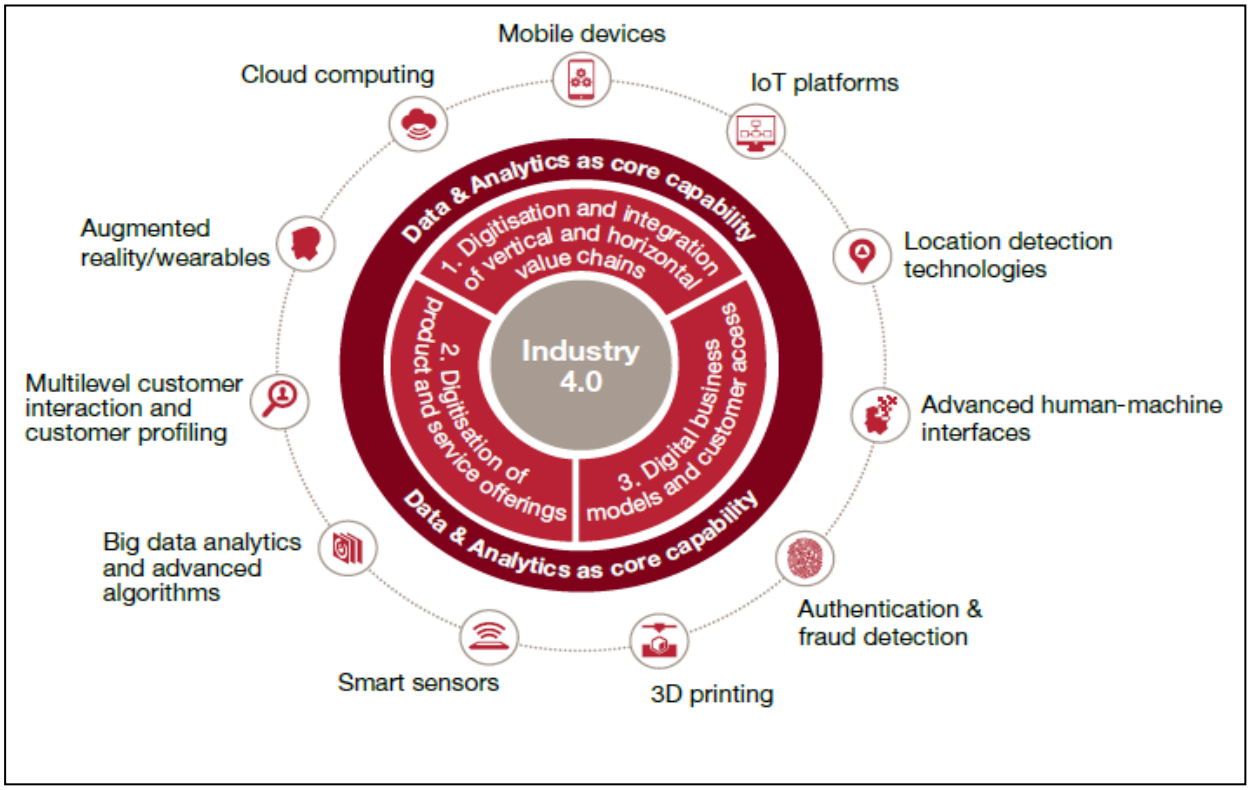
Questo processo passerà attraverso diverse fasi:

(i) *Digitalizzazione e integrazione della catena del valore orizzontale e verticale*. L'industria 4.0 porterà alla digitalizzazione e all'integrazione dei processi verticali passando attraverso l'intera organizzazione, dallo sviluppo all'acquisto del prodotto, mediante sistemi di produzione, logistica e vendita. L'integrazione orizzontale si estenderà al di là delle operazioni interne tra fornitori e clienti su tutta la catena del valore, comprendendo tecnologie di monitoraggio e mappatura dei dispositivi in tempo reale, sistemi di pianificazione integrata ed esecuzione;

(ii) *Digitalizzazione dell'offerta di un prodotto/servizio*. La digitalizzazione dei prodotti comprenderà l'espansione della linea dei prodotti esistenti, per esempio tramite l'aggiunta di sensori intelligenti o di dispositivi di comunicazione che possono essere utilizzati con gli strumenti d'analisi dei dati, nonché attraverso la creazione di nuovi prodotti digitali che si concentreranno su soluzioni completamente integrate. Attraverso nuovi metodi di raccolta e analisi dei dati, le aziende saranno in grado di generare dati relativi al prodotto atti a perfezionare il prodotto stesso, per soddisfare al meglio le crescenti esigenze dei consumatori finali e quindi associeranno ai loro prodotti un ventaglio di servizi;

(iii) *Modelli di business digitali e "accesso" diretto al consumatore*. I principali leader di mercato, stanno già ampliando la loro offerta, essendo in grado di fornire soluzioni digitali importanti come, ad esempio, servizi basati completamente su dati e soluzioni fortemente integrate con le piattaforme. Modelli di business digitali dirompenti saranno spesso incentrati sullo sviluppo di ulteriori ricavi digitali e sull'ottimizzazione delle interrelazioni con il consumatore finale.

Figura 35. Industria 4.0, il quadro delle tecnologie digitali

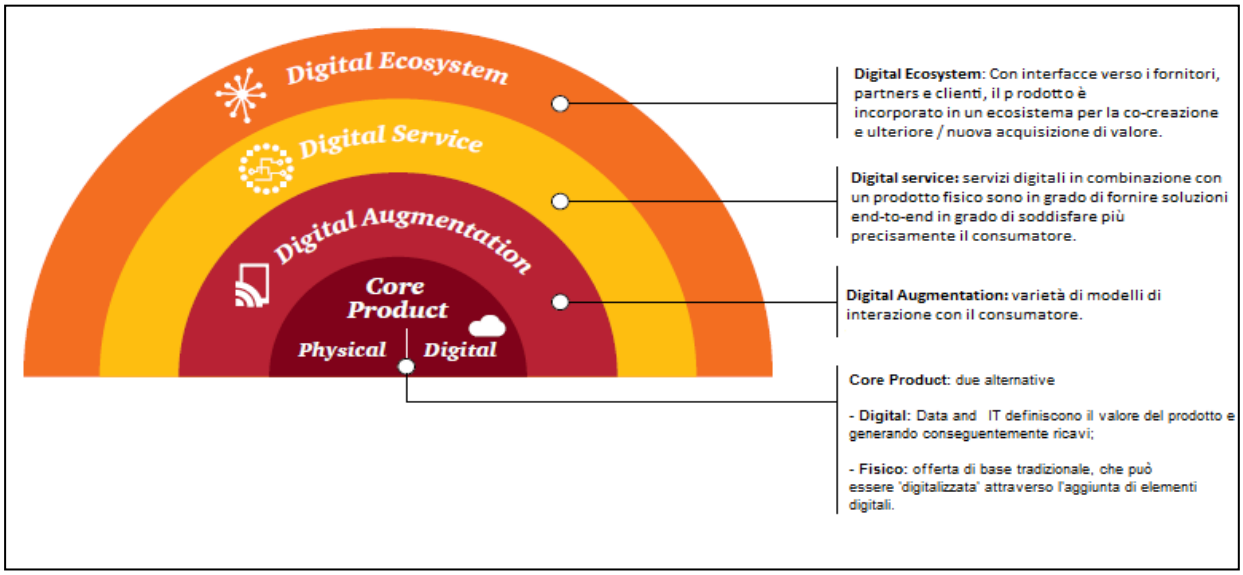


Fonte: PWC, Industry 4.0: Building the digital enterprise

I consumatori saranno dunque al centro dei cambiamenti che interesseranno la catena del valore, i prodotti e i servizi. Prodotti, sistemi e servizi saranno sempre più personalizzati rispetto alle esigenze del consumatore, soprattutto attraverso meccanismi di analisi dei dati. Le imprese necessiteranno di mantenere rapporti diretti con i clienti finali al fine di stimolare la domanda, o almeno di integrarla, soprattutto attraverso l'uso di piattaforme. In questo senso, si sottolinea l'esempio degli aggregatori di taxi, che attraverso l'uso di applicazioni istallate su dispositivi mobili, tramite reti wireless, hanno rapidamente trasformato il settore del trasporto privato. Allo stesso modo, over-the-top fornitori di contenuti sono stati in grado di aggregare, curare, e di fornire contenuti ai consumatori finali sfruttando l'infrastruttura del *cloud over wireline* e le reti wireless a banda larga.

L'immagine che segue evidenzia proprio il mutato scenario e il diverso punto di vista delle imprese della nuova *digital industry*, non più incentrato sul prodotto bensì sulla piattaforma, vista come ponte diretto per arrivare ad interagire con il consumatore finale.

Figura 36. Il passaggio da un approccio product-oriented ad un approccio platform-focused



Fonte: elaborazione ItMedia-consulting su dati PWC

In conclusione, gli effetti positivi della *digital trasformation* sono, senza alcun dubbio, particolarmente significativi sia in termini quantitativi che sostanziali:

- forte impatto sul sistema impresa, che include l'organizzazione dei processi aziendali e le relazioni con tutti gli stakeholder interni ed esterni all'azienda. La trasformazione della catena del valore porta all'ottimizzazione dei processi produttivi attraverso un'integrazione tra i sistemi informativi;
- forte impatto sul sistema economico industriale di un singolo Paese o a livello globale. La trasformazione del modello economico e del meccanismo di creazione di valore alla base dello stesso, comporta un aumento della redditività e dell'efficienza del capitale impiegato del Sistema-Paese;
- le nuove tecnologie (in particolare l'utilizzo degli algoritmi e dell'intelligenza artificiale) consentono un'analisi e una gestione delle informazioni più rapida, precisa ed economica in termini di minori risorse investite, producendo effetti positivi in una moltitudine di settori: dalla ricerca, alla sanità, all'economia, all'energia e all'ambiente, ai trasporti;
- attraverso l'uso dei dati generati dai diversi processi aziendali, le imprese sono capaci di sviluppare nuove soluzioni di business particolarmente innovative e dunque essere maggiormente competitive sul mercato, considerando anche il fatto che l'asse della competizione si è spostato dai costi all'innovazione, alla qualità e alla capacità di personalizzazione dei prodotti;
- i consumatori beneficiano di prodotti e servizi personalizzati e innovativi adatti alle proprie necessità, mediante la predisposizione di offerte costruite ad hoc sulla base delle informazioni sul comportamento/propensione alla spesa e sui bisogni/interessi del cliente;

- l'opportunità di investire in soluzioni aziendali basate sui Big data, acquista un'importanza notevole non soltanto nell'ottica di identificare e soddisfare i bisogni interni all'impresa, ma anche e soprattutto nell'individuare nuovi servizi a valore aggiunto da proporre ai consumatori finali;

- impatto positivo sulla produttività del lavoro, con un ampliamento delle opportunità per i lavoratori maggiormente qualificati e specializzati. A tal proposito, di recente, la società LinkedIn, analizzando milioni di dati e attività di ricerca, relativamente alle assunzioni, sulla propria piattaforma, ha stilato l'elenco delle 10 figure professionali che saranno più richieste dalle aziende nei prossimi anni, tra cui: Cloud and Distributed Computing, Statistical Analysis and Data Mining, Web Architecture and Development Framework, Middleware and Integration Software, User Interface Design,Network and Information Security, Mobile Development, Data Presentation, SEO/SEM Marketing e Storage Systems and Management. Dunque, se da un lato è vero che alcuni posti di lavoro e talune figure professionali diventeranno obsolete, è altrettanto vero che se ne creeranno di nuove, soprattutto grazie all'aumento della competitività tra le imprese, alla crescita dei mercati esistenti e all'introduzione di nuovi beni e servizi.

Tabella 3. SWOT analisi industria 4.0

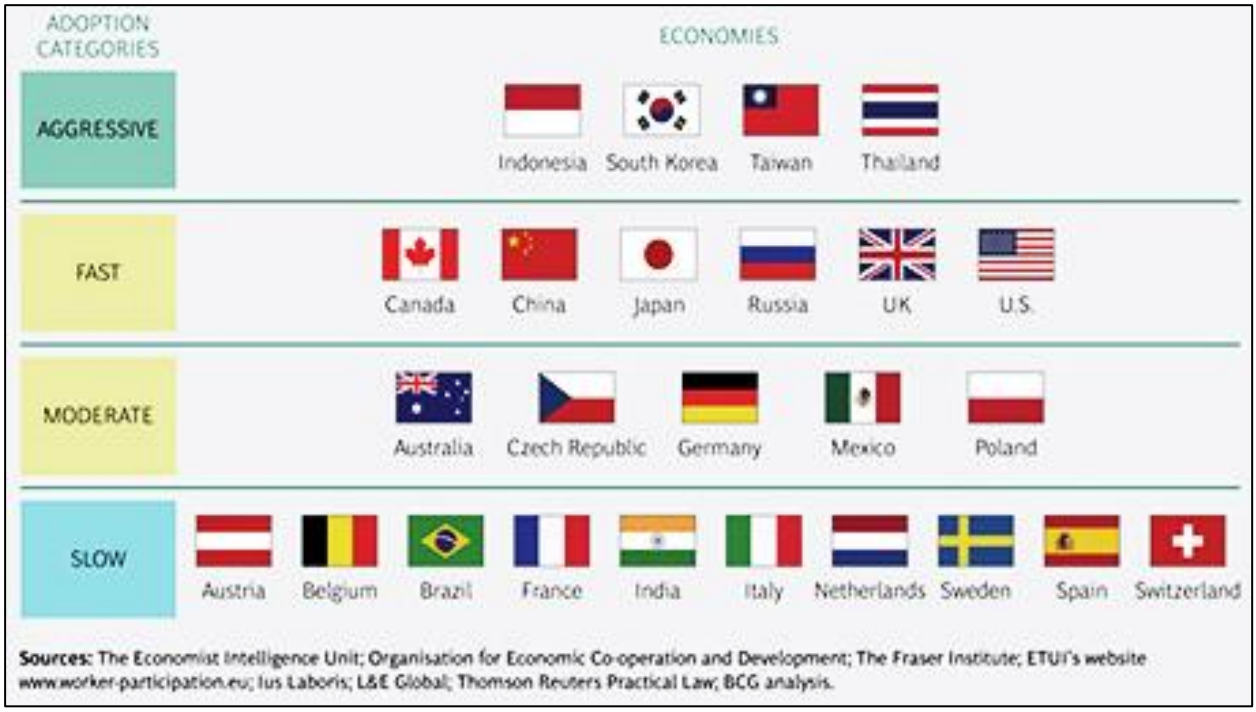
Punti di forza	**Punti di debolezza**
• Aumento della produttività, dell'efficienza, dei ricavi e della competitività su scala globale; • Crescita di personale altamente qualificato e ben retribuito; • Miglioramento del livello di soddisfazione del cliente: personalizzazione e varietà dei prodotti/servizi offerti; • Flessibilità e controllo produzione.	• Forte dipendenza dalla tecnologia e dalla rete; • Dipendenza da una serie di fattori di successo tra cui norme, *framework* coerente, offerte di lavoro coerenti rispetto alle competenze richieste, investimenti e R&S; • Costi di sviluppo e implementazione; • Potenziale perdita di controllo sulle imprese; • Disoccupazione per il personale semi-qualificato; • Necessità di importare manodopera qualificata e integrare le comunità di immigrati.
Opportunità	**Rischi**
• Rafforzare l'industria manifatturiera e rilanciare l'economia; • Sviluppare nuovi mercati guida per prodotti e servizi; • Basse barriere all'ingresso per alcune PMI in mercati emergenti legati alle nuove catene del valore.	• *Cybersecurity*, proprietà intellettuale, data privacy; • Mancanza di consapevolezza e degli strumenti necessari per adattarsi all'*industry 4.0* da parte di lavoratori, PMI, industrie, economie nazionali, che conseguentemente rischiano di rimanere indietro; • Vulnerabilità e volatilità della catena del valore a livello globale; • L'implementazione dell'industria 4.0 da parte di paesi concorrenti potrebbe determinare vantaggi competitivi rispetto a possibili ritardi dell'Europa.

Fonte: ITMedia Consulting

1.3. Overview sugli sviluppi di politica industriale

Gli effetti positivi della *digital trasformation* sono fortemente legati ad una pluralità di fattori: dalla qualità dei meccanismi di trasmissione dei dati, dall'ampiezza di banda a disposizione, dalla capacità di formare *skill* digitali, dai fattori istituzionali abilitanti. Rispetto all'implementazione di questi fattori, gran parte l'Europa e dunque ancor più l'Italia risultano posizionarsi agli ultimi posti della classifica delle principali economie mondiali.

Figura 37. I livelli di implementazione della digital economy a livello mondiale



Fonte: Elaborazione BCG Are your operation ready for a digital revolution

Se da un punto di vista dei tassi di penetrazione dei dispositivi mobili, l'Europa dimostra di essere un terreno fertile per sfruttare appieno i benefici e i vantaggi della quarta rivoluzione industriale; d'altro canto la crisi degli ultimi anni ha posto l'accento sul ruolo fondamentale del settore manifatturiero per il rilancio dello sviluppo economico del Paese. L'Europa sta infatti scontando un profondo deficit competitivo rispetto soprattutto ai paesi emergenti che negli ultimi anni (1990-2014) hanno registrato una forte crescita dei propri settori manifatturieri a livello mondiale.

Uno studio del 2016 di Accenture, *Digital Density*, sul fronte dei cosiddetti "*Fostering Enablers*" colloca l'Italia all'ultimo posto nel ranking delle 17 principali economie mondiali, con ampi gap legati soprattutto a rigidità del mercato del lavoro, scarsa connettività, elevati oneri a carico delle imprese e scarso accesso ai meccanismi che facilitano l'adozione del digitale da parte delle imprese e lo sviluppo delle professioni digitali.

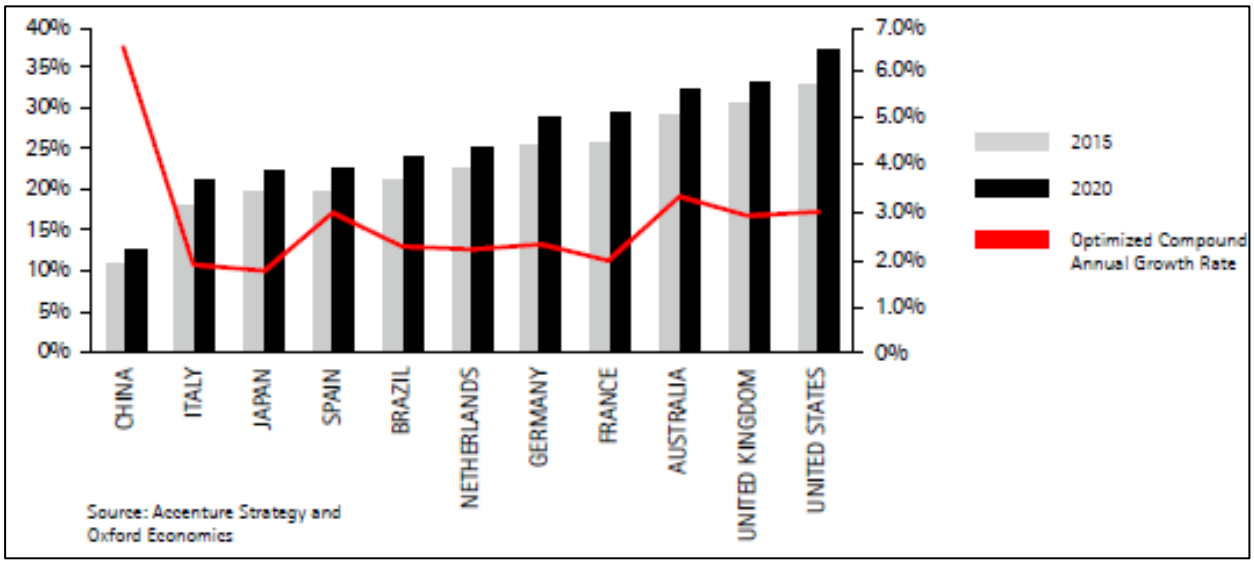
Nonostante le stime riportate nello stesso report di Accenture evidenziano il fatto che l'economia digitale rappresenti circa il 22,5% di quella mondiale, tuttavia le opportunità offerte dal mondo

digitale e dalle nuove tecnologie in termini di business non hanno ancora avuto un pieno sfruttamento. Fino ad ora, soltanto i cd. giganti della tecnologia e le imprese figlie dell'era digitale hanno goduto dei vantaggi propri della *digital disruption* sviluppando nuovi modelli di business su piattaforma. Tale approccio ha fatto sì che queste imprese dominassero il mercato in termini di crescita, profitti e alti tassi di capitalizzazione del mercato, ricavi che sono stati re-investiti all'interno di nuovi ecosistemi digitali.

Oggi, la sfida più grande è dunque quella di capire dove investire per ottenere un miglioramento del prodotto interno lordo del Paese. Si stima infatti che la combinazione per così dire "intelligente" di competenze digitali, tecnologie e investimenti potrebbe aumentare la produttività di un Paese generando 2 miliardi di dollari in più entro il 2020. Prendendo come riferimento l'esempio americano, investimenti in competenze digitali, tecnologie digitali e acceleratori digitali potrebbero far registrare un aumento del PIL del 2,1%, il che equivarrebbe a circa 421 miliardi di dollari nel 2020. Per questo motivo è necessario che i leader aziendali e i *policy maker* effettuino le giuste scelte e prendano le opportune decisioni al fine di favorire la crescita dell'economia del proprio Paese.

La figura 39 mostra la quota "digitale" del PIL rispetto a 11 paesi considerati. Per la maggior parte delle economie di riferimento, la quota digitale del PIL può potenzialmente crescere di circa tre punti percentuali tra il 2015 e il 2020, l'equivalente di un 12,5% di aumento a livello globale. Attualmente, gli Stati Uniti guidano la classifica, con un livello di economia digitale del valore di circa 5.9 miliardi di dollari, pari al 33% del PIL prodotto.

Figura 38. Quota del digitale del PIL paese per paese (2015 and 2020)



Fonte: Accenture Digital disruption: The growth multiplier

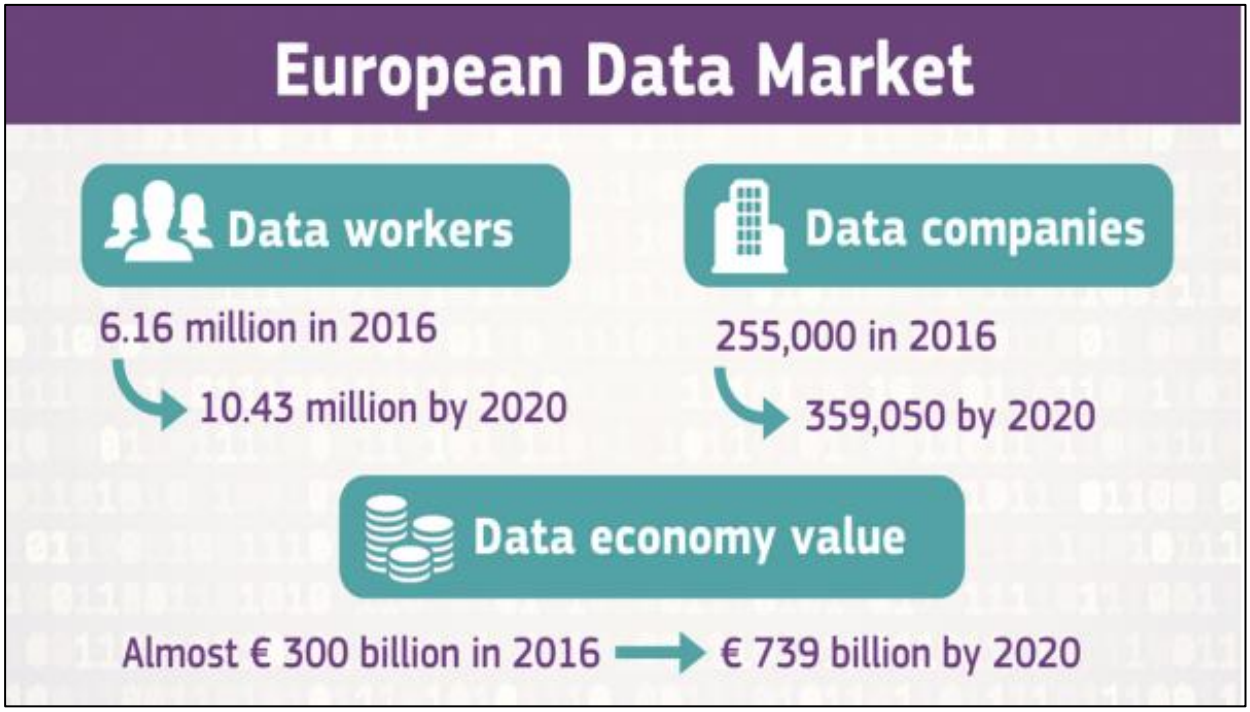
Un recentissimo studio della Commissione Europea[6], compiuto su i 28 paesi dell'Unione, ha descritto i trend di crescita economica guidati dalla data economy. In particolare, lo studio presenta, tra gli

[6] European Commission, *European Data Market study*, 2017. Vedi Final Report - European Data Market Study.

altri, i risultati dell'attività di monitoraggio compiuta su *data workers* (soggetti economici che raccolgono, memorizzano, gestiscono e/o analizzano, interpretano e visualizzano dati), nonché l'equilibrio tra domanda e offerta rispetto alle competenze professionali necessarie per lavorare nell'ambito della *data economy*; imprese (*data companies* e i loro ricavi, *data users companies* e i loro investimenti in tecnologie); il *data market* più in generale e il suo valore economico; la *data economy* e il relativo impatto sull'economia europea; le previsioni a medio termine di tutti gli indicatori basate su traiettorie di mercato alternative. L'obiettivo è quello di promuovere e di accelerare lo sviluppare del mercato unico digitale nell'UE, rappresentando i dati il catalizzatore per la crescita economica, l'innovazione e la digitalizzazione in tutti i settori economici, in particolare per le PMI e le start-up.

Ciascun indicatore è stato analizzato per il periodo compreso tra il 2013-2016 e le previsioni sono state stimate al 2020. Le previsioni a medio termine esplorano tre scenari potenziali alternativi di evoluzione del mercato europeo dei dati, basati su percorsi di sviluppo alternativi determinati da diverse condizioni macroeconomiche e di struttura del mercato: uno scenario di base (tenendo conto della continuazione delle attuali tendenze di crescita e dell'evoluzione delle attuali condizioni strutturali), uno scenario di crescita elevato (dove il mercato dei dati entra in una più veloce traiettoria di crescita grazie a condizioni strutturali più favorevoli) e uno scenario "sfida" (dove il mercato dei dati cresce più lentamente che nello scenario di base, a causa delle condizioni strutturali meno favorevoli e di un contesto macroeconomico meno positivo).

Figura 39.Dimensioni e tendenze dell'economia dei dati dell'UE



Fonte: European Data Market study

Quanto ai *data workers,* lo studio ha mostrato che nel 2015 6 milioni di lavoratori in Europa risultavano occupati all'interno della cd. *data industry*, raggiungendo 6,16 milioni nel 2016. Per quanto riguarda gli sviluppi a medio termine, si stimava che in uno scenario di crescita elevata, il numero dei *data workers* in Europa crescerà fino a 10,43 milioni, ad un tasso di crescita medio complessivo del 14,1% entro il 2020.

Secondo l'analisi, nel 2016 si è verificato un gap tra la domanda complessiva e l'offerta di *data workers* di ben 420.000 posti di lavoro vacanti, pari al 6,2% della domanda totale dei professionisti di dati specializzati. Entro il 2020, l'UE dovrà affrontare un divario di competenze in materia di dati corrispondente a 769.000 posizioni inutilizzate nello scenario di base, concentrandosi in particolare nei grandi Stati membri (in particolare in Germania e Francia).

L'industria dei dati nel suo complesso comprendeva circa 255.000 imprese nel 2016 nell'UE e che sono state capaci di totalizzare ricavi complessivi per 62 miliardi di euro. Secondo le previsioni di scenario di crescita elevata, questa cifra salirà a 359.050 entro il 2020 con un tasso di crescita annuale dell'8,9%.

Il mercato europeo dei prodotti e servizi digitali ("*data market*") nell'UE era di 54,4 miliardi di euro nel 2015 e di 59,5 miliardi di euro nel 2016. Questo trend positivo dovrebbe continuare nei prossimi quattro anni, anche se in percentuali diverse a seconda degli scenari considerati. Secondo lo scenario di crescita elevata, il valore del mercato dei dati UE raggiungerà 106,8 miliardi di euro entro il 2020 con un tasso di crescita annuo del 15,7% dal 2016.

Il valore complessivo dell'economia dei dati è passato dai 247 miliardi di euro nel 2013 a quasi 300 miliardi di euro nel 2016. Secondo le stime, il valore dell'economia dei dati nel 2016 rappresenta quasi 2% del PIL europeo. Entro il 2020 la *data economy* europea dovrebbe aumentare a 739 miliardi di euro con un impatto complessivo del 4% sul PIL in uno scenario di crescita elevata caratterizzato da un ruolo più forte dell'innovazione digitale e con maggiori investimenti ICT complessivi.

Nell'ambito dello scenario sopra delineato, relativamente all'insieme dei cambiamenti descritti, va tuttavia precisato che la sola applicazione delle tecnologie digitali non è in alcun modo sufficiente per cogliere e beneficiare appieno dei vantaggi dell'industria 4.0. In tal senso, risulta necessario sviluppare e adottare standard comuni, su scala europea, per la digitalizzazione dell'industria. Nella prospettiva del Digital single market, strategia europea presentata dalla Commissione il 6 maggio del 2015 comprensiva di una serie di azione mirate e finalizzate a creare le condizioni perché il mercato unico possa meglio adeguarsi all'era digitale, la stessa Commissione ha adottato norme tecniche comuni atte a coordinare le diverse iniziative nazionali di digitalizzazione. Nei paragrafi successivi evidenzieremo gli aspetti salienti della strategia europea, nonché il piano del Governo italiano.

1.3.1 Lo scenario europeo

Nell'ambito del processo di digitalizzazione, gli interventi della Commissione Europea, promotrice di un piano che muoverà complessivamente 50 miliardi di euro fino al 2020, si pongono come obiettivo di garantire alle imprese europee del settore, anche mediante la previsione di una forte spinta agli investimenti congiunti tra settori diversi attraverso partnership strategiche e reti di imprese, il

massimo sfruttamento delle opportunità offerte dalla trasformazione digitale, al fine di renderle maggiormente competitive sul mercato globale.

Molti comparti economici hanno già adottato rapidamente le tecnologie e i processi digitali, tuttavia l'industria europea non ha ancora sfruttato appieno le opportunità digitali per essere competitiva a livello mondiale. In particolare i settori tradizionali (come l'edilizia, il settore agro-alimentare, il comparto tessile e la siderurgia) e le PMI sono indietro per quanto riguarda la trasformazione digitale.

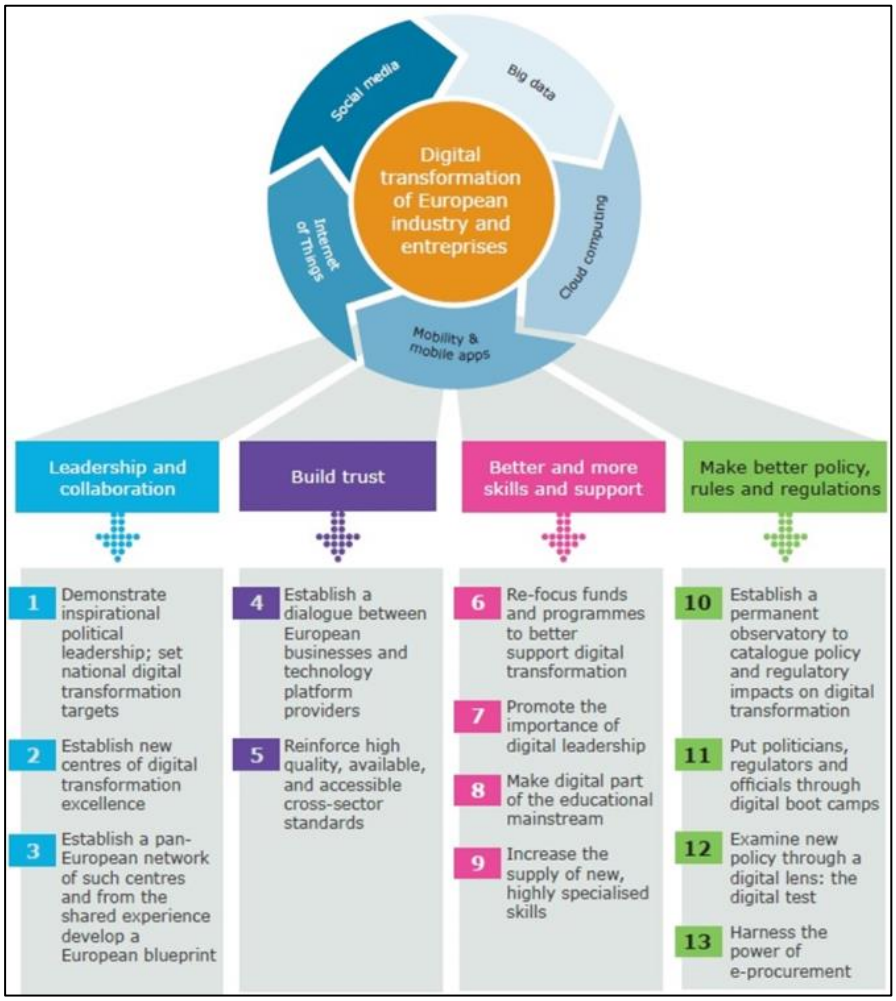
Per questo motivo diversi Stati membri dell'UE, tra cui anche l'Italia, hanno già varato strategie per sostenere la digitalizzazione dell'industria. Tuttavia la predisposizione di un approccio comune in tutta l'area comunitaria avrà certamente un impatto ancor più significativo ed efficiente e contribuirà a facilitare notevolmente il raggiungimento dell'obiettivo, evitando la frammentazione dei mercati. L'azione della Commissione è sostanzialmente riconducibile a sei punti chiave:

- favorire il coordinamento tra iniziative nazionali o regionali per la digitalizzazione dell'industria, garantendo il massimo dialogo tra le parti in causa. Un quadro di *governance* sarà istituito dagli Stati membri e dall'industria;
- investire nelle partnership pubblico-privato incoraggiando l'utilizzo dei fondi Ue, perché diventino realmente delle strutture di aggregazione e degli ecosistemi per le innovazioni industriali digitali.;
- investire 500 milioni di EUR in una rete paneuropea di poli di innovazione digitale (centri di eccellenza nelle tecnologie) in cui le imprese potranno ricevere consigli e sperimentare le innovazioni digitali;
- definire progetti pilota su larga scala per potenziare l'Internet delle cose, i processi produttivi avanzati e le tecnologie in ambiti quali città intelligenti, domotica, automobili connesse o sanità mobile;
- adottare una normativa adeguata alle esigenze future che faciliti il libero flusso dei dati e chiarisca i diritti di proprietà dei dati generati da sensori e dispositivi intelligenti. La Commissione intende inoltre riesaminare le norme sulla sicurezza e l'affidabilità dei sistemi autonomi;
- presentare un'agenda europea per le competenze che contribuisca a dotare i cittadini delle competenze necessarie per i posti di lavoro nell'era digitale.

La necessità di un approccio comune deriva dalla considerazione che nel Digital single market miliardi di oggetti connessi avranno bisogno di poter comunicare tra loro, e perché questo sia possibile sarà necessario lo sviluppo di un linguaggio comune. Le aree prioritarie su cui lavorare agli standard sono

riconducibili a 5: 5G, *cloud computing*, IoT, dati e *cybersecurity*, tutte perfettamente applicabili all'industry 4.0[7].

Figura 40. Le iniziative della Commissione in chiave di digitalizzazione dell'industria europea



Fonte: Commissione Europea

Quanto al tema della connettività, il 14 settembre 2016 la Commissione ha proposto un pacchetto di iniziative e proposte legislative, al fine di garantire in tutto il territorio dell'Unione la massima connettività alla rete internet, in modo da facilitare la piena partecipazione di tutti, cittadini e istituzioni, allo sviluppo della *digital economy*. Gli obiettivi della Commissione sono riconducibili a 5

[7] I documenti di riferimento sono 3 Comunicazioni della Commissione: *Digitalizzazione dell'industria europea. Cogliere appieno i vantaggi di un mercato unico digitale*, doc. COM (2016) 180 del 19.04.2016; Iniziativa europea per il cloud computing – Costruire un'economia competitiva dei dati e della conoscenza in Europa, doc. COM(2016) 178 del 19.04.2016; "Priorità per la normazione delle TIC per il mercato unico digitale", doc. COM(2016) 176 del 19.04.2016.

categorie tra loro sinergiche: la comunicazione "Gigabit Society"[8]; la proposta di direttiva per un nuovo codice europeo delle comunicazioni elettroniche[9]; un piano d'azione per avviare in tutta l'UE, a partire dal 2018, l'implementazione del 5G[10]; un'iniziativa, la cosiddetta WiFi 4 EU[11], tesa ad aiutare le comunità ad offrire a tutti i cittadini punti di accesso wi-fi gratuito; una proposta di Regolamento per il BEREC[12].

In particolare, passando ad un'analisi più di dettaglio, attraverso la Comunicazione "Gigabit Society" la Commissione ha fissato 3 obiettivi (non vincolanti) di connettività al 2025: a) tutti i principali motori socioeconomici (quali scuole, università, etc..) dovrebbero avere accesso a una connettività pari a 1 Gbps in download e upload; b) tutte le famiglie europee, nelle zone rurali o urbane, dovrebbero avere accesso a una connettività che offra velocità di download di almeno 100 Mbps, che può passare all'ordine dei Gbps; c) tutte le aree urbane e le principali strade e ferrovie dovrebbero avere una continuità di copertura delle reti 5G.

In aggiunta, con riferimento alla proposta di un codice delle comunicazione elettroniche europee (CCEE), che imposta a livello UE norme e obiettivi comuni su come l'industria delle telecomunicazioni deve essere regolamentata e che si ispira ad una logica di incentivazione della concorrenza e degli investimenti, questa è orientata al perseguimento di quattro grandi finalità: un incremento della concorrenza e della prevedibilità degli investimenti; un uso migliore delle frequenze radio; una maggiore tutela dei consumatori e la creazione di un ambiente online più sicuro per gli utenti con la predisposizione di regole più eque per tutti gli operatori.

Partendo poi dalla constatazione dell'importanza dello sviluppo del 5G in un contesto sempre più caratterizzato dalla crescita esponenziale del traffico dati – favorita dallo sviluppo dell'IoT e dalla fruizione di contenuti in mobilità – la Commissione ha presentato un piano d'azione per il 5G che prevede l'adozione di standard globali comuni, la predisposizione di un calendario comune per sperimentare il 5G attraverso test paneuropei a partire dal 2018 e coordinare il lancio commerciale del 5G nell'UE nel 2020, l'individuazione delle bande di spettro per il 5G e l'adozione di tabelle di marcia nazionali per l'installazione del 5G in tutti gli Stati membri dell' UE. A tale proposito, nell'ambito dello sviluppo del sistema 5G e con particolare riferimento allo spettro di frequenze da utilizzare, la Commissione Europea nel mese di marzo 2017 ha rilasciato una Comunicazione relativa

[8] European Commission, *Connectivity for a Competitive Digital Single Market - Towards a European Gigabit Society*, 2016.

[9] Proposta di Direttiva che istituisce il Codice Europeo delle Comunicazioni Elettroniche, COM(2016) 590.

[10] *Il 5G per l'Europa: un piano d'azione*, Comunicazione della Commissione, COM(2016) 588.

[11] Regolamento del Parlamento europeo e del Consiglio che modifica i Regolamenti (UE) n.1316/2013 e (UE) n. 283/2014 per quanto riguarda la promozione della connettività internet nelle comunità locali, COM(2016) 589.

[12] Proposta di Regolamento del Parlamento europeo e del Consiglio che istituisce l'Organismo dei regolatori europei delle comunicazioni elettroniche (BEREC), COM(2016) 591.

al 5G Action Plan per invitare gli Stati membri ad individuare entro il 2018 almeno una città dove avviare la sperimentazione del 5G. All'interno di questo Action Plan, sulla base dei regolamenti radio dell'Unione internazionale delle telecomunicazioni, adottati dalla Conferenza mondiale delle radiocomunicazioni del 2015, che prevedono l'assegnazione della banda di frequenza dei 700 MHz ai servizi di trasmissione e ai servizi mobili su base coprimaria, ha identificato la banda 700 MHz come banda pioniera insieme alla 3.43.8 GHz ed alla 26 GHz (24.25 – 27.5 GHz) per usi 5G. In questo contesto, il 17 maggio 2017 è stata adottata la Decisione (UE) 2017/899 del Parlamento europeo e del Consiglio relativa all'uso della banda di frequenza 470-790 MHz nell'Unione. Si tratta di un'iniziativa particolarmente rilevante nella quale vengono riconosciute l'importanza della banda di frequenza dei 700 MHz in un'ottica di armonizzazione e coordinamento a livello mondiale dello spettro per la banda larga senza fili e le potenzialità che la stessa banda possiede in termini di sviluppo di nuovi servizi digitali innovativi nelle zone urbane e in quelle rurali o remote.[13]

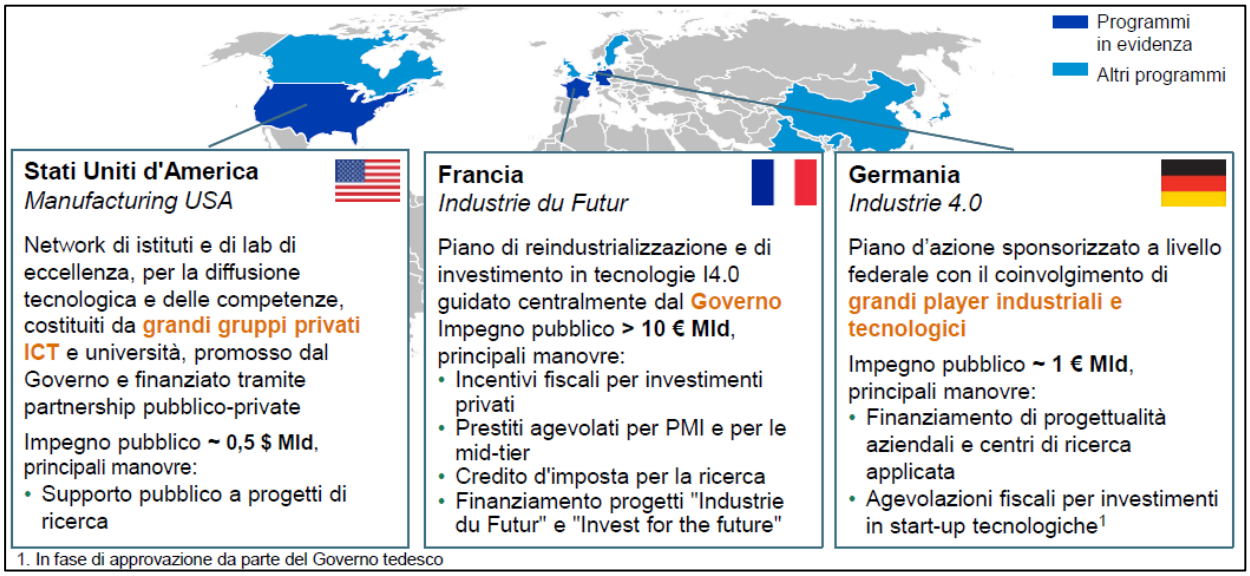
Tra le altre iniziative di interesse sul versante della connettività a supporto dello sviluppo dei nuovi servizi digitali, l'iniziativa WiFi 4 EU persegue l'obiettivo di agevolare le autorità locali – almeno 6.000-8.000 – interessate ad offrire gratuitamente connessioni wi-fi ai cittadini che si trovino, ad esempio, all'interno e in prossimità di edifici pubblici, centri sanitari, parchi o piazze. Si tratta di un'azione importante che prevede una dotazione iniziale di 120 milioni di euro e finanziamenti per l'installazione di punti di accesso locali senza fili disponibili in tempi rapidi.

L'attuazione dell'industria 4.0 nella pratica richiede il coordinamento e la pianificazione su scala europea di una vasta gamma di attività, tra le quali in particolare: il sostegno alla ricerca e a progetti pilota; lo sviluppo di una legislazione in materia di standardizzazione e di dati; una forte attività di formazione sia a livello universitario che di impresa; procedere al completamento del mercato unico digitale; adeguare le imprese ai nuovi modelli di business.

Mentre molti Stati membri stanno portando avanti queste attività in maniera piuttosto mirata, in particolare la Germania e la Francia, in altri paesi come l'Austria, il Belgio, la Danimarca, la Finlandia, l'Italia, i Paesi Bassi, il Portogallo, la Spagna, la Svezia e il Regno Unito nonostante il forte sostegno e interesse, tuttavia solo in pochi casi il programma di governo è così ampio e strutturato come quello di matrice tedesca o francese. La figura 41 mostra un quadro di comparazione tra i programmi adottati dai governi francese e tedesco, perlopiù incentrati sulla predisposizione di agevolazioni fiscali e finanziamenti a tassi agevolati, in rapporto con il modello statunitense in cui lo sviluppo di competenze di alto profilo e il sostegno alla ricerca rappresentano i punti cardine del piano.

[13] Il paragrafo 15 della Decisione (UE) 2017/899 prevede, inoltre, la possibilità per gli Stati membri di ritardare, sulla base di motivi debitamente giustificati, l'autorizzazione all'uso della banda di frequenza dei 700 MHz per i sistemi terrestri in grado di fornire servizi di comunicazione elettronica a banda larga senza fili oltre il termine comune dell'Unione fissato al 2020, per un periodo massimo di due anni.

Figura 41. I principali programmi industria 4.0 nel mondo



Fonte: Dati pubblici

All'interno delle attività della Commissione Europea concernenti il sostegno della ricerca e i progetti pilota, si colloca il programma Horizon 2020[14], il primo piano europeo per la ricerca e l'innovazione investimenti essenziali per il futuro dell'Europa e che pertanto si pongono al centro della strategia Europa 2020 per una crescita intelligente, sostenibile e inclusiva. Il programma è incentrato su tre settori chiave: eccellenza scientifica, leadership industriale e sfide per la società. L'obiettivo è assicurare che l'Europa produca una scienza e tecnologia di classe mondiale in grado di stimolare la crescita economica, in particolare attraverso un finanziamento mirato che contribuirà ad assicurare che le idee migliori vengano commercializzate nel più breve tempo possibile affinché l'intera società europea possa goderne e trarne beneficio.

Nello scenario appena descritto, sul presupposto che i dati rappresentino una risorsa essenziale per la crescita economica, occupazionale e per il progresso sociale, è pacifico ritenere che la loro analisi faciliti i processi decisionali, l'innovazione e la previsione di eventi futuri. Per queste ragioni, l'Europa si propone di sfruttare il potenziale riconosciuto ai dati, ma senza violare i diritti e le libertà degli individui o disincentivare gli investimenti privati nel settore. La Commissione punta alla creazione di un mercato unico digitale efficiente e competitivo. In quest'ottica, risulta necessario identificare le sfide di carattere legale, economico e normativo, e di avviare un intenso dibattito con le parti interessate per una più corretta definizione delle azioni da intraprendere.

Il 10 gennaio 2017, la Commissione ha lanciato una consultazione sulle regole dell'*European data economy,* il pacchetto "*Building the European Data Economy*" composto da una Comunicazione e da uno *Staff Working Document.* Si tratta di documenti di policy che forniscono una panoramica sulle diverse questioni in gioco, nonché sul contesto di riferimento.

[14] Vedi https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_IT_KI0213413ITN.pdf

L'obiettivo è quello di raccogliere informazioni su:

- se e come le restrizioni di localizzazione dei dati, locali o nazionali, ostacolano la loro libera circolazione in Europa;
- se e in che misura i dati anonimi generati dai dispositivi connessi vengono negoziati e scambiati;
- la natura e l'entità di eventuali ostacoli all'accesso di tali dati;
- le modalità con cui affrontare queste barriere;
- le sfide sul tema della responsabilità legate allo sviluppo dell'emergente IoT e della robotica;
- le pratiche e le questioni relative alla portabilità dei dati, l'interoperabilità e la definizione di standard tecnici.

Il pacchetto "*Building the European Data Economy*" si focalizza principalmente sulle problematiche connesse alle limitazioni sul libero trasferimento dei dati, incluse le barriere legali previste per la raccolta e l'elaborazione di dati localizzati, trattando tutta una serie di questioni critiche emergenti legate ai dati, quali: la proprietà, l'accesso, la portabilità, la responsabilità e il riuso. Mentre le questioni sulla responsabilità sono affrontate in questa consultazione in un contesto di *data economy*, verrà successivamente avvita una seconda consultazione distinta, sulla valutazione complessiva circa l'applicazione della direttiva sulla responsabilità del produttore (85/374/CEE). La consultazione in commento non copre alcuna questione legata alle protezioni dei dati personali, che sono invece ampiamente trattate nell'ambito delle nuove norme UE sulla protezione dei dati, così come attraverso la revisione della direttiva *e-privacy*. Anche le problematiche legate all'accesso e al riutilizzo dei dati del settore pubblico sono escluse da questa consultazione, perché saranno affrontate nell'ambito della prossima revisione della Direttiva 2003/98/CE.

La consultazione è indirizzata a: attività di ogni dimensioni; imprese e consumatori di dispositivi connessi; operatori e utenti di piattaforme digitali; data brokers; attività che commercializzano prodotti e servizi basati sui dati; autorità pubbliche; organizzazioni non governative; ricercatori e enti di ricerca; consumatori. La consultazione coprirà ogni settore: dal manifatturiero, all'energetico, all'automobilistico, al sanitario, all'IoT, ecc e terminerà il 26 aprile 2017.

La Commissione già nel 2014 era intervenuta con una comunicazione[15] prodromica alla *digital single market economy,* in cui veniva proposto un piano d'azione coordinato che coinvolgeva gli Stati membri e l'Unione europea al fine della programmazione di tutta una serie di azioni volte ad accelerare l'innovazione, la crescita della produttività e l'aumento della competitività all'interno della cd. *data economy*, oltre che sul mercato globale. In quest'ottica, l'Unione europea intervenne nei seguenti campi:

[15] European Commission, *Towards a thriving data-driven economy"*, 2004. A cui segue poi lo *staff working document*, *"Accompanying the document Towards a thriving data-driven economy*.

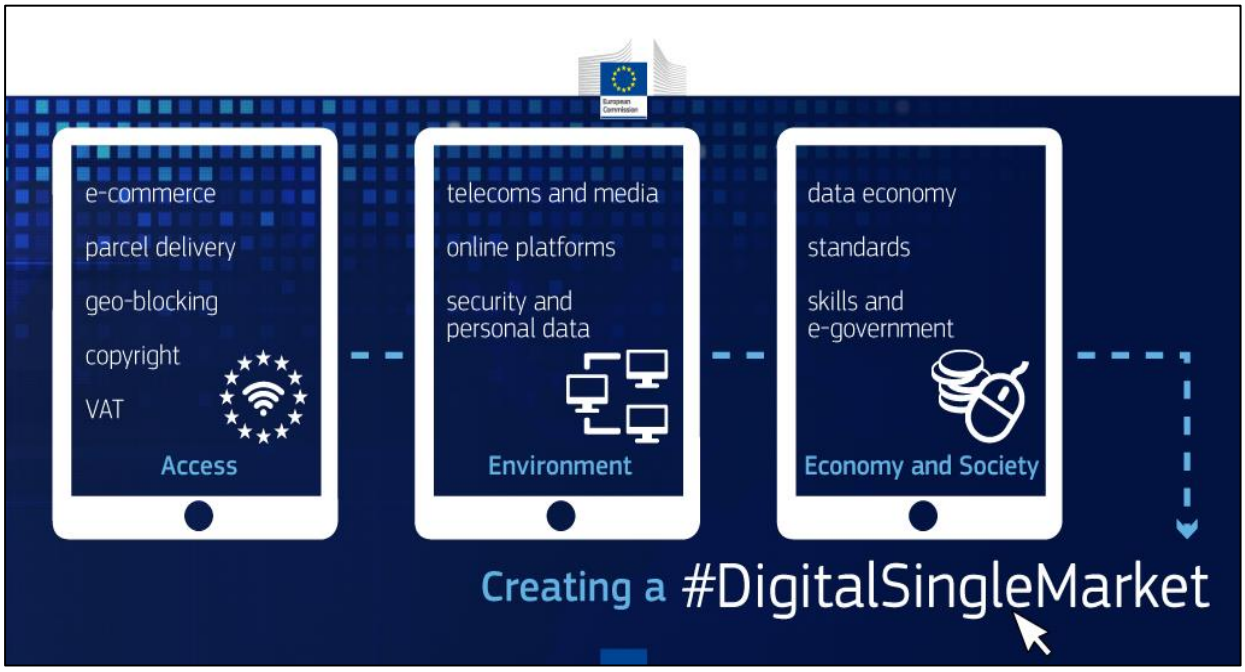
- sostenere le iniziative cd. "*lighthouse*" finalizzate al miglioramento della competitività, della qualità dei servizi pubblici e della vita dei cittadini;
- sviluppare le tecnologie abilitanti, le relative infrastrutture e le competenze, in particolare a vantaggio delle piccole e medie imprese;
- promuovere la condivisione, l'utilizzo e lo sviluppo dei dati pubblici, nonché la ricerca di idonee infrastrutture di dati;
- concentrare la ricerca e gli investimenti pubblici nel settore tecnologico, legale e in altri *bottlenecks*;
- garantire che il quadro giuridico e la policy siano *data-friendly;*
- accelerare la digitalizzazione della PA e dei servizi per aumentarne l'efficienza;
- utilizzare gli investimenti pubblici per assicurare al mercato i vantaggi derivanti dalle tecnologie *data based;*
- concludere rapidamente il processo legislativo di riforma sulla protezione dei dati, la sicurezza delle reti e dell'informazione, e lo scambio di sostegno e la cooperazione tra le autorità di controllo competenti (per esempio per la protezione dei dati, la tutela dei consumatori e la sicurezza della rete).

Le due comunicazioni adottate nel 2014 e nel 2017 trovano le proprie radici nelle idee formulate per la prima volta dall'ex vicepresidente della Commissione Neelie Kroes, nell'ambito di un'iniziativa strategica sulla catena del valore dei dati, lanciata nel novembre 2013[16]. In particolare, questa iniziativa era incentrata sulla definizione di un ecosistema di dati europeo, destinato a stimolare la ricerca e l'innovazione in tutti i settori nei quali i dati potenzialmente troveranno applicazione.

All'interno di questo quadro si inserisce la *Digital Single Market Strategy for Europe* presentata, come detto, nel maggio 2015, con cui la Commissione, attraverso 16 iniziative diverse, ripartite in tre distinti pilastri, si propone di abbattere le barriere regolamentari ai fini della costituzione di un unico mercato al posto dei 28 mercati nazionali esistenti. In particolare, i tre pilastri sono indirizzati a: 1) Migliorare l'accesso ai beni e servizi digitali in tutta Europa per i consumatori e le imprese; 2) Creare un contesto favorevole e parità di condizioni affinché le reti digitali e i servizi innovativi possano svilupparsi; 3) Massimizzare il potenziale di crescita dell'economia digitale.

[16] European Commission DG CONNECT, *A European strategy on the data value chain*, 2013.

Figura 42. Digital Single Market Strategy for Europe



Fonte: Commissione Europea

Sempre nel 2015, il Parlamento europeo è intervenuto sul tema privacy e dati personali attraverso uno studio[17], in cui si è dato conto dello stato del dibattito sul trattamento dei dati personali, considerando gli sviluppi delle strategie comunitarie promotrici di un'economia basata sui dati; nonché dell'attuale riforma del quadro giuridico sulla protezione dei dati personali, nel contesto dell'adozione di un regolamento generale (GDPR)[18]. Considerando inoltre i recenti sviluppi nell'IoT, lo studio mira a dimostrare che l'alto grado di opacità sulle modalità di trattamento dei dati, colpisce direttamente gli individui, inconsapevoli di come saranno impiegati i propri dati personali. Il Parlamento conclude nel senso di ritenere che la promozione di un'economia *data driven* non può sottovalutare gli aspetti di privacy, ma si rende parimenti necessario rafforzare i diritti, per così dire digitali, dei cittadini.

Oltre a ciò, sempre nall'ambito relativo alla protezione dei dati, il 13 settembre 2017 è stata adottata una proposta di regolamento del Parlamento europeo e del Consiglio relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea. Si tratta di un'iniziativa importante che, partendo dalla constatazione dell'enorme valore del mercato dei dati, si prefigge l'obiettivo di migliorare la mobilità dei dati non personali a livello transfrontaliero nel mercato unico, far sì che la facoltà delle autorità competenti di chiedere e ottenere l'accesso ai dati ai fini del controllo regolamentare, quali ispezioni e audit, resti invariata ed, infine, facilitare agli utenti professionali di servizi di archiviazione o di altri servizi di trattamento di dati il cambio di fornitore di servizi e la

[17] European Parliament, Directorate-General for Internal Policies of the Union, *Big Data and smart devices and their impact on privacy*, 2015.

[18] Per maggiori approfondimenti sul GDPR, vedi il paragrafo 5.2.

portabilità dei loro dati, senza generare un onere eccessivo per i fornitori di servizi o distorsioni del mercato.

Negli anni, a partire dal 2011, anno della comunicazione della Commissione sugli "*Open data*"[19] in cui la Commissione sottolinea l'importanza dei dati pubblici/governativi (informazioni geografiche, statistiche, dati meteo, dati provenienti da progetti di ricerca finanziati con fondi pubblici, e libri digitalizzati delle biblioteche) il cui ri-uso ha un enorme potenziale: nella creazione di nuovi prodotti e servizi; nella scoperta di soluzioni nuove e innovative; nel raggiungimento di una maggiore efficienza tra PA; nella promozione della partecipazione dei cittadini alla vita politica e sociale attraverso l'aumento della trasparenza nella gestione governativa. La Commissione è intervenuta in diversi ambiti sempre sul tema della digitalizzazione e dei dati. Sempre nel 2015 meritano di nota gli interventi sui profili di *machine learning*[20]in particolare per la classificazione ed elaborazione dei dati satellitari che per loro natura e caratteristiche quali varietà, volume e parziale struttura sono assimilabili ai big data; la presentazione di una ricerca[21] finalizzata all'implementazione di una piattaforma software applicativa per la sperimentazione scientifica e l'analisi dei dati. L'obiettivo è quello di riuscire ad esplorare, estrarre valore da, e dare un senso ed un'interpretazione di enormi set di dati interconnessi. La piattaforma fa uso di modelli statistici, tecniche esplorative o descrittive, strumenti di visualizzazione, per effettuare ragionevoli interpretazioni dei dati. L'approccio è di tipo multi-fase e mira ad acquisire preliminarmente una chiara visione di come utilizzare l'analisi dei dati, coerentemente rispetto alle proprie esigenze strategiche. Ancora, a chiusura dell'anno si collocano anche gli interventi in materia di archiviazione dei dati personali[22]; delle problematiche di privacy connesse al trattamento dei dati personali[23], principi di privacy e protezione dei dati come valore essenziale dei big data, non solo a vantaggio degli utenti, ma anche per la genuinità della stessa analisi dei dati; e infine della *data security*[24] considerata una dei principali aspetti critici della *data economy*.

[19] European Commission, *Open data: An engine for innovation, growth and transparent governance*, 2011.

[20] *Benchmarking of the symbolic machine learning classifier with state of the art image classification methods*, 2015.

[21] *Collaborative research-grade software for crowd-sourced data exploration*, 2015.

[22] *Study on Personal Data Stores*, commissioned by DG CONNECT to Cambridge University Judge Business School, 2015. Il rapporto rileva che questo modello incentrato sull'utente potrebbe offrire agli individui più convenienza, nuove forme di remunerazione e strumenti di gestione dei dati, oltre a facilitare l'esercizio dei loro diritti alla privacy. Dalla prospettiva del business e della ricerca, potrebbe offrire nuove opportunità di crescita innovativa, consentire una ricerca nuova e più completa, rendere possibile l'analisi dei big data anche alle aziende più piccole e offrire guadagni in efficienza, in particolare ai soggetti operanti nel settore pubblico.

[23] ENISA – European Union Agency for Network and Information Security, *Privacy by design in big data*, 2015.

[24] ENISA – European Union Agency for Network and Information Security, *Big data security. Good practices and recommendations on the security of big data systems*, 2015.

Nel 2016 numerosi sono stati gli interventi della Commissione Europea in settori e aree specifiche, come quello spaziale[25] per quanto concerne lo studio dei dati terrestri e spaziali raccolti dai diversi sensori installati sullo spazio e sulla terra e aventi le stesse caratteristiche dei big data tradizionalmente concepiti (volume, varietà, velocità, valore e veridicità); nel settore del *cloud computing* e dell'IoT[26] per dar conto dello stato dell'arte e del livello di implementazione di queste nuove tecnologie in tutta Europa, fornendo una serie di raccomandazioni per affrontare al meglio le prossime sfide di ricerca e innovazione. Sempre nel 2016 si colloca la Comunicazione della Commissione "*Digitising European Industry Reaping the full benefits of a Digital Single Market*" attraverso cui la Commissione introduce una serie di misure politiche coerenti come parte del pacchetto di modernizzazione DSM per quanto concerne le nuove tecnologie e i servizi pubblici. Il pacchetto comprende altri tre comunicazioni. L'obiettivo della Commissione è quello riuscire a stabilire un quadro di coordinamento tra iniziative nazionali e comunitarie in questo settore e le azioni politiche pertinenti, compresi gli investimenti in innovazioni digitali e infrastrutture, accelerando lo sviluppo degli standard ICT, esplorando le condizioni normative e promuovendo procedure di formazione e aggiornamento della forza lavoro. Le tre comunicazioni di accompagnamento hanno ad oggetto in primo luogo iniziative inerenti allo sviluppo sul tutto il territorio europeo di un sistema cloud[27] integrato e ad alte performance, in grado si assicurare la realizzazione di un'economia fortemente competitiva basata sui dati e sulla conoscenza. L'iniziativa europea Cloud è stata progettato al fine di agevolare la scienza, l'industria e le autorità pubbliche nell'accesso alle infrastrutture di dati di livello mondiale in Europa e dei servizi basati su cloud, essendo, questi ultimi, fattori decisivi per il successo nell'economia digitale. L'obiettivo è quello di sviluppare, per la comunità scientifica, un ambiente sicuro e aperto per l'archiviazione, la condivisione e il riutilizzo dei dati e dei risultati scientifici, la cd. *European Open Science Cloud*, attraverso l'implementazione della potenza di calcolo, la connettività ultra-veloce e le soluzioni cloud ad alta capacità. Partendo inizialmente dalla comunità scientifica, la base di utenti sarà ampliata al settore pubblico e poi all'industria, comprese le PMI, garantendo un adeguato livello di sicurezza, la portabilità dei dati, l'interoperabilità e la conformità ai requisiti giuridici dell'UE, e dunque la creazione di soluzioni e tecnologie di cui beneficeranno tutti i settori dell'economia e della società. Il raggiungimento di questo ambizioso obiettivo richiederà uno sforzo collaborativo da parte di tutti coloro che sono interessati a sfruttare la rivoluzione dei dati in Europa come una componente essenziale della crescita globale.

[25] *Proceedings of the 2016 conference on Big Data from Space*, 2016.

[26] European Commission DG CONNECT, *Definition of a research and innovation policy leveraging Cloud Computing and IoT combination*, 2016.

[27] European Commission, *European Cloud Initiative - Building a competitive data and knowledge economy in Europe*, a cui si allegano gli *staff working documents*, *Implementation of the Action Plan for the European High-Performance Computing strategy* e *Quantum Technologies*, 2016.

La seconda comunicazione riguarda la standardizzazione del settore dell'ICT[28] e si propone di garantire che in futuro tutti i dispositivi, tra cui elettrodomestici intelligenti, *connected cars* o *smart factory* con *supply chain* completamente digitali, saranno in grado di connettersi e condividere i dati gli uni con gli altri, indipendentemente dal produttore, dal sistema operativo, o da qualunque altro dettaglio tecnico. La Commissione si focalizza su due priorità principali. Da un lato, la garanzia di un nuovo approccio agli standard nei settori del: 5G, IoT, *cybersecurity* informatica, Cloud e Big Data. D'altra parte, si cerca di assicurare che tutte le forze in Europa vadano nella stessa direzione, con la standardizzazione come strumento strategico per la politica industriale dell'UE.

Ultima comunicazione del pacchetto è quella che riguarda il cd. e-government[29], un insieme di prescrizioni destinate a supportare i processi amministrativi, migliorare la qualità dei servizi e incrementare l'efficienza interna del settore pubblico, attraverso 20 iniziative che saranno lanciate entro la fine del 2017. Il piano è essenzialmente incentrato sulle necessità di imprese e cittadini, ossia su soluzioni che siano online, transfrontaliere e interoperabili per impostazione predefinita (by default) e soluzioni che siano end-to-end sin dalla progettazione (by design). I servizi pubblici digitali ridurranno gli oneri amministrativi per le imprese e i cittadini, rendendo le loro interazioni con le amministrazioni pubbliche più veloci ed efficienti, più convenienti, trasparenti e certamente meno onerose. Inoltre, integrando le tecnologie digitali nelle strategie di modernizzazione dei governi sarà possibile acquisire ulteriori benefici economici e sociali per la società nel suo complesso. La trasformazione digitale dell'amministrazione diventerà, infatti, un elemento chiave per il successo del mercato unico.

Inoltre, va segnalato un ulteriore documento di approfondimento dello Staff della Commissione sull'IoT[30], in cui il fenomeno viene studiato e analizzato in ogni suo dettaglio, dalla sua definizione ai benefici, alle difficoltà di implementazione, all'analisi del mercato e ai suoi trend di crescita futura, fino anche alle sue applicazioni pratiche (*Smart Homes, Personal Wellness and Wearables, Smart Manufacturing, Smart Energy, Smart Cities, Automated Driving/Smart Mobility,Smart Farming*).

Infine tra i più recenti interventi, prima della presentazione del pacchetto "*Building a European data economy*" dello scorso gennaio, vanno sicuramente ricordati: uno studio[31] della fine del 2016 sugli strumenti giuridici esistenti rispetto al tema della proprietà e dell'accesso ai dati, che individua quale, tra gli ordinamenti giuridici europei, abbia la normativa più conforme agli interessi in gioco, effettuando un'analisi dettagliata sulla direttiva sui segreti commerciali, sulla proprietà intellettuale e sulle leggi settoriali che regolano l'accesso ai dati, con particolare riguardo alle clausole contrattuali

[28] European Commission, *Communication on ICT Standardisation Priorities for the Digital Single Market*, 2016.

[29] European Commission, *EU eGovernment Action Plan 2016-2020 Accelerating the digital transformation of government*, 2016.

[30] *Advancing the Internet of Things in Europe*, Commission *Staff Working Document*, 2016.

[31] European Commission DG CONNECT, *Legal study on ownership and access to data*, 2016.

utilizzate per gestire i diritti legati ai dati, nonché all'adeguatezza del quadro giuridico presente valutando la necessità di una riforma.

Ancora, altro interessante studio riguarda l'applicazione dei big data in un settore ben preciso, quello della sanità[32], in cui vengono fornite una serie di indicazioni per i decisori europei e nazionali, tra i cui i membri della rete *eHealth,* per l'ulteriore sviluppo di una catena del valore dei big data nel campo della salute. Sicuramente, la raccomandazione più importante ha ad oggetto la sensibilizzazione sul valore aggiunto che i big data hanno per un settore come quello della sanità e della salute. Una strategia di comunicazione per incoraggiare il settore della sanità pubblica ad orientarsi verso l'implementazione dell'utilizzo dei big data dovrebbe essere basata sullo stimolare continuamente il dialogo aperto con tutte le parti interessate e i gruppi di pazienti, che potrebbero essere non del tutto favorevoli a tale implementazione, attraverso la creazione di una piattaforma europea per lo scambio di esperienze e la discussione su come poter migliorare l'attuale situazione e su come meglio affrontare le sfide future. Inoltre, in concomitanza con il lancio del pacchetto "*Building the European Data Economy*", è stato pubblicato uno studio[33] che, proprio al fine di garantire la libera circolazione dei dati all'interno del DSM, analizza le restrizioni sulla circolazione dei dati imposte in 8 paesi dell'UE (Repubblica Ceca, Francia, Germania, Italia, Lituania, Lussemburgo, Spagna e Regno Unito). Dallo studio si evidenzia che le restrizioni legali al trasferimento di dati intra-UE sono principalmente poste sotto forma di regolamenti di accesso e requisiti di notifica (ad esempio da parte dei regolatori o delle autorità fiscali) a seconda della tipologia di informazioni. Inoltre, alcune aziende hanno requisiti rigorosi sui cd. "*data residency*" che peraltro non sono basati su restrizioni legali formali. Proprio la mancanza di una garanzia sulla libera circolazione dei dati in Europa desta non poche perplessità ai fini della piena attuazione dei piani e delle diverse iniziative tutte concatenate ed unite per la realizzazione di un unico obiettivo: la creazione di un unico mercato digitale aperto a tutti e a vantaggio di tutti.

Tra le altre iniziative volte alla creazione del Digital Single Market, il 10 maggio 2017 è stata resa nota la relazione finale sull'indagine settoriale sul commercio elettronico (avviata il 6 maggio 2015) che completa le proposte legislative della Commissione sul tema e che perseguiva l'obiettivo di individuare eventuali problemi di concorrenza sui mercati elettronici europei. Analizzando i dati di circa 1.900 imprese che effettuano il commercio elettronico di beni di consumo nonché di contenuti digitali e l'esame di circa 8.000 contratti di distribuzione e di licenza, la relazione ha confermato che la crescita del commercio elettronico nell'ultimo decennio e, in particolare, la trasparenza dei prezzi online nonché la concorrenza di prezzo, hanno inciso significativamente sulle strategie di distribuzione delle imprese e sul comportamento dei consumatori.

Anche sul fronte dell'eGovernment, a marzo 2017 la Commissione Europea ha lanciato un'iniziativa avente ad oggetto l'aggiornamento del quadro europeo di interoperabilità. Questa importante iniziativa è finalizzata ad assicurare che tutte le amministrazioni degli Stati membri seguano un

[32] *Study on Big Data in Public Health, Telemedicine and Healthcare*, 2016.

[33] *Facilitating cross border data flow in the Digital Single Market*, 2017.

approccio equivalente concordato quando mettono a disposizione online i loro servizi pubblici e a ridurre i costi e i rischi connessi all'utilizzo dei dati. Attraverso l'interoperabilità dei servizi digitali, infatti, gli Stati membri, offrendo servizi standardizzati, automatizzati e semplificati, garantiranno l'accessibilità dei loro servizi non solo entro i confini nazionali ma anche tra paesi e settori di attività diversi, agevolando così la comunicazione dei cittadini e delle imprese con la loro amministrazione nazionale e con quelle di altri Stati membri e semplificando l'analisi ed i processi decisionali. Le nuove raccomandazioni, in particolare, sono incentrate soprattutto sull'apertura e la gestione delle informazioni, la portabilità dei dati, la governance dell'interoperabilità e l'erogazione di servizi integrati e delineano un piano d'azione piuttosto articolato riguardante le priorità per sostenere l'attuazione del quadro europeo di interoperabilità dal 2016 al 2020. Complementare a tale iniziativa, vi è la proposta di uno sportello digitale unico, adottata il 2 maggio 2017, con l'obiettivo di diminuire gli oneri amministrativi a carico dei cittadini e delle imprese quando si spostano o svolgono attività commerciali oltre frontiera all'interno del mercato unico. Tramite questo strumento sarà possibile un accesso agevole ai pertinenti servizi di informazione e assistenza in modo tale da dare possibilità agli utenti di completare alcune importanti procedure amministrative online ed alle imprese di adempiere gli obblighi amministrativi (registrazione, archiviazione e aggiornamento dei documenti aziendali) online anche dall'estero.

Sul versante dell'intelligenza artificiale, viste le potenziali criticità a questa connessa, prima tra tutte, la tutela dei dati, il Parlamento europeo, nel febbraio 2017, ha adottato una Risoluzione recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica. In questo importante documento vengono descritti i benefici connessi all'utilizzo crescente delle intelligenze artificiali in termini, ad esempio, di salvaguardia dei lavoratori rispetto alle professioni più faticose o pericolose, ma anche, in generale di impatto sul mondo del lavoro e sulle competenze richieste ai lavoratori.

Infine nel settembre 2017 la Commissione ha svelato il piano strategico per la cybersecurity, partendo da un'analisi dei dati relativi ai reati informatici commessi sul territorio comunitario (e non solo). Si tratta di un'iniziativa fondamentale in un contesto in cui la rete, diventando sempre più centrale per lo svolgimento delle attività socio-economiche e raccogliendo enormi quantità di dati di tutti i tipi, diventa anche il principale bersaglio della criminalità. Il piano, con l'obiettivo di incrementare la difesa, la deterrenza e la resilienza dei sistemi informatici, verte su tre pilastri fondamentali: 1) costruire un sistema europeo resiliente e incrementare il livello di cyber sicurezza dell'Unione; 2) creare una effettiva e univoca risposta in campo penale ai reati informatici, adeguando le pene alla gravità delle azioni criminose; 3) potenziare la stabilità globale attraverso il miglioramento della collaborazione internazionale. Tra gli aspetti più interessanti della proposta, un ruolo di primo piano è rivestito dalla creazione di una Agenzia Europea sulla Cybersecurity – frutto del potenziamento della già esistente Agenzia Europea per la Sicurezza delle Informazioni e delle Reti (ENISA) – alla quale viene attribuito un mandato pieno e permanente, con maggiori strumenti e obiettivi riguardo il suo operato, entro il 2020, anno nel quale scade l'attuale mandato dell'Agenzia. L'obiettivo della Commissione UE è di affiancare la nuova Agenzia alle altre Istituzioni europee e gli Stati membri nella definizione e successiva attuazione di una politica comune sulla sicurezza informatica.

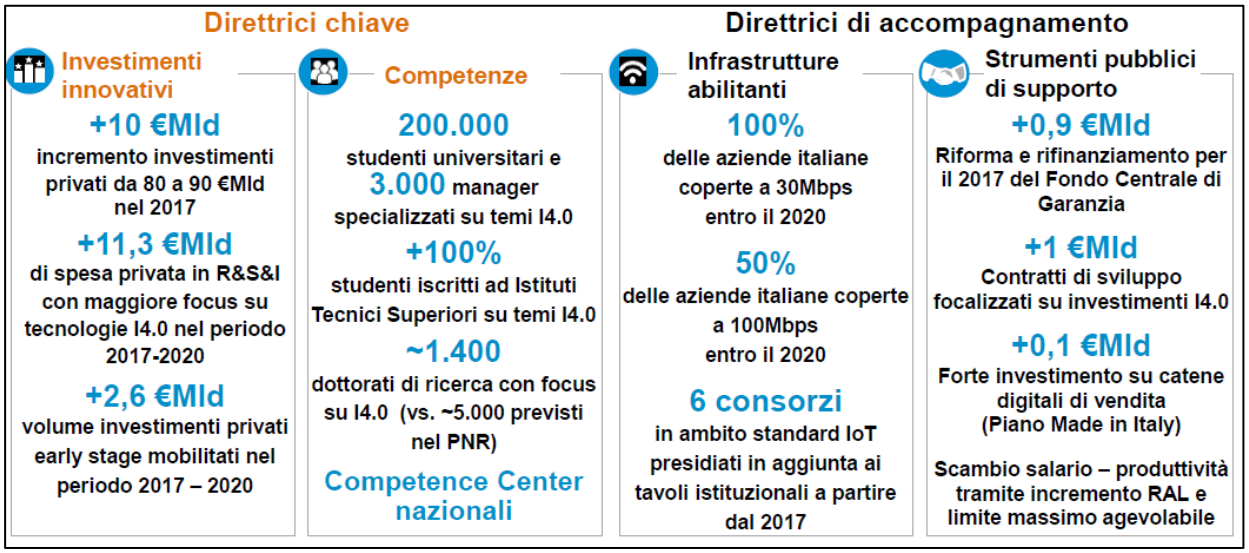
1.3.2 Le iniziative italiane: Piano Nazionale Industria 4.0 e lancio delle sperimentazioni 5G

All'interno dello scenario sopradescritto, in un contesto di particolare fermento, si inserisce il **Piano nazionale Industria 4.0**, presentato a Milano il 21 settembre 2016 dal Presidente del Consiglio e dal Ministro dello Sviluppo Economico. Il Piano rappresenta una guida strategica che getta le basi per attuare una nuova politica industriale incentrata sull'innovazione, che incorpori l'approccio Industria 4.0 e che sia in grado non solo di sostenere il settore manifatturiero del Paese, forza trainante dell'intero sistema economico, ma anche di governare le trasformazioni della società.

Partendo dalla caratteristiche proprie del settore industriale italiano quali: la presenza di pochi e grandi player industriali e ICT in grado di guidare la trasformazione della manifattura nazionale; il limitato numero di capi filiera in grado di coordinare il processo evolutivo delle catene del valore; un sistema industriale fortemente basato sulle PMI; il ruolo chiave ricoperto da importanti poli universitari e centri di ricerca per lo sviluppo e l'innovazione; una forte connotazione culturale dei prodotti finiti, le linee guida che il Governo ha stabilito si sostanziano nella considerazione di operare secondo una logica di neutralità tecnologica; di intervenire mediante azioni d carattere orizzontale non già verticale o settoriale; di operare sui fattori abilitanti; di orientare strumenti esistenti per favorire il progresso tecnologico e la produttività; di coordinare i principali stakeholder ma senza ricoprire un ruolo dirigista.

In particolare, il Piano prevede una serie di direttrici strategiche d'intervento, riportate schematicamente in figura 43, tra cui quelle di incentivare gli investimenti privati su tecnologie e beni 4.0, aumentare la spesa privata in Ricerca, Sviluppo e Innovazione e rafforzare la finanza a supporto di I4.0, Venture capital e start-up.

Figura 43. Le direttrici del Piano nazionale industria 4.0 e gli obiettivi



Fonte: Cabina di regia industria 4.0

Per l'implementazione del presente Piano, il Governo ha previsto investimenti pubblici per 13 miliardi di euro nel periodo 2017-2020 (a cui si aggiungono i 355 milioni di euro per l'implementazione del piano nazionale Scuola digitale e dell'alternanza Scuola-lavoro sui percorsi coerenti col progetto, i 70 milioni di euro destinati alla formazione specialistica, i 170 milioni previsti per il potenziamento dei cluster tecnologici e infine i 100 milioni per i già citati *competence center*. Un totale quindi di ulteriori 700 milioni di euro) che dovrebbero mobilitarne 24 miliardi privati (si evidenzia una crescita di circa 10 miliardi di investimenti privati in innovazione nel 2017 – da 80 miliardi a 90 miliardi – , e di ben 11,3 miliardi di spesa privata in più nel triennio di riferimento per la ricerca e lo sviluppo, oltre ad un incremento di 2,6 miliardi dei finanziamenti privati).

Rispetto alle competenze, l'obiettivo è avere 200 mila studenti e 3 mila manager specializzati sui temi dell'Industria 4.0, raddoppiando il numero degli iscritti agli istituti tecnici superiori focalizzati su questo ambito di specializzazione. A supporto della trasformazione digitale saranno istituiti, sempre a partire dal 2017, *competence center* e *digital innovation hub* nazionali, sei consorzi deputati alla discussione sugli standard dell'IoT e un *roadshow* di sensibilizzazione lungo tutta il territorio nazionale a cui prenderanno parte associazioni, università, aziende testimonial e le più alte cariche istituzionali. Nell'ottica di rafforzare ed incrementare le competenze digitali del paese si collocano anche le iniziative "Scuola Digitale", documento di indirizzo per il lancio di una strategia complessiva di innovazione della scuola italiana e per un nuovo posizionamento del suo sistema educativo nell'era digitale; e "Alternanza Scuola Lavoro" inserita nell'ambito della legge "La Buona Scuola" che ha reso obbligatoria l'alternanza scuola-lavoro nell'ultimo triennio della scuola secondaria di II grado.

Inoltre, tra le misure incentivanti previste, una delle principali è certamente rappresentata dalla proroga del super-ammortamento al 140% sugli investimenti in beni strumentali materiali e dall'introduzione di un iper-ammortamento al 250% sugli investimenti in beni strumentali funzionali alla trasformazione tecnologica e digitale delle imprese legati all'industria 4.0, nonché da una diversa modulazione del credito di imposta per ricerca e innovazione che sarà incrementale, portando l'aliquota della spesa interna fino al 50%, con un credito massimo da 5 fino a 20 milioni di euro.

Merita inoltre pari considerazione la previsione sulle detrazioni fiscali fino al 30% per investimenti fino a un miliardo in start up e PMI innovative, e su altri fondi dedicati all'industrializzazione di idee e brevetti ad alto contenuto tecnologico e al *co-matching* con startup industry 4.0. L'obiettivo è indubbiamente quello di recuperare il gap rispetto agli altri Paesi europei facendo diventare l'Italia quello col più alto tasso di investimento *early stage*.

Infine, va certamente ricordato che nell'ottica di facilitare la cd. quarta rivoluzione industriale, si colloca il piano per la banda ultra-larga, sinergico alla Strategia per la Crescita Digitale ideato per il perseguimento degli obiettivi dell'Agenda Digitale e finalizzato a recuperare il gap infrastrutturale e di mercato, creando le condizioni più favorevoli allo sviluppo integrato delle infrastrutture di telecomunicazione fisse e mobili, con azioni quali:

- agevolazioni tese ad abbassare le barriere di costo di implementazione, semplificando e riducendo gli oneri amministrativi;
- coordinamento nella gestione del sottosuolo attraverso l'istituzione di un Catasto del sotto e sopra suolo che garantisca il monitoraggio degli interventi e il miglior utilizzo delle infrastrutture esistenti;
- adeguamento agli altri Paesi europei dei limiti in materia di elettromagnetismo;
- incentivi fiscali e credito a tassi agevolati nelle aree più redditizie per promuovere il "salto di qualità";
- incentivi pubblici per investire nelle aree marginali;
- realizzazione diretta di infrastrutture pubbliche nelle aree a fallimento di mercato.

In tal senso, all'interno del piano Strategia per la Crescita Digitale e nello specifico nelle direttive di accompagnamento per l'implementazione delle infrastrutture abilitanti, una delle misure particolarmente significative previste è quella relativa all'iniziativa di copertura al 100% delle aziende a 30 Mbps e almeno 50% delle aziende a 100 Mbps, entro il 2020, tramite investimenti pubblici di circa 6,7 miliardi di euro e privati pari a circa 6 miliardi di euro.

La governance del piano industria 4.0 sarà affidata a una cabina di regia gestita in un primo momento da governo e imprese, cui poi successivamente prenderanno parte anche le Regioni. In particolare, la cabina di regia sarà composta dalla presidenza del Consiglio dei ministri, dai ministeri dell'Economia, dello Sviluppo, dell'Istruzione, del Lavoro, delle Politiche Agricole e dell'Ambiente e da una rappresentanza degli atenei tecnici (Politecnici di Bari, Milano e Torino oltre alla Scuola Superiore Sant'Anna di Pisa), dei centri di ricerca, dell'imprenditoria e delle organizzazioni sindacali. La Cabina assumerà il ruolo di una sorta di consiglio di amministrazione avente il compito di controllare e verificare il regolare svolgimento dei lavori e il perfetto conseguimento degli obiettivi prefissati, agendo, se del caso, per apportare eventuali modifiche e correzioni.

Figura 44. Cabina di regia industria 4.0

Fonte: Cabina di regia industria 4.0

Riassumendo, il piano prevede degli strumenti finanziari, più che dei finanziamenti, che saranno certamente fondamentali nelle strategie aziendali al fine di riconvertirsi verso tecnologie nuove e innovative. Il punto chiave è che non c'è bisogno di aiuti alle aziende, ma è necessario che le aziende cambino rotta tornando a investire in R&D e in beni strumentali tecnologicamente avanzati, capaci di analizzare ed elaborare informazioni critiche per il successo aziendale da una mole sempre più crescente di dati a disposizione delle stesse imprese, avendo come prospettiva strumenti finanziari di supporto, quali quelli previsti dal Piano.

Dopo un anno, il Piano ha dimostrato una buona efficacia nel sostenere le imprese che investono in innovazione e avanzamento tecnologico. Nel mese di luglio 2017 si è registrato, infatti, un recupero del 7,7% della produzione industriale dal punto più basso della crisi (settembre 2013). In merito agli investimenti innovativi, tra i beni 4.0, in particolare, la produzione di macchinari è cresciuta del 4% circa da inizio 2016 a luglio 2017, a fronte di un fatturato che nello stesso periodo è aumentato del 15%; Per quanto riguarda gli strumenti pubblici di supporto, il totale finanziato dal Fondo di Garanzia è cresciuto dell'8,9% rispetto allo stesso periodo del 2016 (10,5 miliardi di nuovi finanziamenti bancari) e il totale garantito è cresciuto del 10,7% (7,4 miliardi di garanzie pubbliche); i contratti di sviluppo finanziati sono 102 (la gran parte al Sud), con investimenti previsti pari a 3,6 miliardi di euro e agevolazioni concesse per 1,9 miliardi di euro; inoltre, nei soli primi 7 mesi del 2017, si è registrato un avanzo commerciale di 25,6 miliardi di euro ed una crescita nell'export del 7,6%. Nell'ambito competenze e del lavoro, sono stati aperti 700 bandi di dottorato sul tema Industria 4.0, finanziati i Cluster Tecnologici Nazionali con 1 miliardo di investimento pubblico-privato, coinvolti oltre 21.000 giovani in percorsi formativi ed erogati fino a 3.000 euro di contributo per il tutoraggio in azienda. Tra gli sviluppi futuri, nel 2018 è previsto che al capitolo Industria 4.0 si affiancherà il capitolo lavoro e competenze 4.0 che prevede che le imprese che effettueranno una spesa incrementale in formazione avranno accesso al cosiddetto "credito di imposta su formazione 4.0".

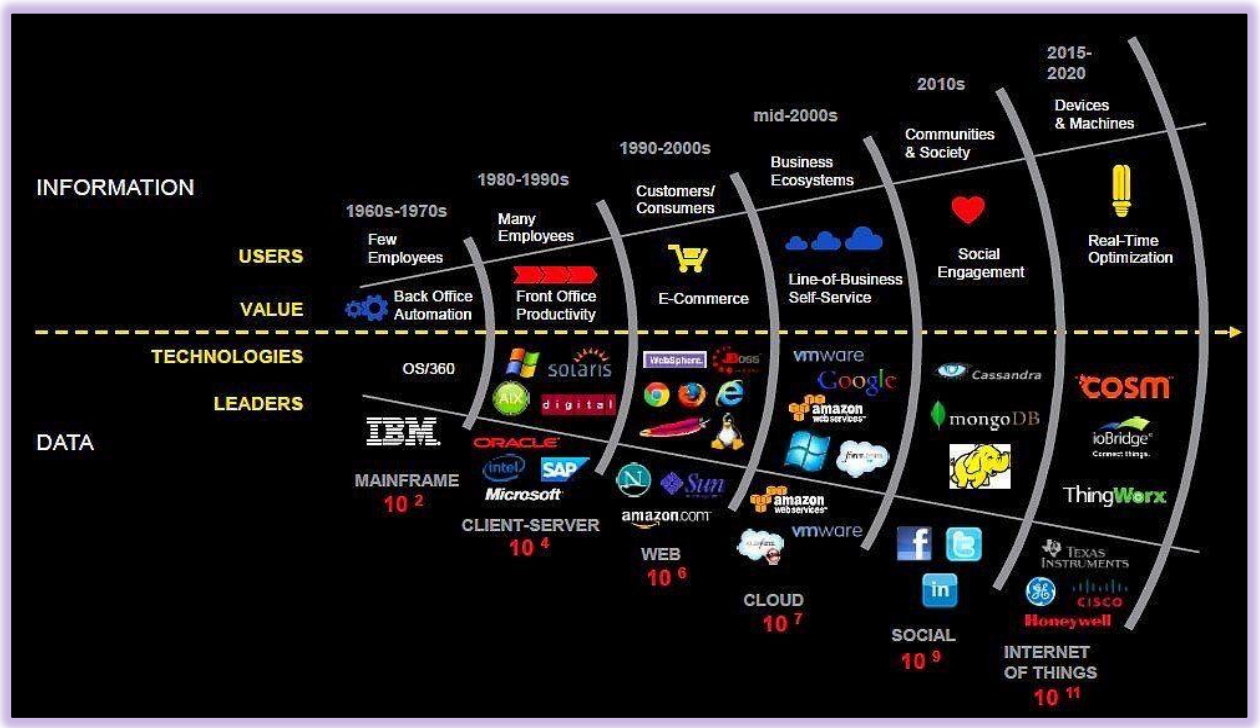
Per quanto riguarda lo sviluppo del sistema 5G, nell'ambito del citato Action Plan, il Ministero dello Sviluppo Economico ha aperto un bando per l'acquisizione di proposte progettuali per la realizzazione di sperimentazioni della rete 5G sia da un punto di vista infrastrutturale che da quello dei servizi che saranno sperimentati. I destinatari del provvedimento sono gli operatori che dispongono dell'autorizzazione per la fornitura di reti e servizi di comunicazione elettronica (o che si impegnino, nella domanda di partecipazione, a conseguirla prima del rilascio dell'autorizzazione provvisoria alla sperimentazione), insieme a diverse tipologie di soggetti – che possono consociarsi in varie forme di aggregazione e partenariato – quali università, enti di ricerca, imprese con competenze nei servizi di comunicazione elettronica, pubbliche amministrazioni e associazioni di categorie (le ultime due purché in forma minoritaria). Ad ogni modo, il soggetto capofila deve essere un operatore di comunicazione. I requisiti dei progetti spaziano dalla realizzazione degli stessi entro specifiche aree, all'impiego efficiente delle frequenze 3700-3800 MHz, all'adozione di soluzioni tecnologiche della famiglia 5G, alla corrispondenza con uno o più casi d'uso tra quelli definiti per il 5G dall'ITU, l'agenzia delle Nazioni Unite che si occupa di definire gli standard nelle telecomunicazioni e nell'uso delle onde radio. Per quanto concerne la durata e l'ambito territoriale, i progetti devono essere realizzati nell'arco di 4 anni nelle seguenti aree geografiche: Area 1: Milano; Area 2: Prato e L'Aquila; Area 3: Bari e Matera.

Capitolo 2. L'economia dei dati

2.1 Il ruolo dei dati nell'ecosistema digitale

Alla luce del quadro fin qui delineato appare evidente come il dato rappresenti il motore della trasformazione nella *digital economy*. Il dato è elemento chiave per lo sviluppo del business ed è proprio per questo motivo che viene considerato "il lubrificante" della nuova economia.

Figura 45. Il nuovo ecosistema digitale



Fonte: ITMedia Consulting, su fonti varie

Non esiste una definizione univoca del termine "dato", inteso in senso lato come una qualsivoglia informazione o come una rappresentazione di più informazioni spesso combinate tra loro e memorizzate all'interno di *data warehouse*. Certo è che non si tratta di un concetto del tutto nuovo. Già nel 2006, il ricercatore di mercato Clive Humby definiva i dati come "il nuovo petrolio". Da sempre poi le informazioni sono parte della realtà d'impresa e dell'intero mercato, ma alla luce dei recenti mutamenti in atto, il ruolo dei dati cambia completamente, diventando parte attiva del cambiamento.

Negli ultimi anni, più che parlare di semplici o singoli dati, si è soliti riferirsi a un concetto molto più complesso e ampio, riconducibile al termine "*big data*", proprio in virtù e con riferimento alle ingenti quantità di dati disponibili all'interno del nuovo ecosistema digitale, prodotti ad alta velocità e provenienti da una moltitudine di fonti, la cui gestione e analisi, come detto, richiedono nuovi e più potenti processori e algoritmi.

Figura 46. Le origini del fenomeno "Big Data"

Fonte: PWC

Il primo elemento di diversità dei dati rispetto al passato è rappresentato dalla loro *quantità*: database ordinari non sono più in grado di gestire il numero sempre crescente di dati, conseguentemente si è reso necessario sviluppare nuovi modelli di database capaci di memorizzare, classificare ed elaborare grandi moli di dati a velocità supersoniche (i cd. *data lake,* vd. par. 2.2.1.1.).

Il secondo aspetto riguarda le modalità di analisi dei dati che, come detto, sono decisamente cambiate, si parla infatti oggi di tecniche di analisi innovative ed avanzate: "*predictive analytics*", "*data mining*" e "*data science*", tecniche queste alla portata di tutti, grazie anche alla presenza di tecnologie *open source* destinate proprio alla loro applicazione. Attraverso tali modalità di analisi è possibile personalizzare una ricerca, un prodotto/servizio, la pubblicità, rivoluzionando completamente le strategie di marketing e in generale di business.

La letteratura prevalente tende a concentrarsi sull'aspetto quantitativo e dunque concernente il *volume* dei dati che, per alcuni autori, rappresenta uno dei principali problemi connessi a questo genere di fenomeno, in ragione del fatto che risulta particolarmente complesso stare al passo con la loro costante crescita esponenziale. Secondo le stime, infatti, ogni giorno ne vengono prodotti circa 2,5 exabyte, un numero elevatissimo.

Tuttavia, il volume non è l'unica caratteristica importante. La *velocità* alla quale i dati vengono generati e resi accessibili è parimenti impressionante. Per citare alcuni esempi, secondo alcune stime del 2016, relativamente ai dati generati sui social, Facebook genera circa 10 miliardi di "like" ogni giorno da parte di 1,09 miliardi di utenti, mentre su Istagram vengono condivise più di 95 milioni di immagini al giorno con una media giornaliera di 4,2 miliardi di "like" e infine su YouTube si registrano oltre 400 ore di contenuti caricati dagli utenti ogni minuto di ogni singola giornata.

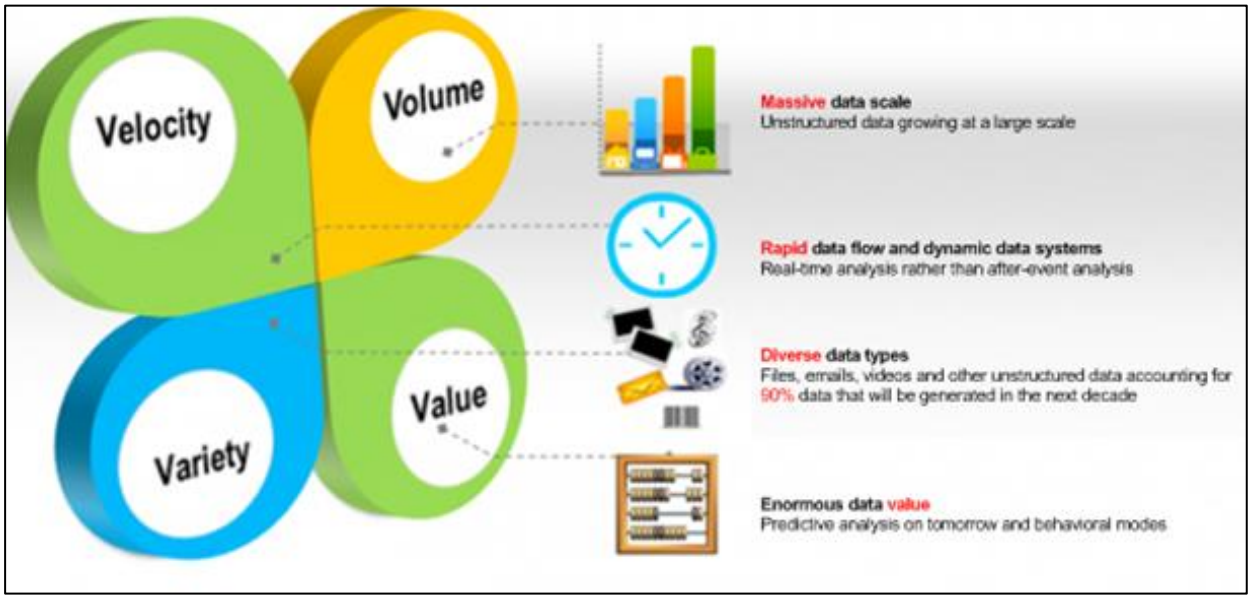
La velocità di sviluppo dei dati tuttavia desta non poche perplessità soprattutto in merito alla loro gestione, in particolare quando tale velocità supera di gran lunga quella di elaborazione. Infatti, perché i dati riescano a garantire le migliori performance, devono essere analizzati a una velocità pari a quella, fulminea, che caratterizza il flusso di informazioni nei *data warehouse*. Nel giro di microsecondi occorre decidere se vale la pena acquisire un determinato dato, valutandone la rilevanza dopo averlo confrontato o combinato con le altre informazioni disponibili.

In alcuni casi, il fenomeno dei *big data* è definito anche in ragione della capacità di analizzare una *varietà* di insiemi di dati non strutturati provenienti da fonti diverse come: registri web, social media, smartphone, sensori e transazioni finanziarie. Ciò richiede la capacità di collegare insiemi di dati diversi, oltre alla capacità di estrarre informazioni da un insieme di dati destrutturati. Il carattere della varietà dei dati fa riferimento essenzialmente alla loro struttura: i dati strutturati entrano in un *data warehouse* già contrassegnati con un tag e sono facili da smistare; oggi, tuttavia, i dati sono per la maggior parte non strutturati, informazioni casuali, difficili da analizzare e gestire.

Queste tre proprietà appena illustrate - volume, velocità e varietà - sono considerate le tre principali caratteristiche dei *big data* e sono comunemente indicate come le tre V. Si tratta di proprietà "tecniche" che sostanzialmente dipendono dalle capacità di memorizzazione dei dati e dalle tecnologie di elaborazione.

Lo sviluppo dell'analisi del fenomeno ha portato a considerare oltre le tre V sopracitate, una quarta V rappresentata dalla caratteristica del *valore*, che dipende fortemente dal crescente potenziale economico e dalla valenza sociale che a loro volta attribuiscono ai dati la qualità di "nuovi fattori di produzione". Il concetto di valore verrà più dettagliatamente trattato nell'ambito del successivo paragrafo.

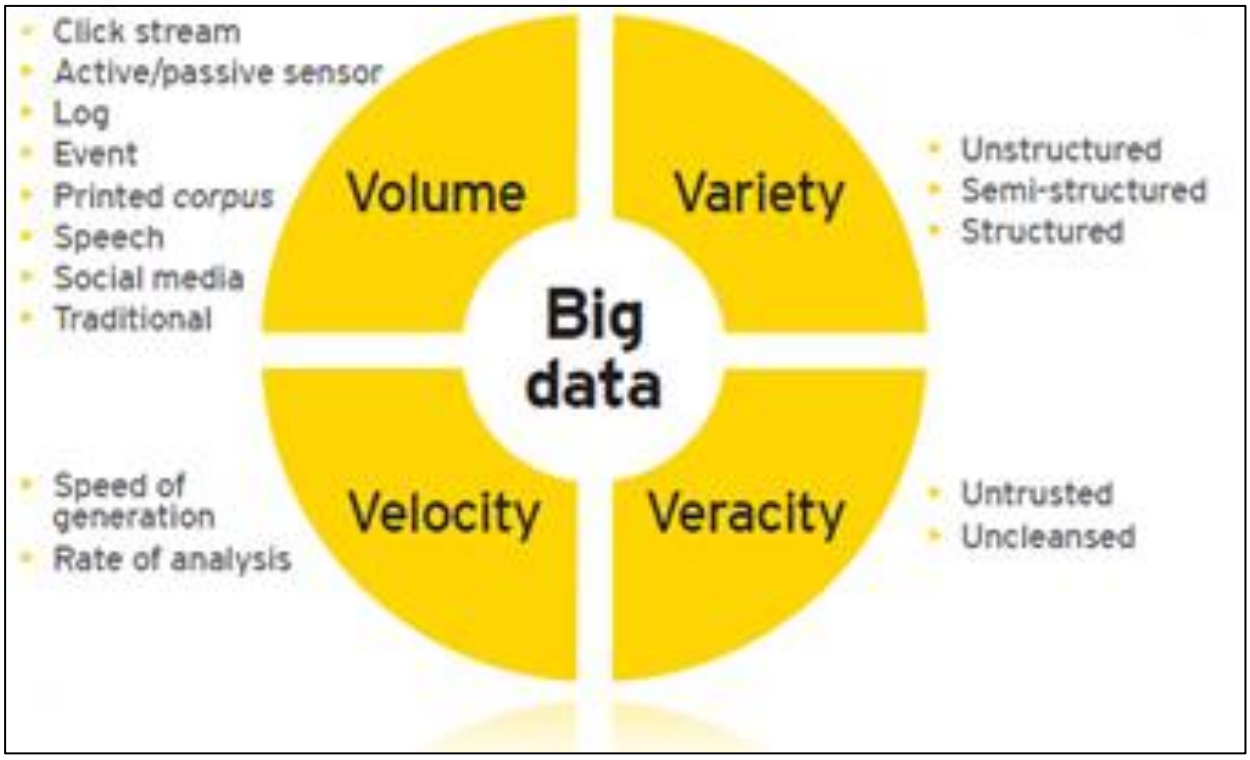
Figura 47. Le caratteristiche dei big data: le quattro "V"



Fonte: IDC

Infine, e più di recente, oltre al valore ovvero in sua alternativa, si considera un'ulteriore V, la *Veracity* (autenticità e affidabilità dei dati), che rappresenta uno dei temi più controversi anche per i correlati aspetti sociali.

Figura 48. La "Veracity"



Fonte: EY

Quanto alle modalità di classificazione esistenti, è possibile distinguere diversi approcci di catalogazione dei dati, con riferimento alla:

(iv) *Tipologia di informazione*: i dati possono essere classificati in base alle diverse tipologie di informazioni che sono in grado di fornire e che, a loro volta, possono avere un diverso grado di utilità per una data impresa. I dati possono, ad esempio, fornire informazioni su individui, sui loro comportamenti, le loro preferenze e la loro posizione geografica; su entità economiche rispetto al fatturato realizzato e al numero delle transazioni commerciali effettuate; ovvero su oggetti, rispetto ad esempio alla posizione in tempo reale o alla velocità di una macchina.

(v) *Struttura del dato*: i dati possono essere differenziati a seconda se siano strutturati o meno. Da questa peculiarità dipende la possibilità di estrarne valore economico e la scelta della modalità necessaria per farlo. Per dati strutturati, ossia dati numerici o di natura finanziaria (dunque tabelle, record, documentazioni di *office automation*, dati M2M, ecc.) è necessario un modello di estrazione che definisca una serie di parametri quali: il settore di riferimento; la tipologia di dati che interessano quel dato settore; e infine la relazione esistente tra i dati stessi. Un esempio di dati strutturati potrebbe essere rappresentato da una base di dati di un

consumatore che contiene informazioni relative al nome, cognome, indirizzo, età, recapito telefonico, ecc. I dati strutturati possono essere più facilmente elaborati e utilizzati per scopi commerciali rispetto a dati destrutturati, soprattutto con riferimento ai metodi tradizionali di elaborazione. Ai dati destrutturati, ossia tutti quei dati che non rientrano in categorie definite (ad esempio testi di e-mail, dati video e audio, dati social, immagini, ecc.), viceversa, non corrisponde un modello specifico di elaborazione e generalmente necessitano di essere trattati da diversi e ancor più innovativi algoritmi perché acquisiscano un determinato valore commerciale. I dati possono anche essere di carattere semi-strutturato: anche in questo caso non esiste un modello predefinito di elaborazione, ma alcuni elementi di differenziazione o campi di appartenenza possono essere individuati attraverso un sistema definito di "marcatore-tipo".

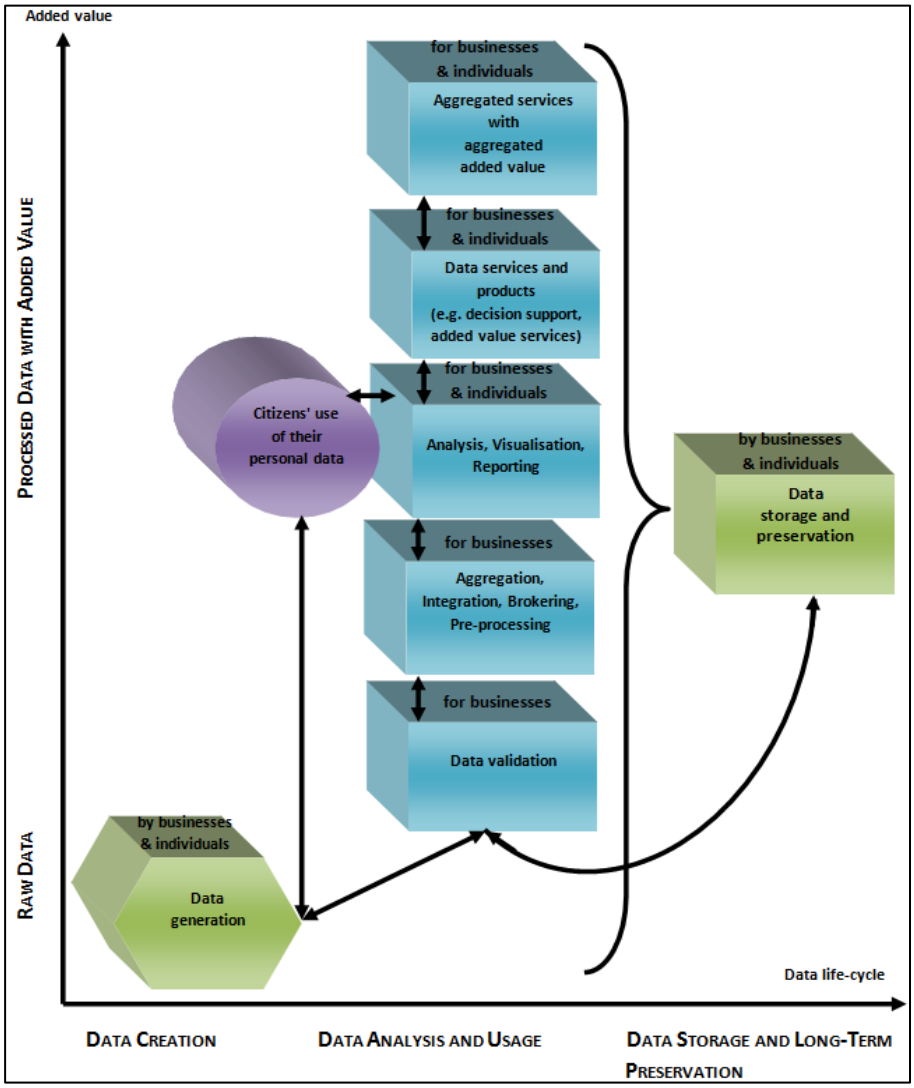
(vi) *Modalità di raccolta*: i dati possono essere raccolti attraverso diverse modalità e a seconda della modalità adottata si registra un diverso impatto sul loro valore economico. Se si considera, ad esempio, il problema della scarsità dei dati. I dati sono perlopiù forniti volontariamente ed attivamente dagli utenti. Ad esempio, per effettuare acquisti online l'utente deve rilasciare tutta una serie di informazioni personali, tra cui indirizzo, dati di pagamento ed e-mail o recapiti telefonici al fine di perfezionare l'acquisto. I social network e le applicazione di messaggistica istantanea multipiattaforma come WhatsApp e Telegram si basano essenzialmente su i dati personali forniti dai loro utenti come nome, indirizzo, sfondo educativo e messaggi personali, foto, video, commenti sulle recenti notizie, le preferenze commerciali ecc. I motori di ricerca basandosi sulle query di ricerca dei rispettivi utenti acquisiscono preziose informazioni rispetto ai loro interessi; una piattaforma video è costruita sui dati di contenuto caricati dagli utenti; una piattaforma di incontri è in grado di proporre un certo individuo ad un altro soggetto solo se gli utenti forniscono informazioni dettagliate e approfondimenti rispetto alle loro vite private. I dati possono essere anche raccolti semplicemente attraverso fonti liberamente disponibili su internet (*open source*) o attraverso lo studio del comportamento dell'utente, pur in mancanza di alcuna informazione personale sull'utente stesso. Esempi del primo tipo sono i cd. motori di ricerca generici che si basano sulla raccolta sistematica e l'elaborazione di ogni pagina del web a loro disposizione - una tecnica nota anche come "*crawling*". La seconda modalità è anche quella più diffusa. Infatti, un gran numero di imprese tracciano il percorso fatto dall'utente da una pagina web all'altra - e talvolta anche all'interno di una singola pagina (tecnicamente è possibile monitorare quale parte di una pagina web viene realmente vista dall'utente) - attraverso tecniche diverse. Ad oggi, queste tecniche sono state ulteriormente sviluppate al fine di incrementare lo studio sul comportamento dell'utente su più dispositivi mobili e non solo sul web. Alcune delle tecniche utilizzate per questo scopo possono però essere facilmente superate dall'utente senza particolari sforzi, altre invece sono difficilmente evitabili. I dati così generati possono anche (tecnicamente) essere facilmente combinati con dati rilasciati volontariamente ad una società, allo scopo di creare profili utente altamente differenziati. Tuttavia, a ben vedere, le leggi sulla privacy possono concretamente limitare il campo di applicazione di questo genere di approccio. Da ultimo, i dati possono essere generati deducendo nuove informazioni utilizzando dati già esistenti. Ad esempio, un sito di e-

commerce potrebbe analizzare i prodotti visualizzati di recente da un utente per dedurre - sia pure con un certo margine di errore – il sesso dell'utente. Allo stesso modo, una società che fornisce diversi servizi web è in grado di combinare i dati degli utenti che abbiano usufruito di questi diversi servizi al fine di ottenere nuove informazioni sul comportamento.

2.1.1 La catena del valore

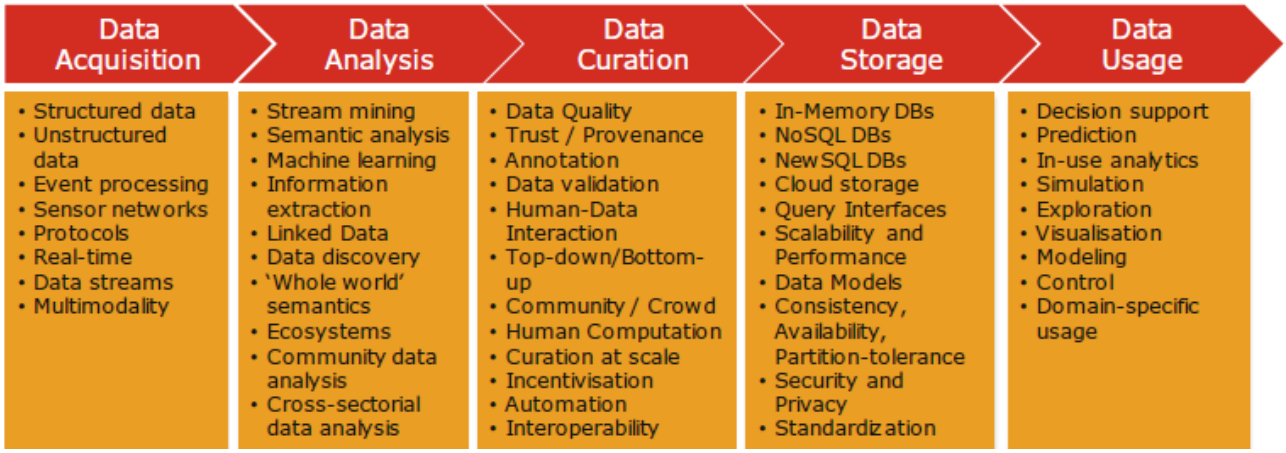
La catena del valore descrive il flusso di informazioni attraverso una serie di step necessari affinché possa essere generato valore dall'insieme di dati considerati. La Commissione Europea già nel 2013 ha ritenuto che la catena del valore dei dati fosse il fulcro dell'economia del futuro, atta a determinare opportunità di sviluppo digitale anche nei settori più tradizionali, quali ad esempio trasporti, servizi finanziari, sanità, manifatturiero, vendita al dettaglio.

Figura 49. Schema del ciclo di vita dei dati e valore aggiunto nella catena del valore dei dati



Fonte: Big data analytics: a survey

Figura 50. La catena del valore dei dati



Fonte: Edward Curry, The Big Data Value Chain, European Data Forum

La catena del valore dei dati, come illustrato in Figura 50, può essere impiegata per coordinare le diverse attività di alto profilo inerenti a un sistema informativo. In particolare:

- *Acquisizione*: è il processo di raccolta, filtraggio, e selezione dei dati prima del loro inserimento in un *data warehouse* o in un qualsiasi altro archivio su cui l'analisi dei dati può essere effettuata. L'acquisizione dei dati rappresenta una delle principali attuali sfide, soprattutto in considerazione dei requisiti infrastrutturali richiesti. L'infrastruttura necessaria a supportare l'acquisizione di dati di grandi dimensioni deve ridurre la possibile latenza nella raccolta dei dati e nell'esecuzione delle query; essere in grado di gestire volumi di transazioni particolarmente elevati; e supportare strutture di dati flessibili e dinamiche;

- *Analisi*: finalizzata a rendere i dati grezzi acquisiti suscettibili di essere utilizzati nell'ambito di un processo decisionale, ovvero in uno specifico settore di utilizzo. L'analisi dei dati comporta l'esplorazione, la trasformazione, e la configurazione dei dati con l'obiettivo di mettere in evidenza quelli maggiormente rilevanti e di sintetizzare ed estrarre informazioni nascoste ad alto potenziale dal punto di vista del business. Aree correlate comprendono il *data mining*, *business intelligence* e il *machine learning*;

- *Cura*: è la gestione attiva dei dati durante il loro ciclo di vita, al fine di garantire il rispetto dei requisiti di qualità necessari per un più efficace utilizzo. Tali processi possono essere classificati sulla base di diverse attività come la creazione di contenuti, la selezione, la classificazione, la trasformazione, la verifica e la conservazione. La cura dei dati viene eseguita da professionisti esperti che sono responsabili del miglioramento dell'accesso al dato e della qualità. I curatori devono inoltre garantire che i dati siano affidabili, individuabili, accessibili, riutilizzabili, e adattabili agli obiettivi prefissati;

- *Archiviazione*: rappresenta la modalità di gestione dei dati in modo modulare atta a soddisfare le esigenze di quelle applicazioni che richiedono un accesso veloce ai dati. Relational Database Management Systems (RDBMS o modelli relazionali) sono stati la

principale e quasi unica soluzione di archiviazione per quasi 40 anni. Tuttavia, le cd. "ACID" (Atomicità, Coerenza, Isolamento e Durabilità), le proprietà logiche che i meccanismi che implementano le transazioni devono avere perché operino correttamente sui dati, e che garantiscono l'indivisibilità della transazione, l'assenza di contraddizione tra i dati archiviati, l'indipendenza dei ciascuna transazione rispetto alle altre, la tracciabilità di tutte le modifiche richieste dalla transazione e compiute sul database di riferimento, risultano, proprio sulla base delle loro caratteristiche, inadatte per gli scenari big data. Tecnologie NoSQL sono stati progettate al fine di introdurre soluzioni di archiviazione basate su modelli alternati che non richiedono uno schema fisso, evitano spesso le operazioni di giunzione e puntano a scalare in modo orizzontale;

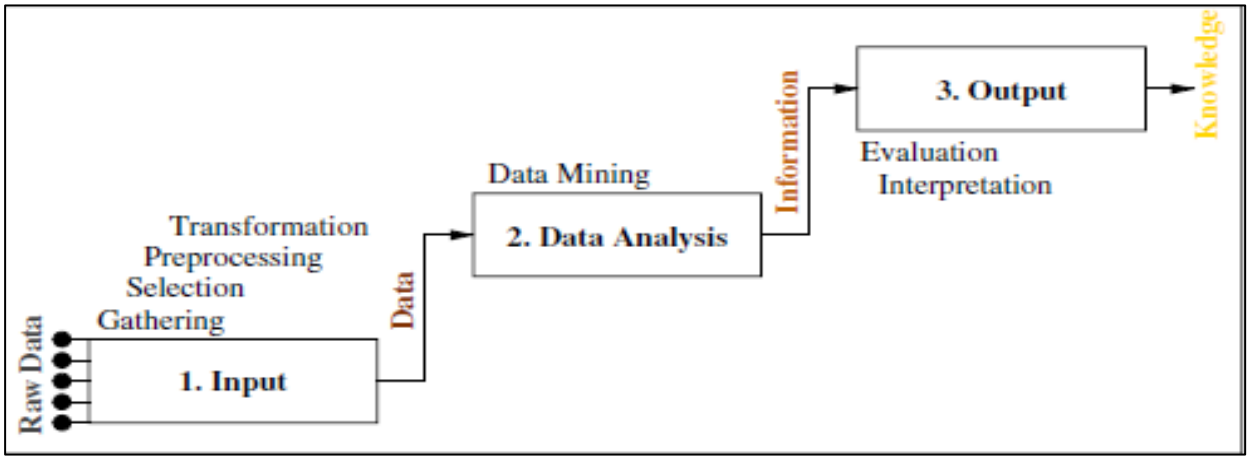
- *Uso*: si riferisce a quell'insieme di attività di business basate sui dati in cui l'accesso, l'analisi e la predisposizione di strumenti atti ad integrare l'analisi dei dati all'interno della attività d'impresa, sono fondamentali. L'uso dei dati nell'ambito dei processi decisionali aziendali può rafforzare la competitività dell'impresa attraverso una riduzione dei costi e un aumento del valore aggiunto.

2.1.2 L'economia dei dati e gli impatti sulle imprese

Il dato dunque rappresenta un *input fondamentale*. Tuttavia, perché possa essere considerato tale e possa poi produrre un output informativo "di successo" è necessario che l'informazione, per così dire grezza, passi attraverso tre distinte fasi, come mostrato nella figura 51: input, *data analysis* e output.

Nella prima i dati destrutturati vengono raccolti, selezionati, pre-elaborati e trasformati al fine di diminuirne la complessità e di migliorare la precisione dei risultati dell'analisi; i dati così selezionati vengono poi analizzati al fine di estrarre l'informazione rilevante attraverso le tecniche sopracitate, per lo più basate sull'utilizzo di algoritmi particolarmente innovativi; infine l'informazione che ne deriva viene valutata e interpretata con l'obiettivo di formare un output di conoscenza coerente rispetto agli obiettivi.

Figura 51. Il processo di scoperta della conoscenza



Fonte: Big data analytics: a survey

Alla base di questo processo vi è un costante miglioramento computazionale in grado di generare migliori processori e unità di elaborazione grafica, combinato con una maggior grado di investimenti in cluster di calcolo enormi.

Tecnologie di *storage* aggiuntivi, tra cui database non relazionali, hanno permesso alle aziende di raccogliere questa enorme ricchezza di contenuti. Tuttavia, gran parte del potenziale è allo stato attuale ancora non del tutto utilizzato. Secondo l'International Data Corporation (IDC)[34], si stima che il mercato delle tecnologie e dei servizi big data crescerà con un tasso composto annuo (CAGR) del 22,6% dal 2015 al 2010, raggiungendo i 58,9 miliardi di dollari nel 2020 – un ritmo di crescita pari a circa 6 volte quello del mercato ICT nel suo complesso. Secondo IDC a livello mondiale la spesa aziendale in *storage* per progetti big data è destinata a crescere con un tasso annuale medio del 22,4% fino al 2020. In particolare, questa spesa andrà a superare i 20 miliardi di dollari nel 2020. In termini di capacità consegnata, tutto questo si tradurrà in un CAGR al 2020 pari quasi al 29%, fino a superare i 92 exabyte nel 2020.

Le nuove applicazioni e soprattutto gli *analytics* hanno comunque le capacità per poter operare anche nei settori più tradizionali che solo fino a poco tempo fa sembravano estremamente lontani dal mondo digitale. Realtà leader negli *analytics*, come Apple, Google, Amazon, Facebook, Microsoft, GE, Baidu, Alibaba Group, e Tencent sono solo alcune delle principali aziende affermatesi a livello mondiale. Queste imprese si sono differenziate grazie alla disponibilità di enormi fonti di dati, *data scientist* altamente qualificati, nonché importanti investimenti in infrastrutture. Stesso discorso può estendersi anche ad altre realtà come Uber, Lyft, Didi Chuxing, Palantir, Flipkart, Airbnb, DJI, Snapchat, Pinterest, BlaBlaCar, Ola, Snapdeal e Spotify, aziende i cui modelli di business sono strettamente legati all'ottenimento di dati e alla loro analisi.

I dati permeano ogni attività svolta da questo genere di imprese: la digitalizzazione delle interazioni con i clienti consente di poter disporre di una ricchezza di informazioni capace di alimentare le dinamiche strategiche, il marketing, le vendite e lo sviluppo del prodotto. Sempre la maggiore disponibilità di dati consente alle aziende di individuare nuovi possibili clienti e di acquisire e sviluppare prodotti e servizi sempre più personalizzati. Inoltre, la digitalizzazione interna alla singola impresa consente l'acquisizione di dati che possono essere utilizzati per rendere le operazioni più efficienti, compresa l'attività di *sourcing* e di gestione della *supply chain* e della logistica; infine, attraverso l'analisi predittiva è possibile ottimizzare i processi di manutenzione delle apparecchiature.

Il valore generato dai dati e dalla loro analisi ha rivoluzionato e completamente ribaltato il tradizionale rapporto tra consumatori e produttori. In passato, le aziende vendevano i propri prodotti ai rispettivi clienti in cambio di denaro e dati di trascurabile valore. Oggi, le transazioni, e in generale ogni interazione con il consumatore, genera preziose informazioni.

[34] International Data Corporation, *Worldwide Big Data Technology and Services 2016-2020 Forecast*, 2015.

Pertanto le aziende sono disposte a offrire servizi gratuiti per ottenerle, e in questo modo, il "cliente" scambia i propri dati in cambio dell'utilizzo gratuito di un prodotto o di un servizio. L'utente è dunque al centro del sistema, i suoi dati hanno un valore primario per le imprese. Si pensi ad Amazon che ha deciso di espandersi a livello internazionale nel settore dei portali video proponendo il servizio Amazon Video Direct capace di combinare tutti i principali modelli di business, dall'hub gratuito e finanziato dall'advertising fino allo streaming ad abbonamento o al noleggio del singolo contenuto, con la chiara intenzione di affermare la presenza del brand in ogni ambito dell'audiovisivo connesso. Allo stesso modo, Apple ha utilizzato i suoi dati unici, infrastrutture e piattaforme per entrare nel mondo della finanza con Apple Pay. Ancora, i giganti cinesi di e-commerce Alibaba, Tencent, e JD.com hanno sfruttato i loro volumi di dati per offrire micro prestiti ai commercianti che operano sulle loro piattaforme. Infine, le banche e le società di telecomunicazioni stanno condividendo dati al fine di proporre sul mercato nuovi prodotti, nonché per migliorare le operazioni fondamentali come la sottoscrizione di un credito, la segmentazione della clientela e la gestione dei rischi e delle frodi.

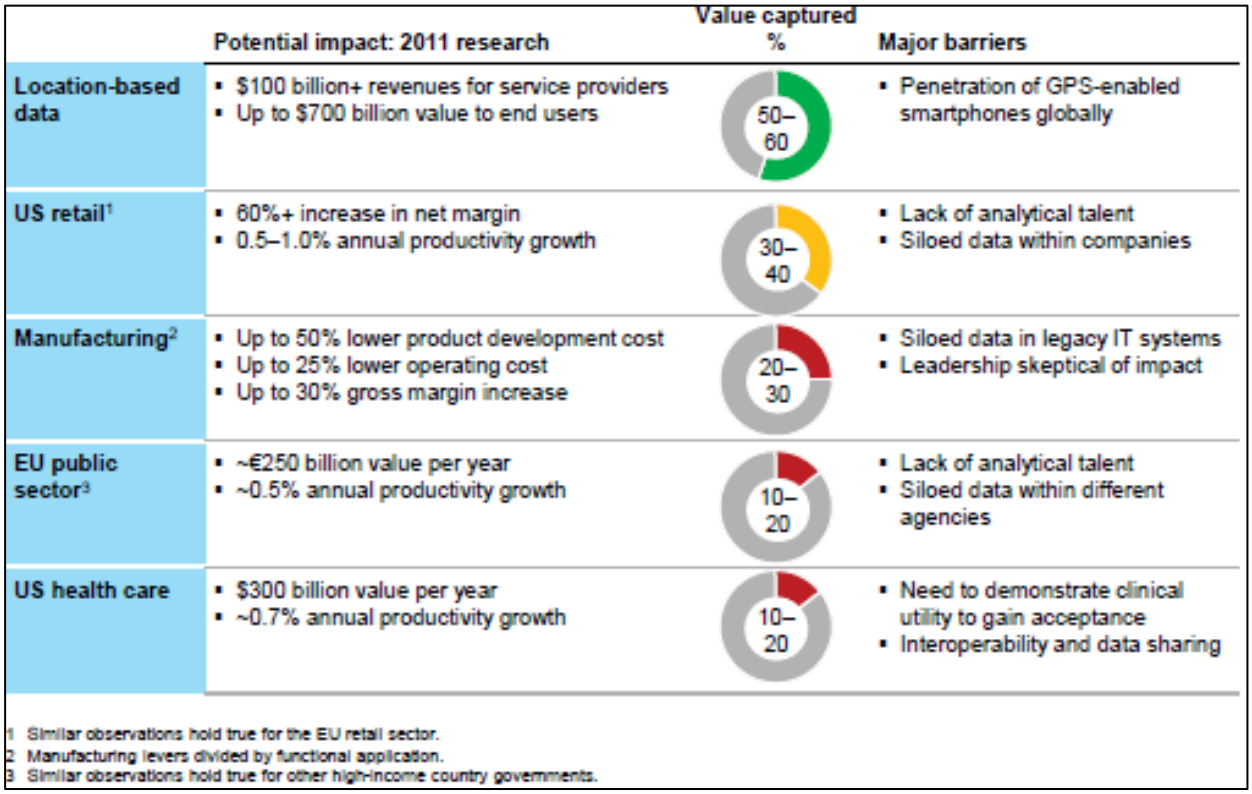
All'interno dello scenario descritto, gli operatori storici al fine di tenere il passo con il ritmo del cambiamento, hanno bisogno di muoversi contemporaneamente in due distinte direzioni. In primo luogo, devono continuare a pensare al futuro, indipendentemente dal fatto che questo significhi entrare in nuovi mercati o cambiare i rispettivi modelli di business. Allo stesso tempo, devono mantenere un focus sull'utilizzo di dati e sugli strumenti di analisi per migliorare il loro *core business*. Quest'attività può comportare l'identificazione di casi di utilizzo specifici, l'ottimizzazione dei processi interni, la costruzione di meccanismi di costante apprendimento e di feedback per migliorare continuamente le proprie performance. Perseguendo questa duplice strategia, tali imprese sarebbero in grado di sfruttare appieno le opportunità di questo cambiamento e di contrastare ad armi pari potenziali "distruttori", tuttavia, va detto, che attuare la suddetta strategia nella pratica è tutt'altro che semplice.

Un recente studio di McKinsey[35], partendo dai risultati di una precedente ricerca del 2011[36], evidenzia come i maggiori progressi si siano registrati nel settore dei servizi basati sulla localizzazione e nella vendita al dettaglio, dove la concorrenza fatta da neo imprese digitali ha spinto gli *incumbent* a reagire e ad innovarsi. Al contrario, nel settore pubblico e nella sanità si sono registrati minori sviluppi in termini di produttività. La figura sottostante riporta sinteticamente i risultati di questo studio ed evidenzia come si registrino progressi irregolari nell'acquisizione di valore dai dati e dalla loro analisi.

[35] McKinsey Global Institute, *The Age Of Analytics: Competing In A Data-Driven World*, 2016.

[36] McKinsey Global Institute, *Big data: The next frontier for innovation, competition, and productivity*, 2011.

Figura 52. I risultati dello studio di McKinsey sull'acquisizione del valore dei dati per settore



	Potential impact: 2011 research	Value captured %	Major barriers
Location-based data	▪ $100 billion+ revenues for service providers ▪ Up to $700 billion value to end users	50–60	▪ Penetration of GPS-enabled smartphones globally
US retail[1]	▪ 60%+ increase in net margin ▪ 0.5–1.0% annual productivity growth	30–40	▪ Lack of analytical talent ▪ Siloed data within companies
Manufacturing[2]	▪ Up to 50% lower product development cost ▪ Up to 25% lower operating cost ▪ Up to 30% gross margin increase	20–30	▪ Siloed data in legacy IT systems ▪ Leadership skeptical of impact
EU public sector[3]	▪ ~€250 billion value per year ▪ ~0.5% annual productivity growth	10–20	▪ Lack of analytical talent ▪ Siloed data within different agencies
US health care	▪ $300 billion value per year ▪ ~0.7% annual productivity growth	10–20	▪ Need to demonstrate clinical utility to gain acceptance ▪ Interoperability and data sharing

1 Similar observations hold true for the EU retail sector.
2 Manufacturing levers divided by functional application.
3 Similar observations hold true for other high-income country governments.

Fonte: McKinsey Global Institute analysis

La maggior parte dei leader aziendali riconoscono la dimensione di opportunità che i dati e la loro analisi offrono e, al contempo avvertono anche il crescere della pressione competitiva. Recenti ricerche hanno dimostrato che investimenti in dati e in capacità di analisi garantiscono in media alti rendimenti: le imprese possono utilizzare queste funzionalità per ottenere incrementi di produttività dal 6 all'8%, che si traducono poi in rendimenti tali da raddoppiare gli investimenti entro dieci anni.

Si tratta di un tasso di rendimento molto più alto rispetto alle altre tecnologie più recenti, superando anche il ciclo degli investimenti dei computer negli anni '80. Tuttavia, va sottolineato come questi alti rendimenti sono in gran parte guidati da solo poche imprese di successo. I primi "innovatori" registrando una crescita più rapida dei profitti operativi, possono a loro volta continuare ad investire nel patrimonio di dati e capacità di analisi, consolidando i loro vantaggi. Facebook, in particolare, ha creato una piattaforma in grado di raccogliere in modo straordinariamente dettagliato dati sui miliardi di utenti in tutto il mondo. Ma non tutti i leader sono nativi digitali. Walmart, GE, Ferrari F1, e Union Pacific sono esempi di aziende in settori tradizionali, i cui investimenti in dati e analisi hanno pagato dividendi significativi sia dal lato dei ricavi che dei costi. Molte altre aziende registrano pesanti ritardi in termini di investimento e innovazione e le barriere, in tali settori, sono rappresentate da limiti di natura per lo più organizzativa. La prima sfida è incorporare dati e analisi in una visione strategica di base. Il passo successivo è lo sviluppo di processi di business e di capacità di costruzione (includendo sia le infrastrutture, sia le competenze).

Il ritmo relativamente lento del progresso in molti casi è indice della presenza di alcuni ostacoli che la maggior parte delle imprese incontrano nel tentativo di integrare l'analisi dei dati all'interno dei loro processi decisionali. Molte imprese hanno risposto alla pressione competitiva, facendo importanti investimenti in tecnologia, senza però adottare le necessarie modifiche organizzative per sfruttare al massimo tali investimenti.

Inoltre, lo sviluppo di soluzioni applicative destinate ai dati non è esente da tutta una serie di problematiche che sostanzialmente vertono sui seguenti aspetti:

- *Descrizione*: molti dataset sono caratterizzati da forti livelli di eterogeneità nella tipologia, nella struttura, nella semantica, nell'organizzazione e nell'accessibilità. La rappresentazione dei dati è funzionale a una più significativa analisi da parte dei computer e ad una più corretta interpretazione da parte dell'utente. Nondimeno, un'impropria descrizione dei dati potrebbe ridurre il valore stesso dei dati originali, inficiando la regolarità dell'analisi;

- *Semplificazione e Compressione*: operazioni di semplificazione e compressione dei dati risultano particolarmente importanti al fine di ridurre il costo indiretto dell'intero sistema sulla base della considerazione che il valore potenziale dei dati non è in alcun modo intaccato;

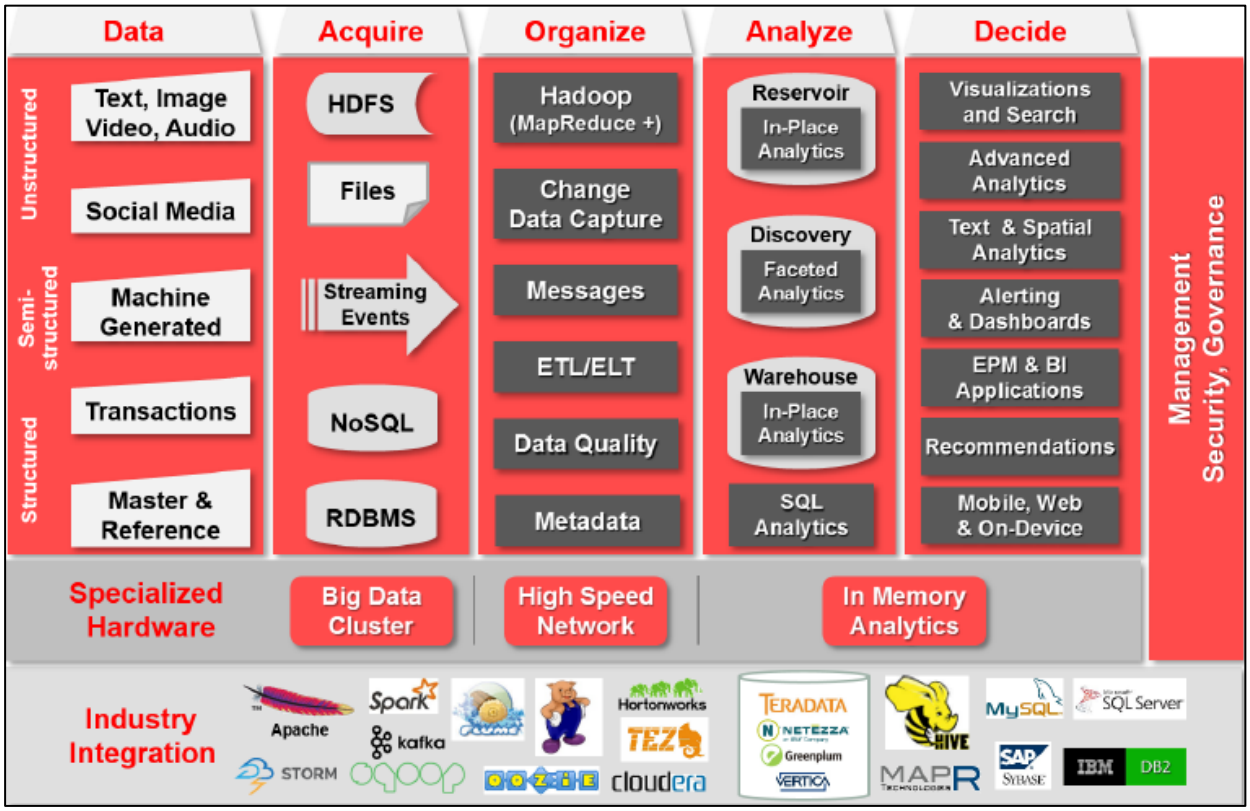
- *Gestione del ciclo di vita*: l'attuale sistema di memorizzazione e archiviazione dei dati rischia di non riuscire a sostenere gli alti livelli di produzione di dati in termini di velocità di produzione e di quantità di dati generati. Generalmente, si sostiene che il valore nascosto dietro una grande mole di dati dipende fortemente dalla loro "freschezza". Pertanto, risulta necessario sviluppare un principio relativo al valore analitico dei dati al fine di decidere quali debbano essere memorizzati e quali invece debbano essere scartati;

- *Meccanismi analitici*: il sistema di analisi dei dati ha il compito di elaborare grandi quantità di dati eterogenei entro un intervallo di tempo limitato. Tuttavia, i modelli relazionali (RDBMS) tradizionali mancano di scalabilità ed espandibilità, non riuscendo a soddisfare i risultati attesi. Viceversa, i modelli non relazionali hanno certamente dei vantaggi nel trattamento dei dati non strutturati, diventando il modello di analisi di riferimento. Tuttavia, anche rispetto ai modelli non relazionali sussistono alcune criticità legate soprattutto alle loro prestazioni, nonché ad alcune loro particolari applicazioni. La soluzione ottimale sarebbe probabilmente quella di trovare un giusto equilibrio tra questi due modelli all'interno di una stessa architettura come mostrato in figura 53.

- *Riservatezza*: l'analisi di dati, soprattutto quelli sensibili, deve essere supportata da adeguate misure di sicurezza per prevenirne l'integrità;

- *Espandibilità e scalabilità dei dataset*: il sistema di analisi di grandi quantità di dati deve supportare set di dati presenti e future. L'algoritmo di analisi deve quindi essere in grado di elaborare dati sempre più estesi e complessi.

Figura 53. Unified Capabilities Information Management



Fonte: Oracle, Information Management and Big Data, Reference Architecture Overview

2.1.3 I dati nel nuovo ecosistema digitale

Secondo quanto appena illustrato, i dati possono essere generati da dispositivi mobili, social network, IoT e molte altre nuove applicazioni, tutte caratterizzate da volume, velocità e varietà (le cd. 3V). Le tecnologie ovvero i fattori abilitanti finora visti, in particolare *cloud computing*, *IoT* e *AI* sono strettamente correlate ai dati stessi.

Relativamente al *cloud computing*, i dati rappresentano l'oggetto delle intense operazioni di calcolo e mettono chiaramente in risalto le capacità di memorizzazione del sistema cloud. L'oggetto principale del *cloud computing* è infatti l'utilizzo di enormi risorse di calcolo e di archiviazione che, attraverso una gestione concentrata, mira a fornire applicazioni di dati. D'altro canto il crescente sviluppo e la diffusione continua di dati ha accelerato lo sviluppo delle tecniche di *cloud computing*. Anche se ci sono molteplici fattori in comune tra dati e *cloud computing*, sussistono almeno due aspetti di differenziazione. In primo luogo, si tratta di due concetti differenti. Il *cloud computing* trasforma l'architettura IT, mentre i dati influenzano i processi decisionali di business. Tuttavia, i dati dipendono dal *cloud computing* essendo quest'ultimo l'infrastruttura fondamentale per il loro regolare funzionamento. In definitiva, mentre il *cloud computing*, con funzioni simili a quelle dei computer e dei sistemi operativi, fornisce le risorse a livello di sistema, i dati operano al livello superiore su supporto del *cloud computing* e assolvono funzioni simili a quelle dei database e dei più efficienti strumenti di elaborazione di informazioni.

Nell'IoT, gli oggetti fisici sono collegati gli uni agli altri e sono costantemente connessi con l'ambiente. Il paradigma dell'IoT prevede infatti l'applicazione di una ingente quantità di sensori all'interno di vari dispositivi, macchine e oggetti fisici. E' dunque evidente come anche questa recente tecnologia sia in grado di dare maggiore impulso alla produzione e alla diffusione dei dati, delle più disparate tipologie: ambientali, geografici, astronomici, logistici, ecc. In questa fase, ciascun dato raccolto può potenzialmente essere impiegato come risorsa da parte di soggetti economici o enti pubblici per la realizzazione di scopi ben diversi rispetto a quelli per i quali quel dato è stato in origine raccolto. Per esempio, le cd. automobili intelligenti o *connected cars* collezionano dati per implementare lo sviluppo del mercato dell'auto a guida autonoma e per fornire servizi di manutenzione migliori, tempestivi e predittivi. Ma le automobili possono anche registrare le abitudini di guida del conducente, informazioni che potrebbero essere di interesse per le compagnie di assicurazione, come pure la posizione geografica del veicolo in un determinato momento temporale, che viceversa potrebbe informare il fornitore di dati geografici di un cambiamento di direzione di un strada a senso unico, ovvero le stesse autorità pubbliche circa il volume di utilizzo e le condizioni del traffico stradale in un certo momento. I benefici, per così dire sociali, dipendenti dai dati aumenteranno con il loro inserimento in un data set molto più ampio, capace di mettere insieme dati provenienti da più fonti, ad esempio da diversi produttori di automobili per ottenere un quadro più completo ed esaustivo rispetto alla situazione del traffico all'interno di una particolare area geografica.

Il carattere innovativo di questo tipo di utilizzo dei dati consiste nel collegare grandi serie di dati al fine di rispondere a bisogni molto diversi tra loro sulla base di semplici correlazioni tra diversi tipi di dati (*data mining*) nell'interesse di singole imprese o del settore pubblico. E' dunque importante sottolineare che i dati generati dall'IoT hanno caratteristiche differenti rispetto alla generalità di dati fino ad ora trattati. Basti pensare alle diverse tipologie che possono essere raccolte, le cui principali peculiarità sono: l'eterogeneità, la varietà, l'alta ridondanza e la mancanza di struttura.

Nonostante, allo stato attuale, i dati generati dall'IoT non rappresentano la quota maggioritaria del totale dei dati presenti nell'ecosistema digitale, si stima che già nel 2030 il numero dei sensori applicati raggiungeranno il trilione, generando un quantitativo di dati tale da rappresentare la quota più alta rispetto alla totalità dei dati presenti. Allo stato, la capacità di elaborazione di dati dell'IoT necessita urgentemente dell'introduzione di tecnologie basate su dati atte a promuovere lo sviluppo stesso dell'IoT. Molti operatori dell'IoT sono consapevoli dell'importanza dei dati in quanto il loro successo dipende fortemente dall'integrazione dei dati e del *cloud computing*. E' stato ampiamente riconosciuto che dati e IoT sono assolutamente interdipendenti e dovrebbero essere congiuntamente sviluppati: da un lato, l'ampia diffusione dell'IoT alimenta la forte crescita dei dati, sia in quantità che in varietà, garantendone al contempo lo sviluppo e l'applicazione; dall'altro, l'applicazione della tecnologia basata sui dati all'IoT accelera anche la ricerca rispetto ai vantaggi e ai modelli di business dell'IoT stessa.

La raccolta di dati dai sensori permetterà infatti ai consumatori di fruire di prodotti intelligenti e servizi innovativi che sostituiranno mano a mano quelli tradizionali. D'altro canto, i dati raccolti in questo settore potranno essere di particolare utilità per gli attori privati in settori aziendali molto diversi, nonché per gli enti pubblici.

Anche nel campo dell'Intelligenza Artificiale, la tecnologia basata sul marketing predittivo è uno strumento estremamente utile, ma potrebbe essere ancor più efficiente. Secondo un recente studio condotto da Conversica[37], la stragrande maggioranza delle aziende non riesce a dar seguito a circa 1/3 dei propri principali obiettivi di business.

Ed è qui che entra in gioco l'AI, come strumento di ausilio e di sostegno. Questi "esseri virtuali" una volta assegnato un nome, un indirizzo e-mail e un titolo, sono in grado di eseguire tutto il lavoro preliminare e assistenziale alla vendita del prodotto/servizio. Questo sistema garantisce che non si registri quella parte di perdita che viceversa un errore umano avrebbe comportato e per questo l'AI consente di creare nuove opportunità di lavoro, diversamente da quanto si possa pensare.

Infatti, ad esempio, quando nel 1990 i bancomat sostituirono la figura dei cassieri, le banche registrarono una maggior livello di efficienza, ogni filiale era in grado di svolgere più lavoro a costi molto più bassi, il che portò ad un maggiore ritorno economico. Conseguentemente, se da un lato le singole banche avevano meno personale; dall'altro, l'incremento dell'utile permise alle aziende bancarie di aprire più filiali e dunque di assumere più dipendenti.

Lo stesso discorso vale per l'AI: se i team di vendita diventano più efficienti, aumentano il loro ROI – e di conseguenza aumentano le opportunità di lavoro. Tutto questo ci dice che l'AI, alimentata da dati, a loro volta gestiti e elaborati da algoritmi, non si sostituisce all'uomo, ma si sostanzia in applicazioni rivolte a rendere più efficienti le operazioni di business a vantaggio delle imprese e dei soggetti che ne fanno parte.

Saper analizzare, aggregare, e visualizzare in maniera leggibile i dati significa cogliere nuove opportunità di business non sfruttate in precedenza, creando quindi importanti vantaggi competitivi attraverso la semplificazione e la maggiore efficienza dei processi[38].

2.1.4 I mercati per la commercializzazione dei dati digitali

Ad oggi i mercati dei dati e, in particolar modo, le piattaforme sulle quali si scambiano dati stanno diventando realtà sempre più popolari[39], soprattutto a vantaggio delle imprese che producono dati a latere di altri processi industriali[40]. Ecco perché non pochi data scientists[41] hanno evidenziato come i

[37] Vedi https://resources.conversica.com/h/i/304887618-2016-sales-effectiveness-report-4ps-of-lead-follow-up

[38] In tal senso, un esempio è rappresentato dal gruppo Ferrovie dello Stato, che, in tempo reale, proprio grazie ai Big Data e agli Analytics ha potuto monitorare e portare a un livello di efficienza maggiore tutta la divisione di interscambio delle merci tra ferro e gomma.

[39] OECD, supra, p. 4, riporta l'opinione espresso da Gregory Manne sulle enormi dimensioni della data *broker industry*.

[40] Thomas L.D.W., Leiponen A., *Big Data Commercialization*, 44 IEEE Eng. Manag. Rev. 74, 2016; Stahl F., Schomm F., Vossen G., Vomfell L., *A classification framework for data marketplaces*, 3 Vietnam J. Comput. Sci. 137, 2016.

potenziali entranti possano ottenere dai dati acquisiti sul mercato informazioni non dissimili da quelle estratte da coloro che raccolgono in prima battuta quei dati.[42]

Guardando a questi mercati, giova ricordare come la commercializzazione dei dati miri a monetizzarne il valore attraverso la loro circolazione, piuttosto che attraverso la loro analisi *in house* volta ideare nuovi e migliori prodotti[43]come nel caso delle imprese *data driven*. Si tratta, infatti, di approcci differenti che in taluni casi possono comunque risultare complementari.

Come descritto approfonditamente in seguito, l'approccio di tipo "*data driven innovation*" mira allo sfruttamento di qualsiasi tipologia di dato all'interno dei propri processi di innovazione per la creazione di valore tramite nuovi beni, servizi, strategie di marketing e più in generale decisioni aziendali realizzate e pianificate sulla base dei dati a disposizione dell'impresa.

Nel caso della commercializzazione dei dati, si punta, invece, a monetizzarne il più possibile il valore attraverso la loro diffusione e circolazione. Lungo tale processo, le imprese possono rivestire differenti funzioni, ossia possono agire come *data suppliers, data managers, data custodians, application developers, service providers* e *data aggregators* anche in rapporto alle caratteristiche del settore.

Segnatamente, soprattutto in mercati dove le informazioni risultano particolarmente frammentate e dove la condivisione dei dati non mette a repentaglio i rispettivi modelli di business, i *data suppliers* forniscono a terzi dati sì riutilizzabili, ma grezzi e poco differenziati, ossia dati che richiedono minimi sforzi e limitate strategie di differenziazione per renderli disponibili. Essi cioè generano valore giusto attraverso la vendita dei dati da loro stessi generati ed organizzati in banche dati – vendita che poi può avvenire secondo diversi schemi, dall'abbonamento, al pay-per-use, alle inserzioni pubblicitarie,

---

[41] Calandrino J., Kilzer A., Narayanan A., Felten E., Shmatikov V., *You might also like: Privacy risks of collaborative filtering"*, 2011; Edelman B., *"How to launch your digital platform,* Harv. Bus. Rev. 90, 2015; Narayanan A., Shmatikov V., *Robust de-anonymization of large sparse datasets*, 2008.

[42] Lambrecht A., Tucker C.E., *Can Big Data Protect a Firm from Competition?*, pp. 5-6, 2015. Gli Autori riportano i seguenti esempi di *datasets* disponibili in commercio, quali «Acxiom, [which] has 'multi-sourced insight into approximately 700 million consumers worldwide' with over 1,600 pieces of separate data on each consumer; Datalogix, [which] asserts that its data 'includes almost every U.S. household' [… and …] the credit-scoring agency Experian, [which] sells frequently updated data on expecting parents, along with income and first-birth information». Inoltre, gli Autori affermano che «Comcast is planning to license TV viewing data collected through set-top boxes and apps [and that] [o]ther companies, such as the Oracle owned Bluekai, sell cookie-based user information online to allow for targeting advertising based on a user's past activities or demographics. [In addition,] Bluekai states that it has data on '750 million unique users per month with an average of 10-15 attributes per user».

[43] Diversi studi si sono focalizzati sui nuovi mercati emergenti di dati, nei quali le intuizioni derivanti da big data e i big data stessi vengono scambiati tra le imprese. In particolare, Gartner ha stimato che il 30% delle imprese punterà sulla monetizzazione dei propri asset informativi a partire dal 2016 – vedi *Gartner Predicts 30 % of Businesses Will Be Monetizing Their Information Assets Directly by 2016*, Online Press Release, 2014.

al freemium[44]. Possono quindi operare quali data suppliers le imprese che, situate in punti di intenso traffico-dati, sono tra le poche (se non le uniche) ad accedere a certi dati, come comunemente accade nei settori quali quello delle telecomunicazioni, dei media e dell'intrattenimento. Un esempio potrebbe essere rappresentato da un'impresa di telecomunicazioni che semplicemente vende i dati generati dai propri utenti ad una seconda impresa che fornisce servizi di informazione sul traffico stradale. In alternativa si potrebbe considerare un'impresa che mette a disposizione i propri dati di consumo prodotti da operazioni commerciali non correlate. Ad esempio, nel caso di un'impresa di trasporto è possibile raccogliere grandi quantità di dati sulle spedizioni dei prodotti a livello globale e creare un'unità aziendale finalizzata alla vendita di dati per attività supplementari e previsioni economiche.

I *data managers* sono invece rappresentati da quelle imprese che catalogano, puliscono e analizzano le informazioni che altrimenti, considerando il formato in cui sono state reperite, non potrebbero essere utilizzate. Queste imprese sono quindi in grado di aggiungere valore ai dati grezzi e semi-grezzi migliorando l'efficienza, l'interpretabilità e in generale la funzionalità dei dati. Ad esempio, attraverso la trasformazione del formato di archiviazione di una risorsa informativa, la traduzione da una lingua all'altra, la piena corrispondenza tra domanda e offerta informativa, tali imprese si assicurano un aumento del valore dei dati. In questo modo, i *data managers* facilitano l'offerta informativa o, più in generale, l'utilizzo di big data, ma non sono essi stessi utenti o riutilizzatori dei dati. Anche queste imprese hanno modelli di prezzo simili a quelli visti per i *data suppliers*, compreso l'abbonamento, il modello *pay-per-view* e la pubblicità.

I *data custodians*, che invece sembrano prediligere soltanto la soluzione per abbonamento, sono imprese che consentono il riutilizzo e la rivendita di big data attraverso la fornitura di un'infrastruttura affidabile. Un importante aspetto di questo modello di business è che agisce come un "*trust framework provider*" che fornisce servizi di gestione delle identità tra utenti e operatori che desiderano utilizzare e/o raccogliere dati. Ad esempio, per quel che riguardano gli aspetti di privacy, i *data custodians* offrono prodotti e servizi che garantiscono agli individui il controllo sulle proprie informazioni personali. Inoltre, i *data custodians* rassicurano anche gli utenti finali e i consumatori dei prodotti big data attraverso la convalida della provenienza, la certificazione e i servizi di audit, al fine di assicurare che l'integrità e la qualità dei dati vengono mantenute dal *sourcing* fino al loro effettivo utilizzo.

Gli *application developers* sono rappresentati da quelle imprese e quei fornitori di software che progettano, costruiscono e vendono applicazioni che consentano la commercializzazione dei big data. Queste organizzazioni possiedono l'esperienza necessaria per sviluppare il software che crea risultati facilmente interpretabili e comprensibili per la mente umana, come visualizzazioni e *mashup*. Questi fornitori di servizi sono soliti instaurare diverse partnership con aziende analitiche e tecnologiche, con fornitori di dati e partner industriali per sviluppare collettivamente soluzioni di analisi innovative per clienti, imprese e utenti finali. Il prezzo per le applicazioni di dati è inteso alla

[44] In questo caso le imprese forniscono dati di base per spingere i consumatori a "impegnarsi" a utilizzare servizi correlati, addebitando un premio per l'accesso a informazioni più dettagliate.

stregua di quello per gli sviluppatori di app, dove un singolo utente o un'organizzazione paga per il diritto di utilizzare lo strumento tecnologico.

Ancora, impiegando svariati modelli di prezzo, i *service providers* sviluppano nuovi servizi basati sui big data, distinti dalla rivendita, dall'analisi o dal re-assemblaggio di dati o dallo sviluppo di specifiche applicazioni. In tal modo, offrono servizi direttamente ai consumatori dei dati, fornendo "intuizioni" da applicarsi a specifici contesti, e che dunque necessitano di un maggior sforzo e una maggiore differenziazione strategica, ma che al contempo generano anche maggiori entrate rispetto alla semplice vendita di dati grezzi. L'utilizzo di *data science*, *data mining*, modelli predittivi e analisi dei risultati, consente alle organizzazioni di offrire nuovi servizi a valore aggiunto e di creare piattaforme per transazioni basate su dati, come ad esempio il *targeting* degli annunci, i pagamenti al dettaglio e la fornitura di informazioni sui clienti. Il più avanzato di questi servizi può rappresentare scenari di tipo "*what if*", in cui i big data acquisiscono maggior valore raccontando un'informazione specifica - ad esempio, quale segmento di clienti è il più redditizio e quali sono le loro abitudini commerciali, gli interessi, la situazione finanziaria e i piani di viaggio degli utenti. In altre parole, i *service providers* possono al contempo aggregare, gestire i dati e creare applicazioni per renderli disponibili al momento giusto e per un particolare scopo.

Infine, nei mercati dei dati si contano i *data aggregators* i quali rappresentano una tipologia specifica di fornitori, in particolare, trattasi di imprese che si concentrano sulla raccolta, l'aggregazione e la ri-determinazione dei big data, incentrati su un tema specifico di settore. I *data aggregators* cercano, incrociano e contestualizzano i dati in modo da trovare le correlazioni, individuare le efficienze o visualizzare eventuali interrelazioni esistenti. Queste intuizioni vengono successivamente fornite come servizi a valore aggiunto alle imprese e ai consumatori e, in alcuni casi, agli enti di governo. L'esempio più comune di *data aggregators* è rappresentato dai servizi di confronto dei prezzi, come il motore di ricerca di viaggi Kayak. Altri esempi includono le start-up sanitarie che integrano dati clinici, di pagamento, di sanità pubblica e i dati comportamentali per aiutare le altre imprese a gestire i costi, come la concessione delle licenze e l'analisi dei dati sui risultati clinici per i pagatori e le autorità di regolamentazione al fine di migliorare il processo decisionale clinico.

Allo stesso modo, la combinazione di dati di posizione con altri tipi di dati comportamentali può fornire un importante valore aggiunto soprattutto per quelle imprese che sanno già a cosa puntare. Quanto ai modelli di prezzo, ve ne sono diversi che possono trovare applicazione, soprattutto in virtù del fatto che spesso i *data aggregators* sono mercati multi-versante che quindi riducono i *bottlenecks* per acquirenti e venditori, facilitando le transazioni. Questi modelli generano valore per gli acquirenti e per i venditori attraverso un miglioramento dell'efficienza del mercato, come il volume delle transazioni, l'efficienza di allocazione delle risorse e una migliore corrispondenza tra offerta e domanda. Questo miglioramento dell'efficienza del mercato si ottiene grazie al sovvenzionamento di un versante del mercato considerato, al fine di trarre profitto da eventuali altri versanti. Il modello di prezzo per le piattaforme più diffuso è l'addebito per l'accesso (ad esempio, l'adesione) o per l'utilizzo. Per quei *data aggregators* che non sono mercati multi-versante (piattaforme) i modelli di prezzo sono simili a quelli visti per i *service providers*.

I modelli di business appena analizzati, fondamentali per la creazione di valore, non sono reciprocamente esclusivi, poiché un'impresa può creare valore utilizzando più di un approccio contemporaneamente. Ad esempio, i *data suppliers* possono fornire dati grezzi, oppure possono trovare modi innovativi per renderli più preziosi e più attraenti per gli utenti a valle. Quindi le organizzazioni con una specifica competenza di base che generano ricchi pool di dati grezzi possono vendere i dati sostenendo bassi investimenti o possono elaborare e analizzare i dati grezzi per fornire servizi dati. Allo stesso modo, i risultati ottenuti dai *data aggregators* hanno molto in comune con i *data service providers*, anche se hanno obiettivi distinti. In particolare, il valore di un'impresa che opera attraverso i dati dipende dalla sua capacità di differenziarsi sul mercato in termini di scala informativa e di scarsa disponibilità dei dati e sulla capacità di combinare elementi differenziati con le fonti tradizionali di informazioni utilizzando *data management* e capacità di analisi. Inoltre, sembrerebbe pacificamente condiviso il fatto che le imprese stanno avanzano lungo le diverse fasi dell'attività di commercializzazione dei dati, della catena del valore, passando dalla semplice vendita e commercializzazione, all'adozione di modelli di business più complessi, come l'aggregazione dei dati e la fornitura di servizi. Questi modelli di business più complessi creeranno entrate maggiori, maggiori opportunità, tuttavia, risulteranno certamente di difficile implementazione.

In tale prospettiva, è cambiata anche la definizione di *big data commercialization,* onnicomprensiva di ciascun modello di business analizzato, non più circoscritta alla sola vendita di dati, intuizioni o servizi, ma caratterizzata da una piena interazione con i dati degli utenti finali allo scopo di eseguire specifiche tipologie di transazioni. L'esempio emblematico è quello di Google, che sfrutta i dati raccolti sulle query di ricerche e attraverso l'attività pubblicitaria per ottimizzare e sfruttare specifiche tipologie di transazioni.

2.1.5 *Data driven innovation*

L'innovazione è un processo iterativo finalizzato alla creazione di nuovi prodotti, processi e servizi mediante l'uso di conoscenze nuove o già esistenti.

Con il termine "*data driven innovation*" si intende lo sfruttamento di qualsiasi tipologia di dato all'interno dei processi di innovazione per la creazione di valore. Lo sviluppo emergente dell'innovazione *data-driven* sta portando alla creazione di beni, servizi, strategie di marketing e decisioni aziendali realizzate e pianificate sulla base dei dati a disposizione delle imprese in tutti i settori industriali. Infatti, secondo anche quanto sostenuto nel 2015 dall'OCSE[45], la richiesta di dati deriva principalmente dalle attività produttive a valle, per le quali il dato rappresenta una risorsa non banale. La stessa OCSE stima una crescita della produttività delle imprese, correlata alla *data-driven innovation* del 5-10% e una riduzione dei costi amministrativi, per gli enti pubblici, del 15%-20% grazie ad una maggiore efficienza, un maggiore gettito fiscale (proveniente, ad esempio, dall'erogazione di servizi personalizzati) ed un minor rischio di frodi o errori.

[45] OCSE, *Data-Driven Innovation Big Data for Growth and Well-Being*, 2015, http://www.keepeek.com/Digital-Asset-Management/oecd/science-and-technology/data-driven-innovation_9789264229358-en#.WT-p3caxUdU#page1.

Secondo il Parlamento Europeo[46], inoltre, la data economy conta per l'1,9% del PIL europeo (1,6% di quello italiano), incidenza che si stima possa crescere al 3,3% nel 2020 (2,7% per l'Italia).

All'interno di questo scenario, il dato si pone come una risorsa fondamentale per offrire agli attori economici l'accesso ad una vasta gamma di prodotti e servizi, sia nel settore pubblico che privato. In altre parole, i dati garantiscono potenzialmente significativi ritorni in termini di economie di scala e di scopo. Come si dirà meglio nel paragrafo successivo, le grandi innovazioni basate sui dati sono implicitamente associate a un modello di catena del valore che descrive le modalità attraverso cui i dati saranno raccolti, organizzati, selezionati, trasformati in prodotti o servizi e distribuiti.

A livello organizzativo, almeno due categorie di iniziative strategiche risulterebbero dal paradigma della *data-driven innovation* e dalla sottostante *data value chain*.

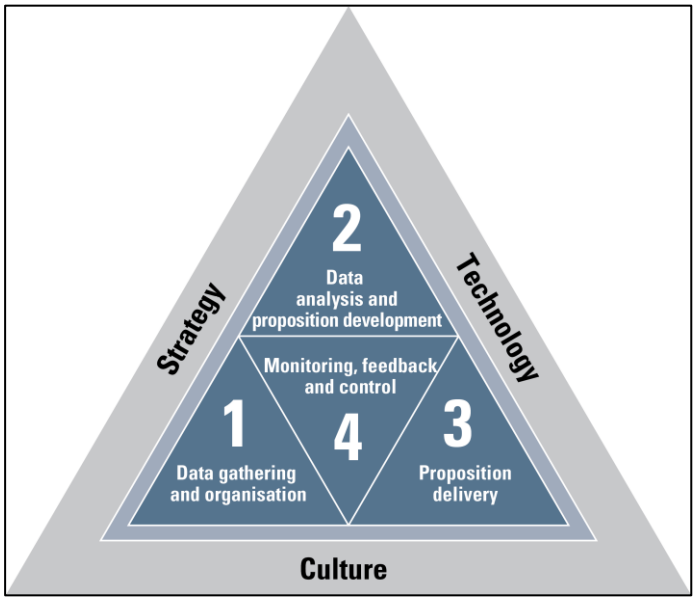
La prima concerne gli aspetti relativi all'organizzazione interna dell'impresa, ed è finalizzata a rendere fruibili le informazioni di cui l'impresa già dispone, allo scopo di accrescere l'efficienza e l'efficacia dei processi e dei servizi aziendali. Al riguardo, si segnala che la maggior parte delle organizzazioni aziendali dispongono di un *Information Management systems* (IMS) non particolarmente funzionale. Molti sono costituiti da molteplici silos frammentati di informazione, ognuno dei quali contiene un numero limitato di dati finalizzato ad una ristretta applicazione. Come risultato della frammentazione e della mancanza di una visione architettonica, semplici modifiche, come il *provisioning* di nuovi dati, nuove relazioni, cambiamenti di gerarchie, ecc, diventano particolarmente complesse, ostacolando l'attività di business e del settore IT. Thomas H. Davenport, in un suo articolo pubblicato nel 2006 nell'Harvard Business Review[47], ha sostenuto che le imprese che sono in grado di rendere operative le decisioni basate sui fatti godranno di ingenti ritorni economici.

Se da un canto, ci sono senza dubbio diversi esempi di come le tecnologie basate sui dati possano ridurre i costi di impresa; dall'altro è altrettanto evidente che i dati portino a focalizzare l'attenzione sulla capacità di analisi dell'organizzazione per trovare prima, e monetizzare poi, le informazioni supplementari, acquisite da nuovi dati o da dati acquisiti a un livello inferiore dell'organizzazione. In questo modo, i dati concentrano l'attenzione del management sull'analisi, accelerando il passaggio verso un approccio di gestione basato sulle informazioni, spesso fonti di ulteriori opportunità per l'analisi o per l'applicazione di analisi al processo decisionale operativo. Le soluzioni IMS possono spesso diventare molto fragili nel corso del tempo, comportando un aumento dei costi operativi e una riduzione della produttività del settore IT. L'approccio basato sui dati, se correttamente implementato, può produrre ampi benefici in questo ambito.

[46] ICom, Convegno *B2C Revolution. Come rendere il digitale un ecosistema di successo per consumatori e imprese*, 2016.

[47] Davenport T.H., *Competing on Analytics*, 2006.

Figura 54. Semplificazione del modello funzionale per l'analisi dei dati



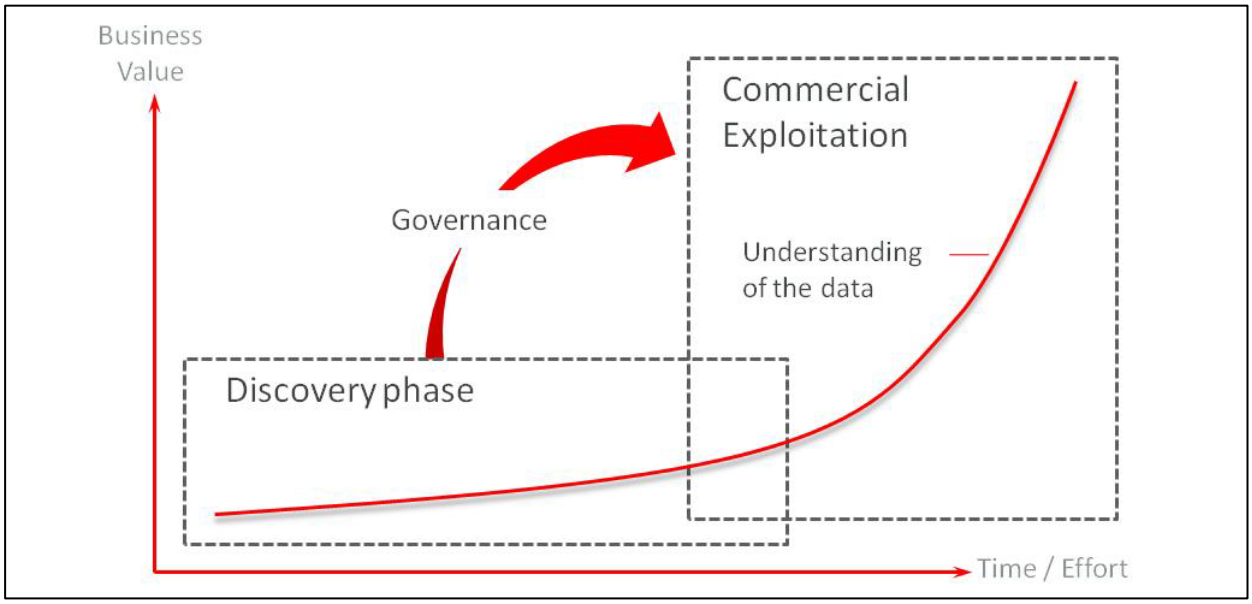
Fonte: Oracle, An Enterprise Architect's Guide to Big Data, Reference Architecture Overview

Il processo, schematizzato in figura 54, mostra come per sfruttare appieno il valore insito nei dati a disposizione di un'organizzazione aziendale. E' fondamentale notare che la portata operativa della soluzione riportata in figura è delimitata da tre dimensioni fondamentali: la strategia, la tecnologia e la cultura aziendale.

Per massimizzare il potenziale dei dati, infatti, è necessario trovare un equilibrio tra queste tre dimensioni. Il tutto non avrebbe alcun senso in mancanza di una strategia di business, che non possa essere supportata da un'idonea organizzazione IT, ovvero dalla presenza di una diffusa cultura aziendale. La mancanza di equilibrio tra queste tre dimensioni può verosimilmente spiegare perché la cultura degli *Analytics* è raramente diffusa all'interno dell'intera organizzazione aziendale, ma è per lo più applicata ad un insieme limitato di problemi o processi aziendali in un ristretto cerchio di dipartimenti.

Per *Analytics* intendiamo il sistema computazionale/statistico di analisi dei dati. Il ruolo del *data scientist* è strettamente legato all'*analytics* e si concentra nell'estrarre valore dai dati mediante un approccio scientifico. Un'altra versione di questo modello funzionale è mostrato in figura 55. Essa evidenzia come, quando guidato da una nuova sfida imprenditoriale e dalla disponibilità di nuovi dati, un *data scientist* è in grado di analizzare questi stessi dati insieme ad altri dati di cui dispone l'impresa, utilizzando una varietà di strumenti e tecniche.

Figura 55. La scoperta e lo sfruttamento economico di nuovi dati



Fonte: Oracle, An Enterprise Architect's Guide to Big Data, Reference Architecture Overview

È importante osservare come le persone, i processi e gli strumenti utilizzati durante la fase di *discovery* saranno diverse rispetto a quelle richieste per la fase di sfruttamento commerciale: c'è uno step obbligatorio per passare dall'una all'altra. Si tratta della Governance che ricopre un ruolo fondamentale per il successo della gestione complessiva delle informazioni. L'obiettivo è di minimizzare il tempo necessario per la fase di *discovery* del processo e ridurre al minimo le distanze tra la scoperta e lo sfruttamento commerciale dei dati.

I Business Intelligence Competency Centre (BICC) continuano a svolgere un ruolo cruciale in molte organizzazioni nell'adozione e nell'aumento del valore e dell'efficacia degli strumenti di BI. Come precedentemente sottolineato, l'adozione di "*analytics*" è spesso frammentaria, limitata ad una piccola cerchia di dipartimenti o a specifiche problematiche del business. Se il valore dei dati è vincolato alla capacità dell'organizzazione di fornire nuove "intuizioni", è chiaro che bisogna porre particolare attenzione ad ampliare l'adozione degli *analytics* a tutta l'organizzazione. L'*Analytical Competency Centre* (ACC) fornisce il contesto organizzativo e l'insieme delle risorse necessarie per il raggiungimento di questo obiettivo.

La seconda serie di iniziative riguarda lo sfruttamento dei dati dei clienti, al fine della predisposizione di forme di relazione e pubblicità personalizzata. Ne sono esempi gli algoritmi utilizzati da Netflix relativamente allo studio delle preferenze sui film e sulle serie televisive dei propri utenti e da Google per quel che concerne le *query* di ricerca.

Negli Stati Uniti, centinaia di aziende utilizzano *open data*, e più in generale le tecnologie basate sui *big data*, come risorse fondamentali per generare valore in diversi settori di mercato, tra cui quello finanziario, dell'istruzione, il settore ambientale, immobiliare, alimentare/agricolo. Tuttavia, il paradigma della *data driven innovation*, se adottato, non garantisce automaticamente la creazione di

valore o di un business di successo. E' fondamentale che l'impresa si doti preliminarmente di una strategia di business, che sulla base di adeguati investimenti tecnologici, sia in grado di sfruttare al meglio il potenziale economico dei dati, associandolo ad un obiettivo ben definito e concretamente realizzabile.

Nel settore sanitario, la grande rivoluzione dei dati è ancora in una fase molto precoce e il maggiore potenziale inerente alla creazione di valore e allo sviluppo del business rimane ancora del tutto sconosciuto. La ragione di questo ritardo sta nell'azione di ostruzionismo promossa da parte del sistema sanitario verso il paradigma *data-driven innovation,* che ostacola di fatto la condivisione e lo scambio di dati. Solo attraverso la stretta collaborazione tra tutte le parti interessate sarà possibile garantire un aumento del valore nel trattamento sanitario del paziente e conseguentemente favorire l'adozione di applicazioni basate sui dati.

Il settore pubblico è sempre più consapevole del valore potenziale che può essere acquisito tramite il paradigma *data driven innovation* in termini di miglioramenti in efficacia ed efficienza e di sviluppo di nuovi strumenti analitici. Le pubbliche amministrazioni generano e raccolgono quotidianamente grandi quantità di dati.

Tuttavia, i principali ostacoli al pieno sfruttamento dei dati come risorsa, peraltro di natura non tecnica, in questo settore sono:

(i) *L'interoperabilità*: frammentazione della proprietà dei dati e dei conseguenti silos di dati

(ii) *Il sostegno legislativo e la volontà politica*: i procedimenti legislativi sono spesso troppo lunghi e lenti per stare al passo con le tecnologie e le opportunità di business in rapida evoluzione

(iii) La privacy e la sicurezza

(iv) *Le competenze*: oltre alla mancanza di tecnici specializzati, c'è una ancor più grave mancanza di conoscenza sul potenziale dei dati da parte di chi fa business

Il settore finanziario-assicurativo è l'esempio più chiaro di industria *data-driven*. I dati rappresentano un'opportunità unica per la maggior parte degli istituti creditizi, lo sfruttamento dei dati dei clienti è in grado di trasformare il loro business, di realizzare nuove opportunità di ricavo, di gestire al meglio il rischio e di rafforzare il rapporto fiduciario con il cliente.

In particolare, occorrerà tener conto di diversi elementi in rapporto agli obiettivi di business:

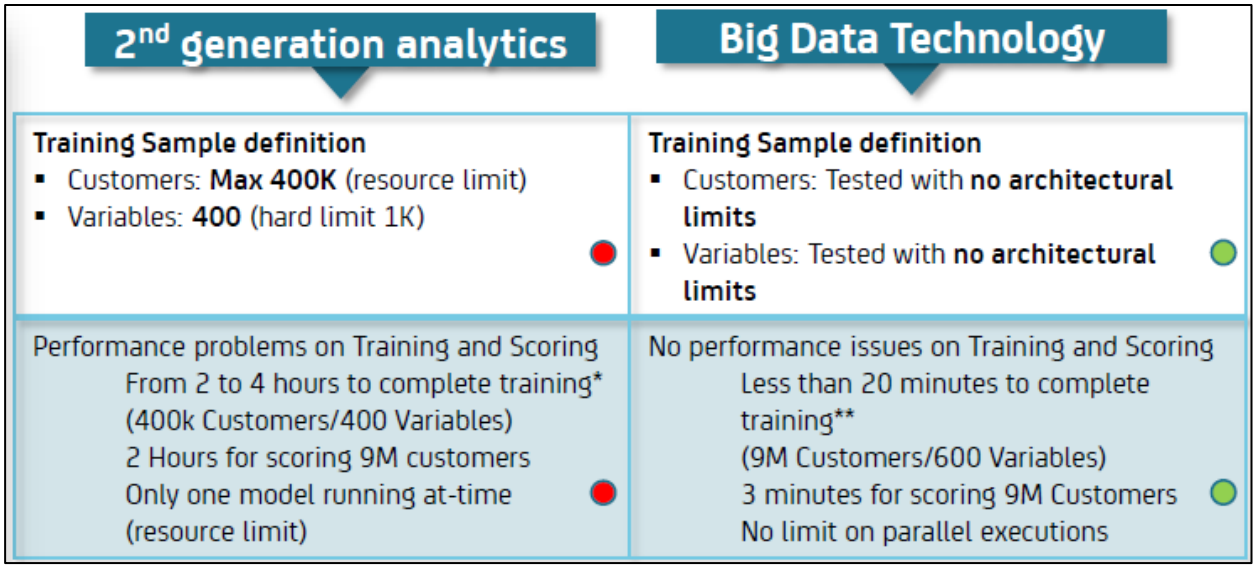
- l'arrivo di una nuova generazione di clienti con aspettative molto alte;
- l'ingresso sul mercato di nuovi concorrenti in grado di offrire servizi sempre più innovativi;
- la presenza dei dati al centro della rivoluzione digitale;
- la crescita esponenziale dei dati in termini di volume, varietà e velocità;

- le aspettative dei clienti rispetto alle capacità degli istituti finanziari di saper gestire i dati finanziari in modo strettamente confidenziale e affidabile;
- la presenza di importanti sfide derivanti da un possibile uso improprio dei dati, dalla loro sicurezza e da asimmetrie informative.

È evidente l'importanza che i dati rivestiranno in tale contesto su entrambi i lati del mercato: i consumatori potranno beneficiare di prodotti e servizi più innovativi e personalizzati; le imprese potranno adottare nuove strategie di business e diventare maggiormente competitive.

Vero è che le società di servizi finanziari da tempo lavorano abitualmente con enormi volumi di dati, provenienti da un'ampia varietà di fonti. Molte imprese dunque utilizzano già da tempo le tecniche di *business intelligence*; ciò che sta cambiando sono le modalità di utilizzo e la tipologia di dati impiegati (passando da dati strutturati a dati non strutturati). Le nuove tecnologie consentono di svolgere le stesse analisi ma in minor tempo, con una maggiore copertura, con maggior precisione e con investimenti molto meno onerosi.

Figura 56. Due generazione di analisi dei dati a confronto



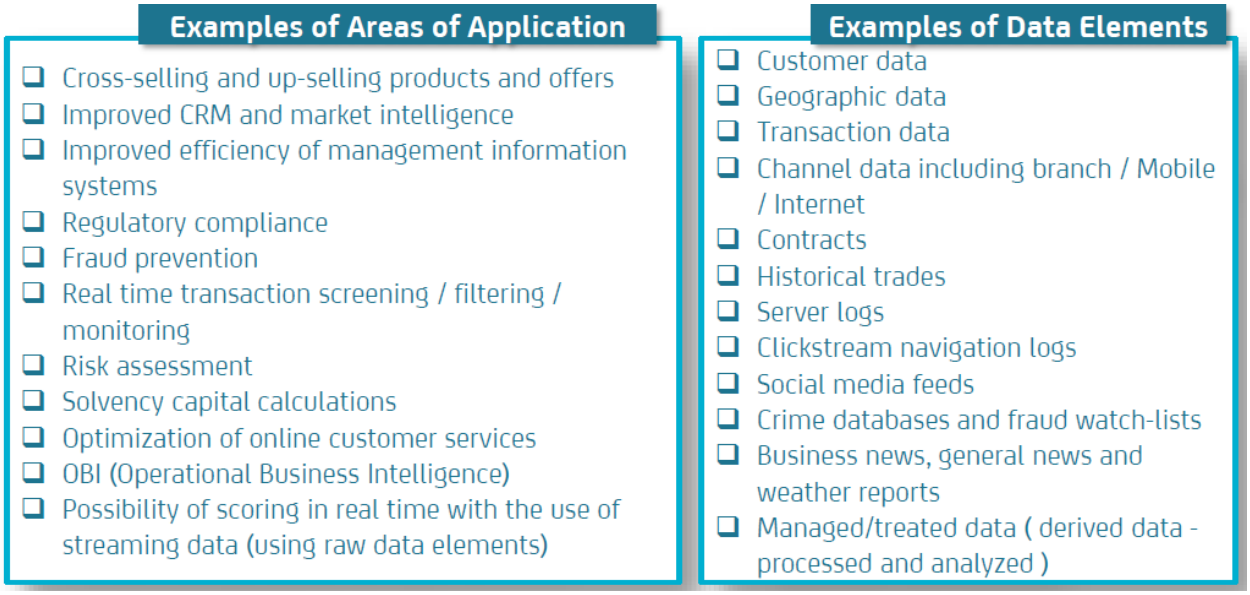
Fonte: Unicredit

I dati rappresentano pertanto l'elemento chiave per assumere decisioni in tempo reale, e in alcuni casi, gli strumenti analitici sono ad oggi gestiti al di fuori dei reparti IT. Alla luce dell'attuale situazione, un fattore critico di successo potrebbe essere rappresentato dalla diffusione della cultura del dato in tutti i settori aziendali. L'implementazione di nuove tecnologie e soluzioni basate sui dati, può infatti essere ostacolata dalle risorse tecniche esistenti: la paura del cambiamento, la mancanza di cultura, o la paura per la sicurezza del posto di lavoro potrebbero tradursi in resistenza al cambiamento, bloccando di fatto le iniziative *data based*. È dunque necessario "educare" le risorse tecniche, esplicitando i vantaggi propri delle soluzioni e delle tecnologie *data based*. Le differenze negli approcci architettonici, il caricamento di dati e i processi ETL (*Extract, transform, load*), la loro

gestione e analisi, ecc devono essere chiariti per comprendere come le nuove soluzioni *data based* si inseriscono nell'architettura informativa complessiva.

Le tecnologie basate sui dati consentono di cambiare i confini dell'impresa, tra quello che è interno e quello che è esterno alla realtà aziendale, favorendo l'inserimento del cliente negli ambiti di logistica interni all'azienda (*customer intimacy*). Studi recenti dimostrano che il 71% delle imprese nel settore dei servizi finanziari a livello globale stanno esplorando le tecnologie *data-driven* e di analisi predittiva e circa il 70% di queste sostengono che i dati sono di fondamentale importanza. Il 50% delle imprese ha nominato uno *Chief Data Officer*. Nel 2015, in questo settore, sono stati investiti circa 6,4 miliardi di dollari in dati, ci si aspetta un aumento di questi investimenti ad un tasso annuo totale del 26% da oggi fino al 2019[48]. Sempre nel 2015 un sondaggio IDC ha affermato che il 35,5% degli istituti considerano i dati come la loro prima priorità per i prossimi anni[49].

Figura 57. Esempi applicativi nel settore financial services & insurance



Fonte: Unicredit

Un concetto chiave è certamente quello che viene definito "*customer intimacy*": seguire l'evoluzione delle esigenze dei clienti, in tempo reale, aiutandoli a raggiungere i rispettivi obiettivi finanziari; capire rapidamente le necessità del cliente e presentare istantaneamente l'offerta più adeguata; informali ed aggiornarli costantemente; identificare meglio le esigenze dei clienti insoddisfatti; migliore utilizzo dei feedback dei clienti per aumentare la qualità del servizio offerto; aiutare i clienti a trarre vantaggio da ecosistemi esterni (ad esempio, la collaborazione con Fintechs); fidelizzazione.

[48] Accenture, Exploring Next Generation Financial Services: The Big Data Revolution, 2016.

[49] IDC, Key success Factors for digital transformation in the Banking Industry, 2015.

In definitiva, gli istituti finanziari hanno la possibilità di sfruttare le soluzioni big data non solo per le loro esigenze interne, ma anche di utilizzare i dati al fine di offrire nuovi servizi a valore aggiunto ai propri clienti finali. Inoltre, possono anche creare nuovi ecosistemi in cui i partner propongono migliori offerte e promozioni sulla base dei comportamenti sull'acquisto dei clienti, analizzati dall'istituto finanziario. I clienti possono quindi beneficiare di una maggiore rispondenza dei servizi rispetto ai propri bisogni e di migliori prezzi. Naturalmente, questi nuovi servizi devono essere progettati in linea con le esigenze di privacy e di sicurezza, e spesso dovranno fornire approfondimenti sulle modalità di aggregazione e su come i dati vengono anonimizzati.

2.1.5.1 La *blockchain*

Proprio nell'ambito del settore bancario-finanziario va menzionata e approfondita un'ulteriore tecnologia che sta recentemente prendendo piede e che è destinata ad essere applicata ben oltre i confini entro i quali si è sviluppata, si tratta della *blockchain*. La tecnologia *blockchain*, inizialmente utilizzata per la gestione della moneta virtuale Bitcoin, è un database distribuito che sfrutta la tecnologia *peer-to-peer,* consentendo lo scambio di valute, la loro tracciabilità e soprattutto di conoscere l'identità dei soggetti che effettuano le relative operazioni. In altre parole è il libro contabile in cui sono registrate tutte le transazioni fatte in Bitcoin dal 2009 ad oggi, e in cui si "registra" l'identità dei titolari di valuta. Il cd. *bank ledger* diventa in realtà un "*distributed ledger*" accessibile da qualsiasi utente che effettui una transazione ed entri quindi a far parte della "catena di distribuzione", cui è affidato il controllo dell'intero sistema o di una parte di esso (tutte le informazioni del 'libro mastro' sono distribuite e condivise da tutti i soggetti del network, cioè da coloro che partecipano alla *blockchain*). Si tratta dunque di un sistema di verifica aperto che non ha bisogno del benestare delle banche per effettuare una transazione. Nello specifico, il protocollo *blockchain* è caratterizzato da una catena distribuita di informazioni tra più nodi di una rete (pubblica o privata) costituita da macchine di varia potenza e connesse tra di loro. Ogni nodo detiene una copia del registro, costituito da una serie di blocchi concatenati, all'interno dei quali sono contenute le informazioni degli scambi avvenuti tra coloro che partecipano alla rete. Il fatto che tali informazioni siano distribuite contemporaneamente, in maniera sincrona su tutti i nodi della rete, rende impossibile l'eliminazione dei dati. Ogni transazione effettuata in questo ambiente è soggetta ad una 'conferma', o 'consenso', espresso attraverso una validazione effettuata risolvendo un problema matematico. Tale conferma viene, dunque, eseguita da un algoritmo condiviso ed accettato nell'ecosistema della rete stessa. Le singole transazioni possono essere costituite da *smart contract*, oggetti informatici le cui condizioni si auto verificano e che abilitano esecuzioni automatiche tra due o più parti, rendendola applicabile ad ogni segmento di mercato. Le transazioni, inoltre, sono caratterizzate da crittografia *hashing*, concetto alla base della *blockchain*, in grado di trasformare ogni dato in una stringa alfanumerica molto complessa, la quale rende inalterabili e criptate tutte le informazioni inserite dentro *blockchain*. In questo modo viene generato un *fingerprint* (cioè un'impronta) di ogni dato all'interno del registro.

Tutto ciò consente, ad esempio, di verificare in qualsiasi momento e in tempo reale l'ammontare di valuta in capo a specifici soggetti senza possibilità di errore in termini quantitativi e di riconoscimento dell'identità. L'operazione è analoga a quella che viene effettuata dalle banche

quando tengono traccia del denaro che transita attraverso i propri conti correnti. E non a caso la *blockchain* è già stata indicata come la tecnologia alla base del prossimo *Internet of finance*. Diversamente dalle piattaforme in capo alle banche, che sono di proprietà delle banche stesse e utilizzabili a livello centralizzato, la *blockchain* è una tecnologia "open" e "free" utilizzabile online da qualsiasi privato e organizzata come un maxi database distribuito. Chiunque può scaricarne una "copia", ossia accedere al sistema: le identità sono protette da crittografia intelligente. La *blockchain* è in grado di assicurare il corretto scambio di titoli e azioni, può sostituire un atto notarile e può garantire la bontà delle votazioni, ridisegnano il concetto di seggio elettorale, proprio perché ogni transazione viene sorvegliata da una rete di nodi che ne garantiscono la correttezza e ne possono mantenere l'anonimato.

Estrapolata dal suo contesto, questa innovativa tecnologia può essere utilizzata in tutti gli ambiti in cui è necessaria una relazione tra più persone o gruppi. Infatti, la *blockchain* non è solo bitcoin, anzi, è una tecnologia capace di avere un impatto rivoluzionario su diversi settori: sanità, tecnologia, finanza, *retail*, media, energia, fino alla pubblica amministrazione. E dopo le banche lo hanno capito aziende del calibro Microsoft, IBM, Samsung e Philips. Ciò che rende la *blockchain* uno strumento dalle molteplici applicazioni, secondo uno studio di Deloitte[50], sono le sue caratteristiche peculiari:

- *Affidabilità*: la *blockchain* non è una struttura governata dal centro ma allarga il cerchio dei partecipanti diretti. Una libertà d'accesso che si traduce in assenza di una plancia di comando. Il sistema, in questo modo, non è solo meno governabile ma anche meno attaccabile. Se uno dei nodi è danneggiato, gli altri continuano a operare saldando la catena, senza perdere alcuna informazione;

- *Trasparenza*: le transazioni sono visibili a tutti I partecipanti. La libertà d'accesso coincide con la trasparenza;

- *Solidità:* è quasi impossibile modificare gli anelli già saldati della *blockchain*. Significa che le informazioni presenti non possono essere manipolate. Una caratteristica che rende le informazioni contenute nella *blockchain* più attendibili e meno esposte alle frodi. Basti pensare alla forza di questa caratteristica in ambito contrattuale;

- *Irrevocabilità:* è possibile effettuare transazioni irrevocabili, in modo da rendere la loro tracciabilità più accurata;

- *Digitale*: documenti, valute, contratti: tutto si trasforma in codice. Senza perdere concretezza. L'applicazione di questa tecnologia diventa così sterminata, sia per estensione geografica, sia per diversificazione.

Sempre Deloitte ha identificato i diversi settori in cui la *blockchain* può avere effetti dirompenti e generare significativi ritorni economici:

[50] Schatsky D., Muraskin C., Deloitte University Press, *Beyond bitcoin, Blockchain is coming to disrupt your industry*, 2015.

- *Finanza:* un sistema distribuito come la *blockchain* elimina la necessità di un'autorità centrale. Un principio che non vale solo per Bce e Fed ma, potenzialmente, per tutti gli intermediari: banche, mercati finanziari, gestori di pagamento. Le transazioni potrebbero essere immediate, più economiche e indubbiamente più sicure. Secondo uno studio di InnoVentures[51], il risparmio in commissioni bancarie potrebbe toccare i 15-20 miliardi di dollari entro il 2022**,** grazie all'incidenza su pagamenti e al trasferimento di denaro a livello internazionale. Ecco perché le banche hanno interesse ad integrare al proprio interno la tecnologia *blockchain*.

- *Assicurazioni:* le agenzie assicurative, così come gli istituti di credito, sono intermediari e, a prima vista, beneficerebbero di un grande potenziale nell'utilizzare la tecnologia *blockchain* soprattutto in termini di semplificazione dei pagamenti di premi e sinistri. Inoltre, le tecnologie *blockchain* potrebbero sostenere la significativa trasformazione digitale in atto nel settore, considerando che gran parte di questa trasformazione si basa sui dati.

- *Media:* per gli editori si apre la porta dei micro-pagamenti senza commissioni. Potrebbe essere un modo per vendere contenuti singoli (a cifre anche molto piccole) e non un intero abbonamento. Il nesso con il diritto d'autore è chiaro, basti pensare alla solidità e all'immutabilità della catena che a sua volta consentirebbe, ad esempio, di accertare in modo sicuro e attendibile la proprietà intellettuale di musica e immagini. Un sistema univoco per verificare una fonte. Più in generale, la *blockchain* può essere applicata al settore dei diritti d'autore. In tal senso, si riporta a titolo esemplificativo la recente acquisizione di Mediachain da parte di Spotify. La startup *blockchain* è nata con l'obiettivo di semplificare la gestione del pagamento delle royalty con l'utilizzo della tecnologia di base Bitcoin. L'operazione rappresenta la soluzione data da Spotify per fronteggiare i problemi di carattere legale connessi al pagamento delle royalties e alla compensazione degli autori e di tutti coloro che rientrano nella filiera dell'industria musicale anche allo scopo di evitare il rischio di ripercussioni finanziarie future. Il team di Mediachain avrà l'obiettivo di lavorare alla gestione di tutte le tematiche transazionali e di remunerazione dei proprietari del copyright nell'ambito della produzione musicale.

- *IoT:* IBM e Samsung stanno guardando alla *blockchain* come strumento per facilitare il dialogo tra gli oggetti dell'IoT. Il sistema distribuito potrebbe adattarsi bene alla necessità di collegare molti dispositivi in aree diverse del globo. Inoltre, la solidità del sistema consentirebbe un trasferimento dei dati più sicuro. La startup Filament sta lavorando proprio al matrimonio tra *blockchain* e IoT. E ha già incassato un round da 5 milioni di dollari che vede tra gli investitori Verizon Ventures.

[51] Santander InnoVentures, Oliver Wyman, Anthemis Group, *Rebooting financial services*, 2015.

- *Retail:* la *blockchain* può fungere da piattaforma di pagamenti alternativa ai contanti, alle carte di credito e di debito nel mercato retail. Questa è la prospettiva più prossima, ma non la sola ipotizzabile. La startup DocuSign, ad esempio, vuole rendere la *blockchain* una piattaforma che faciliti lo scambio di contratti in sicurezza. Ad esempio per l'acquisto di un'auto, registrando il passaggio di proprietà all'interno di un blocco della catena.

- *Sanità:* salute significa dati sensibili. In tal senso, la presenza di un universo sempre più digitale e connesso, rende la (cyber)sicurezza un elemento sempre più importante e necessario. Proprio in quest'ambito può inserirsi la *blockchain*, acquisendo un ruolo determinante. Per fare un esempio, parlando sempre di startup, Factom sta lavorando alla gestione di dati medici, cartelle cliniche e pagamenti di fatture. Tuttavia, non è solo una questione di sicurezza ma anche di trasparenza e condivisione. Infatti, registrare la propria storia clinica su un sistema comune, renderebbe più facile reperire e condividere le informazioni con medici, ospedali e cliniche diverse.

- *Pubblica Amministrazione:* di recente in Italia è stata approvata la legge su *Freedom of information act*[52] che di fatto rende libero, gratuito e aperto a tutti l'accesso all'informazione pubblica e agli atti della PA, ed è stato lanciato il cd. *SPID*[53] *(Sistema Pubblico per la gestione dell'Identità Digitale di cittadini e imprese)*, l'identità digitale che dovrebbe facilitare l'accesso ai dati della PA. Si tratta di un vero e proprio diritto, cui si aggiungono quelli alla trasparenza e alla riservatezza. La *blockchain* potrebbe diventare uno strumento per la gestione delle risorse pubbliche. Il processo non è semplice perché richiederebbe la costituzione di un archivio digitale, da rendere poi facilmente consultabile. Ma i vantaggi potenziali (tra i quali la velocità e la riduzione di frodi e corruzione) sono tali da incentivarne l'implementazione. Non va peraltro dimenticato che, in un sistema non centralizzato, tutti potrebbero godere di un'identità digitale che superi i confini nazionali (basti pensare a migranti e rifugiati). Dunque, la dis-intermediazione introdotta dalla tecnologia *blockchain* potrà modificare le modalità operative consolidate e la Pubblica Amministrazione potrà rendere più semplice ed efficace l'interazione tra cittadino e istituzioni. Occorre tuttavia sottolineare che, alla data, non esiste alcuna soluzione consolidata, siamo nella fase di studio e della ricerca e non è stata ancora individuata una soluzione che abbia sostituito in toto o in maniera consistente, procedure già in uso. Si tratta di applicazioni che vanno dai trasporti alla vendita degli *asset* di Stato, dall'*e-voting* al catasto digitale ed alla distribuzione dei sussidi per il *Welfare*. Tutti progetti pilota, di dimensioni limitate, esempi di ricerca applicata in fase iniziale eppure significativi per una tecnologia orientata alla semplificazione dei processi che non è ancora accompagnata da normative adeguate.

- *Energia:* la *blockchain* potrebbe fare trovare applicazione anche nel settore delle *utility*, sia in termini di piattaforma di pagamento sia per la registrazione di contratti. Un vantaggio

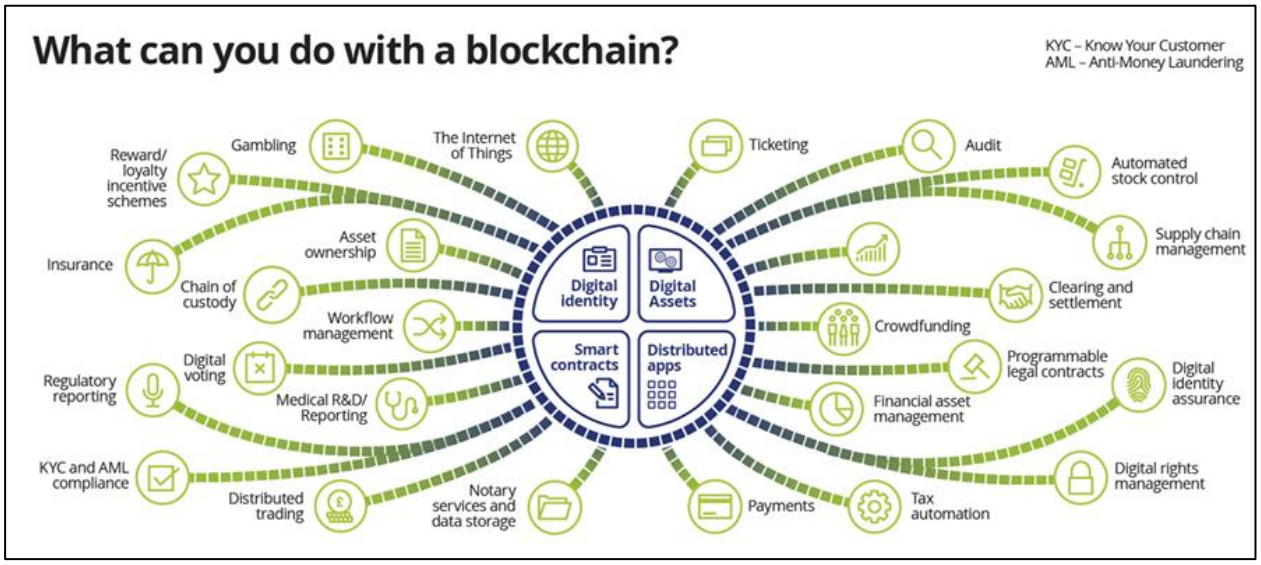
[52] D. lgs 25 maggio 2016, n. 97

[53] DPCM 24 Ottobre 2014

soprattutto per i soggetti (soprattutto nei Paesi in via di sviluppo) che non hanno accesso al sistema bancario e per i piccoli produttori di energia (ad esempio cittadini privati in possesso di piccoli impianti fotovoltaici) Ad ogni modo, anche a livello corporate l'interesse è rilevante come testimoniato dalla prima transazione avvenuta ad inizio ottobre su Enerchain: piattaforma blockchain, implementata da oltre 20 player di rilievo tra cui Enel, E.ON, Total, Vattenfall e Iberdrola, per lo sviluppo di un un mercato europeo decentralizzato per il trading di energia. Enel e E.ON hanno, infatti, finalizzato il primo contratto scambiando elettricità per la prima volta tramite la nuova piattaforma centralizzata effettuando direttamente e in pochi secondi operazioni che di solito richiedono un intermediario centrale. In questo caso, quindi, la possibilità, abilitata dalla blockchain, di effettuare trading diretto senza il coinvolgimento di terze parti, consente di ottenere una riduzione dei costi di acquisto dell'energia elettrica di cui in futuro potranno beneficiare gli utenti finali.

Il grafico seguente esemplifica le diverse applicazioni della *blockchain* potenzialmente attuabili nei diversi settori.

Figura 58. Le potenziali applicazioni della blockchain per settore



Fonte: Deloitte

2.1.5.2 Ulteriori ambiti di applicazione e criticità

Nel settore energetico e dei trasporti, l'innovazione in materia di privacy e riservatezza atta a preservare la gestione e l'analisi dei dati rappresenta un aspetto critico primario per tutti gli *stakeholders* che gestiscono i dati dei clienti, sia che si tratti di *business-to-consumer* che di *business-to-business*. Infatti, in mancanza di una strumentazione che soddisfi tali esigenze di tutela, sarà difficile proporre al cliente un'offerta *data-driven*. La crescente intelligenza incorporata nelle infrastrutture di rete permetterà un'analisi sul campo dei dati al fine di ottenere dati "intelligenti". Questo risulta certamente necessario se si considera che tali analisi richiedono algoritmi molto più elaborati rispetto ad altri settori, come ad esempio quello del *retail*.

L'industria dei media e dell'intrattenimento è stata spesso in prima linea nell'adozione di nuove tecnologie. Le imprese dei media guardano al paradigma *data-driven innovation* come uno strumento per ridurre i costi di funzionamento in un mercato sempre più competitivo, e contemporaneamente, per aumentare le entrate derivanti dalla distribuzione di contenuti.

Non è più sufficiente pubblicare una rivista o mandare in onda nuovi programmi televisivi, gli operatori di oggi devono estrarre valore dai propri asset rispetto a ciascuna fase del ciclo di vita dei dati. Gli operatori media operano in stretta correlazione con i propri utenti e concorrenti, diversamente da quanto accadeva in passato, grazie soprattutto della cd. disintermediazione, i contenuti possono essere generati, condivisi, modificati e ripubblicati da chiunque. Ciò significa che le capacità delle nuove tecnologie basate sui dati, di raccogliere ed elaborare dati diversi provenienti da una moltitudine di fonti, anche in tempo reale, rappresentano un bene preziosissimo per il quale le aziende sono disposte a investire.

In particolare, l'approccio di Netflix, un operatore che fa come noto uso massiccio dei dati, sia per decidere su quali prodotti investire e modellare le proprie offerte, sia per meglio attrarre e soddisfare la domanda da parte dei tanti abbonati ai propri servizi, è del tutto diversa dal tradizionale modello dei demographics.

Netflix privilegia infatti il raggruppamento degli spettatori/abbonati in "*cluster*", definiti quasi esclusivamente sulla base dei gusti comuni (e non delle categorie socio-demografiche) che evidenziano una porzione minima di contenuti che si adatta ai loro profili. Questi potrebbero essere gli stessi per un individuo di Roma così come per uno di Melbourne (anche se, probabilmente, avrebbero accesso a library in parte diverse)[54].

Figura 59. L'intelligenza artificiale nella marketing automation. L'esempio di Netflix

Fonte: Netflix.com

[54] Preta A., Netflix si gioca tutto sui big data, Corcom 13 maggio 2016.

Netflix sembra aver applicato una delle leggi più consolidate in sociologia e psicologia, e cioè che in generale la variazione all'interno di un gruppo sociale è molto più ampia della differenza collettiva tra due gruppi qualsiasi. Dunque se si vuole, ad esempio, fare in modo che aumenti lo streaming dei propri contenuti, sarebbe meglio sfruttare le proprie conoscenze su individui simili in gruppi demografici completamente diversi e solo successivamente provare a disegnare un quadro più ampio e segmentato. Ad esempio, a detta della società, il 90% del volume totale di streaming di un genere tipico di quel paese come le *anime* non proviene dal Giappone. Questo perché quanto piaccia questo genere non è determinato tanto dalla nazionalità, ma dall'attitudine verso il prodotto, in particolare nel caso specifico se l'utente è un *nerd*.

Si tratta di una lezione di notevole importanza di cui tener conto da parte delle imprese e l'emergere di una realtà molto diversa per quanto riguarda la profilazione dei consumatori. Nell'era dei dati, proprio il profilo dei consumatori non può più basarsi su macro-categorie consolidate, ma richiede l'uso di algoritmi basati su criteri più sofisticati, che possono consentire anche delle maggiori semplificazioni.

Gli aspetti giuridici e normativi sul funzionamento di queste tecnologie e sul trattamento dei dati, in particolare dei dati personali degli utenti, non possono essere trascurati, essendo necessario garantire una certa trasparenza. L'Europa dispone di un sistema normativo sulla protezione dei dati più stringente rispetto a quello previsto negli Stati Uniti, il che significa che la privacy come diritto individuale e la tutela della concorrenza a livello globale dovranno essere in qualche modo bilanciate. Gli aspetti legati al rapporto tra privacy e concorrenza saranno oggetto di studio nell'ultimo capitolo del presente lavoro. Anche nel settore delle telecomunicazioni, del *retail* e dell'industria manifatturiera le potenzialità delle tecnologie basate sui dati sono piuttosto evidenti, garantendo la massima operatività ed efficienza.

In definitiva, l'innovazione basata sui dati impatta su tutti i settori dell'economia. Tuttavia, al fine di sfruttare al meglio il potenziale di tale innovazione risulta necessario sviluppare politiche coerenti per l'utilizzo dei dati, attraverso:

- l'istruzione di supporto che si concentra sullo sviluppo delle competenze scientifiche;
- rimuovere gli ostacoli per la creazione di un mercato unico digitale;
- stimolare gli investimenti nelle tecnologie basate sui dati; rendere i dati pubblici accessibili attraverso i cd. *open data* e la rimozione dei silos di dati;
- promuovere lo sviluppo di infrastrutture tecniche competitive;
- sviluppare una legislazione chiara e coerente soprattutto rispetto ai profili di privacy e di sicurezza, della proprietà e del trasferimento dei dati.

Come visto, i cambiamenti tecnologici dell'economia digitale hanno rivoluzionato le possibilità di raccogliere, elaborare e utilizzare i dati in commercio in quasi ogni settore di attività. E' per questo motivo che oggi è rivolta molta più attenzione ai vantaggi economici derivanti dall'utilizzo dei dati per

migliorare la posizione di un'impresa sul mercato. Se da un lato, la raccolta e l'uso di maggiori quantità di dati potrebbe sollevare problematiche di carattere concorrenziale (come sarà ampiamente trattato nel secondo capitolo del presente lavoro), nella maggior parte dei settori considerati tali operazioni rappresentano un meccanismo attraverso cui migliorare la qualità dei prodotti e dei servizi offerti, nonché incrementare l'efficienza economica. I dati possono incidere positivamente sul miglioramento dei prodotti o servizi di un'impresa. Da un lato, questo risultato può essere raggiunto attraverso i cd. "effetti di apprendimento", come nel caso dei motori di ricerca. Si può affermare infatti che più ricerche, unite alla possibilità di osservare su quale risultato della ricerca ciascun utente si sofferma, possano contribuire a migliorare e perfezionare il servizio offerto dallo stesso motore di ricerca e l'attuazione del suo algoritmo di supporto.

Tutto ciò è in grado di incrementare la qualità del risultato della ricerca, che a sua volta può portare ad un aumento degli utenti che utilizzano quel determinato sistema. Similmente, molti software installati su personal computer o smartphone raccolgono informazioni dettagliate sull'uso che si fa di tali prodotti. Esempi importanti sono rappresentati dai browser web e dai sistemi operativi. Inoltre, molti, se non la maggior parte dei siti web raccolgono informazioni dettagliate durante la navigazione dell'utente e utilizzano queste informazioni al fine di individuare quelle parti della pagine web che sono state usate in modo più intensivo ovvero per ridurre al minimo i problemi tecnici. Queste informazioni possono essere utilizzate anche per estendere quelle parti del sito che risultano essere state lette più frequentemente ovvero per accelerare le funzioni più utilizzate.

Al di là degli esempi sopra esposti, è evidente che la disponibilità di un maggior numero di dati è estremamente vantaggiosa per quei servizi in cui il dato è più o meno il prodotto stesso, ovvero è strettamente correlato al prodotto offerto. Quest'ultimo è il caso delle cd. "*matching platforms*" come le piattaforme online per appuntamenti: più soggetti, con il loro profilo personale, rendono il servizio di *dating* più attraente per nuovi individui che sperano di trovare, attraverso quel servizio di incontri, il loro partner ideale. Youtube, come piattaforma video, è probabilmente molto più attraente rispetto a molti dei video che offre. Ad ogni modo, in entrambi i casi, il miglioramento del servizio dipende in parte dalla capacità della piattaforma di consentire una navigazione scorrevole e fluida attraverso questa raccolta di dati; in parte dalla sua abilità nel preservare l'esperienza degli utenti, aumentando l'ampiezza e la profondità della propria offerta.

L'accesso ai dati può anche consentire alle imprese di sfruttare nuove opportunità di business. Riutilizzando dati raccolti nell'ambito di un dato servizio per un diverso scopo, le imprese possono fornire nuovi servizi basati su questi stessi dati. Ad esempio, i dati sulla mobilità generati da operatori di rete mobile e da telefoni cellulari sono utilizzati dai fornitori di servizi di navigazione per mostrare ai propri utenti una migliore visibilità del traffico stradale, suggerendo anche percorsi alternativi meno trafficati.

I dati possono anche essere utilizzati per orientare meglio i potenziali clienti, fornendo loro pubblicità, servizi o prodotti personalizzati. In questo modo, le imprese possono ridurre i loro costi in pubblicità, nonché fissare singoli prezzi proporzionati rispetto alla propensione alla spesa dei consumatori stimata e al grado di sensibilità alle variazioni di prezzo.

La pubblicità online basata sul cd. *behavioral targeting* può servire come esempio di un modello di business sviluppato unicamente grazie ai progressi tecnologici sopra descritti. Per *behavioral targeting* si intende l'offerta di una pubblicità online specifica per un singolo utente, comprensibilmente sviluppata attraverso il profilo dell'utente costruito mediante l'analisi e lo studio delle rispettive abitudini di navigazione. Questo modello di business non esisterebbe senza gli sviluppi citati, sostanzialmente per tre motivi di fondo. Innanzitutto, in mancanza della possibilità di avere piena conoscenza della maggior parte se non di tutti gli aspetti della vita dell'utente attraverso tecnologie di informazione e comunicazione digitale non sarebbe possibile registrare il comportamento di un numero elevato di utenti in modo così dettagliato. Infatti, solo da quando gran parte della popolazione ha iniziato a leggere le notizie online, visualizzare film e programmi on-line, effettuare acquisti online, ascoltare musica in streaming o la radio su internet, commentare politica pubblica, esprimere commenti, pensieri, condividere immagini e video sui social network, ovvero portare sempre con sé un dispositivo che registri costantemente la posizione geografica, è stato possibile registrare puntualmente le azioni di ciascun utente, riuscendo tra l'altro a definire ed individuare con precisione la recettività a specifici messaggi di vendita. In secondo luogo, per riprodurre questo stesso effetto su una scala molto più ampia di utenti, è assolutamente necessaria un ingente capacità e potenza di calcolo e di archiviazione dei dati. In terzo luogo, sono parimenti necessari canali di comunicazione digitali estesi e veloci in modo tale da garantire una distribuzione capillare di tutti questi messaggi di vendita.

Non è ancora chiaro quali saranno, complessivamente, gli effetti economici generati dal fenomeno *data-driven*, tuttavia dagli studi ad oggi disponibili emerge chiaramente l'esplicarsi di un effetto positivo sulla produttività delle imprese:

(i) negli Stati Uniti, secondo uno studio svolto su 330 imprese, le società che adottano un processo di *decision making "data-driven"* registrano una produzione ed una produttività del 5%-6% superiore rispetto a quello che avrebbero se si limitassero agli altri investimenti in ICT;

(ii) nel Regno Unito le imprese che utilizzano i dati sui consumatori risultano essere dall'8% al 13% più produttive;

(iii) un'analisi svolta su un campione di circa 175 milioni di profili LinkedIn, risulta un aumento del 3% nella crescita della produttività, associato appunto ad investimenti in tecnologie Big Data[55].

Complessivamente, questi risultati prevedono un aumento della produttività imprenditoriale pari a circa 5-10% per quei soggetti economici che decidono di adottare un approccio *data-driven*.

A livello di impresa, alcune realtà imprenditoriali già da tempo sono state in grado di modificare il proprio business registrando importanti risultati, come dimostrano gli esempi illustrati di seguito, ripresi da uno studio di Cisco[56]:

[55] OCSE, *Data-Driven Innovation Big Data for Growth and Well-Being*, 2015.

(i) Harrah's ha aumentato i ricavi dell'8-10% analizzando i dati sulla segmentazione dei clienti,

(ii) Amazon ha dichiarato che il 30% dei propri ricavi era riconducibile al suo motore analitico per i consigli ai consumatori.

(iii) Il reparto di gestione dell'inventario di Wal-Mart ha implementato la tecnologia RFID (Radio Frequency IDentification, identificazione con radiofrequenza) per collegare informazioni in tempo reale tra i fornitori e il data warehouse Retail Link. Così facendo, ha ridotto del 16% la frequenza di esaurimento delle scorte.

(iv) FedEx ha ottenuto visibilità in tempo reale con i dati su spedizioni e consumatori provenienti da più di 46.000 centri di distribuzione e supply chain.

(v) Kaiser Permanente ha ridotto del 40% i costi operativi mediante l'integrazione degli archivi medici elettronici e del sistema di gestione dei pazienti interni ed esterni. L'integrazione dei dati clinici e quelli sui costi ha portato alla scoperta degli effetti collaterali del Vioxx e al conseguente ritiro del farmaco dal mercato.

Molteplici sono le applicazione dei dati che forniscono una serie di benefici all'impresa che le adotta:

(i) *Garanzia di ricavi*. L'integrazione dei dati all'interno di una stessa azienda potrebbe perfezionare l'identificazione e prevenire le frodi prima che si verifichino. A trarne beneficio sarebbero soprattutto i settori più soggetti a questo tipo di problemi, ad esempio quello dell'assistenza sanitaria.

(ii) *Attenuazione del rischio.* Attualmente le reti trasportano petabyte di informazioni cruciali per aziende, enti pubblici e consumatori, acuendo sempre di più la vulnerabilità a intrusioni e attacchi alla sicurezza. L'associazione dei dati tra punti geografici e della rete distanti consentirebbe d'individuare modelli sospetti, segnalando immediatamente la necessità di intervenire.

(iii) *Ciclo di vita del cliente.* Le aziende possono concentrarsi su casi di frustrazione dei clienti e offrire una risposta immediata, migliorando così l'esperienza e riducendo il tasso di abbandono. Qualsiasi settore basato sui servizi che tenga ai rapporti con I clienti ne trarrebbe beneficio.

(iv) *Applicazione nel mercato*. I big data consentono migliori servizi di mercato grazie ai sistemi di analisi, ampliando le opportunità di cross-selling e up-selling. Tra i settori che potrebbero beneficiarne maggiormente spiccano banche e commercio su internet. Innovazione dei prodotti. Il feedback dei clienti è fondamentale per lo sviluppo di nuovi prodotti e oggi molte

[56] Gerhardt B., Griffin K., Klemann R., *Come estrarre valore nell'universo frammentato dell'analisi dei Big Data*, Cisco, 2012.

aziende dichiarano già di disporre di maggiori informazioni su ciò che piace e non piace alla loro clientela. L'integrazione di fonti di dati non aziendali quali feed di social network fornirebbe un quadro più completo delle opinioni dei consumatori su un determinato prodotto, portando potenzialmente alla luce la necessità di un nuovo prodotto prima che quest'ultimo venga pensato o progettato.

(v) *Modelli commerciali*. Nuovi algoritmi per la definizione dei prezzi consentiranno nuove opportunità di monetizzazione e interazioni più efficienti con i consumatori. In particolare, i rivenditori al dettaglio potrebbero trarre beneficio da nuovi modelli di prezzi associati alla località specifica e al comportamento dei consumatori.

(vi) *Gestione operativa*. I big data potrebbero aiutare quasi tutte le aziende a ottimizzare l'efficienza. Un provider di servizi, ad esempio, potrebbe migliorare le operazioni quotidiane della propria rete. Un rivenditore al dettaglio, invece, potrebbe dare vita a interazioni più efficienti e redditizie presso il punto vendita. E praticamente tutte le supply chain diventerebbero più armoniose. In generale, un tessuto di informazioni comune migliorerebbe l'efficienza dei processi e fornirebbe una panoramica completa degli asset.

(vii) *Campagne pubblicitarie avanzate*. Più dati le imprese raccolgono sul comportamento dei clienti, più facile è riconvertire tali informazioni in pubblicità. eMarketer è giunta alla conclusione che, attraverso il targeting avanzato basato sui dati, l'efficacia delle campagne pubblicitarie raddoppia. Sono possibili aumenti di più del 50% del costo per mille impressioni (CPM). Per sfruttare ulteriormente questi vantaggi mediante l'analisi dei big data, sarà necessario implementare meccanismi di fiducia in grado di alleviare i vincoli relativi alla privacy e fare in modo che gli annunci con targeting siano in linea con le informazioni sui consumatori presenti in rete. La personalizzazione radicale può potenzialmente generare enorme valore in molti settori di mercato, tra cui l'istruzione, turismo, media, retail, e la pubblicità, così come più in generale nel mercato del lavoro complessivamente considerato (come mostrato nel grafico sottostante). Si tratta di settori in cui il prodotto o il servizio offerto ha un valore differenziato per ciascun consumatore, e le preferenze per caratteristiche distinte influenzano il consumatore nella sua disponibilità a pagare. La possibilità di personalizzare su scala crea la possibilità di andare incontro ad un enorme spettro di esigenze individuali. Altre tecnologie digitali, come la stampa 3D possono consentire alle imprese di portare avanti questa strategia, garantendo una migliore personalizzazione del prodotto e non richiedendo enormi economie di scala. Grazie alle tecnologie *data based* infatti le aziende sono capaci di identificare le migliori opportunità di vendita e di concentrare i loro "sforzi" sui clienti maggiormente propensi all'acquisto. Secondo un recente studio pubblicato su Forbes[57], l'89% dei *marketer* utilizza l'analisi predittiva per migliorare il ROI di vendita.

[57] Vedi http://www.forbes.com/sites/louiscolumbus/2016/01/24/89-of-b2b-marketers-have-predictive-analytics-on-their-roadmaps-for-2016/#6255c47cd291

BOX.2 Programmatic Advertising

Il **Programmatic Advertising** (PA) è un processo automatizzato di compravendita di spazi pubblicitari basato su piattaforme tecnologiche altamente sofisticate. Gli spazi selezionati dagli editori e destinati alla vendita, vengono collocati su particolari piattaforme chiamate SSP (*Supply Side Platfor)* che si occupano di piazzare le offerte sui differenti siti, sulla base dello spazio libero e dell'analisi dei dati riguardanti: il numero di visite registrate, la permanenza degli utenti su un determinato sito, il numero di pagine per singola visita.

Fonte:http://www.ninjamarketing.it/2017/02/03/programmatic-advertising/

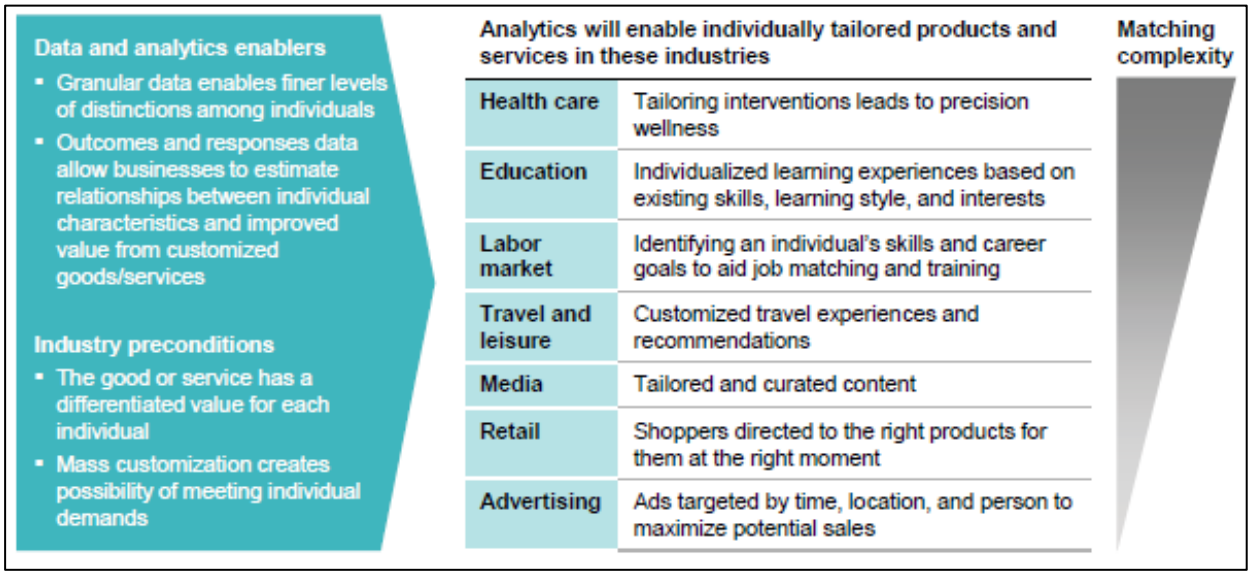
Gli spazi pubblicitari possono essere venduti all'asta sul cd. *open market* man mano che si rendono disponibili o tramite *private deal.* L'acquisto avviene attraverso le piattaforme DSP (*Demand Side Platform*), su cui i *marketer* effettuano le proprie offerte, specificandone le caratteristiche (prezzo, target, posizionamento, ecc.). Il valore aggiunto di questa innovativa tipologia di advertising sta nell'indirizzare la pubblicità su target molto specifici, attraverso l'utilizzo di un mix di dati provenienti da più fonti. In sostanza, il suo più grande potenziale sta nella capacità di mostrare il contenuto che l'utente vuole vedere, nel momento stesso in cui vuole vederlo. A tal fine, sofisticati algoritmi analizzano continuamente i comportamenti online dei consumatori, estraendo il maggior numero di informazioni possibile. In base a dati acquisiti, le imprese in prima persona o avvalendosi di soggetti tersi, provvedono a definire un accurato target della pubblicità, su base sia demografica sia comportamentale.

I dati dunque rappresentano ancora una volta il fulcro dell'innovazione, considerando che la qualità della base informativa sul target raggiungibile costituisce un parametro determinante per la determinazione del prezzo. È proprio in queste decisioni che la tecnologia non sostituisce l'intelligenza strategica del *marketer*, che dovrà sempre: definire precisi obiettivi, individuare il target, progettare forme pubblicitarie coerenti rispetto al messaggio informativo che si intende dare, individuare il posizionamento più efficace e monitorare tempestivamente i risultati, così da intervenire rapidamente con eventuali correttivi.

Fonte:http://www.ninjamarketing.it/2017/02/03/programmatic-advertising/

Si stima che la crescita del PA beneficerà: di un'offerta sempre più ampia in termini di nuovi formati Video e Rich Media, oltre ai tradizionali formati display; degli sviluppi su piattaforme *Mobile* (Smartphone e Tablet); di un maggior grado di differenziazione grazie anche al crescente sviluppo e alla diffusione del *private deal* accanto all'*open market*. Infine, il già citato tema dei dati sarà un altro ingrediente fondamentale per lo sviluppo del PA: le opportunità offerte dal PA aumenteranno, infatti, con l'aumentare della mole e della qualità dei dati sugli utenti, anche attraverso lo sviluppo di player specializzati nella fornitura di dati di qualità (come le DMP – *Data Management Platform,* piattaforme tecnologiche in grado di analizzare i dati relativi al comportamento degli utenti e di trasformali in *real time* in azioni di pianificazione, erogazione e misurazione. Proprio in base a questi dati, i *marketer* potranno poi identificare con maggior accuratezza il target desiderato). In conclusione, il PA non può che essere considerato, allo stato attuale e tenendo conto anche delle sue prospettive di crescita futura, il sistema pubblicitario più economico, più targetizzato e più efficiente.

Figura 60. I vantaggi della "personalizzazione" nei diversi settori



Fonte: McKinsey Global Institute analysis

I benefici non sono però circoscritti unicamente alle imprese: un uso efficace del Big Data Analytics da parte dei governi e delle amministrazioni pubbliche a livello europeo, potrebbe ridurre del 15%-20% i costi amministrativi grazie a una maggiore efficienza derivante dalla maggiore trasparenza, maggiore gettito fiscale (grazie, ad esempio, all'erogazione di servizi personalizzati) e un minor rischio di frodi o errori, legati più direttamente alla capacità di analisi dei dati.

2.2 Dove risiede il valore dei dati

2.2.1 Come massimizzare il valore dei dati

Dato l'ampio ventaglio di possibilità dall'analisi dei big data, al fine di sfruttarne adeguatamente le potenzialità, è particolarmente utile comprendere a quale tipologia di scenario big data una specifica esigenza di business può appartenere. In questa prospettiva si può osservare un processo basato su due step.

Step 1: *Associare l'esigenza di business a un big data type*.

Tale fase consiste nell'effettuare un'analisi qualitativa dell'esigenza o del problema di business che si vuole risolvere, finalizzata all'individuazione di un big data type a questa associabile. Categorizzare i big data problem in big data type rende più semplice l'individuazione delle caratteristiche peculiari di ogni tipo di dati. Tali caratteristiche possono aiutare nella comprensione di come i dati sono acquisiti, come sono elaborati, quanto frequentemente cambiano o sono generati, etc. La tabella seguente presenta alcuni esempi di associazione tra esigenze di business e big data type.

Tabella 4. Abbinamento Business problem e tipologia di dato

Business problem	Big data type	Description
Utilities: Predict power consumption	Machine-generated data	Utility companies have rolled out smart meters to measure the consumption of water, gas, and electricity at regular intervals of one hour or less. These smart meters generate huge volumes of interval data that needs to be analyzed. Utilities also run big, expensive, and complicated systems to generate power. Each grid includes sophisticated sensors that monitor voltage, current, frequency, and? other important operating characteristics.
Telecommunications: Customer churn analytics	Web and social data Transaction data	Telecommunications operators need to build detailed customer churn models that include social media and transaction data, such as CDRs, to keep up with the competition. The value of the churn models depends on the quality of customer attributes (customer master data such as date of birth, gender, location, and income) and the social behaviour of customers. Telecommunications providers who implement a predictive analytics strategy can manage and predict churn by analyzing the calling patterns of subscribers.
Marketing: Sentiment analysis	Web and social data	Marketing departments use Twitter feeds to conduct sentiment analysis to determine what users are saying about the company and its products or services, especially after a new product or release is launched. Customer sentiment must be integrated with customer profile data to derive meaningful results. Customer feedback may vary according to customer demographics.
Customer service: Call monitoring	Human-generated	IT departments are turning to big data solutions to analyze application logs to gain insight that can improve system performance. Log files from various application vendors are in different formats; they must be standardized before IT departments can use them.
Retail: Personalized messaging based on facial recognition and social media	Web and social data Biometrics	Retailers can use facial recognition technology in combination with a photo from social media to make personalized offers to customers based on buying behaviour and location. This capability could have a tremendous impact on retailers? Loyalty programs, but it has serious privacy ramifications. Retailers would need to make the appropriate privacy disclosures before implementing these applications.
Retail and marketing: Mobile data and location-based targeting	Machine-generated data Transaction data	Retailers can target customers with specific promotions and coupons based location data. Solutions are typically designed to detect a user's location upon entry to a store or through GPS. Location data combined with customer preference data from social networks enable retailers to target online and in-store marketing campaigns based on buying history. Notifications are delivered through mobile applications, SMS, and email.

Fonte: International Journal of Innovative Research in Information Security (IJIRIS), Big Data: Review, Classification and Analysis Survey

Step 2: *Analizzare il big data type per classificare le caratteristiche dello scenario big data*

Analizzando il big data type si ricavano le caratteristiche principali che caratterizzano lo scenario big data ricavando importanti indicazioni per lo sviluppo e l'implementazione dell'applicazione. L'immagine seguente mostra alcune caratteristiche da analizzare per ciascun big data type.

Tipologia di analisi: considera se i dati sono analizzati in real time o tramite elaborazioni batch a cadenza programmata (ad esempio ad una data ora o al verificarsi di un dato evento). Tale caratteristica influisce su diversi aspetti relativi a prodotti, tecnologie e hardware da utilizzare nonché sulle sorgenti dati ammissibili o meno. In taluni casi potrebbe essere necessaria una combinazione dei due tipologie suddette: ad esempio nel caso delle applicazioni di rilevazione e prevenzione delle frodi, le analisi devono essere condotte in real time e near real time.

Metodologia di elaborazione: tipologia di tecniche da applicare per l'elaborazione dei dati (modelli predittivi, modelli statistici, query ad-hoc, reportistica, etc.). Dipende fortemente dai requisiti di business che in alcuni casi possono prevedere una combinazione di diverse tecniche. Anche tale aspetto contribuisce a identificare i prodotti e le tecnologie più indicate.

Dimensione e frequenza dei dati: si riferisce alla quantità attesa di dati e alla frequenza con cui questi alimentano l'applicazione. Queste caratteristiche sono essenziali per determinare i meccanismi e i formati di memorizzazione nonché le eventuali operazioni, e relative tecniche e tecnologie, di pre-elaborazione. Dipendono fortemente dal tipo di sorgente dati quali ad esempio on demand, come i dati provenienti dai social media, real-time (dati transazionali, dati metereologici), etc.

Data type: tipologia di dati da elaborare (dati transazionali, dati storici, etc.) Tale aspetto aiuta ad organizzare la memorizzazione dei dati.

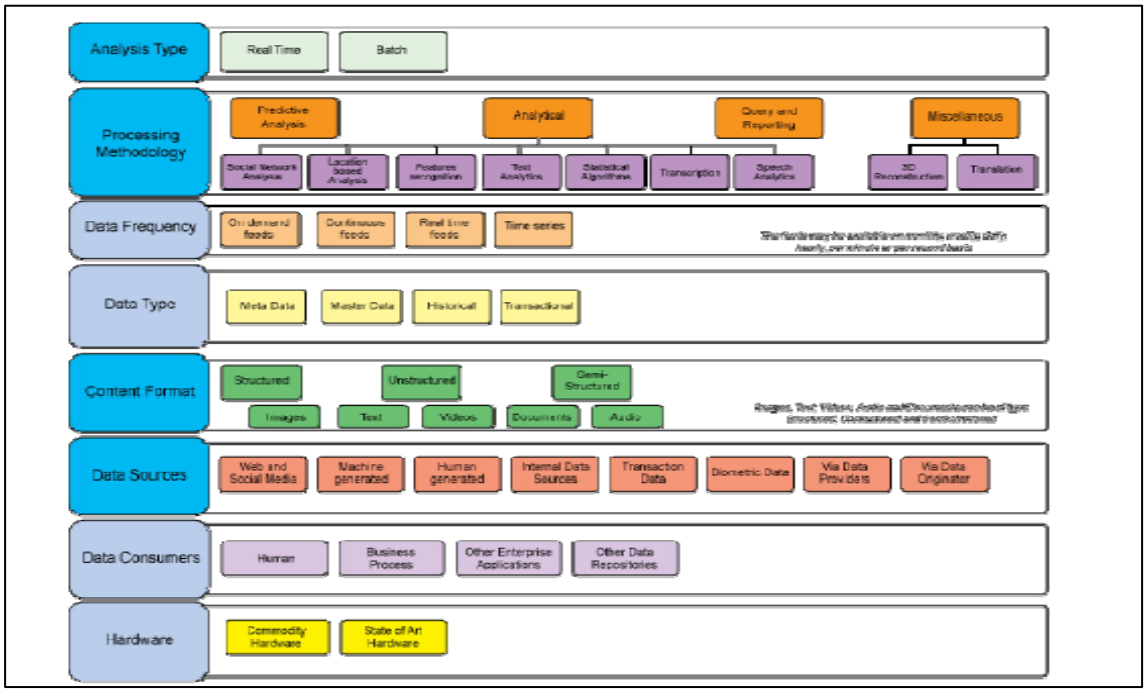
Formato dei dati: strutturati (dati proveniente da un Data Base relazionale ad esempio), non strutturati (audio, video, immagini, etc.) o semi strutturati. Questa caratteristica, molto importante dal punto di vista della risoluzione dell'esigenza di business, determina la modalità di elaborazione da implementare e può vincolare fortemente le tecniche e le tecnologie da utilizzare.

Sorgenti dati: si riferisce a tutte le possibili fonti di dati per l'applicazione o il sistema in esame (web e social media, dati generati da macchine e smart objects, dati generati da persone, etc.) Identificare tutte le sorgenti dati contribuisce a determinare il perimetro d'azione per risolvere l'esigenza di business.

Data consumer: un elenco di tutti i possibili utilizzatori dei dati elaborati (utenti finali, processi di business, altre applicazioni e sistemi, etc.).

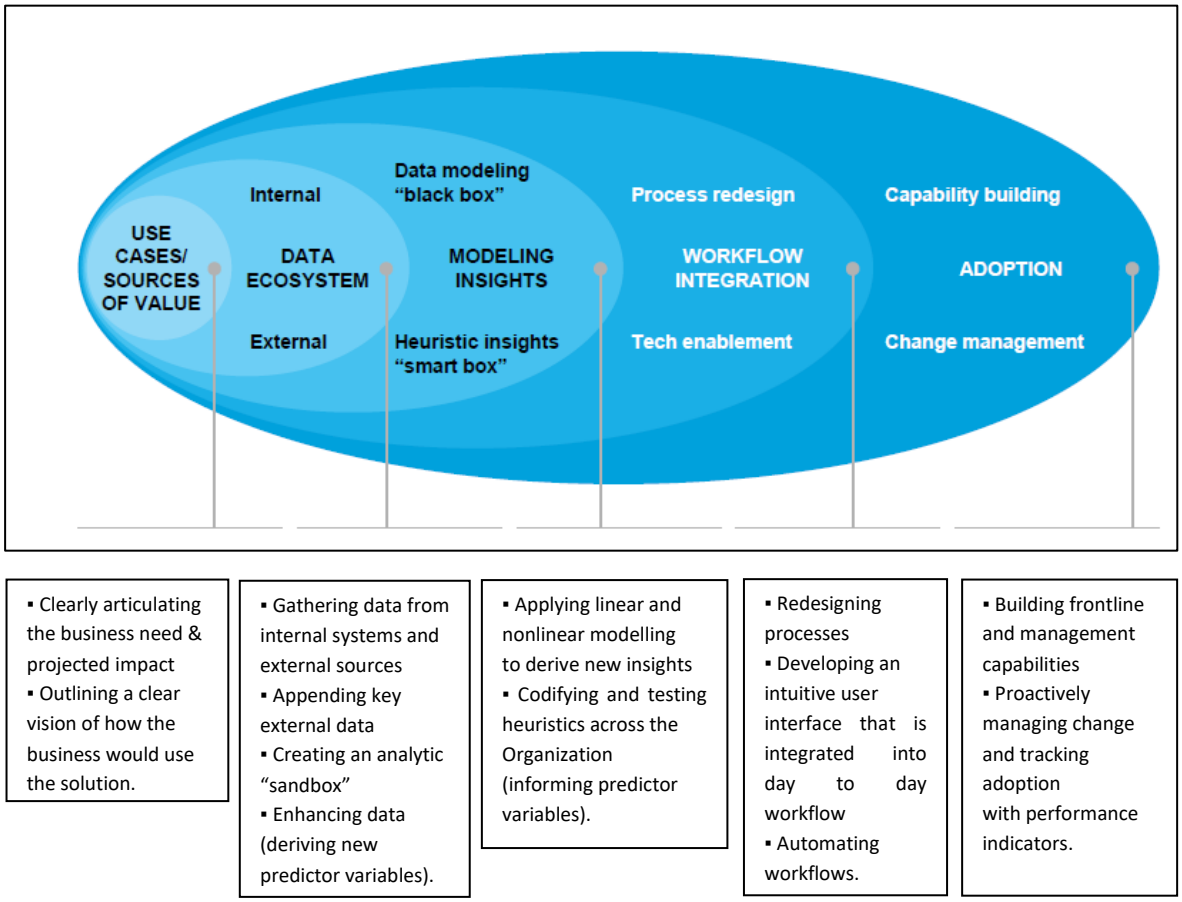
Hardware: tipologia di hardware che verrà impiegata per implementare la soluzione big data. Può essere di tipo commodity a basse prestazioni o ad alte prestazioni.

Figura 61. Classificazione dei Big Data



Fonte: International Journal of Innovative Research in Information Security (IJIRIS), Big Data: Review, Classification and Analysis Survey

Figura 62. I 5 elementi alla base del successo dei dati e della loro trasformazione tramite analitycs



Fonte: McKinsey Global Institute

Dunque se i dati sono diventati una risorsa critica di successo, tuttavia il loro valore è strettamente legato al modo in cui verranno utilizzati e da chi. Un set di dati verosimilmente non ha alcun valore, men che mai può da solo essere impiegato per il lancio di una nuova linea di prodotti; potrebbe però influenzare le entrate di un'impresa, ovvero rappresentare un driver di crescita futura.

La maggiore sfida nell'ambito del processo di analisi dei dati, è rappresentata dalla difficoltà di dover valutare dati spesso profondamente diversi tra loro. Infatti, considerando che ad oggi, i dati certamente non mancano alle aziende – gli analisti stimano una crescita annua del volume di dati del 40%, il 90% dei quali non strutturati – e che le tecnologie tradizionali (alcune risalenti a più di 40 anni fa) non risultano più sufficienti, rispetto alle nuove esigenze, il tradizionale EDW (data *warehouse*) non è più in grado di sostenere la grande varietà e il consistente volume dei dati, senza contare gli alti costi.

Infatti, i dati cd. comportamentali (catturando azioni sia in ambienti digitali che fisici), possono essere valutati assieme a dati transazionali (scaturenti da rapporti d'affari), dati ambientali (condizioni del mondo fisico monitorato e catturato attraverso sensori), dati geo-spaziali, fatti che afferiscono a materiale informativo o di conoscenza (notizie, libri di testo, opere di letteratura, e simili), e registri pubblici. Alcuni dati, come detto, sono strutturati (facilmente rappresentabili in righe e colonne),

altri, come le immagini, audio e i video sono dati destrutturati. I dati possono provenire anche dal web, social media, sensori industriali, sistemi di pagamento, macchine fotografiche, dispositivi elettronici indossabili, ecc. Miliardi di telefoni cellulari, in particolare, sono in grado di memorizzare immagini, video e dati di localizzazione. Dal lato della domanda, i dati possono fornire spunti per diversi usi, alcuni dei quali sono più importanti di altri.

Molti operatori di mercato sono letteralmente "affamati" di dati, scorgendo in essi opportunità di crescita e di implementazione delle proprie prestazioni, in un'ottica di accrescimento della domanda di mercato. La trasformazione del dato grezzo all'utilizzo effettivo e per così dire produttivo dell'informazione consta di una serie di step fondamentali, e la creazione di valore, come detto, può concretizzarsi solo al termine di questo complesso procedimento. Per semplificare, ci si può concentrare su tre categorie di player operanti nel cd. *data ecosystem*, riconoscendo che alcuni di essi potrebbero ricoprire più di un ruolo.

(i) *Produzione e raccolta:* la sorgente e la piattaforma in cui i dati sono inizialmente "catturati";

(ii) Aggregazione: processi e piattaforme che combinano i dati provenienti da più fonti;

(iii) *Analisi:* la scoperta delle intuizioni provenienti dai dati che possono essere attuate.

Figura 63. La creazione di valore all'interno del data ecosystem

	Data generation and collection	Data aggregation	Data analysis
Description	▪ The source and platform where data is initially captured	▪ Combining data from multiple sources ▪ Forming marketplaces for commercial exchange of data	▪ Manipulating aggregated data to develop actionable insights
Factors driving value up	▪ Certain data types will have higher value if collection barriers are extremely high or data cannot be legally shared between parties	▪ Demand growth as more applications are developed ▪ Value will be higher if aggregation is technically challenging or requires a neutral third party	▪ Talent shortage ▪ Deep sector expertise needed to deliver effectively ▪ Close relationship to actual use or implementation clarifies value
Factors driving value down	▪ Growth in available proxies and expansion of open access will increase supply	▪ Technology advances making aggregation easier	▪ Scope could be limited as solutions will be for "vertical" applications
Future trajectory of value	↓	→	↑

Fonte: McKinsey Global Institute analysis

Di solito, le prime opportunità economiche derivano direttamente dalla monetizzazione dei dati. Considerando che i dati sono sempre più semplici da reperire e che i costi di *storage* sono sempre più bassi, la maggior parte di essi sta diventando sempre più una *commodity*. I proxy esistono ora per quei dati che un tempo erano piuttosto scarsi; Google Trends, per esempio, offre un proxy gratuito per i dati cd. *public sentiment* che in precedenza sarebbe stati raccolti attraverso indagini telefoniche.

Tuttavia, vi sono importanti eccezioni alla tendenza alla mercificazione (*commoditization*) dei dati. Quando l'accesso è limitato da barriere fisiche o la raccolta è particolarmente costosa, i dati avranno un proprio valore intrinseco. Google, ad esempio, ha compiuto ingenti investimenti per raccogliere dati per il servizio Google Maps, ha "guidato" per più di sette milioni di km al fine di raccogliere le immagini rilevanti e i dati di posizione[58]. Queste fonti di dati sono pubblici, ma le barriere alla raccolta rimangono estremamente alte. I titolari di questo genere di dati unici, difficili da raccogliere continueranno ad avere la possibilità di monetizzarli direttamente. In questi casi, la natura non-rivale dei dati aiuterà il loro fornitore nell'acquisire maggior valore, dal momento che ci sono più opportunità per monetizzare i dati. Alcune imprese possono mantenere la proprietà di quei dati che hanno la capacità di rimanere "standard industriali". È anche possibile rendere artificialmente scarsi alcune tipologie di dati, ad esempio mediante l'impiego di licenze per evitare che vengano trasferiti ad altri utenti. Un importante caso in cui il valore può accumularsi grazie alla generazione e alla raccolta di dati riguarda il *market making* o le piattaforme di social media con forti effetti di rete. In alcuni settori, un piccolo numero di imprese è in grado di ottenere tale massa critica grazie alla raccolta e alla disponibilità della stragrande maggioranza dei dati generati in questi ecosistemi sul comportamento degli utenti. Il valore nella fase di raccolta dei dati, come in qualsiasi altro mercato, è guidato dalla domanda e dall'offerta. All'aumentare dei dati disponibili, ci si aspetta che la raccolta di specifiche tipologie di dati grezzi generalmente diventerà nel tempo un'operazione di scarso valore, fatta eccezione per quei casi in cui l'offerta o la domanda sono vincolata a particolari tipologie di dati. Tuttavia, in assenza di questi tipi di vincoli di offerta eccezionale, la semplice vendita di dati grezzi può dar luogo a rendimenti decrescenti nel tempo.

Un altro ruolo nell'ecosistema dati attiene all'aggregazione di informazioni derivanti da fonti diverse. In generale, questa capacità sta diventando sempre più accessibile e a costi ridotti, ma questa attività può essere utile quando si applicano determinate condizioni. L'aggregazione dei dati aggiunge valore quando i dati si combinano tra loro generando nuova informazione. Alcune aziende hanno costruito dei modelli di business tali per cui fungono da aggregatori per i concorrenti all'interno di un determinato settore, e questo modello ha il potenziale per generare effetti di rete. Il valore dei servizi di aggregazione sembra dunque destinato ad aumentare solo nei casi in cui l'integrazione di dati provenienti da diverse fonti risulta essere particolarmente complesso, ovvero quando il coordinamento dell'accesso ai dati attraverso diverse fonti rappresenta una barriera. Questo può accadere anche se i dati di riferimento sono "commodizzati" (come per i dati del mercato finanziario) o quando sono di varia e differenziata natura (come per la documentazione sanitaria). Molti dei tradizionali fornitori di dati di marketing (come i fornitori di mailing list) o come i fornitori di servizi di informazione (come ad esempio Westlaw, Bloomberg, e Argo) rientrano in questa categoria e hanno sviluppato rapporti di lunga data con i collettori di dati o hanno immobilizzazioni tecniche impiegate all'aggregazione. Molti di questi aggregatori fungono anche da vere e proprie "guide" nell'universo dati, attraverso la loro profonda comprensione degli ambienti di dati complessi e della normativa sulla privacy sono in grado di consigliare i loro clienti su come gestire al meglio i dati stessi. Le imprese che offrono servizi di aggregazione di dati provenienti da fonti diverse sono in grado di

[58] Miller G., *The huge, unseen operation behind the accuracy of Google Maps*, 2014.

generare valore per una serie di motivi. In primo luogo, queste imprese fungono da "*one-stop shop*" per i dati provenienti da più fonti. In secondo luogo, dati aggregati incrementano il valore in quanto idonei a generare intuizioni migliori. Il *benchmark* della performance di più entità può aiutare a identificare le aree di miglioramento, per esempio. Le valutazioni dei rischi sono più precisi quando incorporano più dati comportamentali e ambientali. La combinazione dei dati sul comportamento di navigazione dell'utente dal dispositivo mobile e dal pc offre un quadro più completo di modelli di consumo di un cliente. Tuttavia, se è vero che l'aggregazione è in grado di produrre un valore significativo, è altrettanto vero che sta diventando particolarmente semplice per gli utenti eseguire molti aspetti di questa funzione autonomamente. Si registra infatti una crescita robusta in nuovi servizi software per l'organizzazione di dati provenienti da diverse fonti interne ed esterne. Di conseguenza, gli utenti finali hanno ora a disposizione gli strumenti meno costosi e più potenti per aggregare i dati.

La terza parte dell'ecosistema dei dati, l'analisi, è quella dove ci si aspetta di vedere le più grandi opportunità per il futuro. La crescente complessità dei dati e della loro analisi necessita sempre più di strumenti e competenze di alto livello per coglierne appieno le opportunità e acquisire valore. Analisi erronee possono distruggere il valore potenziale dei dati, viceversa una buona analisi può generare importanti intuizioni anche da dati apparentemente di scarso valore. Date le dimensioni delle potenziali opportunità, alcune imprese stanno valutando la possibilità di immettersi nel mercato dell'analisi. Aggregatori di dati si propongono sul mercato per offrire servizi di integrazione e di analisi dei dati dei clienti.

Dal lato della domanda, poiché l'analisi è spesso l'ultimo step del processo, il valore generato dai dati e dalla loro analisi è facilmente calcolabile. Questo pone l'impresa che offre servizi di analisi in una posizione favorita per acquisire una parte di questo valore. Mentre in passato le aziende erano spesso incerte su cosa fare di enormi volumi di dati grezzi raccolti, di recente invece, sono maggiormente disposte a investire in analisi approfondite al fine di utilizzare le informazioni acquisite per ottimizzare le vendite, le strategie e altre funzioni aziendali. In tutti i settori, le imprese guardano ai dati come importanti strumenti che, associati a giuste capacità analitiche, offrono nuove opportunità di business. Sul lato dell'offerta, le competenze altamente specializzate necessarie per l'analisi e l'interpretazione dei dati sono attualmente piuttosto scarse.

Anche se si registrano importanti progressi in termini di strumenti e piattaforme disponibili, la necessità di combinare questi strumenti analitici con specifici *skills* continuerà a presentare un collo di bottiglia, accrescendo il valore dell'analisi. Alcune imprese offrono sul mercato software o prodotti e servizi personalizzati di analisi dei dati, registrando i più alti margini netti all'interno dell'intero ecosistema di dati, spesso tra il 20% e il 30%[59]. Tuttavia, il solo acquisto di simili strumenti di analisi non è sufficiente per quelle aziende che non hanno al proprio interno le giuste competenze tecniche, e per tale motivo molte di esse si rivolgono a fornitori di servizi esterni. Va peraltro sottolineato il fatto che le competenze di analisi applicate ad un determinato settore possono non essere necessariamente le stesse richieste per un altro settore. Infatti, mentre l'attività di aggregazione dei

[59] Wikibon, Big Data Vendor Revenue and Market Forecast 2011-2020, 2014.

dati è di tipo orizzontale, passando attraverso diverse tipologie e fonti di dati; viceversa, la loro analisi è di tipo verticale, pertanto ogni ulteriore caso d'uso verticale richiederà una ulteriore competenza e, talvolta, anche nuove tecniche analitiche. Tenendo conto delle opportunità offerte dall'analisi dei dati, talune imprese, in altre parti dell'ecosistema, aggiungono l'analisi proprio per aggiudicarsi una piccola parte di questo segmento ad alto valore di mercato. Nel settore assicurativo, ad esempio, CoreLogic ha utilizzato il proprio patrimonio di dati per sviluppare un sistema di gestione dei rischi catastrofici, analisi questa venduta come prodotto agli assicuratori. Altri esempi possono essere anche imprese come Wood Mackenzie nel settore energetico e Inovalon nel settore sanitario.

Oltre ai prodotti-dati di cui sopra, imprese di raccolta dati e aggregatori di dati stanno offrendo servizi atti ad integrare i propri dati con quelli dei clienti per effettuare particolari analisi *ad hoc*, soprattutto in settori quali quello dell'assistenza sanitaria, che necessitano di analisi su larga scala, ma in cui la maggior parte delle aziende manca delle principali competenze. Start-up innovative, come SparkBeyond stanno integrando l'attività di aggregazione con la capacità di analisi al fine di applicare il ML a specifiche esigenze aziendali. In altri casi, i produttori di software di analisi si espanderanno nel fornire servizi di analisi nelle rispettive aree di competenza. Alcune agenzie di marketing e consulenza, nonché le imprese di business intelligence stanno aggiungendo attività di analisi alle loro offerte per aumentare le proprie capacità. Le imprese che offrono soluzioni tecnologiche integrate stanno infine enfatizzando l'attività di analisi, un po' come ha fatto la IBM con Watson.

Nei settori in cui la maggior parte degli operatori storici sono abituati a fare affidamento su un certo tipo di dati standardizzati per prendere decisioni, introdurre nuovi insiemi di dati e integrandoli con quelli già in uso può letteralmente cambiare lo scenario competitivo. Nuovi entranti con accesso privilegiato a questi insiemi di dati possono rappresentare una concreta minaccia per gli operatori storici. Lo vediamo nel settore assicurativo dove le nuove aziende sono entrate nel mercato con dati telematici che permettono di comprendere meglio il comportamento di guida dell'utente e che combinati con i dati demografici acquisiti all'atto di stipula della polizza conferiscono quel valore aggiunto che consente al *new entrant* di affermarsi sul mercato di riferimento. Attraverso l'IoT, lampade collegate, che rilevano la presenza di persone in una stanza, ovvero che riducono il consumo di energia, generano ulteriori dati utilizzati dai manager per ottimizzare la pianificazione dello spazio fisico. Anche nel campo delle risorse umane, alcune organizzazioni hanno fatto indossare ai propri dipendenti dispositivi elettronici in grado di acquisire dati e riprodurre un quadro delle "vere" reti sociali esistenti nel mondo del lavoro, consentendo a tali organizzazioni di ottimizzare le collaborazioni interne attraverso una serie di cambiamenti negli spazi di lavoro.

In definitiva, il valore dei dati si riduce alla loro unicità e alle modalità di utilizzo. Comprendere il valore di tutti questi piccoli bit di informazione che devono essere raccolti, setacciati, e analizzati è piuttosto complesso, soprattutto perché le imprese non possono fissare il loro valore fino a quando non sono in grado di specificare chiaramente il loro utilizzo, immediato o potenziale. I dati potrebbero potenzialmente rappresentare solo una piccola percentuale delle entrate di un'impresa, ma potrebbero, al contrario, essere anche un fattore chiave della sua futura crescita.

Molte imprese colgono questo potenziale e sono sempre più orientate ad utilizzare i dati per accrescere e migliorare le proprie performance. I dati possono essere impiegati in vario modo come ad esempio:

(i) *Ottimizzazione dei costi e dei ricavi*: le potenziali applicazioni in quest'ambito sono numerose. Sul lato dei costi, i dati possono essere impiegati per analisi predittive ai fini di opere di manutenzione, efficientamento gestionale, nell'attività di approvvigionamento, *supply chain* e pianificazione logistica. Sul lato dei ricavi, le intuizioni derivanti dai dati possono essere utilizzate per entrare in nuovi mercati, per attività di segmentazione dei consumatori, per migliorare le caratteristiche dei prodotti offerti e per rendere più efficaci i canali di distribuzione. I dati derivati dalle macchine e da processi, in particolare dai sensori IoT e dal comportamento dei clienti sono i più utili per le operazioni di ottimizzazione. Viceversa, i dati generati internamente, possono essere utili attività di efficientamento.

(ii) *Marketing e pubblicità*: queste attività in genere si basano su transazioni effettuate dall'utente e sui dati comportamentali aggregati derivanti da più fonti, come profili social, informazioni demografiche, cronologia di navigazione on-line e acquisti precedenti. In quest'ottica possono essere analizzati grandi volumi di dati, fondamentale è l'attività di pulizia e scrematura dei dati ai fini della creazione dell'intuizione di valore. In particolare, imprese pubblicitarie altamente tecnologiche e social media con grandi quantità di dati comportamentali di consumo sono solo alcuni dei player cd. *fastestgrowing* in questo settore.

(iii) *Market intelligence*: molti fornitori tradizionali di servizi di informazione rientrano in questa categoria. Poche aziende sono in grado di generare macro dati autonomamente, la maggior parte dei fornitori in questo settore aggregano dati da fonti esterne. Poiché i dati conferiscono un chiaro valore per i clienti e non è facile replicarlo, può essere venduto direttamente e, in questo caso, il suo valore è più facilmente determinabile.

(iv) *Market-making*: le imprese *market-making*, dalle app *ride-sharing* ai siti di incontri, giocano un ruolo cruciale nel far incontrare le esigenze di acquirenti e venditori. Queste imprese spesso creano piattaforme per raccogliere le informazioni necessarie per consentire una corrispondenza quanto più efficiente ed efficace. In alcuni casi, i dati di segnalazione pura rappresentano tutto ciò che conta. Ma in altri casi, le preferenze, i dati reputazionali (a garanzia dell'autenticità e della qualità dei partecipanti), i dati sulle transazioni e quelli comportamentali sono determinanti. In questo contesto, economie di scala ed effetti di rete svolgono un ruolo fondamentale.

(v) *Training data for artificial intelligence*: *machine learning* e *deep learning* necessitano di enormi quantità di *traning data*. Alcuni vengono generati attraverso simulazioni ripetute, in parte sono generati nella sfera pubblica (come ad esempio i dati catastali e quelli climatici), in parte sono aggregati da una varietà di fonti (quali immagini e video o dati comportamentali). Le imprese che con la loro piattaforma producono enormi quantità di dati

rilevanti possono godere di un vantaggio importante, in quanto le loro offerte avranno più tempo per imparare e generare ulteriori dati, alimentando il circolo virtuoso. Però, proprio perché c'è una grande varietà di possibilità di utilizzo per i diversi utenti, valorizzare i dati in questo contesto può risultare particolarmente impegnativo.

Questi ecosistemi possono poi sovrapporsi. In alcuni casi, lo stesso set di dati può avere molteplici applicazioni, ciascuna con un diverso valore.

2.2.1.1 La raccolta e la condivisione dei dati all'interno dell'impresa: i data lake

Il primo passo nella creazione di valore dai dati e dallo loro analisi è quello di garantirne innanzitutto l'accesso. Questo può essere semplice, in teoria, ma ha dimostrato di essere molto più difficile nella pratica. Non importa quale sia il settore in cui si opera, la maggior parte delle grandi imprese hanno all'interno della propria struttura aziendale, un reparto dedicato alla creazione di silos di dati. Di conseguenza, la condivisione delle informazioni tra più reparti risulta essere particolarmente difficile. Diversi dipartimenti possono avere dati univoci, possono esserci sistemi non integrati o addirittura incompatibili tra loro. A causa di questo genere di barriere per così dire "organizzative", la gestione dei dati può diventare piuttosto costosa, e poche aziende hanno sistemi di dati tali da favorire analisi su larga scala.

Per questo motivo, la maggior parte delle imprese potrebbero trarre vantaggioso acquisire nuove tipologie di dati da fonti esterne per poi combinarle con i propri dati interni. Tecnologie come i *data lake* sono una soluzione ottimale per superare questo genere di problemi. Questi nuovi strumenti sono in grado di semplificare l'accesso ai dati per tutti i reparti aziendali attraverso l'integrazione di tutti i dati disponibili in un unico *repository* facilmente accessibile e flessibile. Inoltre, hanno anche la capacità di memorizzare la stragrande maggioranza delle informazioni in tempo reale, dei flussi continui di dati provenienti dai sensori IoT, così come sono capaci di integrare facilmente le informazioni provenienti da nuove fonti, anche se i dati sono in formati diversi. I *dake lake* possono incorporare dati strutturati provenienti da database relazionali, dati non strutturati provenienti da fonti come ad esempio i file multimediali ed e-mail, o anche dati grezzi. L'intera organizzazione può accedere alle informazioni contenute nel *dake lake* tramite una semplice ricerca. Proprio perché i *dake lake* memorizzano in tempo reale grandi quantità di dati, essi sono in grado di fornire uno spazio ideale per i *data scientist* per eseguire il rilevamento, la costruzione di applicazioni *data-driven* e portare avanti grandi progetti di *analytics*.

Aziende come Amazon, Facebook, e Goldman Sachs hanno adottato approcci come questi per raggiungere significativi vantaggi, nonché per effettuare miglioramenti e creare nuovi prodotti. Capacità di integrazione di mole di dati sono fondamentali per le imprese con problemi di accessibilità o quelle per le quali la scoperta di dati ha particolare rilevanza per la creazione di valore.

Semplificando, i *data lake* vengono definiti come piattaforme:

- per la gestione flessibile dei dati;
- per aggregare i dati cross-silos in un unico insieme;

- per permettere l'esplorazione di tutti i dati.

In questo senso, la piattaforma Data Lake più utilizzata allo stato è Hadoop che:

- permette la scalabilità orizzontale su hardware commodity - permette uno schema di dati variegati ottimizzato in lettura;
- include strati di lavorazione dei dati in linguaggi comuni ;
- grandi referenze (Google e Yahoo).

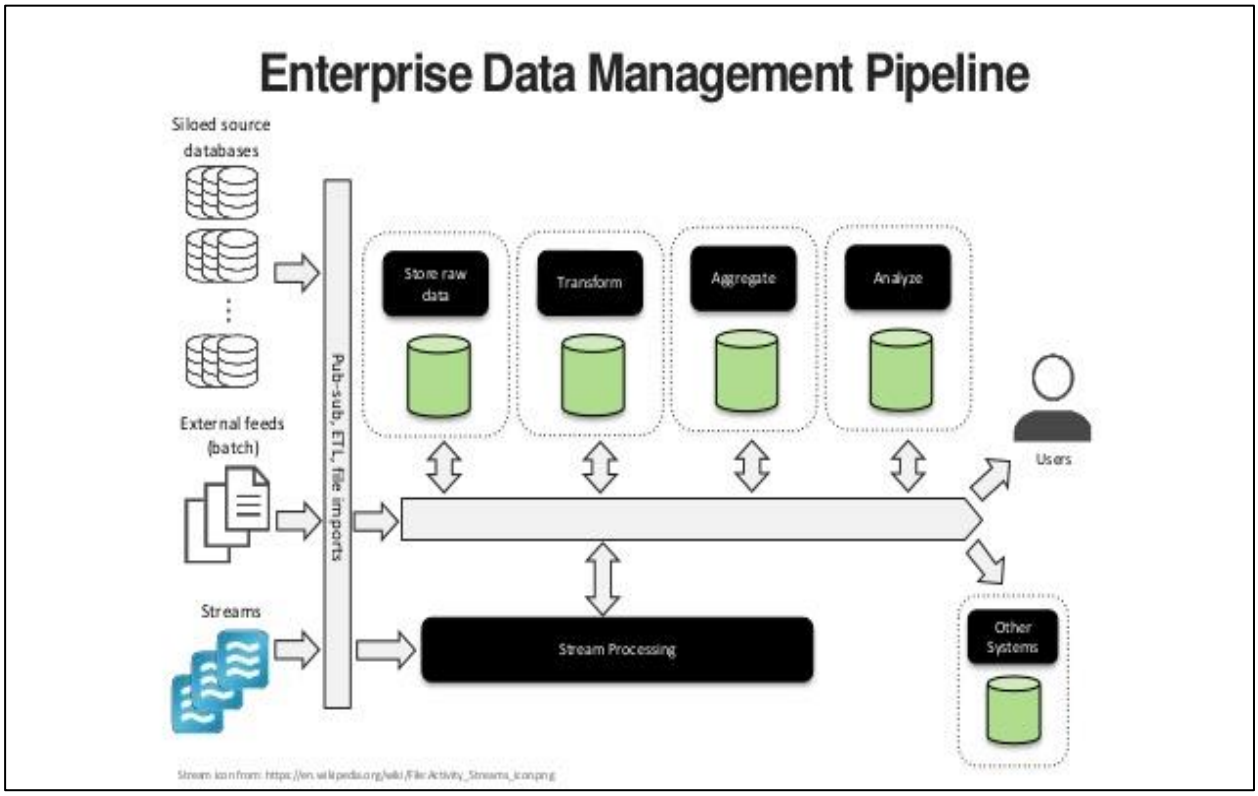
Hadoop ha riscosso grande successo fondamentalmente perché è in grado di gestire dati multi-strutturati al fine della segmentazione dei clienti per campagne di marketing e raccomandazioni; dell'analisi predittiva e della definizione di modelli di rischio. I *data lake* sono disegnati per fornire l'output di Hadoop alle applicazioni online.

Queste applicazioni hanno dei requisiti tra cui:

- latenza di risposta in millisecondi;
- accesso random su un sottoinsieme di dati indicizzato;
- supporto di query espressive ed aggregazione di dati;
- update di dati che cambiano valore frequentemente e in tempo reale.

La creazione dell'output passa attraverso un articolato processo, l'*Enterprise Data Management*, attraverso cui l'impresa definisce con precisione, integra facilmente all'interno della propria organizzazione aziendale e recupera efficacemente i dati al fine delle applicazioni interne o delle comunicazioni esterne.

Figura 64. Il processo di Enterprise Data Management



Fonte: MongoDB

Come si evince dalla figura 56, il processo consta di una serie di fasi:

- *Scrittura massiva:* memorizzazione di dati grezzi-struttura in grado di scalare orizzontalmente ad alta velocità;
- *Lettura massiva:* trasformazione al fine di normalizzare i dati;
- *Aggregazione:* merger di dati attraverso indici di correlazione anche attraverso dati già presenti nel database;
- *Analisi:* attraverso gli analytics, di tutti i dati o di porzioni;
- *Destinazione finale:* utenti (tramite dashboard offline o in tempo reale) o altri sistemi (applicazioni digitali).

Al termine del processo, eventuali scenari applicativi possono trovarsi nella definizione della cd. *Single view of customer*, ovvero l'aggregazione di tutte le informazioni, di cui si dispone, relative ad un cliente al fine di migliorarne il rapporto fiduciario e di acquisire ulteriori ricavi; nelle attività di *Clustering,* ovvero nelle valutazioni e verifiche sui mancati rinnovi di un certo servizio. Tuttavia, è necessario considerare che all'interno del processo alcune operazioni superflue potrebbero inficiarne la buona riuscita.

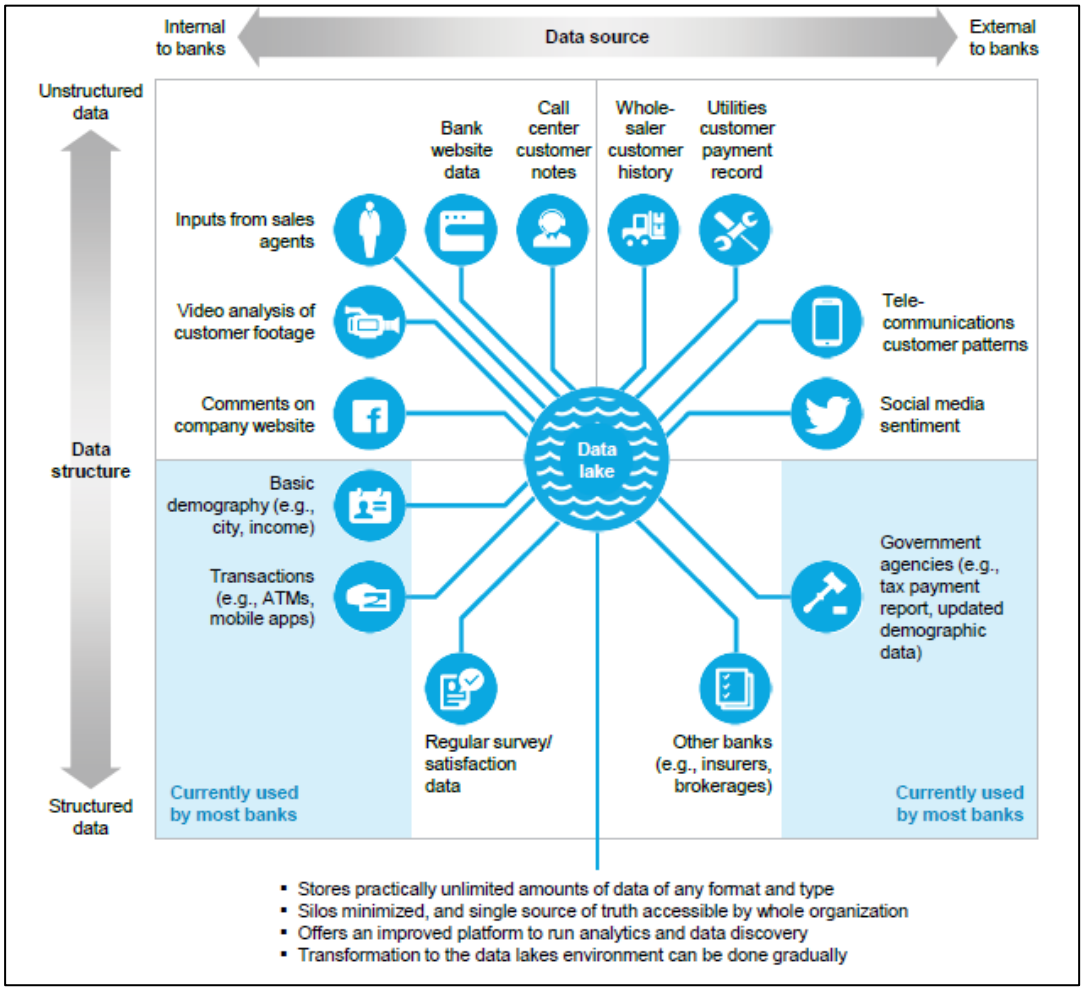
In particolare:

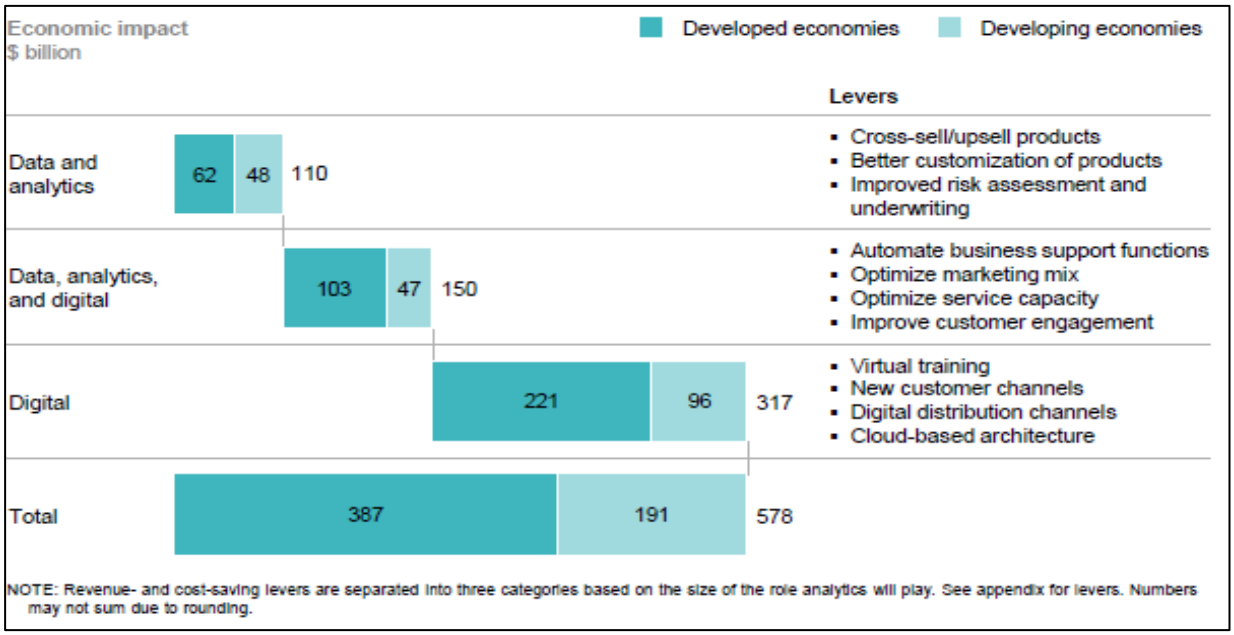
- *join* non necessarie possono causare pessime performance;
- scalare verticalmente risulta particolarmente costoso;
- la predisposizione di uno schema rigido rende difficile il consolidamento di insiemi di dati variabili e non strutturati;
- è fondamentale eliminare durante la fase di aggregazione eventuali differenze nei record;
- i processi soventi possono anche durare per ore ma devono essere sempre sottoposti ad un controllo costante;
- va poi considerato che i dati sono vecchi per prendere decisioni intraday.

Nel settore del *banking retail* i dati sono all'ordine del giorno, informazioni su clienti, transazioni, situazioni finanziarie, dati demografici sono sempre state disponibili e in grandi quantità. Ma pochi istituti hanno pienamente colto il vantaggio che è possibile trarre da questi dati a causa della presenza di barriere interne che limitano l'accesso in tutta l'organizzazione, della qualità variabile dei dati, e talvolta anche dell'incapacità di comprenderne il valore. Ad oggi, il superamento di queste barriere sta diventando un fattore critico, ora che sempre più crescenti nuove fonti di dati possono essere aggiunti ai record delle transazioni esistenti. Queste nuovi tipologie di informazioni includono la messaggistica su piattaforme social, conversazioni con call center, video, ed i dati acquisiti da fonti esterne o da partner.

Inoltre, le banche possono collaborare con operatori di telefonia mobile e rivenditori al fine di integrare la loro visione di ogni singolo cliente. L'analisi dei dati può migliorare l'esperienza del cliente, permettendo alle banche di consolidare i rapporti con i clienti esistenti e attirarne di nuovi. I clienti si aspettano sempre più un'esperienza personalizzata in tutti i canali. Vogliono anche che i rispettivi servizi bancari siano disponibili sulle loro piattaforme preferite; un numero crescente di utenti effettuano i loro pagamenti direttamente tramite applicazioni di messaggistica, per esempio. Strumenti di integrazione come i *data lake* contribuiscono a fornire un approccio olistico definendo il profilo di ciascun cliente, combinando dati socio-demografici con dati comportamentali diversi. Per gli istituti bancari, una delle applicazioni più importanti è la capacità di migliorare la valutazione del rischio e di credito attingendo a dati provenienti da diverse fonti per effettuare previsioni sempre più accurate. Le figure seguenti mostrano come nel settore del *banking retail,* l'utilizzo dei *data lake* consente di distruggere il modello di organizzazione dei dati in silos e di creare un unico *repository* nel quale confluiscono dati nuovi e dati tradizionali; e come l'implementazione di questi strumenti *data driven* può determinare un impatto economico di notevole valore.

Figura 65. I vantaggi dei data lake





Fonte: McKinsey Global Institute analysis

In linea di principio, non ci sono settori in cui la capacità di integrare continuamente nuove fonti di dati di qualsiasi formato e qualità non genererebbe dei miglioramenti. Oltre al *banking retail*, molti altri settori hanno potenzialmente le caratteristiche per fare dell'integrazione dei dati di massa un vantaggio economico. Nel settore pubblico, ad esempio, i *data lake* potrebbero avere un effetto dirompente distruggendo i silos esistenti. Si tratta comunque di un'operazione piuttosto complessa, considerando che spesso i dati appartengono a diversi dipartimenti e funzioni, possono trovarsi su una serie di piattaforme diverse e con differenti tassonomie e requisiti di accesso. La frammentazione e l'assenza di una proprietà centralizzata per le infrastrutture IT a livello nazionale potrebbero quindi rendere il tutto molto più complesso. Ma l'integrazione massiccia dei dati potrebbe creare una nuova esperienza senza soluzione di continuità per l'utente finale, se l'utente è un dipendente pubblico, un business, un cittadino, o un altro ufficio intergovernativo. Le città, per esempio, potrebbero collegare i dati relativi alla salute e all'istruzione nella pianificazione dei loro programmi di assistenza sociale. L'assicurativo è un altro settore l'utilizzo di questi strumenti può fare la differenza, contando su un'ampia gamma di informazioni per la valutazione del rischio in modo più accurato. Le aziende possono ridurre le frodi, migliorare i prezzi e il *cross-selling*, senza contare che l'applicazione di tecniche di analisi come il *data mining* su ampi volumi di dati ordinatamente raccolti in un *data lake* può facilitare anche nella creazione di nuovi prodotti, ma di questo ne parleremo più approfonditamente nel paragrafo successivo.

2.2.2. L'analisi dei dati tramite algoritmi

L'analisi dei dati è un processo di ispezione, pulizia, trasformazione e modellazione di dati con l'obiettivo di evidenziare informazioni rilevanti che suggeriscano conclusioni e supportino le decisioni strategiche aziendali. L'analisi dei dati ha molti approcci e sfaccettature, il che comprende l'applicazione di tecniche diversissime tra loro: il *data mining* si focalizza nella modellazione e scoperta di conoscenza per scopi predittivi piuttosto che descrittivi; la *business intelligence* identifica l'analisi di dati che si basata sull'aggregazione, focalizzandosi sulle informazioni aziendali; nelle applicazioni statistiche, l'analisi dei dati è ripartita in statistica descrittiva, analisi dei dati esplorativa (ADE) e analisi dei dati confermativa (ADC). L'ADE si concentra sullo scoprire nuove caratteristiche presenti nei dati, mentre l'ADC nel confermare o falsificare le ipotesi esistenti.

Gli algoritmi standard di analisi di dati si basano su principi deduttivi e induttivi: all'insieme di dati, si associa generalmente un modello distributivo sul quale vengono realizzate le deduzioni matematiche. Per riadattare i risultati del modello all'insieme di informazioni reali, è però necessario ideare una serie di coefficienti di affidabilità, che rimangono di determinazione aleatoria. Benché queste analisi possano essere molto raffinate, l'interpretazione dei loro risultati può variarne completamente l'affidabilità, o addirittura il senso.

Attualmente, non esistono precedenti strutturali nell'utilizzo di grossi volumi di informazioni per la deduzione di trend e correlazioni nascoste; tuttavia, i pochi esempi a disposizione ne dimostrano l'enorme potenziale applicativo nei settori dell'alta finanza e del trading, nel campo assicurativo e nel marketing avanzato.

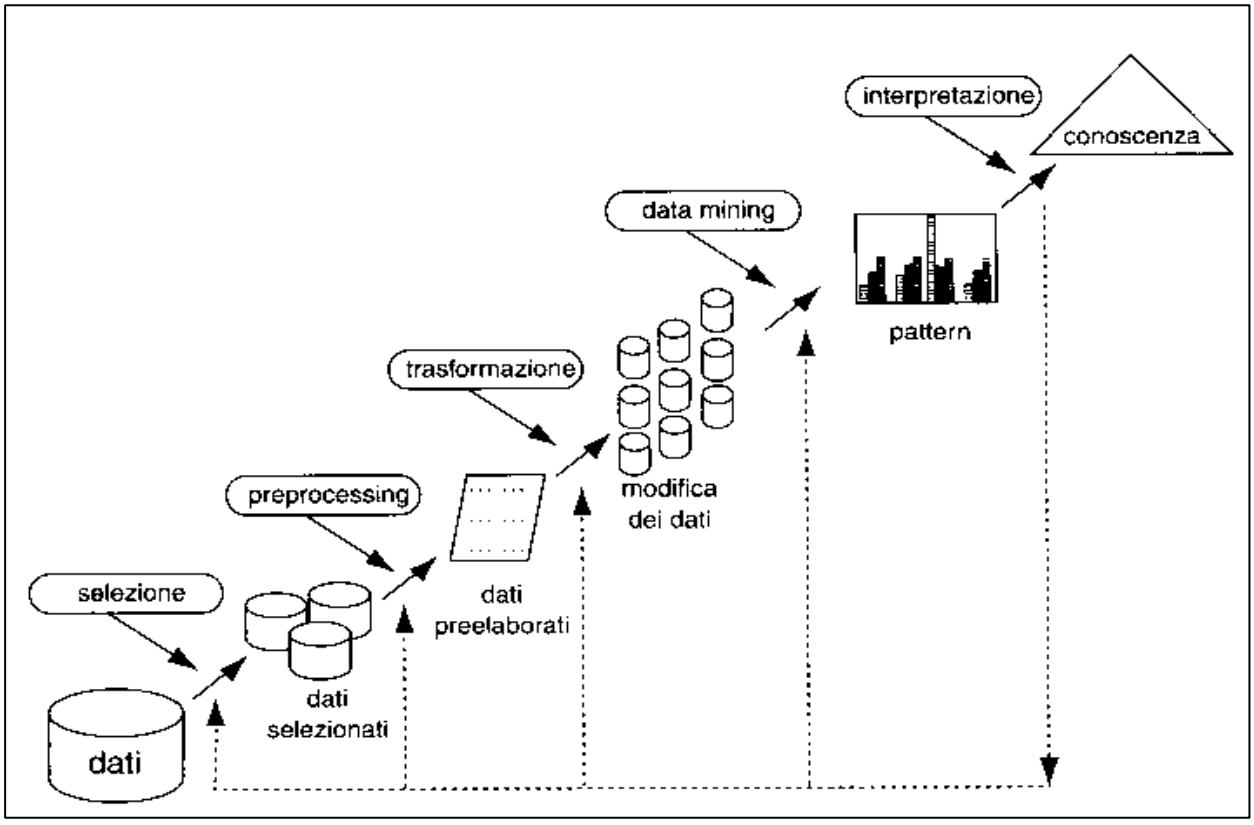
Alcune imprese particolarmente innovative stanno investendo nello sviluppo di una serie di nuovi algoritmi capaci di determinare le leggi di autocorrelazione delle informazioni in maniera deduttiva, senza assoggettare l'analisi a un modello distributivo astratto, ma interpretando le informazioni là dove si trovano: nei dati.

Questo approccio innovativo permette, ad esempio, la creazione di algoritmi capaci di valutare e gestire il rischio estrapolando dalla realtà i modelli matematici necessari, e ha peraltro permesso lo sviluppo di *concept* di *robotrader* per la negoziazione sul mercato Forex (*foreign exchange market)* basati sull'interpretazione quantistica della meccanica finanziaria; nonché di stabilire le variabili di correlazione nel settore della GDO, individuando i tempi e le modalità di promozione di ciascun prodotto grazie all'utilizzo dei predetti modelli matematici.

2.2.2.1 *Data mining*

Venendo più nel dettaglio allo studio delle specifiche tecniche di analisi dei dati, per definizione il *data mining* è il processo di scoperta di relazioni, *pattern* ed informazioni precedentemente sconosciute e potenzialmente utili, all'interno di grandi set di dati. In particolare, un *pattern* indica una struttura, un modello, o in generale, una rappresentazione sintetica dei dati.

Figura 66. Knowledge Discovery in Database (KDD)



Fonte: ITMedia consulting su fonti varie

Dunque il *data mining* si sostanzia in un processo analitico finalizzato a selezionare, esplorare, modellare grosse quantità di dati nella ricerca di relazioni e informazioni non note a priori,

informazioni concretamente tramutabili in azioni commerciali, allo scopo di ottenere un vantaggio economico.

Il risultato della conoscenza è una soluzione direttamente integrabile nei processi decisionali aziendali, derivante da una convergenza tra analisi quantitative e conoscenza del business attraverso un approccio multidisciplinare (sistemi basati sulla conoscenza, sistemi basati sull'auto apprendimento, KDD, informatica, statistica, database, visualizzazione).

I vantaggi dell'applicazione del *data mining* sono molteplici: trattamento di dati di carattere quantitativo, qualitativo, testuale, immagini e suoni; non necessita di analisi aprioristiche da parte del ricercatore, né ipotesi sulla forma distributiva delle variabili; possibilità di elaborare un numero elevato di osservazioni e di variabili; algoritmi ottimizzati per minimizzare il tempo di elaborazione; semplicità di interpretazione del risultato; visualizzazione dei risultati.

Il processo del *data mining*-KDD si suddivide in una serie di fasi:

(i) *definizione e comprensione del dominio applicativo*: individuare le effettive problematiche di business e gli obiettivi da realizzare;

(ii) *creazione di un target data set:* selezione di un sottoinsieme di variabili e di dati o di un campione dei dati;

(iii) *data cleaning e pre-processing:* operazioni per attenuare il rumore nei dati, o degli outlier, selezione delle informazioni necessarie per generare il modello; decisioni sul trattamento dei campi mancanti o incompleti, dei dati rari (con un'eventuale sovra-campionatura) sulla definizione della storicità e sull'aggiornamento dei dati; aggiunta di variabili derivate e indicatori che hanno valori ricavabili da dati già esistenti;

(iv) *data reduction e projection:* definizione della modalità di rappresentazione dei dati secondo gli obiettivi posti, utilizzo di metodi per ridurre il numero delle variabili;

(v) *scelta del ruolo dei sistemi di data mining per l'analisi:* utilizzo dei sistemi di *data mining* per classificazione, regressione, clusterizzazione, etc.;

(vi) *scelta del o degli algoritmi di data mining:* selezione dei metodi per la ricerca di *pattern*, decidendo quali modelli o parametri possono essere più appropriati, integrazione dei metodi di *data mining* scelti con l'intero processo di scoperta della conoscenza;

(vii) *data mining (cuore del processo KDD):* ricerca di modelli di interesse per l'utente, con raffinamenti successivi, presentati secondo definite modalità di rappresentazione (classificazione, alberi di decisione, regressione, cluster analysis…);

(viii) *interpretazione dei modelli identificati:* analisi e verifica dei risultati con possibile retroazione ai punti precedenti per ulteriori iterazioni al fine di migliorare l'efficacia dei modelli trovati;

(ix) *consolidamento della conoscenza acquisita:* integrazione della conoscenza e valutazione delle performance del sistema, mettendo a confronto i risultati con l'effettivo andamento nella realtà dei fatti e produzione della documentazione agli utenti finali o a terze parti interessate.

La novità offerta dagli strumenti di *data mining* non sta nel rinnegare il tipo tradizionale di conoscenza, ma nell'integrare (laddove la cultura aziendale lo permette) i processi decisionali con regole costruite sintetizzando estesi patrimoni informativi. Di conseguenza, i risultati ottenuti attraverso gli strumenti di *data mining* devono essere presentati, comunicati e condivisi con le aree aziendali che ne trarranno i maggiori benefici, privilegiando modelli organizzativi di tipo reticolare.

Le modalità attraverso cui conseguire i risultati prefissati sono riconducibili ai seguenti aspetti:

- esplorazione mediante visualizzazione multidimensionale (*scaling* multimensionale, analisi di regressione logistica, *stepwise*, analisi delle corrispondenze);
- associazione e sequenze (usate nella *market basket analysis* per misurare l'affinità dei prodotti); *clustering* (segmentazione della clientela in gruppi omogenei);
- analisi fattoriale (per determinare il numero di fattori da estrarre e per componenti principali);
- modelli previsionali - di classificazione (alberi di decisione) o reti neurali;
- mappe di Kohonen (reti neurali non supervisionate); algoritmi genetici.

La tabella seguente riporta alcuni esempi applicativi.

Tabella 5. Alcuni esempi applicativi

Esempio	Tipologia di problema	Tecnica adottabile
Quali sono i tre principali motivi che hanno indotto il mio cliente a passare alla concorrenza?	Classificazione	Reti Neurali Decision Tree
Quali sono le fasce di clienti a cui posso offrire nuovi prodotti?	Clustering	Reti Neurali *kohonen. Cluster Analysis
Quali sono le probabilità che un cliente ha aperto un c/c acquisterà anche il prodotto x in breve tempo?	Sequencing	Tecniche statistiche Rule induction
Quali sono le probabilità che un cliente acquisti due prodotti completamente differenti?	Associazione	Tecniche statistiche Rule induction
Quale sarà il prezzo del titolo tra un giorno/mese ecc?	Previsione	Reti neurali Tecniche statistiche

Fonte: Elaborazione di ITMedia Consulting su dati vari

Ciascuna delle tecniche sopracitate presenta dei vantaggi e dei limiti in ragione della loro applicazione, come illustrato in tabella.

Tabella 6. Vantaggi e limiti delle tecniche di data mining

Tecnica	Vantaggi	Limiti
Visualizzazione	L'utente è in grado di visualizzare grandi moli di dati, scoprire relazioni e di testarle.	Richiede un utente esperto in statistica e in grado di utilizzare altre tecniche di Data Mining.
Reti Neutrali	Elevata capacità elaborativa con dati in cui si nascondono relazioni non lineari. Anche con dati incompleti e rumorosi.	Incapacità di spiegare i risultati sebbene sia possibile utilizzare altri sistemi per interpretazioni. Converte dati qualitativi n quantitativi.
Algoritmi generici	Buona capacità previsionale usando dati in cui si nascondono relazioni non lineari.	Incapacità di spiegare i risultati sebbene sia possibile utilizzare altri sistemi per interpretazioni. Converte dati qualitativi n quantitativi.
Logica fuzzy	Può classificare variabili e risultati sulla base di vicinanza alla soluzione.	Numero limitato di fornitori e applicazioni disponibili sul mercato
Decision Tree e Rule induction	Creano regole e modelli sulla base di dati storici. Le regole e i modelli sono trasparenti all'utente e facilmente intepretabili.	Richiedono un tuning ottimale per evitare la produzione di elevati numeri di regole difficilmente interpretabili e gestibili.

Fonte: Elaborazione di ITMedia Consulting su dati vari

E' importante osservare che non esiste una tecnica 'superiore' alle altre, ma ogni tecnica è riferita a determinati obiettivi e tipologie di dati da analizzare. Spesso i migliori risultati per trasformare i dati in informazioni si ottengono attraverso la combinazione di diverse tecniche di analisi.

Per costruire i modelli sulla base dei dati si possono differenziare due tipologie di apprendimento che differenziano le tecniche di *data mining*: *data mining* supervisionato, fondato su algoritmi predittivi; *data mining* non supervisionato, che viceversa sfrutta algoritmi di tipo descrittivo.

In particolare, il *data mining* supervisionato è un approccio *top down* applicabile quando è chiaro l'obiettivo da prevedere, che genera previsioni, stime, caratterizzazioni rispetto al comportamento di alcune variabili target, individuate in funzione di variabili di input. Nei modelli previsionali l'obiettivo è quello di apprendere in modo che la conoscenza acquisita sia applicabile anche in futuro, quindi il modello migliore non è solo quello che presenta migliore efficacia ma quello meglio performante con i dati futuri. Confrontare i risultati del modello con il reale andamento della realtà risulta un passo necessario soprattutto per fare previsioni sui comportamenti futuri.

Le fasi della costruzione e dell'applicazione di un modello previsionale sono:

(i) *La costruzione di un set di apprendimento (training):* il modello viene sperimentato usando dati pre-classificati da un sottoinsieme del set di costruzione (campione dell'universo di riferimento), dove gli algoritmi trovano *pattern* di valore previsionale.

(ii) *La costruzione di un set di valutazione (evaluation):* per verificare il rendimento del modello o per paragonare i risultati di modelli diversi, viene usato un set distinto dagli altri due. Il set di prova consente al modello di generalizzare meglio, evitando il sovra adattamento ai dati, e di stimarne le prestazioni.

(iii) *La costruzione di un set di test:* viene rifinito il modello attraverso un altro sottoinsieme del set di costruzione, per evitare che il modello memorizzi il set di training garantendo che sia più generale e che funzioni meglio con i dati sconosciuti.

(iv) *L'applicazione del modello al set di calcolo (score):* set non pre-classificato e non appartenente al set di costruzione. Non si conoscono i risultati corretti per tali dati e quindi verranno usati i calcoli previsionali per prendere decisioni su di essi. La fase di *scoring* è importante per ottimizzare i risultati di un'operazione commerciale. Lo *scoringsystem* è un insieme di procedure per costruire un modello predittivo in cui l'obiettivo si basa sull'individuazione di una relazione tra una serie di variabili predittive ed una variabile *target*. E' un approccio di analisi incentrato sull'assegnazione ai clienti della probabilità di adesione ad una campagna di marketing, in modo da classificarli per applicare azioni diversificate.

Il *data mining* non supervisionato, viceversa, è un approccio *bottom up* in cui si lascia che i dati stessi indichino un risultato, dove non esiste una variabile target usata per la descrizione e l'individuazione di segmenti. Tale approccio viene spesso applicato nella fase esplorativa per cogliere nelle strutture decisionali un pattern interessane.

2.2.2.2 *Machine Learning*

Nell'ambito del modello previsionale, si colloca la tecnologia del *machine learning* (ML). Diversamente dai programmi software standard, strettamente codificati attraverso regole severe al fine di eseguire determinati compiti, non adattabili a nuove variabili o richieste a meno che un programmatore non li aggiorni con nuove e specifiche regole; il ML utilizza un approccio induttivo per formare una rappresentazione del mondo sulla base dei dati che esamina. E' poi in grado di ottimizzare e migliorare la sua rappresentazione all'aggiungersi di nuovi dati. In questo senso, l'algoritmo "impara" dai nuovi dati inseriti e migliora nel tempo. Dunque, il requisito fondamentale per l'apprendimento automatico è rappresentato da una grande mole di dati, necessari per guidare l'algoritmo. Infatti, il ML è uno strumento che esamina grandi quantità di dati in cerca di modelli, quindi genera il codice che consente di riconoscere quegli stessi modelli in nuovi dati.

Le applicazioni possono utilizzare questo codice così generato per effettuare migliori previsioni. In altre parole, l'apprendimento automatico può aiutare a creare applicazioni più intelligenti. Il funzionamento del ML si basa sostanzialmente su regole statistiche e trova applicazione in diversi ambiti, fungendo da soluzione ad una vasta gamma di problematiche, quali ad esempio la creazione di un software in grado di determinare, con elevata precisione:

(i) se una transazione con carta di credito è da considerarsi fraudolenta;

(ii) la previsione circa le future entrate di un'azienda;

(iii) il rischio sul possibile passaggio di un dato cliente ad altro operatore;

(iv) l'esatto momento in cui un motore a reazione necessita di un intervento di manutenzione;

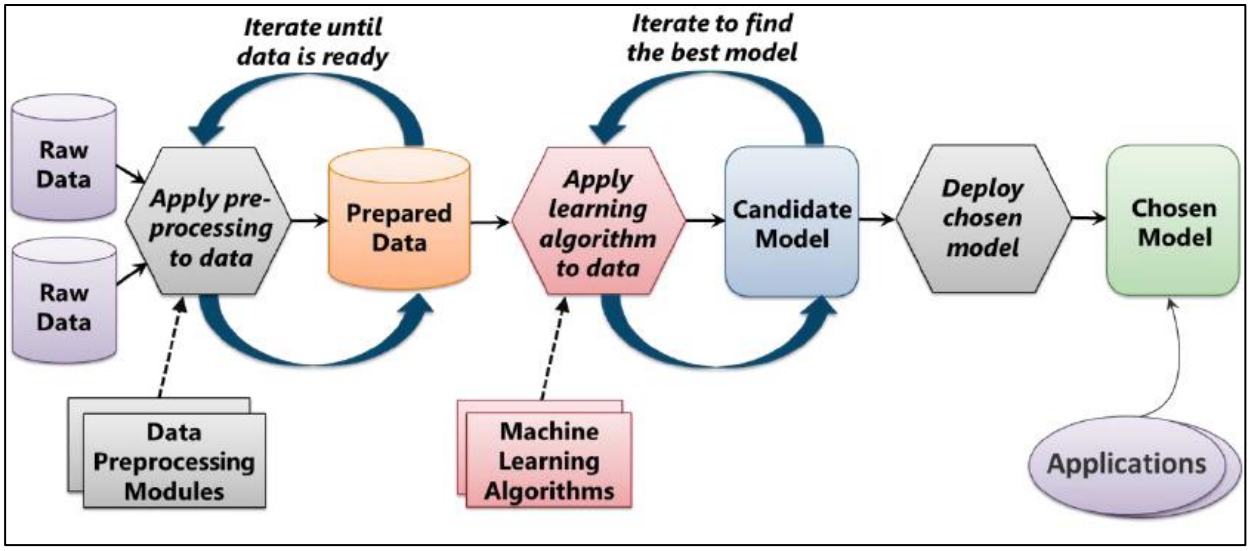
(v) consigli sui film per i clienti;

(vi) effettuare qualsiasi altra analisi su una serie di dati storici disponibili.

Proprio in virtù del fatto che il ML aiuta a predire il futuro, spesso è incluso nella categoria più ampia di analisi predittiva. Tutto ciò che serve a tal fine sono i dati, software di apprendimento per "imparare" da questi stessi dati, e i tecnici che sanno come utilizzare i software impiegati.

La figura 67 mostra come funziona il processo di ML.

Figura 67. Il processo di Machine Learning



Fonte: Introducing azure machine learning

Come si evince dalla figura 67, il processo del ML inizia con i dati, di conseguenza maggiore è il numero di dati di cui si dispone, migliori saranno i risultati ottenibili.

Proprio in ragione del sempre crescente aumento dei dati a disposizioni delle imprese e in generale del mercato, le tecniche di ML stanno diventando sempre più popolari, compreso il cd. *deep learning*.

Tra i più importanti progressi nelle tecniche di apprendimento automatico degli ultimi anni segnaliamo:

(i) *Deep learning*: Questo ramo di ML utilizza le reti neurali profonde con più livelli nascosti. Due dei più comuni tipi di reti neurali profonde sono quella convoluzionale e quella ricorsiva. Reti neurali convoluzionali sono spesso utilizzate per il riconoscimento di immagini, elaborando

un insieme di elementi, ad esempio, per effettuare un collegamento tra un naso, un volto, e, infine, un oggetto/animale vero e proprio. Questa capacità di riconoscere le immagini è un'importante applicazione per lo sviluppo di veicoli autonomi, che hanno bisogno di riconoscere istantaneamente il loro ambiente circostante. Al contrario, le reti neurali ricorsive vengono utilizzate quando l'insieme delle sequenze e il contesto di riferimento sono importanti, come nel caso del riconoscimento vocale o nell'elaborazione del linguaggio naturale. Il *deep learning* è l'esempio più chiaro della confluenza tra grandi moli di dati, potenza di elaborazione, e algoritmi sempre più sofisticati. In realtà, le reti neurali sono state sviluppate decenni fa, ma di certo non disponevano delle enormi quantità di dati e della potenza di elaborazione necessaria per raggiungere le loro piene capacità. Ora che questi ostacoli sono stati superati, i *data scientist* stanno rapidamente sviluppando importantissimi progressi nelle tecniche di *deep learning*;

(ii) *Reinforcement learning*: un'altra forma di ML, prende l'individuazione di una serie di azioni per il raggiungimento di uno specifico obiettivo, tuttavia, non definisce il verso secondo cui tali azioni devono essere intraprese. Gli algoritmi analizzano una vasta gamma di azioni possibili, mentre gradualmente individuano quali sono quelle più efficaci, incorporando in tal modo un elemento di creatività. Questa forma di ML ha trovato diverse dall'imparare come si gioca a scacchi, al miglioramento nella gestione del traffico ai semafori stradali;

(iii) *Ensemble learning*: Questo insieme di tecniche utilizza diversi metodi di ML per ottenere previsioni migliori rispetto a quelle che un qualsiasi metodo potrebbe realizzare da solo. L'insieme di metodi può essere particolarmente utile quando v'è una vasta gamma di possibili ipotesi, facilitando l'individuazione del modello più appropriato. *CareSkore*, per esempio, ha impiegato l'e*nsemble learning* utilizzando la piattaforma Tensorflow di Google per analizzare una serie di dati socio-demografici e comportamentali con l'obiettivo di migliorare l'assistenza sanitaria preventiva.

Queste nuove tecniche sono rese possibili dai nuovi strumenti. *Deep learning libraries* e piattaforme come Tensorflow, Caffe e Theano consentono ai professionisti di integrare rapidamente e facilmente gli algoritmi di *deep learning* all'interno della loro analisi. Spark offre una grande piattaforma di dati che fornisce applicazioni avanzate in tempo reale e analisi predittiva su ampio ventaglio di utilizzi. Apache Hadoop ha reso fruibili efficienti sistemi di storage. Nuovi programmi di ML, API (*application program interfaces*), come quello di Microsoft, consentiranno agli utenti di implementare il ML in nuove aree.

È interessante notare come le basi tecniche del ML hanno più di 50 anni, ma fino a poco tempo fa pochi soggetti, al di fuori del mondo accademico, erano a conoscenza delle sue capacità. Il ML richiede grande potenza di calcolo ed è per questo motivo che originariamente, con riferimento alle prime applicazioni, mancavano le infrastrutture necessarie a che questa tecnica fosse efficacemente implementata.

Attualmente, diverse tendenze convergenti contribuiscono all'aumento di interesse e di attività connesse al ML:

(i) L'implementazione della cd. legge di Moore ha drasticamente ridotto i costi di elaborazione; una massiccia potenza di calcolo è ora ampiamente disponibile ad un costo minimo;

(ii) La disponibilità di algoritmi nuovi e particolarmente innovativi garantisce risultati più veloci;

(iii) *Data scientists* hanno accumulato un importante know-how quanto all'applicazione del ML.

Poiché il processo inizia con i dati, la scelta dei dati su cui lavorare è estremamente critica. Tuttavia, qualunque siano i dati scelti, non capita spesso che siano nelle giuste condizioni per essere utilizzati direttamente.

Per questo motivo il processo di *machine learning* prevede una fase di pulizia dei dati destinata a formare dati "preparati" per la successiva analisi in cui è previsto l'applicazione di un algoritmo. La fase di pulizia dei dati è tutt'altro che semplice e comporta l'applicazione di tutta una serie di moduli di pre-elaborazione dei dati grezzi. Infatti, la creazione di dati "preparati" da una serie di dati grezzi, comporta il dispendio della maggior parte del tempo richiesto per un programma di *machine learning*.

Lo step successivo a quello della formazione dei "prepared data" è quello concernente la definizione della strada più corretta da seguire per il raggiungimento degli obiettivi prefissati. A tal fine, i dati vengono gestiti da algoritmi che vi applicano analisi statistiche. L'obiettivo è quello di determinare quale combinazione di algoritmi di apprendimento automatico e di *prepared data* genera i risultati più utili. Come accennato in precedenza, l'algoritmo di apprendimento automatico agisce su *prepared data* con l'obiettivo di creare un modello. E' importante notare che un modello, tipicamente, non fornisce una risposta secca (affermativa o negativa che sia) ad una domanda ben posta. Al contrario, restituisce una probabilità compresa tra lo zero e l'uno.

Decidere cosa fare di questa probabilità restituita dal modello è di solito una decisione aziendale e non compete ai *data scientists* di turno. Nella maggior parte dei casi, il primo modello realizzato (il cd. *candidate model*) non è mai il migliore. Il *data scientist* infatti tenterà molte combinazioni diverse di algoritmi di apprendimento automatico e *prepared data*, alla ricerca di quella che è in grado di determinare il modello migliore. Ognuno di questi tentativi (vale a dire, ogni iterazione) può essere pensato come un esperimento. Una volta individuata la più giusta combinazione e quindi il modello maggiormente efficace, si passa ad un'ulteriore fase: la distribuzione del modello prescelto. Quest'ultima fase, è il punto di arrivo di tutto il procedimento analizzato, senza il quale l'intera procedura non avrebbe alcuna ragione di esistere. La distribuzione del modello consente alle applicazioni di utilizzare l'algoritmo che ha implemento quel modello, individuando e memorizzando gli schemi predittivi che ricorrono all'interno dei dati trasformati. Se la distribuzione risulta troppo complicata, coloro che hanno creato il modello potrebbero non essere in grado di implementare velocemente nuovi modelli, cosa che è spesso richiesta nel nostro sistema sempre in rapida evoluzione.

Molti aspetti del processo di *machine learning* non sono affatto semplici. Ad esempio, come individuare quali dati grezzi usare; come elaborare i dati grezzi al fine di creare i *prepared data;* quale

combinazioni tra *prepared data* e algoritmi di apprendimento automatico dovrebbe essere utilizzata per creare il modello migliore.

Domande di questo genere vengono generalmente risolte da specialisti di settore, i *data scientist*: esperti con una vasta gamma di competenze nel campo degli algoritmi di apprendimento automatico complessi. Tali professionisti sono ad esempio in grado di individuare quali, tra gli algoritmi presenti, siano capaci di lavorare più efficientemente in situazioni diverse; o ancora si occupano anche della scrittura di codici essendo particolarmente competenti nel campo dello sviluppo di software. Proprio i *data scientist,* anche per le tecniche di ML così come precedentemente visto per quelle di *data mining,* distinguono tra tecniche di ML supervisionato e non supervisionato. Le prime, come visto, richiedono necessariamente una conoscenza preventiva di un dato risultato e forniscono potenti strumenti per la previsione e la classificazione delle informazioni. Spesso, tuttavia, non conosciamo l'esito "finale" di un evento. Per esempio, in alcuni casi di frode, potremmo non sapere se una transazione è fraudolenta prima che sia trascorso un certo tempo, anche piuttosto lungo, dopo il verificarsi dell'evento. In questo caso, piuttosto che tentare di prevedere quali transazioni sono da considerarsi fraudolente, potremmo voler usare il ML per individuare quelle operazioni che risultano quanto meno insolite, in modo da tenerne conto per eventuali ulteriori indagini. Il ML non supervisionato viene invece utilizzato nei casi in cui non conosciamo preliminarmente un risultato specifico, ma si intende estrarre ulteriori informazioni utili dai dati a disposizione.

Le tecniche di ML supervisionato più utilizzate sono:

(i) *Generalized linear models* (GLM) - una forma avanzata di regressione lineare che supporta diverse distribuzioni di probabilità e funzioni di collegamento, consentendo all'analista di modellare in modo più efficace i dati. GLM è un ibrido di statistica classica e il più avanzato ML;

(ii) *Decision Trees* - un metodo di apprendimento supervisionato che memorizza un insieme di regole atte a dividere una popolazione in segmenti sempre più piccoli e che risultano essere omogenei rispetto alla variabile di destinazione;

(iii) *Random forests* – un classificatore d'insieme composto da molti alberi di decisione e che dà in uscita la classe che corrisponde all'uscita delle classi degli alberi presi individualmente. Il processo di classificazione si compone in generale di due sottoprocessi in sequenza: la costruzione del modello a partire da dati già classificati (training)e la classificazione vera e propria, in cui il modello precedentemente creato viene utilizzato per associare a nuovi dati un'etichetta di classe;

(iv) *Gradient boosting machine* (GBM) - un metodo che produce un modello di previsione attraverso la formazione di una sequenza di alberi di decisione, dove gli alberi successivi correggono gli errori di previsione di alberi precedenti;

(v) *Deep learning* - un approccio che si basa su diversi livelli di rappresentazione, corrispondenti a gerarchie di caratteristiche di fattori o concetti, dove i concetti di alto livello sono definiti sulla base di quelli di basso livello.

Le tecniche chiave per il ML non supervisionato sono invece:

(i) *Clustering* - una tecnica atta a raggruppare oggetti che risultano simili tra loro su molti parametri in segmenti, o cluster. La segmentazione della clientela è un tradizionale esempio di raggruppamento. Esistono diversi algoritmi di clustering differenti, il più diffuso è il cd. k-means;

(ii) *Anomaly detection* - il processo di identificazione degli eventi imprevisti o dei possibili risultati. In campi come la sicurezza e la frode, non è possibile indagare in modo esaustivo ogni transazione, è necessario segnalare sistematicamente le operazioni più insolite. Si segnale che la tecnica di *deep learning*, nominata in precedenza, può anche essere utilizzato per il rilevamento di anomalie;

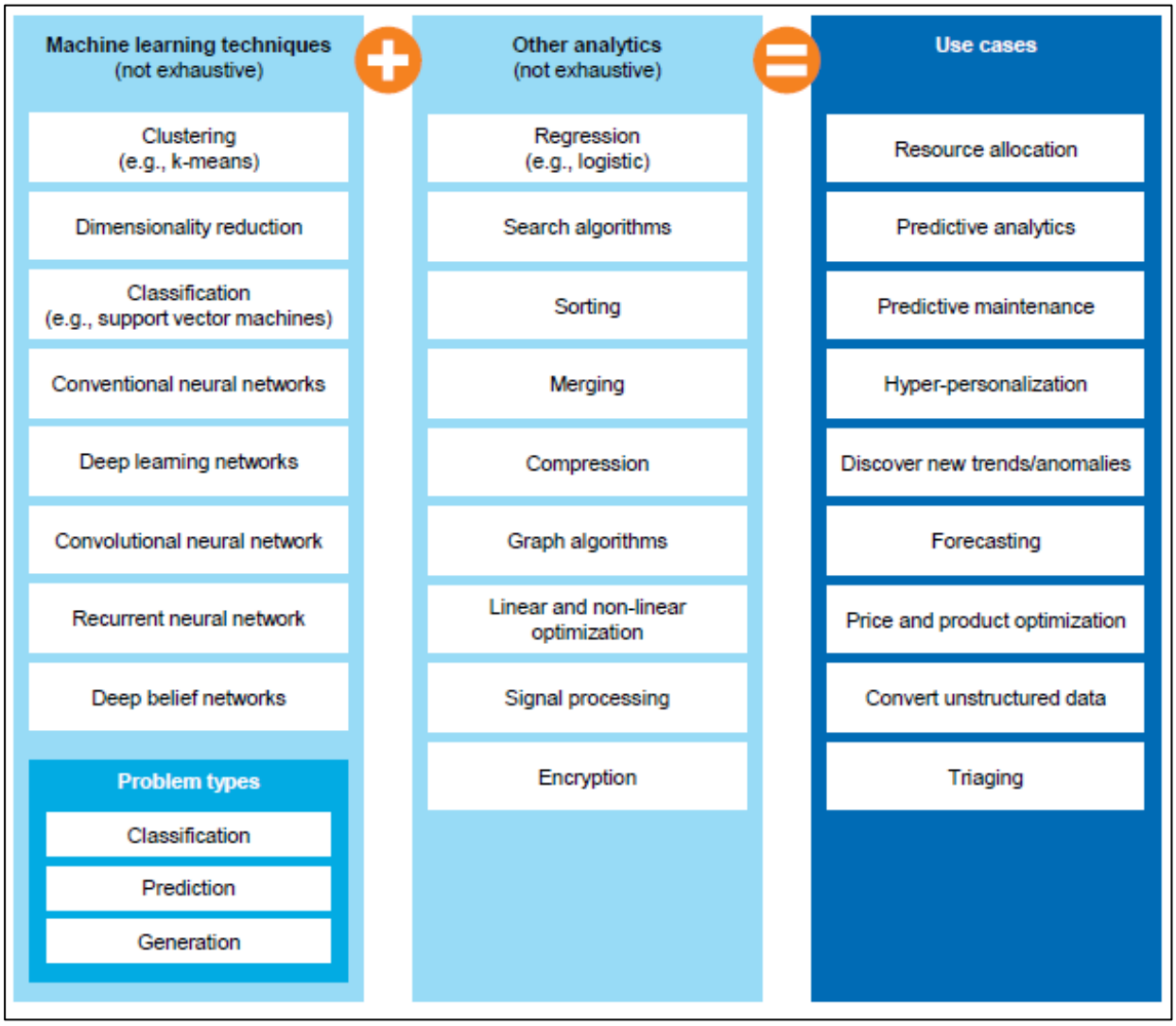
(iii) *Dimension reduction* - il processo di riduzione del numero di variabili considerate. Considerando che le organizzazioni tendono a raccogliere sempre più dati, il numero di possibili informazioni disponibili per formulare previsioni si espande rapidamente. È sufficiente identificare quali dati garantiscono informazioni di valore per un determinato problema. L'analisi delle componenti principali (PCA) valuta un insieme di caratteristiche di base e le riduce a indici che sono indipendenti l'uno dall'altro.

Posto che i dati e la loro analisi possono alimentare la scoperta e l'innovazione e che gli algoritmi sono in grado di supportare e migliorare il processo di *human-decision making*, è altrettanto vero che il ML può, a sua volta, aumentare la potenza di ciascuno degli archetipi sopra descritti.

Questo fondamentalmente per una ragione di fondo: i programmi software convenzionali sono strettamente codificati dall'essere umano che gli impartisce delle istruzioni specifiche circa i compiti di cui necessitano che vengano eseguiti. Al contrario, gli algoritmi di ML sono meccanismi di autoapprendimento, traggono dai dati le informazioni rilevanti senza essere esplicitamente programmati. Il concetto di apprendimento automatico consta nel dare all'algoritmo un massiccio numero di "esperienze" (dati di allenamento) e una strategia di apprendimento generalizzata, per poi lasciare che lo stesso algoritmo identifichi i modelli, le associazioni, e gli approfondimenti del caso, dai dati a disposizione. In sostanza, questi sistemi sono "addestrati" e non programmati.

Il ML può essere combinato con altri tipi di analisi per risolvere una grande fetta di problemi di business:

Figura 68. Tipologie di analisi combinabili con il ML



Fonte: McKinsey Global Institute analysis

Le capacità del ML sono le più adatte per la risoluzione di tre grandi ordini di problemi, come illustrato nello schema sottostante:

(i) Classificazione;

(ii) Predizione/stima;

(iii) Generazione.

Problemi di classificazione riguardano l'osservare la realtà circostante, tra cui l'identificazione di oggetti in immagini e video, e il riconoscimento del testo e dell'audio. La classificazione riguarda anche le associazioni che si trovano nei dati o le segmentazioni in cluster, attività particolarmente utile per segmentare la clientela. Il ML può anche essere usato per predire la probabilità di eventi e

eventuali risultati. Infine, può essere usato per produrre i contenuti, dall'inserire i dati mancanti a generare il fotogramma successivo all'interno di una sequenza video.

Figura 69. Le tre tipologie di problemi risolvibile tramite l'applicazione del ML

Classification	Classify/label visual objects	Identify objects, faces in images and video
	Classify/label writing and text	Identify letters, symbols, words in writing sample
	Classify/label audio	Classify and label songs from audio samples
	Cluster, group other data	Segment objects (e.g., customers, product features) into categories, clusters
	Discover associations	Identify that people who watch certain TV shows also read certain books
Prediction	Predict probability of outcomes	Predict the probability that a customer will choose another provider
	Forecast	Trained on historical data, forecast demand for a product
	Value function estimation	Trained on thousands of games played, predict/estimate rewards from actions from future states for dynamic games
Generation	Generate visual objects	Trained on a set of artist's paintings, generate a new painting in the same style
	Generate writing and text	Trained on a historical text, fill in missing parts of a single page
	Generate audio	Generate a new potential recording in the same style/genre
	Generate other data	Trained on certain countries' weather data, fill in missing data points for countries with low data quality

Fonte: McKinsey Global Institute analysis

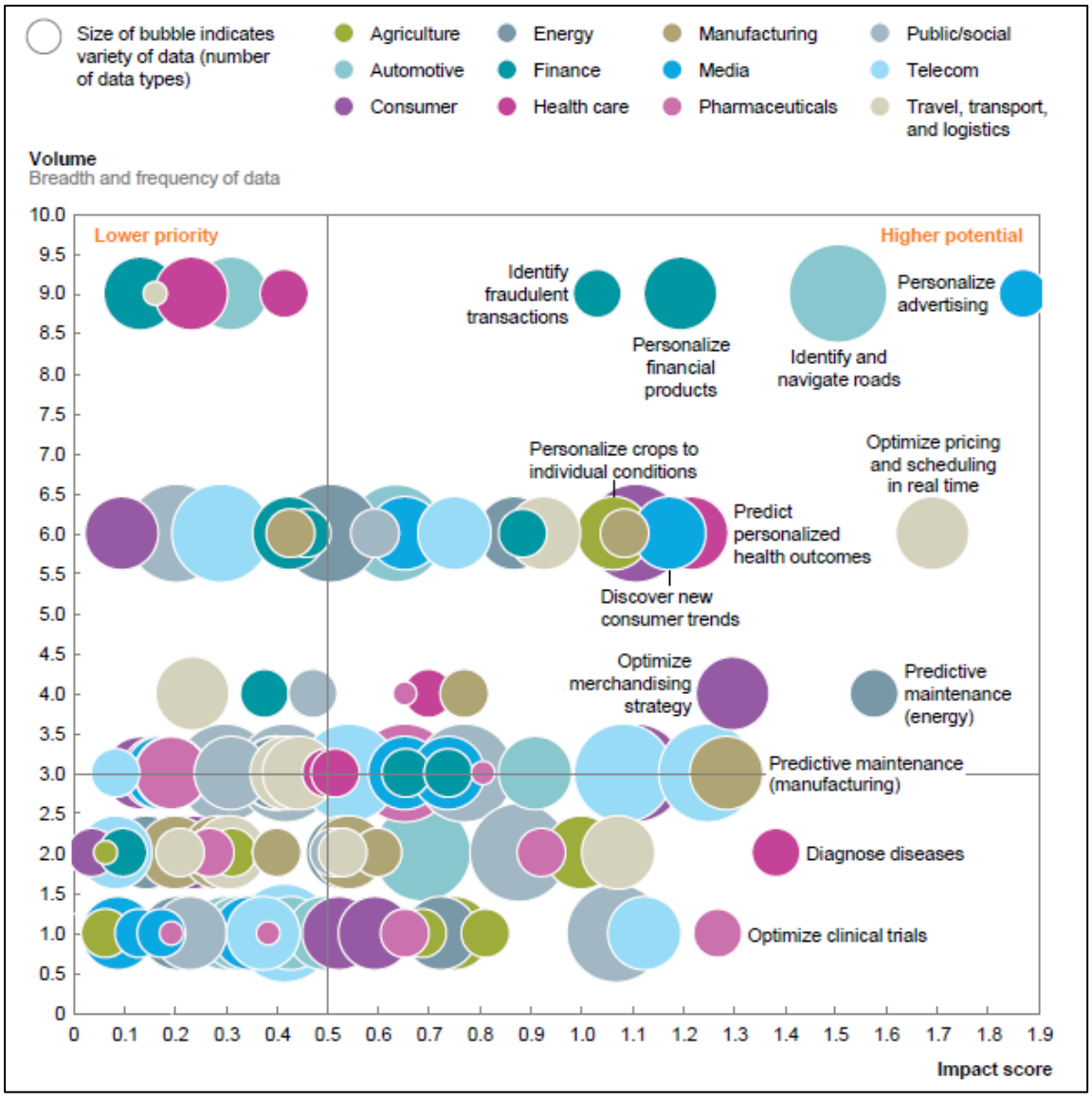
2.2.2.3 Le applicazioni del *Machine learning* e lo *statistical learning*

Il ML può trovare applicazione in un'ampia varietà di ambiti per la risoluzione delle problematiche più disparate, ad esempio: fuori dalle piste, i team della F1 competono in una corsa alle innovazioni tecnologiche per rendere le loro auto più veloci. Top team di F1 investono centinaia di milioni di dollari ogni anno in sviluppo, puntando ad incrementare la tecnologia al fine di aumentare la velocità della auto. Recentemente, sono tre i team di F1 che hanno investito in ML per contenere i costi delle loro divisioni operative che generalmente assorbono più dell'80% delle risorse per lo sviluppo. Basandosi su anni di diversi progetti sui dati, tra cui i registri CAD, dati sulle risorse umane, comunicazioni dei dipendenti, si sono ricercati modelli che avessero influenzato l'efficienza di un determinato progetto individuale. È stato scoperto, per esempio, che anche molti ingegneri o lunghe interruzioni di lavoro generalmente aumentano le ore di lavoro su un determinato progetto dal 5 al 6%, mentre l'uso da parte del team del sistema di documentazione migliora la produttività di oltre il 4%. Nel complesso, questa applicazione ha ridotto il budget dal 12 al 18%, con un risparmio di milioni di dollari. Un'altra applicazione di ML, analisi predittiva, ha dimostrato di essere particolarmente efficace nell'individuazione delle frodi. Ad una grande impresa di assicurazioni auto, gli alti tassi di incidenti per le polizze di nuovi clienti hanno suggerito che le domande venivano depositate per preesistenti danni. Il modello di ML è stato in grado di utilizzare i dati per identificare i diversi gruppi di nuove polizze con il tasso di incidenti di sei volte superiore a quello della media. Questi gruppi hanno costituito la base per definire una nuova strategia di prezzo che ha portato ad un

miglioramento della redditività di oltre il 10%. Ancora, un importante banca inglese attraverso algoritmi di ML ha identificato una serie di operazioni fraudolente con un tasso di errore pari a solo il 10%. In un altro esempio, un noto processore di pagamento ha applicato il ML sui suoi dati riguardanti un numero molto esteso di transazione al fine di identificare il cd. "mule accounts" coinvolto nel riciclaggio di denaro.

Nello schema seguente, è proposto uno studio di McKynsey in cui sono stati tracciati i primi 120 casi d'uso del ML nell'ambito di 12 diversi settori. Sull'asse y sono stati riportati i volume di dati disponibili, mentre sull'asse x è stato indicato il potenziale impatto del ML sulle attività interessate sulla base di indagini condotte da oltre 600 esperti del settore. La dimensione della bolla riflette la diversità delle fonti di dati disponibili.

Figura 70. I settori di applicazione del ML



Fonte: McKinsey Global Institute analysis

Diversi saranno gli usi specifici del settore che uniscono la ricchezza dei dati con un'opportunità, più grandi saranno le bolle nel quadrante in alto a destra del grafico. Queste rappresentano le aree dove le organizzazioni dovrebbero dare priorità all'uso del ML e prepararsi ad una trasformazione senza precedenti. Alcuni dei casi d'uso in cui si registrano le più alte opportunità includono la pubblicità personalizzata; i veicoli autonomi; l'ottimizzazione dei prezzi, il *routing*, e la pianificazione sulla base di dati in tempo reale sulla mobilità e la logistica; predire risultati personalizzati sulla salute; e l'ottimizzazione delle strategie di merchandising nella vendita al dettaglio.

I casi d'uso nel quadrante in alto a destra si dividono in quattro categorie principali. In primo luogo è evidente che la personalizzazione dei prodotti e dei servizi per i clienti in settori quali prodotti confezionati, finanza e assicurazioni, sanità e media, rappresentano un'opportunità che la maggior parte delle aziende non ha ancora sfruttato appieno. La seconda categoria è quella dell'analisi predittiva. Ci si riferisce ad esempi come il *triaging* di chiamate al servizio clienti; la segmentazione della clientela sulla base di modelli di rischio, di abbandono e di acquisto; l'individuazione delle frodi e di anomalie bancaria, ovvero di sicurezza informatica; la possibilità di diagnosticare malattie da scansioni, biopsie, e altri dati. La terza categoria è rappresentata dall'ottimizzazione strategica, che comprende usi come il merchandising e le procedure di ottimizzazione nella vendita al dettaglio, la programmazione e l'assegnazione di lavoratori in prima linea, e l'ottimizzazione di team e altre risorse attraverso aree geografiche e relazionali. La quarta ed ultima categoria è l'ottimizzazione delle operazioni e della logistica in tempo reale, che include impianti e macchinari automatici atti a ridurre gli errori e migliorare l'efficienza, ottimizzando la gestione della catena di fornitura.

Figura 71. L'impatto del Machine Learning nei diversi settori industriali

Fonte: McKinsey Global Institute

Sempre nell'ambito dello stesso studio, McKinsey ha analizzato e misurato l'impatto del ML all'interno dei diversi settori industriali, come mostrato nella mappa riportata in figura 70.

Ai fini di una più approfondita analisi e di una maggiore comprensione degli usi applicativi del ML, si riportano di seguito una serie di esempi applicativi analizzati più nel dettaglio:

- *Prevenzione delle frodi*. Con più di 150 milioni di portafogli digitali attivi, 200 miliardi di dollari in pagamenti annuali, PayPal è leader nel settore dei pagamenti online. Considerando queste cifre, anche i bassi tassi di frode possono avere una certa consistenza. Soprattutto agli inizi, l'azienda stava perdendo circa 10 milioni di dollari al mese per il numero di frodi effettuate. Per affrontare il problema, PayPal ha costituito un top team di ricercatori, che hanno utilizzato le tecniche di ML esistenti al fine di per costruire modelli in grado di identificare i pagamenti fraudolenti in tempo reale;

- *Targeting digital display.* Una società tecnologicamente molto avanzata, Dstillery, utilizza il ML per supportare aziende come Verizon e Williams-Sonoma ad indirizzare in tempo reale la *display advertising* sulle piattaforme. Utilizzando i dati raccolti sulla cronologia di navigazione di un individuo, pagine visitate, clic e acquisti, Dstillery elabora migliaia di previsioni al secondo, gestisce centinaia di campagne alla volta, sovra performando in modo significativo le operazioni di marketing svolte invece da un singolo individuo e registrando un impatto particolarmente efficiente per ciascun dollaro speso;

- *Recommending content*. Per i clienti di Comcast X1, servizio di TV interattiva, la società offre consigli personalizzati in tempo reale sui contenuti, basandosi sulle scelte di visioni precedenti effettuate da ciascun utente. Comcast utilizza tecniche di ML per sviluppare un profilo per ciascun utente in relazione alle rispettive preferenze, raggruppando poi in cluster gli utenti aventi preferenze comuni. Per ciascun cluster, la società tiene traccia e visualizza i contenuti più popolari in tempo reale, in modo da consentire agli utenti di vedere quali contenuti sono attualmente dei trend. Il risultato netto equivale a migliori raccomandazioni, utilizzi più elevate e clienti maggiormente soddisfatti;

- *Costruzione di auto migliori.* Le nuove vetture costruite da Jaguar Land Rover hanno 60 computer di bordo che producono 1,5 GB di dati ogni giorno in più di 20.000 metriche. Gli ingegneri della società utilizzano le tecniche di ML per estrarre i dati e capire come i clienti effettivamente lavorano con il veicolo. Lavorando con i dati reali, i *designers* possono prevedere parziali fallimenti e potenziali problemi di sicurezza; questo li aiuta a costruire veicoli in modo appropriato rispetto alle condizioni attese;

- *Migliori prospettive di targeting*. I *marketers* utilizzano modelli di "propensione all'acquisto" come strumenti per determinare le migliori prospettive di vendita, di marketing e i migliori prodotti da offrire;

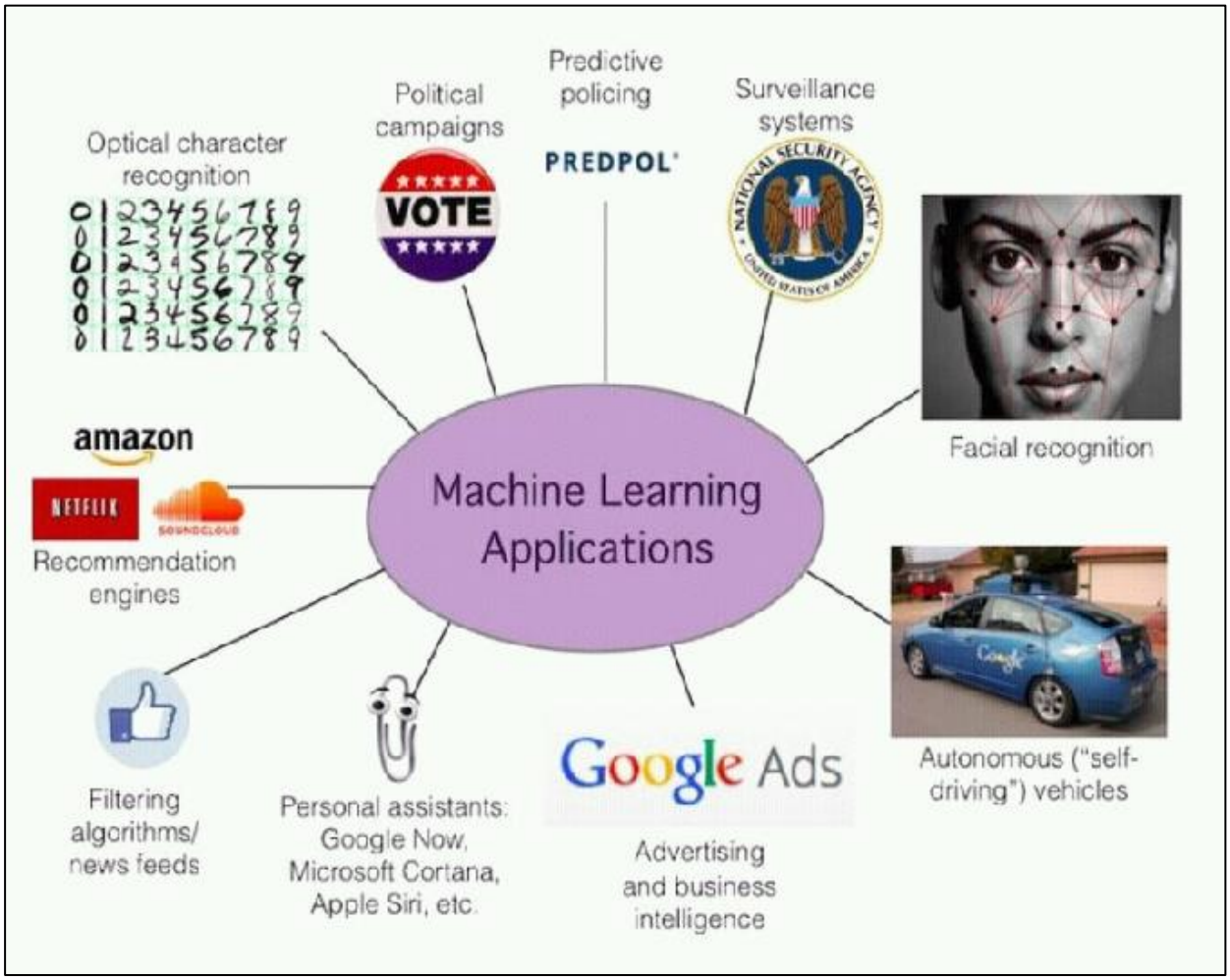
- *Ottimizzazione dei media.* NBC Universal archivia centinaia di terabyte di file multimediali per la distribuzione di TV via cavo internazionale. Una efficiente gestione di questa risorsa on-line

è necessaria per garantire la distribuzione ai clienti internazionali. L'azienda utilizza il ML per prevedere la domanda futura di ciascun servizio televisivo basato su una combinazione di parametri. Sulla base di queste previsioni, l'azienda archivia quei media per i quali è prevista una bassa domanda. Le previsioni effettuate con l'utilizzo del ML sono certamente più efficaci di regole arbitrarie basate su singoli parametri, come l'età del file. Di conseguenza, NBC Universal riduce i costi complessivi di archiviazione pur mantenendo un alto livello di soddisfazione del cliente;

- *Migliorare la fornitura dei servizi di assistenza sanitaria*. Il sistema sanitario Carolinas (CHS) utilizza il ML per costruire punteggi di rischio per i pazienti, utilizzati dai *case manager* nell'ambito delle loro decisioni sul dimettere o meno il paziente. Questo sistema permette un migliore utilizzo di infermieri e case manager, privilegiando i pazienti in base al rischio e la complessità del caso. Di conseguenza, CHS ha abbassato il suo tasso di ri-ricovero dei pazienti dal 21% al 14%;

- *Lotta alla criminalità.* Gli algoritmi sono più di recente utilizzati anche dalle forze di polizia. In città come Los Angeles, Atlanta e Philadelphia, algoritmi " *predictive policing*" esaminano i dati di precedenti crimini per informare gli ufficiali su quali persone e luoghi risultano essere maggiormente a rischio per i possibili crimini futuri. Il più popolare è PredPol, un algoritmo sviluppato dal Dipartimento di Polizia di Los Angeles in collaborazione con le università locali che utilizza dati reali sul dove e quando i crimini sono avvenuti, per poi elaborare una mappa "hotspot" in cui verranno identificati i luoghi in cui probabilmente si verificherà il prossimo crimine. In altre parole, si tratta di un algoritmo segretissimo in grado di predire e localizzare atti delittuosi. Tuttavia, la funzione «anticipatoria» nella quale questi algoritmi s'inscrivono si fonda su criteri d'incriminazione molto labili e sulla soggettività (collettiva o individuale) e non su una base giuridica. Si arriva così a un determinismo sociale e razziale. I dati che vengono analizzati nei complessi modelli matematici che portano alle predizioni non sono infatti soltanto quelli che riguardano i luoghi, le date e le persone coinvolte nei delitti del passato. Vengono considerati anche dati in grado di determinare cluster di potenzialità criminale, legati a fattori prettamente geografici. La questione della predizione del crimine è certamente molto delicata, e l'utilizzo di algoritmi è inquietante soprattutto perché questo tipo di pratiche predittive implica una forma di determinismo, come se la criminalità venisse annoverata fra le disuguaglianze sociali. Dimenticando qualsiasi fattore sociologico che conduce a comportamenti criminali, va tuttavia evidenziato che la tecnologia, se ben usata, è un aiuto prezioso (in particolare se si considera che nelle città dove è in uso questo genere di algoritmo ha determinato una diminuzione della criminalità dal 10 al 30%, a seconda del tipo di crimine) ma farne una scienza esatta (come per esempio assumere che il futuro assomiglia al passato, cosicché i delitti di domani replicheranno quelli di ieri) è restrittivo.

Figura 70. Applicazioni pratiche del Machine learning



Fonte: Elaborazione ITMedia Consulting su dati vari

A conclusione del paragrafo, analizziamo un ultimo aspetto dell'analisi predittiva, ovvero il confronto tra tecniche di auto apprendimento e analisi statistiche.

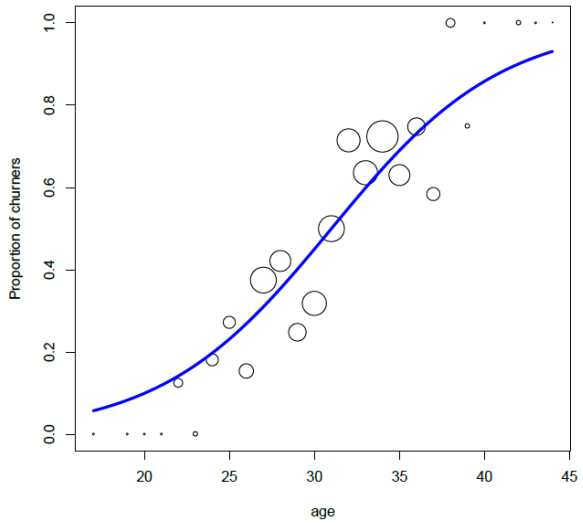
ML e *Statistical Learning* (SL) rappresentano due facce della stessa medaglia. Mentre il ML è finalizzato alla costruzione di algoritmi che possono imparare dai dati; lo SL invece rappresenta quella parte della statistica applicata che emerge come risposta al ML, enfatizzando i modelli statistici e la valutazione dell'incertezza. Le finalità delle due applicazioni sono le medesime: descrivere trend e andamenti, capire il presente, prevedere il futuro; viceversa, il fulcro delle rispettive attività invece cambia: per il ML l'algoritmo e il suo output è centrale, per lo SL i modelli statistici e la quantificazione dell'incertezza sono centrali. Inoltre, va chiarito anche il concetto di *data science* che è sostanzialmente l'attività di estrazione di conoscenza dai dati, attraverso strumenti matematici, statistici, ML, informatici, ingegneristici, fisici.

Esemplificando, proponiamo di seguito alcuni esempi per meglio comprendere la differenza tra i due strumenti di analisi. Partendo dal modello di classificazione statistica, supponiamo la necessità di dover compiere una *churn analysis* nell'ambito di un servizio di telecomunicazione, ossia una

previsione sui clienti che verosimilmente in un certo tempo disdiranno il loro contratto/abbonamento per passare alla concorrenza. Ebbene, un'operazione di questo tipo consiste nell'attribuzione di un'etichetta del tipo "*churn/not churn*" alle informazioni note al momento dell'analisi, con l'obiettivo di prevedere oggi, il possibile *churn* di domani. L'etichetta in questione rappresenta la cd.variabile risposta o target. Le informazioni note sulle quali viene compiuta la classificazione, possono essere rappresentate da variabili numeriche (n. connessioni alla rete, n. chiamate, durata chiamate, età del cliente…) o variabili qualitative (genere del cliente, piano tariffario, handset, ecc) e rappresentano le cd.variabili esplicative o variabili di input. E' evidente che il target non è definito perfettamente dagli input, ma ci sarà sicuramente una componente casuale/aleatoria. Un modello statistico prevede una relazione di tipo:

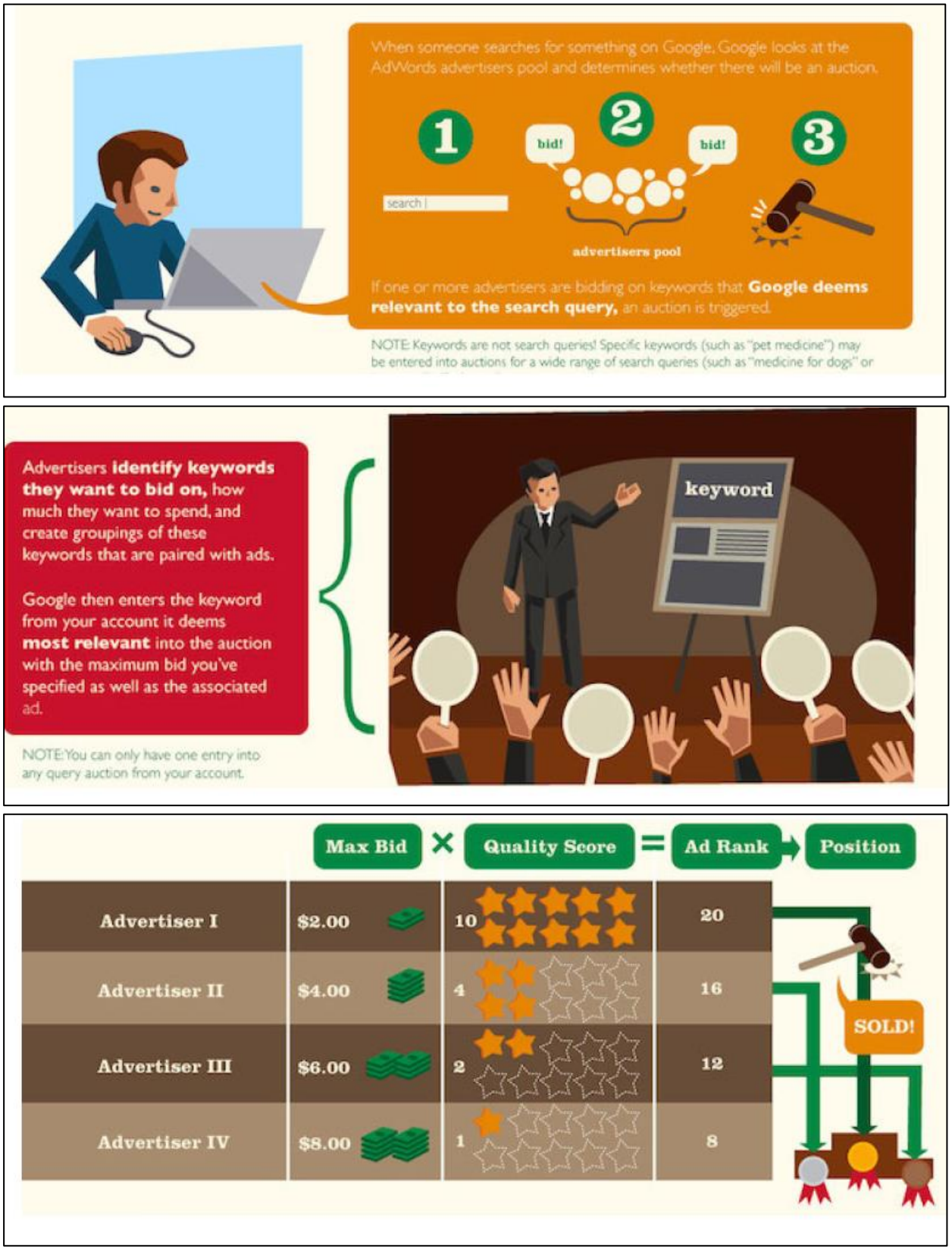
$$Pr(target = churn \mid input) = f(input)$$

Conoscere l'esatta probabilità non mi consentirà di sapere con altrettanta esattezza quale sarà l'etichetta (*churn/not churn)*, pertanto è ragionevole pensare che si etichetterà un determinato cliente come futuro *churn* se la sua Pr(target = churn) sarà ad esempio maggiore del 50%. Inoltre, va detto che la f non è nota ma va stimata (^f) sulla base delle informazioni disponibili ovvero imparando il meccanismo con cui dagli input f ricaviamo l'output. Il grafico seguente mostra una rappresentazione del modello semplificato ovvero della probabilità del *churn* sulla base della variabile età utente.



L'algoritmo di classificazione funziona attraverso una serie di operazioni in parte effettuate dall'elaboratore, in parte effettuate automaticamente. In primo luogo lo statista sceglie su quale output (*churn/not churn)* intende sviluppare l'analisi. In secondo luogo, automaticamente viene utilizzato lo storico per stimare ^f secondo alcuni criteri che minimizzano una qualche misura dell'errore di previsione. A questo punto, sempre automaticamente, si applica ^f a nuovi dati di cui si conoscono unicamente gli input per ottenere una probabilità che, infine, verrà classificata come *churn* se supera il 50%, condizione quest'ultima che lo statista aveva preliminarmente stabilito. A questo punto, per ulteriore esemplificazione del procedimento, consideriamo un caso pratico, il servizio Google AdWords.

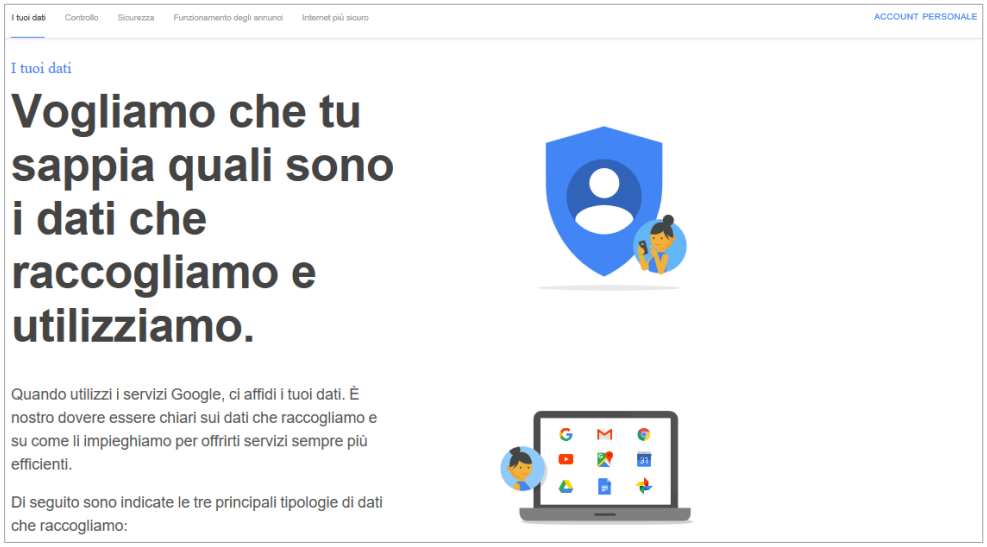
Figura 71. Un caso pratico: il servizio Google AdWords



Fonte: www.hubspot.com

In particolare, nel momento in cui l'utente effettua una ricerca su Google, il motore di ricerca associa automaticamente a quella ricerca una serie di *keyword* e, per ogni cliente che ha delle offerte su quella keyword, associa un indice (quality score) sulla rilevanza dell'inserzione sulla query. A quel punto, Google calcola AdRank (max bid _ quality score = AdRank) e restituisce i risultati. Analizzando questo esempio, il motore di ricerca parte dall'associazione automatica alla ricerca delle *keyword,* in tal senso, la query costituisce l'input e le keyword sono i possibili argomenti. Nello scegliere le *keyword* Google cerca di prevedere la probabilità con cui l'utente intende un certo significato, ed è in ciò sta il problema di classificazione.

L'algoritmo di apprendimento dipende dall'uomo non solo per la scelta del tipo di modello ma anche per quanto riguarda l'identificazione dei dati su cui "allenarlo". La categorizzazione distorta delle etichette da parte dell'uomo porta inevitabilmente in una distorsione del modello di classificazione. Tendenzialmente, la scelta della famiglia di modelli, dei parametri di settaggio e di tutti i dettagli tecnici non sono note negli algoritmi di classificazione degli operatori digitali. Tali elementi rappresentano una parte rilevante della capacità di estrazione di valore dai dati da parte dell'impresa. Ad ogni modo, diversi player presentano informazioni più o meno dettagliate sulle modalità di raccolta e utilizzo dei dati forniti dagli utenti. Google, fornisce ad esempio, nell'account personale di ogni utente una policy di trasparenza[60] sui dati di input che esplicita quali tipologie di dati personali utilizza ai fini della profilazione di ciascun utente.



In conclusione è da ritenere che i dati e le informazioni rappresentino una ricchezza da valorizzare per creare lavoro, imprese e innovazione. La capacità di gestire da questi patrimoni informativi e di estrarre da essi informazioni attraverso gli *analytics* acquisterà sempre più valore. Ne trarranno beneficio le imprese, nel marketing, nei rapporti con i clienti, i fornitori e i consumatori, così come le Pubbliche Amministrazioni che avranno input preziosi per tutto quanto riguarda la pianificazione (territorio, mobilità, ecc.) e l'allineamento dei servizi alla domanda di cittadini e imprese. In

[60] https://privacy.google.com/your-data.html

particolare, alla luce di quanto fin qui descritto, emerge con evidenzia che il valore dei dati non è racchiuso nel dato stesso, ma è frutto di un articolato processo di creazione di valore in cui l'accesso al dato rappresenta soltanto una delle componenti.

Tale processo parte da un piano d'azione sviluppato dall'impresa e finalizzato al raggiungimento di un obiettivo di business ben definito, la cui realizzazione passa attraverso una particolare tecnica di elaborazione e analisi di specifiche tipologie di dati, di cui l'azienda dispone o che potrebbe disporre, strettamente correlata al risultato da conseguire. Ad oggi i dati e gli *analytics* hanno ancor più potenziale nella creazione di valore di quanto ne avessero per le prime imprese pioniere. Le organizzazioni che sono in grado di sfruttare queste capacità in modo efficace saranno in grado di creare un valore significativo e di differenziarsi, mentre le altre si troveranno sempre più in una condizione di svantaggio, incapaci di competere e progredire.

2.2.2.4 *Reccomender System*

Un esempio significativo di applicazione delle tecnologie e degli algoritmi di analisi big data per l'estrazione di valore per il business è rappresentato dai Reccomender System spesso indicati anche con il termine Reccomendation Engine. Si tratta di una classe di sistemi di filtraggio delle informazioni che mira a prevedere il livello di preferenza che un utente darebbe a un dato oggetto o elemento. Diventati sempre più popolari in tempi recenti in contesti di marketing personalizzato e contestuale, tali sistemi abbracciano una vasta area di settori e attività economiche: musica, news, libri, film, servizi finanziari, assicurazioni, ristoranti, persone, etc.

Utilizzati dalle grandi piattaforme web OTT per incrementare l'engagement e il tempo speso online dagli utenti, aumentare le vendite e minimizzare il churn rate, questi sistemi costituiscono una valida alternativa agli algoritmi di ricerca dato che aiutano gli utenti a scoprire oggetti o contenuti che difficilmente potrebbero trovare da soli.

Il principale rischio, di cui tener conto nell'ambito della digital strategy aziendale, risiede nella possibilità di disturbare un utente in una specifica situazione (ad esempio non è consigliabile presentare una reccomendation durante una riunione aziendale, di prima mattina, etc.) in modo tale da influire negativamente sulla web reputation aziendale.

La logica elaborativa alla base della produzione delle reccomendation si basa essenzialemente su due approcci di analisi e filtraggio dei dati: collaborative filtering e content-based filtering. I sistemi content-based sono più semplici da realizzare ma producono reccomendation meno interessanti mentre i sistemi basati sull'approccio collaborative filtering consentono di ottenere risultati molto accurati al prezzo di un livello elevato di complessità sia in termini di algoritmi che di volume di dati richiesti.

Inoltre, nei casi più avanzati questi approcci sono utilizzati in mondo congiunto dando luogo ai c.d. Hybrid Recommender System.

Collaborative filtering

Tale approccio è spesso riferito *model-based* in quanto basato sulla definizione di un modello descrittivo dell'utente sulla base del suo comportamento passato (ad esempio oggetti comprati o selezionati, rating dato a questi oggetti, etc.). A tal fine è necessario raccogliere ed analizzare una grande quantità di dati sulle attività e le preferenze dell'utente. Le reccomendation sono prodotte tenendo conto del modello e delle preferenze di numerosi altri utenti considerati "simili". L'assunzione di fondo è che gli utenti che hanno mostrato un comportamento e preferenze simili in passato continueranno a mostrarlo in futuro, preferendo quindi oggetti simili a quelli preferiti in passato.

La definizione del modello utente sulla base del suo comportamento si avvale della raccolta di dati raccolti in modo implicito ed esplicito.

Esempi di raccolta dati in modo esplicito sono:

- chiedere all'utente di valutare un oggetto sulla base di una scala predefinita di valori;
- chiedere all'utente di effettuare una ricerca;
- chiedere all'utente di classificare un elenco di oggetti/elementi in ordine di preferenza;
- presentare all'utente due oggetti a chiedergli di scegliere il migliore tra i due;
- chiedere all'utente di creare una lista di oggetti preferiti.

La modalità di raccolta dati in maniera implicita appare più interessante e si basa sulle seguenti operazioni:

- osservare e memorizzare gli oggetti che l'utente visualizza sull'online store;
- memorizzare gli oggetti comprati online dall'utente;
- analizzare il numero di volte che l'utente visualizza gli oggetti;
- analizzare l'attività dell'utente sui social network per scoprirne le preferenze.

Uno degli esempi più noti di collaborative filtering reccomender system è sicuramente dato dagli algoritmi utilizzati da Amazon basati proprio sull'assunzione che l'utente che ha comprato un dato prodotto molto probabilmente comprerà un alto prodotto già scelto da utenti "simili". Anche le piattaforme Social quali Facebook e Linkedin applicano algoritmi basati su tale approccio per analizzare la rete di collegamenti di un dato utente al fine di suggerirgli nuovi amici, gruppi etc. Un altro esempio è rappresentato da Last.fm che crea una "stazione radio" di canzoni suggerite analizzando regolarmente la musica che un utente ascolta per poi compararla con quella ascoltata da utenti dal comportamento simile. Vengono così suggerite canzoni mai ascoltate dall'utente cercando tra quelle più ascoltate da utenti con interessi simili.

I principali elementi di attenzione per la massima efficacia dell'approccio suddetto sono:

- *cold start:* questi sistemi richiedono come prerequisito una grande quantità di dati sull'utente al fine di fornire reccomendation accurate;
- *scalabilità:* i sistemi suddetti operano frequentemente in situazioni caratterizzati dalla presenza di diversi milioni di prodotti e utenti per cui è necessario un elevato livello di potenza di calcolo e di risorse computazionali per calcolare le reccomendation;
- *sparsità:* dato la vastità di oggetti venduti su una tipica piattaforma di e-commerce, è molto probabile che il rating dato dagli utenti sia relativo ad un insieme limitato di oggetti polarizzando in qualche modo le reccomendation.

Un modo semplice per descrivere tale approccio è dato dalle seguenti espressioni:

"Utenti che hanno scelto questo prodotto hanno scelto anche questi prodotti";

"Utenti simili a te hanno scelto anche questi prodotti".

Content-based filtering

Questa tecnica si basa sull'utilizzo di una descrizione degli oggetti che tiene conto di uno o più di attributi al fine di suggerire oggetti con caratteristiche simili a quelli preferiti in passato dall'utente. Per ciascun oggetto viene creato un profilo dato da un insieme di attributi e funzionalità. Sulla base del comportamento passato dell'utente viene assegnato a ciascun attributo un peso. Ogni oggetto all'interno del sistema viene così classificato tramite un vettore pesato di attributi. I pesi riflettono l'importanza di ogni attributo o caratteristica per l'utente e possono essere calcolati utilizzando diverse tecniche a partire da una semplice media pesata delle preferenze degli utenti ad approcci più sofisticati basati sul machine learning come reti neurali, cluster analysis, decision tree finalizzati a determinare la probabilità di preferenza di un oggetto per un dato utente. I feedback assegnati direttamente dall'utente, solitamente in termini di "like" o "dislike", sono usati per assegnare un peso maggiore o minore all'importanza di certi attributi.

Un tipico esempio di content-based reccomender system è rappresentato da Pandora Radio. Utilizzando gli attributi descrittivi di una canzone o artista (tratti dagli oltre 400 attributi definiti dal progetto "Music Genome"), il sistema viene alimentato con un seme iniziale di attributi preferiti dall'utente per determinare le canzoni da trasmettere. Le reccomendation sono man mano raffinate avvalendosi del feedback dato dall'utente privilegiando le canzoni o gli artisti che presentanto gli stessi attibuti di quelli oggetto di "like" da parte dell'utente e viceversa.

Una delle principali problematiche di questo approccio risiede nella capacità del sistema di applicare in modo corretto ed efficace le preferenze dell'utente su un dato contenuto ad altre tipologie di oggetti o contenuti. Quando il sistema è limitato a formulare reccomendation sullo stesso tipo di oggetti o contenuti di quelli già utilizzati dall'utente, il valore complessivo del sistema è inferiore al caso in cui è possibile suggerire altri tipi di contenuto o oggetto. Ad esempio, fornire reccomendation in termini di articoli di news sulla base dei dati raccolti navigando su siti di news è sicuramente utile

ma si otterrebbe un valore maggiore se il sistema offrisse reccomendation in termini di altre tipologie di contenuto come musica, video, discussioni, etc.

Diversamente dall'approccio collaborative filtering, non è necessario disporre inizialmente di una grande quantità di dati. Tuttavia tale approccio risulta abbastanza limitato nel perimetro di applicazione in quanto capace di fornire solo reccomendation simili al seme iniziale di attributi, caratteristiche e funzionalità.

La seguente espressione sintetizza in modo semplice la logica di base: "Se ti è piaciuto questo prodotto, potrebbero piacerti anche questi altri prodotti".

La tabella seguente offre un quadro sinottico dei due metodi descritti.

Tabella 7. Modelli di filtering

Approccio	Patrimonio informativo richiesto (volume e varietà di dati)	Perimetro di applicazione	Accuratezza reccomendation	Livello di complessità
Collaborative filtering	Elevato (problema cold start e sparsità)	Ampio	Alta	Alto
Content-based filtering	Ridotto	Ridotto	Bassa	Basso

Fonte: ITMedia consulting

Hybrid reccomender system

I sistemi ibridi combinano i due approcci suddetti per produrre risultati migliori in taluni casi. Tale approccio può essere implementato in diversi modi:

- calcolando separatamente le reccomendation utilizzando i due metodi precedenti e poi combinando i risultati;
- aggiungendo elementi dell'approccio content-based all'approccio collaborative filtering (e viceversa);
- unificando i due metodi in un unico modello.

Diversi studi hanno mostrato che le performance dei sistemi ibridi sono superiori in termini di accuratezza delle reccomendation rispetto a quelle determinate dall'applicazione di un solo modello. Inoltre, i sistemi ibridi consentono di mitigare alcune problematiche tipiche come il cold start o la sparsità.

La principale criticità nell'implementazione dei sistemi ibridi è data dal loro elevato livello di complessità che richiede adeguate risorse in termini di investimenti per integrazione nelle attività di business e ricerca e sviluppo.

Sulla base delle suddette considerazioni, emerge un quadro in cui è ragionevole concludere che le piene potenzialità di questi sistemi possano essere sfruttate da soggetti in possesso di:

- rilevante patrimonio informativo soprattutto in termini di volume di dati e in termini di varietà;
- cospicua e diversificata user base per contenere la problematica della sparsità;
- adeguate risorse di calcolo;
- adeguate risorse per sostenere gli investimenti necessari sia a livello tecnologico (data center, connettività, tecnologie cloud, etc.) che a livello integrazione, ricerca e sviluppo.

Case study: Netflix

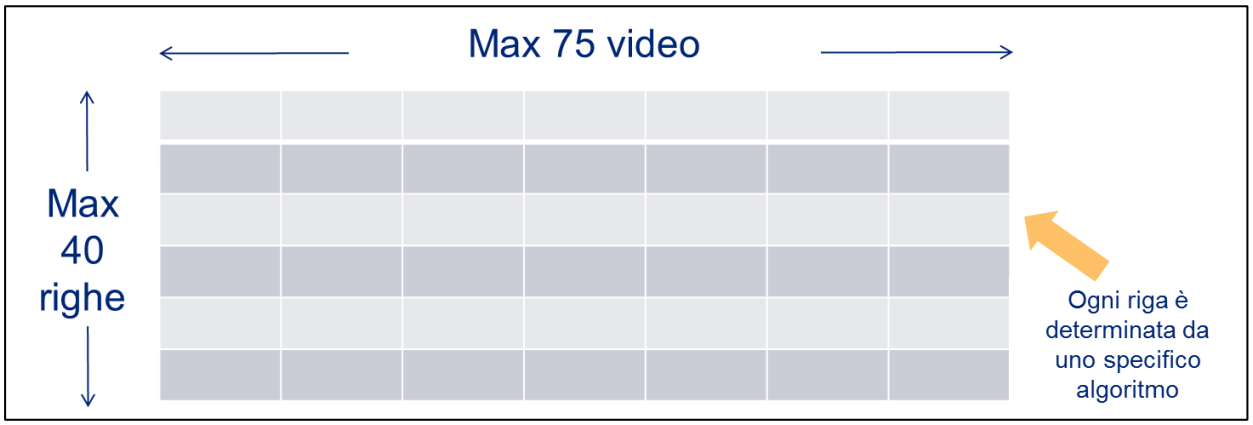
Netflix rappresenta uno dei più importanti casi di applicazioni di tali tecnologie e in particolare dei sistemi ibridi dato che determina le reccomendation sia comparando le abitudini di ricerca e visione tra utenti simili (come da approccio collaborative filtering) sia suggerendo film con caratteristiche simili a quelle che l'utente ha valutato positivamente (in base all'approccio content-based filtering).

Quale principale attore nello sviluppo e nella diffusione della Internet TV, Netflix si è adoperata fortemente nell'innovazione del proprio reccomender system strutturando e ottimizzando progressivamente tutta una serie di sofisticati algoritmi basati su tecniche di statistical learning e machine learning (in termini sia di approcci supervisionati come classificazione e regressione, sia di approcci non supervisionati come clustering). Un grande impulso alla ricerca in questo ambito è stato dato dalla competizione, denominata Netflix Prize, indetta dalla società nel 2006 mettendo in palio un premio di 1.000.000 $ al team capace di produrre reccomendation più accurate del 10% di quelle prodotte dal proprio sistema su un data set di oltre 100 milioni di voti di film. Il premio è stato assegnato il 21 settembre 2009 al team BellKor's Pragmatic Chaos; inoltre, numerosi altri team hanno sviluppato e proposto altri algoritmi che sono stati applicati successivamente in altri settori o mercati. Era in programma una seconda competizione che è stata annullata per via di un accertamento legale da parte della Federal Trade Commission.

Come accennato precedentemente, il reccomender system di Netflix si basa sulla combinazione di diversi algoritmi impiegati per rispondere adeguatamente a diverse esigenze contribuendo alla creazione della Netflix experience. Ricerche di mercato e sui consumatori hanno dimostrato che l'utente medio di Netflix perde interesse nei confronti di un contenuto dopo un intervallo di attesa per la scelta che va dai 60 ai 90 secondi, solitamente avendo dovuto scegliere tra 10 o 20 film su uno o due schermi. In questa situazione è auspicabile che l'utente trovi qualcosa di suo gradimento altrimenti il rischio che abbandoni il servizio cresce in modo sostanziale. In questa situazione il reccomender system deve produrre suggerimenti tali che sia altamente probabile, o meglio ancora certo, che l'utente trovi un contenuto di interesse. La natura stessa dello streaming dei contenuti abilita la raccolta di una grande quantità di dati che descrive esattamente cosa è visto da ogni utente e come viene visto (su che dispositivo, a che ora del giorno, in quale giorno della settimana, con che intensità, etc.); le informazioni raccolte riguardano anche la posizione all'interno del sito in cui il contenuto è stato scelto e anche gli altri titoli suggeriti ma non selezionati dall'utente.

Il reccomender system di Netflix capitalizza adeguatamente questo grande patrimonio di dati soprattutto nella costruzione dell'homepage personalizzata per utente. A prescindere dal dispositivo utilizzato, l'homepage è la parte di Netflix che contiene il maggior numero di numero di reccomendation dando luogo a oltre i due terzi dei minuti visti dall'utente. Il layout della pagina ha una struttura a matrice in cui ogni istanza presentata è una reccomendation e ogni riga contiene video dal "tema" analogo. Ogni riga presenta un'etichetta relativa al tema con l'obiettivo di rendere questo trasparente e più intuitivo per gli utenti. Di solito vi sono circa 40 righe nell'homepage e fino a 75 video per ciascuna riga (questi numeri possono variare sulla base delle caratteristiche hardware del dispositivo utilizzato con l'obiettivo di assicurare la migliore user experience possibile). I video di ciascuna riga sono generalmente determinati da uno specifico algoritmo.

Figura 72. Le righe dell'homepage di Netflix



Fonte: ITMedia consulting

Gli algoritmi utlizzati da Netflix sono di seguito brevemente descritti.

Personalized Video Ranker (PVR)

Questo algoritmo è utilizzato per le righe relative a specifici generi (ad esempio "Romantic Movies", come indicato nella schermata di sinistra nella figura). L'algoritmo ordina in modo personalizzato l'intero catalogo di video (o un suo sottoinsieme selezionato per genere) per ogni utente. Il risultato è dato da un elenco di video ordinati per un dato genere (e ciò spiega perché una riga relativa ad uno stesso genere può mostrare risultati completamente differenti per diversi utenti). I risultati possono essere migliorati combinando l'ordinamento personalizzato sulla base del comportamento dell'utente con una piccola componente legata alla popolarità dei film (utilizzata per determinare le reccomendation della riga "Popular").

Top-N video Ranker

Questo algoritmo produce le reccomendation per la colonna denominata "Top Picks". Il suo obiettivo consiste nel trovare un set limitato tra le migliori reccomendation personalizzate all'interno dell'intero catalogo scegliendo tramite l'utilizzo di specifiche metriche solo i video reputati "migliori" tra quelli calcolati con l'algoritmo PVR.

Figura 73. Personalized Video Ranker (PVR)

Fonte: Netflix, The Netflix recommender system: algorithms, business value, and innovation

Trending Now

Questo algoritmo, che determina la riga dal corrispondente nome, punta a sfruttare la rilevanza temporale di alcuni trend che a volte si hanno in intervalli limitati, da pochi minuti ad alcuni giorni, coniugandola con la personalizzazione. Tale algoritmo classifica adeguatamente due tipologie di trend: quelli che si ripetono periodicamente in base a tradizioni, festività, etc (ad esempio come nel caso di San Valentino); quelli relativi a situazioni o eventi momentanei o improvvisi di varia natura.

Continue watching

Tale algoritmo è utilizzato per contemplare le nuove abitudini di consumo dei contenuti, sia in modo seriale che in modo non episodico tramite la visione di piccoli spezzoni. L'ordinamento dei contenuti visti di recente si basa sulla stima di cosa l'utente intenda continuare a vedere, rivedere o non vedere più. La stima è effettuata tenendo conto di diversi parametri: il punto di abbandono (metà programma, inizio, fine), il dispositivo utilizzato, il tempo intercorso dall'ultima visione, etc.

Video – video similarity

Questo algoritmo è utilizzato per determinare le reccomendation della riga "Because you watched (BYW)". Si tratta di un algoritmo non personalizzato che calcola un elenco di video "gemelli" tra tutti quelli presenti nel catalogo. Si hanno così numerose righe di video "gemelli". La scelta della riga di video "gemelli" da presentare all'utente si basa, invece, su elementi di personalizzazione relativi a molteplici aspetti a partire dai video già visti.

Page generation

Date le numerosissime combinazioni di righe possibili per l'homepage, Netflix utiliza un algoritmo di **Page Generation** che utilizza l'output di tutti algoritmi descritti precedentemente. Per ogni utente tipo vi sono infatti migliaia di righe potenzialmente utilizzabili per l'homepage. Questo algoritmo matematico interamente personalizzato seleziona e ordina in modo ottimizzato le righe più rilevanti per l'utente. Non basandosi su template prefissati, l'algoritmo consente di costruire in modo

flessibile homepage senza una certa tipologia di riga (ad esempio BYW o Popular) per un dato utente, homepage con metà righe dello stesso tipo per un altro utente, etc.

Evidence selection

Si tratta di altri algoritmi orientati a determinare che tipologia di informazioni far vedere nell'area "evidence" posizionata nella parte superiore di ogni contenuto selezionato. Ad esempio, questi algoritmi decidono se mostrare che il film ha vinto un oscar o che il film è simile ad uno visto di recente dall'utente. Altre informazioni riguardano quale immagine mostrare, tra alcune disponibili, per supportare al meglio ogni reccomendation.

Search algorithms

Tutti gli algoritmi suddetti costituiscono il reccomender system di Netflix e influenzano l'80% di tutte le ore di contenuto mostrato sulla piattaforma. Il restante 20% deriva dalle ricerche effettuate dagli utenti. Per ottimizzare anche questo aspetto della user experience, si utilizzano specifici algoritmi basati su numerosi dati tra quelli relativi ai contenuti visti dall'utente e alle ricerche effettuate, che di fatto trattano la ricerca come un caso particolare di reccomendation. Come indicato in figura, ad esempio effettuando una ricerca con la parola "fren" un algoritmo restituisce contenuti che contengono tale parola nel titolo ("Frenemies" in figura); un altro algoritmo tenta di prevedere contenuti relativi ad un concetto desumibile dalla chiave di ricerca ("French movies" in questo caso).

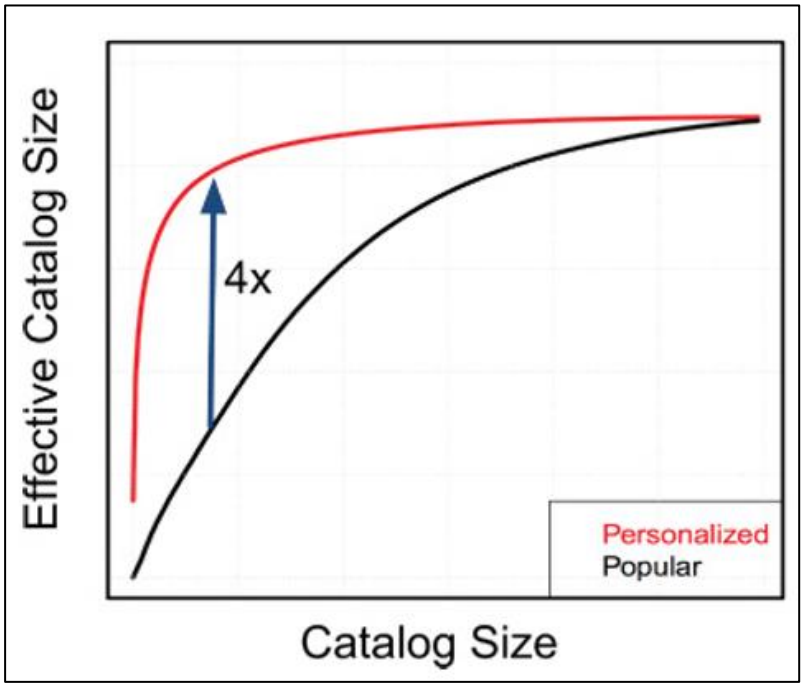
Figura 74. Search algorithms

Fonte: Netflix, The Netflix recommender system: Algorithms, business value, and innovation

Lo sviluppo e il miglioramento continuo del recommender system di Netflix hanno comportato e continuano ad apportare interessanti benefici in termini di valore di business. Tale sistema è infatti essenziale nell'affrontare in modo efficace il c.d. momento della verità ovvero quando un utente inizia una sessione e riesce a trovare qualcosa di interessante in pochi secondi risultando soddisfatto e appagato. Ciò previene efficacemente un'eventuale cancellazione dell'abbonamento. Inoltre, tramite la possibilità di fornire reccomendation mirate per ogni singolo utente, si ha la possibilità di sfruttare pienamente l'intero catalogo di contenuti anche in relazione a quei film di nicchia che quasi sicuramente sarebbero un insuccesso sulla TV broadcast a causa di un'audience troppo bassa tale da trainare adeguati ricavi pubblicitari che giustifichino l'occupazione dello slot temporale utilizzato. Il reccomender system permette quindi di distribuire il consumo di contenuti su un ventaglio molto

ampio di titoli come ben testimoniato dall'ECS (Effective Catalog Size) relativa alla dimensione effettiva del catalogo. Questa metrica descrive quanto diffusa è l'attività di consumo sull'intero catalogo. Se la maggior parte del consumo di contenuti deriva da un singolo video tale valore sarà molto vicino a 1; se invece tutti i video del catalogo generano la stessa quantità di visualizzazioni allora il valore dell'ECS sarà vicino a quello del numero di titoli del catalogo. La curva di colore nero in figura mostra l'andamento dell'ECS in relazione al numero dei video più popolari nel catalogo, spaziando, sull'asse delle ascisse, da 1 a N (con N pari al numero di questi video del catalogo). La curva di colore rosso mostra l'andamento dell'ECS per i primi k video determinati con l'algoritmo PVR. Si osserva che l'effetto della personalizzazione apportata dall'algoritmo PVR consente di ottenere un miglior sfruttamento dei set di video rispettivamente considerati con un picco pari a 4 volte il valore del caso senza il citato algoritmo.

Figura 75. Gli effetti della personalizzazione dell'algoritmo PVR



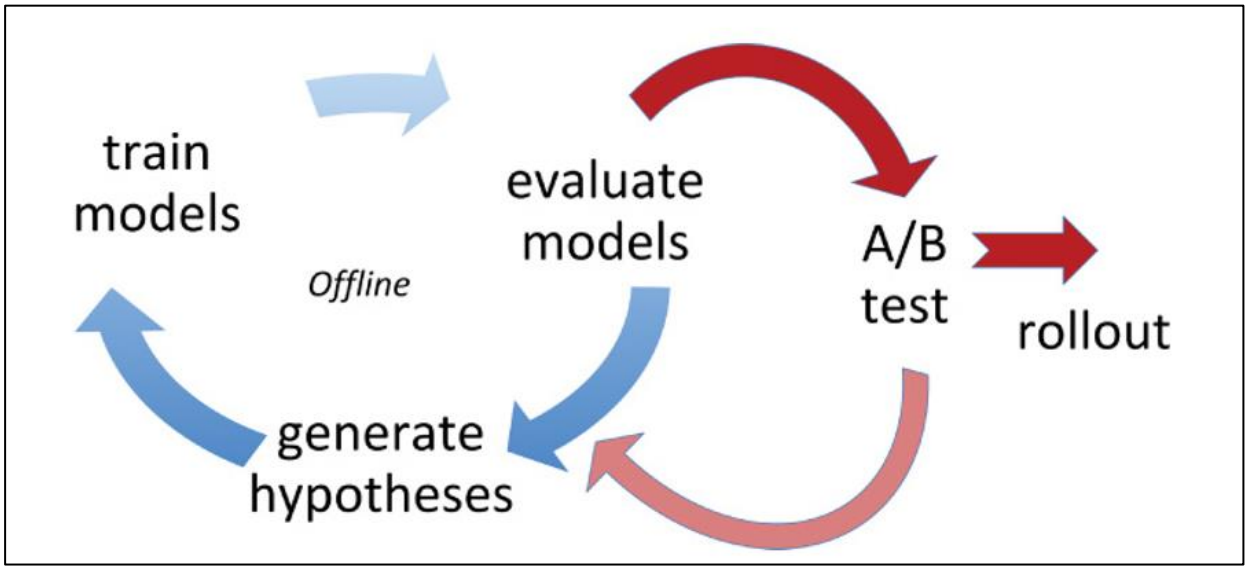
Fonte: Netflix, The Netflix recommender system: Algorithms, business value, and innovation

Grazie alla massimizzazione del catalogo e alla personalizzazione abilitata dalle reccomendation, Netflix è stata capace nel corso degli anni di abbassare il churn rate mensile di diversi punti percentuali giungendo ad un valore attuale molto basso ovvero a singola cifra bassa, dovuto per lo più a problematiche legate al sistema di pagamento piuttosto che alla scelta esplicita degli utenti di cancellare l'abbonamento. Ciò comporta un aumento complessivo del valore nel tempo degli abbonati e riduce il numero di nuovi abbonati che è necessario acquisire per compensare quelli che abbandonano il servizio. Tutto ciò produce, a detta della società, risparmi pari a circa 1 miliardi di dollari l'anno.

L'evoluzione del reccomender system di Netflix si basa su sforzi costanti di innovazione volti sia a ottimizzare gli algoritmi attualmente impiegati che a svilupparne di nuovi. In particolare, il processo

di ricerca e sviluppo è di tipo iterativo-incrementale ed è svolto secondo lo schema riportato in figura. Nella fase off-line, il nuovo algoritmo viene istruito utilizzando specifici train model come da paradigma Machine Learning basati sui dati storici solitamente impiegati per produrre reccomendation accurate. Non appena si iniziano ad intravedere risultati incoraggianti in queste sperimentazioni off-line si prepara un test A/B per il nuovo algoritmo. Il test A/B è sistematicamente utilizzato per valutare in modo casuale e controllato la prestazione ottenuta da ogni algoritmo in relazione all'engagement di medio termine degli utenti rispetto al tasso di cancellazione di abbonamenti. Se il risultato del test è positivo, il nuovo algoritmo entra a far parte, in aggiunta o eventualmente in sostituzione di qualche altro algoritmo, dell'insieme di algoritmi utilizzati dal reccomender system. Se il risultato del test è negativo o simile a quello ottenuto da altri algoritmi impiegati si hanno due possibilità: abbandonare le attività di ricerca e sviluppo su questo algoritmo o al contrario intensificarle per provare a ottenere in seguito un buon risultato nel test A/B.

Figura 76. Il processo di ricerca e sviluppo di Netflix di tipo iterativo-incrementale



Fonte: Netflix, The Netflix recommender system: Algorithms, business value, and innovation

Attualmente le attività di ricerca e sviluppo sono orientate a migliorare e ottimizzare le seguenti problematiche:

- Creazione dell'homepage;
- Scelta delle migliori evidence per supportare adeguatamente le reccomendation;
- Coldstart per i nuovi membri sui quali si hanno pochi dati storici limitando di fatto l'accuratezza delle reccomendation (l'approccio attuale consiste nel proporre a un nuovo utente durante la fase di registrazione una suvery durante la quale gli si chiede di selezionare dei video da un insieme creato automaticamente e utilizzato come input per tutti gli algoritmi);

- Creazione di reccomendation "globali" per supportare l'espansione internazionale di Netflix suggerendo il contenuto migliore in ogni paese in cui l'azienda opera tenendo conto dei vincoli di lingue e sottotitoli disponibili (dato che sarebbe troppo laborioso e costoso tradurre o produrre sottotitoli in tutte le lingue è molto utile sapere le lingue parlate o al limite comprese da ogni utente analizzando le informazioni dei contenuti visionati e delle modalità di visione: con sottotitoli, in lingua originale, etc.).

2.2.2.5 Visual Analytics & Data Visualization

Storicamente la tecnologia mirava a un unico scopo: migliorare la produttività. Gli utenti dovevano investire tempo per imparare ad usarla. Oggi invece i prodotti sono disegnati attorno alle azioni naturali dell'uomo: che sia basata su algoritmi automatici o sul design empatico, la *user experience* è fondamentale. E l'intelligenza artificiale interviene a supporto.

L'aspetto che più colpisce, quando si parla di analisi dei dati, è una sorta di polarizzazione uomo-macchina, che vede da una parte la necessità di "umanizzare" l'analisi, con strumenti sempre più visuali in grado di semplificare la rappresentazione di fenomeni altamente complessi, e dall'altra la forte spinta all'automazione, con le "macchine" in grado di imparare, replicando il processo del pensiero e del ragionamento dell'essere umano. Da una parte la *Visual Analytics*, quindi, dall'altra l'Apprendimento automatico, o *Machine Learning*.

Gli strumenti di analisi visuale vanno incontro alle limitate capacità cognitive dell'essere umano, non in grado di processare la miriade di dati generati in azienda, e, di norma, mettono a disposizione tre funzioni principali: visualizzazione dei dati in formato grafico, suggerimenti sulle azioni da intraprendere (*insight*), comprensione e condivisione delle informazioni. Possono essere usati, quindi, in diverse modalità: per l'esplorazione dei dati, per l'analisi dei dati e la creazione di report, per l'ottimizzazione dei processi o ancora per fare previsioni analitiche su trend futuri.

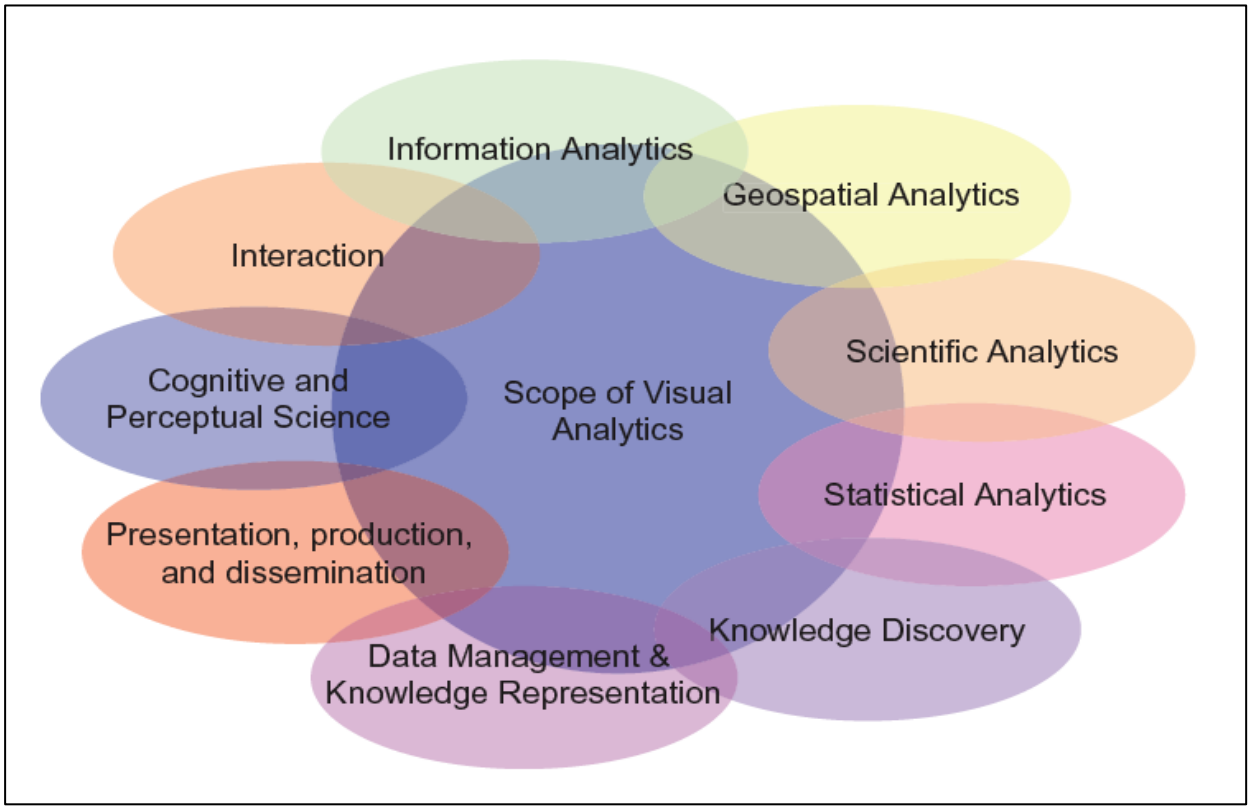
La *Data Visualization* definita "*crucial to unlock the value of data*[61]", rappresenta un importante segmento della "*Visual Analytics*", tecnica di analisi che mira a rendere i dati e l'elaborazione delle informazioni quanto più trasparenti possibili. La *Visual Analytics* ha la capacità di trasformare l'informazione in una opportunità. In particolare: *just as information visualisation has changed our view on databases, the goal of visual analytics is to make our way of processing data and information transparent for an analytic discourse. The visualisation of these processes will provide the means of examining the actual processes in-stead of just the results. Visual analytics will foster the constructive evaluation, correction and rapid improvement of our processes and models and ultimately the improvement of our knowledge and our decisions*[62].

[61] Keim D. A., *Data Drive Innovation open summit*, 2017.

[62] Keim D. A., Kohlhammer J., Ellis G., *Mastering the Information Age Solving Problems with Visual Analytics*, 2010.

La *Visual Analytics*[63] è la scienza del ragionamento analitico supportato da interfacce visive interattive. Oggi, i dati vengono generati ad una velocità incredibile e le capacità di raccoglierli e memorizzarli stanno aumentando ad un ritmo molto più veloce rispetto allo sviluppo delle capacità di analisi. Nel corso degli ultimi decenni, sono stati sviluppati un gran numero di metodi automatici di analisi dei dati. Tuttavia, la natura complessa di molti problemi rende indispensabile includere l'intelligenza umana in una fase iniziale del processo di analisi. Metodi di analisi visiva permettono ai decisori di unire la loro flessibilità, creatività, conoscenza e *background* con le enormi capacità di memoria e di elaborazione dei computer di oggi al fine della risoluzione di problemi piuttosto complessi. Utilizzando interfacce visuali avanzate, gli esseri umani possono interagire direttamente con le capacità di analisi dei dati dei computer, consentendo loro di prendere decisioni ben informate in situazioni complesse.

Figura 77. L'importanza della Visual Analytics



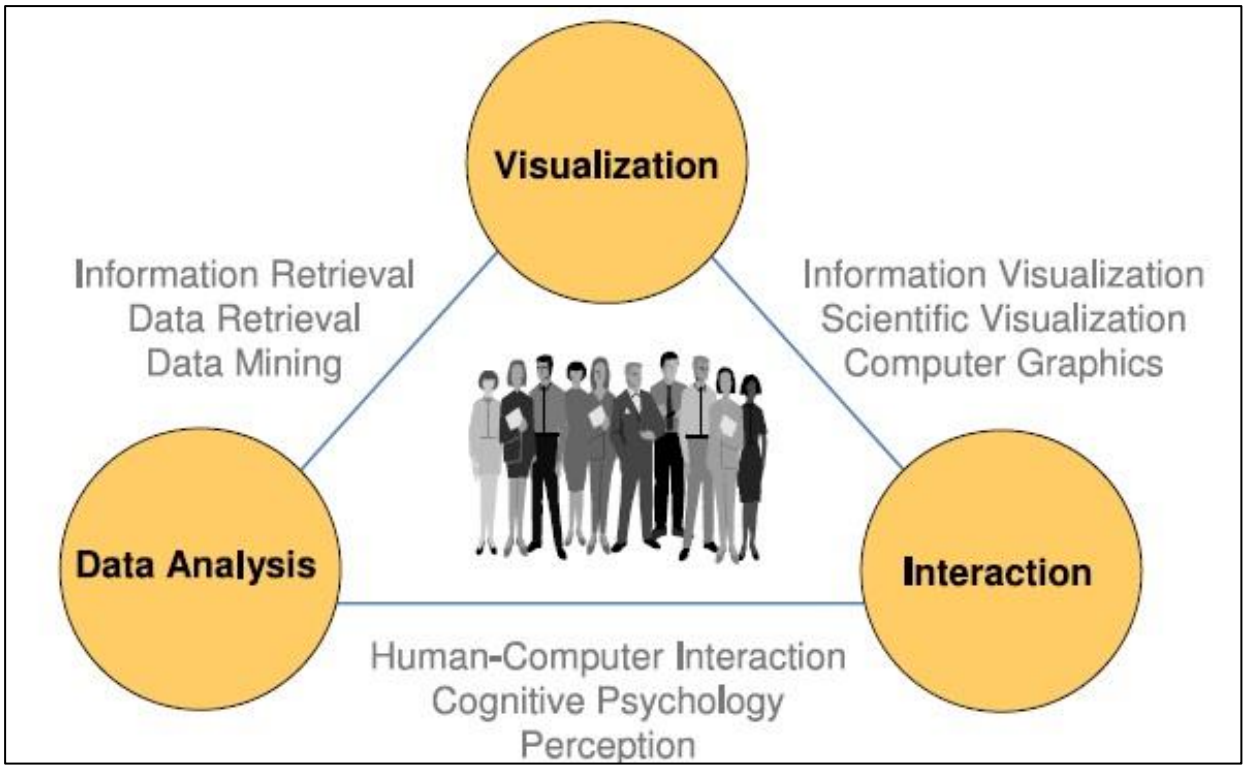
Fonte: Visual-analytics.eu

La *Visual Analytics* può essere vista come un approccio che combina visualizzazione integrale, fattori umani, e analisi dei dati. La figura seguente illustra le aree di ricerca relative alla *Visual Analytics*. Oltre la visualizzazione e l'analisi dei dati, in particolare i fattori umani, comprese le aree di cognizione e percezione, svolgono un ruolo fondamentale nella comunicazione tra l'uomo e il

[63] Thomas J., Cook K., *Illuminating the Path: Research and Development Agenda for Visual Analytics*, IEEE-Press, 2005.

computer, così come nel processo decisionale. Per quanto riguarda la visualizzazione, la *Visual Analytics* si riferisce alle aree di *Information Visualization* e *Computer Graphics*, mentre, per quanto riguarda l'analisi dei dati, si basa su metodologie sviluppate nei settori della *information retrieval*, *data management*, *knowledge representation* e *data mining*.

Figura 78. Le aree di ricerca della Visual Analytics



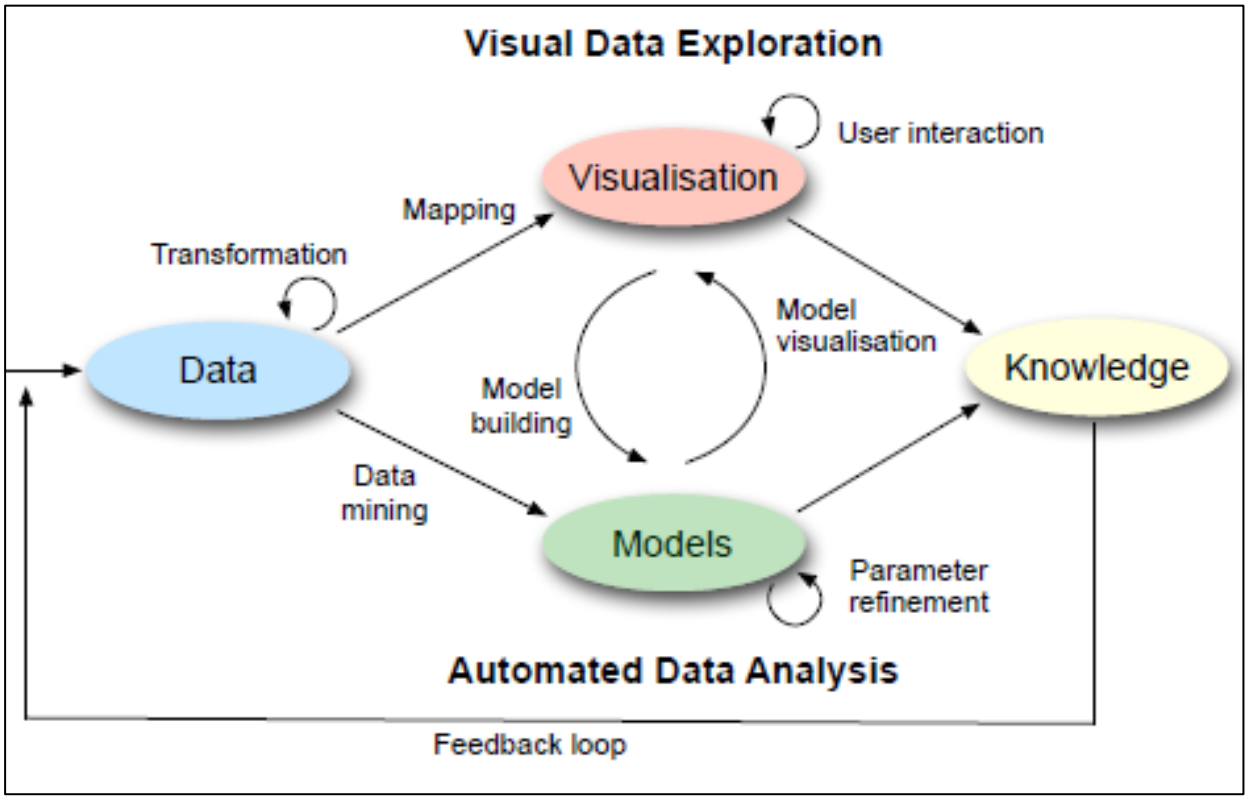
Fonte: visual-analytics.eu

In altre parole, la *visual analytics* crea gli strumenti e elabora le tecniche affinché le persone siano in grado di:

- sintetizzare le informazioni e di muoversi all'interno di una grande massa di dati ambigui, particolarmente dinamici e spesso confligenti;
- rilevare il risultato atteso e scoprire quello inaspettato;
- fornire valutazioni tempestive, difendibili, e comprensibili;
- comunicare queste valutazione ai fini della loro effettiva attuazione.

La figura seguente mostra una panoramica delle diverse fasi (rappresentati in ovali) e delle relative transizioni (frecce) del processo di *visual analytics* che, come detto, è caratterizzato dall'interazione tra dati, visualizzazioni, modelli sui dati, e utenti al fine della scoperta di conoscenza.

Figura 79. Il processo di Visual Analytics



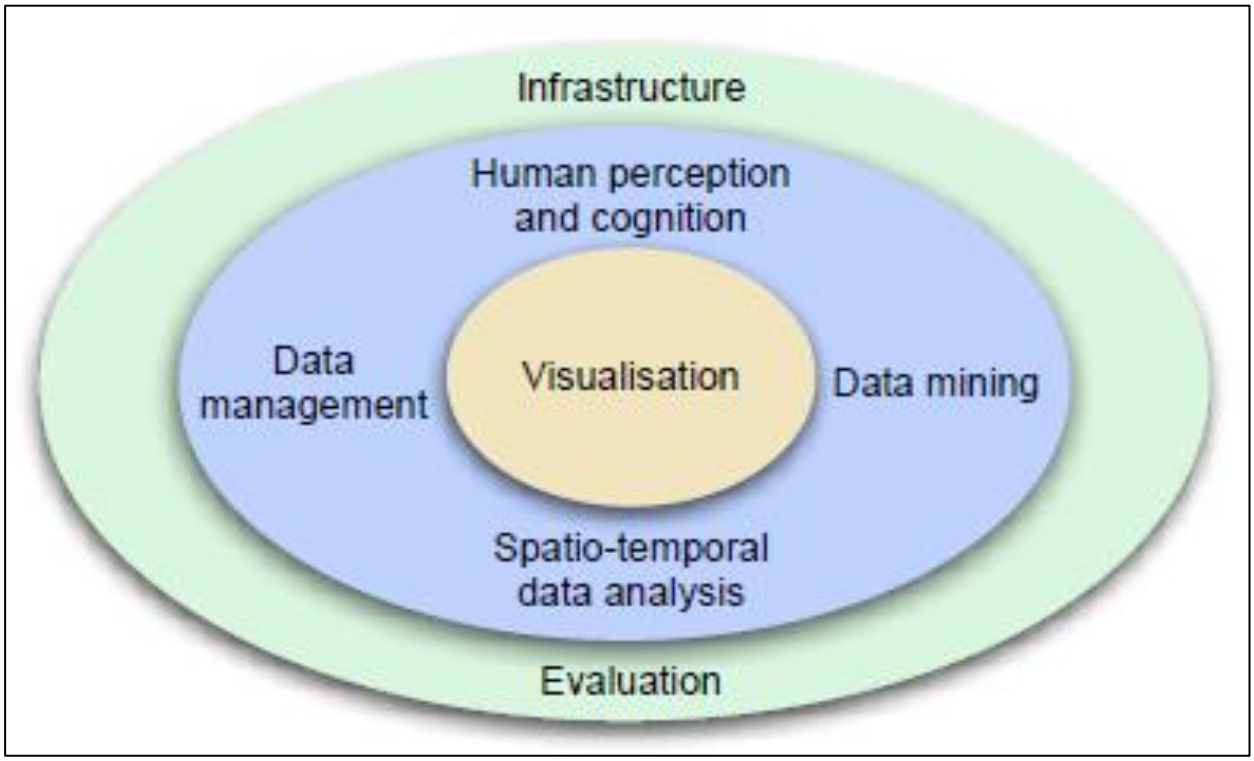
Fonte: visual-analytics.eu

In molti scenari applicativi, fonti di dati eterogenee devono essere integrate prima dell'applicazione dei metodi di *visual* o *automated analysis*. Pertanto, il primo step è spesso costituito dalla pre-elaborazione e trasformazione dei dati al fine di ottenere diverse rappresentazioni per ulteriori ricerche (come indicato dalla freccia *Transformation* in figura). Altri compiti tipici della fase di pre-elaborazione includono anche la pulizia dei dati, la normalizzazione, il raggruppamento, o l'integrazione di fonti eterogenee di dati. Dopo la trasformazione, si può scegliere se applicare metodi di *automatic* o di *visual analysis*. Nel primo caso, verranno applicati strumenti di *data mining* al fine di generare pattern di conoscenza. Una volta che un modello è stato creato, bisogna valutare se perfezionarlo o meno tramite un'integrazione di ulteriori dati. La fase di *Visualization* consente l'interazione tra l'utente e i metodi automatici al fine di modificare i parametri ovvero di selezionare un diverso algoritmo d'analisi. Dunque, il *visulatization model* può essere utilizzato per valutare i risultati generati dai modelli. L'alternanza tra metodi visivi e automatici è caratteristica dei processi di analisi visiva e porta ad un continuo perfezionamento e verifica dei risultati preliminari. In tal senso, risultati fuorvianti in uno step intermedio possono essere scoperti in fase iniziale, portando ad una miglioramento dei risultati e in generale ad una maggiore fiducia. Se la scoperta della *data visualization* dovesse essere eseguita per prima, l'utente dovrà confermare le ipotesi generate da un'analisi automatizzata.

La *Visual analytics* integra scienza e tecnologie di diverse discipline, come mostrato in figura 81. La *visualization* è il cuore del sistema, non rappresenta soltanto lo strumento per comunicare il valore

dei dati o i risultati dell'analisi, ma è anche sempre più utilizzata per monitorare i processi in altre discipline, come la *data management* e il *data mining*.

Figura 80. The disciplines that contribute towards visual analytics



Fonte: visual-analytics.eu

In particolare, la *visualization* è una nuova disciplina di ricerca sviluppatesi nel corso degli ultimi due decenni. Può essere sostanzialmente classificata in visualizzazione scientifica e informativa. La prima si occupa principalmente della visualizzazione tridimensionale di fenomeni in 3D, come flussi o strutture molecolari, dati 3D provenienti dal mondo dell'ingegneria, della biologia, della meteorologia, della cosmologia, e così via. La seconda, invece, è orientata allo sviluppo di metodi di *visualization* per dati astratti come dati di business, dati demografici, dati derivanti da social network e dati scientifici. Va poi sottolineato che oltre alla sempre crescente disponibilità di grandi volumi di dati, uno stesso dato può spesso comprende centinaia di dimensioni. Inoltre, in aggiunta alle classiche tipologie di dati di tipo numerico e testuale, alcune di queste dimensioni possono riferirsi a tipologie di dati particolarmente complesse, quali grafici, video, audio, oltre a quelle ancor più sofisticate definite per la web semantica. I valori dei dati difficilmente possono essere naturalmente mappati all'interno di uno spazio bi o tridimensionale, come accade per la *scientific visualisation* e per alcune tecniche di grafici standard come le rappresentazioni grafiche su assi x e y, grafici lineari che risultano inefficaci per la rappresentazioni di ampi dataset multidimensionali. Proprio per questo motivo sono stati sviluppati nuovi modelli di *visualization*, capaci di rappresentare anche dati non strutturati. Inoltre, alcune forme di *data visualization* utilizzano tecniche di analisi automatica dei

dati, come il *clustering* o la riduzione dimensionale, in fase di pre-elaborazione, prima della visualizzazione vera e propria.

Passiamo ora all'analisi delle altre componenti del sistema:

- *Data management*: la gestione efficiente dei dati di diverse tipologie e qualità rappresenta un fattore chiave della *visual analytics*, in quanto tipicamente individua la serie di dati da immettere nel sistema e che poi dovranno essere analizzati. Generalmente, una condizione necessaria per eseguire qualunque tipologia di analisi dei dati è la presenza di un database integrato e coerente. In particolare, trovare la giusta rappresentazione per diverse tipologie di dati (numerici, testuali, semi strutturati, segnali audio e video, rappresentazioni semantiche) rappresenta un problema chiave delle moderne tecnologie database. Ma la disponibilità di dati eterogenei richiede non solo la capacità di integrare diverse tipologie di dati in diversi formati, ma anche un'importante opera di pulizia dei dati stessi - come nel caso di valori di dati mancanti o imprecisi. Le moderne applicazioni necessitano, in tempo quasi reale, di questa integrazione tra dati e di renderla quanto più automatica possibile. Inoltre, le nuove fonti di informazione come lo streaming, le reti di sensori o l'estrazione automatica di informazioni provenienti da collezioni di documenti di grandi dimensioni (ad esempio, testi, HTML) pongono non pochi problemi di *data analysis*. Le tecniche di *data management* stanno sempre più incrementando gli usi di tecniche di analisi intelligenti e di visualizzazione ai fini dell'ottimizzazione dei processi e della piena informazione dell'utente.

- *Data mining*: come abbiamo visto in precedenza, la disciplina del *data mining* sviluppa metodi computazionali per estrarre automaticamente informazioni preziose da dati grezzi attraverso algoritmi di analisi automatica. Esistono diversi approcci: uno è l'apprendimento supervisionato in cui, sulla base di una serie di campioni di formazione, vengono utilizzati algoritmi deterministici o probabilistici per apprendere modelli di classificazione (o predittivi) di campioni di dati inediti. *Decision trees*, *support vector machines* e *neural networks* sono solo alcuni degli esempi di apprendimento supervisionato; un secondo approccio è quello dell'apprendimento non supervisionato, come l'analisi dei cluster, che sostanzialmente mira ad estrarre la struttura da dati senza che ci sia stata una preventiva conoscenza. Questo genere di soluzioni sono adottate per raggruppare automaticamente moli di dati in classi sulla base di una reciproca somiglianza, ovvero per identificare valori anomali in dati "rumorosi" durante la fase di pre-elaborazione. Altri approcci includono l'associazione di una regola estrazione, ovvero riduzioni di dimensionalità.

- *Spatio-temporal Data Analysis*: i dati spaziali sono dati che si riferiscono al mondo reale, come misurazioni geografiche, dati di posizione GPS e dati generati da applicazioni di telerilevamento. In sostanza, ci si riferisce a tutti quei dati che possono essere rappresentati su una mappa o su un grafico. Individuare relazioni spaziali e modelli all'interno di questi dati è di particolare interesse, ma richiede lo sviluppo di adeguate funzioni di gestione, rappresentazione e analisi. I dati temporali, d'altra parte, sono una funzione del tempo, vale a dire che il valore delle variabili dei dati possono cambiare nel tempo; qui, l'importanza dell'analisi è di identificare modelli, tendenze e correlazioni tra dati nel corso del tempo.

L'analisi di dati spazio-temporali ha aggiunto complessità di scala e incertezza. Ad esempio, è spesso necessario per scalare mappe che individuino modelli più ampi e anche per la localizzazione delle aree, e allo stesso modo per il tempo, l'analisi delle tendenze che si verifichino giornalmente o annualmente. In termini di incertezza, i dati spazio-temporali sono spesso incompleti, interpolati, collezionati in diversi intervalli di tempo o basati su presupposti diversi. Altre questioni relative ai dati spazio-temporali comprendono l'esistenza di relazioni topologiche complesse tra gli oggetti nello spazio, in genere dataset particolarmente ampi e la necessità di avere tipologie di dati specializzate. Inoltre, i dati geo-spaziali sono sempre più disponibili per non esperti che necessitano di strumenti adeguati per trarre vantaggio da questa ricca fonte di informazione.

- *Perception and Cognition*: la percezione e la cognizione rappresentano il lato umano della *visual analytics*. La percezione visiva è il processo di elaborazione delle informazioni provenienti dal mondo esterno attraverso gli occhi ed elaborate dal cervello che le traduce in informazioni più complesse, disponibili per le funzioni cognitive del soggetto. La cognizione è la capacità di capire le informazioni visive e di effettuare inferenze in gran parte basate sul primo apprendimento. La conoscenza del nostro modo di 'pensare visivamente' è molto importante nella progettazione delle interfacce utente e, insieme con l'esperienza pratica nel campo dell'interazione uomo-computer, contribuirà alla creazione di metodi e strumenti per la progettazione di *perception-driven*, tecniche di interazione multimodali per la visualizzazione e l'esplorazione di grandi spazi informativi, nonché per le valutazioni sull'utilizzabilità di questi sistemi.

- *Infrastructure*: l'infrastruttura si occupa di collegare insieme tutti i processi, le funzioni e i servizi richiesti dalle applicazioni analitiche visive al fine di consentire all'utente di esplorazione i dati visualizzati nella maniera più efficiente ed efficace. Quest'operazione di armonizzazione è tutt'altro che semplice posto che le infrastrutture software create dalle diverse tecnologie sono generalmente incompatibili. Per questo motivo, la maggior parte delle applicazioni di analisi visiva sono attualmente costruite su misura come applicazioni *stand-alone*, utilizzando, per esempio, archiviazioni dati in memoria, piuttosto che sistemi di gestione di database. La progettazione di infrastrutture di sistema e software è fondamentale per consentire alle applicazioni di utilizzare con successo le tecnologie più appropriate. Inoltre, il riutilizzo di molti componenti comuni si tradurrà in applicazioni che risulteranno maggiormente adattabili e che verranno costruite più velocemente.

- *Evaluation*: la valutazione è un'operazione molto difficile data la natura esplorativa della *visual analytics*, la portata dell'esperienza dell'utente, la diversità delle fonti dei dati e i compiti veri e propri. Nel campo della *data visualization*, solo di recente la valutazione è divenuta più rilevante. È stato riconosciuto che una comprensione generale delle tassonomie riguardanti i principali tipi di dati e delle attività sono particolarmente importanti per la formazione della *visual analytics research*.

In definitiva, la *visual analytics* sta sempre più diventando una parte essenziale del business. Le aziende e le imprese stanno utilizzando le tecnologie di *data visualization*, al fine di accelerare la loro

performance di business e migliorare i loro processi decisionali. Alcuni dei vantaggi chiave per le imprese riconducibili alla *visual analytics* sono:

- Il miglioramento nell'attività di esplorazione e di analisi di dati, minimizzazione dei costi complessivi;
- semplificazione della mole di informazioni complesse disponibili ai fini dell'ottimizzazione dei processi di *decision making*;
- velocizzazione dei processi di comprensione dei dati e conseguentemente di *decision making*;
- risoluzione dei principali problemi aziendali anche grazie ai significativi miglioramenti nel campo dell'informatica e della memorizzazione dei dati;
- risultati più accurati per decisioni più redditizie;
- identificazione delle diverse tendenze di visualizzazione per una maggiore comprensione delle modalità di presentazione dei dati.

Aziende leader a livello mondiale, università e agenzie governative stanno sempre più implementando e adottando questi strumenti. Dalle più grandi e innovative organizzazioni di tutto il mondo - Procter & Gamble, Apple, Pfizer, Microsoft, Coca Cola, Google, Cornell University, Progressive Insurance, Amazon, Georgetown University, il VA (Veteran's Administration), Blue Cross Blue Shield – al singolo consulente d'impresa, gli strumenti di *visual analytics* sono ormai *mainstream*. Grazie alla sua importanza pratica e ai costi di implementazioni efficaci più imprese sono orientate a mantenere la loro serie di dati in base agli standard di *visual analytics*.

PARTE II. Prospettive di policy

Introduzione

Nella prima parte di questo lavoro si è avuto modo di esaminare il fenomeno big data, distinguendo tra la creazione di voluminosi patrimoni di dati e la loro analisi, nonché descrivendo le differenti tecnologie che, come le piattaforme digitali e la rete delle cose, contribuiscono a generare i dati digitali o che, come il *cloud computing* e l'*intelligent automation*, rendono materialmente possibile la raccolta, gestione e analisi dei big data. Inoltre, nei precedenti capitoli si è apprezzato come le istituzioni europee e nazionali abbiano riconosciuto il significato strategico dei big data, promuovendo delle politiche che abilitino le imprese ad utilizzarli per estrarre conoscenza e ricchezza. In particolare, nelle ultime pagine della prima parte, descrivendo la catena del valore connessa ai big data, si è dato conto di due fenomeni tra loro non mutualmente escludenti, ma tra loro non necessariamente connessi, vale a dire: (a) il processo grazie al quale le imprese possono trasformare i dati in informazioni utili a migliorare i loro processi decisionali, ad individuare sistemi di produzione più efficienti, nonché a ideare nuovi prodotti e servizi da offrire ai consumatori; (b) il processo per effetto del quale le imprese possono utilizzare i dati che detengono per produrre informazione sotto forma di raccomandazioni, recensioni, e risultati della ricerca e così incidere sulla formazione delle opinioni dei consumatori circa prodotti, servizi e mercati. Complessivamente, dunque, nella prima parte di questo lavoro si è riconosciuto come i big data siano destinati non solo a permeare la quotidianità degli individui, ma anche a plasmare l'attività delle imprese, incidendo sulla loro capacità di competere con processi sempre più efficienti e beni sempre più originali e personalizzati, nonché attraverso la produzione di informazione.

Ora, proprio perché fonti di innovazioni[64] tali da incidere anche sulle più tradizionali leve concorrenziali a disposizione delle imprese, i big data non possono non diventare materia di interesse per il diritto antitrust. Questa seconda parte del report sarà dunque dedicata a verificare come le categorie del diritto della concorrenza europeo e nazionale, della regolamentazione e della privacy si intreccino con i big data.

In particolare, nel terzo capitolo, dopo aver brevemente illustrato perché il diritto antitrust si distingue dalla regolamentazione economica, dalla tutela dei consumatori e dalla disciplina delle pratiche commerciali scorrette (paragrafo 3.1) e dopo aver passato in rassegna la letteratura sul rapporto tra diritto antitrust e big data (paragrafo 3.2), si proporrà la visione secondo la quale il diritto antitrust pensa i big data quali input che meritano di essere analizzati caso per caso in rapporto all'uso che le singole imprese ne fanno (paragrafo 3.3). Successivamente, si discuteranno le seguenti questioni: (i) quale sia la relazione tra i big data e il potere di mercato (paragrafo 3.4); (ii) se

[64] Si parla infatti di *"data-driven innovations"* a cui l'OCSE ha dedicato un apposito report, *Data-Driven Innovation Big Data for Growth and Well-Being,* 2015, http://www.keepeek.com/Digital-Asset-Management/oecd/science-and-technology/data-driven-innovation_9789264229358-en#.WT-p3caxUdU#page1.

e quando le imprese possano incontrare delle difficoltà nel formare un patrimonio di big data (paragrafo 3.5); (iii) se e quando i big data possano essere utilizzati per tenere delle condotte imprenditoriali lesive della concorrenza nei mercati dei beni e servizi ideati grazie ai big data, nonché nei mercati dell'informazione (paragrafo 3.6); (iv) se e quando la disciplina a tutela dei dati personali e della privacy possa in qualche misura condizionare anche l'analisi antitrust (paragrafo 3.7); nonché e in ultimo (v) se esistano delle aree di sovrapposizione tra la gestione dei big data svolta con gli strumenti antitrust e gli strumenti della c.d. *consumer protection* (paragrafo 3.8).

Nel capitolo quarto, invece, si valuterà quanto e come gli strumenti della regolamentazione possano aiutare a gestire il fenomeno big data, eventualmente sopperendo alle lacune del diritto antitrust. In particolare, dopo aver osservato come i big data possano rientrare nell'ambito della regolazione delle comunicazioni (4.1) ed esaminato in che misura i big data potrebbero limitare il funzionamento del meccanismo di mercato, producendo esternalità positive e aumentando le asimmetrie informative (paragrafo 4.2), si discuteranno gli strumenti regolatori che possono aiutare a rispondere a questi problemi. Così si parlerà di come i consumatori possano essere istruiti e di come le autorità pubbliche possano contribuire ad aumentare la consapevolezza dei consumatori attraverso forme di diclosure regulation (paragrafo 4.3). Inoltre, proprio alla luce delle difficoltà che il diritto antitrust incontra nell'eliminazione degli switching cost e nella gestione sia dei prezzi personalizzati sia dei dati eventualmente ritenuti essenziali, si discuterà della data portability e degli obblighi di condivisione delle risorse (paragrafo 4.4).

In ultimo, nel capitolo quinto si discuteranno le novità che contraddistinguono la disciplina a tutela dei dati personali e della privacy, pensandole in rapporto allo sviluppo della data economy e all'avvento dei big data. Verranno spiegate nel dettaglio le differenze tra la tuttora vigente Direttiva e il Regolamento che diventerà efficace a partire dal gennaio 2018 (paragrafo 5.2). Per meglio comprendere le direzioni di *policy* intraprese a livello europeo, dapprima si offrirà una lettura dei principi che tradizionalmente regolano la *data privacy* attualizzata sulla base di una nuova ed inedita centralità dell'individuo, rilevabile sia al momento della scelta (consenso) di condivisione delle informazioni personali (paragrafo 5.3), sia in seguito nelle modalità di trattamento dei dati (paragrafo 5.5). Successivamente, il lavoro illustrerà il nuovo approccio - c.d. della *privacy by design*, anche detta *privacy by default* - alle vulnerabilità dei dati personali (paragrafo 5.4). Inoltre si discuterà di come, nel nuovo scenario delineato dal regolamento, il rafforzamento dell'*accountability* di chi compie il trattamento dei dati sia accompagnato da un parallelo *empowerment* del soggetto interessato da leggersi nell'ambito di quel più ampio panorama che scaturisce dall'intreccio tra economia dell'informazione *(information economy)* ed economia della condivisione *(sharing economy)* (paragrafo 5.6). Infine, si guarderà a come la nuova disciplina dei dati personali e della privacy sia sollecitata dal fenomeno big data (paragrafo 5.7). La seconda parte si chiuderà, infine, con un case study sulla responsabilità delle piattaforme per violazione del diritto d'autore.

Capitolo 3. Le questioni di diritto antitrust

3.1 Lo spirito della disciplina antitrust europea e nazionale (cenni)

Chi volesse usare una formula semplice e sintetica per individuare il bene giuridico protetto dal diritto della concorrenza potrebbe affermare che le norme antitrust europee e nazionali – vale a dire gli artt. 101 e 102 del Trattato sul Funzionamento dell'Unione Europea (TFUE), il Regolamento 139/2004 e la Legge 287/90 – vogliono tutelare il corretto funzionamento del mercato per come esso si desume dal variare del benessere dei consumatori. Il diritto antitrust vieta infatti alcuni specifici comportamenti delle imprese – ossia le intese anticompetitive, gli abusi di posizione dominante e le concentrazioni lesive della concorrenza – nella misura in cui essi determinano nel breve periodo una riduzione dell'output disponibile sul mercato e il conseguente aumento dei prezzi di mercato e/o nel medio-lungo periodo una riduzione della qualità dei beni offerti, della loro varietà e/o del loro grado di innovatività.

Secondo la visione dominante, il diritto antitrust, dunque, non è volto a proteggere in via diretta ed immediata una particolare classe di soggetti, quali potrebbero essere i consumatori o le piccole-medie imprese. In linea di massima il diritto antitrust non può agire affinché questi soggetti siano bene informati e compiano scelte consapevoli, a meno che l'informazione non venga pensata alla stregua di un prodotto capace di individuare un mercato – *rectius*, tanti mercati per quanti tipi di informazioni esistono – e quindi il diritto antitrust possa esplorare, e se del caso punire, quelle condotte imprenditoriali che riducono la quantità e la qualità dell'informazione disponibile[65]. Né, fatta eccezione per quei casi particolarissimi riconducibili alla categoria dei c.d. abusi di sfruttamento (cfr. paragrafi 3.5 e 3.7), il diritto antitrust può proteggere consumatori e (piccole-medie) imprese che si trovino in posizioni di debolezza contrattuale, come confermato dalla Corte di Giustizia nella recentissima sentenza *Intel*[66].

In termini generali, infatti, per rispondere a queste esigenze di trasparenza ed equità si è soliti invocare altre forme di intervento dei pubblici poteri nell'economia. Ad esempio, si fa appello alla c.d. *disclosure regulation* che mira a ridurre le asimmetrie informative presenti nel mercato e le situazioni di minore potere contrattuale nelle quali possono versare i consumatori; si richiede l'intervento della disciplina a tutela dei consumatori che vieta comportamenti ingannevoli, aggressivi o comunque tali da essere contrari alla diligenza professionale, nonché idonei a falsare, in misura apprezzabile, il comportamento economico del consumatore medio (cfr. paragrafo 3.8); o ancora si invocano i divieti di concorrenza sleale, di abuso di dipendenza economica e pratiche commerciali scorrette.

[65] Assai chiaro in tal senso Patterson M., *Antitrust Law in the New Economy*, 2017. In particolare, l'Autore sembra voler considerare il caso dei mercati dell'informazione diffusa per il tramite di siti di review, come TripAdvisor e Yelp, o per il tramite dei motori di ricerca, come Google e Bing.

[66] Corte di giustizia, Grande sezione, 6 settembre 2017, causa C-413/14 P, Intel Corp. c. Commissione Europea.

Analogamente, il diritto antitrust non vuole intervenire sulla struttura dei mercati per agire su alcune caratteristiche degli stessi, come potrebbero essere le economie di scala e scopo, gli effetti di rete o i c.d. *winner-takes-all effects*. Anche nell'industria digitale può accadere che un'impresa riesca a dominare un mercato perché ha raggiunto, pur conquistando una massa di consumatori relativamente contenuta, un *tipping point*, ossia un punto oltre il quale l'impresa si trova ad essere preferita dai consumatori più per gli esistenti effetti di rete che per meriti. Tuttavia, finché l'impresa ha conquistato la sua massa "iniziale" di consumatori senza tenere pratiche tali da ledere il *consumer welfare*, il diritto antitrust non può nulla contro questo fenomeno: non può, cioè, annullare gli effetti di rete o l'esistenza di punti di *tipping*. Per gestire e disciplinare tali profili ed eventualmente scongiurare l'ipotesi che alcuni mercati vengano monopolizzati, occorre eventualmente rivolgersi alla regolamentazione economica che, in una prospettiva anche generalista e aprioristica, può anche giungere a disegnare lo sviluppo di uno o più settori industriali.

Analogamente, il diritto antitrust non può disciplinare in via universale e aprioristica il tema dell'accesso ai big data. Laddove si volesse operare in tal senso, occorrerebbe adottare una specifica legislazione sull'accesso, sulla falsa riga di quella approvata con riguardo all'informazione detenuta dal settore pubblico. Inoltre, il legislatore potrebbe anche promuovere l'accesso attraverso deroghe e limitazioni intese come parte di una legislazione globale sulla titolarità dei dati. Ad esempio, lo Stato può promuovere l'accesso come parte della propria politica di finanziamento, ovvero quando concede delle sovvenzioni. Tuttavia, le caratteristiche dinamiche della *data economy* favoriscono piuttosto regimi di accesso autonomo che meglio potrebbero adattarsi allo sviluppo dinamico della *data economy*, la quale probabilmente solo gradualmente sarà in grado di suggerire al legislatore gli ostacoli all'accesso nel corso degli sviluppi dei nuovi modelli di business.

Il diritto antitrust può "soltanto" agire affinché le imprese non si comportino onde impedire ai mercati nei quali operano di funzionare come meccanismi di selezione, ossia come strumenti che, tramite le scelte dei consumatori, premiano il merito imprenditoriale a discapito delle imprese inefficienti e obsolete. Come si vede, dunque, anche il diritto antitrust tutela gli interessi della collettività tutta, dei consumatori e delle piccole-medie imprese; semplicemente, persegue questo obiettivo in via mediata, ossia individuando il proprio bersaglio in quei comportamenti delle imprese che – lo si ripete – riducono il *consumer welfare*.

Ma se l'analisi antitrust si concentra sui comportamenti delle imprese, sono due le vie che si devono percorrere per analizzare il rapporto tra il diritto antitrust e i big data. Dapprima, muovendo da quanto spiegato nella prima parte del nostro report, bisogna comprendere come i big data incidano sull'attività di impresa, ossia sulla produzione di beni, servizi, ed eventualmente informazione (cfr. paragrafo 3.3). Successivamente, occorrerà guardare a se e come le imprese possano utilizzare i big data in modo anticompetitivo, ossia per ledere il benessere dei consumatori (cfr. paragrafi 3.5-3.6). E ciò tenuto conto che, per chi è solito occuparsi di diritto antitrust, le condotte anticompetitive sono condotte illecite.

3.2 Lo stato dell'arte

Come spesso accade quando un fenomeno rivoluzionario come l'avvento dei big data incide in modo pervasivo sui comportamenti umani, il dibattito che vuole appuntarsi su quel fenomeno è animato da molteplici e diversi spunti, non tutti di natura economica o giuridica. Giusto a titolo di esempio, si consideri che in ambito epistemologico si contesta l'idea che la mera analisi dei dati possa effettivamente condurre a una forma di conoscenza vera e oggettiva. Infatti, a fronte di coloro che hanno enfaticamente[67] salutato l'avvento dei big data parlando di «morte della teoria»[68], si contano coloro che associano a loro i vizi tipici dell'induttivismo, ricordando come anche l'elaborazione dei dati richieda l'adozione di scelte soggettive, tutte caratterizzate da un certo margine di errore[69].

Chi poi si occupa di libertà di informazione affronta l'eventualità che forme più o meno evidenti di manipolazione dei dati possano sottrarre agli esseri umani la possibilità di formarsi opinioni indipendenti circa i fatti del mondo[70]. E giacché questo dubbio sorge anche quando quegli individui considerati sono "meri consumatori" interessati a formarsi un'idea autonoma riguardo a questioni di rilievo puramente commerciale, si dubita della loro sovranità. In altri termini, ci si chiede quanto le preferenze che i consumatori si formano sulla scorta dell'informazione raccolta in internet siano effettivamente rappresentative del merito delle imprese nei fatti scelte dai consumatori o se non siano appunto il frutto di manipolazioni. Questione, quest'ultima, che recentemente sta acquistando

[67] Per alcuni la tesi circa la morte della teoria sarebbe volutamente iperbolica e provocatoria: Bollier D., *"The promise and peril of big data"*, 2010, https://www.emc.com/collateral/analyst-reports/10334-ar-promise-peril-of-big-data.pdf.

[68] Anderson C., *The end of theory: The data deluge makes the scientific method obsolete*, Wired, Giugno 2008, https://web.archive.org/web/20150417230631/http://archive.wired.com/science/discoveries/magazine/1607/pb_theory, dove si legge «[t]his is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves»; Steadman I., *Big data and the death of the theorist*, Wired, 2013, http://www.wired.co.uk/article/big-data-end-of-theory; e Williamson B., *The death of the theorist and the emergence of data and algorithms in digital social research*, 2014, https://web.archive.org/web/20150322073355/http://blogs.lse.ac.uk/impactofsocialsciences/2014/02/10/the-death-of-the-theoristin-digital-social-research/

[69] Si considerino, ad esempio, Boyd D., Crawford K., *Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon*, in Information, Communication & Society 2012, p. 662; Timmer J., *Why Cloud cannot Obscure the Scientific Method?*, Ars Technica 2008; e Kitchin R., *Big Data, new epistemologies and paradigm shifts*, Big Data & Society 2014, p. 1.

[70] A titolo di esempio, si consideri Allcott H. and Gentzkow M., *Social Media and Fake News in the 2016 Election*, Journal of Economic Perspectives 2017, p. 211 dove molti altri riferimenti alla letteratura in tema di post-truth e fake news. Con riguardo al nostro paese si segnale un disegno di legge presentato lo scorso febbraio al Senato e il cui esame non è ancora iniziato. Si tratta dell'atto S2688 *Disposizioni per prevenire la manipolazione dell'informazione online, garantire la trasparenza sul web e incentivare l'alfabetizzazione mediatica*.

sempre più rilievo anche in ambito antitrust, giacché la concezione dei mercati come strumenti di selezione si poggia proprio sull'attendibilità delle scelte dei consumatori.

L'interesse per la libertà di autodeterminazione degli individui anima poi anche il dibattito circa il c.d. determinismo dei dati, ossia circa la possibilità che la clusterizzazione degli individui operata sulla scorta dei big data da istituzioni e imprese finisca per imprigionare le persone in categorie che, per di più, potrebbero non rappresentarli fedelmente. In particolare, questa preoccupazione ha già formato oggetto di riflessione negli Stati Uniti dove si teme che l'uso dei big data possa anche perpetrare forme di discriminazione raziale, religiosa, politica e altro a svantaggio delle c.d. categorie protette tutelate dalle leggi sui diritti civili[71].

Va poi ricordato come due ulteriori fenomeni legati al mondo dei dati, quali la *data-driven innovation* e l'*open data*, avessero già attirato l'attenzione di istituzioni e studiosi a livello sovra-nazionale. Infatti, in un lungo scritto dell'OECD sono stati illustrati i nuovi processi e i nuovi prodotti che la conoscenza alimentata dai dati potrebbe ingenerare. Nel documento, in particolare, si evidenzia come *«[d]eclining costs along the data value chain . . . have been a significant driver of the increasing generation and use of data, as well as the accelerated migration of socioeconomic activities to the Internet thanks to the wide adoption of e-services in an increasingly participative web. The resulting phenomenon – commonly referred to as "big data" – signals the shift towards a data-driven economy, in which data enhance economic competitiveness and drive innovation and equitable and sustainable development»*. Analogamente, la ricca letteratura in tema di *open data*[72] e *public sector information*[73] ha sottolineato, rispondendo anche a delle istanze di politica industriale, la necessità che almeno le pubbliche amministrazioni, che negli anni hanno raccolto dati anche personali in ottemperanza ad obblighi di legge, permettano alle imprese di riutilizzare i loro dati per sviluppare nuovi prodotti a base informativa. In questo senso si devono ad esempio ricordare le direttive europee in materia di Public Sector Information.

Ebbene, è proprio in questo contesto così poliedrico che si colloca il dibattito circa il rapporto tra big data e diritto della concorrenza. Nonostante molti scritti possano dirsi di "ampio respiro", ossia

[71] White House, *Big data and differential pricing*, 2015, https://www.whitehouse.gov/sites/default/files/docs/Big_Data_Report_Nonembargo_v2.pdf, nonché FTC, *Big Data. A Tool for Inclusion or Exclusion?*, 2016, https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf.

[72] Vedi tra gli altri, Roman, Charles, *Open Data*, ConLawNOW 19, 2016; Cohn B. L., *Data Governance: A Quality Imperative in the Era of Big Data, Open Data and Beyond***,** ISJLP 811 (2014-2015); Gurin J., *Big Data and Open Data: How Open Will the Future Be***,** ISJLP 691 (2014-2015); Edward S., *Reflections on the Concept of Open Data***,** SCRIPTed 154, 2015; Leucci S., *Preliminary Notes on Open Data Licensing***,** J. Open Access L. 1, 2014; nonché con riguardo al contesto nazionale Pagnatelli V.*,Access, Accessibility, Open Data. The Italian Model of Public Open Data in the European Context*, Giornale di Storia Costituzionale 205, 2016.

[73] Vedi Lundqvist B., Forsberg Y., De Vries M., Maggiolino M.*, Open Data and Competition Law: Some Issues regarding Access and Pricing of Raw Data***,** Masaryk U. J.L. & Tech. 95, 2015 e Maggiolino M., *Il riutilizzo dell'informazione detenuta dal settore pubblico: alcune riflessioni di politica e diritto della concorrenza*, Concorrenza e mercato, 2012, pp. 765 – 802.

finalizzati a offrire un quadro completo, ancorché conciso, dei tanti punti di contatto tra i big data e il diritto antitrust, si possono comunque rintracciare degli ordini di problemi che più spesso occupano le pagine delle riviste giuridiche. In molti si sono interrogati circa i fattori – dai comportamenti strategici degli incumbent alle barriere strutturali e legali caratteristiche di ogni mercato – che potrebbero impedire alle imprese di accumulare big data[74]. Altri – non molti, invero – hanno invece valutato come le imprese potrebbero impiegare i big data per alterare il funzionamento dei mercati in cui operano, se del caso manipolando l'informazione che raggiunge i consumatori, rifiutando di condividere dati essenziali, o ancora usando in modo anticompetitivo la discriminazione di prezzo[75]. A latere della discussione su questi due temi, si collocano poi coloro che hanno intrecciato l'analisi del fenomeno big data con il dibattito circa l'operare delle piattaforme digitali[76] e l'analisi delle questioni di tutela della privacy[77] a vario titolo interessate dai big data.

Più nel dettaglio, vi è chi discute dei comportamenti imprenditoriali che una o più imprese possono tenere per impedire ai concorrenti di accedere ai dati digitali. Ad esempio, alcuni contestano la pratica delle piattaforme digitali di offrire i propri servizi a prezzo zero. Questo prezzo nullo impedirebbe infatti alle imprese concorrenti non solo di entrare nel mercato di quei servizi, ma anche di entrarvi per poi collezionare dati digitali, giacché inizialmente tali concorrenti non riuscirebbero a concentrare da un lato della loro piattaforma un numero sufficiente di inserzionisti pubblicitari per finanziare l'offerta ai consumatori di servizi a prezzo nullo[78]. Altri, invece, evidenziano come si possano configurare delle situazioni in cui le imprese, compresa l'importanza strategica dei dati,

[74] Cfr. *infra*.

[75] Cfr. *infra*.

[76] Vedi Rochet J.C. and Tirole J., *Platform Competition in Two-Sided Markets*, 2003 e *Two-sided markets: a progress report*, 2006; Einav L., Levin J. D., *The data revolution and economic analysis*, 2013; Lerner A. V., *The Role of 'Big Data' in Online Platform Competition*, 2014; Tucker D. S. and Wellford Hill B., *Big Mistakes Regarding Big Data*, 2014; Graef I., *Market Definition and Market Power in Data: The Case of Online Platforms*, 2015; Evans DS and Schmalensee R., *The Antitrust Analysis of MultiSided Platform Businesses*, in RD Blair e DD Sokol, Oxford Handbook on International Antitrust Economics, 2015; Evans DS, *Multisided platforms, dynamic competition, and the assessment of market power for internet-based firms*, 2016; Peng P., *Implications of Big Data for Competition Policy*, 2016.

[77] Vedi Tucker C. E., *Social Networks, Personalized Advertising, and Privacy Controls*, 2011; Goldfarb A. and Tucker C. E., *Privacy Regulation and Online Advertising*, 2011; Harbour and T. I. Koslov, *Privacy in Antitrust: A Relevant Product Market Factor*, 2011 P. J.; Cooper J. C., *Privacy and antitrust: underpants gnomes, the first amendment, and subjectivity*, 2013; Manne G. A. and Ben Sperry R., *The law and economics of data and privacy in antitrust analysis*, 2014 and *The problems and perils of bootstrapping privacy and data into an antitrust framework*, 2015; Acquisti A., Taylor C. and Wagman L., *The Economics of Privacy*, 2016; Deutscher E., *The role of personal data and privacy in merger analysis – a critical reassessment of the EU Commission's Facebook/WhatsApp decision*, 2017.

[78] Vedi Evans DS, *The Antitrust Economics of Free*,2011; Vaidhyanathan S.*,The googlization of everything (And why we should worry)*,2011; Newman N., *Search, Antitrust, and the Economics of the Control of User Data, 2014; Newman J.M., Antitrust in Zero-Price Markets: Foundations*, 2015; Gal M. S. and Rubinfeld D. L., *The Hidden: Costs of Free Goods Antitrust Implication*, 2015; *Access Barriers To Big Data*, 2016.

agiscano intenzionalmente per controllarli. In particolare, le piattaforme digitali avrebbero la capacità di limitare l'accesso ai dati da parte dei concorrenti, impedendo forme di condivisione dei dati, nonché ostacolando le politiche di portabilità[79].

In secondo luogo, si discute delle barriere alla raccolta dei dati oscillando tra due posizioni estreme. Per un verso, vi è chi ricorda come la non-rivalità e l'ubiquità dei dati li renda facili da reperire, a costi di produzione e distribuzione particolarmente contenuti, certo più bassi di quelli esistenti in passato[80]. Al riguardo – si argomenta – basterebbe pensare a come gli utenti di internet e degli smartphone lascino le proprie "tracce digitali" alla mercé di tutte le piattaforme con le quali contemporaneamente interagiscono (c.d. fenomeno del *multi-homing*) per avere una idea di quanto sia semplice e rapido raccogliere dati digitali. Di conseguenza, questa parte della letteratura sottolinea come i dati, di per sé considerati, non posseggano un grande valore e come, anche quando raccolti in grandi volumi, essi siano facilmente reperibili. Nel dettaglio, poi, questa dottrina osserva come anche un'impresa che entrasse in un mercato nel quale è già presente un *incumbent* ricco di dati potrebbe in breve tempo competere ad armi pari, giacché molti business digitali non hanno bisogno di dati storici e poiché, superato un certo ammontare di dati, i dataset si caratterizzano per economie di scala decrescenti[81].

Per altro verso, vi è chi invece sostiene come, per lo meno con riguardo all'operatività delle piattaforme digitali, i dati siano diventati una risorsa imprescindibile, ossia una risorsa senza la quale i modelli di business di queste piattaforme non potrebbero funzionare[82]. Per questo, alcuni autori

[79] Vedi Newman N., *Data Justice- Taking on Big Data as an Economic Justice Issue*, 2015.

[80] Vedi Thierer A., *Relax and Learn to Love Big Data*, 2013; Lerner A. V., *The Role of 'Big Data' in Online Platform Competition*, 2014; Lambrecht A. and Tucker C. E., *Can Big Data Protect a Firm from Competition?*, 2015, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705530.

[81] Lerner A. V., *The Role of "Big Data" in Online Platform Competition*, 2014.

[82] Vedi Evans D.S., *The antitrust economics of two-sided markets*, 2002; Armstrong M., Competition in two-sided markets, 2006, Rand Journal of Economics Vol.37(3): 668-691; Armstrong M. and Wright J., *Two-sided Markets, Competitive Bottlenecks and Exclusive Contracts*, 2006; Kaiser U. and Wright J., *Price structure in two-sided markets: Evidence from the magazine industry*, 2006; Chakravorti S. and Roson R., *Platform Competition in Two-Sided Markets: The Case of Payment Networks*, 2006; Eisenmann T., Parker G., and Alstyne M.W.V, *Strategies for Two-Sided Markets*, 2006; Haigiu A., *Pricing and commitment by two-sided platforms*, 2006, Rand Journal of Economics, Vol.37(3): 720-737; Argentesi E., Filistrucchi L*., Estimating Market Power in a Two-Sided Market: the Case of Daily Newspapers in Italy,* 2007, Journal of Applied Econometrics , Vol. 22(7), pp.1247-1266; Evans D., Schmalensee R., *Markets with Two-sided Platforms*, 2008, Issues in Competition Law and Policy, Vol. 1:667-693; Filistrucchi L*., How many markets are two-sided?*, 2010, Competition Policy International-Antitrust Chronicle, Vol.7; Evans D., *Platform Economics: Essays on Multi-Sided Businesses*, 2011; Haigiu A., Spulber D., *First-party content and coordination in two-sided markets*, 2013, Management Science, Vol.59(4) 933-949; King S.P., *Two-Sided Markets*, 2013; Filistrucchi L., Geradin D., Damme E.V., Affeldt P., *Market definition in two-sided markets: theory and practice, 2014; Newman N., Feature: Search, Antitrust, and the Economics of the Control of User Data*, 2014; Haigiu A, Strategic decisions for multisided platforms, 2014, MIT Sloan Management Review; Graef I., *Market Definition and Market Power in Data: The Case of Online Platforms*, 2015; Grunes A.P. e Stucke M.E., *No Mistake About It: The Important Role of Antitrust in the Era of Big Data*, 2015; Jullien B., *Competition policy for on-line platforms, intervento al Convegno On Line Intermediation and Competition in E-Commerce*, Università Roma Tre, 5 novembre 2015; Autorité de la

osservano come, sebbene allo stato le piattaforme non abbiano accordi, né sistemi di incentivi, che inducono gli utenti a fornire loro i dati in esclusiva, tali imprese siano oggi più che mai orientate verso strategie basate sui dati ai fini dell'ottenimento dell'efficienza operativa, nonché ai fini del mantenimento di un vantaggio competitivo[83]. E, proprio proseguendo lungo questa china, questi ed altri autori evidenziano come almeno alcuni tipi ed insiemi di dati raccolti dalle piattaforme digitali possano diventare risorse inimitabili e comunque inaccessibili a molti[84].

Rispetto a questi due poli del dibattito la dottrina più sofisticata ha compreso la necessità di procedere caso per caso, distinguendo le ipotesi nelle quali i dati sono disponibili e liberamente appropriabili da quelle in cui i dati raccolti da alcune imprese sono non duplicabili e comunque riconoscendo come la raccolta dei dati digitali richieda infrastrutture e competenze costose[85].

A valle di questa letteratura, si collocano invece gli studi sull'uso anticompetitivo dei big data. Al momento, infatti, sembrano tre i temi più discussi. In primo luogo, si valuta se il rifiuto opposto da un'impresa in posizione dominante a concedere il proprio patrimonio di dati possa dirsi abusivo ai sensi della dottrina sull'*essential facilities*. E mentre alcuni qualificano i dati come essenziali e quindi si esprimono a favore di eventuali rimedi nella forma del *duty to deal*[86], altri mostrano – oggi come in passato – non poco scetticismo sugli effetti concorrenziali di una tale misura. Il secondo tema, fortemente dibattuto in dottrina, riguarda l'uso anticompetitivo degli algoritmi e, in particolar modo, di come tali algoritmi possano servire a fare discriminazione di prezzo e colludere. Se il primo profilo sembra al momento non destare particolari pericoli sotto il profilo antitrust, sempre che si scelga di tutelare il *total welfare* in luogo del *consumer welfare*[87], la possibilità che le imprese munite di big data facciano cartello, attraverso opportuni algoritmi che mimano le scelte delle imprese "come se" fossero parte di un cartello, rappresenta una questione tanto pericolosa da essere stata dettagliatamente discussa in un libro[88]. In terzo luogo, alcuni discutono dell'eventualità che i dati

concurrence e Bundeskartellamt, *Competition Law and Data*, 2016. In questa direzione anche EU Commission Staff Working Document, "Online Platforms", di accompagnamento alla Comunicazione su *Online Platforms and the Digital Single Market*, 2016, laddove tra le caratteristiche comuni alla maggior parte delle piattaforme digitali si individua «[the] ability to collect, use and process a large amount of personal and non-personal data in order to optimize, inter alia, the service and experience of each user».

[83] Vedi. Newman N., *Data Justice- Taking on Big Data as an Economic Justice Issue*, 2015.

[84] Graef I*., EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility*, 2016.

[85] Rubinfeld D. L. and Gal M. S., *Access Barriers To Big Data*, 2016.

[86] Sokol D. D. and Comerford R. E., *Antitrust and regulating big data*, 2016.

[87] Maggiolino M., *Big data e prezzi personalizzati*, in Concorrenza e Mercato, 2016, pp. 95-138; OECD, *Price Discrimination*, 2016; Bourreau M., De Streel A. and Graef I., *Big Data and Competition Policy: Market Power, Personalised Pricing and Advertising*, 2017.

[88] Stucke M. E. and Ezrachi A., *Virtual Competition - The Promise and Perils of the Algorithm-Driven Economy*, 2016; OECD, *Algorithms and Collusion-Background Note by the Secretariat*, 2017.

siano utilizzati per produrre informazioni ingannevoli o comunque informazioni tali da favorire le imprese che diffondono quelle informazioni a discapito delle loro rivali[89].

Sotto altro profilo, molta della letteratura in tema di big data ha connesso questo fenomeno a quello delle piattaforme digitali, considerando i primi come un fattore che sarebbe capace di proteggere il potere di mercato di questi soggetti attivi in internet. Il ragionamento può essere così sintetizzato: grazie ai dati raccolti dalle piattaforme, queste ultime si creano una conoscenza più approfondita della domanda e, dunque, non solo riescono a raffinare i propri servizi in relazione alle caratteristiche dei singoli utenti, ma altresì possono offrire agli inserzionisti spazi pubblicitari sempre più idonei a catturare consumatori con preferenze e abitudini di acquisto compatibili con i prodotti e servizi che essi intendono fornire. Viene così a determinarsi un doppio effetto di feedback positivo a favore delle piattaforme che, da un lato, perfezionando i propri servizi, accrescono i loro utenti, i quali a loro volta mettono a disposizione delle piattaforme sempre più dati e, dall'altro lato, migliorando la loro offerta di pubblicità comportamentale, aumentano i loro inserzionisti, i quali dal loro canto sono disposti a pagare sempre più denari perché sono messi in contatto proprio con i consumatori maggiormente propensi ad acquisire i loro beni. Ma vi è di più. L'associazione tra big data e piattaforme ha fatto sì che l'analisi antitrust dei primi si sia intrecciata con la discussione di temi diversi, che tipicamente riguardano le sole piattaforme, come: (*i*) la definizione dei mercati rilevanti quando i modelli di business impiegati sono a due o più versanti; (*ii*) l'apprezzamento del potere di mercato delle piattaforme, quando i prezzi da queste praticati sono nulli; (*iii*) le variabili competitive, come la qualità, rispetto alle quali le imprese si farebbero concorrenza a fronte di beni venduti a prezzi nulli[90]; nonché (*iv*) la comprensione delle strategie di business delle piattaforme.

In secondo luogo, molta letteratura si è sinora concentrata sul rapporto tra diritto antitrust e privacy[91] per dare conto di almeno tre scenari diversi: (*a*) l'eventualità che il diritto antitrust sia utilizzato per sanzionare condotte lesive delle disposizioni a tutela dei dati personali o dei principi che informano quella disciplina, come sembra abbia voluto suggerire il Bundeskartellamt avviando il

[89] Michael Luca et al*., Does Google Content Degrade Google Search? Experimental Evidence*, Harvard Business School Working Paper, 2015, ma su posizioni diametralmente opposte Eldeman B. e Lockwood B., *Measuring Bias in Organic Web Search*, 2011 e Wright J., *Defining and Measuring Search Bias: some preliminary bias*, 2011. Inoltre, Michael Luca, Reviews, *Reputation and Revenues: The Case of Yelp.com*, Harvard Business School 2011 e Patterson M., *Manipulation of product ratings: Credit-Rating Agencies, Google, and Antitrust*, CPI Antitrust Chronicles, 2012.

[90] Grunes A. and Stucke M., *"No Mistake About It: The Important Role of Antitrust in the Era of Big Data"*, 2015; Gal M. S. and Rubinfeld D. L., "Access Barriers to Big Data", 2016; Harbour P.J., *"Dissenting statement, In the matter of Google/DoubleClick"*, 2007, p. 4.

[91] Swire P., *Protecting Consumers: Privacy Matters in Antitrust Analysis*, 2007; Grunes A., *Another Look at Privacy, 2013, 20 Geo. Mason L. Rev. 1107; Pasquale F., Privacy, Antitrust, and Power*, 2013, 20 Geo. Mason L. Rev. 1009; Grunes A. and Stucke M., *No Mistake About It: The Important Role of Antitrust in the Era of Big Data*, 2015; Stucke M. and Ezrachi A., *When Competition Fails to Optimise Quality: A Look at Search Engines*, 2016, 18 Yale Journal of Law and Technology 70.

procedimento contro *Facebook* [92]; (*b*) la possibilità che delle condotte siano considerate anticompetitive perché tali da determinare una riduzione del livello di privacy garantito ai consumatori e, per tale via, una diminuzione della qualità dei beni e servizi offerti sul mercato e una conseguente riduzione del *consumer welfare*[93]; nonché *(c)* l'eventualità che il potere di mercato delle piattaforme si manifesti in rapporto alla quantità di dati che ottengono dai loro utenti giacché, essendo questi la nuova moneta impiegata per gli acquisti in internet, il sovrapprezzo che le imprese con potere di mercato sarebbero solite estrarre dai consumatori si tradurrebbe per l'appunto nell'ottenimento di una quantità extra di dati[94].

In ultimo, si contano alcune voci in dottrina che contemplano l'ipotesi che i big data sollecitino alcune questioni che riguardano, oltre al diritto della concorrenza e alla privacy, anche la tutela del consumatore[95]. Questo dovrebbe essere il caso: (a) dei contratti conclusi per la raccolta dei dati, i quali potrebbero contenere clausole ingannevoli e vessatorie; e più in generale (b) della qualità e della quantità dell'informazione che ad oggi le imprese distribuiscono nel mercato, la quale potrebbe essere ingannevole, lacunosa e fuorviante. Non sono in pochi a credere, infatti, che la capacità delle imprese di manipolare, tramite raccomandazioni, recensioni e risultati della ricerca, l'informazione sul mercato sia inversamente proporzionale alla pratica degli utenti di internet di prendere per buone "le verità" della rete (e non più quelle della televisione).

[92] Ohlhausen M. and Okuliar A., *Competition, Consumer Protection, And The Right [Approach] To Privacy*, 2015; Stucke M. and Grunes A., *Big Data and Competition Policy,* 2016.

[93] Harbour PJ. and Koslov TI., *Section 2 in a Web 2.0 World: An Expanded Vision of Relevant Product Markets*, 2010, 76 Antitrust L J 769, 794, in cui gli autori chiedono di "consider whether a dominant or potentially dominant firm would have the same incentives to adapt its privacy policies either in response to consumer demand or as a reaction to competition from other firms. If achieving a dominant market position might change the firm's incentives to compete on privacy dimensions, this is a consequence that antitrust enforcers might wish to explore further"; Stucke M. and Grunes A., *The Curious Case of Competition and Quality*, 2015; Autorité de la concurrence e Bundeskartellamt, cit., 23-24.

[94] Newman N., *Search, Antitrust, and the Economics of the Control of User Data*, 31 Yale J. on Reg. 401, 2014; Newman N., *The Costs of Lost Privacy: Consumer Harm and Rising Economic Inequality in the Age of Google*, 40 Wm. Mitchell L. Rev. 849, 2013-2014; Newman J. H., *Antitrust in zero price markets: Foundations*, 164 University of Pennsylvania Law Review 149, 2015; Newman J. H., *Antitrust in zero price markets: Applications*, 94 Washington University Law Review 49, 2016; Evans D., *The antitrust economics of free*, 7 Competition Pol'y Int.l 71, 2011; Auer D.,Petit N., *Two-sided markets and the challenge of turning economic theory into antitrust policy*, 2015, https://orbi.ulg.ac.be/bitstream/2268/207347/1/Petit%2C%20Nicolas%2C%20Auer%2C%20Dirk%20-%20Two-Sided%20markets.pdf

[95] Harbour P.J., *Dissenting statement, In the matter of Google/DoubleClick,2007*, p. 4.; Ohlhausen M. and Okuliar A., *Competition, Consumer Protection, And The Right [Approach] To Privacy*, 2015; Monopolkommission, *Competition policy: The challenge of digital markets, Special Report*, 2015; and European Data Protection Supervisor, cit.; Kerber W., *Digital Markets, Data, and Privacy: Competition Law*, Consumer Law, and Data Protection, 2016, 11 Journal of Intellectual Property Law & Practice 856.

3.3 I big data e la necessità di procedere caso per caso

Nella prima parte del presente rapporto si sono raggiunte alcune acquisizioni che rilevano in sede antitrust o, *rectius*, che possono indirizzare l'analisi antitrust quanto al suo oggetto e al suo metodo.

In primo luogo, si è rimarcata la funzione economica dei big data, ossia la loro capacità di produrre valore grazie alla formazione di nuova conoscenza. In altri termini, si è detto che se si vuole cogliere l'importanza strategica dei big data, occorre ricordare che le imprese usano i big data per migliore i propri processi decisionali, rendere più efficienti i loro meccanismi produttivi, distributivi e promozionali, nonché per ideare nuovi beni e servizi, se del caso personalizzati. Ora, questa notazione aiuta chi svolge l'analisi antitrust a fissare un punto: i big data sono una risorsa che si colloca "a monte" delle diverse attività che un'impresa può svolgere servendosene. I big data hanno cioè una natura strumentale e la loro ragion d'essere deve essere quindi colta in termini relazionali, ossia guardando alle soluzioni che grazie ai big data vengono ideate. Insomma, basta soffermarsi sulla catena del valore per come descritta nella prima parte di questo lavoro per comprendere come i big data siano un input, tanto nell'ipotesi in cui un'impresa li impieghi per migliorare i propri processi, quanto nel caso in cui li usi per creare nuovi beni capaci di meglio soddisfare i consumatori, quanto ancora nella circostanza che li utilizzi per elaborare raccomandazioni, recensioni o altro. Giova ribadirlo, dai big data si possono inferire le più disparate informazioni che poi possono essere impiegate per ideare i più diversi prodotti e servizi e/o per formulare i più differenti messaggi da comunicare al pubblico.

Ma anche i big data sono tanti e diversi. Come infatti suggerito nella prima parte di questo studio, sotto il profilo contenutistico, esistono tanti e diversi tipi di big data: non a caso, nessuno ha mai provato a descrivere i big data muovendo dal "cosa" essi catturano sotto forma di sequenze di uno e zero. In altri termini, i dati digitali che vengono raccolti nel mondo non sono evidentemente i medesimi per ogni soggetto che si dispone a raccoglierli con, per giunta, strumenti e tecnologie diversi. Giusto a titolo di esempio, rientrano ugualmente nella categoria dei big data sia i dati astronomici provenienti dalle sonde spaziali, sia i dati sul genoma umano raccolti da imprese che si occupano di mapparlo. Ancora, sono big data sia i dati delle banche che raccontano dell'accesso al credito degli individui, sia i dati delle piattaforme *e-commerce* che riportano delle abitudini di consumo di quegli individui.

Inoltre, le correlazioni e le predizioni che i big data racchiudono e permettono di elaborare, oltre a non essere prevedibili *ex ante*, non sono legate da un rapporto necessariamente biunivoco con i dati digitali confluiti nei big data. Infatti, come osservato nella prima parte del report, l'apparato analitico proprio di ogni impresa può consentire a ciascuna di loro di inferire informazioni diverse dai medesimi dati o di inferire le stesse informazioni da dati differenti. Ad esempio, laddove pare ragionevole supporre che l'analisi di dati astronomici e di dati genetici condurrà a rispondere a interrogativi differenti, non si può escludere che una ricerca circa la propensione a pagare degli individui possa essere condotta con uguale accuratezza processando tanto i dati sull'accesso al credito, quanto i dati sui consumi online. D'altro canto, processando gli stessi dati sugli acquisti in rete, mentre un'impresa potrebbe essere interessata ad individuare l'ora nella quale cresce la disponibilità a pagare degli individui, un'altra impresa potrebbe voler stabilire l'età di coloro che

scelgono di rivolgersi alle piattaforme di *e-commerce*. Insomma, giacché la complessità dei fenomeni consente di registrarli da punti di vista distinti, ma affini, anche big data diversi per dati raccolti potrebbero racchiudere informazioni simili. Analogamente, poiché imprese e amministrazioni possono ben porsi interrogativi differenti, big data che racchiudono dati simili possono condurre a inferenze diverse, ossia a individuare correlazioni e a elaborare predizioni differenti. Ora, anche questa duplice osservazione circa la molteplicità dei big data e delle informazioni che da essi possono essere inferite rileva sotto il profilo antitrust. Essa svela la necessità di appurare caso per caso non solo quanto uno specifico data set, qualificato come big data, offra un'informazione capace di migliorare un particolare processo imprenditoriale, di ideare un determinato prodotto o servizio, o di produrre una certa informazione da rivolgere al mercato, ma anche quanto quell'informazione non possa essere ottenuta altrimenti, ossia analizzando altri insiemi di dati, o offerta da altri operatori. Come sempre accade quando ci si interroga con riguardo a degli input più o meno reperibili e più o meno utilizzabili, compito dell'analisi antitrust sarà cioè capire "cosa serve a chi" e "in quale misura" e perché chi eventualmente reclama quel "cosa" non possa procurarselo da solo e perché – parlando delle informazioni rivolte al mercato – non esistano altri soggetti disposti a produrle e diffonderle.

In terzo luogo, nella prima parte si è evidenziato come il valore dei big data possa risiedere in molte e diverse delle loro caratteristiche. In alcuni settori merceologici, il valore dei big data risiede proprio nel loro volume e nella loro varietà; in altre industrie, il valore di questi insiemi di dati dipende invece da quanto essi vengono aggiornati e in altre ancora in quanto quei dati possono rappresentare la realtà storica. Ecco che allora anche questa acquisizione rammenta a chi svolge l'analisi antitrust come sia necessario esaminare ogni contesto e ogni mercato separatamente, guardando cioè a quali caratteristiche dei particolari big data considerati sono importanti per le attività delle imprese che operano in quel mercato. Del resto, se i big data hanno una natura strumentale, anche il rilievo delle loro tante caratteristiche – come è stato brillantemente notato, i big data sono «huge in volume, high in velocity, diverse in variety, exhaustive in scope, fine-grained in resolution, relational in nature, scalable in size and flexible in composition»[96] – dipenderà dall'uso che le imprese fanno di quei big data. Rovesciando il punto di vista, per ogni bene e servizio che si vorrà produrre le imprese dovranno dotarsi di big data con specificità differenti.

In quarto luogo, si è ricordato come la raccolta, gestione e analisi dei big data richieda alle imprese di dotarsi di adeguate infrastrutture, algoritmi efficaci e opportune competenze informatiche e manageriali. In altri termini, l'utilità a fini economici (oltre che politici, sociali e scientifici) dei big data dipende anche dalle risorse materiali e intellettuali che un'impresa investe nelle attività di acquisizione, pulizia, analisi e rappresentazione dei dati. Inoltre, sempre nella prima parte del lavoro si è spiegato come anche l'analisi dei big data possa condurre a risultati falsi o perché i dati processati erano in qualche misura b*iased*, o perché sono stati commessi degli errori in sede di analisi. Di conseguenza, chi svolge l'analisi antitrust deve ricordarsi che l'implicazione secondo la quale i big data attribuirebbero *sic et simpliciter* un vantaggio informativo a chi li controlla non può che considerarsi una semplificazione. Per dare il senso della complessità del fenomeno occorrerebbe

[96] Kitchin R., *The Data Revolution*, 2014.

piuttosto affermare che detto vantaggio informativo segue eventualmente, ossia caso per caso, dalla combinazione dei big data e delle più efficaci, nonché efficienti, tecniche di analisi degli stessi.

In ultimo, si è osservato come i big data possano caratterizzarsi sia quali risorse capaci di conferire un vantaggio competitivo, sia come strumenti utili a che un'impresa si adatti velocemente ai diversi corsi e sviluppi del mercato. Nondimeno, ancora una volta chi si occupa di analisi competitive e di diritto antitrust sa che questa qualificazione non può considerarsi vera in ogni caso. Se più imprese sono dotate di big data, nessuna godrà di per ciò stesso di un vantaggio competitivo immediatamente connesso alla disponibilità di questi grandi patrimoni di dati. Tutto dipenderà da quanto le imprese sono riuscite a mettere a buon frutto quei big data. Del resto, come evidenziato con specifici esempi nella prima parte di questo documento, potrebbe anche darsi che il successo di un'impresa che pur controlla molti dati digitali non sia da attribuire a questi o alle sue tecniche di analisi, quanto ad altri fattori, come all'idea di business sviluppata o alla tecnologia impiegata per raggiungere i consumatori. Ecco che allora l'analisi antitrust avrà un'ennesima ragione per procedere caso per caso, senza indugiare in facili generalizzazioni o *one-fits-all solutions*.

Peraltro, con riguardo ai vantaggi informativi e competitivi derivanti dai big data, va precisato che proprio in ragione del suo interesse a tutelare il corretto funzionamento del mercato, il diritto antitrust adotti un approccio neutrale nei confronti di questi vantaggi. In altri termini, è nell'interesse del mercato e del benessere dei consumatori che le imprese trovino il modo di ideare prodotti e servizi migliori, che poi serviranno anche per vincere la concorrenza opposta dai loro rivali. Di conseguenza, il diritto antitrust – che non tutela i concorrenti ma la concorrenza – non persegue la formazione di un vantaggio informativo o competitivo in capo a una impresa, a meno che quell'impresa non abbia conseguito quel vantaggio attraverso una pratica – intesa, concentrazione o condotta unilaterale – che riduce il *consumer welfare*.

Ora, l'indicazione di metodo alla quale convergono le tante acquisizioni appena ricordate, ossia l'esigenza secondo la quale l'analisi antitrust dovrebbe svilupparsi guardando al particolare e senza velleità universalistiche, verrà tenuta in considerazione nei successivi paragrafi dove – come anzidetto – si considererà: (i) se e quando le imprese possano avere delle difficoltà a formarsi un patrimonio di big data da utilizzare per migliorare la loro offerta; (ii) se e quando i big data possano creare delle rendite di posizione a favore di alcune imprese, se del caso attraverso la manipolazione dell'informazione; e (iii) se e quando i big data possano essere utilizzati per tenere delle condotte imprenditoriali lesive della concorrenza che si consuma non solo nella produzione di beni e servizi, ma anche nella produzione di informazione.

3.4 I big data e il potere di mercato

Sono principalmente due gli ordini di questioni che in letteratura vengono associate al binomio big data-potere di mercato. Per un verso, ci si chiede se i big data conferiscano potere di mercato o se, per lo meno, contribuiscano a schermarlo. Per altro verso, si discute del caso particolare delle *multi-sided media platforms*, perché ci si interroga sulle modalità da seguire per apprezzare il loro potere di mercato. Di seguito, si procederà seguendo questo ordine di argomenti.

3.4.1 Il potere di mercato nei mercati a valle di prodotti e servizi offerti usando big data

Come osservato sopra, il fatto che una impresa disponga di big data non implica necessariamente che quell'impresa disponga di potere di mercato nel mercato o nei mercati della generazione e raccolta di quei dati. Analogamente, non sembra che si possa stabilire un nesso causale necessariamente vero tra l'esistenza di big data in capo a una impresa e il fatto che essa possa esercitare potere nei mercati che si collocano a valle di quello – o di quelli – della generazione e raccolta dei dati. Più esattamente, in rapporto agli utilizzi dei dati generati e raccolti da un'impresa, sembrano delinearsi almeno due differenti scenari, non mutualmente escludenti.

In primo luogo è possibile, come visto sempre nel terzo capitolo, che le imprese generino/raccolgano dati digitali per rivenderli in tutto o in parte, grezzi e/o rielaborati[97]. In questo caso bisognerà allora verificare se nei successivi mercati della commercializzazione dei dati digitali una o più imprese finiscano per acquistare – in virtù di meriti, fortuna, o altro[98] – potere di mercato. Ma certo quest'ultima analisi rappresenterà una indagine distinta da, e successiva a, quella relativa all'accertamento dell'esistenza di un certo potere di mercato nel mercato della generazione e raccolta dei dati digitali. In altri termini, di volta in volta, occorrerà definire i confini dei mercati per la commercializzazione dei dati, così da verificare come essi si collochino e comportino rispetto al mercato – o ai mercati– per la generazione/raccolta dei dati. E, fatta questa analisi, potrà anche accadere che un eventuale potere nel mercato a monte della generazione/raccolta dei dati digitali si traduca in un sovrapprezzo nel mercato a valle della loro commercializzazione.

In secondo luogo, è possibile che le imprese scelgano di utilizzare "internamente" i dati da loro generati e raccolti per poi ideare nuovi beni e servizi, siano questi delle polizze assicurative personalizzate, dei video-giochi, o dei consigli nutrizionali. In questo caso, bisognerà procedere con l'individuazione dei mercati di questi beni e servizi – mercati in cui le imprese potranno anche vantare un potere di mercato significativo, ma non perché depositarie di enormi volumi di dati, o perché uniche detentrici di una particolare categoria di dati, ma in virtù di quanto i loro beni e servizi superano quelli dei rivali. Di conseguenza, similmente a quanto si diceva sopra, non sembra ragionevole ritenere che i big data in quanto tali possano da soli attribuire a una impresa il potere di agire indipendentemente dai suoi concorrenti nei mercati dei suoi beni e servizi.

Se si vuole raggiungere questa conclusione bisognerà piuttosto argomentare che il possesso dei big data, o di una speciale categoria di dati, è ciò che permette a una impresa di migliorare i suoi beni e servizi al punto di sbaragliare la concorrenza attuale e potenziale nell'offerta di quei prodotti. Ma chi sviluppa il ragionamento in questi termini sta invero articolando due analisi congiuntamente. In termini dinamici, sta guardando a come una condotta unilaterale sia capace di produrre effetti

[97] Cfr. capitolo terzo, paragrafo 4.2.

[98] Si pensi, al riguardo, a quanto affermava Justice Douglas quando, parlando dell'acquisizione del potere di mercato, distingueva il caso in cui essa fosse frutto una condotta escludente e anticompetitiva, dal caso in cui essa seguiva una forma di crescita e sviluppo dell'impresa «as a consequence of a superiorproduct, business acumen, or historicaccident» – *UnitedStates v. Grinnell Corp.*, 384 U.S. 563, 571 (1966).

escludenti ma comunque non anti-competitivi, finché i prodotti offerti risultano migliori di quelli dei rivali – questione di per sé rilevante, e già esaminata, se e solo se l'impresa che tiene questa condotta unilaterale sia già in posizione dominante[99]. In termini statici, chi si sofferma sull'esclusione prodotta dalla possibilità di disporre di big data sta concettualizzando questi ultimi come un vantaggio competitivo e, dunque, come una barriera posta a difesa del mercato a valle, dei beni e servizi realizzati dall'impresa che controlla quei big data.

3.4.2 Barriere all'ingresso nei mercati a valle di prodotti e servizi offerti usando big data

Sono due le circostanze che possono rilevare se ci si chiede se i big data possano servire a una impresa (anche tradizionale) per rendere meno contendibile la propria posizione di mercato.

In primo luogo, può accadere che i big data siano assimilabili a delle risorse ritenute essenziali per la realizzazione di un determinato prodotto e/o servizio – circostanza questa che evidentemente pone l'impresa che detiene quei particolari dati nella condizione privilegiata di controllare un bottleneck e, dunque, di godere di un vantaggio competitivo assoluto che fisiologicamente rende la sua posizione di mercato meno contendibile. Di questa eventualità però si è già discusso con riferimento ai comportamenti che portano all'acquisizione di un input essenziale e si continuerà a dire di seguito in relazione alla dottrina delle risorse essenziali[100]. E ciò perché nella prospettiva statica qui abbracciata il diritto antitrust deve solo limitarsi a constatare la possibile esistenza di eventuali barriere all'ingresso, laddove ci si è interrogati e ci si interrogherà sulla liceità dei comportamenti delle imprese che, di volta in volta, possono consistere in forme di acquisizione dei dati o in rifiuti a condividere quei dati.

In secondo luogo, si può ipotizzare che i big data riducano i costi delle imprese già presenti nel mercato e inneschino dei meccanismi che disincentivano i consumatori dall'abbandonare i prodotti e i servizi dell'impresa che detiene quei big data, andando così a rinforzare il potere di mercato dell'impresa medesima e a rendere più probabile l'eventualità che il mercato rilevante risulti concentrato[101]. Più nel dettaglio si è sostenuto che, al crescere dei dati disponibili a una impresa,

[99] Cfr. capitolo terzo, paragrafo 6.

[100] Cfr. H.A. Shelanski, *Information, Innovation, and Competition Policy for the Internet*, 161 U. Pa. L. Rev. 1663, 1680-81 (2013), dove osserva che «[w]hen customer information is a useful input for a platform and is not equally available to that platform's competitors, the informational advantage can help to entrench market power».

[101] Le imprese che non hanno accesso a flussi di dati comparabili – per dimensione e varietà – a quelli disponibili all'incumbent potrebbero cioè incontrare delle difficoltà nell'entrare e/o nel prosperare nel mercato rilevante. Infatti, proprio a causa delle economie di scala e di scopo derivanti dall'utilizzo di complessi algoritmi, è probabile che le imprese di grandi dimensioni già operanti sul mercato beneficino – rispetto ad imprese di minori dimensioni e alle potenziali entranti – di vantaggi di costo nelle attività di raccolta, stoccaggio e analisi dei dati. In alcune circostanze, tali vantaggi possono rivelarsi sostanziali – cfr. Autorité de la Concurrence, Bundeskartellamt, *Competition Law and Data*, cit., 51 e ss..

risultino maggiori le economie di scala[102], scopo[103] e aggregazione[104] di cui quell'impresa beneficia nella produzione e commercializzazione dei suoi beni e servizi. Inoltre, si dice che i big data schermino la posizione delle imprese che li controllano in ragione del c.d. effetto di retroazione – o *feedback loop*, nell'espressione inglese[105] – che, rispetto alle *multi-sided media platforms*, viene declinato in due differenti versioni[106]. In primo luogo, si considera l'effetto di retroazione dei

[102] In altri termini, si dice che all'aumentare dei dati disponibili diminuiscono i costi medi implicati dalla produzione dei beni e dei servizi ideati sulla scorta di quei dati. Al riguardo, va però ricordato che la semplice esistenza di economie di scala non implica, di per sé, che i provider di dimensioni maggiori abbiano un potere di mercato vicino a quello di monopolio,né che questi abbiano agito in modo anticompetitivo, né che i consumatori o il processo competitiviosiano stati danneggiati. E nemmeno si può affermare che questi grandi provider debbano essere sottoposti a controlli più stringenti da parte delle autorità antitrust proprio in virtù delle economie di scala. Infatti, sottoporre tali provider a regolazione più stringente per la mera esistenza di economie di scala equivarrebbe ad imporre loro una tassa..

[103] In altri termini, si dice anche marginalizzare i rivali delle imprese che raccolgono dati. Si pensi, ad esempio, a Facebook. Da un versante, migliorando la sua offerta di pubblicità comportamentale, il social network vede aumentare gli inserzionisti determinati a comprare spazi pubblicitari a fronte di sempre più denari. Dall'altro versante, perfezionando i propri servizi offerti a prezzo nullo, Facebook accresce il numero di utenti, i quali a loro volta non solo, in ragione del già menzionato effetto di rete indiretto, attraggono sempre più inserzionisti, ma altresì mettono a disposizione del social network sempre più dati, consentendo così che il ciclo ricominci.

[104] EDPS, *Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy*, *preliminary opinion* (2014), § 66, dove silegge «[p]owerful or dominant undertakings are able to exploit 'economies of aggregation' and create barriers to entry through their control of huge personal datasets alongside proprietary software which organises the data». Inoltre, EU Commission Staff Working Document, *Online Platforms* 2*,* COM (2016) 288, https://ec.europa.eu/digital-single-market/en/news/commission-staff-working-document-online-platforms. Si veda anche Newman, *Search, Antitrust,* cit., 451-452, dove silegge«Information is itself a commodity and particular companies can, in Stiglitz's words, "appropriate the returns to creating information" for economic advantage in the market in ways not seen in traditional commodities. This leaves a company like Google with the economic incentive, pushed by advertisers seeking their own information advantage in marketing to users, to increasingly violate user privacy to appropriate more and more information to solidify its economic dominance, all while making it less economically possible for potential competitors to challenge the company. This concentrated control of user data gives such companies more than just a competitive advantage vis-a-vis their competitors. Companies know the intimate details about the preferences of consumers and this skews markets against consumers who lack similar data about those companies or their options in negotiating terms with those companies».

[105]H.R. Varian , Joseph Farrell, Carl Shapiro, *Introduzione all'economia dell'informazione* (2005).

[106]Cfr. C. Argenton, J. Prüfer, *Search Engine Competition With Network Externalities,* 8 J. Competition L. & Econ. 73 (2012); UK Competition & Markets Authority, *The commercial use of consumer data* § 3.56 (2015); Monopolkommission, *Competition policy: The challenge of digital markets* §§ 2.2.1 e 2.2.3 (2015); DanielRubinfeld, Michal Gal, *The Hidden Costs*, cit., 17-22; N.-P. Schepp, A. Wambach, *On Big Data and Its Relevance for Market Power Assessment* 121-122(2016); M.E. Stucke, A.P. Grunes, *Big Data and Competition Policy* 337 (2016). Inoltre, Autorité de la concurrence and Bundeskartellamt, *Competition Law and Data* 30 (2016), dove si legge«[w]hile dynamic competition could be strong enough to mitigate competition law concerns related to static market power on at least some markets, this balance should be evaluated on a case-by-case basis and the difficulties for new undertakings to enter and grow on online or digital markets should not be underestimated and therefore be carefully assessed, along with the possible role of data collection in this regard»; Graef, *Market Definition*, cit., 484-486; F. Thépot, *Market Power in Online Search and Social-*

confronti degli utenti, ossia la circostanza per cui al crescere dei dati raccolti, aumenta la qualità dei beni offerti e dunque il potere attrattivo dell'impresa nei confronti dei consumatori e dei loro dati, così che l'impresa, ottenendo sempre più dati in via continuativa, instaurerà un circolo virtuoso a proprio vantaggio[107]. In secondo luogo, si individua l'effetto monetario di retroazione, ossia il meccanismo per cui, basandosi su una base dati di sempre migliore qualità, le piattaforme riuscirebbero a creare una pubblicità comportamentale sempre più capace di cogliere le necessità degli inserzionisti paganti, per così instaurare un secondo circolo virtuoso teso ad ottenere sempre più ricavi con cui potenziare le proprie tecniche di raccolta e analisi dei dati e, dunque, la qualità dei propri servizi. Ad esempio, di questo duplice effetto di retroazione beneficerebbero i motori di ricerca che vedono aumentare la qualità dei loro risultati all'aumentare delle ricerche eseguite[108], le piattaforme di social networking[109], nonché imprese di distribuzione di beni e contenuti le cui raccomandazioni risultano sempre più accurate all'accrescersi del numero di dati raccolti[110].

Networking: A Matter of Two-Sided Markets, 36 *World Competition* 195, 202-203 (2013) dove però questa espressione non vieneutilizzata.

[107] Così G. Pitruzzella. *Big Data, Competition And Privacy: A Look From The Antitrust Perspective*, 23 Concorrenza e Mercato 19 (2016), il quale osserva che «larger companies have access to larger datasets and therefore can offer more successful services to consumers, which in turn allow them to collect even more information and data. This self-reinforcing mechanism may be similar to a network effect driving market concentration». Inoltre, Pasquale, *Privacy, Antitrust*, cit., 1015, il quale ha osservato che l'incapacità dei concorrenti di uguagliare la base dati delle imprese in posizione dominante, li pone in una posizione di perenne svantaggio: a causa del «self-reinforcing data advantage of dominantfirms, there is unlikely to be muchcompetition in search and social networking».

[108] Per migliorare il loro algoritmo, i più noti motori di ricerca come Google, Yahoo e Bing raccolgono e memorizzano i dati sugli utenti, tra cui la data e l'ora della ricerca, la posizione dell'utente (in base all'indirizzo Internet Protocol (IP)) e i dati relativi alle query di ricerca che gli utenti hanno cercato nonché i collegamenti successivamente cliccati. Questi registri di ricerca vengono utilizzati dai motori di ricerca per migliorare l'accuratezza dei loro risultati di ricerca guardando, ad esempio, in quale lingua, da quale posizione geografica e in quale ora del giorno un utente entra in un particolare query di ricerca –cfr. Oren Bracha, Frank Pasquale, *Federal SearchCommission? Access, Fairness and Accountability in the Law of Search*, 93 Cornell L. Rev. 1149, 1181 (2008), dove siosservaappuntoche «[t]he more searches an engine gets, the better able it is to sharpen and perfect its algorithm»; e Kristine Laudadio Devine, *Preserving Competition in Multi-Sided Innovative Markets: How Do You Solve a Problem Like Google?*, 10 N.C. J.L. & Tech. 59 (2008).

[109] I fornitori di social network sono in grado di migliorare la qualità del loro servizio agli utenti aumentando la rilevanza delle interazioni sociali e dei contatti suggeriti che vengono visualizzati a un determinato utente. Un algoritmo viene utilizzato per selezionare quali elementi di informazioni sono più rilevanti per un utente specifico. Ad esempio, le storie che vengono visualizzate in un feed di notizie di un utente su Facebook sono influenzate dalle connessioni e dall'attività di tale utente sulla piattaforma.

[110] Nell'ambito del funzionamento delle piattaforme di commercio elettronico, la raccomandazione o il sistema di raccoglitori svolgono un ruolo centrale. La raccolta e l'analisi dei dati relativi al comportamento degli acquirenti degli utenti, al loro carrello virtuale e agli articoli che hanno visto, gradito o valutato, consente alla piattaforma di meglio prevedere i prodotti ai quali gli utenti sono maggiormente interessati in base alla loro somiglianza con altri utenti appartenenti ai medesimi cluster.

Tuttavia, con riferimento alla consistenza delle economie generate dai dati e con riferimento alla portata dei descritti effetti di retroazione la dottrina si è divisa, dando luogo ad un dibattito assai vivace, ancorché spesso intrecciato a valutazioni non immediatamente legate ai soli dati, quanto più genericamente connesse alle tecnologie dell'informazione e, in particolare, alle piattaforme digitali multi-versante, anche quando queste non offrono prodotti e servizi a prezzo nullo[111]. In particolare, anche sulla scorta di alcuni casi empirici, si è negata l'esistenza di un *first mover advantage*[112], esaltando il fenomeno del *multi-homing*, evidenziando come gli effetti di rete siano unidirezionali[113] e ricordando come anche l'accumulo dei dati digitali soffra di rendimenti decrescenti[114]. Così si è arrivati a negare l'ipotesi secondo la quale i mercati digitali si concentrerebbero[115], portando ad esempio casi di imprese che oggi hanno acquistato importanti quote di mercato senza essere state le *prime* ad operare nei relativi mercati digitali[116].

[111] Cfr. OECD, *Role and Measurement,* cit., 3 dove si legge l'opinione espressa da Hal Varian secondo cui i cc.dd. "user feedback loops"assomigliano al "learning-by-doing", un fenomeno caratterizzante le funzioni di produzione di quasi tutti i settori industriali.

[112]Cfr. Andres V. Lerner, *The Role of Big Data in Online Platform Competition* 58-61 (2014); e Andrei Hagiu, Simon Rothman, *Network Effects Aren't Enough*, Harv. Bus. Rev. 64 (2016), dove si legge «[a]s a result [of the network effects], entrepreneurs often mistakenly assume that they need to reach the exponential growth phase as quickly as possible. However, a headlong rush to fast growth is often unnecessary and can even backfire, for several reasons. Entrepreneurs should really focus on being the first to create a liquid market in their segment. The winner in the marketplace is the first one to figure out how to enable mutually beneficial transactions between suppliers and buyers—not the first one out of the gate. Indeed, many prominent marketplace businesses were not first movers: Airbnb was founded more than a decade after VRBO, Alibaba was a second mover in China after eBay, and Uber'sUberX copied Lyft's peer-to-peer taxi business model». Si vedano, inoltre, in termini più generici Anja Lambrecht, Catherine Tucker, *Can Big Data Protect a Firm from Competition?* 7 (2015), nonché Andrei Hagiu, Joshua D. Wright*, Multi-Sided Platforms,* 43 *International Journal of Industrial Organization* 162 (2015).

[113] Lerner, *The Role*, cit.,6, dove si dice chiaramente che «[a]lthough advertisers value the ability to access users, user demand for a platform is not substantially driven by the availability of advertisements. The fact that cross-platform network effects are essentially one-sided fundamentally weakens or eliminates the possibility of a feedback loop that locks users and advertisers to a dominant platform. If a smaller rival or new entrant offers a better service to users, network effects do not inhibit users from switching and, once users switch, advertisers likely would follow».

[114]Cfr. ancora Lerner, *The Role*, cit., 20, il quale rimanda a W. Terdoslavich, *Big Data & The Law Of Diminishing Returns*, INFORMATION WEEK (2015), http://www.informationweek.com/big-data/big-data-analytics/big-data-and-the-law-of-diminishing-returns/d/d-id/1323310

[115] Cfr. Lerner, *The Role*, cit.,46, dove sileggerisce che «[c]ontrary to claims that online markets are prone to "tip" to dominant platforms because of the collection of user data, the relatively short history of the Internet is filled with examples of "dominant" platforms or providers that subsequently were displaced by new firms—many new entrants have prospered, and many large online providers thought to be dominant at the time have failed or declined in the face of new competition».

[116] Cfr. Lerner, *The Role*, cit., 4-5; Lambrecht, Tucker, *Can Big Data*, cit., 12-15; Daniel D. Sokol, Roisin Comerford, *Antitrust And Regulating Big Data*, 23 George Mason Law Review 119, 1136 (2016), dove I casi considerati sono quelli di Facebook, WhatsApp, Candy Crush Saga, Uber, AirBnB e Tinder, nonché di imprese che non sono native digitali come Walmart, GE, Ferrari F1, e Union Pacific.

Tuttavia, questo pur interessante dibattito è frutto di un fraintendimento, ossia dell'idea che si possano stabilire ex ante le caratteristiche dei costi che un'impresa deve sostenere per fare ingresso in un mercato. Invero, tra gli studiosi di diritto antitrust è del tutto condivisa l'idea secondo cui ogni mercato abbia le proprie barriere da valutare caso per caso. Per questo, è ragionevole ritenere che di volta in volta le autorità antitrust dovranno stabilire se ed in che misura un certo patrimonio di dati operi per schermare l'attività commerciale di un'impresa, se del caso insieme ad altri fattori legati – come è probabile – all'impiego di tecnologie dell'informazione.

Questione diversa è invece quella che vuole le imprese dotate di big data difese dalla concorrenza dei potenziali rivali in ragione dei c.d. *lock-in effects* generati dai dati personali presenti tra quei big data. In particolare si ritiene che, anche laddove dei consumatori volessero cambiare fornitore di beni e servizi, non sarebbero incentivati a farlo perché, così facendo, potrebbero passare a un fornitore che, privo di quei dati, risulterebbe meno capace di proporre un'offerta ritagliata sulle loro preferenze. Tuttavia, come si vedrà nel paragrafo 5.5, in virtù del nuovo regolamento in materia di privacy e delle differenti disposizioni che facilitano la circolazione dei dati personali e l'apertura dei dati che personali non sono, questo stato di cose è venuto a modificarsi, per lo meno all'interno dell'Unione Europea.

3.4.3 I mercati digitali, le piattaforme *multisided* e i prezzi zero

L'altro aspetto, ricordato in precedenza, nel rapporto big data / potere di mercato si riferisce a una dimensione particolare del fenomeno, legata alle cosiddette *multisided platforms*.

In generale i big data si caratterizzano per essere beni non rivali, con costi marginali di riproduzione e distribuzione trascurabili. Inoltre, talune imprese, secondo una dottrina prevalente, beneficiano di economie di scala e di scopo, derivanti dagli alti costi fissi d'infrastruttura non recuperabili (*sunk cost*), collegati alla gestione, organizzazione ed elaborazione di enormi masse di dati.

Come si vedrà più in dettaglio nel paragrafo 3.5, per quanto riguarda i primi due aspetti, il fatto che il loro consumo da parte di un individuo non ne riduca la disponibilità per gli altri e al contempo la perdita nel digitale della distinzione tra originale e copia e la sua riproducibilità con costi marginali pari a zero, sebbene non risolva alla radice il problema dell'accessibilità, determina certamente un atteggiamento di maggiore favore in una prospettiva pro-concorrenziale.

Il terzo elemento, che caratterizza anche alcuni dei mercati "tradizionali" dell'informazione di seguito analizzati, determina potenziali barriere all'ingresso, sulle quali però la dottrina economica in materia di dati non fornisce interpretazioni univoche. Secondo alcuni autori[117] infatti i dati presentano ritorni di scala decrescenti, per cui ogni unità di dato aggiunto da un certo punto in poi riduce il suo valore fino a non aggiungerne più alcuno. Tutto ciò, collegato alla specificità della elaborazione in rapporto ai fini da perseguire e la sua volatilità, come evidenziato nella prima parte del lavoro, renderebbe meno efficienti le economie di scopo e più in generale non razionale dal punto di vista del titolare del

[117] Lerner A. V., *The Role of 'Big Data' in Online Platform Competition*, 2014; Varian H. R., Tucker D. S., Wellford H. B., *Big Mistakes Regarding Big Data*, 2014.

dato, in posizione dominante, utilizzare il proprio potere di mercato per porre in essere condotte unilaterali (abusi) per ridurre o falsare la concorrenza. D'altra parte, secondo una diversa corrente di pensiero[118], su alcuni utilizzi specifici dei dati come nell'analisi predittiva basata sulle probabilità derivante da un numero maggiore di osservazioni, più alta è la misura della probabilità e quindi dei dati disponibili, tanto più è affidabile la previsione: una sorta di "learning by doing" a tutto vantaggio dell'utilizzatore, che spiegherebbe ad esempio, nel caso di Google, la superiorità del proprio motore di ricerca.

In tutti i casi, pur nella diversità dei punti di vista, che l'attività economica faccia affidamento sui dati non è un fenomeno nuovo. Anche nella cosiddetta "old economy" i dati degli utenti erano una fonte essenziale d'informazione per le imprese, come il prevedere gli acquisti in ragione dei gusti e delle preferenze del consumatore.

Da questo punto di vista i mercati digitali ripropongono questioni già analizzate in altri mercati, in particolare quelli multi-versante, in merito ai quali la teoria economica ha offerto importanti contributi negli ultimi anni, al fine di meglio interpretare il fenomeno e fornire una strumentazione più adeguata al mutato quadro di riferimento.

Prima però di sviluppare questa analisi occorre precisare chiaramente che ci troviamo di fronte al caso particolarissimo delle imprese che raccolgono big data e offrono i loro beni non solo a prezzi nulli, ma altresì impiegando modelli di business a più versanti. Questo è un caso "particolarissimo" non perché si voglia disconoscere il ruolo che le c.d. *two-sided media platforms* giocano nell'economia digitale, ma perché i problemi che attengono all'accertamento del loro potere di mercato non sono connessi alla detenzione dei big data, ma alla scelta di queste imprese di praticare prezzi nulli all'interno di quelli che, per giunta, sono modelli di business che aggregano due o più domande.

In altri termini, le difficoltà connesse all'apprezzamento del potere di mercato si pongono ogni volta che una impresa decide di vendere i suoi beni a prezzi nulli, di là dal fatto che essa detenga anche dei big data. In più, questi problemi si potenziano quando quell'impresa colloca tale vendita all'interno di un modello di business a più versanti perché, in presenza di questi modelli, la definizione del mercato rilevante vacilla e, quindi, rende ancor più complessa la determinazione del potere di mercato. Ma, ancora una volta, il problema della definizione del mercato rilevante in presenza di modelli di business multi-versante è questione concettualmente autonoma rispetto al tema dei big data.

Il punto di convergenza tra i fenomeni dei "big data", degli "zero-price goods" e delle "multi-sided platforms" è dunque di natura puramente empirica: capita, cioè, che imprese come Google o Facebook offrano i loro servizi di searching e di social networking a prezzi nulli, finanziando questa loro offerta con i dati dei loro utenti e i ricavi provenienti dagli inserzionisti pubblicitari interessati all'attenzione di quegli stessi utenti. Nondimeno, nella realtà non solo si danno imprese che detengono sì big data, ma che offrono i loro prodotti a prezzi positivi, attraverso modelli di business

[118] Autorité de la concurrence e Bundeskartellamt, *Competition Law and Data*, 2016.

che si rivolgono ad una sola domanda. Ad esempio, appartengono a questa categoria le compagnie aeree, quelle assicuratrici, le aziende ospedaliere, le banche, nonché tutte quelle imprese "tradizionali", come le compagnie petrolifere o le imprese che producono elettrodomestici, che proprio in questi anni stanno associando ai propri beni i sensori della rete delle cose.

Di più, nella realtà esistono imprese che praticano prezzi nulli senza raccogliere dati o senza ricorrere a modelli multi-versante, così come esistono delle imprese multi-versante che applicano prezzi positivi in ogni mercato in cui operano. A titolo di esempio, si consideri che i prezzi nulli caratterizzano strategie di pricing come il cross-selling e il bundling che non prevedono né la raccolta di dati, né l'esistenza di una seconda domanda interdipendente. Allo stesso modo, i sistemi di carte di credito sono sì multi-mercato, ma non applicano prezzi nulli.

Complessivamente, dunque, non pare corretto esaurire i temi connessi alla relazione tra i big data e il potere di mercato con la discussione del potere di mercato delle piattaforme che offrono alcuni dei loro servizi a prezzi nulli, avendo adottato un modello di business multi-versante, è comunque utile inserire, con tali premesse, questa trattazione all'interno del dibattito sulla definizione e potere di mercato, con una focalizzazione sull'efficacia della strumentazione antitrust nei mercati digitali caratterizzati dall'innovazione[119].

3.4.3.1 I mercati a più versanti

Come noto un mercato rilevante è costituito da un prodotto (o un insieme di prodotti) e un'area geografica in cui è venduto, su cui è possibile esercitare un potere di mercato. La definizione di mercato rilevante implica un'analisi di sostituibilità per individuare i vincoli concorrenziali che sono in grado di limitare il potere di mercato dell'impresa sottoposta all'indagine antitrust. Il test più noto per effettuare tale analisi è lo SSNIP test (*Small, but Significant and Non-Transitory Increase in Price*), denominato anche test del monopolista ipotetico. L'idea sottostante lo SSNIP test è che la sostituibilità tra due prodotti è tale da includerli nello stesso mercato rilevante se per l'ipotetico monopolista non fosse profittevole alzare il prezzo di uno dei due prodotti in maniera contenuta ma protratta nel tempo.

Una corretta definizione del mercato riveste pertanto grande importanza nell'analisi antitrust: la definizione dell'ampiezza del mercato rilevante può influenzare in maniera determinante le condizioni di concorrenzialità che in esso vengono riscontrate e dunque l'individuazione o meno di una condotta anti-competitiva all'interno dello stesso[120]. Ad esempio praticare un prezzo inferiore ai costi marginali costituisce tipicamente una fattispecie abusiva nelle imprese ad un versante.

[119] Schumpeter J. A., *Capitalism, Socialism and Democracy,* 1994 [1942], London: Routledge. pp. 82–83.

[120] Filistrucchi L., Gerardin D., Van Damme E., Affeldt P., *Market Definition in Two-Sided Markets: Theory and Practice*, 2014, Journal of Competition Law and Economics, Volume 10 (2), pp. 293-339.

Ciò non è necessariamente vero nei mercati a più versanti, in cui due (o più) distinti gruppi di utenti (versanti) interagiscono attraverso un intermediario o piattaforma, che coordina la loro domanda, favorendo la relazione tra i due versanti[121].

Questi mercati si caratterizzano per alcune caratteristiche quali:

(i) *la presenza di esternalità indirette tra i due gruppi di utenti*. L'utilità che gli utenti di un versante traggono dal consumo di un prodotto o servizio aumenta all'aumentare del numero degli utenti di altri versanti che acquistano un prodotto diverso, ma collegato alla stessa piattaforma;

(ii) *non neutralità della struttura dei prezzi*: il modo in cui i prezzi sono distribuiti tra gli utenti dei due versanti influisce sulla domanda e il livello di transazione. La struttura dei prezzi su due versanti è importante tanto quanto il loro livello per determinare la domanda e l'output;

(iii) al gruppo di utenti che genera il più alto livello di effetti di rete indiretti sarà applicato un prezzo relativamente più basso, un prezzo inferiore al costo marginale, o addirittura sotto lo zero (*versante sussidiato*). Il contrario accade agli utenti dell'altro versante (*versante profit-making*).

Se pertanto una caratteristica dei mercati a più versanti è la struttura di prezzi "sbilanciati", per ritornare all'esempio precedente, praticare un prezzo inferiore ai costi marginali al lato del mercato che presenta minore esternalità non necessariamente costituisce una fattispecie abusiva, ma al contrario potrebbe rientrare nella ottimale logica di massimizzazione del profitto delle imprese[122].

Di conseguenza, quando si discute di attività di impresa multiversante, l'effetto di alcune pratiche o accordi può essere particolarmente complessa: è probabile che si sottovaluti alcuni effetti o si

[121]La prima analisi dei mercati multiversante risale ad uno studio di William Baxter, capo della divisione antitrust del DOJ, sul mercato delle carte di pagamento. Lo studio rilevò che il successo del business delle carte di pagamento dipendeva dall'unione della domanda dei titolari delle carte di credito, disposti a pagare con le carte, e dei commercianti disposti ad accettarle come metodo di pagamento. Il concetto di mercato a più versanti è stato poi ripreso in un successivo studio di Rochet J-C., Tirole J., *Platform Competition in Two-Sided Markets*, 2003, *Journal of the European Economic Association*, 1:4, pp. 990-1029. Tale documento rivelò che molti business in diversi settori potevano essere analizzati attraverso l'approccio del mercato a più versanti: a titolo esemplificativo basti pensare ai sistemi operativi dei computer, alle console per video games, ai giornali, ai centri commerciali, alle carte di credito, ai locali di incontri, ecc.

[122] A complicare ulteriormente ulteriormente l'analisi, c'è anche la opposta visione di parte della dottrina economica, che sostiene come in determinati mercati caratterizzati da forti economie di scala o comunque non sufficientemente concorrenziali, come sarebbero quelli digitali *multisided* qui considerati, in realtà il prezzo ottimale non è quello prevalente, in quanto non concorrenziale. In questo caso uno SSNIP test potrebbe definire erroneamente un mercato rilevante più grande di quello effettivo anche quando quest'ultimo è un effettivo monopolio. Questa fattispecie è nota in letteratura come *Cellophane Fallacy.* Su questo tema, tra gli altri, Nelson P., *Monopoly Power, Market Definition and the Cellophane Fallacy*, U.S. Department of Justice, 2015.

enfatizzi eccessivamente l'impatto di tali effetti su un determinato gruppo di utenti, trascurando gli altri gruppi presenti[123].

Per questo motivo, ai fini di un'analisi del mercato a due versanti, è necessario esaminare gli effetti che la condotta contestata ha sulla concorrenza del mercato nel suo complesso[124]. Esaminando solo un versante del mercato si finirebbe sicuramente col distorcere il risultato di tale analisi e si potrebbe arrivare a sanzionare condotte legittime che di fatto aumentano, piuttosto che diminuire, il benessere dei consumatori[125].

Ancora, nei mercati audiovisivi che presentano tipicamente le caratteristiche del mercato multiversante, i consumatori scelgono tra i vari servizi, gratuiti o a pagamento, bilanciando tra la migliore qualità e prezzi più alti che le sono associati; in altre parole, il pagamento è fatto per evitare l'utilità negativa derivante da spot pubblicitari, o per vendere l'attenzione agli inserzionisti, facendo loro pagare per il contenuto gratuito. Se riflette una differenziazione di prodotto, un prezzo più alto viene ricaricato per una più elevata qualità e per la fruizione di un contenuto considerato di maggior valore[126].

Ne discende che una differente strategia di prezzo rende indicatori tradizionali come l'indice di Lerner, utilizzato abitualmente per misurare le quote di mercato, non validi nel contesto di mercati a due versanti. Nelle attività di impresa a due o più versanti, le esternalità indirette di rete saranno materializzate dagli operatori della piattaforma, dando quindi luogo ad un effetto redistributivo tra i

[123] In proposito, nel settembre 2016, la Corte d'Appello degli Stati Uniti ha emesso una importante decisione in cui ha riconosciuto e applicato rilevanti principi economici riguardanti l'analisi antitrust di singole condotte di una impresa nei mercati a due versanti. Il Secondo Circuito ha invertito una decisione del febbraio 2015 della Corte Distrettuale di New York, che aveva considerato le «non-discriminatory provisions (NDPs)» di Amex, delle misure che vietavano ai commercianti di fare in modo che i clienti usassero altre carte di credito con più basse commissioni, restringendo irragionevolmente il mercato e violando la sezione 1 dello Sherman Act. In particolare, la Corte Distrettuale rilevava che l'uso da parte di Amex di quelle prescrizioni, creava una situazione in cui i circuiti delle carte di credito avevano pochi incentivi ad applicare commissioni più basse, e che presumibilmente limitava la concorrenza interbrand all'interno di questi stessi circuiti. Il Secondo Circuito, tuttavia, ha stabilito che l'analisi della Corte Distrettuale fosse erroneamente concentrata su un solo lato del mercato. Di conseguenza, il Secondo Circuito ha invertito le conclusioni della Corte Distrettuale rispetto alla circostanza che Amex possedesse un significativo potere di mercato e che le sue clausole avessero un effetto negativo reale sulla concorrenza nel suo complesso. (Docket No. 15-1672 - http://www.scotusblog.com/wp-content/uploads/2017/07/16-1454-opinion-below.pdf)

[124] Filistrucchi L., *How many markets are two-sided?*, Competition Policy International - Antitrust Chronicle (2010), Vol.7; reprinted in Competition Policy International-Antitrust Chronicle (2011), Vol.1.

[125] Come ricordano Calvano e Jullien,"è necessario sviluppare test adeguati per questi mercati", in Calvano E. e Jullien B., *Issues in on-line advertising and competition policy: a two-sided market perspective*, p. 11, 2011. Per un ulteriore approfondimento del tema vedi anche Sutton J., *Sunk costs and market structure: price competition, advertising, and the evolution of concentration"*, 1991, MIT Press, Cambridge, Mass. ISBN 0-26-219305-1.

[126] Preta A., *Televisione e mercati rilevanti, Vita e Pensiero*, 2012.

diversi gruppi di agenti. In generale, i sussidi incrociati sono ampiamente utilizzati, soprattutto quando le elasticità dei prezzi differiscono tra i diversi gruppi di utenti[127].

Ad esempio, nella maggior parte delle piattaforme sostenute dalla pubblicità (tra cui i media tradizionali come i giornali, la free-to-air tv, le pagine gialle, i motori di ricerca, le piattaforme di social network e la maggior parte degli altri servizi internet) i ricavi dell'intera piattaforma provengono principalmente da un solo gruppo di utenti: gli inserzionisti. Di conseguenza, i ricavi pubblicitari vengono utilizzati per sovvenzionare i servizi per i consumatori dall'altro lato della piattaforma. La ragione è duplice: in primo luogo, la domanda degli utenti finali è generalmente più elastica, in particolare su internet esistono numerosi servizi sostituibili gratuiti, il valore aggiunto di pagare in più non è normalmente significativo (a meno che non siano sostituibili, come il contenuto premium); in secondo luogo, gli inserzionisti valorizzano la quantità di consumatori a cui possono avere accesso sulla piattaforma, la bassa elasticità dei prezzi determina un prezzo relativamente alto da parte degli inserzionisti[128].

In definitiva dal quadro qui delineato si possono trarre le prime importanti conclusioni[129]:

(i) La *definizione del mercato* diventa sempre più complessa. Ciò risulta dalla competizione tra piattaforme multi-versante, effetti di rete e sbilanciamenti - tipping - di mercato, convergenza tecnologica, etc.

(ii) Il *potere di mercato* dipende non solo dalla concorrenza a livello orizzontale. In particolare nei mercati *multi-sided* va misurato in un contesto più ampio che tenga conto dei diversi lati del mercato.

(iii) Alcuni *indicatori* usati nella misurazione del potere di mercato, in contesti di concorrenza statica sul prezzo, possono perdere di significato (e.g. l'indice di Lerner)[130]. Poiché i costi marginali dei prodotti *innovation intensive* (digitale) possono essere insignificanti, il margine di profitto elevato può essere considerato un fattore necessario per poter sostenere l'investimento.

(iv) Le prevalenti strategie di sussidio incrociato nelle *piattaforme multi-versante*, sia scelte di «prezzi sotto-costo», sia l'esistenza di un elevato margine di profitto, possono risultare ingannevoli per comprendere la sussistenza del potere di mercato.

[127] Calvano E. and Jullien B., *Issues in on-line advertising and competition policy: a two-sided market perspective*, 2011 e più di recente da Preta A., Peng P., *"Competition and innovation in digital markets"*, working paper EALE Poland, 2017

[128] Calvano E. and Jullien B., *cit.*, 2011.

[129] Per un approfondimento dei vari aspetti qui sinteticamente analizzati, vedasi ITMedia Consulting, LUISS DREAM *Il video e la terza rivoluzione di internet*, 2016 http://www.itmedia-consulting.com/DOCUMENTI/video_internet.pdf

[130] Sidak J. G. & Teece D. J., *Dynamic competition in antitrust law*, 2009.

(v) Adottare un *concetto di welfare* che consideri performance e sviluppo del settore: il surplus del consumatore si riferisce non ad un singolo gruppo di utenti, bensì alla somma del surplus di ogni agente, in ogni versante.

(vi) Infine, se l'analisi non è corretta e l'intervento è troppo vigoroso, si può avere una situazione **"***winner curse*[131]*"*, in cui il premio del vincitore, cioè *l'impresa più efficiente, viene espropriato, sanzionando condotte legittime che di fatto aumentano, piuttosto che diminuire, il benessere dei consumatori.*

(vii) L'effetto *lock-in* infine può essere ridotto da altri fattori: i costi di *switching* per il consumatore sono ridotti se non nulli, rispetto ai mercati tradizionali; prevalenza del *multi-homing*[132]; la differenziazione del prodotto (basata su discriminazione di prezzo) in mercati *two sided* può consentire di estrarre dal consumatore il valore più alto, massimizzando il welfare totale.

3.4.3.2 Prezzi zero

Una seconda criticità e ulteriore insegnamento che emerge dal modello dei mercati multiversante è che spesso essi si basano sulla fornitura di servizi gratuiti o cd. "*zero price*" su un lato del mercato. La possibilità di offrire servizi gratuiti deriva come abbiamo visto dalla capacità delle piattaforme multiversante di internalizzare gli effetti di rete diretti e indiretti e le divergenti elasticità della domanda dei loro due /o più gruppi di utenti. Infatti, le piattaforme spesso sono in grado di fornire i loro servizi a costi marginali ridotti e di sovvenzionare la loro attività imponendo prezzi più elevati sugli altri gruppi di utenti meno sensibili alle variazioni di prezzo.[133]

Anche in questo caso il fenomeno non è caratteristico dei soli mercati digitali e in particolare di quelli multiversante, visto il crescente interesse verso i *non-price effects*, come emerge anche dalle Horizontal Merger Guidelines della FTC, quando si fa espresso riferimento a "price increase or corresponding non-price change such a reduction in product quality or service"[134].

[131] Thaler R. H., *Anomalies: The Winner's Curse*, 1988, Journal of Economic Perspectives 2 (1): 191–202.

[132] Sul punto si veda in particolare: Evans D.S ,Shmalensee R., *The Antitrust Analysis of Multi-Sided Platform Businesses* in Blair R. - Sokol D. (a cura di), Oxford Handbook on International Antitrust Economics, Oxford University Press, 2013. Nel modello di concorrenza dinamica delle piattaforme multi-sided secondo i due Autori ciò che rileva è il numero di piattaforme che i singoli agenti utilizzano su ogni versante, le differenze nel numero di piattaforme usate sui due versanti e la capacità di ogni agente di un versante di determinare la scelta della piattaforma per l'altro versante. Ne discende che assume particolare rilevanza il *multi-homing*, ossia l'utilizzo da parte di un agente di più piattaforme, poiché in un contesto altamente competitivo come quello digitale ciascuna piattaforma tende a sviluppare specifiche strategie per attirare l'attenzione dei consumatori.

[133] Sul tema oltre ai numerosi lavori di Evans, vedi anche Peitz M. – Waldfogel J., *Oxford Handbook of the Digital Economy*, Oxford, 2012, e Calvano E., Jullien B. *cit*., 2011.

[134]U.S. Department of Justice and the Federal Trade Commission, *Horizontal Merger Guidelines*, 2010, pp. 10 e 11, available at https://www.ftc.gov/sites/default/files/attachments/merger-review/100819hmg.pdf

Il prezzo zero in ogni caso aggiunge un ulteriore elemento di complessità, dal momento che il metodo convenzionale di definizione dei mercati – il test del monopolista ipotetico o SSNIP test – si basa sui prezzi monetari e su un determinato aumento degli stessi (5-10%) per individuare i vincoli concorrenziali in un determinato mercato. Inoltre il prezzo non rappresenta il solo strumento analitico e talvolta neppure il più appropriato per analizzare la concorrenza nei mercati multiversante, anche laddove costituisce un valido parametro concorrenziale.

Laddove non esistono infatti beni omogenei ma caratterizzati da forte differenziazione (come ad esempio nel caso dell'informazione), la qualità diventa un ulteriore elemento nella definizione dei mercati rilevanti allo stesso modo del prezzo. Il problema però è che sebbene l'importanza della qualità sia indiscutibile e le autorità antitrust considerino la qualità rilevante allo stesso modo del prezzo, ci sono diverse ragioni per cui esse raramente analizzano gli effetti sulla qualità allo stesso modo degli effetti sui prezzi.

In primo luogo, come è ovvio, la qualità è un concetto soggettivo e conseguentemente più difficile da definire e misurare dei prezzi. In aggiunta, la teoria microeconomica offre un aiuto limitato nel prevedere quanto i cambiamenti nel livello di competizione in un mercato influenzano la qualità ed è lasciato spesso all'analisi empirica determinare quanto la qualità cambierà in risposta ai vari livelli di concorrenza nel contesto di particolari mercati. Tuttavia, ai nostri fini, sebbene le differenze nella qualità sono più difficili da misurare e quantificare dei prezzi, il comportamento del consumatore può costituire un utile indicatore sulle preferenze di quest'ultimo quando cambia la qualità[135].

Lo SSNIQ test in questo senso è dunque una variante del test di prezzo che si concentra su un'ipotetica "piccola, ma significativa e non transitoria riduzione della qualità". Esamina il cambiamento una volta che la qualità è ridotta (piuttosto che quando il prezzo è aumentato) e può includere sia l'incremento nell'efficienza dinamica che la riduzione dei costi. Laddove il consumatore paga per il bene gratuito in un'altra valuta, come l'attenzione o l'informazione, e tali costi possono essere quantificati, il test può essere riferito ai cambiamenti nei costi (SSNIC)[136].

Ne discende che se in linea di principio il test del monopolista ipotetico è realizzabile anche in mercati con prezzi non monetari, la sua applicazione in tali casi come abbiamo visto non è semplice. In molti mercati a prezzi zero, ad esempio nei mercati digitali, la qualità del prodotto viene raggiunta principalmente attraverso costi fissi di ricerca e sviluppo, mentre la differenza tra il costo marginale per la realizzazione di un prodotto di alta qualità anziché un prodotto di scarsa qualità è generalmente nulla. Di conseguenza, è improbabile che le imprese che godono di potere di mercato in almeno alcuni mercati a prezzi zero avrebbero scelto di esercitare tale potere riducendo la qualità. Dato che ciò comporterebbe una riduzione dei costi trascurabile, la perdita attiva dei clienti

[135] OECD, *The role and measurement of quality in competition analysis*, 2013, available at http://www.oecd.org/competition/Quality-in-competition-analysis-2013.pdf ; FTC, *Big Data. A Tool for Inclusion or Exclusion?*, 2016.

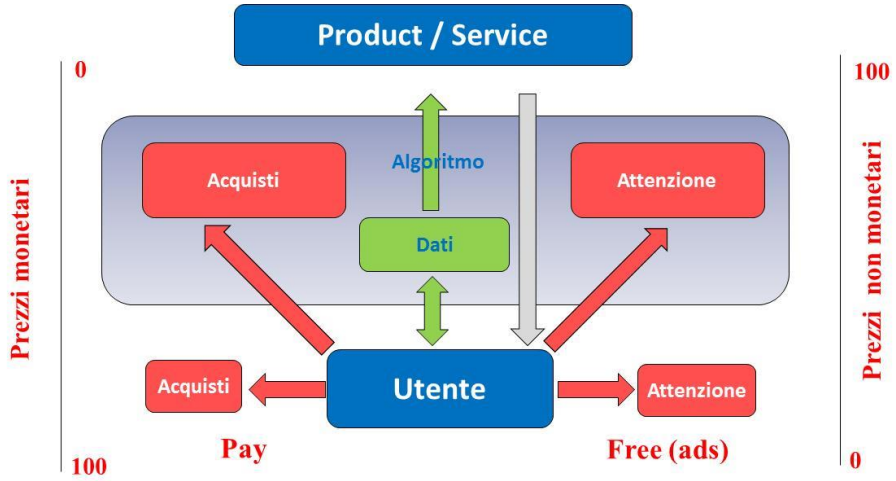
[136] Newman J., *Antitrust in Zero-Price Markets: Applications*, Washington University Law Review, Vol. 94, 49 2015.

probabilmente renderebbe un SSNIQ irrazionale. Di conseguenza, i test SSNIQ sono tendenzialmente più facili da utilizzare in mercati con prezzi monetari, in cui i costi marginali variano sostanzialmente in linea con i livelli di qualità e più complicati quando non vi sono[137].

Ciò non toglie però che, come abbiamo visto in precedenza, nei mercati finanziati dalla pubblicità e caratterizzati da una concorrenza per attenzione/base utenti, a parità di prezzo, nullo, la concorrenza sulla qualità mostra dinamiche simili a quelle della tradizionale concorrenza di prezzo: un aumento del potere di mercato (attenzione) potrebbe incentivare una riduzione della qualità (più pubblicità), sfruttando gli effetti di rete[138].

3.4.3.3 *Data driven digital markets*

Figura 813. Data driven digital markets



Fonte: ITMedia Consulting

Alla luce di ciò, ci si domanda se si può applicare la stessa logica dei "tradizionali" mercati (multiversante) a prezzo zero basati sull'attenzione ai nuovi mercati digitali basati sui dati.

Per un verso il dato non può essere considerato un bene omogeneo, ma al contrario come un prodotto differenziato[139]. Infatti la natura multidimensionale dei dati si evince dalle differenti caratteristiche degli stessi (strutturati e non strutturati) e dalla pletora di tipologie differenti, che vanno dai dati scannerizzati dei supermarket ai dati personali, ai metadata geografici. In più, come già ampiamente descritto nel primo capitolo, si possono differenziare i dati sulla base della qualità (dati grezzi rispetto ai data set), il livello in cui sono organizzati e analizzati, le informazioni che forniscono, e così via.

[137] Newman J., *Antitrust in Zero-Price Markets: Foundations*, University of Pennsylvania law review, 2015

[138] Preta A., *cit,* 18, 2012. Vedi inoltre, Polo M., *Concentration and Competition in Media Markets*, EUI, 2010.

[139] Shaked A., Sutton J., *Relaxing Price Competition through Product Differentiation*, 1982, Review of Economic Studies, Vol. 49:3-13.

In questo senso i dati, nei mercati con particolari caratteristiche come quelli legati alle piattaforme digitali *multisided*, non si sostituiscono all'attenzione o al prezzo monetario e dunque non rappresentano un ulteriore versante del mercato, ma intervengono per lo più come input per orientare le dinamiche degli stessi, su uno piuttosto che sull'altro lato del mercato.

Dato infatti un mercato del prodotto con particolari caratteristiche (multi-sided) e con prezzi monetari, caratterizzati da una differenziazione di prodotto, quali ad esempio quelli televisivi in presenza di offerte Pay TV o di Video on Demand (Netflix), una riduzione della qualità legata alla riduzione della privacy potrebbe tradursi all'atto pratico, come abbiamo visto nella prima parte, in un mark-up, in un ricarico, sulla quantità di informazione e dati raccolti e quindi anche a una possibile riduzione del prezzo (monetario) del servizio all'utente finale.

In altre parole nel mercato tv, nelle offerte con prezzo positivo (Pay TV), il pagamento è fatto per evitare l'utilità negativa derivante da spot pubblicitari, e in quelle a prezzo zero (Free TV), per vendere l'attenzione agli inserzionisti, facendo loro pagare per il contenuto gratuito[140]. Nel caso di servizi di video streaming data driven, un prezzo più alto, paragonabile a quello di un'offerta pay tradizionale, verrebbe scambiato per un maggiore prezzo non monetario, ossia una maggiore quantità/combinazione di dati e un conseguente minore livello di privacy, tale da ottenere a un prezzo monetario più basso (10 euro contro 20-30) la fruizione di un contenuto considerato di maggior valore (senza pubblicità).

Anche nei mercati a prezzi zero, la concorrenza sui dati mostrerebbe le stesse dinamiche: una riduzione della privacy dovuta al maggior numero di dati estratti e combinati avrebbe lo stesso valore dell'incremento dei prezzi nei mercati tradizionali. Qualora questo avvenga senza una riduzione dell'attenzione, e dunque della pubblicità, mantenendo lo stesso prezzo monetario (0), ciò potrebbe costituire un indicatore di potere di mercato[141].

In entrambi i casi si dovrebbe accettare l'idea di considerare esercizio del potere di mercato la duplice capacità di un'impresa di estrarre significative quantità di dati dai consumatori e, quindi, di ridurre il livello di privacy di cui essi godrebbero, pur senza violare alcuna disposizione in materia di tutela dei dati personali. Il prezzo zero nei mercati *multisided*, peraltro allo stesso modo di quanto avverrebbe nel caso del prezzo monetario ridotto, potrebbe dunque essere considerato un prezzo del monopolista.

Questo fenomeno verrebbe più facilmente misurato in presenza di fusioni o acquisizioni. Se due entità operanti su uno stesso mercato a prezzo zero si uniscono, i concorrenti dovrebbero o arrivare ad estrarre dai loro utenti le medesime quantità di dati con la conseguente riduzione della privacy dei loro consumatori, o dovrebbero pagare direttamente l'utente per avere accesso agli stessi dati per competere.

140 Calvano E., Polo M., *Strategic Differentiation by Business Models: Free-to-Air and Pay-TV,* CSEF Working Paper, 2015.

141 Deutscher E., *The role of personal data and privacy in merger analysis*, 2017.

D'altro canto però, oltre alla difficoltà di applicazione, al pari dei test di qualità visti in precedenza, ciò che rileva è in primo luogo la differente natura del dato, che non si sostituisce nei casi considerati ad alcuno dei versanti del mercato siano essi monetari o non monetari (prezzo e attenzione/audience). Ma anche, ammesso e non concesso, ci trovassimo in presenza di un mercato dei dati, in tutti i casi, al fine di considerare eventuali esercizi di potere di mercato (abusi), bisognerà valutare da un lato gli effetti di rete e l'assenza di eventuali sostituti (servizi concorrenti), e all'opposto l'assenza di costi di *switching*, il *multi-homing*, l'aumento nella qualità dei servizi per il consumatore (*learning by doing*), la non rivalità e soprattutto l'elasticità della domanda. Quest'ultima in particolare al momento appare molto elastica alla privacy[142].

Di conseguenza, anche utilizzando una diversa strumentazione, i risultati, dal punto di vista pratico, potrebbero non essere diversi da quelli di una "tradizionale" analisi antitrust. In questo senso uno dei temi chiavi rimane quello delle barriere all'accesso, così come approfondito nel paragrafo che segue.

Infine, per un approfondimento dei vari aspetti collegati al tema del rapporto tra antitrust e privacy si rimanda al paragrafo 3.7.

3.5 I mercati della generazione e raccolta dei dati digitali: l'accesso ai big data

Sono tre i quesiti sottesi al tema dell'accessibilità ai dati digitali da parte delle imprese: ci si può chiedere se l'accesso ad alcuni dati digitali possa risultare, in via di fatto o in punto di diritto, impossibile o meno; ci si può interrogare su quanto questo accesso, quando possibile, sia più o meno oneroso; e ci si può domandare quanto le imprese – purché in possesso, singolarmente o congiuntamente, di un certo, se non significativo potere di mercato – possano tenere delle condotte nella forma di intese, concentrazioni o pratiche unilaterali al fine di rendere l'accesso a certi dati digitali impossibile o, per lo meno, costoso.

Nei prossimi paragrafi tenteremo di rispondere ai suddetti interrogativi muovendo da quanto già osservato nella prima parte di questo lavoro, ossia avendo constatato che:

(i) le fonti dei dati digitali sono molteplici: oltre a tutti gli oggetti digitali, si contano internet e, in tempi più recenti, la rete delle cose;

(ii) queste sorgenti di dati digitali possono collocarsi all'interno dei più disparati settori merceologici – dall'industria metallurgica a quella della cura della persona – a dimostrazione di come i dati digitali si caratterizzino per i contenuti più disparati, personali e non[143]. Ecco perché, al pari dei dati analogici, i dati digitali non costituiscono un *unicum* in ragione del loro contenuto e, quindi, non possono fare ingresso nel discorso antitrust come se individuassero una categoria di risorse omogenee;

[142] Tucker C., *Social Networks, Personalized Advertising, and Privacy Controls*, 2011, available at http://www.econinfosec.org/archive/weis2011/papers/Social%20Networks,%20Personalized%20Advertising,%20and%20Privacy%20Cont.pdf.

[143] FTC, *Big data. A tool for inclusion or exclusion?*, 2016, p. 2.

(iii) proprio sulla scorta di quanto detto con riguardo alle fonti dei dati digitali, ben si comprende perché si è soliti parlare di pervasività dei dati digitali[144]. In altri termini, in un contesto come quello della quarta rivoluzione industriale che ha eletto il codice binario a linguaggio per la rappresentazione del mondo, nulla fa credere che questa tendenza a utilizzare l'elettronica digitale potrebbe subire un'inversione;

(iv) i costi marginali della produzione e distribuzione dei dati digitali sono quasi nulli[145] e i dati digitali sono risorse non-escludibili e non-rivali[146]. In altre parole, una volta acquisite le infrastrutture e le competenze tecniche necessarie per generarli, gestirli e analizzarli, i dati digitali possono essere prodotti e distribuiti senza che questo incrementi in modo significativo i costi di impresa. Inoltre, una volta che questi dati digitali vengono diffusi, nessuno può impedire a dei terzi di usarli e quei dati possono essere sfruttati indefinitamente e allo stesso tempo da molte imprese, senza che per questo essi perdano il loro valore;

(v) infine, la molteplicità delle fonti suggerisce come di uno stesso evento si possano fornire tante registrazioni digitali per quanti sono gli strumenti digitali che possono percepire quell'evento. Ad esempio, molte applicazioni per cellulari possono mappare simultaneamente la geolocalizzazione degli individui[147]. Questo di per sé suggerisce come il grado di sostituibilità dei set di dati – ossia il fatto che i dati non siano rivali e quindi che i singoli dati possano essere teoricamente trovati in diversi data set – dipenda dalle circostanze concrete, inclusa la natura stessa delle informazioni contenute nei dati.

In questo contesto, sembra dunque che le imprese siano nella condizione non solo di generare, ma anche di ricevere e trattenere a sé enormi volumi di dati, anche quando essi sono generati dagli individui che con quelle imprese interagiscono a qualche titolo. Ecco perché si condivide l'opinione di chi sostiene che le imprese che vogliano disporre di dati digitali non dovranno far altro che sostenere i costi fissi necessari ad acquisire le necessarie infrastrutture e competenze tecniche, ben sapendo che, invece, i costi marginali implicati dalla generazione e raccolta di quei dati saranno quasi nulli[148].

Tuttavia, si obietterà che quanto appena detto non implica che una impresa possa avere sempre e comunque accesso a qualsiasi insieme o flusso di dati. Ora, sebbene paia difficile immaginare che esistano degli input accessibili a qualsiasi impresa in ogni possibile stato del mondo, resta vero che le

[144] Tucker C., *The Implications of Improved Attribution and Measurability for Antitrust and Privacy in Online Advertising Markets*, 20 Geo. Mason L. Rev. 1025, 1030, 2013.

[145] Shapiro C., Varian H. R., *Information rules: A strategic guide to the network economy,* 1999.

[146] Varian H. R., Tucker D. S., Wellford H. B., *Big Mistakes Regarding Big Data, 2014; Lerner A. V., The Role of 'Big Data' in Online Platform Competition*, 2014, http://ssrn.com/abstract=2482780

[147] Almuhimedi H., Schaub F., Sadeh N., Adjerid I., Acquisti A., Gluck J., Cranor L., Agarwal Y., *Your location has been shared 5,398 times!: A field study on mobile app privacy nudging*, in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 787–796, 2015.

[148] Shapiro C., Varian H. R., *Information rules: A strategic guide to the network economy,* 24 (1999).

affermazioni sull'ampia disponibilità dei dati digitali meritano di essere sottoposte a un ulteriore vaglio.

Così in ciò che segue constateremo in primo luogo che le imprese possono dover sostenere dei costi per raccogliere insiemi di dati digitali – insiemi di dati che di volta in volta avranno le caratteristiche (volume, varietà, etc. etc.) che le imprese ritengono essere utili onde realizzare i beni che vogliono offrire. Se ci si dispone in questa prospettiva si procede allora descrivendo quelle che, in gergo antitrust, possono considerarsi le barriere all'accesso ai mercati della generazione e raccolta dei dati digitali, le quali poi – e come sempre – possono essere di vario tipo. Possiamo ad esempio verificare l'esistenza di barriere naturali, quando notiamo che la raccolta dei dati sulle precipitazioni di acqua piovana in una data regione dipende soltanto dall'installazione in quella stessa regione di una centralina di registrazione. Oppure, possiamo constatare l'esistenza di barriere strutturali, quando ad esempio ricordiamo che le imprese interessate a raccogliere dati digitali dovranno dotarsi delle necessarie infrastrutture e competenze e, in particolare, se vorranno agire tramite delle piattaforme online, dovranno considerare i *network effects* di cui godono le imprese loro rivali e già *incumbent*. Ancora se un'unica impresa avesse il diritto esclusivo ad esercitare un'attività per effetto della quale raccoglie alcuni tipi di dato, allora il mercato di quei dati sarebbe difeso da barriere legali. Ora, il diritto antitrust non può agire contro le barriere naturali e strutturali che proteggono un particolare mercato o che generalmente proteggono certi tipi di mercato. In altri termini, laddove delle scelte regolatorie possono prevedere una distribuzione alternativa di spese e oneri[149], il diritto antitrust non può azzerare i costi che le imprese devono sostenere per dedicarsi a una certa attività. Infatti, l'indagine circa queste barriere serve al diritto antitrust a fini descrittivi, ossia per individuare i mercati rilevanti e apprezzare il potere di mercato che le imprese in essi già presenti detengono. Per dirla con altre parole, l'individuazione delle barriere naturali e strutturali serve a comprendere quale sia la portata dei vincoli competitivi che potrebbero costringere l'azione degli *incumbents*.

In secondo luogo, discuteremo di come le imprese con i loro comportamenti contribuiscano a aumentare i costi che i loro concorrenti sostengono per ottenere dati digitali. Ebbene, purché ne ricorrano i presupposti, il diritto antitrust può intervenire in relazione a questi comportamenti. Il diritto antitrust può cioè sottoporre al vaglio le condotte imprenditoriali che impediscono ai loro concorrenti di avere accesso ad alcuni insiemi di dati, per poi giudicare se esse siano anche anticompetitive, ossia tali da ledere il *consumer welfare*. Più chiaramente, cioè, il diritto antitrust potrà vietare le pratiche che determinato l'insorgere di quelle che siamo soliti chiamare "barriere strategiche" solo quando il descritto effetto escludente prevarrà sugli eventuali effetti procompetitivi prodotti dalle medesime pratiche.

[149] Direttiva 2003/98/Ce del Parlamento Europeo e del Consiglio del 17 novembre 2003 relativa al riutilizzo dell'informazione del settore pubblico la quale impone, all'art. 3, che gli Stati membri provvedano affinché i documenti in possesso degli enti pubblici siano resi disponibili per via elettronica e resi riutilizzabili a fini commerciali o non commerciali.

3.5.1 Accesso esclusivo ai dati digitali, risorse essenziali e potere di mercato

Se i dati corrispondono alle registrazioni che un individuo fa dei fatti del mondo, allora pare possibile che quell'individuo (eventualmente un'impresa) sia l'unico ad avere accesso a dei dati quando egli risulti anche, in punto di diritto o in via di fatto, l'unico soggetto capace di assistere all'evento o di disporre della tecnologia necessaria a percepirlo. E ciò che i dati considerati siano analogici o digitali, giacché la differenza tra queste categorie di dati è di natura meramente tecnica, attenendo, come si diceva nell'introduzione, a nulla più che alle modalità (continua o discreta) secondo cui si rappresenta la realtà.

Così, la direttiva sulla Public Sector Information suggerisce che le pubbliche amministrazioni possano essere le uniche depositarie di grandi moli di dati raccolti nell'adempimento delle loro funzioni pubbliche[150]. Analogamente, ragionevolezza vuole che i dati generati dai sistemi informativi di una specifica impresa, oppure i dati prodotti dagli strumenti o dagli oggetti intelligenti di quella impresa, o ancora il flusso di dati che detta impresa eventualmente raccoglie tramite i suoi siti, siano considerati come naturale e automatico appannaggio di quella sola e unica impresa. A titolo di esemplificazione, si pensi ad un'impresa che è l'unica a possedere i dati circa gli orari di lavoro dei suoi dipendenti; oppure, si consideri il caso di una impresa che, occupandosi di servizi meteorologici, sia l'unica a registrare il volume delle precipitazioni atmosferiche che si verificano in una determinata località. O, ancora, si può valutare il caso di una impresa che, producendo frigoriferi intelligenti, sia l'unica a detenere i dati relativi al cibo immagazzinato nei suoi frigoriferi, così come l'ipotesi – tanto citata in dottrina – di un social network che risulta l'unico a raccogliere i dati sulle specifiche interazioni che gli individui hanno avuto sulla sua piattaforma[151]. Ebbene, in tutte queste e molte altre occasioni, un'impresa che fosse interessata ad avere quei particolari dati – e non altri, eventualmente con essi sostituibili – dovrebbe certamente richiederli alle amministrazioni o alle imprese depositarie degli stessi[152].

Nondimeno, come spiega la teoria della preclusione anticompetitiva, l'appena descritta circostanza non dovrebbe ingenerare allarme. *Rectius*, l'esistenza di una esclusiva di fatto – ma si supponga anche l'esistenza di una, per quanto remota, esclusiva di diritto – su dei dati digitali non dovrebbe portare a concludere né che quei dati siano, per ciò stesso, anche delle risorse importanti per la

[150] Cfr. V. Zeno Zencovich, *Uso a fini privati dei dati personali in mano pubblica*, Il diritto dell'informazione 197 (2003) e i diversi contributi sul tema raccolti in Daniela Tiscornia (a cura di), *Open data e riuso dei dati pubblici*, Informatica e diritto (2011).

[151] D. Boyd, K. Crawford, *Six provocations for big data,* 2011, http://ssrn.com/abstract=1926431; H.A. Shelanski, *Information, innovation, and competition policy for the Internet*, 161 U. Pa. L. Rev. 1663, 1679, 2013; D. Geradin, M. Kuschewsky, *Competition law and personal data: Preliminary thoughts on a complex issue,* 2013, http://ssrn.com/abstract=2216088.

[152] T. Cowen, *Big data as a competition issue: Should the EU Commission's approach be more careful?*, ENLR 14, 16 (2016), dove l'Autore ricorda che «[m]ultiple sources of data may provide multiple players with equivalent access to data as a raw material, and there might not be a competition issue if all have access to each other's sources».

produzione dei prodotti/servizi a valle, né che le imprese che vantano quell'esclusiva siano *naturaliter* in posizione dominante.

Sotto il primo profilo, quello inerente all'importanza dei dati preclusi, va innanzitutto osservato che i dati digitali in possesso di alcune imprese possono assai frequentemente vantare dei sostituti, giacché gli individui, non essendo vincolati da alcuna esclusiva con le imprese, utilizzano molti siti e strumenti in grado di raccogliere dati, determinando quello che viene di solito denominato come l'effetto multi-homing. Inoltre, questi dati non sono rivali, ossia il fatto che una impresa li collezioni non impedisce ai suoi rivali di fare altrettanto[153]. In aggiunta, come sempre la teoria della preclusione anticompetitiva ricorda, perché sussistano gli estremi dell'illiceità è necessario che il fatto di non potervi accedere incrementi in modo significativo i costi dei rivali o impedisca loro di produrre i loro beni e servizi. Per questo, il giudizio circa l'importanza di una risorsa preclusa dovrebbe essere svolto non solo in relazione all'accessibilità della stessa, ma altresì *in rapporto a ciò che con quella risorsa si vuole produrre*[154]. Insomma, il fatto che un'impresa detenga il controllo esclusivo su alcuni dati non significa – o, *rectius*, non prova di per sé – che i suoi rivali non possano accedere ad altri input ugualmente capaci di soddisfare la loro domanda. Se, ad esempio, un'impresa volesse conoscere le abitudini alimentari degli individui, non dovrebbe necessariamente accedere ai dati di un produttore di frigoriferi intelligenti, ma potrebbe rivolgersi a dei supermercati. Allo stesso modo, se Facebook esplora i suoi dati per dedurre le tendenze politiche degli iscritti, Google potrebbe fare lo stesso considerando i dati provenienti dai suoi motori di ricerca. Ancora, se una biblioteca analizza i propri dati per approssimare le preferenze di letturadei propri utenti, un'associazione di editoripotrebbe elaborare delle inferenze paragonabili interrogando i dati sulle vendite dei propri libri. In parole povere, il fatto che un'impresa vanti un'esclusiva su alcuni dati non dimostra di per sé che quei dati non ammettano sostituti e che siano, di conseguenza, necessari ai suoi concorrenti per la realizzazione di beni e servizi.

Riguardo poi all'apprezzamento del potere di mercato in rapporto all'eventualità che un'impresa controlli alcuni dati in via esclusiva, pare opportuno riprodurre qui un ragionamento solitamente fatto quando si discute del potere di mercato delle imprese che vantano di un diritto di proprietà

[153] Lerner A. V., *The Role of 'Big Data' in Online Platform Competition*, 2014, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2482780; A. Lambrecht, C.E. Tucker, Can Big Data Protect a Firm from Competition? 4-5 (2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705530, dove si legge «the unique economic characteristics of data mean that its accumulation does not, by itself, create a barrier to entry, and does not automatically endow a firm with either the incentive or the ability to foreclose rivals, expand or sustain its own monopoly, or harm competition in other ways», nonché D. Sokol, R. Comerford, Antitrust and regulating big data, 23 Geo. Mason L. Rev 1129, 1135 (2016), secondo i quali «[b]oth theory and actual cases support a finding that the characteristics of data are such that larger online firms cannot foreclose rivals from replicating the benefits of Big Data they enjoy, and that Big Data in the hands of large firms does not necessarily pose a significant antitrust risk».

[154] Cfr. Cowen, Big data, cit.,16, dove l'Autore ricorda che «[w]hether, in any individual case, data is a freely available raw material, and a commodity, or a critical component vitally necessary for competitors and the efficient functioning of technology markets, may be difficult to determine and requires close scrutiny in each case».

intellettuale su dei beni immateriali. In quel contesto, si suole dire che se il bene immateriale individua da sé un mercato rilevante, allora il diritto esclusivo su quella risorsa immateriale determina per ciò stesso l'esistenza di un monopolio di mercato a favore dell'impresa titolare di quel diritto esclusivo. Così, parlando di dati, viene spontaneo osservare che l'esclusiva su una particolare categoria di dati origina una posizione dominante – se non un vero monopolio – nel mercato della generazione e raccolta di quei dati, soltanto quando proprio quei dati, e magari dei sostituti imperfetti degli stessi, individuano da sé un mercato rilevante, non ammettendo dei veri sostituti capaci di rispondere a una specifica domanda di mercato. Ad esempio, riguardo al contesto italiano, si potrebbe ricordare il caso *Cerved*, in cui l'Autorità Garante individuò il mercato delle «informazioni commerciali per la cui generazione sono stati utilizzati i dati forniti in via istituzionale dalle camere di commercio»[155]. L'Autorità infatti scriveva che «le informazioni camerali rappresentano … un unicum e non sono perciò sostituibili con dati provenienti da altre fonti». E in tale mercato «[l]a Cerved detiene una posizione dominante sul mercato dei servizi di informazione commerciale di origine camerale in quanto è pressoché inevitabile che le sue concorrenti, le società private di informazioni, per operare efficientemente su tale mercato, acquisiscano le informazioni grezze direttamente dalla Cerved stessa. … A nulla rileva l'obiezione, avanzata dalla Cerved, che nessuna clausola di esclusiva è presente nei contratti da essa stipulati con le CCIAA. A parte la surriferita clausola contrattuale imposta dalle CCIAA agli stampatori, è indubbio, infatti, che l'attuale situazione vede la Cerved in qualità di unica controparte delle CCIAA, né è verosimile che altra impresa possa nell'immediato futuro avviare un sistema concorrente».

Allontanandosi da questo estremo, però, se un certo gruppo di dati non individua da sé un mercato, ossia se quei dati possono essere sostituiti con altri dati in rapporto allo scopo per cui sono domandati, allora bisognerebbe constatare che il controllo esclusivo su quei dati, anche quando ulteriormente schermato da dei diritti di privativa[156], non determina alcun monopolio di mercato, ma al più individua delle barriere all'ingresso in quel mercato[157]. Ad esempio, se esiste una domanda di dati circa la geolocalizzazione degli individui, diversi pacchetti di dati ottenuti tramite l'utilizzo di svariate applicazioni per cellulari potrebbero rispondere in modo tra loro fungibile a questa domanda. Pertanto, sebbene ogni impresa che abbia sviluppato l'applicazione avrà un'esclusiva, di fatto o di diritto, su quei dati, tale esclusiva non sarà di per sé sufficiente a conferire all'impresa una posizione dominante. Al più, come si accennava, l'esclusiva potrebbe rappresentare una barriera all'ingresso nei mercati della generazione e raccolta dei dati circa la geolocalizzazione degli individui.

Complessivamente, dunque, parte della dottrina che si è interessata dei big data è sembrata voler creare una relazione di causa-effetto tra la circostanza che un'impresa controlli questi input e

[155] AGCM, *A4 – Ancic/Cerved*, Provv. n. 452, 10 aprile 1992. In tema anche C.E. Mezzetti, *Dati pubblici ed abuso di posizione dominante*, Giur. It., 548 (2006).

[156] Cfr. Drexl, Designing competitive, cit..

[157] Per una ricognizione circa la nozione di barriere all'ingresso cfr. OCSE, Barriers to entry, DAF/COMP(2005)42, https://www.oecd.org/fr/daf/concurrence/abus/36997651.pdf, dove un'ampia bibliografia sul punto.

l'eventualità che tale impresa goda di potere di mercato[158]. Nondimeno, per esplorare detta questione bisogna superare la generalità della affermazione e comprendere in quali mercati un'impresa che controlla dei big data potrebbe per ciò stesso sottrarsi, in misura più o meno significativa, alla concorrenza esercitata dalle sue rivali. Nei mercati della generazione e raccolta dei dati, il fatto che una impresa disponga di big data non implica necessariamente che quell'impresa disponga di potere di mercato nel mercato o nei mercati della generazione e racolta di quei dati.

3.5.2 Le barriere all'accesso ai big data

Al pari di qualsiasi altra impresa interessata ad approvvigionarsi di input, le imprese determinate a ottenere dei dati digitali possono o decidere di integrarsi verticalmente, entrando nei mercati della generazione e raccolta dei dati o scegliere di acquisire quei dati dal mercato.

Guardando dapprima all'ipotesi dell'integrazione verticale, va da subito osservato che, chiaramente, questo tipo di strategia comporta necessariamente un costo, che gli esperti di diritto antitrust sono soliti definire "barriere all'ingresso" perché questi costi lavorano come ostacoli che i potenziali entranti devono superare per iniziare a operare in un dato mercato. Inoltre – altrettanto chiaramente – gli studiosi antitrust sanno che l'integrazione verticale è più o meno costosa secondo le caratteristiche specifiche dell'industria e dei mercati in gioco. Pertanto, discutendo della generazione dei big data, non è un caso che molti studiosi abbiano osservato che le imprese interessate debbano sopportare costi fissi per ottenere la tecnologia necessaria – sia essa il cloud computing, una rete di sensori o una piattaforma digitale. Ad esempio, un'impresa interessata a raccogliere dati sulla precipitazione atmosferica in una posizione specifica dovrà installare i sensori necessari in quella posizione, oppure un'impresa che vuole produrre per sé un flusso di dati sulla geolocalizzazione dei singoli dovrebbe sviluppare la propria domanda per dispositivi portatili e mobili.

È inoltre probabile che gli alti costi fissi insieme ai costi zero marginali implicati dalle produzioni digitali, comportino economie di scala e possibilità per quelle imprese che già operano nel mercato considerato. Pertanto, si è sostenuto con vigore che i gestori operanti nei mercati di generazione e raccolta di dati possono anche beneficiare di queste economie di scala e di portata. Vale a dire, possono trarre vantaggio dal fatto che abbiano già raccolto abbastanza clienti e abbiano differenziato abbastanza i loro beni in modo da spalmare tali costi fissi tra questi clienti e tra questi prodotti differenziati[159].

Ancora, poiché ogni mercato può essere riparato da barriere che corrispondono ai costi che le imprese devono sostenere per rispettare le norme applicabili, vero è che si è anche osservato come i mercati per la generazione di big data possano essere protetti da barriere legali rappresentate dai costi che le imprese devono sostenere per, ad esempio, adeguarsi alla disciplina a tutela dei dati

[158] Cfr. Autorité de la Concurrence, Bundeskartellamt, cit., 3, dove, discutendo di come i dati siano scambiati a fronte di servizi, si dice «[t]his has spurred new discussions about the role of data in economic relationships as well as in the application of competition law to such relationships, in particular as regards the assessment of data as a factor of market power».

[159] Rubinfeld e Gal, cit., 14-17.

personali o a tutela dei consumatori. Infine, in termini generali è sempre vero che, come si vedrà nel prossimo paragrafo 3.4.3, le imprese possono costruire delle barriere strategiche adottando dei comportamenti escludenti.

Tuttavia – per ciò che conta osservare in questa sede – l'appena svolta analisi delle barriere all'ingresso non è nuova o limitata ai mercati digitali. Questa stessa analisi potrebbe essere eseguita per le imprese che producono energia idroelettrica o plastica. Pertanto, la possibilità – o, se si vuole, la certezza – che queste barriere (naturali, strutturali, legali o anche strategiche) esistano non dovrebbe giustificare alcun allarme sulla produzione di grandi dati - almeno non una forma di paura più grave di quella che normalmente caratterizza i mercati dell'energia idroelettrica e della plastica (!).

Infine, queste barriere a protezione dei mercati nei quali si generano dati digitali non vanno confuse con la possibilità che gli stessi big data costituiscano una barriera a protezione dei mercati "a valle", dove vengono prodotti servizi e prodotti ideati con l'ausilio dell'informazione estratta da quei big data, come già spiegato nel paragrafo 3.4.2.

3.5.3 Condotte imprenditoriali che accrescono i costi all'accesso: le concentrazioni

In linea teorica, qualsiasi impresa operante in un qualunque mercato può tenere comportamenti tesi ad aumentare gli altrui costi e, tra i tanti, i costi che i suoi rivali sostengono per raccogliere dati. Discutendo dapprima di concentrazioni sottoposte al vaglio delle autorità antitrust[160], occorre osservare che nel corso dell'ultimo decennio in Europa e negli Stati Uniti si sono consumate, ad un

[160] Poiché alcuni dei prodotti e dei servizi commercializzati dalle piattaforme digitali sono offerti a prezzi nulli, può capitare che anche le imprese digitali di grande successo come WhatsApp abbiano, quantomeno inizialmente, dei fatturati piuttosto contenuti. Ad esempio, la concentrazione tra Facebook e WhatsApp di cui si dirà oltre è stata esaminata dalla Commissione Europea non perché i superasse le soglie comunitarie, ma perché essa riguardava le giurisdizioni di almeno tre Stati Membri – Cfr. CE, *Facebook/WhatsApp*, Case No COMP/M.7217, §§ 9-10, http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf. Ma se una caratteristica precipua del mondo digitale come i prodotti e i servizi offerti a prezzi nulli rende il parametro del fatturato non perfettamente rappresentativo del valore dell'impresa coinvolta nella transazione, si potrebbero adottare delle soluzioni alternative. Piuttosto che ricorrere a regole ad hoc per i mercati digitali – pare infatti difficile coniare una definizione di questi mercati – si potrebbe avallare una forma di controllo ex post delle operazioni sotto soglia – come è previsto negli Stati Uniti, dove la FTC ha infatti esaminato e vietato solo in un secondo momento la concentrazione tra PowerReviews e Bazaarvoice – cfr. infra. In alternativa, laddove si volessero conservare i vantaggi legati al sistema di notifica e valutazione preventiva, si potrebbero definire delle soglie calcolate in termini di volumi delle transazioni, oppure di valore delle acquisizioni. Non è dunque un caso se il 31 marzo 2017 il Bundesrat, il Consiglio Federale Tedesco, ha adottato il nono emendamento alla GesetzgegenWettbewerbsbeschränkungen, la legge antitrust tedesca, per introdurre una nuova soglia che estende l'obbligo di notifica anche alle operazioni di concentrazione il cui valore supera i 400 milioni di euro. Si ritiene infatti che, nonostante i fatturati contenuti delle parti, queste transazioni possano comunque avere un impatto significativo sulla concorrenza nell'industria high-tech e digitale, nonché nei comparti ad alto tasso di innovazione come quello farmaceutico - Cfr. http://kluwercompetitionlawblog.com/2017/06/07/new-merger-control-threshold-germany-beware-ongoing-transactions/

tasso peraltro crescente[161], diverse fusioni e operazioni di acquisto del controllo per effetto delle quali imprese che avevano raccolto importanti dataset sono state incorporate da altre imprese già in possesso di big data. Trova dunque conferma l'idea che anche le operazioni di concentrazione possano annoverarsi tra gli strumenti utilizzabili dalle imprese per avere accesso ai dati digitali, ossia per approvvigionarsi di un input[162]. In questo senso depone, ad esempio, la decisione *Microsoft/Yahoo*[163], relativamente alla quale la Commissione Europea ha osservato come l'autorizzata concentrazione consentisse alle due imprese di disporre di una scala di input sufficienti a contrastare la concorrenza di Google Search[164]. Ma, più in generale, in questo senso si spiega perché la Commissione Europea, non diversamente dalle autorità statunitensi, abbia sempre valutato le concentrazioni in materia di big data non solo guardando ai loro effetti coordinati, ma altresì – e come è uso fare quando si valutano i profili verticali delle concentrazioni – considerando se le concentrazioni determinassero una forma di preclusione anticompetitiva[165]. In altri termini, la Commissione ha verificato se le concentrazioni portassero le entità risultanti dalle stesse a divenire dei *gatekeepers*[166].

Tuttavia, nell'Unione Europea il "cumulo" dei dati non è mai stato giudicato motivo sufficiente per ritenere che la concentrazione portasse a una violazione dell'art. 2 del reg. 139/2004, ossia a una sostanziale riduzione della concorrenza e, in particolare, alla costituzione o al rafforzamento di

[161] Cfr. OCSE, *Data-driven*, cit., 21.

[162] Cfr. Graef, *Market definition*, cit.,476-77.

[163] CE, 18 febbraio 2010, caso M.5727, Microsoft/Yahoo! Search Business, http://ec.europa.eu/competition/mergers/cases/decisions/M5727_20100218_20310_261202_EN.pdf. Un'analisi dettagliata degli effetti sugli inserzionisti, gli utenti, gli editori e i distributori aveva infatti dimostrato che: (i) in diverse aree Microsoft e Yahoo non erano in concorrenza in misura significativa tra loro prima della transazione (ad esempio nella ad-search) e che (ii) in altre aree Microsoft e Yahoo continueranno a concorrere almeno in una certa misura post-transazione (ad esempio sull'esperienza utente). Inoltre, quando si sovrapponevano, le quote di mercato delle parti erano inferiori al 10%, ossia decisamente minori di quelle del loro maggiore.

[164] Id., §§ 140-143. Analogamente, durante le proprie indagini circa la partnership tra Microsoft e Yahoo!, il Dipartimento di giustizia statunitense accettò l'argomento delle parti secondo cui l'accumulo di dati avrebbe portato ad un aumento di efficienza e, di conseguenza, avrebbe reso l'offerta di Microsoft competitiva rispetto a quella di Google, l'impresa di maggiore successo nel mercato.

[165] I profili verticali di una concentrazione devono essere valutati quando le imprese interessate dalla operazione operano a livelli diversi della catena di approvvigionamento. Ad esempio, quando un produttore di un determinato prodotto procede ad una concentrazione con uno dei suoi distributori, oppure quando una impresa che produce a valle dei beni intermedi o finali acquista il controllo di una impresa sua fornitrice, ossia di una impresa che produce i fattori di produzione che impiega nella realizzazione dei beni poi venduti a valle – CE, Orientamenti relativi alla valutazione delle concentrazioni non orizzontali a norma del regolamento del Consiglio relativo al controllo delle concentrazioni tra imprese, GU C 265, 18 ottobre 2008, 7.

[166] Cfr. R.H. Weber, *Data Portability and Big Data Analitycs. New Competition Policy Challenges*, 23 Concorrenza e Mercato 59, 65 (2016).

posizione dominante[167]. Visto l'esito delle proprie indagini, la Commissione ha cioè sempre scartato l'eventualità che le operazioni sottoposte al suo giudizio consentissero all'impresa risultante dalla concentrazione di ottenere una forma di "controllo" su quei dati che i rivali avrebbero potuto utilizzare nei propri processi produttivi. Ad esempio, nel 2007 la Commissione ha autorizzato la concentrazione verticale *TomTom/Tele Atlas*[168], per effetto della quale TomTom acquistava una importante banca di mappe digitali, avendo dunque escluso che il famoso produttore di navigatori potesse avere un incentivo a privare i suoi concorrenti orizzontali dell'accesso al mercato a monte delle mappe digitali[169].

Analogamente, nel 2008, non divergendo dal giudizio espresso dalla FTC[170], la Commissione ha considerato l'operazione *Google/DoubleClick* compatibile con il mercato dell'Unione Europea perché

[167] L'articolo 2 del regolamento (CE) n. 139/2004 del Consiglio, del 20 gennaio 2004, relativo al controllo delle concentrazioni tra imprese (1) («regolamento sulle concentrazioni») prevede che la Commissione valuti le concentrazioni che rientrano nel campo d'applicazione del regolamento stesso per stabilire se siano compatibili o meno con il mercato comune. A tal fine la Commissione deve valutare, a norma dell'articolo 2, paragrafi 2 e 3, se una concentrazione sia atta ad ostacolare in modo significativo una concorrenza effettiva, in particolare creando o rafforzando una posizione dominante nel mercato comune o in una parte sostanziale di esso. – Regolamento (CE) n. 139/2004 del Consiglio, del 20 gennaio 2004 (GU L 24 del 29.1.2004, pag. 1).

[168] Commissione Europea, 14 maggio 2008, caso COMP/M.4854, disponibile all'indirizzo http://ec.europa.eu/competition/mergers/cases/decisions/m4854_20080514_20682_en.pdf.

[169] Nondimeno, la Commissione aveva comunque (e correttamente) verificato l'esistenza di importanti barriere all'ingresso nel mercato a monte delle mappe digitali – barriere che conferivano all'impresa risultante dall'operazione un vantaggio rispetto a rivali come Google e Microsoft i quali, nel 2007, pur potendo disporre di loro mappe digitali, non avevano chi le arricchisse di commenti e modifiche muovendosi di persona lungo i differenti percorsi. Al tempo, infatti, le osservazioni dei consumatori che avrebbero potuto impiegare le mappe di Google e Microsoft non vennero considerati una risorsa sufficiente a sostituire il lavoro degli esperti di Tele Atlas – cfr. Commissione Europea, 14 maggio 2008, caso COMP/M.4854, http://ec.europa.eu/competition/mergers/cases/decisions/m4854_20080514_20682_en.pdf.

[170] Nello stesso senso anche FTC, 20 dicembre 2007, caso 071-0170, 12-13, https://www.ftc.gov/system/files/documents/public_statements/418081/071220googledc-commstmt.pdf. Qui la FTC notò come l'operazione non sottraesse ai rivali di Google l'accesso ai dati relativi ai consumatori, giacché i concorrenti di Google disponevano già di un proprio bottino di dati da continuare ad accrescere e utilizzare nel tempo. In breve, né i dati disponibili a Google, né i dati disponibili a DoubleClick, potevano considerarsi un input essenziale per fare ingresso o operare nel mercato dell'intermediazione pubblicitaria. In commento a questo caso anche Edith Ramirez, Deconstructing the Antitrust Implications of Big Data, Keynote Remarks of FTC Chairwoman Edith Ramirez1 43rd Annual Conference on International Antitrust Law and Policy Fordham Competition Law Institute New York, NY September 22, 2016, 8, dove si legge che la decisione Google/DoubleClick «underscores the point that one firm's stores of consumer data may not be more valuable than another firm's, even though they are collected from different platforms. On the facts there, the collections of consumer data were functional substitutes for one another. As long as Google could not prevent other firms from accumulating their own consumer data, there was not any competitive concern». Probabilmente valutazioni analoghe, legate cioè all'accessibilità da parte dei concorrenti al mercato delle consumer information, sono quelle che hanno determinato la FTC a considerare compatibile con il mercato, peraltro al termine della c.d. "fase uno", la concentrazione Alliance Data System/Conversant, la quale vedeva fondersi due imprese depositarie di dati sui consumatori poi impiegati per definire le campagne di marketing. Nondimeno, diverse associazioni dei consumatori chiesero alla FTC di rivedere la sua decisione, dato che la concentrazione consentiva «the consolidation of twopowerful sets of consumer data» e per tale via poneva a

«the combination of data about searches with data about users' web surfing behaviour is already available to a number of Google's competitors today. … Competitors may also purchase data or targeting services from third parties. … Data is also available from internet service providers, which can track all of the online behaviour of their users, following them to every website they visit»[171]. In particolare, in questa decisione la Commissione, a differenza che in *TomTom/Atlas* dove definì il mercato della mappe digitali[172], abbia elaborato la suddetta analisi della preclusione senza dover definire un mercato rilevante per gli user data. Tuttavia, sebbene questa divergenza sia stata da alcuni evidenziata, non vi è nulla di eterodosso in questo approccio. In primo luogo, per prassi l'impatto competitivo delle concentrazioni si articola considerando cosa ne sarebbe delle dinamiche di mercato che riguardano una certa risorsa a operazione conclusa. Ma se la risorse non sono commercializzate già da prima dell'operazione e non saranno oggetto di commercializzazione successivamente, non ha senso svolgere la suddetta analisi comparativa[173]. In secondo luogo, poiché nel diritto antitrust quello che conta non è la forma della analisi, ma il risultato alla quale essa conduce, l'aver testato ed escluso che l'entità nascente dall'operazione Google/DoubleClick potesse avere gli incentivi e le capacità per impedire ai rivali di avere accesso agli user data è stato giustamente considerato bastevole[174].

Anche in occasione della ben più recente concentrazione *Telefonica UK/ Vodafone UK/ Everything Everywhere*, la Commissione ha autorizzato l'operazione dopo aver verificato come il tipo di informazione disponibile alle imprese coinvolte nella joint venture fosse ugualmente accessibile ai suoi concorrenti, che peraltro ne facevano già ampio uso[175].

Ancora, anche in *Facebook/WhatsApp* la Commissione ha riconosciuto la liceità dell'acquisizione, sottolineando come la presenza di numerosi e importanti operatori capaci di collezionare dati circa i consumatori avrebbe impedito a Facebook di avere il controllo del prezzo delle inserzioni pubblicitarie, anche nell'ipotesi in cui Facebook avesse utilizzato i dati di WhatsApp a fini pubblicitari. La Commissione, cioè, ha escluso che la concentrazione comportasse un danno al benessere del

---

serio rischio la privacy dei consumatori – https://www.democraticmedia.org/sites/default/files/FTCletter110514.pdf. Nell'avanzare queste richieste, però, le associazioni dei consumatori non facevano vestire alla FTC i panni dell'agenzia antitrust, ma quelli dell'autorità deputata a tutelare i consumatori e, dunque, la loro privacy.

[171] CE, 11 marzo 2008, caso COMP/M.4731, § 365.

[172] Par. 17-38 COMP/M.4854.

[173] Vedi *Orientamenti relativi alla valutazione delle concentrazioni orizzontali a norma del regolamento del Consiglio relativo al controllo delle concentrazioni tra imprese* (2004/C 31/03), par. 9. http://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52004XC0205(02)&from=EN

[174] Par. 193-230 COMP/M.4731

[175] Commissione Europea, 4 settembre 2012, caso COMP/M.6314, Par. 543 e 544.

consumatore perché «there will continue to be a large amount of Internet user data that are valuable for advertising purposes and that are not within Facebook's exclusive control»[176].

A ben vedere in relazione a questo caso non sono mancate osservazioni critiche. Nuovamente, alcuni hanno mosso dei rilievi alla scelta della Commissione di non procedere alla definizione del mercato dei c.d. user data, ossia il mercato dei dati relativi agli utenti delle due piattaforme[177]. Tuttavia, e come anzidetto, la Commissione ha correttamente escluso di definire un mercato dei dati degli utenti perché, come spiegato nella medesima decisione, né prima né dopo la concentrazione i patrimoni di dati di Facebook e WhatsApp e le loro tecniche di analisi degli stessi formavano, o avrebbero formato, oggetto di commercializzazione[178]. In altri termini, secondo gli attuali standard in materia di concorrenza, una corretta definizione del mercato rilevante richiede l'esistenza di una offerta e di una domanda di un prodotto o di un servizio[179]. Poiché i dati degli utenti costituiscono solo un prodotto intermedio e non vengono venduti o scambiati, non esiste alcuna domanda e offerta di tali dati, non è possibile determinarne la sostituibilità e, dunque, non è possibile raggrupparli in uno o più mercati rilevanti. Di conseguenza, anche un esercizio nel senso indicato dai sostenitori della riportata critica avrebbe condotto la Commissione ad osservare che nel mercato degli user datale parti non detenevano, né avrebbero detenuto, alcuna quota di mercato né prima, né dopo l'operazione[180].

Certo, per supportare la critica mossa alla Commissione si potrebbe comunque tentare di definire un mercato per i dati utilizzati internamente dalle imprese. Ad esempio, vi è chi ha proposto di attribuire all'attività di raccolta dei dati il fatturato dei beni realizzati analizzando quei dati[181]. Nondimeno, di là dal trovare azzardata questa ultima ipotesi di lavoro per il grado eccessivo di approssimazione alla quale condurrebbe, ad essa si oppone comunque un contro-argomento di natura sostanziale che si potrebbe sviluppare nel modo seguente. È ben noto che in rapporto ai casi che si sviluppano lungo "la verticale", obiettivo delle autorità antitrust è impedire che la condotta esaminata alteri il funzionamento del mercato sottraendo un input importante ai rivali dell'impresa indagata – nel caso

[176] Commissione Europea, 3 ottobre 2014, caso COMP/M.7217, §§ 188-189

[177] In scia con la dissenting opinion della Commissioner Harbour nel caso Google/DoubleClick,cfr. Pamela J. Harbour e Tara I. Koslov, *Section 2 in a Web 2.0 World: An Expanded Vision of Relevant Product Markets*, 76 Antitrust L.J. 769 (2010), le quali ritengono appropriato definire un input market per i data, separato da quello relativo ai servizi alimentati dai data (nel caso Google/DoubleClick, tale mercato, secondo gli autori, sarebbe rappresentato o dai data conseguiti tramite il motore di ricerca oppure dai data utilizzati per il behavioral advertising).

[178] Commissione Europea, caso COMP/M.7217, cit., §§ 70-72.

[179]Cfr. J. Almunia, *Competition and personal data protection. Privacy Platform event: Competition and Privacy in Markets of Data*, Brussels, 26 Novembre 2012;

[180] Cfr. D.S. Tucker, H.B. Wellford, *Big mistakes regarding big data*, Antitrust Source 4-5 (2014), dove le A. riconoscono che non si può dare un mercato dei dati se quei dati non vengono commercializzati.

[181] Analogamente, Graef, *Market definition*, cit., 492-495, la quale propone di definire un mercato potenziale per i data in aggiunta a quelli effettivi per i servizi offerti a utenti e inserzionisti.

di specie, una impresa risultante dalla concentrazione. A tal fine, nella sua indagine la Commissione ha stabilito quanto l'input oggetto dell'operazione fosse, oltre che commercialmente importante, effettivamente *precluso* ai concorrenti, per poi escludere questa evenienza giacché i rivali di Facebook e WhatsApp potevano comunque accedere a tante altri fonti di dati sugli individui da utilizzare nei loro affari. Tuttavia, laddove la Commissione avesse invece individuato un mercato degli user data, non si vede sotto quale aspetto la sua analisi sarebbe mutata: nel mercato degli user data la Commissione avrebbe infatti individuato, oltre alle quote di mercato delle parti, le quote dei rivali come Yahoo! e Apple che già avevano accesso ai dati considerati alternativi a quelli di Facebook e WhatsApp. E così, quella che si sarebbe trasformata in una concentrazione orizzontale nel mercato dei dati degli utenti, si sarebbe comunque rivelata incapace di ledere in modo sostanziale la concorrenza.

In secondo luogo, vi è chi ha evidenziato come la Commissione avrebbe sbagliato nell'escludere l'esistenza di una forma di preclusione anticompetitiva[182]. Secondo questa dottrina, infatti, l'informazione estraibile dalla combinazione dei dati di Facebook e WhatsApp sarebbe diversa dalla – non sostituibile con la – informazione derivabile dall'analisi dei dati di Apple o Yahoo!. Questa critica, dunque, ha l'indubbio merito di ricordare come il valore dei dati si radichi non nei dati stessi, ma nell'informazione che da questi dati si inferisce. Pertanto, al di là del risultato da appurarsi caso per caso, procedere nel senso indicato dalla dottrina non sarebbe scorretto, anche se esporrebbe le autorità antitrust a un grado di aleatorietà assai elevato. Infatti, se le risorse di cui occorre determinare reperibilità e sostituibilità non sono i dati, ma le informazioni estraibili dagli stessi, allora non è chiaro in che misura bisognerebbe considerare come tali informazioni dipendano anche dagli algoritmi utilizzati dalle differenti imprese. Inoltre, si potrebbe argomentare che il termine del confronto non dovrebbe essere l'informazione derivabile dai big data in quanto tale, ma l'informazione necessaria a produrre i prodotti che si vogliono realizzare.

Infine, un terzo ordine di critiche ha riguardato la teoria del danno considerata dalla Commissione in riferimento all'aggregazione dei patrimoni di dati controllati da Facebook e WhatsApp[183]. Si è cioè argomentato che la Commissione avrebbe mancato di valutare le conseguenze anticompetitive di quello che si può definire "l'effetto aggregazione". Più chiaramente, la dottrina ha evidenziato che, per effetto della combinazione dei big data di Facebook e WhatsApp, l'impresa risultante dalla concentrazione abbia acquisito il potere di derivare un numero e una qualità di informazioni sui suoi utenti superiori a quelle che le due imprese considerate separatamente avrebbero inferito lavorando singolarmente sui loro rispettivi patrimoni di dati. Inoltre, onde collocare all'interno del ragionamento antitrust questa notazione, la medesima dottrina ha aggiunto che la capacità, così acquisita, di meglio profilare i clienti abbia messo l'impresa risultante dalla concentrazione nella condizione di offrire beni e servizi meno rispettosi della privacy dei consumatori, ad evidente svantaggio della qualità di quegli stessi beni, nonché del complessivo benessere del consumatore che

[182] Cfr. Cowen, *Big Data*, cit., 22.

[183] Ad onor del vero, queste medesime critiche sono state rivolteanchealla FTC con riguardo alla concentrazione tra Google e Double Click.

da quella qualità dipende. Ebbene, nel paragrafo 3.7 si analizzerà meglio questo approccio che fa convergere il diritto antitrust e la disciplina a tutela della privacy attraverso la riflessione circa la qualità dei beni.

Da ultimo, anche guardando anche all'esperienza statunitense, si contano dei casi in cui le autorità hanno riscontrato dei profili anticompetitivi con riguardo alle operazioni di concentrazione tra imprese dotate di big data. Ad esempio, in *Google/ITA* Google, intendendo potenziare le capacità del proprio motore di ricerca, si accingeva ad acquisire una società, ITA, che aveva sviluppato un software per individuare le combinazioni di viaggio più appetibili per i consumatori, e che allo stato era impiegato da molte compagnie aeree, agenzie di viaggio e siti di ricerca verticale. In altri termini, come notava la medesima DOJ, «Google is acquiring a critical input not previously owned by a company that is a horizontal competitor to users of ITA»[184]. Per questo l'operazione venne considerata tale da poter ledere il meccanismo concorrenziale a tutto svantaggio degli operatori rivali che indicavano agli utenti possibili soluzioni di viaggio. Evidentemente, dunque, la conclusione raggiunta non aveva nulla a che fare con i dati posseduti dalle due imprese, ma con il software per la ricerca delle migliori combinazioni di viaggio – software sì considerato un input essenziale che la concentrazione avrebbe sottratto ai rivali di Google. Solo in un punto dell'accordo transattivo con cui si chiuse il caso, il DOJ osservò che «access to airlineseat and booking class information is a critical input to a P& S system. To ensure that Defendants [Google/ITA] do not restrict access to this crucial information, Section V.A prohibits Defendants from entering into agreements with an airline that restricts the airline's right to share seat and booking class information with Defendants' competitors, unless one or more airlines enter into exclusive agreements with a competitor. Subject to certain limitations, Sections V.B-C require Google to make available to [online travel Intermediaries] any seat and booking class information Defendants obtain for use in Google's new flight search service»[185].

Similmente, anche in *Bazaarvoice/Powerreviews*[186] il DOJ non ha prestato particolare attenzione ai dati delle imprese, sebbene queste ne possedessero migliaia, masi è confrontata con una "semplice" concentrazione orizzontale tra i primi due operatori nel mercato statunitense delle piattaforme che offrono «ratings and reviews» alle imprese di e-commerce. Al netto di qualsiasi valutazione circa l'enormità di dati che le due imprese avrebbero finito per condividere, la DOJ è dunque intervenuta per impedire la creazione di una posizione dominante, che per giunta sarebbe stata difesa da importanti barriere all'ingresso, in un contesto in cui le imprese che si ritenevano essere potenziali

[184]U.S. v. Google Inc. and ITA Software, Inc., Competitive Impact Statement, 8 aprile 2001, 2, https://www.justice.gov/atr/case-document/file/497671/download

[185]United States v. Google Inc, Complaint, Case No. 1:11 cv-00688, 12-13 (D.D.C. Apr. 8, 2011), https://www.justice.gov/atr/case-document/complaint-115.155 e United States v. Google Inc,Competitive Impact Statement, 1:11-cv-00688, 13-14 (D.D.C. Aug. 4, 2011),http://www.justice.gov/atr/cases/f269600/269620.pdf.

[186]US v. Bazaarvoice, Inc., Case No. 13-cv-00133 (N.D. Cal. 2014), https://www.justice.gov/atr/case-document/file/488846/download.

entranti, come Amazon, Oracle o Facebook, non erano in realtà per nulla interessate ad entrare nel mercato dei ratings e delle review[187].

E tornando all'esperienza europea, anche essa conta di una concentrazione tra imprese depositarie di big data che è stata autoritzzata previa l'adozione di impegni. Tuttavia, in *Microsoft/LinkedIn* le perplessità manifestate dalla Commissione – e che hanno poi determinato l'adozione degli impegni – nascevano non dai dati aggregati, ma dalla "tradizionale" sovraesposizione che, legandosi ad un sistema operativo dominante come Windows, LinkedIn (non diversamente da Explorer o Windows MediaPlayer) avrebbe ottenuto rispetto ai suoi concorrenti orizzontali[188]. Anche in *Microsoft/LinkedIn* la Commissione ha cioè escluso che si potesse creare alcun bottleneck nel mercato dei dati dei consumatori utilizzabili a fini pubblicitari, e ciò non solo per le limitate quote di mercato delle parti, o perché anche prima della concentrazione le parti non cedevano i loro dati a terzi, ma soprattutto perché, secondo la Commissione, molti altri dati circa i consumatori e gli utenti di internet rimanevano nel mercato a disposizione dei concorrenti di Microsoft e LinkedIn[189].

Complessivamente, dunque, l'esperienza accumulata sinora in materia di concentrazioni – esperienza che potrebbe modificarsi in altri contesti di mercato[190]e avere ad esempio ad oggetto dati che

[187]US v. Bazaarvoice, Inc., Case No. 13-cv-00133 (N.D. Cal. 2014), https://www.justice.gov/atr/case-document/file/488846/download, nonchéU.S. v. Bazaarvoice, complaint, 10 gennaio 2013, § 56, dove silegge «Bazaarvoice's syndication network is a formidable barrier to entry in the market for PRR platforms. As more manufacturers purchase Bazaarvoice's PRR platform, the Bazaarvoice network becomes more valuable to retailers because it will allow them to gain access to a greater volume of ratings and reviews. Similarly, as more retailers purchase Bazaarvoice's PRR platform, the Bazaarvoice network becomes more valuable for manufacturers because it will allow them to syndicate content to a greater number ofretail outlets. The feedback between manufacturers and retailers creates a network effect that is a significant and durable competitive advantage for Bazaarvoice».Dello stesso tenore anche il caso ReedElsevier NV, et al., File No. 081-0133 (Sept. 16, 2008), https://www.ftc.gov/sites/default/files/documents/cases/2008/09/080916reedelseviercpcmpt.pdf, dove la fusione tra i due maggiori fornitori di «electronic public record services to U.S. law enforcement customers» è stata autorizzata solo previe alcune misure di disinvestimento, perché diversamente avrebbe eliminato la concorrenza tra le due imprese – cfr. https://www.ftc.gov/news-events/press-releases/2008/09/ftc-challenges-reed-elseviers-proposed-41-billion-acquisition.

[188]CE, 21 ottobre 2016, caso M.8124, § 338, dove silegge «[a]s explained in Sections (i) and (ii) above, post-transaction, the merged entity is likely to have both the ability and the incentive to foreclose competing providers of PSN services both by (i) pre-installing a LinkedIn application on Windows PCs; and by (ii) integrating LinkedIn features into Office, while denying access to Microsoft APIs. As also explained above, each of those practices would likely lead to an increase in the size and usage of LinkedIn's PSN platform in a way that competing providers of PSN services would be unable to match».

[189] Id., §§ 179-180.

[190] Ramirez, *Deconstructing the Antitrust*, cit., 8, dove sileggeche «[w]hether there is a competitive advantage associated with access to a large volume of data will depend on the particular set of facts, including the specific market at issue. We therefore have to analyze big data issues on a case-by-case basis and be on the lookout for transactions and conduct that seem inconsistent with competition on the merits»

anziché fungere da input, fungono da beni finali[191]– ha quasi sempre escluso l'eventualità che imprese, pur depositarie di big data, riuscissero, aggregando i propri patrimoni di informazioni, a impedire che i propri concorrenti abbiano accesso a dati sufficienti per produrre beni e servizi in concorrenza con i loro. D'altro canto, resta vero che le sopra menzionate decisioninon hanno mancato di sollecitare due ordini di critiche, tra loro connesse.

3.5.4 (Segue) Condotte unilaterali: la raccolta di dati come attività economica

Come sopra ricordato, il diritto antitrust si applica ai comportamenti delle imprese. In particolare, ai sensi degli artt. 101-102 TFUE e 2-4 della legge 287/90 si definisce impresa qualsiasi entità, indipendentemente dalla natura pubblica o privata della sua proprietà e indipendentemente dalla fonte pubblica o privata dei suoi finanziamenti, che svolga un'attività economica, ossia che offra «beni o servizi *su un determinato mercato*»[192].

Ecco perché, secondo alcuni[193], la circostanza che vuole che alcune delle piattaforme operanti in internet – come Facebook e Google – non offrano sul mercato i dati da loro raccolti impedirebbe di qualificare proprio la loro attività di raccolta dati alla stregua di un'attività economica e, di conseguenza, escluderebbe che con riguardo a tale e non ad altre attività[194] queste piattaforme fossero considerate imprese ai sensi del diritto antitrust. E, probabilmente, un discorso analogo potrebbe farsi riguardo ai dati raccolti dagli oggetti intelligenti in possesso di terze parti, poiché si osserva che quei dati sono innanzitutto impiegati dai produttori di tali oggetti e non necessariamente offerti al mercato.

---

[191] Cfr. https://www.ftc.gov/news-events/press-releases/2010/05/ftc-challenges-dun-bradstreets-purchase-competing-education-data.

[192] Enfasi aggiunta. Cfr. CG, 25 ottobre 2001, causa C-475/99, Firma AmbulanzGlöckner contro LandkreisSüdwestpfalz, ECLI:EU:C:2001:577, § 19; CG, 18 giugno 1998, causa C-35/96, Commissione delle Comunità europee contro Repubblica italiana, ECLI:EU:C:1998:303, § 36; nonché CG, 11 luglio 2006, causa C-205/03 P, FederaciónEspañola de Empresas de Tecnología Sanitaria (FENIN) contro Commissione, ECLI:EU:C:2006:453, § 25. Inoltre, con riguardo alla irrilevanza della proprietà e delle fonti di finanziamento, cfr. CG, 23 aprile 1991, causa C-41/90, Klaus Höfner e Fritz Elser contro MacrotronGmbH, ECLI:EU:C:1991:161, § 21, dove si legge «nel contesto del diritto della concorrenza, […] la nozione di impresa abbraccia qualsiasi entità che esercita un' attività economica, a prescindere dallo status giuridico di detta entità e dalle sue modalità di finanziamento». La medesima affermazione è stata ripresa dalla Corte in molte sentenze, tanto da diventare una clausola quasi di stile – cfr. a titolo di esempio, CG, 12 luglio 2012, Compass-DatenbankGmbH contro RepublikÖsterreich, causa C-138/11, ECLI:EU:C:2012:449, § 18.

[193] In talsenso, Darren S. Tucker, Hill B. Wellford, *Big mistakes regarding big data*, 5 (2015), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2549044, dove silegge «[b]y definition, there can be no competition where there are and will be no sales».

[194] Con riferimento alla nozione antitrust di impresa vale il principio di divisibilità, in ragione del quale una medesima entità che svolge sia attività economiche sia attività non economiche può definirsi impresa in relazione alle prime e non alle seconde – cfr. sul punto AG Jacobs, Opinione, 28 gennaio 1999, Albany International BV contro StichtingBedrijfspensioenfondsTextielindustrie, causa C-67/96, ECLI:EU:C:1999:28, § 207.

Nondimeno, questa tesi non pare condivisibile per due ordini di motivi. In primo luogo, la nozione di attività economica per come desunta dalla giurisprudenza europea non deve essere interpretata letteralmente: quando cioè la Corte di Giustizia parla di attività rivolte al mercato non vuole espungere dall'ambito di applicazione dei diritto antitrust i beni e i servizi che non sono commercializzati, ma intende evidenziare la necessità che quei beni e servizi siano prodotti *rispondendo alla logica di mercato*[195] – quella logica di mercato che dunque giustifica un'eventuale applicazione delle norme di concorrenza[196]. Prova ne sono le attività che non appartengono all'ambito di applicazione del diritto antitrust. Ad esempio, per quanto il diritto antitrust non elegga il perseguimento del profitto a requisito essenziale dell'attività imprenditoriale[197], resta vero che non sono considerate attività economiche le attività che riflettono scopi e modi solidaristici, eminentemente perché un agente economico razionale non troverebbe conveniente svolgerle[198]. Ma le piattaforme digitali raccolgono dati nel pieno rispetto della logica di mercato[199], ossia con lo scopo di perseguire un duplice fine schiettamente commerciale[200]: estrarre conoscenza per migliorare la propria offerta di prodotti e servizi e ottenere una (ulteriore) contropartita – i dati – per i servizi da loro offerti, anche a prezzo nullo. Analogamente, gli oggetti intelligenti della rete sono pensati per restituire ai loro produttori i flussi di dati necessari a ideare servizi personalizzati – al punto che si parla dei fenomeni della *servification*[201] e dell'*everything-as-a-service*[202] – e a promuovere la

[195] Jones, Sufrin, *EU Competition Law*, cit., 127 dove, per evidenziare la differenza tra imprese e enti che perseguono un fine pubblico, si legge «entitiesengaged in economicactivities must respect the principles of competition». Inoltre, OkeogheneOdudu, The Boundaries of EC Competition Law: The Scope of Article 81 45 (2006), il quale osserva come la nozione di impresacara al diritto antitrust sia «justified by a presumption underlying the rules of the private sphere that its occupants are self-interested».

[196] P. Fattori, M. Todino, *La Disciplina della Concorrenza in Italia 63* (2010).

[197] CG, 8 novembre 1983, Cause riunite 96-102, 104, 105, 108 e 110/82, N.V. IAZ International Belgium e altri contro Commissione, ECLI:EU:C:1983:310 e CG, 30 aprile 1974, Causa 155/73, Giuseppe Sacchi, ECLI:EU:C:1974:40.

[198]Cfr. Jacobs, Opinione, CG, 21 settembre 1999, Causa C-67/96, Albany International BV controStichtingBedrijfspensioenfondsTextielindustrie, ECLI:EU:C:1999:28, § 311, il quale scrive «the Court has generally adopted a functional approach. The basic test is therefore whether the entity in question is engaged in an activity which could, at least in principle, be carried on by a private undertaking in order to make profits».

[199]Cfr. D. Sokol, RoisinComerford, *Antitrust and Regulating Big Data*, 23 Geo. Mason L. Rev. 1129, 1134 (2016), dove silegge «[t]he monetization of the data in the form of targeted advertising sales for antitrust purposes is not suspect or harmful, but rather "economically rational, profit-maximizing behavior»

[200] D.J. Solove, Privacy and power: Computer databases and metaphors for information privacy, 53 Stanford Law Review 1393, 1448 (2001), dove l'A. argomentache «under the market approach, this practice can be justified as an information trade».

[201]Cfr. Magnus Lodefalk, *Servicification of Swedish Manufacturing* (2010), http://www.kommers.se/Documents/dokumentarkiv/publikationer/2010/skriftserien/report-2010-1-servicification-of-swedish-manufacturing.pdf

[202] Cfr. per l'uso di questo termine OCSE, Data-driven, cit.,21 e 27-28.

condivisione dei beni. Di conseguenza, non vi è ragione per ritenere che questa attività di raccolta dei dati non possa formare oggetto del vaglio antitrust, ossia che non rappresenti una attività economica.

In secondo luogo, e in termini più generali, non bisogna dimenticare che il diritto antitrust si è sinora sviluppato secondo i canoni dell'interpretazione funzionale, ossia con il fine ultimo di garantirsi il più ampio ambito di applicazione possibile. Pertanto, si presume che la più autorevole dottrina non sarebbe incline ad accogliere un'interpretazione letterale della nozione di attività economica che consentisse alle imprese – siano queste piattaforme digitali o imprese attive nella rete delle cose – di eludere i divieti antitrust semplicemente evitando di commercializzare i big data da loro raccolti.

Stabilito dunque che la raccolta di dati costituisce un'attività economica ai sensi della disciplina antitrust anche quando quei dati non sono offerti al mercato, occorre ora stabilire a quali condizioni le modalità secondo le quali detta attività viene svolta potrebbero costituire una forma di abuso di sfruttamento o una specie di abuso escludente e anticompetitivo.

3.5.5 (Segue) Condotte unilaterali di sfruttamento

La lett. a) degli artt. 102 TFUE e 3 l. 287/90 vieta che un'impresa in posizione dominante imponga alle proprie controparti delle condizioni contrattuali inique[203]. Giacché la raccolta di dati avviene in ragione di un accordo raggiunto tra gli utenti di un sito e l'impresa che gestisce lo stesso, ovvero in ossequio a un contratto concluso tra il produttore di un oggetto intelligente e i suoi clienti, si potrebbe argomentare che anche questa attività sia abusiva quando i termini che ne definiscono le modalità sono tali.

Esistono solo poche decisioni, peraltro di chiarezza non proprio adamantina[204], nelle quali la Corte di Giustizia e la Commissione Europea hanno analizzato la nozione di equità. Ad esempio, si rinvengono alcune sentenze e decisioni risalenti in cui l'ingiustizia delle clausole analizzate è stata rintracciata in due fatti: nella circostanza che tali clausole non fossero funzionali al raggiungimento dello scopo dell'accordo; e nella loro natura immotivatamente restrittiva della libertà delle parti. In particolare, in *SABAM*, la Corte stabiliva che poteva costituire sfruttamento abusivo di una posizione dominante «il fatto che una società incaricata [della gestione] dei diritti d'autore … impon[esse] ai suoi membri obblighi *non indispensabili* al raggiungimento dello scopo sociale e tali quindi da *limitare in modo iniquo il libero* esercizio dei diritti d'autore»[205].

[203] Cfr. in relazione alla disciplina delle condizioni contrattuali ingiustificatamente gravose, PinarAkman, *The concept of abuse in EU competition law*, 149 (2012).

[204] Kati Cseres, *Towards a European model of economic justice: the role of competition law,* 427, in Hans W. Micklitz (a cura di), The many concepts of social justice in European private law (2011)

[205]Enfasi aggiunta. CG, 27 marzo 1974, Causa 127-73, Belgische Radio en Televisie e société belge des auteurs, compositeurs et éditeurs contro SV SABAM e NV Fonior, EU:C:1974:25, § 15. In senso conforme, poi, le successive decisioni della Commissioni inerenti la GEMA, la società tedesca per la gestione collettiva dei diritti d'autore – cfr. CE, 2 giugno 1971, 71/224/CEE, IV/26.760, GEMA; CE, 6 luglio 1972, 72/268/CEE, IV/26.760

Anche più di recente, l'elemento oggettivo di questa interpretazione – ossia l'associazione tra iniquità e l'assenza di un rapporto di mezzo a fine tra le imposte clausole contrattuali e lo scopo del contratto – è stato evidenziato, oltre che in *Tetra Pak II*[206], in *Duales System Deutschland* (*DSD*), dove la Commissione ha deciso che un'impresa in posizione dominante nel mercato tedesco per il riciclaggio degli imballaggi, titolare del marchio *Green dot*, abusava del proprio potere imponendo alle controparti un onere non necessario, ossia il versamento di una royalty per l'utilizzo del marchio, anche quando i clienti effettivamente non usufruivano dei servizi identificati da quel marchio[207]. Così, la Commissione stabiliva esplicitamente che «[s]i è in presenza di condizioni contrattuali non adeguate quando l'impresa dominante *non rispetta il principio della proporzionalità*»[208]. E, in effetti, già in *United Brands* la Corte di Giustizia aveva eletto l'assenza di proporzionalità a parametro per giudicare iniqua la scelta dell'impresa in posizione dominante di interrompere una relazione commerciale con quei clienti che avevano supportato la campagna pubblicitaria di un concorrente. La Corte, infatti, chiarì come questa reazione si fosse rivelata abusiva perché non «proporzionata alla minaccia [procurata]» una volta che questa era stata apprezzata in rapporto alla forza economica delle imprese coinvolte nella competizione[209].

Inoltre, nella prassi applicativa della Commissione si rinvengono alcuni casi nei quali il carattere iniquo delle condizioni contrattuali è ricondotto sia alla loro *opacità*, quando la mancanza di trasparenza impedisce alle controparti dell'impresa in posizione dominante di comprendere i termini effettivi della proposta commerciale loro offerta, sia alla *debolezza* psicologica, economica o finanziaria in cui quelle condizioni costringono dette controparti. In modo emblematico nel caso *Michelin II* la Commissione ha infatti giudicato iniquo un programma di sconti attuato dall'impresa francese perché, in primo luogo, esso costringeva «[i] rivenditori … nell'impossibilità di conoscere con certezza il prezzo d'acquisto finale dei pneumatici Michelin». Inoltre, i tempi in ragione dei quali

---

GEMA II, GU L 166, 24 luglio 1972, 22; CE, 4 dicembre 1981, 82/204/CEE, IV/29.971, GEMA-Satzung, GU L 094, 8 aprile 1982, 12.

[206] CE, 24 luglio 1991, 92/163/CEE, IV/31.043 – Tetra Pak II, GU L 72, 18 marzo 1992, 1, §§ 105-108, dove la Commissione evidenzia chiaramente che «al cliente vengono imposte numerose obbligazioni senza nesso con l'oggetto del contratto». La decisione è stata confermata in primo e secondo grado – cfr. T, 6 ottobre 1994, Causa T-83/91, Tetra Pak International SA contro Commissione, ECLI:EU:T:1994:246 e CG, 14 novembre 1996, Causa C-333/94 P, Tetra Pak International SA contro Commissione, ECLI:EU:C:1996:436.

[207] Diffusamente sul caso R. O'Donoghue, J. Padilla, *The Law and Economics of Article 102 TFEU 853* (2013), I quali concludono che il test sullaproporzionalità in sostanzarichiede «a balancing between the object of the contract, the terms of the contract, and the contractor's justification of those terms» – 856.

[208] CE, 20 aprile 2001, 2001/463/CE, Caso COMP D3/34493, DSD,GU L 166, 21 giugno 2001, 1, § 112. La decisione è stata poi confermata in primo e secondo grado – cfr. T, 24 maggio 2007, Causa T-151/01, DerGrünePunkt - Duales System DeutschlandGmbH contro Commissione, ECLI:EU:T:2007:154 e CG, 16 luglio 2009, Causa C-385/07 P, DerGrünePunkt - Duales System DeutschlandGmbH contro Commissione, ECLI:EU:C:2009:456.

[209] CG, 14 febbraio 1978, Causa 27/76, United Brands Company e United Brands Continentaal BV contro Commissione, ECLI:EU:C:1978:22, § 190.

erano calcolati e liquidati gli sconti obbligavano i rivenditori a operare in uno stato di pericolosa precarietà, economica e finanziaria, alla quale nessuno avrebbe soggiaciuto «in un contesto di concorrenza non falsata». Infine, la Commissione scriveva assai esplicitamente che il carattere non equo del sistema di sconti «risulta[va] non solo dal fatto che i rivenditori erano messi in una situazione di debolezza psicologica al momento del negoziato, ma anche dal fatto che non erano in grado di fondarsi, in tale occasione, su una valutazione attendibile dei loro prezzi di costo e quindi di determinare liberamente la loro strategia commerciale»[210]. Analogamente, come è stato notato in dottrina[211], seguendo questa interpretazione si sarebbero potute ravvisare delle clausole ingiustificatamente gravose nel caso *Alsatel*, che pure non è stato deciso giacché l'impresa interessata non deteneva alcuna posizione dominante. Qui però le parti, d'accordo con la Commissione, lamentavano come «le clausole imposte dall'attrice, nei contratti da essa conclusi, relativamente alla durata e al prezzo sarebbero di fatto costitutive di un abuso di posizione dominante [a causa de]il carattere indeterminato del prezzo …, il fatto che esso [fosse] fissato unilateralmente … e l'automatico rinnovo del contratto per 15 anni»[212].

Complessivamente, dunque, se la nozione di equità di cui alla lett. a) degli artt. 102 TFUE e 3 l. 287/90 cattura clausole ingiustificatamente slegate dallo scopo del contratto, inutilmente limitative dell'altrui libertà, sproporzionate, unilateralmente imposte o gravemente opache[213], esistono margini perché questa disposizione possa essere applicata con riguardo ai contratatti che governano l'attività di raccolta dei dati.

Ad esempio – in assenza di una prassi decisionale in materia si possono formulare solo ipotesi – un'impresa in posizione potrebbe tenere un comportamento iniquo se richiedesse più dati di quelli necessari all'offerta del servizio offerto, sia esso di social networking, di ricerca, o di commercio online. E ciò non tanto e non solo perché un'impresa che tratta dati personali – e solo dati personali – agisce a dispetto del principio della minimizzazione dei dati, se richiede più dati di quelli necessari all'offerta di un bene o servizio. Questo comportamento rappresenterebbe infatti una violazione della privacy che dovrebbe però essere tenuta distinta dal ragionamento antitrust, onde salvaguardare l'autonomia di quest'ultimo[214]. La iniquità di una richiesta eccessiva di dati dovrebbe

[210] CE, 20 giugno 2001, 2002/405/CE, COMP/E-2/36.041, PO — Michelin, GU L 143, 31 maggio 2002, 53, §§ 220-224.

[211]Cfr. Akman, *The concept*, cit.,155-157.

[212] CG, 5 ottobre 1988, Causa 247/86, Société alsacienne et lorraine de télécommunications et d'électronique (Alsatel) controNovasam SA, ECLI:EU:C:1988:469, §§ 9-10. Circa la natura iniqua di clausole decise unilateralmente si veda ancora il caso Michelin II, dove la Commissione con riferimento a un premio concesso ai rivenditori di pneumatici osservava «[i]l premio di progresso era non equo poiché, malgrado il suo carattere apparente di convenzione, costituiva di fatto un obbligo di aumento progressivo degli pneumatici Michelin sul mercato, imposto unilateralmente dal fabbricante» – cfr. COMP/E-2/36.041, PO — Michelin, § 265.

[213]Cfr. Akman, *The concept*, cit., 156 e 157.

[214] Cfr. Art. 7, comma 4, del Regolamento (UE) 2016/679, 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la

invece appurarsi in ragione di quanto stabilito dalla giurisprudenza in relazione al principio di proporzionalità e al rapporto di causa-effetto che dovrebbe sussistere tra clausole imposte e obiettivo del contratto, ancorché si potrebbe obiettare, come si dirà in seguito, che un'impresa in posizione dominante esercita parte del suo potere proprio estraendo una quantità sovra-concorrenziale di dati[215].

Ancora, in dottrina si è sostenuto che dovrebbe considerarsi iniqua ai sensi della disciplina dell'abuso di posizione dominante tanto l'assenza di trasparenza circa i termini dello scambio per effetto del quale gli utenti delle piattaforme digitali ottengono beni e servizi a fronte dei loro dati, quanto la difficile intellegibilità di questi termini. A causa di detta opacità, i consumatori sarebbero infatti impossibilitati a comprendere la quantità e il tipo di informazioni esattamente trasferiti alle imprese e, più in generale, le caratteristiche delle transazioni così concluse finirebbero per essere state unilateralmente definite dalle imprese, come se queste avessero proposto delle offerte "prendere o lasciare"[216]. Ebbene, ricordato che in questa direzione si sta indirizzando il procedimento *Facebook* avviato dal Bundeskartellamt[217], sembra esistano margini per perseguire queste clausole come delle violazioni antitrust alla luce della precedente giurisprudenza sulla iniquità delle clausole unilateralmente imposte ed opache.

Ad ogni modo, di là dalle ipotesi che si possono formulare, pare importante che, nell'esprimere il proprio giudizio sulla natura anticoncorrenziale dei termini contrattuali adottati da una piattaforma, un'autorità antitrust mantenga una autonomia di giudizio, ispirandosi ai principi che informano il diritto della concorrenza e quindi senza mutuare orientamenti e valutazioni proprie di altre discipline, come quella a protezione dei dati personali e della privacy[218].

---

direttiva 95/46/CE (regolamento generale sulla protezione dei dati), GU L 119, 4 maggio 2016, 1. La disposizione infatti recita che «Nel valutare se il consenso sia stato liberamente prestato, si tiene nella massima considerazione l'eventualità, tra le altre, che l'esecuzione di un contratto, compresa la prestazione di un servizio, sia condizionata alla prestazione del consenso al trattamento di dati personali non necessario all'esecuzione di tale contratto».

[215]Cfr. quarto capitol, § 4.2.

[216]Cfr. I. Graef, *Market definition and market power in data: the case of online platforms*, 38 World Competition 473, 490 (2015), dove si legge«[c]ontrary to usual economic transactions, users as suppliers of data cannot determine the amount and type of information they want to supply and do not have any influence on what they will get in return. Instead, the providers of the online services unilaterally decide what type and which amount of data will be retrieved and impose their practices on users as a take-it-or-leave-it offer». Sembrano avere una posizione affine anche A. Mantelero, *Competitive value of data protection*, 3 International Data Privacy Journal 229, 232 (2013), e P.M. Schwartz, *Property, Privacy, and Personal Data*, 117 Harvard Law Review 2056, 2081(2004).

[217] Comunicato stampa, http://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_Facebook.html?nn=3591568

[218]Cfr. H. Kalimo, K. Majcher, The *concept of fairness: linking EU Competition and data protection law in the digital market place*, 42 E.L.Rev. 210, 226, 2017.

3.5.6 (Segue) Condotte unilaterali che escludono i concorrenti dall'accesso ai dati digitali

Una condotta, anche diversa da quelle previste alle lett. da b) a d) degli artt. 102 e 3, si dice escludente e anticompetitiva quando rafforza il potere di mercato dell'impresa in posizione dominante e lede il benessere dei consumatori. E ciò perché si riconosce che un'impresa, pur in posizione dominante, che vede crescere, consolidarsi e perdurare il suo potere grazie a condotte che incrementano la quantità, la qualità o il grado di innovazione dei beni disponibili sul mercato non può essere considerata contraria al normale funzionamento del mercato, il quale opera come un meccanismo selettivo a tutela delle imprese meritorie.

Ora, si è già detto di come l'attività di generazione e raccolta di dati digitali effettuata dalle piattaforme che operano in internet e dai produttori di oggetti intelligenti non necessariamente impedisca ai loro rivali di dotarsi di dati sostituibili.

Nondimeno, anche volendo supporre, per amor di argomento, che questa attività venga qualificata come preclusiva, non pare corretto sostenere che essa sia, inoltre, anticompetitiva, ossia tale da ledere il benessere del consumatore. E ciò perché, come si diceva nel primo capitolo, tramite i dati raccolti le imprese acquistano una conoscenza utile a migliorare i propri beni e servizi, nonché a offrire al mercato informazioni sotto forma di raccomandazioni, recensioni, o risposte a eventuali ricerche. Ad esempio, i motori di ricerca possono utilizzare i dati per prevedere l'andamento dei prezzi dei voli; Amazon e molteplici altri siti di e-commerce utilizzano i dati sugli acquisti passati per formulare raccomandazioni commerciali personalizzate; le piattaforme di social networking utilizzano i dati raccolti da parte degli utenti per suggerire amici, celebrità, pagine aziendali, articoli ai quali gli utenti potrebbero essere interessati; e i media online utilizzano la cronologia di navigazione e le informazioni personali per raccomandare altri articoli[219]. In altri termini, è opinione diffusa in dottrina[220] e in giurisprudenza[221] che i prodotti e i servizi ideati grazie all'analisi dei dati siano apprezzati dai consumatori i quali, peraltro, li ottengono a prezzo nullo.

[219]Cfr. D.D. Sokol, R. Comerford, *Antitrust and Regulating Big Data*, 23 Geo. Mason L. Rev. 1129, 1135 (2016); A. Acquisti, H.R. Varian, Conditioning Prices on Purchase History, 24 Marketing Sci. 367, 368 (2005); A. Goldfarb, C. Tucker, *Privacy Regulation and Online Advertising,* 57 Mgmt. Sci. 57, 61 (2011); R.H. Bork, J.G. Sidak, What *Does the Chicago School Teach About Internet Search and the Antitrust Treatment of Google?*, 8 J. Competition L. & Econ. 663, 667 (2012); G. Manne, J. Wright, *Google and the Limits of Antitrust: The Case against the Case Against Google*, 34 Harv. J.L. & Pub. Pol'y 171, 244 (2011). Per una differente opinione si veda invece N. Newman, *Search, Antitrust, and the Economics of the Control of User Data*, 31 Yale Journal on Regulation 401, 409-410 (2014).

[220]Si vedano, oltre agli autori citati alla precedente nota, gli scritti di M.A. Salinger, R. J. Levinson, *Economics and the FTC's Google Investigation*, 46 Rev. Industrial Org. 25, 47 (2015); P. Simon, *How to Get Over Your Inaction on Big Data*, Harv. Bus. Rev. (2014), https://hbr.org/2014/02/how-to-get-over-your-inaction-on-big-data-2; Russell Walker, *From Big Data To Big Profits: Success With Data And Analytics* 141-142 (2015); A.S. Edlin, R.G. Harris*, The Role of Switching Costs in Antitrust Analysis: A Comparison of Microsoft and Google*, 15 Yale J.L. & Tech. 169, 177 (2013); T. Körber, *Common Errors Regarding Search Engine Regulation – And How to Avoid Them*, 36 European Competition Law Review 239, 240 (2015); Lerner, The role, cit.,10-12. In particolare,

Nondimeno, vi è chi associa anche alla raccolta dei dati degli effetti anticompetitivi.

Alcuni ritengono che tali effetti siano riconducibili alla pubblicità di cui gli individui diventerebbero destinatari. Tuttavia, da un lato non c'è motivo per credere che la pubblicità targettizzata sia più fastidiosa di quella generalista[222]; d'altro canto, se si ritiene che questa pubblicità sia più persuasiva[223], bisognerà poi sostenere che l'influenza pubblicitaria sia in sé dannosa, oltre ad argomentare che una pubblicità più efficace sia peggiore di una pubblicità meno performante.

Altri valutano che l'attività di raccolta dei dati produca effetti anticompetitivi giacché riduce la privacy di cui godono gli individui. Nondimeno, questa eventualità si consuma se le modalità di raccolta non si sono adeguate alle regole che proteggono il trasferimento dei dati personali. Diversamente, bisognerebbe argomentare che ad essere anticompetitivo sarebbe l'effetto di aggregazione che si produce a vantaggio delle imprese che raccolgono dati. Ma, come si vedrà nel sesto capitolo, questa rappresenta una nuova teoria del danno competitivo che merita di essere esaminata in via specifica.

3.6 L'uso anticompetitivo dei big data

Le imprese possono fare un uso anticompetitivo dei risultati che l'analisi dei big data permette di inferire o, per lo meno, questa è la riflessione che al momento alimenta il dibattito internazionale. In particolare, si discute dell'eventualità che i big data facilitino la collusione tra le imprese e che i big data permettano alle imprese in posizione dominante di operare forme anticompetitive di

---

l'A. da ultimo citato spiegabrillantementeche «the ability to monetize effectively creates incentives to attract users by offering high-quality services at low or even zero prices. This is because the greater the advertising revenues that a provider can earn from each user, the greater the benefits of attracting and retaining users. Thus, the ability to earn greater advertising revenues enhances competition for users, creating incentives for providers to invest in improving the quality of services offered and to offer those services to users at low or zero prices. The targeting of advertisements, including through the collection and utilization of user data, therefore generates huge consumer benefits. Conversely, restricting the ability of online providers to collect and utilize data from users to target ads would inhibit competition for users and lead higher quality-adjusted prices for online services» – a 15.

[221] Cfr., a titolo di esempio, T, 11 dicembre 2013, Cisco Systems, Inc. e Messagenet SpA contro Commissione europea, causa T-79/12, ECLI:EU:T:2013:635 eCE, 3 ottobre 2014, caso COMP/M.7217, §§ 47, 73 e 90.

[222]Cfr. Kyle Bagwell, *The Economic Analysis of Advertising* 1706, Mark Armstrong, Robert Porter (a cura di), Handbook of Industrial Organization (2007), dove infattisilegge «when a firm advertises, consumers receive at low cost additional direct (prices, location) and/or indirect (the firm is willing to spend on advertising) information. The firm's demand curve becomes more elastic, and advertising thus promotes competition among established firms. As well, advertising can facilitate entry, as it provides a means through which a new entrant can publicize its existence, prices and products. The suggestion here, then, is that advertising can have important pro-competitive effects».

[223]Cfr. sulpunto Nathan Newman, *The Costs of Lost Privacy: Consumer Harm and Rising Economic Inequality in the Age of Google*, 40 William Mitchell L. Rev. 849, 857 (2014), dove l'A. considerandoilcasodeimutui subprime mortgages, osserva «Google's profiling of its users for advertisers allows the kind of predatory marketing we saw in the subprime housing bubble globally and in a range of other sectors».

discriminazione di prezzo. Inoltre, acceso è il dibattito circa l'eventualità che un'impresa in posizione dominante violi il diritto antitrust rifiutandosi di condividere i propri big data. Di seguito, dunque, procederemo con questa analisi.

3.6.1 I big data e la collusione

3.6.1.1 La nozione di intesa (cenni)

Come noto, due o più imprese danno luogo a una intesa ogniqualvolta sostituiscono una forma di coordinamento a quello che dovrebbe essere un agire autonomo e indipendente. In particolare, poi, tale coordinamento può assumere una forma esplicita e tangibile, come nei casi di accordo, oppure una forma assai più rarefatta, come nelle ipotesi di pratiche concordate, le quali si rintracciano anche in presenza di meri scambi di informazioni strategiche, giacché si presume che l'informazione così scambiata verrà utilizzata internamente dalle imprese per decidere la propria azione ed eventualmente assumere condotte parallele. Se dunque esiste una intesa quando le imprese raggiungono un "meeting of minds", occorre comprendere in che misura i big data possano incidere su questo fenomeno.

A tal fine, si considereranno gli algoritmi che vengono utilizzati per estrarre informazioni dai big data e come questi algoritmi, per un verso, incidono sulla struttura dei mercati e, per altro verso, supportano o plasmano i comportamenti delle imprese.

3.6.1.2 Gli algoritmi e le condizioni strutturali che favoriscono la collusione

Secondo la teoria economica, la collusione è più semplice da attuare quando si danno alcune condizioni: ad esempio, quando la domanda di mercato non è stabile nel tempo e, dunque, facilmente prevedibile; quando nel mercato operano poche imprese dalle dimensioni, capacità e costi simili; quando si danno alte barriere all'ingresso; quando il mercato è trasparente e dunque ciascuna impresa può osservare i comportamenti dei suoi rivali; nonché quando i beni e servizi scambiati sono omogenei. Di conseguenza, un regolatore – ma non un'autorità antitrust che, come ricordato nel secondo capitolo, non interviene sulla struttura dei mercato – potrebbe domandarsi se la raccolta, l'analisi e l'uso dei big data possa incidere sulle predette variabili, andando a mutare la fisionomia dei mercati e rendendoli più inclini ad accogliere forme di collusione.

Ebbene, dapprima occorre osservare che, come sostenuto sin dal primo capitolo di questo scritto, l'uso di big data e algoritmi dovrebbe condurre le imprese a meglio e più innovare, nonché a differenziare la loro offerta in ragione delle preferenze dei singoli consumatori. E questo dovrebbe rendere più complessa e improbabile la collusione, giacché le divergenze su base persino individuale di beni e servizi dovrebbero quasi impedire che si individui una offerta omogenea su cui tutte le imprese dovrebbero poi convergere.

Inoltre, se la diffusione degli algoritmi non dovrebbe incidere sull'andamento della domanda,resta comunque incerto quale effetto potrebbe produrre sul numero delle imprese presenti nel mercato e sul livello delle barriere lì esistenti[224].Non è ancora chiaro cioè come gli algoritmi possano influenzare

[224] Cfr. OECD, *Algorithms and Collusion*, cit., §43.

queste due caratteristiche strutturali. Per un verso, si constata un dato empirico: al momento gli algoritmi impiegati per segmentare i consumatori, migliorare la qualità dei prodotti/servizi offerti, determinare le dinamiche di prezzo, sono utilizzati in mercati che contano pochi rivali – come quelli dei motori di ricerca, delle agenzie di prenotazione viaggi, delle compagnie aeree, del trasporto stradale e dei *social networking* – nonché in mercati caratterizzati da alte barriere all'ingresso, quali economie di scala, economie di scopo e effetti di rete. Nondimeno, sotto il profilo teorico resta difficile stabilire se gli algoritmi siano la causa o l'effetto di questi effetti.

Piuttosto, si può e si deve affermare che il ricorso sempre più comune agli algoritmi rende il numero dei concorrenti attivi nel mercato un fattore meno rilevante per la collusione. Infatti, se nei mercati tradizionali il ridotto numero di concorrenti facilita la collusione perché semplifica le operazioni necessarie al suo compimento, ossia l'individuazione dei termini del comune coordinamento, il monitoraggio di eventuali deviazioni e l'applicazione di meccanismi sanzionatori, nei mercati nei quali le imprese fanno un uso pervasivo degli algoritmi quelle stesse operazioni diventano meno costose e rischiose, per quanto grande possa essere il numero degli agenti economici attivi nel mercato. In altri termini, la potenza computazionale degli algoritmi consente di identificare l'equilibrio collusivo e di monitorarne il rispetto, per quante numerose posso essere le imprese coinvolte nel mercato.

Infine, ha trovato autorevole sostegno l'idea secondo cui gli algoritmi favorirebbero la collusione aumentando la trasparenza del mercato e aiutando le imprese a più velocemente reagire alle altrui scelte commerciali[225]. Sotto il primo profilo, si sostiene infatti che la crescente tendenza delle imprese a raccogliere e analizzare dati – ed, eventualmente, la decisione di alcune di esse di condividere i propri big data[226] – potrebbe condurre a un mercato *asimmetricamente trasparente*, ossia a un mercato in cui tutte le condotte delle imprese diventano intellegibili per tutti coloro che dispongono di big data, vale a dire per le imprese stesse. Questa esclusione dei consumatori dal novero dei beneficiari della trasparenza avrebbe poi con l'effetto – certo, da non considerarsi scontato[227] – di rendere più semplice la collusione, giacché faciliterebbe l'individuazione del prezzo di

[225] *Cfr. Bundeskartellamt, Autorité de la concurrence, Competition law and data (2016),* https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.html

[226] Cfr. Maurice E. Stucke, Ariel Ezrachi, *Artificial Intelligence & Collusion: When Computers Inhibit Competition* (2015) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2591874 e Id., Virtual competition, cit..

[227] La trasparenza dei mercati produce infatti effetti ambivalenti. Se, come detto nel testo, essa può rendere più probabile la collusione, per altro verso essa offre alle imprese rivali l'opportunità di fare benchmarking, ossia di meglio competere cercando di battere le offerte dei concorrenti per prezzo e qualità. In particolare, lontani dalle ipotesi spurie del hub and spoke (cfr. infra), la mera trasparenza c.d. verticale, ossia l'eventualità che soggetti che si collocano a differenti livelli della filiera produttivo-distributiva dispongano di informazioni circa i loro fornitori e clienti, produce efficienza, perché consente a questi agenti economici di prendere decisioni fondate, capaci di minimizzare i rischi e comportamenti opportunistici che si possono configurare nelle relazioni verticali. Cfr. diffusamemte sul punto Simonetta Vezzoso, Competition Policy in a World of Big Data (2016), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2717497. Diversamente, la trasparenza c.d. orizzontale produce effetti ambivalenti – cfr. Florian Wagner-von Papp, *Information share Agreements* 130, Damien Geradin, Ioannis Lianos (a cura di), *Research Handbook on EU Antitrust Law* (2013); Matthew Bennett, Philip Collins, *The law and economics of information sharing: the good, the bad and the ugly*, European Competition Journal 311 (2010); e Viktoria H.S.E. Robertson, *A counterfactual on Information Sharing: The Commission's Horizontal Guidelines* 2011 Applied to the Leading Cases, World Competition 459, 488 (2013).

cartello, nonché il controllo e la punizione di eventuali deviazioni dallo stesso, all'insaputa dei medesimi consumatori. Sotto il secondo profilo, invece, si evidenzia come l'avvento dell'economia digitale abbia rivoluzionato la velocità alle quali le imprese possono assumere decisioni di business. A differenza dei mercati tradizionali in cui le rettifiche dei prezzi sono tendenzialmente operazioni costose e richiedono determinate tempistiche di attuazione, i prezzi dei mercati on-line possono in linea di principio essere cambiati di frequente a seconda delle necessità dell'operatore. Questo semplificherebbe dunque la collusione, abbattendo i costi di quelle azioni di ritorsione resesi necessarie nei confronti delle imprese che hanno deviato dall'equilibrio collusivo.

Complessivamente, dunque, non pare possibile stabilire se l'uso di algoritmi chiamati a lavorare sui big data renderà i mercati più proni alla collusione o meno. E ciò anche perché la tecnica attuale che, come si diceva, si ritiene per natura neutrale consente anche di elaborare algoritmi per potenziare le capacità di scelta dei consumatori e le abilità investigative delle autorità antitrust. I primi algoritmi – conosciuti anche come "maggiordomi digitali" – permettono, persino comunicando direttamente con gli algoritmi delle imprese ed eseguendo "in prima persona" le operazioni di acquisto, di individuare le esigenze da soddisfare e di confrontare e prevedere i prezzi e le qualità dei beni atti a soddisfarle, così da aiutarei consumatori a prendere decisioni meglio informate, avendo peraltro ridotto i loro costi di ricerca e transazione[228]. I secondi, ossia gli algoritmi investigativi, monitorando i comportamenti delle imprese e gli andamenti di mercato, possono riconoscere gli equilibri di collusione.

Al momento, però, soprattutto questi ultimi strumenti investigativi non paiono ancora sviluppati, né diffusi, ragion per cui sembrano altri due i profili che meritano di essere approfonditi: in primo luogo, occorre guardare a se e come gli algoritmi potrebbero consentire alle imprese di meglio coordinare i propri comportamenti e, in secondo luogo, serve verificare se l'attuale nozione di intesa sia in grado di catturare queste – per alcuni, nuove – forme di coordinamento delle condotte imprenditoriali.

3.6.1.3 Gli algoritmi e i comportamenti delle imprese

Proprio perché gli algoritmi consistono in una serie di comandi necessari al raggiungimento di determinati fini, le imprese possono congegnare i più diversi algoritmi per perseguire i più disparati obiettivi o svolgere le più differenti funzioni. Di conseguenza, non è possibile – e non sarebbe neanche corretto – qualificare in termini antitrust la classe degli algoritmi in sé considerata. Occorre, piuttosto, operare dei distinguo che, a ben vedere e inevitabilmente, scontano la contingenza del momento[229].

Inoltre, si consideri l'assai esaustiva comunicazione in materia anche di scambi di informazione – cfr. CE, *Linee direttrici sull'applicabilità dell'articolo 101 del trattato sul funzionamento dell'Unione europea agli accordi di cooperazione orizzontale*, GU C 11, 14 gennaio 2011, 1.

[228] Cfr. Michal S. Gal, Niva Elkin-Koren, *Algorithmic Consumers*, 30 Harvard Journal of Law and Technology 309 (2017).

[229] In tal senso si sono comportati anche i primi che in dottrina hanno studiato il tema – cfr. Ariel Ezrachi, *The Competitive Effects of Parity Clauses on Online Commerce* (2015), https://ssrn.com/abstract=2672541; Stucke, Ezrachi, *Virtual Competition*, cit.; Id., Two Artificial Neural Networks Meet in an Online Hub and Change the

Ad esempio, si può riconoscere che se due o più imprese dovessero condividere degli algoritmi per la determinazione dei loro prezzi, non sarebbe affatto eterodosso considerare questa pratica alla stregua di un cartello[230]. In altri termini, se *tradizionalmente* le imprese violano gli artt. 101 e 2 quando condividono le modalità attraverso le quali calcolano i loro prezzi[231], pare ragionevole ritenere che il medesimo illecito si realizzi quando due o più imprese pongono in comune i propri pricing algorithms[232], che in fondo non sono altro che le istruzioni per la determinazione del prezzo. In tal senso, depone peraltro il recente caso statunitense *David Topkins* dove il Dipartimento di Giustizia ha contestato l'esistenza di un cartello a un individuo che, operando su Amazon, ha progettato e condiviso con altri venditori i rispettivi algoritmi di prezzi dinamici, programmati per agire conformemente al loro accordo[233].

Analogamente, se un'impresa svelasse il suo algoritmo ai concorrenti si potrebbe configurare una intesa qualora essi non se ne dissocino pubblicamente. In altri termini, lo scenario in esame potrebbe ricondursi a quello di un invito a colludere[234] che, non essendo stato espressamente rifiutato, integra gli estremi di uno scambio di informazioni strategiche e dunque di una pratica concordata. E ciò *a fortiori* qualora le imprese che hanno ricevuto l'algoritmo abbiano poi tenuto comportamenti paralleli in esecuzione alle informazioni così circolate [235].

Future (of Competition, Market Dynamics and Society) (2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2949434.

[230] Per l'ipotesi relativa al caso del c.d. hub and spoke, cfr. paragrafo 1.4.

[231] Cfr. T, case T-48/00,Corus, ECLI:EU:T:2004:219, § 82.

[232] Cfr. Chawla, Hartline, Kleinberg, *Algorithmic Pricing*, cit.

[233] Cfr. DOJ, *Former E-Commerce Executive Charged with Price Fixing in the Antitrust Division's First Online Marketplace Prosecution* (2015), https://www.justice.gov/opa/pr/former-e-commerce-executive-charged-price-fixing-antitrust-divisions-first-online-marketplace; Le Chen, Alan Mislove, Christo Wilson, An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace (2016), http://www.ccs.neu.edu/home/amislove/publications/Amazon-WWW.pdf.

[234] Negli Stati Uniti l'invito a colludere, evidentemente un atto unilaterale che sfugge alla Sec. 1 dello Sherman Act, può essere perseguito agendo ai sensi della Sec. 5 del FTC Act. Più esattamente, affinché la fattispecie sia integrata, evidentemente in assenza della prova di un accordo, la FTC deve mostrare o che l'impresa sotto indagine aveva intenzione di agire al fine di ridurre la concorrenza sui prezzi o che essa non può produrre alcuna giustificazione oggettiva per il comportamento tenuto – cfr. E. I. du Pont de Nemours & Co. v. F.T.C., 729 F.2d 128 (2d Cir. 1984). Di conseguenza, l'impresa sotto indagine può considerarsi autrice dell'illecito se ha agito con l'intento di realizzare un equilibrio collusivo o se era consapevole che questo sarebbe stato l'esito del suo comportmanto.

[235] Non a caso l'OCSE riporta una celeberrima frase di Posner, secondo il quale «[i]f a firm raises price in the expectation that its competitors will do likewise, and they do, the firm's behaviour can be conceptualized as the offer of a unilateral contract that the offerees accept by raising their prices». Questa affermazione però si riferisce alla nozione di concerted action, la quale richiama la nostra precedente concezione di pratica concordata che riconosceva nei comportamenti paralleli una componente costitutiva – Cfr. In Re High Fructose Corn Syrup Antitrust Litigation Appeal of A & W Bottling Inc. et al., 295 F3d 651, 653 (7th Circ. 2002).

Di là da queste ipotesi di condivisione delle istruzioni per calcolare i prezzi, è poi anche possibile che le imprese utilizzino gli algoritmi – e, in particolare, i c.d. algoritmi di segnalazione – per svolgere le trattative necessarie al raggiungimento di un equilibrio collusivo. Detti algoritmi, infatti, inviano continuamente offerte per aumentare i prezzi e, allo stesso tempo, controllano i segnali inviati dagli altri concorrenti; quando tutti gli algoritmi inviano lo stesso segnale, allora è come se tra loro fosse intervenuto un accordo e le imprese fissano il prezzo così concordato fino a quando non viene avviata una nuova trattativa per un altro prezzo di cartello. Ebbene, anche in questo scenario sembrano sussistere proprio tutti gli estremi perché si configuri una ipotesi di cartello. Del resto, se la segnalazione reciproca di ipotesi di prezzo costituisce un illecito ai sensi dell'art. 101 quando realizzata off-line, come ribadito dalla Commissione Europea nel recente caso *Container Shipping*[236], non si intravedono ragioni per cui la medesima condotta non dovrebbe essere qualificata come una intesa quando eseguita online, per mezzo di macchine e strumenti automatici.

In modo equivalente, è anche possibile che gli algoritmi vengano impiegati dalle imprese per dare esecuzione alle proprie decisioni, che queste siano di natura collusiva o meno[237]. Per ciò che qui più interessa, potrebbe cioè accadere che due o più imprese scelgano di utilizzare degli algoritmi – denominati anche messengers o algoritmi di monitoraggio[238] – per verificare l'esecuzione di un accordo di cartello e sviluppare forme di rappresaglia in caso di eventuali deviazioni. Di conseguenza, giacché l'uso degli algoritmi sarebbe così finalizzato ad eseguire la congiunta e convergente volontà collusiva delle imprese, di nuovo pare corretto che le autorità antitrust procedano in modo *tradizionale*, qualificando l'operare dell'algoritmo come una pratica strumentale e accessoria alla violazione principale. Più esplicitamente, resta che la nozione giuridica di intesa potrà trovare facile applicazione anche in questo scenario e, conseguentemente, non sarà particolarmente complesso "sanzionare", magari in misura maggiorata[239], l'uso dell'algoritmo che ha facilitato l'illecito coordinamento[240]. Assai chiaro in proposito un recente caso europeo che ha riguardato delle agenzie

[236] CE, 7 luglio 2016, Caso AT.39850, Container Shipping, GU C327, 6 settembre 2016, 4, dove la Commissione ha espresso preoccupazione circa un sistema di annunci per effetto del quale le parti regolarmente comunicavano al mercato, sui loro siti internet, sulla stampa o in altri modi, gli aumenti futuri dei prezzi praticati per i servizi di trasporto marittimo in container di lungo percorso. Secondo la Commissione detti annunci, anziché aiutare la clientela a meglio organizzare i suoi traffici, consentivano alle parti di esaminare le reciproche intenzioni di prezzo e di coordinare i loro comportamenti, testando i possibili effetti degli stessi e comunque riducendo l'alea competitiva. La decisione è stata conclusa con impegni.

[237] Cfr. Competition and Markets Authority, Price-fixing: guidance for online sellers (2016), https://www.gov.uk/government/publications/price-fixing-guidance-for-online-sellers.

[238] Cfr. OECD, Algorithms and Collusion, cit., § 23.

[239] Non mancano infatti i casi nei quali la gravità della violazione del diritto antitrust è stata inferita anche dalla pervicacia e dalla determinazione profuse dalle imprese per realizzarla – cfr. CE, 15 luglio 2015, AT.40098 Blocktrains, § 83.

[240] Cfr., Ariel Ezrachi, Maurice E. Stucke, *Algorithmic Collusion: Problems and Counter-Measures* § 5 (2017), https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP/WD%282017%2925&docLanguage=En, i quali scrivono che «from a legal and policy perspective, this scenario is unremarkable. Technology in this case does not affect the scope and application of the law».

di viaggio lituane che, impiegando il medesimo sistema di prenotazione online, avevano ricevuto dall'amministratore di sistema un invito a non fissare sconti superiori ad un certo livello e avevano poi visto implementata una restrizione tecnica che, ancorché aggirabile, concretamente limitava in quel senso gli sconti applicabili alle prenotazioni eseguite tramite il sistema. Di là dalla circostanza che ha voluto questo invito proposto tramite un software, la Corte di Giustizia ha correttamente osservato che, come nel mondo offline – e, qui si aggiunge, come in un celeberrimo caso statunitense[241] – in questo scenario si sarebbe configurata una intesa qualora si fosse accertato che ciascuna agenzia avesse materialmente ricevuto il messaggio e non si fosse da questo pubblicamente dissociata[242].

Discorso invece ben più complesso è quello che riguarda la c.d. collusione algoritmica, ossia il caso delle imprese che impiegano i c.d. predictable agent o parallel algorithms per stabilire le proprie strategie, se del caso reagendo in tempo reale alle altrui scelte commerciali. In questo scenario ogni operatore sviluppa il proprio algoritmo in modo autonomo e indipendente e, soprattutto in presenza di alcune caratteristiche di mercato[243], può inoltre sapere che, giacché i suoi concorrenti potrebbero fare altrettanto, il mercato nel suo complesso potrebbe allinearsi verso un equilibrio collusivo. Ad esempio, in una strategia "follow-the-leader", un'impresa potrebbe fare un'offerta collusiva implementando un algoritmo che imita in tempo reale il prezzo del leader di mercato, mentre il leader potrebbe accettare l'offerta aumentando il prezzo in reazione ai concorrenti algoritmo. Analogamente, laddove un'impresa implementasse una strategia, autonomamente decisa, di "tit-for-tat", l'equilibrio cooperativo potrebbe raggiungersi ancor più velocemente, non appena i rispettivi algoritmi riconoscessero che il prezzo di cartello è la condotta che i rivali "infliggono" loro e che dunque essi devono "restituire"[244]. Per un verso, quindi, vi è chi ritiene di poter rintracciare una forma di concertazione nel fatto che le imprese sarebbero consapevoli di come l'uso di questi algoritmi, pur unilateralmente disegnati, sarebbe capace di produrre interdipendenza, quasi che il meeting of minds fosse sostituito dal meeting of algorithms. Per altro verso, però, non si capisce

[241] Cfr. Theatre Enterprises v. Paramount Distributing, 346 U.S. 537 (1953), dove il fatto che i cinema concorrenti avessero ricevuto il medesimo invito a colludere dallo stesso distributore di film fu considerato un indizio di una pratica concertata perché tutte le copie di quelle lettere recavano gli indirizzi di tutti i cinema.

[242] CG, 21 gennaio 2016, Caso C-74/14, Eturas, ECLI:EU:C:2016:42. Parte della sentenza si dedica a discutere della necessità di provare che le imprese abbiano effettivamente ricevuto l'invito a colludere – circostanza che deve essere provata dalle autorità antitrust.

[243] Fermo restando quanto detto sopra, con riguardo alla possibilità che il numero delle imprese conti di meno, resta vero che la collusione è più probabile quando il mercato concentrato, prodotti omogenei, prezzi online, meccanismi di ritorsione credibili, come la velocità di ritorsione e le transazioni sono frequenti, ripetute, regolari e piccole. Infatti, Ezrachi, Stucke, *Algorithmic Collusion*, cit., §14 scrivono «the industry-wide use of algorithms, given the speed and enhanced transparency, could expand the range of industries susceptible to collusion beyond duopolies to perhaps markets dominated by 5 or 6 players, as we illustrate below».

[244] Il nome della strategia corrisponde all'espressione italiana pan per focaccia, perché consiste nel pensare che ogni impresa reagisca alle condotte delle proprie rivali ripetendo la loro medesima strategia. Ecco perché essa viene anche denominata come la strategia della ritorsione equivalente – cfr. Robert Axelrod, *The Evolution of Cooperation*, Basic Books (1984).

come questo scenario dovrebbe considerarsi diverso dal caso lecito dell'impresa che, nell'era analogica, avesse utilizzato una strategia "calcolata a mano" per adattarsi *intelligentemente* ai comportamenti delle sue concorrenti[245]. In altri termini, non v'è dubbio che degli algoritmi sviluppati su dei big data producano risultati incommensurabilmente più accurati di quelli ipotizzati da alcune persone e, quindi, non v'è dubbio che l'interdipendenza stabilmente gestita da degli algoritmi che per giunta reagiscono in tempo reale ad eventuali shock sia assai più pericolosa di una interdipendenza, contingente e momentanea, indovinata da alcuni esseri umani. Tuttavia, questa valutazione non incide – o non dovrebbe incidere – sul principio per cui un comportamento unilaterale, autonomo e indipendente non dovrebbe costituire una intesa. E ciò a meno che, come si dirà nel sesto capitolo, non si riformi l'attuale nozione di intesa, per consentire che anche ipotesi di collusione tacita vengano ricondotte nell'ambito di applicazione degli artt. 101 e 2.

È invece diversa la questione che attiene un'altra specie di algoritmi, i c.d. algoritmi automatici che "da soli fanno cartello". Lo scenario considerato – forse al momento ancora futuribile – è quello di imprese che fissano come obiettivo quello della massimizzazione del profitto senza escludere l'ipotesi di una collusione e di macchine che, attraverso l'auto-apprendimento e meccanismi di sperimentazione, determinano autonomamente dei prezzi che sì massimizzano il profitto, ma che al contempo sono prezzi di cartello. In altre parole, esiste il rischio che alcuni algoritmi dotati di una potente capacità predittiva, attraverso l'apprendimento automatico e la riproduzione automatizzata delle azioni proprie e di altri attori di mercato (macchine o esseri umani che siano), potranno colludere senza la necessità di un intervento umano a tal fine predisposto. Il tema che dunque si pone non riguarda la nozione di intesa cara al diritto antitrust ma, più in generale, se l'azione dei robot e l'intelligenza artificiale possano essere sottoposti a scrutinio e reprimenda. Ora, scartata l'ipotesi di un non intervento che individuerebbe ampie sacche di impunità, la dottrina evidenzia che o si attribuisce la responsabilità del comportamento eventualmente illecito allo strumento artificiale, oppure si sceglie di continuare a ritenere responsabili coloro che quel robot hanno congeniato e/o stanno utilizzando[246]. Al momento, questa ultima opzione sembra quella preferita nell'Unione Europea[247], ma non negli Stati Uniti[248] dove alcuni paventano come gli individui di fatto perdano il controllo degli algoritmi[249].

---

[245] In tal senso anche Ezrachi, Stucke, *Algorithmic Collusion*, cit, § 74, dove si legge «'pure' forms of tacit collusion which result from a unilateral rational reaction to market characteristics would not normally trigger antitrust liability». Inoltre, Joseph E. Harrington, Jr., *Developing Competition Law for Collusion by Autonomous Price-Setting Agents* 30 (2017), https://ssrn.com/abstract=3037818, il quale afferma «[i]n conclusion, it seems unlikely that, according to existing jurisprudence, collusion by [autonomous agents] could result in the Örms deploying those [autonomous agents] being found in violation of section 1 of the Sherman Act».

[246] Cfr. Salil K. Mehra, *Antitrust and the Robo-Seller: Competition in the Time of Algorithms* (2015), http://www.minnesotalawreview.org/wpcontent/uploads/2016/04/Mehra_ONLINEPDF1.pdf;

[247] Cfr. sul punto quanto affermato da Margrethe Vestager, *Algorithms and competition*, discorso tenuto press oil Bundeskartellamt, 18th Conference on Competition, Berlino, 16 Marzo 2017, https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/bundeskartellamt-18th-conference-competition-berlin-16-march-2017_en, dove la Commissaria ha affermato che «[t]he challenges that automated systems create are very real. If they help companies to fix prices, they really could make our economy work less well for everyone else. (…) So as competition enforcers, I think we need to make

3.6.1.4 Il caso particolare delle piattaforme della sharing economy

L'uso degli algoritmi è stato discusso anche in relazione alle intese tra imprese che operano a diversi livelli della filiera produttivo-distributiva. E, analogamente a quanto osservato in rapporto alle intese orizzontali, si è constatato come gli algoritmi possano, oltre che rendere più efficiente l'integrazione verticale a beneficio dei consumatori, anche supportare la realizzazione di una strategia anticompetitiva di fissazione e/o raccomandazione dei prezzi di rivendita. Gli algoritmi cioè potrebbero aiutare le imprese a individuare quali prezzi di rivendita fissare e a verificare che tutte le parti dell'accordo vi si attengano. Così – e di nuovo – si deve constatare come non siano gli algoritmi in sé a essere anticompetitivi, quanto le funzioni per l'assolvimento delle quali essi sono stati predisposti.

Nondimeno, l'uso nel contesto verticale dei pricing algorithms induce a riflettere su un ulteriore scenario: quello delle piattaforme della c.d. sharing economy che, come Uber, offrono a coloro che prestano servizi tramite la piattaforma, gli autisti nel caso di Uber, un comune sistema di determinazione del prezzo[250]. Ebbene, ricordato che l'algoritmo condiviso potrebbe anche rivelarsi in sé pro competitivo, andando ad individuare non un prezzo di cartello, ma il prezzo di mercato che dipende dalle fluttuazioni istantanee della domanda e dell'offerta[251], resta che la condivisione di

---

it very clear that companies can't escape responsibility for collusion by hiding behind a computer program».«[…] businesses also need to know that when they decide to use an automated system, they will be held responsible for what it does. So they had better know how that system works». Inoltre, ha aggiunto che «[o]ne could wonder whether pricing algorithms could, without explicit instructions to do so, engage in explicit collusion with each other. This would require algorithms to achieve a sense of communication (for concerted practices) and bargaining and mutual commitment (for reaching "agreement"). If this is or were to become possible in the future, the firms using such algorithms would remain liable for their behaviour. It is up to the firms using algorithms to ensure that their algorithms do not engage in illegal behavior». In tal senso anche Sathya S. Gosselin, April Jones, Annabel Martin, *Are Antitrust Laws Up to the Task? A US/EU Perspective on Anti-Competitive Algorithm Behavior* (2017) www.hausfeld.com/news/us/are-antitrust-laws-up-to-the-task-a-us-eu-perspective-on-anti-competitive-a, dove si legge «It is no defense to suggest that algorithms, programmed for autonomy, have learned and executed anticompetitive behavior unbeknownst to the corporation. The software is always a product of its programmers – who of course have the ability to (a¢ rmatively) program compliance with the Sherman Act».

[248] Cfr. DOJ, *Algorithms and Collusion* - Note by the United States 6 (2017), dove si legge «Absent concerted action, independent adoption of the same or similar pricing algorithms is unlikely to lead to antitrust liability even if it makes interdependent pricing more likely. For example, if multiple competing firms unknowingly purchase the same software to set prices, and that software uses identical algorithms, this may e§ectively align the pricing strategies of all the market participants, even though they have reached no agreement».

[249] Cfr. Mehra, *Antitrust and the Robo-Seller*, cit..

[250] Maurice E. Stucke, Ariel Ezrachi, *How Pricing Bots Could Form Cartels and Make Things More Expensive, Harward Business Review* (2016), https://hbr.org/2016/10/how-pricing-bots-could-form-cartels-and-make-things-more-expensive.

[251] Uber, ad esempio, utilizza un sistema di individuazione dei prezzi che viene detto dinamico, perché esso varia in ragione di alcuni fattori quali: (a) le metriche interne all'impresa (come accade quando un sito web che ad esempio vende biglietti per spettacoli ed eventi sportivi cambia il suo prezzo a causa della basso traffico che registra); (b) i prezzi dei concorrenti (quando l'impresa cambia i prezzi per tenere il passo di alcuni concorrenti);

questo algoritmo da parte di tanti e diversi fornitori di servizi limiterebbe la concorrenza tra gli stessi. Questi soggetti cioè non avrebbero la libertà di praticare prezzi diversi (anche inferiori) rispetto a quelli indicati dall'algoritmo, il quale potrebbe essere pensato proprio per evitare una guerra di prezzi tra detti utenti. Si verrebbe così a riprodurrebbe il modello collusivo del *Hub and Spoke*, ossia il caso di un centro – il c.d. hub – che favorisce la collusione orizzontale tra i tanti agenti economici – i c.d. spoke – che fanno affidamento sulle medesime istruzioni per sviluppare la rispettive strategie di tariffazione. Ad esempio, negli anni '90 il DOJ ha esaminato il caso delle compagnie aeree statunitensi che erano riuscite a concertare i loro prezzi utilizzando un centro di coordinamento esterno che diramava, ancorché tramite sofisticati meccanismi di segnalazione, le tariffe che ogni impresa avrebbe dovuto praticare[252].

Tuttavia, almeno con riguardo all'ipotesi delle piattaforme di sharing economy[253], occorre riflettere sull'eventualità che i soggetti che offrono servizi per il tramite di dette piattaforme siano

---

(c) l'incontro in tempo reale di domanda e offerta (come spesso accade con riguardo ai prezzi di biglietti aerei e stanze di albergo); e (d) eventuali altri elementi esterni (come le condizioni climatiche, in caso di trasporti). Famoso, al riguardo, è il sistema di dynamic pricing impiegato da Uber – Cfr. https://newsroom.uber.com/guest-post-a-deeper-look-at-ubers-dynamic-pricing-model/. Ora, non sembra esserci ragione per considerare lesivo del diritto antitrust questo modo di fare il prezzo. Anzi, esso mima il caso di scuola di un mercato dove i prezzi (ma anche le quantità, ossia il numero di "drivers" disponibili) si aggiustano quasi istantaneamente in funzione della domanda e dell'offerta. Del resto, «[w]e must look at the price system as such a mechanism for communicating information if we want to understand its real function – a function which, of course, it fulfills less perfectly as prices grow more rigid» – Cfr. Friedrich A. Hayek, *The use of knowledge in society*, 35AmericanEconomic Review 519, 526 (1945). Diversamente, si potrebbe invece profilare un tema di correttezza nel caso di un acquirente (ad esempio, una mamma) che ha scelto di utilizzare un determinato servizio (impiegare Uber affinché i figli fossero accompagnati a casa dopo la scuola) quando il prezzo dello stesso era più basso (si supponga 40 euro) di quello che le verrebbe invece praticato al momento del godimento del servizio (si supponga 400 euro, perché nel frattempo si è scatenata una tempesta di neve e i "drivers" in circolazione sono pochi). Tuttavia, finché l'acquirente è posto nelle condizioni di poter recedere dal servizio (sempre a titolo di esempio, Uber chiede ai propri clienti di riconfermare la richiesta del servizio se nel frattempo il prezzo dello stesso è aumentato), anche un aumento del prezzo che potrebbe giudicarsi eccessivamente elevato non dovrebbe qualificarsi come vessatorio ai sensi dell'art. 33 del codice del consumo.

[252] Cfr. Severin Borenstein, *Rapid Price Communication and Coordination: The Airline Tariff Publishing Case* 310, John E. Kwoka Jr., Lawrence J. White (a cura di), The Antitrust Revolution: Economics, Competition and Policy (1999).

[253] Non è semplice risalire alle origini dell'espressione shairng platform economy. Ciò che però appare chiaramente è che solo dieci anni fa l'avvento di queste piattaforme per la condivisione venne accolto associandolo a una serie di valori positivi: segnatamente, come una alternativa al consumismo e al modello tradizionale della proprietà privata, ossia come una nuova forma di economia sostenibile che limitava gli sperperi e ostacolava l'anonimato della produzione e della società di massa. Su dette piattaforme gli individui hanno infatti l'opportunità di ri-progettare il modo in cui consumavano, possedevano e lavoravano. Invece di consentire alle imprese di vendere i propri beni e servizi, tali piattaforme permettevano agli individui di beneficiare dell'accesso comune a prodotti e competenze. Su quelle piattaforme gli individui potevano cioè condividere, barattare, prestare, affittare, regalare e scambiare risorse che altrimenti sarebbero rimaste inaccessibili e sottoutilizzate. Il "consumo collaborativo" diventava dunque – stando sempre alla reazione entusiastica che ingenerò nei primi anni del nuovo millennio – uno strumento per ridurre costi e sprechi, per risparmiare tempo, per semplificare le relazioni produttive, nonché un mezzo per fare amicizia e diventare cittadini più consapevoli. Tuttavia, il termine "condivisione" non sembra il più accurato per descrivere il modello di business impiegato da piattaforme come Uber. Vero, Uber aggrega lavoratori sotto-impiegati e

effettivamente considerati imprese ai sensi del diritto antitrust; e di più che essi siano considerati imprese distinte dalla piattaforma medesima. Se, ad esempio, si ritenesse che gli autisti che utilizzano Uber per offrire i propri servizi non perseguono un interesse economico in conflitto con, o anche solo divergente da, quello della piattaforma di Uber, allora coerentemente con la tradizione antitrust, quegli autisti e la piattaforma dovrebbero qualificarsi come un unico agente economico o, se si preferisce, come un unico centro di interessi economici. E questa conclusione dovrebbe portare ad escludere che tra la piattaforma e gli utenti/prestatori di servizi possa prodursi alcuna forma di intesa, perché a priori non si potrebbero configurare degli interessi divergenti ricondotti a unità per effetto della concertazione[254].

3.6.1.5 Possibili soluzioni rispetto alla collusione algoritmica

Sotto il profilo antitrust è dunque questo ultimo scenario a sollevare maggiori preoccupazioni, le quali al momento potrebbero essere affrontate in due modi. Si potrebbe seguire la tesi di Kaplow, che rinverdisce quella di Posner, circa la necessità di colpire gli equilibri collusivi in quanto tali, ossia indipendentemente dall'esistenza e dalla prova di una previa intesa tra le imprese. In altri termini, una via per la radicale soluzione del problema di una collusione non più riconoscibile e non più riconducibile all'idea dell'incontro tra due o più volontà, potrebbe trovarsi nella eliminazione della nozione di intesa, ossia nella neutralizzazione dell'elemento giuridico interno all'art. 101, con l'effetto finale di appunto colpire non tanto i comportamenti delle imprese che conducono a equilibri collusivi, quanto la mera esistenza di questi equilibri.

Lontano da questa scelta così estrema, si rinvengono altre opzioni. Ad esempio, le autorità potrebbero utilizzare dei simulatori, chiamati "collusion incubator", che sappiano riconoscere se e quando l'evolversi del mercato segnala l'esistenza di una forma di collusione. Diversamente, si potrebbe predisporre la soluzione del c.d. "Antitrust by design", la quale dovrebbe consistere nella scelta di punire le imprese che, in sede di programmazione dei loro algoritmi, non inseriscono delle linee di codice che ordino alle macchine di scartare gli scenari collusivi. In altri termini, si tratterebbe di rintracciare la volontà collusiva non nella definizione della strategia operativa, non negli

consumatori insoddisfatti e, quindi, rimedia alle inefficienze consentendo agli individui di comunicare e collaborare in modo più efficace ed efficiente. Inoltre, a differenza di un negozio online, Uber non fornisce direttamente beni o servizi. Tuttavia, Uber e gli autisti lavorano per perseguire un profitto e Uber non opera come un mero portale web o come un'applicazione mobile che ospita le numerose transazioni che si verificano tra consumatori e autisti. Al contrario, Uber fornisce il sistema di pagamento elettronico, garantisce che i prezzi funzionino in modo dinamico, organizza gli elenchi degli autisti disponibili, addebita una tassa per ogni scambio e stabilisce standard minimi di qualità per i conducenti. Di conseguenza, e come si diceva, pare che l'etichetta sharing economy non sia la più indicata per descrivere cosa Uber sia. In proposito e in tema di sharing economy si vedano da ultimo Edmondo Mostacci, Alessandro Somma *Il caso Uber. La sharing economy nel confronto tra common law e civil law* (2016).

[254] In altri termini, il diritto antitrust riconosce l'esistenza di una o più imprese quando riscontra una pluralità di interessi economici che un'eventuale intesa potrebbe ridurre a uno – cfr. Piero Verrucoli, La nozione di impresa nell'ordinamento comunitario 396 (1977); Enzo Moavero Milanesi, *Diritto della concorrenza dell'Unione Europea 55* (2004); Antonio Spadafora, La nozione di *impresa* nel diritto comunitario, Giust. civ. 283 (1990); e Gianvito Giannelli, Impresa pubblica e privata nella legge antitrust 121 (2000).

intervenuti scambi di informazione, ma nella determinazione di non impedire alle macchine di individuare gli equilibri collusivi. Tuttavia, in questo caso, la tecnologia potrebbe minare l'efficacia dell'intervento. Vengono infatti utilizzati algoritmi sempre più avanzati e complessi che, di conseguenza, sono più difficili da verificare. Inoltre, sorge il dubbio che anche questa via elimini dal quadro dell'analisi la multilateralità delle condotte di cui all'art. 101: un'impresa che non vieta al suo algoritmo di contemplare la collusione, agisce autonomamente, non necessariamente di comune accordo.

Nondimeno, proprio per i limiti di queste soluzioni, ancora si discute dell'opportunità di qualsiasi intervento[255].

3.6.2 I big data e i prezzi personalizzati

Secondo la teoria economica, si definisce personalizzato il prezzo che uguaglia la massima cifra che un possibile acquirente di un determinato bene è disposto a, ed è nelle condizioni di, pagare per ottenerlo. L'esplosione del numero di dati che descrivono i comportamenti delle persone e la realtà empirica, nonché l'incredibile aumento delle capacità computazionali degli attuali sistemi informatici, hanno reso le imprese contemporanee capaci di individuare (o di ben approssimare) la disponibilità a pagare dei loro acquirenti senza dover ricorrere a negoziazioni individuali e, di più, senza che tali clienti e consumatori siano informati o siano pienamente consapevoli di questa nuova capacità delle imprese. Ad esempio, piattaforme di *e-commerce*, ad esempio, già da diversi anni, utilizzano algoritmi per la determinazione automatica dei prezzi di vendita. Queste piattaforme consentono ai venditori di segmentare il mercato attraverso delle dinamiche di prezzo. Tali modalità di tariffazione sono ampiamente utilizzare nel settore dei viaggi, prenotazione alberghiere, vendite al dettaglio, nello settore dello sport e del divertimento, come pure nel settore assicurativo. Gli algoritmi di prezzi dominano anche le vendite online di beni - l'ottimizzazione del prezzo sulla base degli stock disponibili e della domanda attesa.

Ebbene, rispetto a queste considerazioni, si individuano due diverse questioni che potrebbero trovare qualche rilievo giuridico.

La prima riguarda l'eventualità che gli acquirenti cui vengono offerti dei prezzi personalizzati non sappiano né del processo *data driven* e *data based* attraverso il quale quei prezzi sono stati formati, né della loro natura di "prezzi diversi per ogni compratore". Una forma di duplice ignoranza, quest'ultima, che potrebbe sollecitare qualche riflessione con riferimento alla natura eventualmente scorretta della pratica dei prezzi personalizzati. La seconda questione giuridica che si potrebbe considerare concerne invece sia il rispetto delle regole in materia di tutela dei dati personali che dovrebbero guidare la formazione dei dataset ai quali i sopra menzionati strumenti di analisi vengono applicati, sia l'eventualità che l'analisi dei big data consenta di associare, ancorché in via predittiva, ad ogni individuo un'identità digitale – fatta, per esempio, di previsioni su preferenze e vincoli di bilancio – non soggetta al controllo dell'individuo stesso. Questi profili, però, esulano dall'ambito di

[255] A. Ezrachi & M. E. Stucke, *Algorithmic Collusion: Problems and Counter-Measures*, 21-23 June 2017.

studio del presente lavoro, che invece vuole evidenziare in che misura i prezzi personalizzati siano perseguibili in termini antitrust.

A tal fine, giova partire con una considerazione di stampo economico. In economia, l'espressione discriminazione di prezzo viene utilizzata per indicare tutti quei casi di concorrenza imperfetta[256] nei quali una impresa offre il medesimo bene a prezzi differenti e maggiori del costo marginale di quel prodotto o servizio[257]. Infatti, fatta eccezione per le imprese che discriminano sulla scorta della propensione al rischio degli individui, scopo precipuo delle altre imprese che optano per questa strategia è quello di calcolare dei prezzi che, anziché dipendere dai costi di produzione e commercializzazione, derivano dalla domanda e, in particolare, dalla disponibilità a pagare (anche detta, prezzo di riserva) dei loro diversi acquirenti. Tuttavia, non è affatto semplice individuare il prezzo di riserva di clienti e consumatori o, meglio, non è così comune per un'impresa disporre delle informazioni necessarie a conoscere le preferenze e i vincoli di bilancio dei propri compratori.

In tale ottica possiamo individuare, sotto il profilo economico, tre diverse modalità attraverso cui le imprese riescono ad acquisire le informazioni rilevanti per la determinazione del prezzo di riserva:

i) «discriminazione di primo grado» o «prezzo personalizzato»[258], è il caso in cui l'impresa riesce a fare a *ciascuno* dei suoi acquirenti un prezzo differente, come avviene nel caso dei rivenditori di automobili che offrono uno stesso modello a prezzi diversi al termine di trattative individualizzate, oppure come accade nell'ipotesi della c.d. asta olandese;

[256] Come noto, in concorrenza perfetta le imprese, proprio perché prive di potere di mercato, non possono praticare prezzi differenziati, ma devono applicare un prezzo uniforme pari al loro costo marginale. Di conseguenza, affinché si possa concepire un'ipotesi di discriminazione di prezzo, occorre supporre che le imprese interessate dal fenomeno detengano una qualche forma, ancorché limitata e di breve periodo, di potere di mercato, ossia una qualche capacità di continuare a profittare pur facendo un prezzo diverso dal loro costo marginale. Infatti, la dottrina economica maggioritaria riconosce il potere di mercato così inteso, come uno dei tre presupposti (insieme all'assenza di arbitraggio e alla conoscenza, ancorché approssimata, del prezzo di riserva degli acquirenti) richiesti affinché la discriminazione di prezzo possa essere praticabile – Vedi in via specifica sul potere di mercato Stole M. A., *Price discrimination and imperfect competition,* 2003, p. 1, disponibile all'indirizzo http://web.mit.edu/14.271/www/hio-pdic.pdf, Areeda P., Hovenkamp H., *Antitrust law. An analysis of antitrust principles and their application, Boston*, Aspen Publishers, IIB, 2007, 517, p. 150; Carbonneau S., McAfee P., Mialon S., *Price discrimination and market power*, Emory economics, 2004, p. 413, nonché Levine M. E., *Price discrimination without market power*, 19 Yale J. on Reg. (2002) p. 1. Per quanto invece concerne le condizioni generali, affinché la discriminazione di prezzo sia una strategia praticabile, si vedano Varian H.R., *Price Discrimination*, in Schmalensee R., Willig R.D., *Handbook of industrial organization*, Vol. 1, 1989, p. 600 e P. Belleflamme, Peitz M., *Industrial Organization. Market and Strategies*, Cambridge, Cambridge University Press, 2015, p. 198.

[257] Clerides S., *Price discrimination with differentiated products: definition and identification*, 20 International journal of industrial organization, 2002, p. 1385; Stigler G.J., *The theory of price*, New York, The Macmillan Company, 1966, p. 209.

[258] Pigou A.C., *The economics of welfare,* London, Macmillan and Co., 1920 *e* Shapiro *C.,* Varian *H.R., Information rules: a strategic guide to the network economy*, Boston, Harvard Business School Press, 1999 e sono ormai entrate nel lessico comune.

ii) «discriminazione di secondo grado» o del «*menu pricing*», in questo caso l'impresa definisce i prezzi da praticare in rapporto a delle caratteristiche associabili al suo bene, per poi attendere che i compratori rivelino i loro prezzi di riserva proprio compiendo una scelta rispetto a quelle caratteristiche (esempi ne sono gli sconti quantità, i prezzi praticati dalle aziende di trasporto pubblico che prevedono importi maggiorati per i biglietti acquistati nelle ore di punta, oppure delle compagnie aeree che prevedono prezzi ridotti per chi acquista i biglietti con largo anticipo, i beni pacchetto);

iii) «discriminazione di terzo grado» o di «*group pricing*», attraverso cui un'impresa sceglie di assumere un criterio verificabile come l'età dell'acquirente, la sua occupazione, o la sua provenienza geografica, a indice della ricchezza del compratore e/o del suo desiderio per un determinato prodotto o servizio, per poi utilizzare questo indicatore per separare *ex ante* i propri acquirenti in categorie diverse di compratori alle quali infine applicherà prezzi differenti (esempi di questo tipo di pratica sono le tasse calmierate che le università offrono a studenti con redditi inferiori a una certa soglia, gli sconti che le sale cinematografiche praticano a bambini e anziani perché, in entrambi i casi, vale la convinzione secondo cui questi soggetti avrebbero una capacità di spesa ridotta, la tecnica del *geobloking* basata sull'assunto secondo cui l'appartenenza degli acquirenti a determinate aree geografiche approssimerebbe la loro disponibilità a pagare).

Come detto, l'uso di internet e delle tecnologie digitali, la collezione dei big data e lo sviluppo di tecniche di analisi sempre più potenti e sofisticate hanno permesso alle imprese di raggruppare i propri clienti e consumatori in modo così preciso e dettagliato da approssimare in maniera significativa lo scenario della discriminazione perfetta, riducendo notevolmente le difficoltà insite nella raccolta di informazioni rilevanti per la determinazione dei prezzi di riserva. Di più, la diffusione della rete, degli smart phone, delle nuove tecnologie abilitanti largamente analizzate nella prima parte del presente lavoro, e l'ampio ricorso da parte delle piattaforme digitali ai modelli di business c.d. «*two-sided*» e «*ad-supported*» ha ad oggi decuplicato le fonti di dati da utilizzare per estrarre informazioni circa gli acquirenti. Le analisi condotte sui big data potrebbero, infatti, consentire alle imprese di abbandonare forme grossolane di segmentazione della clientela determinate da poche variabili, a cluster sempre più piccoli di individui con molteplici caratteristiche comuni, così da formulare predizioni sempre più vicine i loro prezzi di riserva[259].

Onde comprendere se i prezzi personalizzati possano sussumersi nella più ampia fattispecie dei prezzi discriminatori, occorre elaborare due osservazioni.

[259] I prezzi personalizzati potrebbero essere dunque assimilati a una forma assai sofisticata di discriminazione di terzo tipo, piuttosto che ad una vera e propria discriminazione perfetta. Nondimeno, la differenza tra queste forme di discriminazione passa proprio per la misura dell'eterogeneità dei consumatori che rimane non osservabile dalle imprese – vedi Stole, *Price discrimination*, cit., p. 3. Così, se per effetto delle analisi dei big data tale eterogeneità può essere mappata e quindi conosciuta dalle imprese, allora l'equivalenza con la discriminazione perfetta è sostenibile. Questione diversa è invece osservare che, mentre i criteri che guidano la «tradizionale» discriminazione di prezzo di terzo tipo sono solitamente espliciti e conosciuti dai consumatori, i parametri che dovrebbero plasmare la discriminazione di terzo tipo basata sui big data restano nascosti negli algoritmi delle imprese. Ma questo, anziché rilevare sotto il profilo allocativo e distributivo, potrebbe al più importare con riguardo ai temi della trasparenza e della simmetria informativa.

In primo luogo, quanto all'impatto dei prezzi personalizzati sul consumer welfare, va detto che la discriminazione di prezzo in concorrenza non costituisce una minaccia per l'efficienza dei mercati e la distribuzione della ricchezza. Diversamente, essa diventa problematica quando operata da monopolisti e imprese in posizione dominante. Meglio, la discriminazione perfetta del monopolista rappresenta comunque una opzione preferibile rispetto al divieto di discriminazione dunque al prezzo di monopolio giacché, sebbene il monopolista si appropri di tutto il surplus dei consumatori, si produce un aumento dell'output (e quindi del numero di consumatori che riescono a godere del bene) che non si avrebbe qualora fosse praticato il prezzo di monopolio. Seguendo le fila di questo ragionamento, dunque, anche la discriminazione perfetta delle imprese in posizione dominante è preferibile a una situazione di divieto di discriminazione e dunque di prezzo sovra-competitivo da posizione dominante, qualora si accerti che tale discriminazione perfetta abbia prodotto un aumento dell'output.

In secondo luogo, venendo all'analisi più prettamente giuridica, occorre muovere dal *dato letterale*, ossia dal fatto che per il diritto antitrust la discriminazione di prezzo può costituire o una pratica unilaterale, che verrà allora esaminata in quanto posta in essere da un'impresa in posizione dominante, o una pratica multilaterale, che verrà vagliata in quanto convenuta dalle imprese parti di un'intesa.

In quest'ultimo caso, qualora due o più imprese *decidessero di comune accordo* di praticare gli *stessi* prezzi ancorché dissimili, le autorità antitrust si troverebbero a fronteggiare un cartello sui prezzi, senza che fosse dunque necessario ai fini dell'applicazione del divieto procedere oltre con l'accertamento della sussistenza degli elementi costitutivi dell'illecito discriminatorio[260]. Pertanto, la determinazione congiunta dei medesimi prezzi personalizzati dovrebbe qualificarsi come un'intesa sui prezzi e, conseguentemente, dovrebbe essere vietata dato il suo oggetto anticompetitivo e la sua incapacità di soddisfare le condizioni di cui al terzo paragrafo dell'art. 101 TFUE e all'art. 4 l. 287/90. E ciò – lo si ribadisce – a prescindere dall'eventualità che i soggetti passivi di un tale comportamento fossero consumatori o imprese.

Invece, con riguardo alla fattispecie abusiva, giova ricordare che un'impresa in posizione dominante abusa del suo potere quando applica «nei rapporti commerciali con gli altri contraenti condizioni dissimili per prestazioni equivalenti, determinando così per questi ultimi uno svantaggio per la concorrenza». Pertanto, i «prezzi dissimili», le «prestazioni equivalenti» e lo «svantaggio competitivo» sono le componenti essenziali delle condotte di prezzo unilaterali che si volessero caratterizzare come discriminatorie.

Ebbene, posto che le autorità antitrust non si sono mai soffermate largamente su tali elementi, perché hanno spesso goduto dell'agio di rintracciarli in comportamenti già oggetto di divieto per

[260] Ai sensi della lett. (c) dell'art. 101 del TFUE – anche essa ripresa pedissequamente dalla lett. (d) dell'art. 2 della l. 287/90 – un'intesa è anticompetitiva quando consiste nell'applicare «nei rapporti commerciali con gli altri contraenti, condizioni dissimili per prestazioni equivalenti, così da determinare per questi ultimi uno svantaggio nella concorrenza».

altre ragioni[261], il caso dei prezzi personalizzati fissati unilateralmente, riguardando per definizione «prezzi dissimili» applicati ai diversi acquirenti che comprano il medesimo bene, sembra proprio riprodurre l'ipotesi dei prezzi dissimili offerti a fronte di prestazioni equivalenti.

Tuttavia, se si procede oltre e si riflette sulla nozione di «svantaggio competitivo», l'assimilazione tra i prezzi personalizzati e i prezzi discriminatori non pare più così immediata o, per lo meno, merita qualche ulteriore precisazione. Infatti, è diffusa in dottrina la tesi secondo cui lo «svantaggio competitivo» potrebbe consumarsi solo tra *imprese*[262], ossia agenti economici che: (*i*) possono essere coinvolti in un gioco concorrenziale che può vederli, se del caso, soccombere e che (*ii*) possono, a causa del più alto prezzo pagato, disporre di minori risorse da investire nel gioco concorrenziale, ossia da investire in efficienza, qualità e innovazione. I consumatori colpiti da un prezzo più alto, invece, potrebbero sì subire un pregiudizio, ma rispettivamente tale svantaggio: (*i*) non potrebbe qualificarsi come «competitivo», perché per definizione i consumatori non svolgono alcuna attività economica e, dunque, non possono considerarsi in concorrenza tra loro; e (*ii*) equivarrebbe a una mera riduzione di reddito incapace di influire sull'utilizzo delle risorse che incidono sul normale andamento del mercato.

Secondo questa linea interpretativa, dunque, dacché solo le imprese potrebbero essere i soggetti discriminati ai sensi delle lett. c) degli artt. 102 TFUE e 3 l. 287/90, gli unici prezzi personalizzati perseguibili a norma di quelle disposizioni sarebbero quelli praticati da un'impresa in posizione dominante a "danno" di altre imprese operanti a monte o a valle. I prezzi personalizzati offerti da un monopolista all'insieme dei suoi consumatori, invece, non sarebbero sussumibili nella classe delle condotte discriminatorie quale prevista dalle discipline europea e nazionale dell'abuso di posizione dominante.

Ora, sono almeno due le riflessioni che si innestano su questa duplice conclusione.

---

[261] Bellamy & Child, *European Union Law of Competition*, Oxford, Oxford University Press, 2013, p. 800 e Faull J., Nikpay A., *The EU law of competition*, Oxford, Oxford University Press, 2014, p. 524 e 533.

[262] Pappalardo A., *Il diritto comunitario*, cit., p. 177. In tal senso anche, Faull, Nikpay, *The EU law of competition*, cit., p. 387, dove si legge «discriminatory abuses, in which competition is harmed by discriminatory prices … charged … by the dominant undertaking on an intermediate market with the effect of placing certain suppliers or customers of the dominant undertaking at a 'competitive disadvantage». Inoltre, Bellamy, Child, *European Union Law*, cit., p. 803, i quali parlano senza tema di smentita di imprese clienti che operano nei mercati a valle, quando non discutono di un caso di discriminazione preclusiva. In questo senso, anche Libertini M., *Diritto della concorrenza dell'unione europea*, Giuffré, Milano, 2014, p. 326, dove si legge «[l]a formulazione testuale di questo divieto di discriminazione chiarisce che l'ambito del divieto riguarda solo i rapporti fra imprese e non le condizioni di offerta ai consumatori finali (per i quali non potrebbe concepirsi uno 'svantaggio nella concorrenza'). Ciò non significa che le discriminazioni di prezzo a livello di consumatori finali siano sempre lecite, bensì solo che è riconosciuta una maggiore libertà di azione all'impresa dominante. Questa potrebbe pertanto lecitamente praticare, nell'ambito di una propria politica "di immagine" degli sconti ai consumatori finali aderenti a una certa associazione, o a quelli appartenenti a certe fasce d'età, ecc.». Inoltre, ancorché commentando l'art. 101, si veda Frignani A., *Intese discriminatorie*, cit., p. 242, il quale scrive «[v]a detto anche tra questi "contraenti" non dovrebbero esserci i consumatori finali perché non sono soggetti attivi della concorrenza, ma semmai soggetti passivi».

In primo luogo, se le imprese discriminate fossero al tempo stesso rivali e clienti dell'impresa in posizione dominante, si profilerebbe un'ipotesi di *primary line injury*, motivo per cui le autorità antitrust, più che soffermarsi sulla natura discriminatoria dei prezzi personalizzati, guarderebbero soprattutto al loro carattere escludente e anticompetitivo in ragione delle teorie del danno anticompetitivo che si è soliti elaborare in relazione a imprese in posizione dominante che siano altresì integrate verticalmente o siano multi-prodotto. Del resto, nei più recenti casi di condotte di prezzo discriminatorie come *Tomra*[263] e *Intel*[264] la Commissione si è astenuta dal considerare i profili discriminatori, limitandosi ad apprezzare quelli escludenti e anticompetitivi. Ancor più esplicitamente, nel caso *Deutsche Bahn Energie*[265] la Commissione ha esplicitamente scelto di analizzare la condotta qualificandola come un'ipotesi di *margin squeeze* con ciò astenendosi dal considerare eventuali elementi discriminatori, diversamente da quanto aveva fatto nel precedente caso *Deutsche Bahn*[266]. Assai chiaramente, poi, la Corte di Giustizia ha osservato che «il fatto che una pratica di un'impresa che detiene una posizione dominante possa … essere qualificata come 'discriminazione tramite i prezzi', vale a dire l'applicazione di prezzi diversi a clienti diversi o a diverse categorie di clienti per prodotti o servizi i cui costi sono gli stessi o, al contrario, l'applicazione di un unico prezzo a clienti per cui i costi dell'offerta variano, non può, di per sé, suggerire l'esistenza di un abuso diretto all'esclusione di un concorrente»[267]. Così, nell'ipotesi di *primary line injury* il complesso accertamento degli ambivalenti effetti che, come si diceva nel paragrafo precedente, la discriminazione perfetta produce nei mercati intermedi e finali non dovrebbe allarmare, perché esso andrebbe comunque a integrare l'apprezzamento degli effetti che la preclusione determina sul benessere.

Diversamente, se le imprese discriminate fossero solo meri clienti o meri fornitori, si verificherebbe un'ipotesi di *secondary line injury* nella quale potrebbero rilevare i modelli economici che descrivono se e quando la discriminazione nei mercati intermedi – anziché consumarsi in una forma di puro sfruttamento – lede il benessere sociale e il benessere del consumatore e, per questo, distorce la concorrenza integrando gli estremi per uno «svantaggio competitivo». In questa circostanza, cioè, le autorità antitrust interessate a scorgere nella condotta discriminatoria l'origine di una distorsione concorrenziale dovrebbero innanzitutto evidenziare che i prezzi personalizzati non hanno determinato un aumento dell'output e che, ad esempio, hanno anche incentivato l'impresa monopolista a investire i suoi extra-profitti in attività mirate a schermare le sue rendite di posizione.

[263] T, causa T-155/06, Tomra Systems, Racc. 2010, p. II-04361 e CG, causa C-549/10P, Tomra et al., disponibile sul sito www.curia.europa.eu.

[264] Decisione della Commissione, COMP/37.990, Intel, 13 Maggio 2009, disponibile all'indirizzo http://ec.europa.eu/competition/sectors/ICT/intel_provisional_decision.pdf

[265] Decisione della Commissione, COMP/39.678 and COMP/39.731, Deutsche Bahn AG, GU 2003, L 1, p. 1.

[266] CG, causa C-436/97P, Deutsche Bahn AG, Racc. 1999, ECR I-2387.

[267] CG, causa C-209/10, Post Danmark, Racc. 2002, ECR I-172, §30.

La seconda riflessione, invece, riguarda l'eventualità che i prezzi personalizzati praticati da un'impresa in posizione dominante alla volta dei consumatori possano essere comunque puniti ai sensi dell'art. 102 e dell'interpretazione *funzionale* che si è soliti offrire dello stesso. Ebbene, in non pochi casi la Commissione ha applicato proprio l'art. 102(c) per impedire delle condotte discriminatorie che colpivano in via prioritaria i consumatori. Segnatamente, ciò è accaduto quando i comportamenti discriminatori, essendo dettati dalla provenienza geografica, dalla nazionalità e/o dal domicilio dei soggetti discriminati, sono stati ritenuti capaci di pregiudicare il processo di formazione del *mercato unico*, ricreando surrettiziamente le barriere geografiche tra gli Stati membri[268]. In altri termini, vero è che l'interesse alla tutela dell'integrità del mercato europeo – obiettivo di *policy* chiaramente perseguito dalle autorità europee nell'applicazione del diritto della concorrenza – ha indotto la Commissione a seguire un orientamento particolarmente severo e in qualche misura slegato dai testi legislativi o, per lo meno, da quella che è al momento la loro interpretazione dominante[269]. Così – si potrebbe argomentare – se in sede europea si scegliesse di eleggere a bene giuridico degno di tutela l'interesse dei consumatori a non vedere drenate tutte le proprie risorse a beneficio dell'impresa in posizione dominante, una forzatura della lett. c) analoga a quella dettata da ragioni geografiche – o comunque un'applicazione dell'art. 102 che ben sappiamo non prevede un numero chiuso di pratiche abusive – potrebbe sostenersi e quindi consentirebbe l'applicazione del divieto di là dalla lettera della norma.

Del resto, un generico divieto dei prezzi personalizzati risponderebbe alla tradizionale interpretazione degli abusi discriminatori, ossia all'idea secondo cui, ponendosi anche in un rapporto di species a genus rispetto al ben più generale principio europeo di eguaglianza o non-

[268] Si considerino, a titolo di esempio, CG, causa C-18/93, Corsica Ferries Italia v Corpo dei Piloti del Porto di Genova, GU 1994, p. I-1783, § 45, dove il sistema tariffario impiegato dall'impresa in monopolio legale preveda, a parità di servizio, delle riduzioni di tariffe a vantaggio delle sole navi italiane che operavano tra porti situati nel territorio nazionale; CG, causa 27/76, United Brands, Racc. 1978, p. 207 e CG, causa C-333/94P, Tetra Pak II, Racc. 1197, p. 5951. In questi ultimi due casi le imprese in posizione dominante sono state sanzionate per aver applicato prezzi diversi in distinti Stati Membri e per aver contestualmente tenuto delle condotte capaci di impedire l'arbitraggio tra gli acquirenti collocati nei differenti paesi europei. Discutendo di mera discriminazione di prezzo pare allora interessante notare come in occasione delle menzionate sentenze parte della dottrina abbia sostenuto come in United Brands e Tetra Pak II il vulnus competitivo risiedesse non tanto nella applicazione di prezzi differenti, quanto nelle condotte volte a compartimentare il mercato unico, impedendo (o disincentivando) gli scambi transfrontalieri che i delta di prezzo avrebbero invece potuto alimentare – vedi Faull J., Nikpay A., *The EU law of competition*, Oxford, Oxford University Press, 2014, p. 527. Non a caso, in altre decisioni la Commissione ha sanzionato proprio quegli schemi di incentivi e sistemi di sconto che avevano il blocco delle importazioni e delle esportazioni come obiettivo precipuo – vedi CG, causa C-310/93P, British Plasterboard, Racc. 1995, p. 865. Da ultimo, Decisione Commissione, OPCOM/Romanian Power Exchange, GU L 314, p. 7, in cui discriminatoria è stata qualificata la scelta dell'impresa che, oltre a dominare il mercato rumeno dell'energia elettrica, gestiva la borsa elettrica romena di imporre l'obbligo di aprire una partita IVA in Romania alle imprese stabilite nella UE che volessero concludere i propri affari partecipando alla borsa elettrica romena.

[269] Vedi CG, causa C-333/94P, Tetra Pak II, Racc. 1197, p. 5951; Commissione, 1998 Football World Cup, GU L55, p. 5, § 100 e Decisione della Commissione, COMP/35.141, Deutsche Post AG, GU L 125, 05/05/2001, p. 27. Critico nei confronti di questa impostazione Frignani A., *Intese discriminatorie*, cit., p. 329.

discriminazione[270], l'art. 102 dovrebbe tutelare la *correttezza* e l'*equità* nei rapporti negoziali, impedendo dunque che un soggetto dominante *sfrutti* gli agenti economici (consumatori e imprese) con cui si relaziona. In altri termini, vero è che attualmente il diritto antitrust è chiamato a tutelare il normale funzionamento del mercato per come tale funzionamento può essere apprezzato, nell'assenza di un paradigma che lo descriva nel dettaglio, dai risultati che esso produce, ossia dal modo in cui esso incide sul benessere totale *e* sul benessere dei consumatori. Tuttavia, mentre le variazioni del benessere totale danno conto di come il mercato operi nell'allocare prodotti e ricchezza, le variazioni del benessere dei consumatori segnalano come il mercato distribuisca gli stessi. Ecco che, allora, la discriminazione perfetta dell'impresa in posizione dominante applicata nei confronti dei consumatori pone un problema di *politica del diritto* che sino ad oggi non era stato necessario risolvere in modo draconiano: chiede cioè di chiarire se a essere protetto debba essere il mercato quale strumento di allocazione delle risorse, oppure il mercato quale strumento di distribuzione delle stesse. Se infatti nella maggior parte dei casi questa scelta non si impone perché la natura lecita o illecita di una condotta non muta a seconda che si scelga di usare il benessere totale o il benessere dei consumatori a indice dello stato di salute del mercato considerato[271], nel caso della discriminazione perfetta questa equivalenza non regge. Qualora a essere tutelato sia il benessere totale, i prezzi personalizzati dell'impresa in posizione dominante rivolti ai consumatori dovranno dirsi leciti, perché essi consentono di aumentare la quantità di prodotti disponibile a tutti. Diversamente, qualora a essere tutelato sia il benessere dei consumatori, i prezzi personalizzati dovranno essere considerati anticompetitivi, perché – come si diceva – consentono all'impresa in posizione dominante di drenare per l'intero il surplus dei consumatori.

Certo, è poi vero che queste valutazioni tutte interne al diritto antitrust non vanno confuse con considerazioni di carattere più generale, connesse ad esempio a quanto i consumatori siano consapevoli della natura personalizzata dei prezzi o del fatto che essi vengano computati in modo robotizzato, sulla scorta di dati anche personali. Questi profili attengono altre discipline, come la tutela dei consumatori e la tutela della privacy.

3.6.3 I big data e il rifiuto a contrarre

Nell'Unione Europea la disciplina del rifiuto a contrarre segue quanto stabilito nei casi Magill, IMS e Microsoft, ossia un'impresa in posizione dominante viola il diritto antitrust, quando: (i) l'input richiesto è essenziale; (ii) il rifiuto determina una riduzione della concorrenza nel mercato secondario; (iii) l'input richiesto serve a produrre un'innovazione, a fare ingresso in un nuovo mercato, a consentire lo sviluppo tecnico; e (iv) l'impresa non può vantare alcuna giustificazione oggettiva.[272] Ed in effetti la Commissione ha proprio richiamato questi casi, oltre a Huawei[273],

[270] CG, causa C-497/99P, Irish Sugar plc, Racc. 2001, p. I-5333, § 240. Pappalardo A., *Il diritto comunitario*, cit., p. 510, che considera questa «incontestabilmente» la finalità della norma; Bishop M., *Price discrimination under article 86. Political economy in the European court*, in 44 MLR 1981, p. 282.

[271] Motta M., *Competition Policy*, Cambridge University Press, Cambridge, 2009, pp. 20-22.

[272] Nel complesso, ad oggi, la dottrina del rifiuto a contrarre statunitense è inapplicabile quando il rifiuto: i) non è anticompetitivo; ii) è giustificabile; e iii) quando soddisfare la richiesta dei concorrenti richiederebbe

nell'affermare che «*general competition law is applicable in the context of data-driven business models and therefore it may be invoked to claim a wider access to data held by one economic operator*».[274]

Tuttavia, l'applicazione ai dati di questa dottrina europea sull' *essential facilities* (EFD) sembra piuttosto problematica anche secondo quella parte della dottrina che ne condivide l'attuazione[275].

Da subito, va osservato che le sopra menzionate condizioni sono state elaborate dalla giurisprudenza comunitaria con riferimento agli input protetti da diritti di proprietà intellettuale. Ora, non riferendosi al caso *Bronner* e citando soltanto i casi *Magill*, *IMS Health*, *Microsoft* and *Huawei,* lo Staff Working Document della Commissione sembra supporre che le stesse condizioni debbano applicarsi anche ai dati non coperti da IPR. Del resto, nel caso Microsoft sia la Commissione Europea, sia il tribunale hanno proceduto in tal senso per non doversi esprimere circa gli eventuali diritti di proprietà intellettuale che avrebbero potuto proteggere le APIs di Windows.

Inoltre, procedendo passo passo, occorre soffermarsi sull'indispensabilità della risorsa richiesta. Secondo *Bronner*, l'accesso ad un input non può essere ritenuto indispensabile se non esistono "ostacoli tecnici, giuridici o persino economici, tali da renderne impossibile, o addirittura irragionevolmente difficile la riproduzione"[276]. La Corte di Giustizia ha aggiunto che, per dimostrare la mancanza di un'alternativa potenziale reale, risulterebbe necessario quantomeno dimostrare che non sia economicamente sostenibile creare la risorsa su una scala comparabile a quella dell'impresa che controlla il prodotto o il servizio esistente[277]. Questa considerazione suggerisce che l'impresa richiedente dovrebbe effettuare lo stesso investimento dell'impresa dominante. Ciò significa che i dati accessibili perché resi liberamente disponibili o perché commercializzati non possono essere considerati indispensabili, così come non possono essere considerati tali le informazioni derivanti dall'elaborazione e da qualsiasi tipo di analisi di tali dati. Diversamente, come si diceva nel precedente paragrafo 3.3, altri dati potrebbero non essere considerati altrettanto accessibili.

---

all'impresa in posizione dominante di condividere una risorsa-input non essenziale; ovvero di iv) produrre un bene (o una combinazione di beni) che il monopolista non impiega in alcun modo; o infine v) di costituire una nuova joint venture con il concorrente.

[273] Corte di Giustizia dell'Unione Europea, 16 luglio 2015, caso C-170/13.

[274] Commission Staff Working Document on *The free flow of data and emerging issues of the European data economy*, di accompagnamento alla Comunicazione della Commissione *Building a European data economy*, 2017". Vedi anche, la Comunicazione della Commissione Europea, *Online Platforms and the Digital Single Market*, 2016, laddove la Commissione considera il rifiuto a concedere l'accesso ad un data business essenziale come una delle più comuni presunte questioni sottolineate anche dalla consultazione pubblica sulle pratiche commerciali scorrette praticate dalle piattaforme.

[275] Colangelo G., Maggiolino M., *Big Data as Misleading Facilities, European Competition Journal* (in corso di pubblicazione) dove ulteriori riferimenti alla letteratura in materia.

[276] Corte di Giustizia dell'Unione Europea, 26 novembre 1998, nota 18, par. 44.

[277] Ibid., par. 45-46.

Quanto al requisito dell'esclusione della concorrenza effettiva in un mercato secondario, la sua realizzazione nello scenario dei dati solleva alcune questioni pratiche. Come detto, questa condizione è soddisfatta nell'ipotesi in cui l'impresa che detiene l'input essenziale sia già operante nel mercato a valle e, rifiutandosi di concederne l'accesso, miri a garantirsi l'intero mercato. Tuttavia, questa pratica non è frequente nell'ipotesi di rifiuto alla condivisione dei dati, perché raramente le imprese che detengono i dati operano anche nel mercato secondario.

Il rispetto poi del requisito del nuovo prodotto è ancora più complesso e problematico. A questo proposito, due elementi fondamentali distinguono la giurisprudenza europea. In primo luogo, i ricorrenti non hanno voluto accedere a tutti i dati di cui i convenuti dispongono, ma ad un insieme ben identificato di dati destinati a servire un ambito specifico. In secondo luogo, i convenuti hanno voluto accedere ai dati, non per ipotizzare come utilizzarli, ma per produrre nuovi beni (innovazioni) già identificati. Pertanto, la Commissione e le corti UE non hanno considerato tali dati come risorse essenziali in termini generici, ma in relazione alla creazione di alcune innovazioni specifiche. Pertanto, una risorsa è essenziale quando serve - unicamente - ad offrire un prodotto o un servizio specifico, ossia quando esiste una relazione causa-effetto tra la risorsa e il bene o il servizio che il concorrente intende realizzare, utilizzando quella stessa risorsa. Inoltre, il nuovo requisito del prodotto mira a garantire la caratteristica innovativa del bene o del servizio, creando un equilibrio tra gli interessi privati alla tutela del diritto di proprietà intellettuale e l'interesse generale per la protezione della concorrenza e dell'innovazione.

A riprova della difficoltà a verificare questo requisito, si potrebbe inoltre sostenere che quando è l'offerta dei dati a determinare la domanda di beni e servizi, le imprese non possono sapere a priori i beni che andranno a produrre attraverso l'utilizzo dei dati prima di avere accesso agli stessi. Di conseguenza, l'applicazione del requisito del nuovo prodotto ai big data solleva due questioni specifiche: come sarebbe possibile che un concorrente del monopolista sostenga di aver bisogno dei big data del rivale se non sa che genere di informazioni quei big data rivelino? Ancora, come potrebbe questo rivale del monopolista garantire l'offerta di determinati beni basata sull'utilizzo dei big data del monopolista, se non conosce le informazioni racchiuse in questi dati? Non bisogna infatti dimenticare che una delle caratteristiche rivoluzionarie dei big data è che vengono raccolti disordinatamente. Non sono raccolti e organizzati per rispondere a determinate domande di ricerca. Al contrario: in primo luogo, derivano da molte e diverse fonti; quindi, vengono poi analizzati per acquisire un valore informativo. Ma se questo è vero, come è possibile che un concorrente affermi di aver bisogno di questi dati e soltanto di questi, per offrire un certo bene sul mercato, se non può sapere quali informazioni possono essere acquisite attraverso la loro elaborazione?

Peraltro, anche se i big data fossero considerati essenziali, l'imposizione di un obbligo a contrarre porrebbe non pochi problemi.

3.6.3.1 Le difficili licenze obbligatorie

Le licenze obbligatorie rappresentano il rimedio tradizionale contro questo tipo di comportamento anticoncorrenziale. Tuttavia, la condivisione forzata richiede che le autorità antitrust si comportino come pianificatori centrali e assumano gli oneri di un regolatore, un ruolo per cui potrebbero non

essere adatte. Ipoteticamente, anche se i requisiti dell'EFD fossero stati soddisfatti, le licenze obbligatorie per i dati sarebbero difficili da gestire per diversi motivi[278].

In primo luogo, dovrebbero essere individuati i dati soggetti all'obbligo di condivisione. In ogni caso un obbligo di condivisione presuppone l'identificazione di un insieme ben definito di risorse. Tuttavia, ciò non è applicabile alla condivisione dei dati poiché l'impresa richiedente non è in grado di conoscere ex ante il set di dati necessario nell'intero set di dati grezzi di cui dispone l'*incumbent*. Infatti, si tratterebbe di considerare una certa quantità di dati di dimensioni elevate e sconosciute, il cui contenuto non è specificato. Pertanto, la questione riguarda la possibilità che un concorrente possa rivendicare la necessità di un determinato data set per produrre e offrire un bene o un servizio, senza essere in grado di sapere, prima di averne accesso, quale è l'esatto insieme di dati di interesse.

Inoltre, cosa succederebbe se questi dati venissero aggiornati in tempo reale? Il valore dei dati diminuisce notevolmente nel tempo, dunque ha una durata limitata. Se i dati sono quindi considerati essenziali per la concorrenza, tali per cui debbano essere condivisi, l'impresa richiedente deve poter godere di dati per così dire "freschi" e aggiornati. Ciò quindi equivarrebbe a sostenere che ci dovrebbe essere un continuo, illimitato flusso di dati dal monopolista a coloro che ne facciano richiesta?

Un'altra importante preoccupazione riguarda i termini e le condizioni per le licenze. Nella comunicazione "*Building a European Data Economy* ", la Commissione Europea ha avviato una consultazione pubblica volta a discutere, tra le diverse questioni, la possibilità di remunerazioni basate su termini equi, ragionevoli e non discriminatori (F/RAND). Nondimeno, come l'esperienza accumulate in tema di *standard setting organizations* mostra, i termini F/RAND sono difficili da definire ed apprezzare[279], anche quando (o soprattutto quando) letti alla luce della decisione Huawei e delle modalità di negoziazione da essa individuate. Ad esempio, è difficile immaginare come potrebbe essere determinata la royalty di eventuali licenze obbligatorie, a causa dell'incertezza circa le informazioni contenute nei big data, ossia circa le nuove informazioni che possono essere tratte dai dati esistenti attraverso l'estrazione dei dati e le grandi analisi dei dati. Ma se queste informazioni non si conoscono il valore dei dati e, dunque, il prezzo appropriato da pagare per l'accesso sono altrettanto difficili da stabilire. Il cosiddetto paradosso informativo rende particolarmente difficile l'accordo sul prezzo dell'accesso alle informazioni nei negoziati contrattuali.

In ultimo, se l'attenzione è posta sui dati utente delle piattaforme digitali, tali piattaforme come possono essere autorizzate a condividere i dati, se si tratta principalmente di dati personali? [280] Secondo le leggi sulla protezione dei dati personali, gli utenti condividono i propri dati personali

[278] Colangelo G., Maggiolino M., *Big Data as Misleading Facilities*, cit., dove ulteriori riferimenti sul punto.

[279] Swanson D.G. and Baumol W.J., *Reasonable and Nondiscriminatory (RAND) Royalties, Standards Selection, and Control of Market Power,* 73 Antitrust L.J. 1, 2005.

[280] Sulle implicazioni antitrust-privacy nei mercati digitali, vedi Colangelo G. e Maggiolino M., *Data Protection in Attention Markets: Protecting Privacy through Competition?*, 8 Journal of European Competition Law and Practice 363 (2017).

sapendo chiaramente come verranno utilizzati nello specifico dalla piattaforma. Nel caso in cui, un concorrente della piattaforma digitale dovesse richiedere quegli stessi dati, la piattaforma in che modo potrebbe condividerli senza violare le leggi sulla privacy? Chiedere agli utenti di dare un largo consenso a favore di qualsiasi concorrente della piattaforma digitale per eventuali richieste di dati sarà contraria ai principi che informano le leggi sulla privacy.

3.7 I big data: tra privacy e antitrust

Il rapporto tra il diritto antitrust e la disciplina a tutela dei dati personali e della privacy può essere declinato lungo diverse direttrici, alcune delle quali portano a conclusioni difficili da revocare in dubbio.

In primo luogo, vero è che le nuove regole in materia di portabilità dei dati[281], le quali riguarderanno i dati personali volontariamente trasferiti e quelli tracciati[282], aumenteranno la concorrenza tra le imprese che analizzano quei dati personali per produrre conoscenza e valore. Meglio, la possibilità per gli individui di trasferire questi dati produrrà due effetti congiunti. Innanzitutto, essa ridurrà il fondamento empirico dell'argomento secondo cui i consumatori sarebbero "vincolati" alle imprese – ed, in particolare, alle piattaforme digitali – perché queste deterrebbero i loro dati e sulla scorta di quelli migliorerebbero i propri servizi. Infatti, i consumatori che vorranno utilizzare i servizi altrui, trovandoli preferibili per altre ragioni, non saranno più "penalizzati" dall'eventualità di essere poco conosciuti dalla nuova impresa. Grazie al diritto alla portabilità, questa potrà ottenere i loro dati personali e quindi ritagliare il servizio sulle esigenze di questi consumatori. Inoltre, e conseguentemente, la portabilità dei dati ridurrà la concorrenza sui dati personali, e dunque anche la fondatezza empirica dell'argomento secondo cui sarebbe la mera detenzione di questi dati a costituire un vantaggio competitivo. Le imprese dovranno piuttosto competere su altre leve concorrenziali, come le tecniche di analisi dei dati che consentono di meglio conoscere i consumatori e dunque di meglio soddisfarli.

In secondo luogo, vero è che sarebbe del tutto fisiologico un intervento antitrust che, nei mercati per la commercializzazione di prodotti e servizi "privacy-friendly" o di tecnologie "privacy-enhancing", fosse teso a sanzionare eventuali strategie capaci di ridurre la quantità e la qualità di quei beni e di quelle tecnologie. In questo scenario, cioè, l'azione delle autorità antitrust sarebbe del tutto ortodossa: esse non sarebbero chiamate a sviluppare alcuna particolare teoria del danno, perché quei comportamenti sarebbero semplicemente riduttivi del consumer welfare di coloro che acquistano prodotti e servizi "privacy-friendly" e di tecnologie "privacy-enhancing". Del resto nelle decisioni *Facebook/WhatsApp* e *Microsoft/LinkedIn* la Commissione ha osservato per la prima volta

[281] Vedi Art. 20, Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE, anche detto "Regolamento Generale Sulla Protezione Dei Dati".

[282] Vedi Article 29 Data Protection Working Party, *Guidelines on The Right to Data Portability*, 13 dicembre 2016, p. 8.

quanto si accennava sopra, ossia che i consumatori potrebbero vedere la privacy come un fattore significativo che incide sulla qualità dei beni e dei servizi disponibili su internet. Di conseguenza, la Commissione ha anche evidenziato che le piattaforme digitali possono competere per fornire prodotti o servizi più o meno rispettosi della privacy[283]. Nondimeno, va precisato che in quelle decisioni la Commissione non ha elaborato alcuna conclusione a partire da queste affermazioni, né le ha discusse ulteriormente. In particolare, la Commissione non ha valutato se i mercati considerati fossero sufficientemente concorrenziali per fornire agli utenti il livello ottimo di prodotti e servizi *privacy-friendly* o di tecnologie privacy-enhancing. In altre parole, la Commissione non ha individuato alcun legame tra il potere di mercato delle piattaforme sottoposte a scrutinio e la qualità e varietà degli strumenti e dei servizi da queste forniti a tutela della privacy dei loro utenti. Né la Commissione ha trovato alcuna connessione tra tale potere di mercato e la quantità ottimale di dati personali che le imprese coinvolte avrebbero dovuto raccogliere. Tuttavia, non c'è nulla di eterodosso in questo approccio della Commissione. L'individuazione del livello ottimale di soluzioni *privacy-friendly* o di tecnologie privacy-enhancing, o ancora la determinazione del livello ottimale di dati personali che delle piattaforme dovrebbero raccogliere non è un problema del diritto antitrust. Esso, come si diceva all'inizio di questa seconda parte, deve intervenire quando il mercato si muove verso uno scenario peggiore rispetto a quello di concorrenza effettiva. Il diritto antitrust non agisce per promuovere mercati più concorrenziali – questo potrebbe, al più, essere un problema per i regolatori[284]. Nei casi *Facebook/WhatsApp* e *Microsoft/LinkedIn* non c'era dunque alcuno spazio per l'intervento della Commissione, perché essa aveva escluso l'unico profilo rilevante sotto il profilo antitrust, ossia che quelle concentrazioni influissero sull'offerta di prodotti, servizi e tecnologie capaci di meglio tutelare la privacy.

In terzo luogo, vero è che si potrebbe dare una circostanza in cui una medesima condotta fosse plurioffensiva, ossia tale da ledere al contempo il corretto funzionamento del mercato e le disposizioni a tutela dei dati personali. In questo caso, esisterebbero i presupposti affinché le due discipline intervenissero congiuntamente. Questa è ad esempio la visione che sembra aver determinato il Bundeskartellamt ad istruire il caso Facebook. Pare infatti che a Facebook, qualificata come impresa in posizione dominante nel mercato dei social network, si contesti di aver abusato del suo potere di mercato imponendo ai suoi utenti dei *terms of use* che al contempo violerebbero non solo le norme a tutela dei dati personali, ma anche la lett. a) dell'art. 102 TFUE. Se le accuse fossero fondate, e al netto di qualsiasi valutazione circa il principio del *ne bis in idem*, Facebook avrebbe violato le due discipline, posto che evidentemente il Bundeskartellamt agirebbe solo per la parte che gli compete, ossia qualora i termini chiamati a governare il trasferimento dei dati degli utenti fossero abusivi perché iniqui e ingiustificatamente gravosi.

Certo, nelle more della decisione, sorge una perplessità e serve elaborare una precisazione.

[283] Facebook/WhatsApp, fn 21, para 87 and Microsoft/LinkedIn, fn 21, nt. 330.

[284] Monopolkommission, fn 12, paras 306-311.

La perplessità riguarda l'eventualità che la violazione delle norme sulla privacy diventi il parametro per misurare quando una condizione contrattuale in materia di dati personali diventi ingiustificatamente gravosa ai sensi dell'art. 102 (a). In altri termini, nelle more della decisione, viene da pensare al rischio che il giudizio antitrust circa l'iniquità dei termini contrattuali si "appiattisca" sulla valutazione di privacy, quasi creando un automatismo, ossia l'idea che una violazione della disciplina a tutela dei dati personali rappresenterebbe per ciò stesso un abuso, quando posta in essere da una impresa in posizione dominante, nonché – se l'automatismo trovasse applicazione – un'intesa restrittiva della concorrenza, qualora la violazione della privacy fosse decisa congiuntamente da due o più imprese, oppure una ragione per considerare anticompetitiva una concentrazione, qualora l'impresa risultante dall'operazione fosse in grado di imporre termini contrattuali lesivi della privacy. Ora, non vi è chi non veda come l'esistenza di questo automatismo farebbe della plurioffensività delle condotte circa i dati un elemento puramente formale, subordinando l'applicazione o meno del diritto antitrust alla verifica dell'esistenza di una posizione dominante, di una intesa o di una concentrazione. Ecco perché sarebbe meglio se l'eventuale iniquità delle condotte di un soggetto che raccoglie dati personali (e non solo) fosse ricondotta alla autonoma nozione antitrust di unfairness per come essa è desumibile dalla giurisprudenza in tema di art. 102(a)[285].

La precisazione riguarda l'ipotesi che i termini dell'accordo per effetto del quale un individuo trasferisce i suoi dati personali, ancorché rispettosi della disciplina a tutela di quei dati, siano invece giudicati "autonomamente" lesivi delle regole antitrust o – come si vedrà nel prossimo paragrafo – delle norme a tutela dei consumatori. In altri termini, di là dall'intersezione tra privacy e antitrust, è sempre possibile riconoscere una violazione indipendente delle disposizioni antitrust – e in particolare dell'art. 102, lett a) – che possa essere integrata da una condotta connessa al mondo dei dati personali. Questo, del resto, è quello che sembrano preconizzare le autorità antitrust francese e tedesca quando scrivono: «privacy policies could be considered from a competition standpoint whenever these policies are liable to affect competition, notably when they are implemented by a dominant undertaking for which data serves as a main input of its products or services. In those cases, there may be a close link between the dominance of the company, its data collection processes and competition on the relevant markets, which could justify the consideration of privacy policies and regulations in competition proceedings»[286].

Ma qui terminano i "punti fermi" – o qualificabili come tali – dell'interfaccia tra diritto antitrust e tutela della privacy. Tema molto più complesso è quello che attiene alla convergenza tra diritto antitrust e disciplina della privacy che si attuerebbe attraverso la teoria del danno al consumer welfare che aggancerebbe una riduzione della privacy alla riduzione della qualità dei beni offerti dalle imprese. Più chiaramente, lo scenario che ora si considera è quello di un comportamento

[285] Kalimo H., Majcher K., *The concept of fairness: linking EU competition and data protection law in the digital marketplace, European Law Review*, 2017, p. 210, dove si legge di come la nozione di fairness sia stata declinata, nel corso degli anni e rispetto a singoli casi, in termini di «absolute necessity, onesidedness, equality, proportionality, balance of interests, transparency, objectivity and certainty» – p. 225.

[286] Autorité de la concurrence & Bundeskartellamt, *'Competition Law and Data*', report, 2016, pp. 23-24

imprenditoriale che, anche senza ledere alcuna disposizione in materia di tutela dei dati personali, riduca il livello di privacy che complessivamente si può garantire agli individui e la questione che dunque si pone è se in questo scenario un siffatto comportamento debba essere considerato anticompetitivo perché lesivo del consumer welfare. Del resto, in questo senso sembrano indirizzarsi le considerazioni critiche indirizzate alla volta della concentrazione *Facebook/WhatsApp*[287]. Vero, come visto nel paragrafo 3.5.4, quella concentrazione non poteva dirsi capace di ledere in modo significativo la concorrenza. Tuttavia, l'impresa nascente da quella operazione, combinando i big data delle due imprese, vedeva amplificarsi la propria capacità di conoscere e profilare i suoi utenti, ossia – in una battuta – vedeva accrescersi la sua capacità di invadere la loro privacy.

Ora, coloro che vorrebbero qualificare come anticompetitive questa ed altre condotte analoghe sostengono che un comportamento che aumenta la capacità delle imprese di ledere la privacy, riduce il livello di privacy offerto agli utenti e diminuisce altresì la qualità dei prodotti e servizi offerti sulla scorta dei dati così raccolti e utilizzati. E giacché – questi proseguono – il *consumer welfare* si riduce quando la qualità dei prodotti e servizi offerti sul mercato diminuisce, una condotta che abbassa il livello di privacy presente nel mercato è una condotta anticompetitiva[288].

Tuttavia, non sono poche le voci in dottrina che contestano alcuni passaggi di questo ragionamento[289] ed, in particolare, tre punti: (*i*) la tesi secondo cui la dimensione della qualità connessa al livello di privacy garantito dovrebbe fare premio rispetto alle altre proprietà (come la velocità e l'accuratezza) di un prodotto o servizio che evidentemente contribuiscono anche esse a determinare la qualità complessiva di un bene; (*ii*) l'idea che sempre questa dimensione della qualità connessa al livello di privacy garantito debba prevalere anche sulle altre variabili da cui dipende il consumer welfare, quali prezzo, output, varietà e innovazione; nonché (*iii*) la possibilità che, tecnicamente, una singola dimensione della qualità di un prodotto possa essere apprezzata da una autorità antitrust e confrontata con le altre variabili da cui dipendono tanto la qualità complessiva del bene considerato e quanto il più generale *consumer welfare*.

Inoltre, sino ad oggi la Commissione Europea non ha mai utilizzato questa *theory of harm*, né ha mai scelto di usare il diritto antitrust per tutelare i beni giuridici connessi alla disciplina sulla protezione dei dati personali. Già nel 2006, in *Asnef-Equifax*, mentre discuteva di un caso di scambio di informazioni ai sensi dall'art. 101 del TFUE, la Corte di giustizia scriveva, «*any possible issues relating*

[287] COMP/M.7217.

[288] Stucke M.E. and Ezrachi A., *When Competition Fails to Optimise Quality: A Look at Search Engines*, 18 Yale Journal of Law and Technology 70, 2016; Grunes A.P. and Stucke M.E., *No Mistake About It: The Important Role of Antitrust in the Era of Big Data*, 14 Antitrust Source 1, 2015; Grunes A.P., *Another Look at Privacy*, 20 Geo. Mason L. Rev. 1107, 2013; Pasquale F., *Privacy, Antitrust, and Power*, 20 Geo. Mason L. Rev. 1009, 2013; Swire P., *Protecting Consumers: Privacy Matters in Antitrust Analysis*, Center for American Progress, 2007.

[289] Manne G.A. and Sperry R.B., *The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework*, 5 CPI Antitrust Chronicle, 2015; Cooper J.C., *Privacy and Antitrust: Underpants Gnomes, the First Amendment, and Subjectivity*, 20 Geo. Mason L. Rev. 1129, 2013; Sokol D.D. and Comerford R., *Antitrust and Regulating Big Data*, 23 Geo. Mason L. Rev. 1129, 2016.

to the sensitivity of personal data are not, as such, a matter for competition law, they may be resolved on the basis of the relevant provisions governing data protection».[290]

Successivamente, nel 2008, con riguardo alla concentrazione *Google/DoubleClick* ha riaffermato, coerentemente con quanto già stabilito dalla FTC,[291] una netta e chiara separazione tra l'ambito di applicazione del diritto antitrust e il raggio di azione delle regole in materia di privacy e protezione dei dati personali. La Commissione ha chiaramente affermato che la sua decisione «*refers exclusively to the appraisal of [the] operation with Community rules on competition, namely whether the merger … does not impede effective competition in the common market. … [The] [d]ecision is without prejudice to the obligations imposed onto the parties by Community legislation in relation to the protection of individuals and the protection of privacy with regard to the processing of personal data. … [Therefore] [i]rrespective of the approval of the merger, the new entity is obliged in its day to day business to respect the fundamental rights recognized by all relevant instruments to its users, namely but not limited to privacy and data protection*».[292] In altri termini, attribuendo la difesa dei dati personali alla disciplina a tutela degli stessi, la Commissione ha escluso la sua competenza circa le ripercussioni sulla privacy che la combinazione dei big data di Google e DoubleClick avrebbe potuto causare.

Più recentemente, tra il 2014 e il 2016, la Commissione ha tenuto il medesimo approccio nelle decisioni *Facebook/WhatsApp* e *Microsoft/LinkedIn*[293]. Ancora una volta – per quanto qui più interessa – la Commissione non ha preso in considerazione alcun profilo connesso alla privacy[294]. In *Facebook/WhatsApp* ha osservato che *«[a]ny privacy-related concerns flowing from the increased concentration of data within the control of Facebook as a result of the Transaction do not fall within the scope of the EU competition law rules but within the scope of the EU data protection rules*»[295]. Analogamente, in *Microsoft/LinkedIn* la Commissione ha chiaramente affidato alle norme sulla privacy con il compito di proteggere i dati personali degli utenti. In primo luogo, la Commissione ha affermato che *«[the] data combination could only be implemented by the merged entity to the extent it is allowed by applicable data protection rules»*.[296] La Commissione ha anche evidenziato come il

[290] Case C-238/05 Asnef-Equifax, [2006] ECR I-11125, par. 63.

[291] Federal Trade Commission, File No. 071-0170, Google/DoubleClick [2007], 2 dove si legge che «the sole purpose of federal antitrust review of mergers and acquisitions is to identify and remedy transactions that harm competition. Not only does the Commission lack legal authority to require conditions to this merger that do not relate to antitrust, regulating the privacy requirements of just one company could itself pose a serious detriment to competition in this vast and rapidly evolving industry».

[292] Google/DoubleClick, cit., para 368.

[293] Facebook/WhatsApp (Case COMP/M.7217) Commission Decision [2014] OJ C417/4; e Microsoft/LinkedIn (Case COMP/M.8124) Commission Decision [2017].

[294] Facebook/WhatsApp, cit., paras 188-189, and Microsoft/LinkedIn, cit., paras 179-180.

[295] Facebook/WhatsApp, cit., para 164.

[296] Microsoft/LinkedIn, cit., para 177.

nuovo regolamento in tema di protezione dei dati personali *«may further limit Microsoft's ability to have access and to process its users' personal data in the future since the new rules will strengthen the existing rights and empowering individuals with more control over their personal data»*.[297] *In secondo luogo, la Commissione ha discusso i profile antitrust connessi ai due casi, «[a]ssuming such data combination is allowed under the applicable data protection legislation»*.[298] Di nuovo, dunque, la Commissione ha scelto di riferirsi alle regole sulla privacy per proteggere i dati personali degli individui e il controllo che le persone hanno sulla loro identità digitali.

Da ultimo, occorre ricordare un'ulteriore forma di convergenza tra la disciplina della privacy e il diritto antitrust, anch'essa evidenziata in commento alla concentrazione *Facebook/WhatsApp* e relativa alla possibilità già menzionata che una maggiore capacità di estrarre dati dagli utenti equivalga ad imporre un prezzo sovra-competitivo e sia per questo indice di un maggiore potere di mercato dell'impresa risultante dalla concentrazione[299].

A ben vedere, la decisione sulla fusione *Facebook/WhatsApp* non aveva trattato questo particolare aspetto, anche perché si era basata sulle dichiarazioni delle parti, secondo le quali esse non avrebbero combinato i loro dati. Per questo, dunque, rispetto alla quantità di informazioni estrapolate ed estrapolabili da ogni utente, la situazione post-fusione sarebbe stata identica a quella antecedente alla fusione.

Tuttavia, alla luce anche dei successivi accadimenti, se si ritiene che i dati estratti dagli individui siano la "moneta" con la quale quegli individui pagano i servizi e i prodotti offerti dalle piattaforme, allora, come abbiamo visto in precedenza, un aumento dei dati estratti ed estraibili dai consumatori corrisponde ad un aumento dei prezzi nei mercati tradizionali.

Nella fattispecie, l'impatto della fusione in questione, che coinvolge due operatori primari come Facebook e Whats App, è certamente più ampio di quello delle altre imprese di communications services che utilizzano user data. In tal caso, in termini antitrust, ciò che rileva è quanto una riduzione della privacy (analogamente a una riduzione della qualità o ad un aumento del prezzo monetario) non possa essere "sfidata" da dei concorrenti potenziali, quando la nuova entità è schermata da imponenti barriere all'ingresso.

La Commissione, nella sua decisione, non definisce il mercato degli user data e quindi non considera questa concentrazione sommando orizzontalmente le loro quote in quel mercato. L'idea alla base del ragionamento sottostante è che analizzando i mercati separati della concentrazione non si rilevava alcun impatto anti-competitivo, perché in ciascun mercato esisteva una valida offerta concorrenziale, e in ciascuno di quelli individuati (3) c'era la prevalenza di uno dei due soggetti e il secondo era marginale o secondario.

[297] Microsoft/LinkedIn, cit. 21, para 178.

[298] Microsoft/LinkedIn, cit. 21, para 179.

[299] Deutscher E., *The role of personal data and privacy in merger analysis*, 2017.

L'obiezione che emergeva era duplice: per un verso non si teneva conto della ragione per cui avveniva la concentrazione (se era così quale era il vero vantaggio di farla); per un altro non si teneva conto del mercato dei dati, che rappresentava la motivazione vera per cui questa operazione avveniva, dal momento che la stessa Facebook si era affannata a rassicurare la Commissione che non avrebbe integrato le informazioni dei due gruppi di utenti (cosa che poi si è verificato successivamente al via libera della Commissione)[300].

In tutti i casi, indipendentemente dalle soluzioni adottate, il tema della privacy appare sempre più correlato a profili concorrenziali e regolamentari e si prospetta come una delle prospettive di analisi che saranno maggiormente esplorate d'ora in avanti.

3.8 I big data: tra consumer protection e antitrust (cenni)

Tradizionalmente il rapporto tra la tutela del consumatore e il diritto antitrust viene inteso in termini di complementarità[301]. Da un lato, la politica della concorrenza guarda al mercato dal lato dell'offerta, preoccupandosi cioè che, attraverso la concorrenza tra imprese, i consumatori possano beneficiare della più ampia gamma di prodotti e servizi possibili ai prezzi più bassi possibili. Dall'altro lato, la politica dei consumatori si avvicina al mercato dal lato della domanda, garantendo che i consumatori siano in grado di esercitare intelligentemente ed efficacemente le scelte che la concorrenza consente loro di compiere[302].

Pertanto, il diritto antitrust non è solito occuparsi di quanto i consumatori siano ben istruiti, informati e consapevoli. In altri termini – di là dai già considerati casi di manipolazione dell'informazione e di imprese in posizione dominante che impongano condizioni contrattuali ingiustificatamente gravose (cfr. paragrafi 3.2 e 3.2) – il diritto antitrust non guarda quanto

[300] L'obiezione che emergeva era duplice: per un verso non si teneva conto della ragione per cui avveniva la concentrazione (se era così quale era il vero vantaggio di farla); per un altro non si teneva conto del mercato dei dati, che rappresentava secondo i fautori del mercato degli user data, la motivazione vera per cui questa operazione avveniva, dal momento che la stessa Facebook si era affannata a rassicurare la Commissione che non avrebbe integrato le informazioni dei due gruppi di utenti (cosa che poi si è verificato successivamente al via libera della Commissione). Vedasi Deutscher E., "op. cit.", EUI, 2017.

[301] In tal senso anche il discorso dell'allora commissario Kuneva, la quale affermava che « it is clear that competition policy and consumer policy are complementary. Competition is a necessary condition for consumer welfare. Consumer policy in turn enhances competition by empowering consumers to exercise the power of choice» – OECD, *The Interface Between Competition And Consumer Policies*, DAF/COMP/GF(2008)10, p. 106, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP/GF(2008)10&docLanguage=En. Inoltre, Hadfield G.K., Howse R., and Trebilcock M. J., *Information-Based Principles for Rethinking Consumer Protection Policy*, 21 Journal of Consumer Policy 1998, p. 150 che hanno infatti affermato «antitrust policy focuses on the structure of markets, and consumer protection policy focuses on the structure of consumer transactions», nonché Amato G., Schwitzer H., Denozza F., Stallibrass D., e Nicita A. raccolti in *Tutela della concorrenza e tutela dei consumatori*, Merc. conc. reg., 2009, 2.

[302] OECD, *The Interface Between Competition And Consumer Policies*, DAF/COMP/GF(2008)10, p. 8 e 19, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP/GF(2008)10&docLanguage=En.

ponderate, fondate e responsabili siano le scelte dei consumatori. Questo compito è piuttosto svolto dalla disciplina posta a tutela dei consumatori ed eventualmente dalla disclosure regulation.

Non è infatti un caso che l'Autorità Garante abbia recentemente fatto ricorso alla disciplina delle pratiche commerciali scorrette per sanzionare WhatsApp. In particolare, nel procedimento PS10601, l'Autorità ha comminato a WhatsApp una sanzione di 3 milioni di euro per aver indotto gli utenti di WhatsApp Messenger ad accettare integralmente i nuovi termini di utilizzo del servizio ed, in particolare, per averli spinti a condividere i propri dati con Facebook, facendo loro credere che sarebbe stato altrimenti impossibile proseguire nell'uso dell'applicazione[303]. Analogamente, nel secondo procedimento CV154 ancora in corso la Autorità sta considerando la presunta natura vessatoria di alcune clausole del modello contrattuale sottoposto all'accettazione dei consumatori che vogliano usufruire dell'applicazione WhatsApp Messenger[304].

I casi appena menzionati inducono dunque a ritenere come la disciplina a tutela dei consumatori potrebbe supportare la tutela della privacy e, in particolare, contribuire a che i consumatori meglio controllino i propri dati personali e la propria identità digitale più e meglio di quanto non riesca a fare il diritto antitrust.

[303] Cfr. casi PS10601, 11 maggio 2017. La condotta in esame è stata attuata attraverso una procedura in-app di accettazione dei nuovi Termini di Utilizzo tramite la quale: (i) si rappresentava agli utenti la necessità di tale accettazione, entro 30 giorni, a pena di dover interrompere la fruizione del servizio; (ii) non si informavano gli utenti della possibilità di poter negare il consenso alla condivisione dei dati con Facebook; e (iii) si offriva l'opzione opt-in come quella pre-selezionata. Inoltre, una volta accettati integralmente i termini, l'esercizio dell'opt-out sembrava assai difficile. Coloro che erano già utenti alla data della modifica dei termini (25 agosto 2016) avevano, invece, la possibilità di accettare "parzialmente", potendo decidere di non fornire l'assenso a condividere le informazioni del proprio account con Facebook e continuare, comunque, a utilizzare l'applicazione.

[304] Cfr. CV154, 11 maggio 2017. In particolare, ad essere oggetto di scrutinio sono le seguenti previsioni contrattuli: (a) esclusioni e limitazioni di responsabilità in capo a WhatsApp molto ampie e assolutamente generiche; (b) la possibilità di interruzioni del servizio decise unilateralmente da WhatsApp senza motivo e senza preavviso; (c) il diritto generico esercitabile da WhatsApp di risolvere il contratto/recedere in qualsiasi momento e per qualsiasi motivo e non consentire più all'utente l'accesso/utilizzo dei servizi, senza prevedere un analogo diritto per il consumatore; (d) il diritto generico esercitabile da WhatsApp di introdurre modifiche, anche economiche, dei Termini di Utilizzo senza che nel contratto vengano preventivamente indicate le motivazioni e senza neppure prevedere modalità per informarne in maniera adeguata l'utilizzatore, unitamente alla previsione del meccanismo di "silenzio assenso" che fa discendere l'accettazione dei nuovi Termini anche solo dalla mera inerzia inconsapevole dell'utente; (e) la scelta della legge dello Stato della California quale unica legge applicabile al contratto e la scelta dei tribunali californiani come unici fori competenti; (f) un generico diritto esercitabile da WhatsApp di recedere dagli "ordini" e di non fornire rimborsi per i servizi offerti, senza precisare in modo chiaro il contesto in cui tali operazioni si esplicherebbero; (g) la generale prevalenza del contratto scritto in lingua inglese, in caso di conflitto con la versione tradotta in lingua italiana (accettata dall'utente), senza prevedere la prevalenza dell'interpretazione più favorevole al consumatore, a prescindere dalla lingua in cui la clausola è redatta.

Capitolo 4. La regolamentazione economica e di settore

4.1 Big data e regolazione delle comunicazioni

Nei capitoli precedenti si è visto come i big data stiano divenendo uno strumento sempre più essenziale per le imprese per riuscire a competere nel mercato. Avere ed elaborare un'ampia quantità di dati è divenuto fondamentale in diversi settori per rendere più efficienti i processi produttivi e personalizzare servizi e beni offerti ai consumatori. L'interesse del diritto antitrust per i big data nasce proprio dal sempre più ampio impiego di questo strumento, che ha reso necessario osservare i casi in cui le imprese ne fanno un utilizzo inappropriato, alterando così una corretta dinamica concorrenziale e producendo un danno al consumatore finale o ad altre imprese.

In ogni caso l'intento dell'antitrust non è tanto finalizzato a disciplinare il funzionamento del mercato, quanto piuttosto a intervenire ex-post in caso venga ravvisato un illecito, punendo eventuali comportamenti ritenuti lesivi del *consumer welfare* e per questo giudicati meritevoli di sanzione. Vi sono, tuttavia, determinati settori nei quali, per le loro particolari caratteristiche, viene considerato più opportuno un intervento ex-ante, teso a correggere preventivamente situazioni potenzialmente dannose per l'utente finale. In questi casi interviene la regolazione economica che, a differenza dell'antitrust, cerca di favorire il corretto funzionamento del mercato mediante una serie di regole che impongono agli operatori comportamenti ritenuti necessari per l'instaurarsi di una concorrenza virtuosa e dinamica.

Essendo l'azione del regolatore, dunque, più "invasiva", si ritiene necessaria solamente in circostanze particolari, quelle in cui si reputa necessario un suo intervento, al fine di ripristinare le corrette dinamiche del mercato. Ciò si verifica quando si rileva la presenza di fallimenti di mercato, ovvero in tutti quei casi in cui il sistema economico non è in grado di raggiungere la piena efficienza in quanto i fattori impiegati non sono utilizzati in modo ottimale.

Inoltre ipotizzare un intervento regolatorio sui big data rappresenta a sua volta una situazione *sui generis*, in quanto si andrebbe a regolare non tanto un mercato, quanto un input, utilizzato in modo trasversale in diversi settori. I big data, occorre ricordare, non vanno a costituire un mercato a sé stante, ma, per via della loro grande efficacia e diffusione, possono incidere in maniera considerevole sulle dinamiche concorrenziali dei diversi settori e solo in questa chiave si può considerare appropriato l'intervento del regolatore.

4.1.1 Le comunicazioni elettroniche

In generale l'azione del regolatore si esercita su un intero settore economico. In particolare in quei settori strategici, caratterizzati dalla presenza di reti (telecomunicazioni, trasporti, energia), la regolazione ha assunto un ruolo preponderante nel definire una serie di regole che garantissero l'accesso a condizioni eque e non discriminatorie nei mercati a valle. In questo senso l'interrelazione sempre più ampia tra i settori economici spinge nell'economia digitale, per le caratteristiche approfondite nella prima parte, al superamento della tradizionale regolazione verticale, in favore di

una orizzontale e sempre più convergente tra le diverse industries coinvolte, collegate dal tema della connettività e caratterizzate da sempre più ampie interrelazioni (smart cities, connected cars, AI, ecc..).

Per questo motivo con riferimento ai big data, per via della loro crescente importanza nell'economia digitale e con le premesse appena ricordate, ci si chiede se debbano essere soggetti o meno a una regolazione ad hoc volta a disciplinarne l'utilizzo.

L'Autorità per le Garanzie nelle Comunicazioni è fra le autorità di regolazione di settore quella che guarda con particolare interesse ai big data e ai loro effetti sulla concorrenza nei mercati. Fra le sue prioritarie funzioni infatti, ai sensi del Codice delle Comunicazioni elettroniche (d.lgs 1 agosto 2003, n. 259), c'è quella di garantire un accesso al mercato secondo criteri di obiettività, trasparenza, non discriminazione e proporzionalità per i servizi di sua competenza. Fra di essi rientrano tutti i servizi di comunicazioni fissi e mobili, i servizi media audiovisivi e i servizi postali.

Ognuno di essi si inserisce all'interno di settori interessati allo sviluppo dei big data, in quanto caratterizzati da un elevato numero di consumatori che, essendo interconnessi tramite la piattaforma IP alle reti digitali, costituiscono non solo grandi utilizzatori dei servizi in rete, ma anche tra i principali fornitori di dati.

In questo senso l'AGCOM è particolarmente interessata a comprendere il fenomeno dei big data e soprattutto se un eventuale uso anti-competitivo da parte dei soggetti in posizione dominante possa generare fallimenti di mercato ed in caso affermativo creare un insieme di regole finalizzate ad evitarlo[305].

Un aspetto fondamentale è legato al tema della trasparenza e in particolare, secondo quanto previsto nel Codice delle Comunicazioni elettroniche, all'obiettivo di contribuire allo sviluppo dei mercati *"promuovendo la diffusione di informazioni chiare, in particolare garantendo la trasparenza delle tariffe e delle condizioni di uso dei servizi di comunicazione elettronica accessibili al pubblico"* (d.lgs n. 259/2003, art. 5, let. d).

Poiché i big data non sono un mercato e i servizi che ne fanno largo uso vengono offerti spesso a un prezzo non monetario *(zero price)* anche in cambio dei dati personali, il tema della trasparenza, unito a quello della consapevolezza del consumatore, diventa un elemento chiave della possibile azione del regolatore di settore, così come più ampiamente illustrato nei paragrafi che seguono.

[305] Proprio per questo motivo l'AGCOM ha avviato il 30 maggio 2017, con la delibera 217/17/CONS, un'indagine conoscitiva riguardante proprio il tema dei Big Data, a dimostrazione dell'interesse dell'autorità di analizzare le eventuali criticità regolatorie che potrebbero essere generate dall'uso dei Big Data. Alcuni dei temi rilevanti che sono stati individuati, quali la tutela di diritti individuali e collettivi fondamentali, come la concorrenza, la privacy e le garanzie del pluralismo informativo (del. 217/17, punto 14) riguardano l'attività di autorità differenti, motivo per cui l'indagine conoscitiva vedrà coinvolta oltre l'AGCOM anche il Garante per la protezione dei dati personali e l'Autorità garante della concorrenza e del mercato (AGCM).

4.1.2 Il pluralismo informativo

Altro compito dell'AGCOM è quello di promuovere il pluralismo informativo. Tale questione è particolarmente delicata con riferimento al ruolo dei big data, perché se da un lato è vero che nel mercato digitale e dei social media l'offerta di informazioni tende ad essere sempre più profilata, basata su algoritmi che si adeguano e si conformano (spesso confermando) alla nostra visione del mondo, d'altro canto l'adeguamento di questo nuovo contesto alla regolazione del pluralismo appare estremamente complesso, richiedendo un cambiamento di paradigma che, a partire dal caso italiano, non trova applicazione nei principi giuridici che sono alla base della regolamentazione attuale.

Storicamente infatti il legislatore, in un'epoca caratterizzata dalla scarsità delle fonti d'informazioni e dalla centralità della televisione, al fine di assicurare il pluralismo delle voci, l'espressione della libera manifestazione del pensiero, e, in tal modo, il fondamentale diritto del cittadino all'informazione[306], ha ritenuto opportuno garantire e tutelare sia il pluralismo interno, che quello esterno.

Mentre il primo riguarda esclusivamente il servizio pubblico radiotelevisivo, operante inizialmente in regime di monopolio, chiamato ad assicurare la pluralità delle voci e dei punti di vista al proprio interno, e che nel tempo con l'apertura alla concorrenza, ha assunto una rilevanza sempre più limitata, sul secondo la regolamentazione in Italia trova come fonte primaria l'art. 43 del "Testo unico dei servizi di media audiovisivi e radiofonici" (TUSMAR)[307], che dispone: l'obbligo di notifica di eventuali intese o concentrazioni in capo a tutti gli operatori del Sistema Integrato delle Comunicazioni (SIC); la verifica, da parte di Agcom, della non costituzione di posizioni dominanti nel SIC o nei singoli mercati che lo compongono, previa individuazione del mercato rilevante in conformità ai principi di concorrenza e tenendo conto, fra l'altro, oltre che dei ricavi, del livello di concorrenza all'interno del sistema, delle barriere all'ingresso nello stesso, delle dimensioni di efficienza economica dell'impresa (comma 2); il divieto, in capo al medesimo fornitore di contenuti, di acquisire autorizzazioni per diffondere più del 20 per cento del totale dei programmi televisivi o più del 20 per cento dei programmi radiofonici irradiabili su frequenze terrestri in ambito nazionale (comma 7); il divieto, in capo ai soggetti tenuti all'iscrizione nel registro degli operatori di comunicazione, di conseguire ricavi superiori al 20 per cento dei ricavi complessivi del SIC (comma 9)[308].

[306] Corte Cost. Sent. n. 420 del 1994.

[307] D.lgs n. 177 del 31 luglio 2005.

[308] Tali ricavi sono identificati come quelli derivanti dal finanziamento del servizio pubblico radiotelevisivo al netto dei diritti dell'erario, da pubblicità nazionale e locale anche in forma diretta, da televendite, da sponsorizzazioni, da attività di diffusione del prodotto realizzata al punto vendita con esclusione di azioni sui prezzi, da convenzioni con soggetti pubblici a carattere continuativo e da provvidenze pubbliche, da offerte televisive a pagamento, dagli abbonamenti e dalla vendita di quotidiani e periodici inclusi i prodotti librari e fonografici commercializzati in allegato, nonché dalle agenzie di stampa a carattere nazionale, dall'editoria elettronica e annuaristica anche per il tramite di internet, "da pubblicità on line e sulle diverse piattaforme anche in forma diretta, incluse le risorse raccolte da motori di ricerca, da piattaforme sociali e di condivisione" [D.L. 63/2012] e dalla utilizzazione delle opere cinematografiche nelle diverse forme di fruizione del pubblico (comma 10).

Il TUSMAR, vietando la posizione dominante sulla base di quote di mercato, si concentra, quindi, su una tipica problematica antitrust, inclusa la definizione del mercato rilevante, vietando peraltro *la mera posizione dominante* e non l'abuso della stessa. I principi concorrenziali, inoltre, devono essere applicati in maniera strumentale alla tutela del pluralismo (questo significa anche la possibilità di più operatori in posizione dominante).

L'applicazione di strumenti tipici dell'antitrust nell'ambito della protezione e tutela del pluralismo appare peraltro poco convincente, dal momento che l'analisi della dominanza può semplicemente rifarsi ad una logica di ricavi, laddove il pluralismo nei media non può non tener conto di ulteriori elementi distintivi riguardanti tra l'altro la qualità dell'informazione che giunge ai consumatori[309], i tempi di esposizione, l'autorevolezza dell'operatore anche in chiave di affidabilità e reputazione[310].

Tuttavia, nonostante i numerosi tentativi effettuati anche a livello europeo e le più recenti iniziative della Commissione[311], il legislatore ritiene la pluralità delle fonti l'unico elemento in grado di assicurare il pluralismo nel mercato dell'informazione, attraverso la mera concorrenza dal lato dell'offerta. Il nodo pluralismo / concorrenza rimane irrisolto, l'applicazione concreta estremamente complicata, anche alla luce delle difficoltà di intervento in un ambito così controverso)[312].

Questi obiettivi elementi di debolezza finalmente messi in crisi con l'ingresso e il ruolo sempre più pervasivo dei social media. Da questo punto di vista l'utilizzo dei dati e di nuove modalità di accesso all'informazione possono rappresentare non il problema in sé, ma una possibile soluzione alla *relazione pericolosa* concorrenza / pluralismo.

Lato antitrust editori e social media competono per il tempo sempre più ampio speso dai consumatori online per accedere alle informazioni, per fidelizzare gli utenti acquisendone i dati e per essere remunerati dalla pubblicità.

---

[309] Cfr. Questo approccio è tendenzialmente condiviso all'interno dell'Unione Europea. A tale proposito si segnala Centre for Media Pluralism and Media Freedom, *European Union competencies in respect of media pluralism and media freedom*, 2013; nel quale il gruppo di ricerca dello European University Institute of Florence, in un lavoro realizzato per la Commissione Europea, attraverso una review sulle principali misurazioni del pluralismo e della libertà dei media, suddividendo tali misurazioni tra indicatori quantitativi e qualitativi.

[310] D'altro canto l'inadeguatezza dei soli strumenti dell'analisi antitrust per misurare il pluralismo è dimostrata dal fatto che la salvaguardia di questi diritti esula in alcuni ambiti dal campo d'applicazione del diritto della concorrenza (vedi servizi pubblici radiotelevisivi), essendo regolata da altri differenti istituti, nel rispetto dei Trattati europei (Trattato di Amsterdam in particolare) e della Costituzione Italiana. Per un approfondimento del tema, si veda Strozzi G. Mastroianni R., *Diritto dell'Unione Europea. Parte istituzionale. Dal trattato di Roma al trattato di Lisbona,* Giappichelli, 2011.

[311] European Commission, *Next steps against fake news: Commission sets up High-Level Expert Group and launches public consultation*, Brussels, 13 November 2017, https://ec.europa.eu/digital-single-market/en/news/public-consultation-fake-news-and-online-disinformation.

[312] AGCOM, *Indagine sulle posizioni dominanti*, 2016. https://www.agcom.it/documents/10179/5041493/Delibera+310-16-CONS/04871326-f0f9-4d03-b4fb-49907be63d77?version=1.0

A questo si aggiunge l'ulteriore dimensione legata alla domanda, che comporta da un lato costi elevati per gli editori per mantenere un alto livello di qualità del loro prodotto in grado di attrarre i lettori, mentre concorrenti come Facebook possono contare su una molto più ampia gamma di notizie, incluse quelle false (fake news), determinate da un meccanismo di conferma (confirmation bias) che spinge l'utilizzatore a ricercare le notizie che rafforzano le proprie opinioni e il proprio punto di vista, piuttosto che la veridicità delle stesse. In questo senso si parla di post-verità[313].

La dimensione antitrust sarebbe pertanto relativa ai potenziali comportamenti anti-competitivi dei social media e delle piattaforme online: a) nella prioritizzazione delle notizie per l'utente; b) nell'acquisizione nel tempo di una posizione dominante in grado di determinare le relazioni economiche (il prezzo) con i concorrenti/fornitori di notizie. Ad esempio attraverso il servizio *Instant Articles*, Facebook potrebbe assumere nel tempo il ruolo di gatekeeper nell'accesso online alle informazioni anche degli editori da parte degli utenti, in una strategia volta a trasferire al social media una posizione dominante nell'offerta delle notizie online e conseguentemente nella relazione con i singoli editori.

In concreto, il primo aspetto, dipendendo il processo da algoritmi, rientra nella più generale problematica trattata nel cap. 3 a cui si rinvia, mentre il secondo appare di difficile applicazione almeno fino a quando il prezzo non presenti connotazioni monopolistiche, molto diverse dalla situazione attuale dove la scelta del servizio è vantaggiosa per entrambi gli attori, poiché gli utenti che "cliccano" al servizio sono il 20% in più degli altri articoli, con una percentuale che sale a oltre il 30% in più nel mobile e a un revenue sharing particolarmente favorevole per gli editori.

La questione riguarda dunque soprattutto problematiche di tipo regolamentare, che vanno oltre le questioni di concorrenza.

In particolare il tema riguarda in primo luogo il buco nero del pluralismo informativo, per come è stato fin qui regolamentato, e cioè dal lato della domanda e in particolare al rapporto tra qualità dell'informazione, verità (o veridicità) e democrazia[314].

In particolare, a seguito delle ultime elezioni americane e dei risultati del referendum sulla Brexit, sono aumentate le critiche alle informazioni che circolano tramite internet. Nello specifico i gestori dei portali sono stati ripetutamente accusati di manipolare le informazioni mediante algoritmi che filtrano le notizie in base ai dati ricevuti dagli utenti[315]. A tale proposito si parla anche di camera

[313] Il primo utilizzo di questa espressione si fa risalire al drammaturgo Steve Tesich nel 1992, che lo utilizzò per descrivere la guerra Serbo/Croata sulla rivista Nation. Più di recente, nel 2016, l'Oxford Dictionary ha eletto Post verità come parola dell'anno. La definizione che ne dà è "relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief"

[314] Nicita A., *Pluralism in the digital age: public service media and technological changes*, Relazione al Convegno Eurovisioni, Ottobre 2017.

[315] Cfr. Drexl J., *Economic Efficiency versus Democracy: On the Potential Role of Competition Policy in Regulating Digital Markets in Times of Post-Truth Politics*, 2016. Damien G. and Ioannis L., "Competition Policy: Between

dell'eco[316] o bolle di filtraggio[317], intendendo con questi termini l'effetto del sistema di personalizzazione dei risultati di ricerche su siti, che rischiano di amplificare o rafforzare idee e credenze ripetute all'interno di un sistema predefinito (echo chamber) ed isolare l'utente da informazioni che sono in contrasto con il suo punto di vista, come se fosse in una bolla. Questo sistema viene accusato talvolta di incidere sull'opinione politica e di alterare così un corretto processo democratico[318].

In tal senso, "se la normativa attuale ha sin qui ritenuto che fosse sufficiente la concorrenza tra chi parla nel mercato delle idee a garantire il pluralismo e, attraverso di esso, l'affermazione di fatti veritieri, oggi, il destino della relazione tra pluralismo e verità in Rete è affidato alla domanda di informazione, al ruolo attivo di chi ascolta, alla disponibilità a ricercare la qualità nell'informazione e a mettersi in discussione prima di discutere. A "depolarizzare", come scrive Sunstein, sottraendo al dominio esclusivo dell'emozione la ricerca della verità"[319] .

In realtà, muovendo da un contesto, come notato in precedenza, di superamento dei limiti derivanti dalla scarsità e da un aumento drammatico dei canali di informazione, che sta modificando e trasformando la relazione tra produttore e utilizzatore, gli strumenti attuali consentono di superare molte della difficoltà derivante del sistema informativo precedente (e della sua inadeguata regolazione).

Al contempo, se l'effetto eco, e il ruolo dei big data in questo contesto, tende a moltiplicare a dismisura processi comportamentali pre-esistenti (confirmation bias) lato domanda[320], ciò avviene storicamente anche dal lato dell'offerta, con editori tradizionali che si posizionano sul mercato ricercando pubblici specifici, caratterizzati da opinioni ben definite e che accedono spesso a un'unica fonte d'informazione, a cui offrono contenuti in grado di attrarli e fidelizzarli.

Equity and Efficiency", Cambridge: Cambridge University Press, 2017; Max Planck Institute for Innovation & Competition Research Paper No. 16-16. Available at SSRN: https://ssrn.com/abstract=2881191

[316] Quattrociocchi W., Scala A., Sunstein C. R., *Echo Chambers on Facebook*, 2016, https://ssrn.com/abstract=2795110

[317] Il termine è stato coniato da Eli Pariser nel suo saggio *The Filter Bubble: What the Internet Is Hiding from You,* 2011.

[318] Dutton, William H. and Reisdorf, Bianca Christin and Dubois, Elizabeth and Blank, Grant, *Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States*, May 1, 2017. Quello Center Working Paper No. 5-1-17. Available at SSRN: https://ssrn.com/abstract=2960697

[319] Thaler R.H., Sunstein C. R., *Nudge: Improving decisions about health, wealth, and happiness*, TC Leonard - Constitutional Political Economy, 2008.

[320] Nicita A., *L'insostenibile leggerezza della post-verità*, prefazione al volume *Extra Media* a cura di Enzo Argante, Egea, 2017.

Inoltre l'idea che spesso emerge anche nei teorici dell'economia comportamentale, di una dimensione monolitica della domanda, non trova conferma in alcune decisioni giurisprudenziali e studi recenti.

Interessanti spunti provengono in tal senso dalla giurisprudenza nordamericana in tema di marchi. Nella fattispecie si sostiene come un consumatore online, avendo una diversa consapevolezza di uno offline, risulta meno vulnerabile rispetto a notizie ingannevoli o poco attendibili. In particolare, nel considerare la responsabilità secondaria di un ISP per la contraffazione di un marchio realizzata da inserzionisti tramite keyword advertising (ossia attraverso l'impiego come parole chiave, nell'ambito di un servizio di posizionamento su internet, di segni identici o simili a un marchio, senza che il titolare di quest'ultimo abbia fornito il proprio consenso), la Corte d'appello per il Nono Circuito in Network Automation v. Advanced Systems Concepts ha messo in risalto come, nell'ambito della valutazione del rischio di confusione per il consumatore medio, la diffusione delle nuove tecnologie richieda un approccio flessibile[321].

Nel contesto di internet, afferma la Corte, il diritto non può essere ancorato ad assunzioni e a formule rigide ed il rischio di confusione può essere significativamente ridotto dalla crescente sofisticazione degli utenti di internet («the default degree of consumer care is becoming more heightened as the novelty of the Internet evaporates and online commerce becomes commonplace [Internet users] fully expect to find some sites that aren't what they imagine based on a glance at the domain name or search engine summary»). Allo stesso modo, la medesima Corte, nel caso Toyota Motor Sales v. Tabari, ha affermato che, al fine di valutare il rischio di confusione, il consumatore da assumere come parametro di riferimento è «a reasonably prudent consumer accustomed to shopping online»[322]. In particolare, «reasonable, prudent and experienced Internet consumers are accustomed to [...] exploration by trial and error. They skip from site to site, ready to hit the back button whenever they're not satisfied with a site's contents. They fully expect to find some sites that aren't what they imagine based on a glance at the domain name or search engine summary».

Naturalmente il caso non si riferisce direttamente all'oggetto in questione, ma evidenzia comunque come cambiando il contesto di riferimento (scarsità / abbondanza) il tema del pluralismo informativo non possa essere limitato a una lettura tradizionale dell'offerta e della domanda, ma richieda strumenti sofisticati e per certi versi future proof per potere essere oggetto di una regolamentazione efficace e orientata agli obiettivi. In questo senso la possibilità di facilitare l'accesso all'informazione e di allargare la base dei potenziali utenti/cittadini consapevoli costituisce una finestra di opportunità, nella quale l'uso dei dati non costituisce per sé un pericolo per la democrazia, ma al contrario una condizione, insieme ad altre, per una più ampia diffusione della conoscenza e del pluralismo delle idee in una società complessa come quella digitale.

[321] 638 F.3d 1137 (9th Cir. 2011).

[322] 610 F.3d 1171 (9th Cir. 2010).

4.2 I big data tra monopoli naturali, esternalità positive e asimmetrie informative

Come la teoria economica ricorda, i comportamenti delle imprese non sono gli unici fattori che possono alterare il funzionamento del mercato. Può anche accadere che caratteristiche precipue dei processi produttivi e degli agenti economici coinvolti negli scambi impediscano il pieno dispiegarsi del gioco concorrenziale. E' in presenza di questi ultimi fallimenti di mercato che la regolazione trova la sua giustificazione economica, candidandosi a svolgere un ruolo complementare (e talvolta alternativo) alla disciplina antitrust.

Market failure per eccellenza è la presenza di un monopolio naturale, il quale ricorre ogniqualvolta la struttura dei costi rende razionalmente conveniente l'esistenza di un'unica impresa incaricata della produzione. Trattasi, ad ogni modo, di una ipotesi che non sembra riguardare i big data: sebbene le tecnologie dell'informazione possano comportare forme di concorrenza imperfetta, parrebbe che i costi delle attività di raccolta, generazione e analisi dei dati digitali possano essere minimizzati anche quando le imprese attive nel mercato sono più di una.

Diversamente, si sostiene che nei mercati per la generazione e raccolta dei big data, le imprese godano di una esternalità positiva radicata nell'ignoranza dei consumatori che determina un eccessivo trasferimento di dati dai consumatori verso le imprese, vista la incapacità dei primi di internalizzare parte dei benefici che l'uso dei dati comporta. I consumatori, infatti, non sarebbero consapevoli di molti dei meccanismi che governano l'economia digitale[323], quali il valore dei dati in generale, la possibilità che esistano differenti privacy policy, i processi di clusterizzazione su cui si fondano le elaborazioni e l'offerta delle imprese, nonché la circostanza per cui i beni che consumano a prezzo nullo sono scambiati a fronte dei loro dati e della loro attenzione[324]. Pertanto, si argomenta, non solo i consumatori non saprebbero calcolare il livello di dati che dovrebbero mettere a disposizione delle imprese per massimizzare la propria utilità, ma altresì che il mercato fallirebbe

[323] Howard A. Shelanski, *Information, Innovation, and Competition Policy for the Internet*, 161 U. Pa. L. Rev. 1663, 1691 (2013): «failure to read or understand privacy policies remains a possible source of market failure». Al riguardo, si veda Jan Whittington, Chris Jay Hoofnagle, *Social Networks and the Law: Unpacking Privacy's Price*, 90 *N.C. L. Rev*. 1327, 1357 (2012), dove gli A. osservano che «[a]merican consumers profoundly misunderstand the rules underlying these transactions; they do not understand the terms of trade», nonché Alessandro Acquisti, Jens Grossklags, *What Can Behavioral Economics Teach Us About Privacy?*, Alessandro Acquisti e altri (a cura di), *Digital Privacy: Theory, Technologies And Practices* 363 (2008), dove si elencano i fattori che spiegherebbero perché i consumatori avrebbero la tendenza a sotto-stimare il valore dei loro dati personali. Tra questi fattori si contano la razionalità limitata, il c.d. "*optimism bias*" e le asimmetrie informative. Altri studi hanno poi mostrato come la sensibilità ai temi della privacy varierebbe in relazione al reddito e al livello di scolarizzazione – cfr. Il-Horn Hann et al., *Online Information Privacy: Measuring the Cost-Benefit Trade-off*, 23 Int'l Conf. On Info. Systems 1 (2002) – e infine non sono stati pochi coloro che hanno evidenziato l'esistenza del c.d. paradosso della privacy, per cui la tutela dell'ordine pubblico e il diritto alla privacy non sarebbero obiettivi congiuntamente perseguibili. A questo ultimo proposito si rinvia a Patricia A. Norberg et al., *The Privacy Paradox: Personal Information Disclosure Intentions Versus Behaviors*, 41 J. *Consumer Aff*. 100, 101-02 (2007). In tempi più recenti, però, si contano coloro che ritengono come l'interesse per la privacy stia aumentando – cfr. Avi Goldfarb, Catherine Tucker, *Shifts in Privacy Concerns*, 102 *Am. Econ. Rev.* 349 (2012).

[324] Newman, *Zero-Price: Foundations*, cit., 179: «there is no analogue to a "price tag" for attention (or information) costs».

perché raggiungerebbe l'equilibrio in corrispondenza di un livello sub-ottimale (perché eccessivo) di dati.

Ulteriore fonte di fallimento per il mercato, correlata all'ipotesi precedente, è quella rappresentata dalle asimmetrie informative. In questo senso si potrebbe aggiungere che i big data limitino il funzionamento del mercato contribuendo ad aumentare le asimmetrie informative che separano imprese e consumatori. Grazie a i big data, infatti, le imprese sono nelle condizioni di accedere a molte più informazioni rispetto a quelle disponibili ai consumatori: l'elaborazione dei dati consente loro di comprendere preferenze e prassi di cui gli stessi individui possono non essere consci[325].

Secondo la più recente dottrina[326], questa limitata abilità di discernimento dei consumatori, inoltre, si amplificherebbe nei mercati dove le merci sono preferite principalmente in ragione della loro qualità, anche perché offerte a prezzi nulli[327]. In questi mercati la capacità dei consumatori di essere "giudici" delle imprese sarebbe decisamente limitata quando si verificano alcuni condizioni: quando, anche per effetto di comportamenti imprenditoriali volutamente complessi da decifrare[328], i prezzi non

[325] Questa è la tesi sostenuta, ad esempio, da Nathan Newman, *The Costs Of Lost Privacy: Consumer Harm And Rising Economic Inequality In The Age Of Google*, 40 *William Mitchell Law Review* 849, 856, 2014, il quale, onde mostrare come lo scambio tra i servizi di Google e i dati delle persone non seguirebbe le logiche di mercato, scrive «that the market is functioning in such a case requires (1) that those users properly value the benefits they receive from Google, (2) that they properly price their personal information and the opportunity cost of giving it up, and (3) that there are no economic byproducts of Google's monopoly control of user data that reduce consumer welfare more generally. […] [T]here is strong evidence that users do not properly calculate any of those three factors in the exchange with Google».

[326] Maurice E. Stucke, Ariel Ezrachi, *The Curious Case of Competition and Quality* 18-19, 2014, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2494656.

[327] Cfr. Alexandre Gaudeul, Bruno Jullien, *E-commerce, two-sided markets and info-mediation*, in Internet and Digital Economics (E. Brousseau and N. Curien eds.), 2007, 268. Vedi anche European Commission Staff Working Document, *Online Platforms,* di accompagnamento alla Comunicazione della Commissione Europea *Online Platforms and the Digital Single Market* 2, 2016, at 2; nonché Nathan Newman, *Feature: Search, Antitrust, and the Economics of the Control of User Data*, 31 Yale J. on Reg. 401, 451-452 (2014), dove si legge «Information is itself a commodity and particular companies can, in Stiglitz's words, "appropriate the returns to creating information" for economic advantage in the market in ways not seen in traditional commodities. This leaves a company like Google with the economic incentive, pushed by advertisers seeking their own information advantage in marketing to users, to increasingly violate user privacy to appropriate more and more information to solidify its economic dominance, all while making it less economically possible for potential competitors to challenge the company. This concentrated control of user data gives such companies more than just a competitive advantage vis-a-vis their competitors. Companies know the intimate details about the preferences of consumers and this skews markets against consumers who lack similar data about those companies or their options in negotiating terms with those companies».

[328] Eugenio J. Miravete, *The Doubtful Profitability of Foggy Pricing* 2-3 (2004), http://ssrn.com/abstract=618465; Xavier Gabaix, David Laibson, *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*, 121 *Q. J. Econ*. 505, 505-08, 2006; Oren Bar-Gill, Elizabeth Warren, *Making Credit Safer*, 157 *U. Pa. L. Rev*. 1, 27-28, 2008, nonché AdiAyal, *Harmful Freedom of Choice: Lessons from the Cellphone Market*, 74 *Law and Contemporary Problems* 91, 94, 2011.

riescono a segnalare la qualità dei prodotti/servizi scambiati in un mercato[329]; quando il favore per un bene dipende dalla sua qualità; quando tale qualità deve essere sperimentata prima di essere compresa (*experience good*); e quando, comunque la percepiscano, i consumatori non sono in grado di rapportare detta qualità ai prezzi[330].

Complessivamente, dunque, questa duplice forma di ignoranza/inconsapevolezza dei consumatori "fonti di dati" e dei consumatori "destinatari di informazioni e beni" impedirebbe la possibilità che la pressione competitiva operi da pungolo sia nei mercati della generazione e raccolta dei dati, sia nei mercati della commercializzazione dei beni e servizi realizzati dalle imprese che controllano quei dati[331].

Dinanzi alle circostanze ed ai fattori appena illustrati che consentirebbero alle imprese attive nei mercati per la generazione e raccolta dei big data di acquisire potere di mercato, il ruolo e gli strumenti della regolazione differiscono da quelli del diritto della concorrenza. Come noto, l'applicazione del diritto antitrust non risponde, almeno in linea di principio, alle logiche proprie della regolazione e, quindi, non si traduce in un intervento sui mercati per scongiurarne i fallimenti, ossia, nel caso in esame, per consentire a tutti gli agenti economici di disporre della medesima informazione ed essere sufficientemente istruiti e consapevoli. Al più il diritto antitrust potrebbe intervenire contro quei comportamenti che amplificano questo iato informativo attraverso forme di manipolazione dell'informazione estratta dai big data e diffusa al mercato.

A dire il vero, nel contesto dei big Data, anche sul versante regolatorio i margini di intervento (quantomeno quelli auspicabili) appaiono limitati. E' possibile certamente ipotizzare interventi tesi a

[329] Si è, infatti, soliti pensare che esista una correlazione positiva tra prezzi e qualità, per cui un prodotto di qualità elevata sarebbe venduto a un prezzo maggiore rispetto al prezzo di bassa qualità – cfr. OECD, *The Role and Measurement of Quality in Competition Analysis* 44, 28 Ottobre 2013, disponibile all'indirizzo http://www.oecd.org/competition/Quality-in-competition-analysis-2013.pdf e William M. Sage, Peter J. Hammer, *Competing on Quality of Care: The Need to Develop a Competition Policy for Health Care Markets*, 32 *U. Mich. J.L. Ref.* 1069, 1078-1088, 1999. Nondimeno, questa relazione di proporzionalità diretta non è sempre vera e non solo quando i prezzi praticati sono nulli, ma anche quando i prezzi praticati sono positivi. Si pensi, al riguardo, ai prodotti di largo consumo o ai farmaci che recano marchi a fronte dei medesimi beni che invece circolano con etichette commerciali (c.d. private labels) o sono commercializzati dai genericisti – cfr. OECD, *Role and Measurement*, 79 e Deven R. Desai, Spencer Weber Waller, *Brands, Competition and the Law*, 2010 *Brigham Young U. L. Rev.* 1425, 2010. Inoltre, Stucke, Ezrachi, *The Curious Case*, cit., 18-19, i quali notano che «[t]he link between price and quality is … more complex than one might expect. Quality and price at times are positively correlated, whereby higher price signals better quality».

[330] In altri termini, i consumatori non sarebbero sempre in grado di comprendere il rapporto qualità-prezzo e di decidere, ad esempio, se una variazione in diminuzione del prezzo sia controbilanciata dalla diminuzione della qualità– cfr. Stucke, Ezrachi, *The Curious Case*, cit., 18-19: «consumers may not always respondas the agencies expect them to – not because of unforeseen bias or heuristic but rather because of the information landscape in which consumers operate and *their ability to analyze and decode that information*» (enfasi aggiunta).

[331] Drasticamente sul punto Frank Pasquale, *Privacy, Antitrust, and Power*, 20 *Geo. Mason L. Rev.* 1009, 1015-16 (2013), il quale sostiene che i mercati basati sull'estrazione di dati sarebbero così «disfunctional» da non poter essere curati dal mero diritto antitrust, ma da richiedere una forma di vera regolazione.

mitigare le asimmetrie informative attraverso una puntuale definizione degli adempimenti richiesti per l'utilizzo dei dati personali degli utenti. In questo caso ci muoveremmo nel solco dell'approccio tradizionalmente seguito per le questioni relative alla privacy e, più in generale, al deficit informativo dei consumatori-utenti. Tale approccio, tuttavia, negli anni si è dimostrato non particolarmente efficace: gli interventi con i quali si è aggravato l'obbligo di informare per le imprese hanno finito per tradursi in adempimenti prevalentemente formali, inadeguati a fronteggiare il tema sostanziale di fondo dell'adeguata informazione del consumatore e delle scelte consapevoli compiute dagli stessi. Nell'ambito digitale, il quadro complessivo è ancor più desolante. Come è stato brillantemente rilevato, il vero problema della tutela dei dati personali risiede nel suo essere lettera morta[332]. La convinzione che il consenso possa rappresentare il principale elemento di legittimazione anche nel contesto digitale espone ogni intervento normativo all'esito tragico dello sforzo di Sisifo, specie se si considera che non si intravedono modelli di business alternativi rispetto a quelli che ricavano profitti dalla profilazione degli utenti e dalla pubblicità.

Di segno diverso sono le misure di recente implementate in Europa per favorire la mobilità degli utenti tra le piattaforme riducendo gli *switching costs* attraverso la garanzia della *data portability* (cfr. paragrafo 5.5). Una siffatta misura è tesa ad alimentare la competizione tra le piattaforme e potrebbe risultare funzionale a favorire una concorrenza sulla qualità dei servizi: in assenza di rilevanti costi di spostamento per via della portabilità dei dati, si deve, infatti, ipotizzare che, a parità di condizioni economiche (a parità di prezzi nulli) le piattaforme saranno incentivate a far leva sull'altro parametro di competizione (la qualità).

Poco auspicabili risulterebbero, invece, interventi sulle politiche di prezzo delle piattaforme. E' opportuno, infatti, ricordare che a caratterizzare i mercati in questione sono innanzitutto le interrelazioni e l'interdipendenza tra i versanti delle piattaforme, le quali devono portare «both sides on board», ossia intercettare un numero sufficiente di agenti economici su ciascuno dei versanti coinvolti in modo da assicurarsi una massa critica adeguata per alimentare gli effetti di rete indiretti. Nel caso, ad esempio, delle piattaforme che forniscono servizi di ricerca o di *social networking*, nel facilitare l'incontro tra utenti ed inserzionisti pubblicitari, seguendo il tradizionale modello degli *advertising-supported media,* mentre ai primi i servizi vengono offerti gratuitamente, ai secondi viene applicata una tariffa per gli spazi pubblicitari messi a disposizione per il *behavioural targeting.* Questa funzione obiettivo impone sovente alle piattaforme di praticare prezzi asimmetrici ai gruppi presenti sui diversi lati (*skewed pricing*), che non riflettono il costo effettivo del servizio offerto ad un particolare gruppo di utenti ma tengono conto delle differenze nell'elasticità della domanda. La non neutralità della strutturazione dei prezzi applicati dalla piattaforma è funzionale alla necessità di alimentare entrambi i versanti, sicché un eventuale intervento regolatorio teso a proibire i prezzi nulli rischierebbe di minare non solo il *business model* ma l'esistenza stessa di una piattaforma[333].

[332] B.J. Koops, *The trouble with European data protection law*, 4 International Data Privacy Law 250, 2014.

[333] Per un approfondimento del tema, vedi i paragrafi 3.4.3.1 e 3.4.3.2.

4.3 Dei consumatori più consapevoli del valore dei dati

Come già sottolineato, all'interno di una comune transazione economica, un soggetto acquista un bene o servizio in cambio di una determinata somma di denaro, così che è il prezzo che caratterizza la transazione a rendere il consumatore consapevole della convenienza dello scambio medesimo, ossia a consentirgli di valutare razionalmente se i benefici attesi dall'acquisto saranno superiori o meno al suo costo. Di conseguenza, l'eventualità che in alcuni mercati digitali i beni e servizi offerti siano disponibili a prezzo nullo blocca il funzionamento di questo meccanismo, sebbene quei prodotti siano scambiati a fronte di dati e, in particolare, di dati personali. Ecco perché, come si diceva, i consumatori corrono un duplice rischio: (a) di non sapere che l'uso dei beni e servizi offerti a prezzo nullo li coinvolge invero in una forma di scambio in cui, come si diceva nel precedente paragrafo 3.5.3, l'attenzione e i dati personali fungono da valuta[334]; e (b) di non riuscire ad apprezzare il valore dei dati personali barattati a fronte dei beni e servizi consumati.

Pertanto, si potrebbe discutere dell'opportunità di creare un sistema all'interno del quale i consumatori siano più consapevoli delle dinamiche che occorrono nel caso particolarissimo in cui interagiscono con quelle media multi-sided platforms sopra descritte nel paragrafo 3.4.3.1[335]. Infatti e come alcuni esperimenti dimostrano[336], dei consumatori più consci del valore che i dati hanno per loro e le imprese sarebbero il presupposto necessario a che fossero le regole del mercato a gestire il trade-off[337] tra disclosure e protezione dei dati personali[338].

---

[334] Cfr. Jan Whittington & Chris Jay Hoofnagle, *Unpacking Privacy's Price*, 90 N.C. L. REV. 1327, 2012, available at http://www.nclawreview.org/2012/10/unpacking-privacys-rule/

[335] Infatti, va comunque ricordato come i problemi qui esaminati non riguardino la generalità dei mercati dove le imprese raccolgono Big Data, ma soltanto i mercati dove dette imprese, sfruttando dei modelli di business multi-sided, propongono a un versante della piattaforma beni e servizi a prezzi nulli.

[336] Questi, se per un verso, hanno originato risultati piuttosto variabili che complessivamente hanno mostrato quanto i consumatori attribuiscano scarso valore alla loro privacy, per altro verso hanno mostrato come gli individui, una volta messi al corrente del rischio di perdere o diffondere proprie informazioni private, siano tendenzialmente disposti a negoziare la rinuncia alla loro privacy a fronte di benefici economici. A tale proposito Il-Horn Hann et. al. (2002) hanno cercato di comprendere il costo che gli individui associano al rilascio delle proprie informazioni on-line, valutando fra i $29.18 e i $36.47 la capacità di un sito web di accedere alle informazioni per usi secondari e fra i $39.83 to $49.78 il divieto di usare i dati per scopi secondari. Savage and Waldman (2013) hanno stimato che le persone sono disposte a pagare non più di pochi dollari per impedire alle loro applicazioni di raccogliere i propri dati dallo smartphone. Preibusch (2015) ha, invece, valutato che per una cifra inferiore a mezzo penny le persone renderebbero una singola ricerca online privata. Strahilevitz and Kugler (2016) si sono interrogati sul prezzo che le persone sarebbero disposte a pagare per evitare un controllo automatizzato dei contenuti delle loro e-mail, stimandolo a 15$ l'anno. L'azienda Kaspersky Lab ha, infine, recentemente condotto un'indagine per stimare il valore dei dati memorizzati per i possessori di smartphone, simulando una situazione in cui gli intervistati si trovavano fra scegliere se pagare o perdere i propri dati. In questo caso i partecipanti avrebbero speso mediamente 373,36 € per il recupero dei dati.

[337] Acquisti, A., *The economics of personal data and the economics of privacy*, Joint Working Party for Information Security and Privacy (WPISP) and Working Party on the Information Economy (WPIE) Roundtable, background paper 3. Paris: Organisation for Economic Co-operation and Development (OECD), 2010.

Del resto, una consapevole apertura dei dati personali a vantaggio di imprese e altri soggetti terzi porta dei benefici sia a chi ottiene quei dati, sia a chi li fornisce. I primi possono, infatti, diventare più efficienti nella gestione delle operazioni ed aumentare i ricavi riuscendo a tracciare le caratteristiche dei consumatori, conoscendo così le loro preferenze e la loro disponibilità a pagare. E i secondi, al contempo, riescono ad avere un servizio migliore ricevendo offerte personalizzate ed accedendo gratuitamente a determinati contenuti.

Nondimeno, l'apertura dei dati personali comporta dei costi. Si considerino, per esempio, quelli sostenuti dalle imprese quando i dati vengono violati, utilizzati in modo improprio o raccolti in modi che i consumatori ritengono troppo intrusivi. Per i consumatori i costi possono essere, invece, sia di natura economica, come quelli dovuti alla maggiore capacità delle aziende di aumentare il proprio surplus a discapito della domanda, che di natura emotiva e psicologica, in caso per esempio di furto d'identità[339]. Ciò detto, in dottrina si contano anche le posizioni di coloro che si oppongono a una maggiore disclosure a favore dei consumatori: vari studi si sono interrogati sul tema, con conclusioni tendenzialmente unanimi sul fatto che aumentare la consapevolezza dei consumatori sul valore dei dati non sia una soluzione sempre positiva[340]. L'educazione del consumatore viene per lo più vista come un'attività che non è di per sé sufficiente a creare un bilanciamento ideale fra l'esigenza di proteggere i dati e la condivisione di informazioni, venendo per lo più talvolta anche sconsigliata[341]. Interventi di sensibilizzazione possono infatti aiutare a superare alcuni problemi, ma certo non azzerano i rischi connessi a possibili violazioni della privacy e della sicurezza. Inoltre, in alcuni casi si è notato come troppe informazioni per i consumatori possano portare a effetti negativi, risultando pesanti e opprimenti[342].

Infine aumentare la consapevolezza del consumatore sul modo in cui vengono raccolti e analizzati i dati potrebbe aiutare le aziende a creare un rapporto di fiducia con i clienti, mitigando così ex-ante eventuali problemi di privacy[343]. L'educazione del consumatore potrebbe, dunque essere concepita come un modo per fidelizzare il cliente, sebbene anche in questo caso la letteratura non sia del tutto

---

[338] Cfr. Laudon, K. C., *Markets and Privacy*, Communications of the ACM (39:9), pp. 92-104, 1996 e Varian, H. (1997), *Economic Aspects of Personal Privacy*, in U.S. Department of Commerce, Privacy and Self-Regulation in the Information Age, 1997.

[339] Cfr. Stone, E. e Stone, D., *"Privacy in organizations: Theoretical issues research findings, and protection mechanisms"*, Research in personnel and human resources management 8 (3), 349-411, 1990.

[340] Acquisti et al. (2017) hanno analizzato la letteratura sul ruolo degli interventi educativi in materia di privacy e sicurezza nella letteratura informatica, notando come per la maggior parte degli autori viene riconosciuta a tali interventi un'importanza limitata.

[341] Cfr. Acquisti, cit., (2010).

[342] Cfr. Ben-Shahar, O. e Schneider, C. E., *The failure of mandated discourse,* U. Pa. Law Rev. 159, 3 (2010, 647), 2010.

[343] Culnan M. J. and Armstrong P.K., *Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation*, School of Business, Georgetown University, Washington, D.C. 20057-1008, 1999.

favorevole. Educare i consumatori non limita, infatti, il rischio di un eventuale uso erroneo dei dati ex-post ed attiva, inoltre, una serie di sfide, che le aziende non sempre riescono ad affrontare con successo[344].

Così in caso di avviamento di un programma educativo sui big data, bisognerebbe infatti stabilire, fra le molte cose, chi sarà il responsabile, in quale modo finanziare le attività e trovare un sistema che induca i consumatori a perdere del tempo per essere istruiti su tematiche, che generalmente non destano particolare interesse. Sono dunque molti i fattori che potrebbero rendere un'iniziativa formativa di questo tipo fallimentare.

Complessivamente, però, sebbene riequilibrare la conoscenza che individui e imprese hanno circa il valore dei dati possa implicare costi e rischi, cercare per lo meno di ridurre le asimmetrie informative dovrebbe essere in ogni caso visto come uno dei principali obiettivi della regolazione. È per questo che sono state recentemente avviate alcune iniziative educative finalizzate a far conoscere maggiormente ai più giovani i mercati dei dati digitali e a diffondere la consapevolezza dell'importanza di un utilizzo responsabile della rete internet, per evitare che informazioni private possano essere acquisite e diffuse illegalmente[345]. Programmi educativi possono, quindi, talvolta aiutare anche se devono essere usati con cautela e valutando sempre costi e benefici associati.

4.3.1 La disclosure regulation

Come spiegato in precedenza, pretendere di annullare le asimmetrie informative fra consumatori e imprese sarebbe un obiettivo, oltre che molto costoso, quanto mai complesso e di difficile realizzazione. Ciò non toglie che i consumatori avrebbero bisogno di maggiori tutele per evitare che le imprese possano approfittarsi della loro situazione di vantaggio conoscitivo.

Uno dei modi con i quali si cerca di riequilibrare lo squilibrio informativo è tramite la cosiddetta *disclosure regulation*, ovvero ponendo una serie di obblighi in capo alle imprese, ma anche alle pubbliche amministrazioni, per rendere disponibili a consumatori e cittadini una serie di informazioni[346]. Le autorità di regolazione si sono più volte espresse a tale proposito, per esempio l'Agcom è spesso intervenuta per rendere le caratteristiche dell'offerta commerciale dei prezzi degli

[344] Cfr. Whittington e Hoofnagle, cit. (2012).

[345] Fra le iniziative si segnala il progetto "Digitali e Responsabili", attivato da Google nel luglio 2017 con il patrocinio dell'Autorità per le Garanzie nelle Comunicazioni e del Garante per la protezione dei dati personali e con la partecipazione di Polizia postale e Altroconsumo. Nel 2015 Google e Altroconsumo si erano, inoltre, già resi partecipi dell'iniziativa "Vivi internet al sicuro" finalizzata a sensibilizzare gli utenti a un uso corretto degli strumenti di tutela della privacy e della sicurezza. Il Miur ha invece attivato nel 2016 un gruppo di lavoro sui Big Data, in cui cerca di esortare a una maggiore offerta formativa sul tema ritendendo importanti azioni per "educare al valore del dato nello sviluppo delle competenze digitali fin dalla scuola primaria" (link: http://www.istruzione.it/allegati/2016/bigdata.pdf).

[346] Per un inquadramento generale sui problemi legati alla disclosure regulation si veda: Di Porto F., *La regolazione degli obblighi informativi le sfide delle scienze cognitive e dei big data*, Editoriale Scientifica, ISBN 978-88-9391-095-8, 2017 ; Ogus A, "Regulation. Legal Theory and Economic Forms", Oxford University Press, Oxford, 1994, p. 121 ss.

operatori di telefonia più trasparenti[347]. La *disclosure regulation* non riguarda, tuttavia, solo l'azione delle autorità di settore, ma include anche tutte quelle situazioni in cui le pubbliche amministrazioni, riconoscendo il valore sociale della conoscenza e, dunque, l'importanza di rendere accessibili un maggior numero di informazioni, impongono delle leggi per stabilire determinati obblighi informativi. Un esempio tipico di *disclosure regulation* è quello dell'etichettatura dei prodotti alimentari, sulla quale è intervenuta l'Unione Europea con il Regolamento n. 1169/2011, ponendo una serie di norme per fare trascrivere le varie informazioni in modo semplice e chiaro per i consumatori[348]. Norme sugli obblighi informativi sono, inoltre, poste in diversi ambiti quali quello sanitario, per fare in modo che un trattamento medico venga preceduto da un consenso del paziente[349], o nei mercati finanziari, finalizzate a trasmettere informazioni chiare, corrette e non fuorvianti ai clienti.

L'utilizzo dei big data, andando ad incrementare le asimmetrie informative, aumenta per certi aspetti la necessità di regolare gli obblighi informativi. Alcuni studiosi sono, tuttavia, convinti che un intervento del regolatore in materia sia spesso inopportuno e che sia, quindi, preferibile lasciare gli operatori di mercato liberi di scegliere quali informazioni rendere o meno trasparenti[350]. Sebbene la *disclosure regulation* può, infatti, essere utile per dare la possibilità ai consumatori di diventare più consapevoli circa le caratteristiche di un bene o servizio acquisito, spesso gli effetti prodotti non sono quelli sperati. Considerando che imporre dei doveri alle imprese implica nella maggior parte dei casi un aumento dei costi, è dunque opportuno valutare attentamente le eventuali problematiche associabili alla *disclosure regulation* per capire se questo sia un mezzo idoneo per alleviare i fallimenti di mercato generati dalle asimmetrie informative.

Vi sono innanzitutto delle difficoltà legate al fatto che le informazioni fornite obbligatoriamente, sebbene riportate in modo chiaro e comprensibile, non attraggono spesso l'attenzione dei consumatori che, generalmente, non leggono nemmeno quanto scritto. Il tema è noto in letteratura come *"no-reading problem"*[351] ed è stato più volte oggetto di esperimenti che hanno dimostrato come le persone non leggano il più delle volte le informazioni riportate, sebbene queste possano

[347] In attuazione dell'articolo 1, comma 2, della legge n. 40/2007, il quale prevede che "L'offerta commerciale dei prezzi dei differenti operatori della telefonia deve evidenziare tutte le voci che compongono l'offerta, al fine di consentire ai singoli consumatori un adeguato confronto", l'Autorità ha adottato una serie di seguenti delibere. Per maggiori informazioni è possibile consultare il link: https://www.agcom.it/trasparenza

[348] Girolami M., *Etichettatura, informazioni e rimedi privatistici nella vendita di prodotti alimentari ai consumatori*, in Nuove Leggi Civ. Comm., 2014.

[349] Salerno F., *Consenso informato in medicina e qualità soggettive del paziente,* in Giur. it., p. 277-281, 2014.

[350] Fra di essi si segnalano soprattutto Omri Ben-Shahar e Carl E. Schneider, che hanno approfondito il tema in cari testi, nei quali rimarcano il non funzionamento degli obblighi informativi e il fatto che questi ultimi possano produrre più danni che benefici. Cfr. Ben-Shahar O. e Schneider C.E., *More Than You Wanted to Know: The Failure of Mandated Disclosure*, Princeton University Press, 2014.

[351] cfr. Ayres I. e Schwartz A., *The no-reading problem in consumer contract law*, in 66 Stanford l. rev., 2014, p. 545-609; Ben-Shahar O., *The myth of the 'opportunity to read',* in Contract Law, in Europ. Rev. of contract l., 2009, p. 1-28.

essere anche importanti[352]. La scarsa attenzione dei consumatori diminuisce i benefici connessi alla regolazione degli obblighi informativi, rendendo così l'aumento dei costi per le imprese meno giustificabile.

Le informazioni rilevate poi, non sempre riescono ad essere comprese dai consumatori. Si cerca a volte di ovviare al problema semplificando e standardizzando le informazioni, ma questo potrebbe non essere sufficiente, specialmente per tematiche, quali quelle legate al mondo dei big data, che tendono a essere più complicate per un osservatore comune rispetto a quelle riguardanti prodotti e servizi per lui più tangibili[353]. Termini tecnologici o di uso non comune possono, infatti, spesso confondere il lettore[354]; inoltre, alcune delle informazioni riportate dovrebbero essere continuamente adattate rispetto ai continui progressi tecnologici per non risultare inattuali.

Anche qualora il lettore avesse una cultura adeguata a comprendere le varie informazioni divulgate, il tempo necessario per la loro lettura sarebbe in alcuni casi comunque troppo lungo[355]. Si è accennato per di più in precedenza come un eccesso informativo possa generare un effetto addirittura controproducente per i consumatori[356]. Fornire troppe informazioni non dà infatti all'individuo la possibilità di assimilarle e valutarle, facendo sorgere così il rischio che suggerimenti o input esterni possano condizionare oltremodo la sua scelta.

Secondo altri osservatori una regolazione degli obblighi informativi non sarebbe necessaria, in quanto una scarsa trasparenza potrebbe produrre alle imprese un danno reputazionale tale da indurle a fornire in ogni caso le informazioni più rilevanti e di maggiore interesse per i

[352] Bakos, Yannis; Marotta-Wurgler, Florencia; and Trossen, David R., *Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts,* 2014. New York University Law and Economics Working Papers. Paper 195. http://lsr.nellco.org/nyu_lewp/195. Lo studio ha stimato tra 0,05% e 0,22% la frazione di utenti che accede ai contratti per le vendite online, la maggior parte dei quali spende in media un tempo troppo basso per poter dire che abbiano realmente letto i termini dell'accordo. La ricerca si basava sull'analisi del comportamento di navigazione in internet di utenti che acquistano software online da 90 società differenti. La metodologia utilizzata ha consentito di valutare sia se gli utenti avevano o meno fatto accesso agli End User License Agreements (EULAs), sia il tempo di durata dell'eventuale accesso.

[353] Cfr. Guido Noto La Diega; Ian Walden*, Contracting for the 'Internet of Things': Looking into the Nest,* 2016.

[354] Reidenberg et al. (2016) sostengono che le norme sulla privacy in particolar modo sono spesso scritte in un modo che si presta a confondere il lettore. Cfr. Reidenberg, Joel R., Jaspreet Bhatia, Travis D. Breaux, and Thomas B. Norton, *Ambiguity in Privacy Policies and the Impact of Regulation*, Journal of Legal Studies 45:S163–S190, 2016.

[355] McDonald and Cranor (2008) hanno cercato di dimostrare questo fatto empiricamente, misurando il tempo impiegato da una persona media per leggere le informative incontrate sulla privacy. Secondo le loro stime ognuno dovrebbe impiegare 76 giorni per leggerle tutte ed il tempo perduto costerebbe all'economia circa 781 miliardi di dollari. Cfr. McDonald, Alecia M., and Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, I/S: A Journal of Law and Privacy for the Information Society 4:540–65, 2008.

[356] Cfr. Ben-Shahar, O. e Schneider, C. E., *The failure of mandated discourse,* U. Pa. Law Rev. 159, (3), 2010.

consumatori[357]. Un esempio a tale riguardo è quello di Google, che (come già evidenziato nel paragrafo 2.2.2.3) ha ritenuto opportuno informare ogni utente circa i dati personali utilizzati ai fine della profilazione, mediante una apposita pagina dedicata alle politiche di privacy[358]. Strumenti di questo tipo possono migliorare l'affidabilità del servizio, generando maggiore sicurezza negli utenti che saranno, quindi, più portati poi a utilizzarlo. Non sempre, tuttavia, la concorrenza riesce a generare un livello di trasparenza adeguato, per cui il fatto che la reputazione aziendale possa dipendere anche dalle informazioni trasmesse spontaneamente dall'azienda non esclude per forza un intervento del regolatore.

I limiti sopra evidenziati della *disclosure regulation* inducono a pensare che un intervento del regolatore, qualora ritenuto opportuno, dovrebbe ad ogni modo tenere conto del principio di proporzionalità, per evitare l'insorgere di oneri oltremodo gravosi per le imprese. Non è però sempre facile comprendere i costi connessi agli obblighi informativi, motivo per cui l'applicazione della disclosure regulation dovrebbe essere valutata caso per caso, considerando che talvolta le imprese potrebbero adottare comportamenti trasparenti autonomamente, senza la necessità di porre una regolazione rigida a riguardo. L'efficacia degli obblighi informativi è, inoltre, collegata a una maggiore comprensione da parte dei consumatori dell'importanza delle informazioni rilevate, per la quale saranno probabilmente necessari degli interventi più incisivi di sensibilizzazione.

4.4 Una regolazione complementare al diritto antitrust

Il diritto antitrust non può intervenire sulle caratteristiche strutturali dei mercati, come i c.d. switching cost; non può agire compiutamente sui prezzi personalizzati, né può applicare facilmente la dottrina dell'essential facility. Mentre il tema dei prezzi personalizzati potrebbe essere gestito solo attraverso una qualche forma di regolazione dei prezzi che in non pochi hanno già ricordato come una soluzione da economia pianificata[359], le altre due questioni hanno o stanno trovando dei rimedi più puntuali e meno invasivi.

4.4.1 Switching cost e data portability

A volte la concorrenza nei mercati viene limitata dalle difficoltà degli utenti di cambiare operatore o di utilizzare un servizio differente. Può esserci in alcuni casi addirittura una strategia aziendale tesa a impedire che gli utenti possano trasferire i propri dati ad un concorrente per non perdere la clientela. In questi casi le imprese, alzando i cosiddetti switching cost, ostacolano la concorrenza ed impediscono l'ingresso di nuovi operatori. Tali costi sono tanto più elevati quanto più è difficile per un utente passare da un fornitore di servizio a un altro. Switching cost molto alti possono provocare un effetto di lock-in, consistente nel fatto che gli utenti, "bloccati" da un determinato operatore,

[357] Schwartz A. and Wilde L., *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, University of Pennsylvania Law Review 127:630, 1979.

[358] https://privacy.google.com/your-data.html

[359] Carlton D. W., *Roundtable on price discrimination*, Directorate for financial and enterprise affairs Competition Committee, 2016, https://one.oecd.org/document/DAF/COMP/WD(2016)82/en/pdf

sono costretti ad utilizzare i suoi servizi indipendentemente dalla potenziale maggiore qualità offerta da un gestore diverso. Il grado di lock-in è, quindi, direttamente proporzionale allo spessore degli switching cost. Soprattutto per i servizi che si affidano ai dati degli utenti in maniera rilevante, può risultare piuttosto agevole per le aziende limitare la capacità di cambiare fornitore, impedendo il trasferimento dei dati alla concorrenza[360], motivo per cui la Commissione Europea è recentemente intervenuta sulla questione introducendo il diritto alla portabilità dei dati personali (data portability)[361] e trattando il tema anche nella bozza di regolamento riguardante il flusso di dati non personali[362].

Migliorare la data portability consente ai consumatori di avere un maggior controllo sui propri dati personali, dandogli una maggiore capacità di usufruirne in caso di bisogno, per esempio quando questi sono indispensabili per cambiare il fornitore di un servizio. Ciò ridurrebbe l'effetto lock-in e darebbe maggiore linfa al mercato, favorendo l'ingresso di altri concorrenti e la creazione di nuovi e più innovativi modelli di business. È questa la visione della Commissione Europea, che nella valutazione di impatto del GDPR[363] sottolinea come la portabilità sia *"a key factor for effective competition"*[364] ed evidenzia l'importanza di trasmettere dati fra i vari fornitori di servizi: "*the possibility to move data from one service provider to another would increase competition in some sectors, e.g. between social networks, and could also make data protection an element in this competition, when users decide to move away from a service they do not consider appropriate in terms of data protection"*[365].

Vista la necessità di garantire la portabilità dei dati, si può discutere se sia più opportuno un intervento ex-ante del regolatore, oppure se la questione debba rimanere di competenza

[360] Cfr. James F. Ponsoldt J.F. and David C. D., *A Comparison Between U.S. and E.U.Antitrust Treatment of Tying Claims Against Microsoft: When Should the Bundling of Computer Software Be Permitted?*, 27 NW. J. INT'L L. & BUS. 421, 448, 2007.

[361] La data portability è andata a costituire un diritto specifico per via dell'art. 20 del GDPR (Reg. UE 2016/679), inoltre è stata recentemente oggetto di linee guida europee da parte dell' "Article 29 Data Protection Working Party" (WP 29). E' attualmente in corso un dibattito fra la CE e il WP29, proprio in merito agli effetti derivanti dall'applicazione del diritto alla portabilità dei dati, come verrà meglio spiegato nel capitolo sulla privacy al paragrafo 4.4.1.

[362] European Commission, *Proposal for a regulation on a framework for the free flow of non-personal data in the European Union*, 13 settembre 2017, file:///C:/Users/monica/Downloads/ProposalforaRegulation%20(1).pdf.

[363] Commission Staff Working Paper — Impact Assessment accompanying the General Data Protection Regulation and the Directive on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data ("Impact Assessment report"), SEC(2012) 72 final.

[364] Impact Assessment report, p. 43.

[365] Impact Assessment report, p. 28.

dell'antitrust. Quest'ultima può ovviamente intervenire, ma solo qualora ravvisasse un abuso di posizione dominante, secondo quanto previsto dall'art. 102 TFUE[366]. L'antitrust dovrebbe, quindi, prima di porre rimedio ad eventuali abusi, ravvisare la presenza di un'impresa con significativo potere di mercato. Ciò limita fortemente il suo intervento, dal momento che problemi di data portability sono spesso presenti anche in mercati non caratterizzati dalla presenza di un'impresa dominante. Per questo motivo è più indicato l'intervento del regolatore di settore, che nel caso dell'AGCOM si è attivata per garantire la portabilità del numero telefonico in caso di richiesta di cambio operatore, evitando così comportamenti ostruzionistici da parte del fornitore del servizio precedente.

La questione della data portability ha avuto poi progressivamente sempre maggiore rilievo con la diffusione dei big data e dei servizi digitali. Fra i settori più coinvolti sul tema vi è sicuramente quello bancario. Per ridurre le barriere all'accesso per i TTP e consentire alle nuove società tecnologiche finanziarie (FinTechs) di competere risulta infatti necessario trasmettere una serie di dati degli utenti, posseduti dalle banche tradizionali, in modo agevole, sicuro ed economico. Per queste finalità si sta promuovendo l'utilizzo di API, elemento fondamentale per applicare correttamente la nuova direttiva PSD2[367], come già osservato in precedenza. Un altro esempio che dimostra l'importanza della portabilità dei dati per accrescere l'offerta di servizi ed aumentare la concorrenza è quello di Uber, società che rapidamente è passata dall'essere una piccola startup a una multinazionale di rilievo. Tutto ciò si è reso possibile grazie all'utilizzo dei dati, trasferiti sempre tramite API, di aziende partner quali Google Maps per individuare i clienti e i tracciati dei driver, Cloud Messaging di Google per la messaggistica immediata e PayPal's Braintree per il pagamento[368].

La data portability può riguardare due aspetti differenti. Il primo concerne la possibilità dei consumatori di ottenere una copia dei propri dati, posseduti dai fornitori di servizi, in modo rapido ed in un formato tale da consentire loro di riutilizzarli senza difficoltà. Questo permetterebbe ad esempio a chi utilizza un social network di ottenere i dati là inseriti, per ritrasferirli in un social network differente. Un servizio che già consente agli utenti di scaricare i propri dati è Google Takeout. Con questo strumento, Google offre ai suoi utenti la possibilità di scaricare i dati creati nel suo social network Google+ in diversi formati, in modo da fornire la possibilità di importare facilmente tali dati in altri servizi internet[369].

[366] Graef, Inge and Verschakelen, Jeroen and Valcke, Peggy, (2013*), Putting the Right to Data Portability into a Competition Law Perspective*, Law: The Journal of the Higher School of Economics, Annual Review, pp. 53-63. Available at SSRN: https://ssrn.com/abstract=2416537.

[367] Direttiva 2015/2366/UE.

[368] PWC, (2016), *Catalyst or threat? The strategic implications of PSD2 for Europe's banks*, https://www.strategyand.pwc.com/media/file/Catalyst-or-threat.pdf .

[369] Per una descrizione del servizio Google Takeout, è consultabile il link https://support.google.com/takeout/answer/2508459?hl=en&ref_topic=2508503.

Il secondo aspetto della data portability è rappresentato, invece, dalla capacità delle aziende di ottenere, in caso di richiesta dei clienti, i dati direttamente dal precedente fornitore di un servizio. Tale processo dovrebbe avvenire senza soluzione di continuità (*seamless*) dal primo servizio al secondo, ovvero evitando che le difficoltà nel traferire i dati possano interrompere la fornitura del servizio ed indurre i consumatori ad evitare il cambio di operatore. Tuttavia, affinché ciò possa essere possibile, bisogna prima rendere un trasferimento diretto di dati tecnicamente fattibile e disponibile per le varie imprese. Tale processo andrebbe, però, ad aumentare i costi per le imprese, motivo per cui è giusto comprendere in quali circostanze e modalità sia preferibile migliorare la data portability.

Un intervento del regolatore potrebbe, per esempio, obbligare tutte le imprese di un settore a raccogliere dati in dei formati standard in grado di facilitarne la condivisione e il trasferimento. Soprattutto per le aziende di piccole dimensioni questa operazione potrebbe tuttavia essere oltremodo costosa, andando a costituire un onere sproporzionato. Le start-up che non dispongono delle risorse necessarie per rispettare gli eventuali obblighi posti potrebbero, inoltre, essere scoraggiate da una regolazione troppo stringente ed evitare di entrare nel mercato. Bisogna considerare poi come, in alcuni casi, la capacità di raccogliere dati tramite i quali avere delle informazioni esclusive possa essere un fattore determinante per incentivare le imprese ad innovarsi e ad essere più efficienti[370]. A tale proposito la Commissione si è espressa, oltre che nel regolamento sui dati personali di cui si dirà nel capitolo 5, anche nel draft sui dati non personali, proponendo un approccio basato sull'autoregolamentazione, ovvero incoraggiando i fornitori di servizi cloud a sviluppare codici di condotta riguardanti le condizioni in base alle quali gli utenti possano trasferire i dati ad altri operatori e tornare nei loro ambienti IT[371].

In linea di massima, quindi, per quanto la data portability sia un fattore ritenuto essenziale per promuovere la concorrenza, bisogna considerare una serie di trade-off per i quali non risulta sempre conveniente estendere la regolazione sul tema a tutti i settori. Una buona prassi potrebbe essere quella di osservare prima di un eventuale intervento il grado di lock-in di un determinato mercato, per capire se l'eventuale aumento dei costi e degli oneri per le imprese possa essere in quel caso giustificato da una riduzione degli switching cost. Solo qualora l'effetto netto risultasse positivo, una regolazione della data portability potrebbe essere considerata necessaria per aumentare il benessere collettivo e migliorare la concorrenza. Bisogna in ogni caso segnalare come nei settori che utilizzano un ampio quantitativo di dati, politiche volte a promuovere l'interoperabilità e la data portability si ritengano spesso preferibili, in quanto potrebbero avere un impatto particolarmente significativo sulla concorrenza e sulla dinamicità del mercato[372].

---

[370] Cfr. Swire, Peter and Lagos, Yianni, *Why the Right to Data Portability Likely Reduces Consumer Welfare: Antitrust and Privacy Critique,* May 31, 2013. 72 Maryland Law Review 335; Ohio State Public Law Working Paper 204.

[371]Per maggiori informazioni consulatre il link: https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data.

[372] Walden, Ian and Luciano, Laise Da Correggio, *Ensuring Competition in the Clouds: The Role of Competition Law?*, April 7, 2011, osservano che l'estensione del diritto di data portability può per avere un impatto più

4.4.2 L'apertura imposta dei dati

Nei paragrafi precedenti si è detto di come sia complesso e difficile applicare la dottrina dell'essential facility. Laddove, dunque, si volesse ovviare a questo limite si potrebbe configurare un'altra forma di intervento dei pubblici poteri nell'economia, ossia si potrebbero prevedere degli obblighi di apertura dei dati.

Certo, quest'ultima sarebbe una soluzione drastica, che andrebbe molto probabilmente a creare problemi ancora più gravi. Sorgerebbero innanzitutto dei dubbi circa il rispetto della privacy, motivo per cui rendere alcuni dati personali automaticamente pubblici risulterebbe quasi di sicuro vietato. Da un punto di vista economico le imprese perderebbero, inoltre, da un giorno all'altro il vantaggio competitivo acquisito mediante anni di investimenti e di accumulo di competenze specifiche, il che vorrebbe dire che non vi sarebbero più in futuro incentivi per i privati a raccogliere nuovi e migliori dati per rendere più efficienti le proprie attività.

In altri termini, il tema dell'obbligo generalizzato a condividere i dati è stato più volte affrontato dagli studiosi che hanno rimarcato come bisognerebbe sempre tenere presente il trade-off fra i benefici di una maggiore apertura dei dati, derivanti da una maggiore facilità nello svolgere analisi o acquisire notizie, e i danni che si rischiano di provocare rinunciando alla loro segretezza. I ricercatori aggiungono che valutare adeguatamente il trade-off diviene spesso particolarmente complicato, in quanto alcuni dati possono produrre danni o benefici se combinati o meno con altre informazioni[373]. Si ritiene per questo opportuna un'attenta valutazione caso per caso dei tipi di dati e delle condizioni alle quali risulti più o meno conveniente intervenire mediante una regolazione che solleciti la condivisione dei dati[374].

I dubbi sopra menzionati sulla privacy fanno sì che non sia generalmente consona una totale apertura al pubblico, sebbene talvolta potrebbero essere applicate delle particolari condizioni per rendere le procedure di information disclosure meno invadenti per i consumatori[375]. Si potrebbero, per esempio, utilizzare degli pseudonimi o associare le persone non a un nome, ma a un codice per evitare che il soggetto interessato possa essere facilmente riconosciuto. Tale prassi deve però essere effettuata con accortezza, in quanto non sempre in questo modo si riesce a garantire l'anonimato. Ne è una dimostrazione il caso della società AOL, la quale aveva rilasciato informazioni sulle persone

---

significativo sulla concorrenza nel settore del cloud computing rispetto a interventi riguardanti misure di concorrenza più tradizionali. Il documento è disponibile al link https://ssrn.com/abstract=1840547.

[373] Narayanan A., Huey J., Felten E.W., *A Precautionary Approach to Big Data Privacy*, in: Gutwirth S., Leenes R., De Hert P. (eds) Data Protection on the Move. Law, Governance and Technology Series, vol 24. Springer, Dordrecht., 2016, link: http://randomwalker.info/publications/precautionary.pdf.)

[374] Jansenn K., Hugelier S., *Open Data: A New Battle in an Old War Between Access and Privacy?*, in Digital Enlightenment Yearbook 2013 190, 199, 2013.

[375] Borgesius F. Z., van Eechoud M., Gray J., *"Open data, privacy, and fair information principles: towards a balancing framework"*, 2015.

che avevano effettuato ricerche mediante il loro sito, identificando ognuna con un numero diverso. Pochi giorni dopo, tuttavia, dei giornalisti hanno scoperto il nome di alcune persone associate ai dati forniti, evidenziando così la carente privacy del sistema adottato[376].

In alternativa alcune informazioni potrebbero essere fornite obbligatoriamente, ma solo se riportate in forma aggregata per renderle completamente impersonali. In questo caso si risolverebbero i problemi di privacy, sebbene i dati anonimi tendano ad essere meno utili per il riuso e, quindi, ad avere un valore molto inferiore[377]

Per quanto riguarda invece i dati ai quali sono associate informazioni personali specifiche, quali nome, indirizzo o numero di telefono, si tende, invece, a ritenere come questi non dovrebbero mai essere pubblicati senza un esplicito consenso da parte del soggetto coinvolto[378]. Vi possono però essere delle situazioni nelle quali l'interesse pubblico può prevalere sulla privacy, motivo per cui si può ritenere necessario rendere trasparenti determinate informazioni. Anche in queste circostanze si cerca ad ogni modo di applicare delle condizioni che ne impediscano il riutilizzo per finalità diverse da quelle strettamente necessarie a salvaguardare il pubblico interesse.

Bisogna tuttavia considerare che certi dati, una volta pubblicati soprattutto su internet, difficilmente possono essere poi occultati e rischiano così di perdere per sempre la loro segretezza[379]. Le politiche sull'information disclosure tendono per questo a riguardare la capacità di operatori e istituzioni attivi in particolari settori di avere a disposizione certi dati, seppur rispettando determinate condizioni, ma generalmente non prevedono una libertà assoluta ad accedere a informazioni personali. Si può in questo modo cercare di ponderare la necessità di rispondere a uno specifico bisogno della collettività con il desiderio del singolo di non rendere di pubblico dominio dati che lo riguardano.

Ciò detto, a livello internazionale si contano dei provvedimenti in materia di open data. La Francia è stato il primo paese ad aver previsto una legislazione specifica sul tema nel più ampio contesto di una riforma della normativa sulla privacy e della regolazione della digital economy. Il 7 ottobre 2016 è infatti entrato in vigore il French Digital Republic Act[380], tramite il quale si è voluta estendere l'applicazione di politiche di open data, ridefinendo il concetto di dato di pubblico interesse. Secondo la nuova legge, determinate informazioni devono essere rese accessibili sia se possedute da soggetti pubblici, sia da soggetti privati quali i detentori di concessioni pubbliche o di servizi sussidiati.

---

[376] Barbaro M. and Zeller T. Jr., *A Face Is Exposed for AOL Searcher*, No. 4417749, N.Y. TIMES, Aug. 9, 2006.

[377] Zevenbergen B., Brown I., Wright J., Erdos D., *Ethical Privacy Guidelines for Mobile Connectivity Measurements*, Oxford Internet Inst. 1, 11, Nov. 2013, http://www.oii.ox.ac.uk/research/Ethical_Privacy_Guidelines_for_Mobile_Connectivity_Measurements.pdf.

[378] Wendy Hall et al., *Open data and charities*, nominettrust.org 1, 16, July 2012, http://www.nominettrust.org.uk/sites/default/files/Open%20Data%20and%20Charities.pdf

[379] Arvind Narayanan et. al. (2016), vedi nota 30.

[380] Loi n°2016-1321 du 7 octobre 2016 pour une République numérique.

Il Frecnh digital Act ha poi facilitato l'accesso a determinati dati, anche di privati, da parte della Insee (National Institute of Statistics and Economic Studies), per migliorare la ricerca statistica. Aumentare la capacità per le agenzie nazionali statistiche di accedere ai dati privati viene d'altronde considerato necessario dalla maggior parte dei paesi europei[381]. Anche il Regolamento Generale sulla Protezione dei Dati (GDPR) è intervenuto a tale proposito, derogando alle comuni norme in tema di utilizzo di categorie particolari di dati personali, qualora il loro trattamento fosse *"necessario a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici"* (Reg. 2016/679, art. 9, co. 2, let. j).

Prendendo riferimento dal succitato regolamento europeo, la Germania ha approvato il 5 Luglio 2017 il cosiddetto German Data Protection Amendment Act (GDPAA), con il quale ha reso possibile l'utilizzo di dati sensibili, oltre che per fini statistici, anche per ragioni di interesse pubblico soprattutto in ambito medico-sanitario, come migliorare la medicina preventiva, la diagnosi medica, i trattamenti sanitari e sociali, gli accordi con i professionisti del settore. La capacità di accedere e trattare i dati è, tuttavia, possibile solo una volta prese attente misure di protezione per evitare un'eccessiva diffusione dei dati stessi. Tali misure devono essere adeguate alle circostanze del singolo caso e possono includere accorgimenti tecnici e organizzativi, pseudonimi, o la nomina di un responsabile della protezione dei dati[382].

Il miglioramento della salute pubblica è uno dei principali motivi per i quali può essere richiesto libero accesso a informazioni private. Oltre alla nuova normativa tedesca, anche in altri casi si è resa necessaria una maggiore apertura ai dati. Per esempio l'European Centre for Disease Prevention and Control ha ritenuto opportuno accedere ai dati dei cellulari durante un'epidemia di legionella, per definire la posizione e la possibile diffusione della malattia; anche in altri episodi per malattie trasmesse dalla tossicità dell'aria, da acqua contaminata o animali infetti. Altri obiettivi legati al pubblico interesse per i quali può essere ritenuto necessario l'utilizzo dei dati privati, possono riguardare situazioni dove è opportuno salvaguardare la sicurezza delle persone, per esempio in caso di catastrofi naturali o crisi umanitarie[383].

Vi sono comunque molti altri settori interessati ai big data fra cui l'urbanistica, i trasporti, l'energia e l'agricoltura. Bisogna però tenere presente che raccogliere e fornire dati può essere a volte molto

[381] La Commissione Europea ha organizzato un workshop il 26 Giugno 2017, proprio per discutere sull'accesso ai dati posseduti dai privati per gli istituti pubblici. Nel corso del lavoro i rappresentanti danese, spagnolo, francese, olandese e sloveno hanno di comune accordo richiamato la necessità di un maggior capacità delle agenzie statistiche di avere accesso ai dati. Cfr. *DG CONNECT draft report: 26 June 2017 workshop on access to privately-held data for public bodies*, link: https://ec.europa.eu/digital-single-market/en/news/workshop-access-public-bodies-privately-held-data-public-interest

[382] Cfr. Artcolo Bird & Bird LLP, *Germany is the first EU Member State to enact new Data Protection Act to align with the GDPR,* 2017, link: https://www.lexology.com/library/detail.aspx?g=2a2ff4e6-e0cd-4ca8-b0fc-2f8dd7d3f617

[383] Fra le società di comunicazione si segnala come Vodafone abbia vari progetti in corso in diversi settori in cui utilizza i propri dati per il bene pubblico, secondo quanto riferito dalla Direttrice della Data Science Research Vodafone Nuria Oliver in occasione del workshop CE.

costoso, specialmente se il monitoraggio avviene in tempo reale e vengono richiesti, quindi, continui aggiornamenti, motivo per cui l'accesso ai dati non può sempre essere imposto.

Gli interventi di obbligo a diffondere i propri dati devono, dunque, tenere presente la necessità delle aziende di recuperare i costi delle varie operazioni, ottenendo tra l'altro anche una congrua remunerazione del lavoro fornito. Il tema è di forte attualità, specialmente per ciò che riguarda l'efficientamento dei servizi e l'evoluzione delle città nel futuro. Spesso si richiama infatti il concetto di "smart city", per intendere un uso delle nuove tecnologie per migliorare i servizi e le strategie di pianificazione in ambito urbano. Per poter attuare progetti "smart" vi è però bisogno di migliorare la capacità di accedere ed elaborare dati sempre aggiornati. Le differenze di know how portano però a pensare che sia molto più conveniente per le istituzioni collaborare con soggetti privati, rispetto ad autoprodurre un determinato servizio di analisi e raccolta di dati. Per tale ragione è prevedibile un rapporto sempre più stretto fra i fornitori di big data e le pubbliche amministrazioni[384]. Tuttavia imporre a soggetti privati di rendere disponibile la loro banca dati o i loro algoritmi di elaborazione gratuitamente potrebbe essere una misura compromettente per gli equilibri del mercato, che potrebbe finire per produrre un danno oltre che per le imprese anche per i cittadini stessi. Favorire forme di Partenariato Pubblico-Privato potrebbe, invece, essere la chiave di successo per incentivare la crescita di città sempre più all'avanguardia ed in linea con le esigenze della collettività.

Infine si sottolinea come il settore che ha fatto i maggiori passi avanti sul tema dell'information disclosure è probabilmente quello della finanza. Con la recente Payment Service Directive (PSD2), gli Stati Membri dovranno infatti recepire all'interno dei propri ordinamenti nazionali una serie di norme entro il 13 Gennaio 2018, che renderanno irreversibile il passaggio in Europa verso un sistema di open banking[385]. L'obiettivo principale che il Parlamento europeo vuole raggiungere con questa direttiva è quello di rendere più facile, veloce e meno costoso per i consumatori pagare per beni e servizi. Per fare ciò verrà richiesto alle banche europee di aprire le proprie infrastrutture di pagamento e i dati di clienti agli altri fornitori di servizi finanziari. In questo modo si vuole favorire la concorrenza in un mercato che ha visto negli ultimi anni l'inserimento di nuovi attori, che sono stati spesso in grado di sfruttare al meglio le nuove tecnologie digitali.

Il sistema che le banche dovranno probabilmente utilizzare per fornire l'accesso agli altri fornitori di servizi finanziari, inclusi i nuovi TPP (Third Party Providers)[386], è quello delle API (Application Programming Interfaces), le quali rappresentano indubbiamente un modello interessante per trasmettere dati in modo sicuro, ma anche rapido ed economico. Considerando che le maggiori difficoltà riscontrate per l'information disclosure sono quelle di bilanciare privacy e trasparenza

[384] Cfr. Brauneis R., Goodman E. P., *Algorithmic Transparency for the Smart City*, 2017, retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012499

[385] Cfr. PWC, cit. (2016).

[386] In Germania, ad esempio, Sofort è diventato in meno di 10 anni leader nei pagamenti e-commerce con più 2 milioni di transazioni al mese e 35.000 merchant. Insieme a Trustly ed a altri TPP sta modificando il contesto tradizionale del business dei servizi bancari creando nuovi business model con i quali le banche si devono per forza confrontare.

facendo al contempo attenzione ai costi, le API sono probabilmente uno strumento da analizzare con attenzione per valutare un modo adeguato per favorire l'accessibilità ai big data e renderli disponibili per l'utilizzo nei settori socialmente più rilevanti.

4.4.3 L'apertura dei dati posseduti dalle pubbliche amministrazioni

Se per l'utilizzo dei dati posseduti da imprese private appare evidente la necessità di garantire una copertura dei costi ed una remunerazione del capitale investito, discorso differente può essere applicato per le banche dati detenute dal settore pubblico. Considerando come quest'ultimo non abbia come finalità principale quella dell'utile di esercizio, potrebbe infatti ritenere opportuno il rilascio di dati a fronte di un prezzo marginale o nullo pur di soddisfare un interesse collettivo. Una maggiore apertura dei dati pubblici potrebbe, infatti, produrre una serie di effetti positivi sia da un punto di vista economico, per via dei delle maggiori informazioni rese disponibili, soprattutto a start-up innovative e piccole/medie imprese, ma anche sociale dovute all'atteso aumento dell'occupazione ed alla maggiore credibilità che le istituzioni potrebbero acquisire nei confronti dei cittadini rendendo le loro informazioni più trasparenti.

La Commissione Europea sta recentemente affrontando questo tema, il quale è stato inserito all'interno del dibattitto concernente gli obiettivi posti nella Comunicazione *"Building a European Data Economy"* del 10 gennaio 2017[387]. Nello Staff Working Document che accompagna il documento si prevede per l'appunto di analizzare possibili modifiche a regole e regolamenti attualmente in vigore per favorire un libero flusso di dati all'interno dell'Unione. In particolare la Commissione si propone di preparare per la primavera del 2018 un'iniziativa per regolare l'accessibilità e il riutilizzo di dati pubblici e finanziati con fondi pubblici. La revisione della normativa vigente deve comunque basarsi su una valutazione della legislazione esistente, motivo per cui è iniziato nell'agosto 2017 un *Impact assessment* in merito alla Direttiva 2013/37/UE[388] relativa al riutilizzo dell'informazione del settore pubblico, anche detta "*PSI Directive*" (Public Sector Information Directive)[389].

Una possibile revisione sulla direttiva potrebbe consistere nell'estenderne il campo di applicazione allo scopo di favorire un maggiore accesso ai dati. Si potrebbe, per esempio, consentire il riutilizzo dei dati detenuti da enti del settore pubblico (come le società di servizi pubblici) attualmente esclusi dal campo di applicazione della direttiva PSI, a causa della natura commerciale o industriale delle loro attività, oppure dei dati detenuti da enti privati titolari di concessioni pubbliche[390].

[387] Per maggiori approfondimenti su *Building a European Data Economy* si veda il paragrafo 1.3.1.

[388] La Direttiva 2013/37/UE, attualmente in vigore, ha modificato la precedente Direttiva 2003/98/CE.

[389] Per maggiori informazioni sulla consultazione, si veda il link: https://ec.europa.eu/info/consultations/public-consultation-review-directive-re-use-public-sector-information-psi-directive_en.

[390] European Commission, *Review of the Directive on the re-use of public sector information (Directive 2003/98/EU)*, 3 agosto 2017, https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information.

Il tema che va però a incidere maggiormente sulla regolazione economica dei dati pubblici è quello che riguarda una tariffa congrua a fronte della quale concedere il riutilizzo delle banche dati. Secondo i sostenitori dei benefici dell'open data, tariffe troppo alte costituiscono un limite alla diffusione dell'informazione, motivo per cui i dati dovrebbero essere rilasciati a prezzi nulli o irrisori. L'attuale Direttiva PSI prevede a tale proposito che informazioni contenute in documenti del settore pubblico possano essere riutilizzate a fronte di un corrispettivo *"limitato ai costi marginali sostenuti per la loro riproduzione, messa a disposizione e divulgazione"*. Vi sono, tuttavia, una serie di casi che fanno eccezioni a tale norma, ovvero quando gli enti pubblici *"devono generare utili per coprire una parte sostanziale dei costi inerenti allo svolgimento dei propri compiti di servizio pubblico"* o nel caso di *"biblioteche, comprese le biblioteche universitarie, musei e archivi"*. Resta salva, inoltre, la facoltà degli Stati membri di stabilire che, in via eccezionale, l'utilizzo di determinati documenti possa generare utili. In tutti questi casi la tariffa massima applicabile non deve produrre comunque entrate superiori ai costi di raccolta, produzione, riproduzione e diffusione, maggiorati di una congrua remunerazione per gli investimenti effettuati.

Per valutare gli effetti di un eventuale intervento regolatorio sembra a questo punto opportuno osservare nuovamente i trade-off fra protezione dei dati e *information disclosure*, analizzando però questa volta la questione con riferimento ai dati delle pubbliche amministrazioni. Alcune norme inserite nella Direttiva PSI, come quelle che impongono agli enti pubblici di concedere il riutilizzo dei dati, sono state poste nell'ottica che determinati dati fossero di proprietà esclusiva del settore pubblico, per via della scarsa convenienza per soggetti privati di investire sulla loro generazione e raccolta. Tuttavia si è visto come il valore dei dati sia notevolmente aumentato, generando un sempre maggiore interesse da parte delle imprese a produrre dei database sempre più ampi, spesso anche superiori qualitativamente e quantitativamente a quelli pubblici[391]. Nell'economia attuale si è per questo creata una necessità sia di tutelare un mercato "a valle", costituito dalle imprese fornitrici di beni e servizi che necessitano dei dati per essere più efficienti, ma anche un mercato "a monte" formato da tutti quei soggetti pubblici o privati che investono nella raccolta e generazione di dati. Estendere oltremodo il campo di applicazione della direttiva potrebbe, quindi, scoraggiare la formazione e lo sviluppo di banche dati sempre più innovative ed aggiornate. Oltre ciò regolare una tariffa in base al costo marginale di duplicazione del documento finirebbe per produrre in un mercato potenzialmente concorrenziale gli effetti di un prezzo predatorio[392], creando così una insuperabile barriera all'accesso. Bisognerebbe, dunque, valutare in precedenza sia l'eventuale presenza di potenziali entranti, che i reali benefici derivanti dall'imporre l'apertura dei dati mediante un intervento regolatorio su larga scala.

[391] Dragos, D.C., Neamtu B, *Reusing public sector information — policy choices and experiences in some of the Member States with an emphasis on the case of Romania*, in European Integration Online Papers, 2009, p. 4, hanno focalizzando l'attenzione sui mezzi utilizzati per generare PSI, osservando come all'interno della PSI si potrebbe distinguere tra una specie di informazione, per così dire "privilegiata", che può essere ottenuta solo esercitando i pubblici poteri, e una specie di informazione che, al contrario, non sarebbe affatto totale appannaggio degli enti pubblici, potendo essere ottenuta anche da imprese appartenenti al settore privato.

[392] Vedi Maggiolino M., *I riutilizzo dell'informazione detenuta dal settore pubblico: alcune riflessioni di politica e diritto della concorrenza*, in Concorrenza e Mercato, 2012, pp. 765-802.

Il diritto a richiedere una congrua remunerazione per gli investimenti dedicati alla raccolta di dati viene per certi aspetti garantito, inoltre, anche dalla Direttiva 96/9/CE sulla tutela giuridica delle banche di dati, detta Direttiva Database. Essa è recentemente divenuta oggetto di una consultazione pubblica, la quale si è svolta tra il 24 maggio e il 30 agosto 2017. Lo scopo della consultazione era quello di capire in particolare l'impatto della tutela *sui generis* prevista a favore dei costitutori di database. La tutela *sui generis* si inserisce all'interno della normativa europea sul diritto d'autore, stabilendo che *"Gli Stati membri attribuiscono al costitutore di una banca di dati il diritto di vietare operazioni di estrazione e/o reimpiego della totalità o di una parte sostanziale del contenuto della stessa, valutata in termini qualitativi o quantitativi, qualora il conseguimento, la verifica e la presentazione di tale contenuto attestino un investimento rilevante sotto il profilo qualitativo o quantitativo"* (art. 6, Dir. 96/9/CE).

Nell'ambito della consultazione sono state poste ai rispondenti varie domande in merito agli effetti della direttiva, fra le quali è stato chiesto di stabilire se l'attuale applicazione del diritto *sui generis* fosse o meno appropriata. Si rileva a tale proposito come gli intervistati hanno risposto a riguardo in maniera contrastante, dividendosi fra chi ritiene la protezione troppo ampia e chi, invece, la considera non sufficiente. Ad esempio, la grande maggioranza degli intervistati nel settore della ricerca e accademico ritiene che le banche dati del settore pubblico non dovrebbero essere protette e, quindi, auspicano a una loro esclusione dalla tutela *sui generis*. Al contrario, la maggioranza degli intervistati nei servizi di trasporto e IT ritiene che i database costituiti da dati raccolti automaticamente o generati da una macchina debbano avere maggiore protezione[393]. Opinioni divergenti sono state raccolte anche in merito all'impatto della direttiva sugli investimenti per la creazione, l'aggiornamento e la manutenzione dei database. Ciò sembra, quindi, confermare il fatto che le valutazioni sull'apertura delle banche dati, sia detenuti da soggetti privati che pubblici, debbano essere effettuate con cautela, analizzando caso per caso la convenienza di un eventuale intervento regolatorio in merito.

[393] European Commission, *Synopsis Report on the responses to the public consultation on the evaluation of Directive 96/9/EC on the legal protection of databases*, 2017.

Capitolo 5. Privacy: economia dei dati e protezione dei dati personali

5.1 La tutela dei dati personali nel nuovo eco-sistema digitale

Il tema della protezione della privacy e dei dati personali è apparso spesso nelle riflessioni e nei dibattiti degli ultimi anni, soprattutto nel contesto del social networking e della profilazione degli individui, nella loro triplice valenza di consumatori, utenti e cittadini. È infatti evidente che, a fronte dei benefici derivanti dalla società della conoscenza, vi sia altresì il riconoscimento che la raccolta e l'analisi di crescenti quantità di dati possano dimostrarsi lesive della dignità dell'individuo e facilitino pratiche discriminatorie a suo detrimento. Rischi che si aggravano per i soggetti più vulnerabili e che, peraltro, promettono di non rallentare con il consolidarsi dell'Internet of Things e dei Big Data, innovazioni tecnologiche che si inquadrano all'interno di un contesto sociale ed economico nel quale la tutela dei dati personali ed i principi che si pongono alla sua base manifestano sempre più le loro fragilità. Pensiamo, ad esempio, come ogni giorno gli oggetti che utilizziamo, gli spazi pubblici che attraversiamo, gli spazi privati che dimoriamo, diventino sempre più "intelligenti". Ciò sta accadendo perché, come osservato nella prima parte del presente lavoro, questi oggetti sono arricchiti di tecnologie, sistemi informatici e algoritmi (la cosiddetta IA, Intelligenza Artificiale) in grado di imparare dall'esperienza e di adattarsi in maniera evolutiva ai comportamenti delle persone, alle condizioni dell'ambiente, al mutare dei contesti.[394] Questo elemento dell'innovazione tecnologica è in grado di innescare implicazioni enormi tanto sul funzionamento di questi oggetti, quanto su di noi, anche sul piano dei nostri diritti e libertà.

Lo sviluppo incessante della tecnologia digitale, con il moltiplicarsi delle grandi banche dati, il consolidamento del dominio dei brand tecnologici nell'economia globale,[395] e la massiccia condivisione di contenuti mediante i social media, richiedono inevitabilmente anche un aggiornamento, o finanche un ripensamento, delle forme di tutela dei dati personali. In questo nuovo scenario l'identità individuale è infatti definita dalle caratteristiche che altri soggetti ci attribuiscono, come risultato dei nostri comportamenti o comunque della nostra presenza nel web.[396] Dinnanzi ai modi nuovi in cui si costruisce l'identità individuale nella società della conoscenza, che manifesta una

[394] AgID (Agenzia per l'Italia Digitale) – *Task Force sull'Intelligenza Artificiale*, White paper IA, 2017, https://whitepaper-ia.readthedocs.io/it/latest/, ultimo accesso 12 ottobre 2017.

[395] Brand Finance, *Brand Finance Global 500 2017,* febbraio 2017, http://brandfinance.com/images/upload/global_500_2017_locked_website.pdf, ultimo accesso 12 ottobre 2017.

[396] Si pensi, ad esempio, alla profilazione a fini pubblicitari o alla cosiddetta FOMO (Fear of Missing Out), ossia la paura di essere tagliati fuori e che può indurre gli utenti ad essere consumati dal bisogno ossessivo di controllare ciò che gli altri fanno. Cfr. Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V., *Motivational, emotional, and behavioral correlates of fear of missing out*, Computers in Human Behavior, Vol. 29, 2013, 1814.

crescente dipendenza dalle altre persone e dall'ambiente esterno nel quale viviamo,[397] l'Unione Europea ha scelto di rafforzare la dimensione individuale della tutela dei dati personali attribuendogli lo status di diritto autonomo e fondamentale all'interno del Trattato di Lisbona.[398] Al contempo, tuttavia, nel panorama politico ed accademico continua a prendere forma una rinnovata dimensione collettiva[399] della tutela dei dati personali, con particolare riguardo agli impatti etici e sociali dell'uso dei dati derivanti dalla *big data analytics*.[400]

Lo sguardo all'*information privacy* ci restituisce così un'istantanea a tratti ben definita, a tratti quasi enigmatica. Tra gli elementi che si possono mettere a fuoco, appare in primo piano il Regolamento UE 2016/679 (al quale d'ora in poi per comodità si farà riferimento con la locuzione "Regolamento" o GDPR, *General Data Protection Regulation*), in vigore dal 24 maggio 2016 ed efficace a partire dal 25 maggio 2018. Dagli oltre quattro anni di tortuosi negoziati è nato un quadro regolamentare che mantiene i caposaldi dell'*information privacy* europea enucleati dalla precedente Direttiva 95/46/CE (alla quale d'ora in poi per comodità si farà riferimento con la locuzione "Direttiva madre")[401] e aggiunge ulteriori elementi di tutela, senza costituire un vero elemento di discontinuità. Attualizzando il *toolkit* per la protezione dei dati personali per mezzo di norme giuridiche comuni a tutti gli Stati Membri, promette di fornire una disciplina più efficace per la protezione dell'individuo e più efficiente per lo sviluppo del *Digital Single Market.*

In ragione di ciò, si è scelto di mantenere come filo conduttore della presente analisi il fenomeno del crescente ricorso alle tecnologie *data intensive* e al trattamento digitalizzato delle informazioni nei nostri sistemi economici e sociali. Esso permette di offrire uno sguardo al futuro integrando gli elementi di novità del Regolamento di imminente ingresso con i principali aspetti sui quali lo sviluppo tecnologico sfiderà le modalità di presidio e di tutela delle informazioni personali. Muovendo da un passaggio introduttivo e logicamente anteriore della progressiva ridefinizione della sfera privata, sempre più dilatata, sovrapposta e confusa con quella pubblica (par. 1), questo lavoro offre un confronto di prospettiva e di disciplina tra la normativa tuttora vigente della Direttiva e quella del nuovo Regolamento (par. 2). Per meglio comprendere le direzioni di *policy* intraprese a livello europeo, si propone anzitutto un'attualizzazione dei tradizionali principi che regolano la *data privacy*

[397] J.D. Lasica, *Identity in the Age of Cloud Computing: The Next-Generation Internet's Impact on Business, Governance and Social Interaction*, The Aspen Institute, Washington DC, 2009, 1.

[398] Trattato di Lisbona che modifica il trattato sull'Unione europea (TUE) e il trattato che istituisce la Comunità europea (TCE) (GU C 306 del 17.12.2007); entrato in vigore in data 1 dicembre 2009.

[399] Si veda infra par. 7.

[400] Mantelero A., *From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era*, in Taylor L., Floridi L., van der Sloot B. (a cura di), *Group Privacy. New Challenges of Data Technologies*, Springer International Publishing, 2017, 139-158.

[401] Direttiva 95/46/CE del Parlamento Europeo e del Consiglio del 24 ottobre 1995 relativa alla tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati. Essa ha rappresentato il primo intervento legislativo in materia di dati personali nell'Unione Europea e ha attribuito a questa area geopolitica un ruolo prominente nel panorama internazionale della loro difesa,

sulla base di una nuova, inedita, centralità dell'individuo, rilevabile sia al momento della scelta (consenso) di condivisione delle informazioni personali (par. 3), sia in seguito nelle modalità di trattamento dei dati (par. 5). Nel prosieguo, il lavoro illustra il nuovo approccio c.d. *by design/by default* alle vulnerabilità dei dati personali che richiede la valutazione dei rischi sin dalla progettazione del trattamento e che determina un rafforzamento di responsabilità per i soggetti coinvolti nel trattamento stesso (par. 4). Infine, il nuovo approccio europeo ai dati personali, in cui il rafforzamento dell'*accountability* di chi compie il trattamento è accompagnato da un parallelo *empowerment* del soggetto interessato, viene letto nell'ambito del panorama ben più ampio scaturente dall'intreccio tra economia dell'informazione *(information economy)* ed economia della condivisione *(sharing economy)* (par 6). In particolare, la parte conclusiva è dedicata ai big data e alle nuove vulnerabilità di privacy determinate dalla loro analisi, che si contraddistinguono per la moltitudine di soggetti coinvolti e le peculiari possibilità di inferenza informativa (par. 7).

5.2 Le principali novità introdotte dal GDPR

L'analisi del nuovo Regolamento richiede anzitutto di soffermarsi sulla sua diversa natura giuridica rispetto alla Direttiva madre, che abroga. Benché, infatti, la Direttiva madre fosse un atto di diritto dell'Unione Europea vincolante, e diretto sia agli Stati membri che ai singoli, il Regolamento introduce una disciplina in materia di protezione dati uniforme e valida in tutta Europa. Essendo "obbligatorio in tutti i suoi elementi e direttamente applicabile in ciascuno degli Stati membri",[402] è pertanto definito "self-executing" in quanto le sue disposizioni sono immediatamente applicabili senza il bisogno di alcuna intermediazione, quale un atto nazionale di recepimento o di attuazione.

Gli effetti di questa "differenza sistemica" tra Regolamento UE 2016/679 e Direttiva 95/46/CE non riguardano soltanto gli adempimenti richiesti agli Stati membri, i quali non devono adoperarsi per introdurre nuove norme europee nell'ordinamento interno, bensì sono chiamati ad adeguare le proprie normative interne al mutato regime europeo e garantirne la sua piena applicabilità, ma si riflettono principalmente sull'intera articolazione interna della disciplina. È bene infatti precisare che alla sostituzione della Direttiva madre, e del suo sistema di avvicinamento delle legislazioni nazionali, con il Regolamento, e il suo assetto legislativo incentrato sulla diretta applicabilità delle regole europee, si accompagna – come naturale conseguenza – una maggiore completezza e chiarezza dei contenuti normativi. Lo sforzo compiuto dal Legislatore europeo è ampiamente giustificabile e derivante dalle notevoli divergenze applicative emerse a livello sia legislativo sia giurisprudenziale dal 1996 ad oggi tra i diversi Membri dell'Unione. Questa scelta promette di attenuare di non poco i problemi emersi durante gli oltre vent'anni di vigenza della Direttiva madre ma potrebbe manifestare come contraltare l'introduzione di maggiore rigidità regolamentare, dove l'accresciuta precisione definitoria di principi ritenuti eccessivamente vaghi potrebbe rivelarsi sproporzionata in una materia dove l'influenza della innovazione tecnologica è particolarmente accentuata.[403]

[402] Articolo 288, comma 2, TFUE (trattato sul funzionamento dell'Unione europea).

[403] Pizzetti F., *Privacy e il diritto europeo alla protezione dei dati personali. Il Regolamento europeo 2016/679*, 2016, Giappichelli, 4, 152-153.

Esaurita questa doverosa premessa e volgendo lo sguardo al cuore della disciplina, il Regolamento si distingue come un intervento verso l'uniformità sostanziale, non solo formale, del sistema di tutela europeo. Ponendo l'accento sui piani dei diritti dell'individuo, dei doveri delle imprese e delle Pubbliche Amministrazioni, e inevitabilmente dei loro rapporti, l'impianto giuridico del Regolamento si differenzia da quello della Direttiva madre per offrire maggior solidità e azionabilità alla tutela dei dati personali. Nell'analisi che segue si esaminerà l'impianto delle disposizioni del Regolamento evidenziando il loro impatto sull'economia dei dati. In particolare si approfondiranno quelle nuove norme a tutela della riservatezza e dell'identità personale che, mutando i meccanismi volti a garantire la protezione dei dati personali, determinano anche notevoli cambiamenti nelle garanzie all'individuo, nell'attribuzione degli obblighi di tutela, e nella redistribuzione dei poteri di controllo.

5.2.1 Estensione dell'ambito territoriale

L'Articolo 3 del Regolamento introduce considerevoli novità rispetto all'ambito di applicazione territoriale della disciplina europea. Queste novità, è doveroso sottolineare, hanno quale presupposto le elaborazioni giurisprudenziali compiute della Corte di Giustizia Europea[404] e gli approfondimenti poi svolti dal "Gruppo di lavoro articolo 29"[405] in questo ambito, e ne rappresentano la mera codificazione.

In primo luogo, il GRDP estende il proprio campo di applicazione territoriale rispetto alla Direttiva rovesciando la concezione tradizionale del principio di stabilimento e sancendo l'applicabilità delle sue disposizioni ai Titolari[406] e ai Responsabili[407]del trattamento che si trovano nell'Unione, a prescindere dal luogo in cui sia effettuato il trattamento dei dati personali. In secondo luogo, il Regolamento si applica ai Titolari e ai Responsabili non stabiliti nell'Unione nel caso in cui il trattamento abbia ad oggetto dati personali di Interessati[408] che si "trovino" (anche virtualmente)

[404] Pizzetti F., *Privacy e il diritto europeo alla protezione dei dati personali. Il Regolamento europeo 2016/679*, 2016, Giappichelli, cit., 80, con riferimento a CGCE 13 maggio 2014, Google inc e Google Spain c. Agencia Española de Protección de Datos e Mario Costeja Gonzales.

[405] Il "Gruppo dell'articolo 29" (Working Party Article 29 o WP29, Comitato europeo della protezione dei dati nel GRDP) è un organismo con funzioni consultive composto da un rappresentante dell'Autorità Nazionale di controllo designato da ciascuno Stato membro e da un rappresentante dell'Autorità per le Istituzioni e gli organismi comunitari nonché da un rappresentante della Commissione. L'articolo 29 della Direttiva prevedeva, appunto, vari compiti da affidare ai membri dei Garanti nazionali, che quindi si riuniscono allo scopo.

[406] Il Titolare del trattamento è "la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali; quando le finalità e i mezzi di tale trattamento sono determinati dal diritto dell'Unione o degli Stati membri, il titolare del trattamento o i criteri specifici applicabili alla sua designazione possono essere stabiliti dal diritto dell'Unione o degli Stati membri" (GRDP, Articolo 4, n.7).

[407] Il Responsabile del trattamento è "la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che tratta dati personali per conto del titolare del trattamento (GRDP, Articolo 4, n.8).

[408] L'Interessato è la "persona fisica identificata o identificabile si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il

nell'Unione, se riguarda l'offerta di beni o servizi e/o il monitoraggio dei comportamenti di Interessati all'interno dell'UE. In altri termini, il GRDP produce un'efficacia "extraterritoriale" poiché sarà applicabile sia se l'Interessato si trovi realmente o virtualmente nel territorio europeo, sia se il Titolare e/o il Responsabile del trattamento siano stabiliti nell'Unione, ed anche se il trattamento viene effettuato all'esterno dell'Unione stessa.

Si giunge così ad un'innovazione che riflette la volontà di rispondere alle nuove esigenze della società dinnanzi alla rapida evoluzione del digitale e della globalizzazione, e "punta a rispondere alle sfide poste dagli sviluppi tecnologici e dai nuovi modelli di crescita economica, tenendo conto delle esigenze di tutela dei dati personali sempre più avvertite dai cittadini dei Paesi dell'Unione europea".[409]

5.2.2 Il diritto individuale alla protezione dei dati personali

Nel 2009, con l'emanazione del Trattato di Lisbona, è entrata in vigore la Carta dei diritti fondamentali dell'Unione Europea. L'Articolo 8 della Carta, per la prima volta, riconosce il "diritto alla protezione dei dati personali" come diritto fondamentale e autonomo, separato dal "rispetto della propria vita privata e familiare" ex Articolo 7. La creazione di questo diritto, in quanto diritto distinto dal diritto alla privacy, è senza dubbio significativa, e rappresenta un *unicum* dell'ordinamento giuridico europeo, essendo assente da altri strumenti internazionali a tutela dei diritti umani.[410]

Questo solenne riconoscimento non era invece presente nella Direttiva madre del 1995 giacché le sue norme erano pressoché incentrate sui doveri di diligenza ed i principi di *good governance* destinati ai Responsabili del trattamento dei dati, che si riscontravano operativamente nelle due macro categorie degli obblighi di trasparenza e di liceità del trattamento.[411] Questo perché la Direttiva nasceva in un'epoca nella quale le tecnologie informatiche stavano producendo una prima trasformazione nella raccolta delle informazioni, abilitando istituzioni pubbliche e private alla loro organizzazione in banche dati elettroniche e, quindi, ad un più agevole impiego dei dati per fini di analisi e studio rispetto al convenzionale trattamento analogico[412]. Ne deriva, pertanto, che essa

---

nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale (GRDP, Articolo 4, n.1).

[409] Garante per la protezione dei dati personali, *Guida al nuovo regolamento europeo in materia di protezione dei dati personali*, 2016, http://194.242.234.211/documents/10160/5184810/Guida+al+nuovo+Regolamento+europeo+in+materia+di+protezione+dati, ultimo accesso 18 ottobre 2017.

[410] Lynskey O., *Deconstructing data protection: The 'Added-Value' of a Right to Data Protection in the EU Legal Order*, International and Comparative Law Quarterly 63(3), 2014, 569–597.

[411] van der Sloot B., *Do data protection rules protect the individual and should they? An assessment of the proposed General Data Protection Regulation*, International Data Privacy Law 4(4), 2014, 307, 318-320.

[412] L'identità personale è il risultato di un processo: un processo di identificazione. Le caratteristiche costitutive dell'identità personale sono offerte dall'ambiente sociale rilevante, che attribuisce un certo significato a tali

nasceva dall'esigenza di disciplinare la crescente capacità di elaborazione digitale dei dati personali all'interno del territorio dell'Unione e si sostanziava nella definizione delle sue modalità, improntate a principi di liceità e correttezza, per assicurare la libera circolazione dei dati personali tra gli Stati Membri. In questo modo, l'interesse pubblico a beneficiare dell'impiego di soluzioni digitali per l'elaborazione dei dati, tra cui quelli personali, si contemperava con la "libertà da vincoli irragionevoli alla costruzione della propria identità personale".[413] Non era, dunque, un'epoca in cui esistevano gli attuali dispositivi di uso quotidiano come smartphone e app, che oggigiorno accompagnano ogni individuo in ogni spostamento, negli spazi di socialità a cui prende parte e nelle sue svariate attività di acquisto. Internet, che pure era una tecnologia disponibile, era appannaggio soltanto dell'1 per cento della popolazione europea.[414] In questo arco temporale si è prodotta una radicale metamorfosi dei processi di produzione e distribuzione dell'informazione, che ha dato vita ad una società dove gli individui e le istituzioni possiedono canali comunicativi totalmente inediti, dove il singolo pubblica liberamente (più o meno consapevolmente) i propri dati personali sulle piattaforme on-line e social media, e dove sia le imprese private che le autorità pubbliche hanno la possibilità di utilizzare considerevoli quantità di dati. Di fronte a tutto ciò, come risposta ai rischi di insufficiente tutela nel trattamento dei dati personali, si è andata affermandosi un'inedita centralità dell'individuo, per riequilibrare le asimmetrie di ruolo e potere che si producono tra l'Interessato, attore centrale e produttore di dati personali, e gli stakeholder che li commercializzano.[415] È un fenomeno recente, affiorato nel decennio scorso e poi sfociato nell'elaborazione del principio di autodeterminazione informativa *(Information Self-Determination*).[416] Detto principio, che presuppone la possibilità del singolo di decidere, al di fuori da pressioni esterne, se ed entro quali limiti rendere noti fatti legati alla propria vita personale, è alla base del riconoscimento del diritto autonomo alla tutela dei dati personali nel Trattato di Lisbona, con una manifesta consacrazione all'interno del GDPR. Proprio qui,

---

caratteristiche dell'identità come il sesso, l'etnia, o la lingua. Ne consegue che la costruzione dell'identità individuale è un processo di identificazione con qualcuno, o più, dei modelli disponibili nell'ambiente sociale: cfr. Pino G., *L'identità personale*, in Trattato di biodiritto, diretto da S. Rodotà e P. Zatti, vol. I, Ambito e fonti del biodiritto, a cura di S. Rodotà e M. Tallacchini, Giuffrè, Milano, 2010, 297-321.

[413] Agree P.E. – Rotenberg M., *Technology and Privacy*. The New Landscape, 2001, MIT Press, Cambridge; Massachusetts, 7.

[414]Reding V., *Outdoing Huxley: Forging a High Level of Data Protection for Europe in the Brave new Digital World*, Speech at Digital Enlightenment Forum, 18 giugno 2012, http://europa.eu/rapid/press-release_SPEECH-12-464_en.htm, ultimo accesso 12 ottobre 2017.

[415] L'informazione rivelata dall'Interessato potrebbe essere utilizzata dal Titolare per fornirgli il miglior servizio, ma potrebbe implicare anche conseguenze negative che variano da errori involontari nel trattamento dei dati alla discriminazione. Tutte conseguenze che derivano dalla rivelazione volontaria di dati personali ma che l'Interessato è raramente in grado di valutare e prevedere. Sull'argomento, è emblematico il paragone tra il rilascio di dati personali e la firma di un assegno in bianco: Varian H.R., *Economic Aspects of Personal Privacy, Privacy and Self-regulation in the Information Age*, National Telecommunications and Information Administration, 1996, p. 1.

[416] Rouvroy A – Poullet Y., *The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy*, in Reinventing Data Protection, 2009, Springer, Dordrecht. 45-76.

al Considerando 1 del nuovo Regolamento che apre l'intera nuova disciplina in materia, si ricorda infatti lo status di diritto fondamentale attribuito alla "protezione delle persone fisiche con riguardo al trattamento dei dati di carattere personale".

E' bene sottolineare sin d'ora che il riconoscimento del diritto fondamentale alla tutela dei dati personali non rappresenta un mero atto formale, un blando e indiretto sostegno alla costruzione di uno spazio di tutela rafforzato dinnanzi alle questioni via via poste dal "digital tsunami".[417]La presenza sullo sfondo di questo nuovo diritto soggettivo ha, tra l'altro, influenzato anche le sembianze del Regolamento, in cui sono stati introdotti nuovi diritti a favore dell'Interessato, dettagliati obblighi a carico del Titolare del trattamento, accresciuti livelli di *enforcement*, una decisa puntualizzazione dei principi già esistenti nella Direttiva e nuovi strumenti a tutela della sicurezza del trattamento.

5.2.3 Il trasferimento dei dati verso Paesi terzi o Organizzazioni internazionali

Oggi il trasferimento transfrontaliero dei dati riguarda molte organizzazioni, pubbliche e private, piccole e grandi, essendo ormai assai comune affidarsi, per esempio, a servizi di IT online come il *cloud computing*.Il flusso di dati tra paesi è dunque una realtà che cresce velocemente, ancor più del commercio internazionale di beni e servizi.[418]

L'Unione pone tuttavia dei vincoli rispetto al trasferimento dei dati personali al di fuori del territorio europeo. Questo perché il trattamento dei dati personali è tradizionalmente considerato uno dei tratti distintivi del sistema giuridico europeo, al quale si contrappongono approcci tuttora distanti adottati in altri paesi. Già la Direttiva madre vietava il trasferimento verso i paesi che non rispondevano agli standard di adeguatezza europei (principio dell'*adequacy)*, ma il Regolamento introduce criteri per la valutazione di adeguatezza *(adequacy test)* ancora più puntuali.[419]In primo luogo, allarga l'ambito di applicazione soggettiva del *test* assoggettandovi anche le organizzazioni internazionali. Inoltre precisa che l'autorizzazione al trasferimento dei dati personali all'interno del Paese terzo può essere circoscritta anche ad un suo territorio, o persino ad uno o più settori specifici (Articolo 45 n.1).Il catalogo dei requisiti di adeguatezza che la Commissione prende in considerazione per valutare l'adeguatezza è inoltre decisamente più ampio e più dettagliato nel quadro del GRDP, includendo lo stato di diritto, il rispetto dei diritti umani e delle libertà fondamentali, l'accesso delle autorità pubbliche ai dati personali, l'esistenza e l'effettivo funzionamento di una o più autorità di controllo indipendenti, gli impegni internazionali in relazione alla protezione dei dati personali

[417] The Future Group, *Freedom, Security, Privacy: European Home Affairs in an Open World*, 2008, http://www.statewatch.org/news/2008/jul/eu-futures-jha-report.pdf, ultimo accesso 12 ottobre 2017.

[418] Cory N., *Cross-Border Data Flows: Where Are the Barriers, and What Do They Cost?*, 2017, https://itif.org/publications/2017/05/01/cross-border-data-flows-where-are-barriers-and-what-do-they-cost, ultimo accesso 12 ottobre 2017.

[419] Con riferimento all'adequacy test e all'approccio europeo al trasferimento dei dati transfrontalieri dei dati personali, cfr. Kuner, C., *Reality and Illusion in EU Data Transfer Regulation Post Schrems*, German Law Journal, Vol. 18, 2017, 881.

(Articolo 45 n.2). Le decisioni della Commissione rimangono valide per un massimo di 4 anni e possono essere modificate, sospese o abrogate (Articolo 35 n.3). L'intento è di garantire un'efficace protezione dei dati personali soggetti alle norme dell'Unione ed evitare l'impiego di espedienti per sottrarsene; le prossime, concrete, misure della Commissione definiranno l'equilibrio tra questa esigenza e la capacità delle organizzazioni di affidarsi a decisioni di adeguatezza a lungo termine.

In mancanza di una decisione di adeguatezza, il GRDP mantiene sostanzialmente intatti i meccanismi di trasferimento dei dati creati dalla Direttiva nel caso di trasferimenti garantiti dal Titolare o dal Responsabile del trattamento. Ne cambia tuttavia le modalità, riducendo gli oneri di adempimento agli obblighi di legge. In primo luogo, il GRDP riconosce espressamente e disciplina l'uso delle Norme Vincolanti d'Impresa (anche note come *Binding Corporate Rules -BCR)*per legittimare il trasferimento transfrontaliero di dati personali. Le BCR sono strumenti impiegati dai gruppi di società per trasferire dati personali dall'Unione ad un Paese terzo, nel caso in cui il trasferimento avvenga tra società appartenenti allo stesso gruppo (Articolo 47).Mediante il Regolamento, ora la normativa europea elenca dei requisiti – necessari ma non sufficienti – per l'approvazione delle BCR ed i contenuti obbligatori.[420]Inoltre l'Articolo 46 n.2 ammette che l'impiego di clausole tipo di protezione dei dati *(Standard Data Protection Clauses)*, codici di condotta e altri meccanismi di certificazione possa costituire garanzia adeguata (Articoli 40 e 42). Questo significa che il trasferimento transfrontaliero è possibile se i Titolari e i Responsabili del trattamento stabiliti in un Paese terzo assumono poi contrattualmente l'impegno di applicare le garanzie precedentemente sottoscritte attraverso l'adesione ad un codice di condotta o ad uno schema di certificazione. Soprattutto, non è più necessario il parere positivo di un'Autorità di controllo che legittimi il trasferimento come, ad esempio, attualmente previsto in Italia dall'Articolo 44 del d.lgs. 196/2003(il cosiddetto "Codice Privacy").L'autorizzazione dell'autorità di controllo sarà invece ancora indispensabile nei casi in cui un Titolare decida di impiegare clausole contrattuali *ad-hoc*(cioè non riconosciute come adeguate tramite decisione della Commissione Europea) oppure accordi amministrativi stipulati tra autorità pubbliche – una delle novità introdotte dal regolamento (Articolo 46 n.3).

È bene evidenziare l'ingresso di queste novità come un deciso miglioramento rispetto alla disciplina del *data transfer* stabilita dalla Direttiva madre. Ciò nonostante, in questo ambito, le complessità e le incertezze rimangono numerose, come peraltro emblematizzate dalla sentenza della Corte di Giustizia dell'Unione europea sul caso Schrems.[421] La sentenza che ha posto fine al *Safe Harbour,* lo "storico" accordo fra Stati Uniti ed Unione Europea volto ad individuare le imprese che si impegnavano ad offrire standard di tutela ritenuti "adeguati" dalla Commissione, ha sì condotto ad una rapida ridefinizione dei rapporti tra i due paesi attraverso il nuovo "Privacy Shield", ma la sua

[420] L'elenco indicato al riguardo nel paragrafo 2 dell'Articolo 47 non è esaustivo e, pertanto, potranno essere previsti dalle autorità competenti, a seconda dei casi, requisiti ulteriori.

[421] Causa C-362/14 avente ad oggetto la domanda di pronuncia pregiudiziale proposta alla Corte di giustizia dell'Unione Europea, ai sensi dell'articolo 267 TFUE, dalla High Court (Corte d'appello, Irlanda), con decisione del 17 luglio 2014, pervenuta in cancelleria il 25 luglio 2014, nel procedimento Maximillian Schrems contro Data Protection Commissioner, con l'intervento di: Digital Rights Ireland Ltd.

capacità di tenuta rispetto agli standard di protezione stabiliti dal nuovo Regolamento è ancora tutta da verificare.[422]

5.2.4 Le Autorità di controllo

Il riconoscimento di un diritto soggettivo alla protezione dei dati personali e l'estensione dell'ambito di applicazione territoriale della disciplina ha altresì richiesto al Legislatore europeo la ri-considerazione dei meccanismi di controllo e la ri-organizzazione delle Autorità nazionali di controllo.

Di fronte all'evenienza, assai frequente, in cui un medesimo trattamento di dati sia operato dallo stesso Titolare in più di un paese dell'Unione e coinvolgendo Interessati di diversi Stati membri, il nuovo Regolamento ha voluto individuare un'unica Autorità di controllo di riferimento. Lo ha fatto introducendo il meccanismo del c.d. *one-stop-shop*[423]che contempla l'individuazione di un'Autorità di controllo capofila (*Lead Authority)* con riferimento al luogo dello stabilimento principale (e/o unico) del Titolare o del Responsabile che effettui trattamenti in più Stati membri. Le decisioni di una *Lead Authority* saranno vincolanti anche per i trattamenti effettuati dal Titolare (e dal Responsabile) negli altri Stati Membri e per quei trattamenti che utilizzano i dati personali di cittadini di altri paesi dell'Unione. Questo al fine di garantire maggior equità ai cittadini di ogni Stato membro in relazione alla tutela di un diritto fondamentale quale risulta essere quello alla protezione dei dati personali e di evitare che medesime violazioni possano essere oggetto di ricorsi decisi diversamente nei diversi Stati.

Più in generale, il Regolamento amplifica la disciplina assegnata alle Autorità nazionali. Lo fa inserendo due nuovi capitoli (Cap.VI e Cap. VII) rispettivamente dedicati alle Autorità di controllo e al sistema della cooperazione tra l'Autorità di controllo capofila e Autorità di controllo interessate, con

[422] Sica S., D'Antonio V., *Verso il Privacy Shield: il tramonto dei Safe Harbour Privacy Principles*, 137-168, in La Protezione Transnazionale dei Dati Personali (Resta G., Zeno-Zencovich V., a cura di), 2016, Roma Tre Press, http://romatrepress.uniroma3.it/ojs/index.php/PTD/article/view/3 (ultimo accesso, 18 ottobre 2017).

[423]"The 'one-stop-shop' principle, together with the consistency mechanism, is one of the central pillars of the Commission proposal. According to this principle, when the processing of personal data takes place in more than one Member State, one single supervisory authority should be competent for monitoring the activities of the controller or processor throughout the Union and taking the related decisions. The proposal states that the competent authority providing such one-stop-shop should be the supervisory authority of the member state in which the controller or processor has its main establishment. The Council expressed its support for the principle that, in important transnational cases, the regulation should establish a "one-stop-shop" mechanism in order to arrive at a single supervisory decision, which should be fast, ensure consistent application, provide legal certainty and reduce administrative burden. This is an important factor to enhance the cost- efficiency of the data protection rules for international business, thus contributing to the growth of the digital economy."Consiglio dell'Unione Europea, *Data protection: Council supports "one-stop-shop" principle*, Luxembourg, 7 October 2013, 14525/13,

http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/jha/138924.pdf (ultimo accesso 18 ottobre 2017).

lo scopo di non escludere alcuna Autorità di controllo eventualmente coinvolta nel trattamento.[424] Questa maggiore attenzione che il Regolamento dedica alle Autorità di controllo è una evidente conseguenza della natura *self-executing* del Regolamento, che implica la necessità di regolare nel dettaglio anche il funzionamento, le prerogative e i rapporti tra le diverse Autorità di controllo nazionali.

Infine, il "gruppo di lavoro Articolo 29" verrà sostituito dal Comitato europeo per la protezione dei dati (Articolo 68).Intende configurarsi come organismo dell'UE, dotato di personalità giuridica e composto dal responsabile di un'Autorità di controllo di ciascuno Stato membro e dal Garante europeo della protezione dei dati, o dai rispettivi rappresentanti. Il Comitato, a titolo esemplificativo e non esaustivo, è tenuto a consigliare la Commissione Europea in merito a qualsiasi questione relativa alla protezione dei dati personali dell'UE, comprese eventuali proposte di modifica del Regolamento, e pubblicare (i) linee guida, (ii) raccomandazioni e (ii) *best practice* al fine di promuovere l'applicazione coerente del GRDP.

5.2.5 Rimedi e sanzioni

Un ulteriore profilo che caratterizza il nuovo Regolamento e lo distingue dalla precedente disciplina della Direttiva madre è quello attinente al sistema dei rimedi e delle sanzioni.

Nel Regolamento trovano conferma i rimedi già presenti nella Direttiva, ma – anche in questo caso – con una codificazione più puntuale.[425] Figura il principio della responsabilità risarcitoria per il cd. "danno da trattamento" arricchito da una più precisa definizione dei meccanismi di ripartizione del risarcimento tra Titolare, Co-Titolare e Responsabile del trattamento, con previsione specifica di azioni di regresso reciproche e meccanismi di esonero (Articolo 82). Sia per quanto concerne i reclami presso le Autorità di controllo sia i ricorsi rivolti all'Autorità giurisdizionale, il Regolamento attribuisce all'Interessato che ritenga di aver subito trattamenti illegittimi la scelta tra lo Stato membro dove risiede abitualmente e il luogo in cui si è verificata la presunta violazione (Articoli 77 e 79). Vi è un chiaro rafforzamento della protezione dell'Interessato anche grazie al riconoscimento della facoltà di dare mandato a presentare reclamo e/o a ricorrere da parte di organizzazioni/associazioni che abbiano tra gli scopi statutari la difesa dei diritti dei propri associati (Articolo 80).

Il sistema sanzionatorio previsto nel Regolamento appare notevolmente inasprito rispetto a quello della Direttiva madre e a quello adottato nel nostro ordinamento dalla normativa di recepimento

[424] Cfr. Considerando 10 che, in assoluta coerenza, dispone un margine d'azione per gli Stati membri con riferimento al mantenimento o all'introduzione di disposizioni più specifiche, al fine di rendere più adattabili le prescrizioni del nuovo Regolamento.

[425] Nell'attuare la disciplina della Direttiva madre, il Codice Privacy italiano aveva definito un sistema costituito da sanzioni amministrative (Articolo 163) e penali (Articolo 169) in caso di trattamento illecito di dati personali e omessa adozione di misure di sicurezza. Ulteriori sanzioni sono previste ex Articolo 38 dello Statuto dei Lavoratori, legge n.300/1970, in caso di violazione della riservatezza negli ambienti di lavoro tramite controlli a distanza.

(Articoli 161 e ss del Codice Privacy).Prevede che l'Autorità di controllo competente abbia il potere di imporre sanzioni amministrative per un importo pecuniario anche commisurato alle percentuali del fatturato lordo mondiale dell'impresa e con soglie massime estremamente elevate (Articolo 83).[426]Nella determinazione del *quantum* è previsto che si debba tenere conto di determinati indici quali, ad esempio, (i) la natura, la gravità e la durata della violazione, (ii) il carattere doloso o colposo della stessa, (ii) le misure adottate dal Titolare. È bene infine notare che le sanzioni varieranno a seconda del trasgressore, essendo previste differenze di trattamento tra persona fisiche e imprese.Il Regolamento lascia invece autonomia agli Stati membri in merito alle sanzioni penali applicabili, che dovranno comunque assicurare effettività, proporzionalità e l'efficacia dissuasiva (Articolo 84). Quanto all'Italia, rimane pertanto da verificare come verrà modificato il reato di trattamento illecito dei dati personali attualmente previsto all'Articolo 167 del Codice Privacy.

5.2.6 Gli obblighi a carico del Titolare (e del Responsabile) del trattamento

Per contribuire ad una rafforzata protezione dei dati personali, la vera trasformazione si rinviene nell'assetto delle responsabilità del Titolare del trattamento, a cui peraltro si affiancano nuove figure, espressione di nuovi ruoli e funzioni organizzative.

Il GDPR divide la responsabilità dei dati personali tra il Titolare del trattamento e il Responsabile del trattamento(Articolo 4). Come nella previgente Direttiva, il Titolare del trattamento *(Data Controller)* rimane la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali (Articolo 4, n.7). Il Regolamento conserva anche il ruolo del Responsabile del trattamento *(Data Processor),*figura che diviene però obbligatoria e che si riferisce a quella persona fisica o giuridica, autorità pubblica, servizio o altro organismo che tratta dati personali per conto del Titolare del trattamento tramite nomina documentata per iscritto e assoggettamento al potere di controllo e disciplinare (Articolo 4, n.8). Nel nuovo Regolamento non viene invece contemplato l'Incaricato del trattamento, figura invece espressamente riconosciuta nel Codice Privacy. Pur in assenza di questo riconoscimento esplicito, il Regolamento non ne esclude la presenza e, dunque, non dovrebbe essere fonte di criticità per gli Stati membri – come l'Italia – che lo volessero conservare questa figura. Anzi, appare legittimata tale possibilità dalla facoltà riconosciuta al Titolare di autorizzare Terzi al

[426] Nel GRDP non sono previsti minimi edittali e le sanzioni amministrative pecuniarie sono distinte in tre fasce di gravità (Articolo 83, n. 4,5,6). Nella prima si contemplano sanzioni fino a un massimo di 10.000.000 euro, o per le imprese, fino al 2% del fatturato, per violazione degli obblighi del titolare e del responsabile del trattamento; violazione degli obblighi dell'organismo di certificazione; violazione degli obblighi dell'organismo di controllo. Nella seconda si contemplano sanzioni fino a un massimo di 20.000.000 euro, o per le imprese, fino al 4% del fatturato, per la violazione dei principi di base del trattamento comprese le condizioni relative al consenso; la violazione dei diritti degli interessati; i trasferimenti di dati a un destinatario in un paese terzo o organizzazione internazionale; la violazione di qualsiasi obbligo relativo a specifiche situazioni di trattamento; l'inosservanza di un ordine, di una limitazione provvisoria o definitiva di trattamento o di un ordine di sospensione dei flussi dei dati dell'Autorità di Controllo. Nella terza si contemplano sanzioni fino a un massimo di 20.000.000 euro, o per le imprese, fino al 4% del fatturato, per l'inosservanza di un ordine da parte dell'Autorità di controllo derivante dai suoi poteri correttivi (di cui all'Articolo 58, n.2).

trattamento, persone sia fisiche che giuridiche che operano sotto la sua diretta responsabilità o di quella del Responsabile (Articolo 4, n.10).

Rispetto alla Direttiva madre, il GRDP presenta un capovolgimento di prospettiva destinando la parte più significativa della disciplina agli adempimenti dei Titolari e dei Responsabili del trattamento e alle misure di sicurezza adottabili o da adottare.[427]In particolare, gli obblighi di Titolari e Responsabili sono ora rafforzati al fine di assicurare la conformità del trattamento alle norme del Regolamento e fornire livelli di sicurezza adeguati al rischio per i diritti e le libertà dell'Interessato.

Ecco, così, fare ingresso numerose disposizioni in materia di trasparenza, benché il principio di trasparenza non sia stato concepito *ex novo* e costituisca il risultato di un'evoluzione normativa che ha avuto quali due pilastri fondanti il principio di correttezza prescritto dalla Convenzione n. 108 del Consiglio d'Europa[428]e il principio di lealtà prescritto dalla Direttiva madre.[429] Il GRDP, tuttavia, al Considerando 39,afferma che "il principio della trasparenza impone che le informazioni e le comunicazioni relative al trattamento di tali dati personali siano facilmente accessibili e comprensibili e che sia utilizzato un linguaggio semplice e chiaro" e, all'Articolo 5 a), individua la trasparenza come terzo requisito, insieme alla correttezza ed alla liceità, che deve essere posto alla base di qualunque tipo di trattamento. Non solo, il principio di trasparenza si traduce anche in obblighi a carico del Titolare e del Responsabile del trattamento con riferimento alle modalità di trattamento dei dati. Esso richiede che le informazioni fornite all'Interessato: (i) siano rese con un linguaggio semplice e chiaro, soprattutto nel caso di minori; (ii) abbiano sempre forma scritta, essendo l'informativa in forma orale ammessa solo quando ciò è richiesto dall'Interessato e l'identità di questi possa essere provata con altri mezzi; (iii) includano, per esempio, (i) le finalità del trattamento; (ii) il periodo di conservazione dei dati personali *(data retention)* (ii) l'intenzione di trasferire i dati a un Paese terzo (Articolo 14).

Gli Articoli 33 e 34 del GRDP introducono un'ulteriore novità nello scenario degli adempimenti dovuti dal Titolare del trattamento: l'obbligo di Notifica/Comunicazione di *data breach,* ossia di violazione dei dati personali dell'Interessato a causa di violazione dei propri sistemi informatici. In ipotesi di *data breach,* in virtù dell'Articolo 33, il Titolare del trattamento (dopo essere eventualmente venuto a conoscenza dal Responsabile del trattamento, che è comunque tenuto ad informarlo senza ingiustificato ritardo) è chiamato a notificare la violazione all'Autorità di controllo competente,

[427] Pizzetti F., *Privacy e il diritto europeo alla protezione dei dati personali. Il Regolamento europeo 2016/679*, 2016, Giappichelli, cit., 154.

[428] La Convenzione del 28 gennaio 1981, n. 108, è il primo strumento internazionale cogente in materia dati personali. Cfr. Articolo 5 dove si indica che "personal data undergoing automatic processing shall be obtained and processed fairly".

[429] Secondo il Considerando 38 il "trattamento leale dei dati presuppone che le persone interessate possano conoscere l'esistenza del trattamento e disporre, quando i dati che le riguardano sono forniti direttamente da loro, di un'informazione effettiva e completa in merito alle circostanze della raccolta".

sempre senza ingiustificato ritardo[430] e, possibilmente, entro 72 ore dal momento in cui ne è venuto a conoscenza, a meno che sia improbabile che la violazione dei dati personali presenti un rischio per i diritti e le libertà delle persone fisiche. Detta Notifica ha dei contenuti obbligatori: deve "descrivere la natura della violazione dei dati personali compresi, ove possibile, le categorie e il numero approssimativo di Interessati in questione nonché le categorie e il numero approssimativo di registrazioni dei dati personali in questione; comunicare il nome e i dati di contatto del responsabile della protezione dei dati o di altro punto di contatto presso cui ottenere ulteriori informazioni più specifiche; descrivere le probabili conseguenze derivanti dalla violazione dei dati personali e […] le misure adottate dal Titolare del trattamento o di cui lo stesso ne propone l'adozione per porre rimedio alla violazione dei dati personali e anche, se del caso, per attenuarne i possibili effetti negativi". L'Articolo 34,rubricato "Comunicazione di una violazione dei dati personali all'Interessato", prevede poi un obbligo di Comunicazione a carico del Titolare del trattamento alle persone Interessate, ma soltanto nel caso in cui la violazione dei dati personali sia suscettibile di presentare un rischio *elevato* per i diritti e le libertà delle persone fisiche. La Comunicazione deve avere contenuto simile a quella della Notifica ma il suo linguaggio deve essere particolarmente semplice e chiaro. Vi sono però anche eccezioni all'obbligo di Comunicazione, che sono azionabili qualora "il Titolare del trattamento abbia attuato le misure tecniche e organizzative adeguate di protezione e tali misure erano state applicate ai dati personali oggetto della violazione, in particolare quelle destinate a rendere i dati personali incomprensibili a chiunque non sia autorizzato ad accedervi, quali la cifratura; [oppure] il Titolare del trattamento abbia successivamente adottato misure atte a scongiurare il sopraggiungere di un rischio elevato per i diritti e le libertà degli interessati [ed infine nel caso in cui la] comunicazione richiederebbe sforzi sproporzionati. In tal caso, si procede a una comunicazione pubblica o a una misura simile, tramite la quale gli interessati vengono informati con analoga efficacia." In altri termini, sussistono talune condizioni affinché l'obbligo di Comunicazione all'Interessato si determini: si richiede una situazione di rischio elevato (e non di mero rischio) per coinvolgere l'Interessato e abilitarlo a prendere le necessarie precauzioni, e si rinuncia al suo coinvolgimento quando il Titolare abbia già preso le necessarie precauzioni. Tutto ciò probabilmente per ovviare al pericolo di causare inutili allarmismi a fronte di danni soltanto potenziali ed evitare una intensa attività di comunicazione tra Titolare e Interessato, anche quando non sia in sé necessaria.

Per le imprese o organizzazioni con più di 250 dipendenti, l'Articolo 30 del Regolamento introduce il nuovo obbligo per il Titolare e il Responsabile del trattamento (e, ove applicabile, per i loro rispettivi rappresentanti) di tenere, in forma scritta, anche in formato elettronico, un Registro delle attività di trattamento. Il Registro deve includere a) nome e dati di contatto del titolare del trattamento e, ove applicabile, del contitolare del trattamento, del rappresentante del titolare del trattamento e del

[430] Il requisito dell'ingiustificato ritardo che riguarda sia la Comunicazione della violazione dal Responsabile al Titolare, sia dal Titolare all'Autorità di controllo (Articolo 33) e all'Interessato (Articolo 34) consegue a quanto stabilito nel Considerando 85, secondo il quale, "una violazione dei dati non affrontata tempestivamente può causare danni […] alle persone fisiche, ad esempio perdita del controllo dei dati personali che li riguardano o limitazione dei loro diritti, discriminazione, furto o usurpazione d'identità, perdite finanziarie, decifratura non autorizzata della pseudonimizzazione, pregiudizio alla reputazione, perdita di riservatezza dei dati personali protetti da segreto professionale o qualsiasi altro danno economico o sociale significativo alla persona fisica interessata."

responsabile della protezione dei dati; b) finalità del trattamento; c) descrizione delle categorie di interessati e delle categorie di dati personali; d) categorie di destinatari a cui i dati personali sono stati o saranno comunicati, compresi i destinatari di paesi terzi od organizzazioni internazionali; e) ove applicabile, trasferimenti di dati personali verso un paese terzo o un'organizzazione internazionale, compresa l'identificazione del paese terzo o dell'organizzazione internazionale e, per i trasferimenti a persone aventi un legittimo interesse - ammesso soltanto su richiesta di tali persone o qualora tali persone ne siano i destinatari - la documentazione delle garanzie adeguate; f) ove possibile, i termini ultimi previsti per la cancellazione delle diverse categorie di dati; g) ove possibile, una descrizione generale delle misure di sicurezza tecniche e organizzative ex Articolo 32. Il Registro, su richiesta, deve essere messo a disposizione dell'Autorità di controllo.

5.3 I principi che regolano il trattamento dei dati personali

Nella Direttiva madre del 1995la disciplina del trattamento dei dati personali era incentrata sui principi cardine definiti dallo schema dei *Fair Information Practice Principles (FIPPs),*[431] delle linee guida dell'OCSE del 1980[432] e della Convenzione del Consiglio d'Europa del 1981.[433]Detta centralità permane anche nel nuovo Regolamento ed i principi rimangono quelli di sempre:

(i) Liceità e correttezza (Articoli 5. N.1 a) e 6)

Il trattamento è lecito solo alle condizioni individuate dall'Articolo 6, ossia quando l'Interessato ha espresso il proprio consenso informato al trattamento dei dati per una o più specifiche finalità, o quando il trattamento è necessario all'esecuzione di un contratto di cui l'Interessato è parte, o ancora quando il trattamento è necessario per adempiere un obbligo legale a cui è soggetto il titolare del trattamento. Infine il trattamento è lecito quando lo stesso è necessario per la salvaguardia degli interessi vitali dell'Interessato o di un'altra persona fisica ovvero quando è necessario per l'esecuzione di un compito di interesse pubblico o per il perseguimento del legittimo interesse del Titolare del trattamento. In questa ultima circostanza, è tuttavia richiesto un bilanciamento tra gli interessi del Titolare del trattamento e gli interessi, i diritti e le libertà fondamentali dell'Interessato.

---

[431] Cfr. United States Department of Health, Education and Welfare, Records, Computers and the Rights of Citizens, 1973, https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CBwQFjAAahUKEwjtlf6aqM7IAhUDTBQKHfTmC-g&url=http%3A%2F%2Fwww.justice.gov%2Fopcl%2Fdocs%2Frec-com-rights.pdf&usg=AFQjCNFnbUqerP4R4ZlaiVJCvXddJL5EOw&cad=rja (ultimo accesso, 12 ottobre 2017). Cfr. anche Privacy Act of 1974, 5 U.S.C. § 552a.

[432] Cfr. OECD, *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, 23 settembre 1980, http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm (le lineeguidadell'OECDsono state recentementeaggiornatenel 2013, cfr.http://www.oecd.org/internet/ieconomy/privacy-guidelines.htm) (ultimo accesso, 12 ottobre 2017).

[433] Cfr. Consiglio d'Europa, *Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale*, Trattato n. 108, Strasburgo, 28 gennaio1981, http://www.coe.int/it/web/conventions/full-list/-/conventions/treaty/108 (ultimo accesso, 12 ottobre 2017).

(ii) Trasparenza (Articolo 5)

Oltre ad essere un obbligo a carico del Titolare e del Responsabile del trattamento, la trasparenza è una caratteristica fondamentale che deve avere il trattamento in sé, nonché un diritto dell'Interessato. Per tali ragioni è richiesto che siano trasparenti le modalità con cui sono raccolti e utilizzati i dati personali. Devono essere facilmente accessibili e comprensibili le informazioni e comunicazioni relative al trattamento, tra cui l'identità del Titolare del trattamento, le finalità del trattamento, ed i diritti dell'Interessato.

(iii) Limitazione delle finalità dei dati (Articolo 5, N.1, b)

Rimane valido il principio secondo cui i dati personali devono essere raccolti per finalità determinate, esplicite e legittime, e che successivamente devono essere trattati in una modalità che sia compatibile con tali finalità. Il trattamento dei dati per finalità diverse da quelle per le quali sono stati inizialmente raccolti dovrebbe essere consentito solo se compatibile con tali iniziali finalità. È invece possibile l'ulteriore trattamento ai fini di archiviazione nel pubblico interesse o di ricerca scientifica o storica o a fini statistici.

(iv) Minimizzazione dell'uso dei dati(Articolo 5, N.1, c)

Il principio della data minimisation richiede che i dati debbano essere sempre adeguati, pertinenti e limitati a quanto necessario per il perseguimento delle finalità per cui sono raccolti e trattati.

(v) Esattezza dei dati(Articolo 5, N.1, d)

Il significato autoevidente di questo principio è che i dati devono essere sempre esatti e aggiornati. Ne consegue che eventuali inesattezze richiedano sempre una tempestiva rettifica o cancellazione.

(vi) Limitazione della conservazione (Articolo 5, N.1, e)

Il cosiddetto principio della data retention richiede infine che i dati siano conservati soltanto per il tempo necessario al raggiungimento delle finalità per le quali sono trattati.

Il Regolamento esplicita poi il nuovo principio di Integrità e Riservatezza(Articolo 5, N.1, f) esigendo che i dati debbano essere trattati in maniera da garantire una sicurezza adeguata, il che prevede l'adozione di misure di sicurezza tecniche ed organizzative adeguate per proteggere i dati stessi da trattamenti non autorizzati o illeciti, dalla loro perdita o distruzione o dal danno accidentale.

Il Regolamento ha portato anche diverse novità con riferimento al principio del consenso ma, al Considerando 171, si precisa anzitutto che "qualora il trattamento si basi sul consenso a norma della direttiva 95/46/CE, non occorre che l'interessato presti nuovamente il suo consenso, se questo è stato espresso secondo modalità conformi alle condizioni del […] Regolamento, affinché il Titolare del trattamento possa proseguire il trattamento". Dunque, a chi è ad oggi di fatto allineato con le previsioni del Regolamento, non è richiesto di raccogliere di nuovo tutti i consensi.

Ai sensi dell'Articolo 7, il consenso dell'Interessato al trattamento dei dati personali deve essere, come da previgente Direttiva, preventivo e inequivocabile, anche quando espresso attraverso mezzi

elettronici. Per trattare i dati sensibili, il Regolamento prevede che il consenso debba essere esplicito. Viene esclusa ogni forma di consenso tacito oppure ottenuto proponendo a un Interessato una serie di opzioni già selezionate. Il consenso può essere revocato in ogni momento ma i trattamenti effettuati fino a quel momento dal Titolare sulla base del consenso rimarranno comunque legittimi. I fornitori di servizi internet e i social media, dovranno richiedere il consenso ai genitori o a chi esercita la potestà genitoriale per trattare i dati personali dei minori di 16 anni (Articolo 8). Sono altresì menzionati i criteri da rispettare per la redazione delle informative per ottenere un valido consenso "informato" (Articolo 12). Le informazioni sulle modalità di trattamento dei dati devono essere "concise, trasparenti, intellegibili e facilmente accessibili, con linguaggio semplice e chiaro." Sono caratteristiche essenziali che possono essere raggiunte anche grazie all'utilizzo di moduli, schemi e disegni idonei a informare e diffondere le notizie sulla correttezza dei trattamenti (Articoli 14 e 14bis).

Giova precisare che l'attenzione riposta dal Regolamento all'informativa per il consenso contribuisce a rendere effettivo e a potenziare anche il "diritto di accesso" ai dati personali, già previsto nella Direttiva madre e declinato con maggior puntualità nel Regolamento. Nelle nuove norme ex Articolo 12 e 15, qualora l'Interessato eserciti il diritto di accesso ai propri dati, il Titolare sarà tenuto a informare l'Interessato entro un mese dal ricevimento della richiesta e in caso di ritardo il Titolare sarà obbligato a darne giustificazione precisandone i motivi e precisando la facoltà di proporre reclamo all'Autorità di controllo e/o ricorso all'autorità giudiziaria. Più in generale, nel quadro della *data privacy,* con il Regolamento l'Informativa diviene un chiaro strumento volto a garantire che la volontà dell'Interessato possa trovare effettiva espressione. Ne è la premessa la *Cookie Law* in vigore dal giugno 2015 che, al netto di talune eccezioni, non permette installare cookie prima di: (i) aver predisposto e mostrato all'utente una cookie policy; (ii) aver predisposto e mostrato all'utente un banner informativo, e (iii) aver richiesto il consenso agli utenti[434]. Ciò significa che la legittimità di un'azione di trattamento dei dati personali, come l'installazione di un cookie di profilazione, diviene subordinata alla volontà dichiarata dall'Interessato, con il passaggio dall'*opt-out* all'*opt in.*

5.4 I nuovi strumenti a tutela della sicurezza delle informazioni: Accountability, Data protection by design e by default, DPIA e DPO

Il principio di *accountability*[435] prevede che "tenuto conto della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, nonché dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche, il titolare del trattamento mette in atto misure tecniche e organizzative adeguate per garantire, ed essere in grado di dimostrare, che il trattamento è effettuato conformemente al presente regolamento. Dette misure sono riesaminate e aggiornate qualora necessario". Peraltro il nuovo Regolamento non si limita a stabilire questa responsabilità in termini astratti. Al contrario compie un passo deciso verso il Titolare del trattamento, specificando la

[434] La Direttiva 2009/136/CE ha modificato la direttiva 2002/58/CE (E-Privacy), imponendo al gestore del sito web di informare l'utente del fatto che fa uso dei cookie sul sito e, in caso di cookie di profilazione, ad ottenere il consenso preventivo all'uso degli stessi.

[435] Cfr. Articolo 24 n.1.

necessità di tenere in considerazione la natura, l'ambito di applicazione, il contesto e le finalità del trattamento (Articolo 24, n.1), chiedendo che queste valutazioni si traducano sia nella scelta delle misure tecniche e organizzative da adottare nel contesto aziendale sia nell'attuazione di *privacy policy* conformi (Articolo 24 n.2).

Non solo, la tutela dei dati personali deve essere ora inclusa dal Titolare all'interno della progettazione e sviluppo dei sistemi di trattamento dei dati, siano essi hardware che software. Esiste una lunga lista di pratiche e tecnologie che possono essere utilizzate per eliminare, o comunque ridurre, i problemi e le vulnerabilità a cui sono soggetti i dati personali. Una classica suddivisione distingue questi strumenti in reattivi e proattivi, in base al fatto che limitino soltanto i danni nel momento in cui l'esito dell'attacco è compiuto (strumenti reattivi, come le PET – *Privacy-Enhancing Technologies),* o anticipino situazioni di pericolo e abilitino azioni per evitare che l'attacco si produca (strumenti proattivi, come la *Privacy-by-Design).* L'Articolo 25, rubricato *"Data protection by design and by default"* nel suo testo inglese, abbraccia il complesso di contromisure possibili includendo sia gli strumenti reattivi che proattivi. Il Regolamento domanda ai Titolari di non limitarsi a delegare a soggetti terzi la gestione dei dati personali tramite sistemi esterni, avulsi dall'organizzazione aziendale, bensì di affrontare l'esigenza di protezione delle persone fisiche con riguardo al trattamento dei dati personali direttamente all'interno della progettazione delle tecnologie informatiche, delle pratiche commerciali, delle infrastrutture di rete, e più in generale nelle scelte strutturali.[436] La scelta di quali strumenti adottare può essere valutata dal Titolare sotto vari punti di vista (validità temporale, utilizzi, peculiarità, costi di attuazione) ma sempre tenendo in dovuta considerazione le probabilità di rischio del trattamento da compiere in base al contesto e alla finalità (Articolo 25 n.1).

L'Articolo 25 n.2 cita poi alcune PET, a mero titolo esemplificativo, come termine di riferimento delle misure adottabili, come la pseudonimizzazione dei dati e la minimizzazione della loro raccolta. Il Considerando 78, fa però specifico riferimento alla pseudonimizzazione come misura volta a garantire il rispetto dei principi sulla protezione dei dati personali.[437] Proprio in virtù dell'importanza attribuita dal GDPR a tale soluzione tecnica, è opportuno delinearne gli elementi principali, nonché le differenze sostanziali rispetto ad altre misure[438].Con il termine pseudonimizzazione, si fa riferimento all'impossibilità di attribuire determinati dati personali ad un Interessato specifico senza l'utilizzo di informazioni aggiuntive. Per questa ragione, ai sensi del Considerando 26, i dati pseudonimizzati devono essere trattati come informazioni su una persona fisica non identificata, ma identificabile e dunque il Titolare dovrà rispettare i principi relativi alla normativa vigente. Tuttavia, l'applicazione

[436] Cavoukian, A., *Privacy by Design: The 7 Foundational Principles*, 2010, https://www.ipc.on.ca/wp-content/uploads/Resources/pbd-implement-7found-principles.pdf, ultimo accesso 12 ottobre 2017.

[437] Cfr. considerando 78, Regolamento 2016/679 del Parlamento Europeo e del Consiglio "[...] tali misure potrebbero consistere, tra l'altro, nel ridurre al minimo il trattamento dei dati personali, pseudonimizzare i dati personali il più presto possibile [...]".

[438] Sul tema si veda anche D'acquisto G.-Naldi M., *Big data e Privacy by design. Anonimizzazione, pseudonimizzazione e sicurezza*, Giappicchelli, Torino, 2017, 41-171.

della pseudonimizzazione permette al Titolare del trattamento di effettuare un trattamento legittimo senza la previa richiesta del consenso dell'Interessato a condizione che le informazioni aggiuntive di cui sopra, siano conservate separatamente e siano soggette a misure tecniche e organizzative volte a garantire la non attribuzione a una persona identificata o identificabile, così come previsto dall'Articolo 32 n.1 e dal Considerando 28 del Regolamento. Distinta dalla pseudonimizzazione è l'anonimizzazione. Essa si configura come il risultato del trattamento di dati personali volto a impedire irreversibilmente l'identificazione dell'Interessato e consiste nell'utilizzo di una o più tecniche (criptazione, generalizzazione, *hashing,* permutazione) progettate per rendere impossibile l'identificazione di un determinato individuo sulla base dei dati che lo riguardano. Così come previsto dal citato Considerando 26 del Regolamento, i principi di protezione dei dati, nonché le misure a tutela dell'Interessato, si applicano a tutte le informazioni relative ad una persona fisica identificata o identificabile: i dati resi anonimi non rientrano nell'ambito di applicazione della legislazione in materia di protezione dei dati e pertanto il loro trattamento non necessita né del consenso né delle ulteriori misure individuate dalla legge.

Aldilà della pseudonimizzazione e della *data minimisation,* il GDPR non disciplina specificamente alcuna misura tecnica volta a garantire il rispetto dei *principi di privacy by design & by default*.Il GDPR lascia ampio margine di discrezionalità al Titolare nell'adozione di quelle misure che, in seguito ad una valutazione sulle finalità, sulla tipologia del trattamento e sui rischi per gli interessati, vengano ritenute idonee a rispettare i principi previsti per la protezione dei dati personali.[439]Ciò premesso, la virata compiuta dal Regolamento verso il principio della PbD,benché sfuggente nei suoi effettivi contorni, rappresenta un cambio di paradigma per i Titolari del trattamento, che ora sono chiamati a considerare la protezione dei dati sin dall'inizio del processo di pianificazione del trattamento, e come suo perno. Questa nuova prospettiva è la base di partenza per un sistema più olistico ed integrato del trattamento dati, siano essi personali o non personali, da parte dei Titolari.[440]

Un approccio, quest'ultimo, che si contrappone a quello precedente, per la potenziale capacità di adattarsi, ad esempio, anche all'analisi dei big data. Il termine "big" in "big data" è innovativo nella misura in cui punta ad incrociare la dimensione quantitativa e quella qualitativa del dato, facendo uso di algoritmi sempre più in grado di comprendere i dati ed estrarne relazioni semantiche con altri dati, ottenere informazioni attraverso qualità ritenute pertinenti e attendibili (componente qualitativa), all'interno di una mole crescente di dati (componente quantitativa). Il trattamento dei big data si distingue pertanto per una pluralità pressoché illimitata di finalità e di tipologie di soggetti coinvolti. Tale componente innovativa dei big data apparentemente stride sia con le classiche modalità di presidio dei dati personali sia con quelle più recenti fornite dalla PbD, gravitando

[439] Attraverso il combinato disposto dell'Articolo 25 con il Considerando 74 è tuttavia possibile delineare due obblighi che sorgono in capo al Titolare del trattamento. In prima istanza, l'adozione di "misure tecniche e organizzative adeguate per garantire, ed essere in grado di dimostrare, che il trattamento è effettuato conformemente al presente regolamento". In secondo luogo, la capacità di dimostrare e documentare l'efficacia e l'adeguatezza delle misure adottate.

[440] Klitou D., *Privacy-Invading Technologies and Privacy by Design: Safeguarding Privacy, Liberty and Security in the 21st Century*, 2014, Springer, 260-268.

entrambe attorno alla protezione dell'individuo e del singolo dato personale[441], alla centralità dell'utente[442], e alla data minimisation. Seppure, dunque, questi rimangano i caratteri dell'impianto regolatorio europeo, è bene ricordare come la PdD non sia invece necessariamente legata ad un approccio statico della tutela, centrato sulla responsabilità del Titolare e ferma sulle condizioni iniziali di raccolta delle informazioni. Da un lato è sì vero che la PbD richiede di incorporare il requisito della tutela della privacy in fase di progettazione del trattamento ma non si esaurisce in questo momento e si estende alle successive fasi del effettivo trattamento. Dall'altro richiede un cambio di approccio alla tutela, orientandola alla specifica vulnerabilità del dato e alle relative tipologie di rischio. E' esattamente in questo senso che la PbD, con la sua capacità di essere un sistema di tutela robusto ma flessibile, è in grado di rispondere alle esigenze di indeterminatezza e ampliamento semantico dei big data.

Un altro elemento che sembra segnare un rafforzamento della tutela dei dati personali è costituito dall'introduzione del Responsabile per la protezione dei dati personali, una delle novità principali del Regolamento. L'Articolo 37 prevede l'obbligatorietà della sua nomina quando: a) il trattamento è effettuato da un'autorità pubblica o da un organismo pubblico, eccettuate le autorità giurisdizionali quando esercitano le loro funzioni giurisdizionali; b) le attività principali del titolare del trattamento o del responsabile del trattamento consistono in trattamenti che, per loro natura, ambito di applicazione e/o finalità, richiedono il monitoraggio regolare e sistematico degli interessati su larga scala; oppure c) le attività principali del titolare del trattamento o del responsabile del trattamento consistono nel trattamento, su larga scala, di categorie particolari di dati personali di cui all'Art. 9 o di dati relativi a condanne penali e a reati di cui all'articolo 10.

Con riferimento alla prima fattispecie, ai sensi dell'Articolo 3(26) del D.Lgs. 12.4.2006, n. 163, l'organismo di diritto pubblico è "qualsiasi ente, anche in forma societaria, istituito per soddisfare specificatamente esigenze di interesse generale, aventi carattere non industriale o commerciale, che sia dotato di personalità giuridica e la cui attività sia finanziata in modo maggioritario dallo Stato, dagli enti pubblici territoriali o da altri organismi di diritto pubblico; oppure la cui gestione sia soggetta al controllo di questi ultimi, o da ultimo, il cui organo d'amministrazione, di direzione o di vigilanza sia costituito da membri dei quali più della metà sia designata dallo Stato, dagli enti pubblici territoriali o da altri organismi di diritto pubblico". Per la qualificazione di organismo di diritto pubblico è necessario che i tre requisiti sopra enucleati sussistano in maniera cumulativa, come

[441] van der Sloot, *Do Data Protection Rules Protect the Individual and Should They? As Assessment of the Proposed General Data Protection Regulation*, International Data Privacy Law 4(4), 2014, 307, 323.

[442] Il riferimento è ai 7 Foundational Principles pubblicati nel 2005 da Ann Cavoukian, Information and Privacy Commisioner canadese, dopo anni di studio e progetti. Il Principio 7 "Respect for User Privacy", considera predominante l'interesse dell'individuo 'by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options'. Cfr. Cavoukian, A., *Privacy by Design: The 7 Foundational Principles*, cit.

ribadito sia dalla giurisprudenza europea, sia da quella nazionale[443].Posto in ogni caso che un organismo pubblico è un soggetto la cui esistenza e il cui funzionamento sono disciplinati da norme di diritto pubblico, la determinazione dei confini entro i quali far rientrare tali enti è lasciata all'intervento dei singoli Stati membri. Tale rilievo lascia tuttavia aperta la questione sulla natura di organismo pubblico di tutte quelle persone giuridiche che, pur essendo disciplinate dal diritto privato, svolgono funzioni di pubblica utilità (trasporto, organismi di disciplina professionale etc.).

Con riguardo alle lettere b) e c), è opportuno definire cosa si intenda con le espressioni "attività principali" e "su larga scala". Il Gruppo di lavoro articolo 29, rinviando al Considerando 97 del GDPR, definisce le attività principali del titolare come quelle che "riguardano le sue attività primarie ed esulano dal trattamento dei dati personali come attività accessoria",[444] ossia le operazioni essenziali al raggiungimento degli obiettivi perseguiti dal titolare o dal responsabile. L'espressione "su larga scala" invece non è esplicitamente definita dal GDPR, che rimanda tuttavia al Considerando 91 per l'enunciazione di alcuni criteri che possano coadiuvare il titolare nella determinazione dei casi in cui un trattamento è effettuato su larga scala. In particolare, a tal proposito, devono essere tenuti in considerazione: (i) il numero di interessati dal trattamento espressi in percentuale alla popolazione di riferimento; (ii) il volume dei dati trattati; (iii) la durata e la persistenza del trattamento; (iv) la portata geografica del trattamento. Ai fini dell'obbligatorietà della nomina, la lettera b) prevede che le attività principali riguardino il "monitoraggio sistematico e il tracciamento". Rientrano in questa categoria tutti quei trattamenti che prevedono la profilazione o il controllo costante dei comportamenti dell'utente *(behavioural advertising*, servizi connessi alla domotica e legati all'Internet of Things, programmi di fidelizzazione etc.). A titolo esemplificativo, si pensi alla società di sicurezza che si occupa della videosorveglianza di una serie di centri commerciali. La sorveglianza è l'attività principale, ma la stessa è strettamente legata al trattamento di dati personali. Inoltre, tale trattamento comporta un monitoraggio sistematico degli interessati su larga scala. Per questi motivi, la società dovrà nominare un *Data Protection Officer*. Con specifico riferimento alla lettera c), la nomina del DPO è obbligatoria ogniqualvolta l'attività principale dell'impresa sia il trattamento su larga scala i dati sensibili (Articolo 9) o giudiziari (Articolo 10). È opportuno ricordare, che ai sensi del nuovo Regolamento, anche i dati biometrici si configurano come dati sensibili.[445] Un caso riconducibile alla fattispecie indicata è quello della clinica privata che, in quanto titolare del trattamento di dati sanitari dei propri pazienti sarà tenuta alla nomina del Responsabile per la protezione dei dati personali. I dati biometrici sono oltretutto legati allo sviluppo dell'IoT. Amazon Echo, lo speaker intelligente sviluppato da Amazon per rispondere alle domande degli utenti dialogando con la voce, utilizza il machine learning e l'intelligenza artificiale attraverso il riconoscimento vocale.

[443] Il termine è apparso per la prima volta nella Direttiva Europea n. 89/440/CEE del 21.7.1989, con lo specifico intento di ricondurre nel raggio di applicazione della disciplina europea in tema di appalti quei soggetti operanti negli Stati membri.

[444] Cfr. Gruppo di lavoro Articolo 29, *Guidelines on Data ProtectionOfficers ('DPOs')*, 13 dicembre 2016, come rivisto e adottato in data 5 aprile 2017.

[445]Cfr. Articolo 4 n.14.

L'Articolo 37 n.4 del GDPR prevede poi che "nei casi diversi da quelli di cui al paragrafo 1, il titolare e del trattamento, il responsabile del trattamento o le associazioni e gli altri organismi rappresentanti le categorie di titolari del trattamento o di responsabili del trattamento possono o, se previsto dal diritto dell'Unione o degli Stati membri, devono designare un responsabile della protezione dei dati".Nel disposto normativo si evince la chiara volontà del legislatore di raccomandare l'adozione del DPO anche nei casi che esulino dalla sua obbligatorietà[446]. Il *Data Protection Officer* è una figura che opera innanzitutto nell' interesse del Titolare ed è da questi designato in funzione delle qualità professionali, in particolare della conoscenza specialistica della normativa e delle prassi in materia di protezione dei dati personali e della capacità di assolvere i compiti a cui è preposto. Il DPO non è dunque una figura di controllo e vigilanza "contro" il Titolare, ma che ha il dovere di operare nel suo interesse, inteso nel senso più ampio di coadiuvare le esigenze di profitto dell'impresa con il rispetto della normativa in materia di protezione dei dati personali. Deve essere ricordato che quando il Titolare nomina un DPO, o perché ricorrono i casi di nomina obbligatoria o per scelta propria, è responsabile di tale atto. La mancanza di competenze relative alle attività che il DPO dovrà ricoprire costituisce un elemento del quale il titolare potrà essere chiamato a rispondere sia rispetto all'autorità di controllo che, in caso di danno, di fronte a quella giudiziaria.

Come previsto dall'Articolo 39, il *Data Protection Officer* è incaricato almeno dei seguenti compiti:

(i) informare e fornire consulenza al titolare del trattamento o al responsabile del trattamento nonché ai dipendenti che eseguono il trattamento perlopiù in merito agli obblighi derivanti GDPR;
(ii) sorvegliare l'osservanza delle norme relative nonché delle politiche del titolare o del responsabile in materia di protezione dei dati personali;
(iii) fornire, se richiesto, un parere in merito alla valutazione d'impatto sulla protezione dei dati e sorvegliarne lo svolgimento ai sensi dell'Articolo 35;
(iv) cooperare con l'autorità di controllo; e
(v) fungere da punto di contatto per l'autorità di controllo per questioni connesse al trattamento, tra cui la consultazione preventiva di cui all'Articolo 36, ed effettuare, se del caso, consultazioni relativamente a qualunque altra questione.

Nell'esercizio delle sue funzioni, il DPO deve godere di autonomia e indipendenza, non può ricevere alcuna istruzione, né essere rimosso o penalizzato dal titolare o dal responsabile per l'adempimento dei propri compiti e deve essere necessariamente collocato in una posizione apicale dell'organigramma aziendale.[447] Con riferimento al regime di responsabilità, autonomia e

[446] Infatti, se fino agli anni ottanta la privacy era intesa come mero diritto alla riservatezza, oggi tale concetto si è evoluto in quello di «autodeterminazione informativa» , ossia di tutela del contenuto informazionale di ciascun interessato. Cfr. Califano L., *Privacy: affermazione e pratica di un diritto fondamentale*, Editoriale Scientifica, Napoli, 2016, 69. In tal senso, la presenza sempre più diffusa di professionisti della protezione dei dati personali in grado di relazionarsi con le autorità di controllo e di collaborare con il titolare al fine di garantire modalità semplificate per l'esercizio dei diritti da parte degli interessati si colloca in piena coerenza con il dettato normativo appena analizzato.

[447] Cfr. Articolo 38 n.3.

indipendenza, tale figura potrebbe essere assimilata a quella del responsabile del servizio di prevenzione e protezione (RSPP) disciplinata nell'ordinamento giuridico italiano dal D.Lgs 81/2008.[448] Non è inverosimile ipotizzare che la giurisprudenza dei prossimi anni in materia di *Data Protection Officer* ripercorra gli stessi orientamenti enunciati con riferimento al RSPP.

Oltre alla definizione di questi obblighi a carico del *Data Processor Officer,* il GRDP espressamente richiede una maggiore attenzione ai livelli di sicurezza del trattamento e, per questo, un'adeguata analisi del rischi. Benché una preventiva analisi dei rischi fosse comunque implicita nella Direttiva in quanto funzionale all'adozione delle misure di sicurezza richieste dalla legge, con gli Articoli 5 f) e 32 del GRDP si richiede ai Titolari del trattamento una valutazione dei rischi e la conseguente adozione di Misure Tecniche e Organizzative Adeguate (e non più "minime" come nella precedente disciplina della Direttiva madre). Inoltre, per processi "pericolosi" per i dati personali, sempre secondo il nuovo Regolamento(Articoli 35 e 36)sarà obbligatoria una valutazione d'impatto sulla protezione dei dati (*Data Protection Impact Assessment – DPIA,* anche nota come *Privacy Impact Assessment – PIA).*[449]

LA DPIA si configura come una chiara estensione del *principio di accountability* poc'anzi enunciato[450]. La DPIA deve essere effettuata ogni qualvolta "un trattamento possa presentare un rischio elevato per i diritti e le libertà delle persone fisiche". Il Gruppo di lavoro Articolo 29 ha recentemente elaborato una serie di criteri[451] che mirano a coadiuvare il titolare del trattamento nella fase di individuazione della necessità di svolgere tale adempimento. In particolare, essi sono:

(i) "La valutazione o l'assegnazione di un punteggio", incluse la profilazione e i trattamenti con finalità predittiva, in particolare sugli aspetti concernenti le performance lavorative dell'interessato, la situazione economica, la salute, le preferenze o interessi personali, l'affidabilità o il comportamento, la posizione o gli spostamenti.

(ii) "Decisioni automatizzate con effetti giuridici o ugualmente significativi": trattamento che mira a determinare decisioni su interessati e che produce effetti giuridici riguardanti la persona fisica.

(iii) "Monitoraggio sistematico": trattamento posto in essere al fine di osservare, monitorare o controllare gli interessati. Sono inclusi all'interno di tale parametro anche i dati raccolti attraverso un controllo sistematico di una zona accessibile al pubblico. Questo tipo di monitoraggio si configura come un criterio in quanto gli interessati potrebbero non essere a conoscenza di chi sta raccogliendo i loro dati e di come essi saranno utilizzati. Inoltre,

[448] Cfr. Decreto-legge 2008, n. 81, in materia di *Testo unico sulla salute e sicurezza sul lavoro*.

[449] Gruppo di lavoro articolo 29, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, wp248rev.01, 4 aprile 2017, come rivisto ed adottato in data 4 ottobre 2017.

[450] Trovato C. A, *La valutazione d'impatto privacy. Obblighi per le imprese e casi concreti*, in https://www.dpoinnovation.it/2017/10/05/524/.

[451] Gruppo di lavoro articolo 29, nota 51.

potrebbe essere impossibile per le persone evitare di essere oggetto di tale trattamento in spazi pubblici abituali (o accessibili al pubblico).

(iv) "Dati Sensibili": tale criterio include le categorie particolari di dati ai sensi dell'Art. 9 del GDPR nonché i dati personali relativi alle condanne penali o ai reati.

(v) "I dati trattati su larga scala".

(vi) "Set di dati che sono stati abbinati o combinati", ad esempio provenienti da due o più operazioni di trattamento effettuate per finalità diverse e/o da differenti titolari in modo tale da superare le ragionevoli aspettative dell'interessato.

(vii) "I dati riferiti a soggetti vulnerabili": il trattamento di questo tipo di dati può richiedere una DPIA a causa dello squilibrio di potere tra il titolare e l'interessato, ossia quest'ultimo può non avere gli strumenti sostanziali per consentire, od opporsi, al trattamento dei propri dati.

(viii) "L'utilizzo innovativo o l'applicazione di soluzioni tecnologiche o organizzative", quale ad esempio la combinazione dell'uso di impronte digitali e riconoscimento del volto per un migliore controllo sull'accesso fisico dei soggetti.

(ix) Quando il trattamento, per sua stessa natura, "impedisce agli interessati di esercitare un diritto o utilizzare un servizio o usufruire dei benefici di un contratto"[452].

Nelle citate Linee guida, il WP 29 afferma che, nella maggior parte dei casi, il titolare è tenuto a svolgere una DPIA quando una singola operazione di trattamento o più operazioni simili tra loro, riuniscano almeno due dei suddetti criteri. Tale assunto non deve essere inteso come assoluto. Vi possono essere infatti circostanze in cui "un titolare può considerare che un trattamento che racchiuda uno solo dei criteri necessiti comunque di una valutazione d'impatto sulla protezione dei dati personali"[453]. È inoltre opportuno procedere alla verifica sul fatto che il trattamento analizzato non rientri nella lista che l'autorità di controllo ha la facoltà di redigere ex Articolo. 35(5) contenente le operazioni che non necessitano di DPIA. Il procedimento è sempre attivato dal titolare del trattamento che può agire autonomamente o su consiglio del DPO(Articolo 35.2) il quale, se nominato, deve obbligatoriamente essere coinvolto nello svolgimento della valutazione d'impatto.

In seguito alla verifica sulla necessità di effettuare la DPIA, nonché all'avvio del procedimento, il titolare dovrà procedere all'individuazione del rischio privacy inteso nella duplice natura di "sicurezza dei dati" (probabilità e gravità di un'eventuale violazione di dati) e "rischio per i diritti e le libertà delle persone fisiche".[454] È proprio in questa duplice natura che trova la propria dimensione compiuta il diritto alla protezione dei dati personali. Forse per un eccesso di affidamento verso un

[452] Rispetto alla precedente versione delle Linee guida sulla DPIA adottata il 4 aprile 2017, è stato rimosso il criterio relativo al "trasferimento dei dati al di fuori dei confini dell'Unione Europea". Con tutta probabilità l'Article 29 Working Party ha fatto seguito alle indicazioni della Commissione sul Privacy Shield pubblicate il 18 ottobre 2017. European Commission - Press release, EU-U.S. Privacy Shield: First review shows it works but implementation can be improved, in http://europa.eu/rapid/press-release_IP-17-3966_en.htm.

[453] Cfr. Gruppo di lavoro articolo 29, nota 51,p.11.

[454] Per ulteriori spunti di riflessione in merito a nuove forme di tutela e valorizzazione legate ai dati personali, Nunziante E., *Big Data. Come proteggerli e come proteggerci. Profili di tutela tra proprietà intellettuale e protezione dei dati personali*, in Law and Media WorkingPaper Series, no. 06/2017.

approccio informatico, nella totalità dei documenti ufficiali in materia si focalizza l'attenzione quasi esclusivamente sull'analisi dei rischi incombenti sui dati, anziché su quelli relativi ai diritti e alle libertà delle persone fisiche, dimenticando che il diritto alla protezione dei dati personali è sempre stato inteso dal legislatore europeo come diritto fondamentale dell'individuo.[455] la cui tutela non può evidentemente esaurirsi nell'applicazione di misure di sicurezza tecniche e organizzative. La fase successiva all'individuazione dei rischi rappresenta il cuore pulsante della DPIA, ossia la predisposizione delle misure atte a mitigare le possibili minacce individuate. Il titolare sarà pertanto tenuto ad applicare una serie di misure non di carattere esclusivamente tecnico e organizzativo al fine di evitare un probabile *data breach,* ma anche di natura economica e talvolta etica con lo scopo di ridurre l'impatto che il trattamento ha sulle libertà e sui diritti fondamentali delle persone fisiche. Come previsto dall'Articolo 36 del GDPR, "il titolare del trattamento prima di procedere al trattamento, consulta l'autorità di controllo qualora la valutazione d'impatto sulla protezione dei dati a norma dell'articolo 35 indichi che il trattamento presenterebbe un rischio elevato in assenza di misure adottate dal titolare del trattamento per attenuare il rischio". Dunque, qualora nonostante l'applicazione delle misure individuate, sussista un rischio residuo di alto livello, il titolare sarà tenuto a procedere alla consultazione dell'autorità di controllo. Si pensi a tal proposito ai trattamenti di dati sanitari effettuati su larga scala o alla determinazione dei profili creditizi degli interessati attraverso sistemi automatizzati. La DPIA comporterà per il titolare del trattamento un obbligo di adozione di misure non soltanto di natura tecnica e organizzativa, ma dovranno essere ponderati gli effetti di natura economica ed etica in relazione alla scelta di certe misure. A titolo esemplificativo, di pari passo con la diffusione dei dispositivi che, installati sulle autovetture, registrano la posizione e lo stile di guida del conducente, le compagnia assicurative potrebbero arrivare a definire il profilo modello del "guidatore accorto". I soggetti il cui stile di guida si avvicinerà a tale modello potrà beneficiare di sconti significativi. Questo avrà un effetto non solo nei confronti di chi avrà aderito a questa iniziativa, ma progressivamente potrà comportare degli effetti discriminatori anche sulla restante parte della popolazione i quali potrebbero trovarsi a pagare un prezzo più alto perché non si uniformano al modello o, semplicemente hanno deciso di non installare il sensore per esigenze di tutela della propria privacy[456].

5.5 I nuovi diritti dell'interessato: il diritto all'oblio e la data portability

Un'altra forma o livello di garanzia pertiene la permanenza delle informazioni riconducibili ai nostri dati personali e il diritto alla loro cancellazione, anche noto come diritto all'oblio. Dinnanzi all'implacabile memoria collettiva di internet,[457] la giurisprudenza ha elaborato un nuovo diritto dell'individuo ad essere "dimenticato" dalle banche dati, dai mezzi di informazione, o dai motori di ricerca. Il Nuovo Regolamento attua il riconoscimento su base legislativa del diritto all'oblio all'Articolo 17. In particolare, l'Interessato ha diritto di chiedere che siano cancellati o deindicizzati

[455]Cfr. Articoli 7 e 8, Carta dei diritti fondamentali dell'Unione Europea (2000/C 364/01) , Parlamento europeo, dal Consiglio e dalla Commissione.

[456] Morozov E., *L'ingenuità della rete*, Codice Edizioni, Torino, 2011.

[457] Rodotà S., *Il diritto di avere diritti*, Editori Laterza, 2012, 404.

(eliminati dai motori di ricerca) e non più sottoposti a trattamento i suoi dati personali (i) che non siano più necessari per le finalità per le quali sono stati raccolti; (ii) quando abbia ritirato il consenso o si sia opposto al trattamento o il trattamento dei dati personali non sia altrimenti conforme al Regolamento. La stessa norma menziona anche alcuni casi in cui il diritto all'oblio non sussiste, ad esempio, i casi in cui il trattamento sia necessario per l'esercizio del diritto alla libertà di espressione e di informazione o per l'adempimento di un obbligo legale. La cancellazione potrà essere pretesa "senza ingiustificato ritardo"

Un'altra delle principali novità introdotte dal nuovo Regolamento è il diritto alla portabilità dei dati, di cui si è già parlato nel precedente paragrafo 4.4.1. Come si ricorderà, ai sensi dell'Articolo 20, sarà consentito a ciascun cittadino europeo di ottenere gratuitamente e in formato strutturato, di uso comune e leggibile da qualsiasi dispositivo elettronico, i propri dati personali e di richiederne il trasferimento, gratuito e senza impedimenti, ad un altro titolare del trattamento. Il diritto alla portabilità, consentirà quindi a ciascun cittadino europeo non solo di accedere ai propri dati personali, ma di trasferirli per scopi ulteriori, e potenzialmente diversi, rispetto a quelli per i quali sono stati inizialmente raccolti, da un fornitore di servizi ad un altro. L'introduzione di questo diritto ha una portata significativa in quanto traccia una via chiara per la risoluzione delle problematiche connesse al passaggio dei dati tra diversi device, obbligando i fornitori di servizi a rinunciare alla vischiosità commerciale connessa con il c.d. Technology Lock-in.

Il diritto alla portabilità dei dati si pensa che porterà all'apertura di scenari competitivi del tutto nuovi in una molteplicità di settori economici.

L'istituto è stato recentemente oggetto di linee guida europee predisposte e aggiornate da ultimo il 5 aprile 2017 da parte del WP29. L'interpretazione del perimetro dei dati "portabili" secondo gli esperti del WP29 ricomprenderebbe un insieme molto ampio di essi, in cui rientrerebbero anche quelli relativi all'osservazione dell'attività dell'utente nell'interazione con il titolare. Tra i dati portabili vi sarebbero, ad esempio, i nomi delle canzoni ascoltate, dei film visti, dei luoghi registrati per effetto di sistemi di geolocalizzazione. Gli unici dati non trasferibili liberamente sarebbero quelli generati direttamente dal titolare del servizio (come ad esempio il profilo di solvibilità del cliente per un istituto di credito) ovvero quelli raccolti in ottemperanza ad un obbligo di legge. Secondo la Commissione Europea, al contrario, la norma andrebbe interpretata in senso restrittivo, considerando come trasferibili solo i dati forniti volontariamente dall'interessato, ossia i dati anagrafici, le foto, i documenti, i commenti ed altre informazioni intenzionalmente caricate, escludendo dall'applicabilità del diritto alla portabilità tutti i dati che discendono dalla mera "osservazione" del comportamento dell'utente. A giudizio della Commissione, l'interpretazione fornita dal WP29 sarebbe dunque eccessivamente ampia, nonchè contraria allo spirito del legislatore comunitario che avrebbe formulato la norma pensando solo ai dati forniti volontariamente dall'utente e non a tutti gli altri "metadati". In particolare, la Commissione paventa il rischio che l'estensione dell'obbligo della portabilità anche ai dati di traffico rischi di gravare soprattutto sulle aziende di dimensioni più piccole, esponendole, in caso di violazione, a sanzioni esorbitanti (fino a 20 milioni di euro o al 4% del fatturato).

Sullo sfondo si gioca la partita del mercato della concorrenza ed anche quella del confronto tra autorità di controllo. La multa da 110 milioni di euro comminata a Facebook per l'illecita integrazione dei dati personali degli utenti del social network con quelli di Whatsapp, lascia pensare ad un terreno di confronto basato sulla titolarità dei dati e sulla loro libera circolazione, con importanti ripercussioni sulla competitività delle imprese nel mercato interno. L'esercizio della portabilità dei dati aumenterà significativamente la possibilità per gli utenti di passare facilmente da un fornitore di un servizio ad un altro, portando in dote, in modo totalmente gratuito, tutto il proprio patrimonio storico informativo relativo ai servizi acquistati. Si tratta di informazioni che hanno un valore rilevante, e la loro portabilità potrebbe ridurre di fatto il vantaggio accumulato negli anni da alcuni operatori i quali raccolgono e trattano da tempo tali dati. Tra i settori più interessati dal diritto alla portabilità, il primo sarà senz'altro quello bancario che, con il recepimento, previsto per il gennaio del 2018, della direttiva sui servizi di pagamento (PSD2), si aprirà a nuove linee di attività e alla concorrenza di nuovi operatori. L'obbligo per le banche, previsto dalla direttiva sui servizi di pagamento, di rendere disponibili i dati dei propri clienti per servizi di pagamento terzi, consentirà a nuovi operatori, con il consenso dell'interessato, di concorrere con gli istituti di credito su nuove linee di attività, offrendo servizi aggiuntivi derivati dall' aggregazione di dati proveniente da settori differenti (ad es. da servizi di risparmio, servizi assicurativi, consumi elettrici, consumi di carburante).

Al contempo, la possibilità per l'utente di trasferire agilmente tutti i propri dati ad un diverso fornitore di servizi riporterà al centro della scena l'interessato, imponendo un'accelerazione nel processo di digitalizzazione delle aziende. Con il consenso preventivo dell'utente, una significativa mole di informazioni ora depositate presso diversi titolari del trattamento, potrebbe infatti essere trasferita, consentendo ulteriori concentrazioni di dati in capo ai fornitori di servizi esistenti (la norma non impone la cancellazione dei dati in capo al soggetto trasferente, ma solo il loro trasferimento in relazione alle richieste dell'interessato) ovvero l'ingresso nel mercato di nuovi player, favoriti proprio dalla portabilità dei dati.

L'introduzione a livello europeo del diritto alla portabilità dei dati sembra quindi rappresentare una vera svolta in molti settori. Le autorità di controllo si troveranno presumibilmente a cercare di definirne la portata applicativa solo successivamente all'applicazione del regolamento. Il WP29, nelle citate linee guida, si è espresso nel senso di un'applicazione non generalizzata del diritto, in ragione della necessità di tener conto delle norme di ciascun settore ove il diritto dell'Ue o del singolo Stato membro preveda una qualche altra forma di portabilità dei dati, con la necessità di valutare caso per caso le condizioni specifiche di applicabilità della norma.

I rischi per l'interessato d'altra parte non mancano. Essi si concentrano in via principale sull'effettiva sicurezza del processo di trasferimento dei dati, ma anche sulla sempre più frequente coartazione del consenso dovuta a pratiche commerciali aggressive. Infatti, "nel rapporto competitivo e osmotico tra i diversi livelli di gestione della rete (infrastruttura, accesso, servizi), un ruolo di equilibrio essenziale è occupato dalla necessità di garantire una tutela flessibile ed articolata dei dati che circolano in un "mercato" senza confini geografici e di elaborazione, sia in senso protettivo e qualitativo, che ancora in termini concorrenziali"

Sul fronte della sicurezza, si segnala come il GDPR preveda che l'interoperabilità dei dati tra diversi fornitori di servizi sia agevolata attraverso la scelta da parte del titolare del trattamento di formati standard per il salvataggio dei dati, ovvero l'implementazione di sistemi che, con il consenso dell'interessato e salvo l' eccessiva complessità del procedimento, consentano il trasferimento diretto dei dati tra fornitori di servizi senza dover essere previamente acquisiti dall'interessato stesso. I dati trasferiti dovranno essere completi di tutte le informazioni semantiche necessarie affinchè siano utilizzabili anche nel nuovo sistema (tag, metatag, etc). Il loro trasferimento potrà avvenire sfruttando API *(Application Programming Interface)* dedicate.

Secondo il WP29 l'utente dovrà presumibilmente esercitare il controllo delle proprie informazioni attraverso apposite *dashboard*, pannelli di controllo simili a quelli attualmente disponibili su Google o Facebook, che consentiranno all'interessato di interagire con il titolare dei dati ed autorizzare la loro condivisione con soggetti terzi. Tra i soggetti autorizzati a ricevere i dati potranno esserci, oltre a server controllati direttamente dall'interessato, anche i concorrenti o i fornitori di servizi alternativi del titolare del trattamento di provenienza. La gestione di trasferimenti di dati anche molto imponenti richiederà il previo accertamento dell'identità dell'interessato, nonchè politiche di contenimento dei rischi finalizzate ad evitare che i dati, comprensivi anche di informazioni relative ai soggetti terzi (ad esempio quelle riferite al destinatario di un'email o del ricevente di pagamenti) possano essere trasferiti a chiunque non garantisca adeguate tutele.

Come è facile comprendere, la dialettica in relazione ai profili di applicazione di questo nuovo diritto non è destinata ad arrestarsi. Nei prossimi anni si apriranno scenari del tutto nuovi per gli operatori del mercato, anche coloro che sino ad oggi hanno scarsamente valorizzato le informazioni dei propri utenti. Proprio in considerazione della portata, sino ad oggi apparentemente sottovalutata, degli effetti applicativi di questo diritto, è in corso un dibattito a distanza tra la Commissione Europea, rappresentata dal capo di gabinetto del commissario alla giustizia Věra Jourová ed il gruppo di esperti europei della privacy riuniti nel cosiddetto Article 29 Working Party. Le decisioni interpretative che verranno assunte in merito alla quantità e qualità dei dati "portabili" avranno effetti importanti in relazione alla conservazione sul mercato di alcuni servizi, e allo sviluppo di nuovi, oltre a conferire a tutti i consumatori una nuova consapevolezza sul valore dei dati che gli stessi generano attraverso l'interazione con le piattaforme online.

5.6 La condivisione dei dati e gli accordi di data sharing nella c.d. "Membership economy"

La condivisione dei dati tra gli individui secondo modelli giuridici fondati sulla nozione di "piattaforma" (nelle note forme di SoftwareAsAService, PlatformAsAService e InfrastructureAsAService), pone all'interprete il difficile compito di trovare un equilibrio tra i diversi diritti ed obblighi riconosciuti dal GDPR ai soggetti che fanno parte, dal lato attivo e passivo, del trattamento dei dati personali[458].

[458] Si pensi ad esempio alle difficoltà che si sono poste in applicazione della Direttiva 2009/136, attuata in Italia con il decreto legislativo 28 maggio 2012, n. 69, in relazione ai cookies delle c.d. "terze parti" e alla difficoltà di disciplinare

Uno degli aspetti più controversi è quello di coordinare le finalità privacy con l'oggetto del contratto che costituisce la base giuridica del trattamento e del consenso dell'utente-interessato.

Da un lato, il legislatore europeo ha riaffermato nel nuovo regolamento il diritto fondamentale dell'individuo alla tutela della propria vita privata[459], ponendo il consenso come uno degli elementi cardine (ma non l'unico) del sistema. Il consenso dell'interessato è corredato e arricchito da diritti conoscitivi (informativa e accesso) e di controllo (limitazione al trattamento, diritto di revoca, opposizione, rettifica, cancellazione e portabilità dei dati).

Dall'altro lato, il GDPR riafferma in capo ai titolari e ai responsabili del trattamento, oltre ai menzionati obblighi in tema di sicurezza e di tutela dei diritti dell'interessato, nuove possibilità di utilizzo dei dati che derivano da un rapporto giuridico tra questi ultimi e l'interessato. Il titolare può trattare i dati anche al di fuori di un preventivo consenso dell'interessato, purchè il trattamento sia effettuato nell'ambito della tutela di propri interessi legittimi e questi non prevalgano sugli interessi e le libertà dell'interessato (c.47), ovvero nel caso in cui l'utilizzo dei dati per finalità diverse da quelle originariamente previste possa considerarsi compatibile con le finalità originarie (secondo un giudizio di compatibilità effettuato dal titolare sulla scorta di una serie definita di condizioni art. 6.(4)).

Il rapporto tra consenso, prestazione contrattuale e finalità rimane quindi con il nuovo regolamento il paradigma sul quale misurare la liceità dei trattamenti effettuati dal titolare, ma si apre all'utilizzo dei dati anche per altre finalità, ove sia possibile individuare una base giuridica adeguata all'ulteriore trattamento che non ponga rischi per la tutela dei diritti dell'interessato.

In questo modo la normativa europea sembra cercare di conciliare i diritti dell'interessato con quelli del titolare del trattamento, che aspira a trattenere (e quindi a trattare) il maggior numero di dati personali possibile. Stiamo in effetti assistendo all'evoluzione del mercato digitale verso quella che è stata definita la c.d. "*membership economy*"[460], caratterizzata dalla graduale trasformazione del rapporto negoziale tra consumatore e professionista da un rapporto di natura prevalentemente puntuale (un contratto ad esecuzione immediata tipico di acquisti di beni o servizi online), ad un rapporto ad esecuzione continuata, di durata indeterminata (anche grazie alla tecnica dei rinnovi e dei pagamenti automatici.

Nell'ambito della c.d. "*forever transaction*" il rapporto tra titolare ed interessato nasce con la registrazione del consumatore ad un servizio o ad una piattaforma il cui oggetto non è normalmente definito in modo specifico: ciascun titolare aspira ad essere il gateway, il rivenditore di altri prodotti e servizi secondo una causa contrattuale mista ed un oggetto che spesso non è chiaramente determinato o indeterminabile (con tutti i derivanti rischi di nullità del rapporto per assenza di un requisito fondamentale o a causa della violazione del diritto di informazione del consumatore previsto dall'art. 5 del Codice del Consumo).

[459] Si tratta del considerando 1 e 4 del GDPR che richiamano gli artt. 7 e 8 della Carta dei diritti fondamentali dell'Unione Europea (2000/C 364/01).

[460] Kellman Baxter R., *The membership Economy*, Nobles, 2015

Sotto il profilo del diritto alla riservatezza, la registrazione quale "*member*" ad una piattaforma (come Amazon, Google o anche solo il sito di Trenitalia o del Corriere della Sera) implica normalmente per l'utente la necessità di fornire un insieme significativo di informazioni personali, anagrafici e di pagamento molto superiore a quelli che sarebbero stati necessari per l'esecuzione della prestazione tipica originaria, ovvero leggere un giornale o salire su un treno. A questi dati si aggiungono spesso i dati di movimento o di informazioni relative alla salute dell'interessato raccolti attraverso forme di tracciamento sul web (dai *cookies* ai *finger prints*), dispositivi mobili di proprietà dell'utente (cellulari o tablet) ovvero, più di recente, grazie a device mobili ad attivazione vocale forniti dalla piattaforma stessa (si pensi a Google Home Assistant e Alexa).

Questi strumenti consentono alla piattaforma di aggregare una mole significativa di dati (big data appunto) il cui trattamento (inclusa la conservazione) è riconducibile ad almeno due evidenti, e complementari finalità: (i) la prima, quella di fornire all'interessato il servizio inizialmente richiesto nell'alveo dell'esecuzione di un obbligo contrattuale (ii) la seconda, quella di beneficiare di un servizio migliore ed integrato grazie al monitoraggio continuo delle sue scelte di consumo, della sua posizione e stili di vita al fine di usufruire di servizi terzi. E' questo ad esempio il caso della compagnia area che, associata alla vendita di un volo, propone servizi ulteriori di natura alberghiera o di trasporto: i dati dell'interessato sono comunicati dal titolare della piattaforma ai diversi fornitori di servizi in ragione di una richiesta espressa dell'interessato. Ciascun fornitore, in questo caso, acquisisce e tratta i dati in qualità di titolare autonomo del trattamento per il servizio richiesto dall'utente, ma, allo stesso tempo, potrebbe essere tentato dall'idea di conservare informazioni relative al servizio prestato e all'interazione con l'utente degli altri partner commerciali, inclusi soprattutto i partner tecnologici che hanno sviluppato e amministrano l'eventuale *device* mobile grazie al quale l'interessato comunica le proprie informazioni.

Ove le informazioni dell'interessato siano acquisite contestualmente da più titolari, ma per finalità solo parzialmente coincidenti si potrebbe ritenere verificata l'ipotesi di contitolarità "asimmetrica" prevista dall'art. 26 del GDPR[461]. Si tratta di un caso di raccolta congiunta dei dati personali che dovrebbe essere chiaramente manifestata all'interessato al fine di consentirgli di esprimere un consenso libero ed informato sulle finalità ed i mezzi del trattamento. Il problema si pone per i dati che non sono necessari ad alcuna immediata finalità del titolare del trattamento, ma che se aggregati dal titolare del trattamento potrebbero consentire di sviluppare nuovi prodotti o servizi. Ad esempio, nel caso del citato vettore aereo, la società che opera il servizio e che ha richiesto la registrazione dell'utente potrebbe ipoteticamente essere interessata a monitorare il grado di utilizzo di servizi alberghieri in una data località dallo stesso servita ed acquisire direttamente o tramite accordi di data sharing, i dati relativi al servizio di ospitalità offerto al medesimo cliente, compresi i dati relativi alle preferenze alimentari, alle bevande consumate, alla temperatura della stanza, le persone con le quali l'interessato ha viaggiato. Questi dati potrebbero essere utilizzati per sviluppare nuovi servizi o integrare alcuni fornitori terzi.

[461] Sul punto Bolognini, *Il Regolamento privacy Europeo*, p. 136 dove ulteriori riferimenti. Sul punto si veda anche il WP29 nell'opinion del n. 1/2010 ed il provv. del Garante Italiano del 10.10.2013[2680653].

Si tratta di casi, invero, sempre più osservabili nelle dinamiche di mercato se si considerano applicativi legati alla domotica, come Google Home, o all'esperienza alberghiera (come il servizio Keyless offerto da Starwood) o alla sanità. Il loro sviluppo dipende dall'esistenza di un applicativo che raccoglie i dati degli utenti, li conserva e li condivide con altri fornitori di servizi. Il titolare della piattaforma, salvo le ipotesi del *mere conduit* tipico del Internet Service Provider, potrebbe decidere di operare come responsabile del trattamento (è il caso ad esempio dei "*developers license agreements*" di piattaforme come Google Play o Itunes) e conservare solo i dati legati alla propria finalità, cancellando ovvero anonimizzando la restante parte (ad esempio per i dati relativi al funzionamento del dispositivo) ovvero raccogliere e conservare il maggior numero possibile di dati, inclusi quelli che in apparenza non sono necessari per lo svolgimento del servizio[462]. Per raggiungere questo risultato, nel pieno rispetto dei paletti fissati dal GDPR, dovrebbe cercare di ampliare l'oggetto del contratto, in modo da ricomprendere attività allo stesso connesse e non direttamente collegate, ovvero "generalizzare" il servizio[463] ricomprendendo fattispecie inizialmente non considerate dall'utente/interessato(es. un istituto di credito potrebbe raccogliere i dati non solo per l'erogazione del mutuo richiesto, ma richiedere il consenso al trattamento dei dati personali nell'ambito della gestione di un rapporto di consulenza economico-finanziario per la gestione del risparmio dell'utente da effettuarsi attraverso una piattaforma informativa via web)

È evidente che queste situazioni, ancorché frequenti, si pongono al confine con il rispetto del requisito della libertà e della specificità del consenso[464], dettato dall'art. 4 e dal principio di minimizzazione del trattamento dettato dall'art. 5 del GDPR: le informazioni raccolte devono essere adeguate, pertinenti e limitate alla finalità originaria del trattamento.

Il principio di trasparenza delle informazioni che regola il rapporto tra il titolare e l'interessato, come previsto dal considerando 39 del GDPR, potrebbe ritenersi violato ove l'informativa riportasse generiche finalità di trattamento dei dati personali anche in relazione alla definizione del periodo di conservazione dei dati personali che dovrebbe essere limitato al minimo necessario[465].

Come messo in evidenza dal WP29 nell'opinion n. 3/2013 la natura del rapporto instaurato definisce infatti l'ampiezza, in termini di quantità e qualità dei dati, del trattamento consentito e la

[462] Un ulteriore effetto positivo di questa tendenza è legata alla progressiva riduzione dei costi connessi con la creazione di infrastrutture specifiche per ciascun servizio (si pensi ad esempio alla riduzione del numero delle app utilizzate dall'utente per i servizi di viaggio, normalmente divise tra quelle relative ai voli e quelle relative all'ospitalità, cosi come ai considerevoli vantaggi derivanti dall'utilizzo di sensori collocati in prodotti di consumo che possono essere utilizzati per finalità di sicurezza o controllo accessi).

[463] Prassi peraltro assai diffusa nei termini e condizioni utilizzati dalle app e servizi che circolano sul web.

[464]Il considerando 43 del GDPR precisa che il consenso si presume "non sia stato liberamente espresso se non è possibile esprimere un consenso separato a distinti trattamenti di dati personali, nonostante sia appropriato nel singolo caso".

[465] Per una disamina del tema si veda Hacker F., *Personal data, exploitative contracts, and algorithmic fairness: autonomous vehicles meet the Internet of Things*, 2017, International Data Privacy Law, 2017, Vol. 0, No. 0, 1; E Mik, *The Erosion of Autonomy in Online Consumer Transactions,* 2016, 8 Law, Innovation and Technology.

ragionevolezza dei dati che possono essere trattati. La sua definizione non dipenderà unicamente dai termini e condizioni accettati dall'utente al momento della registrazione, ma dovrà tener conto dell'oggetto del contratto, così come le dichiarazioni legali scambiate tra le parti, della prassi generalmente prevista nel contesto in questione e nella specifica relazione commerciale. In ragione del tipo di rapporto e di contesto cambiano le aspettative dell'interessato e con esse il grado di compatibilità di altri trattamenti con ulteriori finalità: più inaspettato o sorprendente sarà l'utilizzo ulteriore, più probabilmente il trattamento rischierà di essere considerato incompatibile[466].

Il medesimo giudizio di compatibilità e ragionevolezza richiesto all'interprete si può richiamare nei casi in cui i dati vengano ceduti, condivisi o scambiati successivamente alla loro raccolta autonoma da parte di diversi titolari. La condivisione in questi casi avviene senza l'immediato intervento dell'interessato e discende dalla stipula di accordi di data sharing che possono avvenire sulla base di diverse tipologie contrattuali. Si tratta di accordi tra diversi titolari del trattamento, come definito dal "*Data sharing code of practice*" emanato nel 2016 dall'ICO, relativi alla condivisione delle informazioni degli utenti derivanti da accordi aventi causa atipica: tra questi si possono annoverare accordi di comarketing, cessioni o affitto di azienda, fusioni, acquisizioni. Nel caso, poi, la condivisione riguardi dati sensibili o che possano avere un impatto sull'esercizio dei diritti o delle libertà dell'individuo, la condivisione sarà condizionata alla verifica dell'esistenza di un consenso espresso. In ogni caso, al di fuori dei casi in cui la condivisione dei dati avvenga a favore di soggetti pubblici o in dipendenza di previsioni normative, l'interessato, prima del trattamento delle proprie informazioni, dovrà essere informato delle finalità della condivisione dei dati tra i diversi contitolari, i soggetti e le circostanze nell'ambito delle quali potranno avere accesso ai dati, la tipologia dei dati condivisi (in termini di qualità, accuratezza, rilevanza, usabilità), i livelli di sicurezza nella conservazione dei dati e la durata di tale conservazione, le procedure riconosciute agli individui per esercitare i propri diritti. Allo stesso tempo l'accordo tra i titolari dovrà disciplinare l'efficacia e le conseguenze sui dati nel caso di risoluzione dell'accordo tra contitolari ed eventuali penali previste nel caso di violazione dei diritti dell'interessato o delle previsioni dell'accordo.

Si tratta di presidi, quelli appena elencati, che nel regolare la circolazione dei dati, postulano ampi margini interpretativi da risolvere all'insegna di profili etici connessi con il trattamento dei dati[467]. Questo perchè la logica della circolazione dei dati non poggia sulle stesse basi di quella della

[466] In questo senso depone l'ultima parte dell'art. 6 del GDPR che espressamente contempla "l'utilizzo di dati raccolti per finalità ulteriori rispetto la finalità per la quale i dati personali sono stati inizialmente raccolti" e la subordina alla verifica da parte dello stesso titolare del trattamento di una serie di elementi tra cui: a) l'esistenza di un nesso tra le finalità per cui i dati personali sono stati raccolti e le finalità dell'ulteriore trattamento previsto; b) il contesto in cui i dati personali sono stati raccolti, in particolare relativamente alla relazione tra l'interessato e il titolare del trattamento; c) la natura dei dati personali, specialmente se possano considerarsi dati sensibili; d) le possibili conseguenze dell'ulteriore trattamento previsto per gli interessati; e) l'esistenza di garanzie adeguate, che possono comprendere la cifratura o la pseudonimizzazione.

[467] European Economic and Social Committee, *The ethics of Big Data: Balancing economic benefits and ethical questions of Big Data in the EU policy context*, 2017.

circolazione di diritti reali, ma su quella ben più ampia della tutela delle libertà individuali[468]. E' quindi richiesto uno sforzo interpretativo significativo ed ulteriore da parte delle autorità di vigilanza, ma anche da parte di una nuova classe di giuristi capaci di bilanciare l'interesse ad un "rapporto contrattuale senza fine" con il diritto alla riservatezza dell'individuo visto in una dimensione oltre che normativa, "etica[469]. E' questo, infatti, l'approccio che una certa corrente dottrinale, formatasi anche nel nostro Paese, recentemente propone al fine di superare il carattere formalistico che ha in origine caratterizzato questa materia e la tutela dei diritti deli interessati.

Al posto della mera "Privacy by Design" questi autori immaginano di utilizzare un approccio etico, definito come "Ethical Design", per la definizione del paradigma di correttezza nel trattamento dei dati degli interessati. L'approccio etico alla privacy postula la definizione di infrastrutture e tipologie di trattamento tali da assicurare il rispetto di scelte e principi etici che garantiscono, a prescindere da qualsiasi valutazione economica e dalle previsioni contenute nel GDPR o nelle decisioni delle autorità nazionali, una tutela massima dei diritti e delle libertà degli individui. Questo si traduce nello sviluppo di applicativi che consentono all'utente maggiori e più pervasivi controlli sui dati trattati e sulle modalità con le quali le informazioni sono condivise con terze parti. I maggiori costi associati alle garanzie offerte poggerebbero, d'altra parte, sull'interesse generato nel mercato per questa tipologia di prodotti e servizi, con risvolti significativi sia in termine di valore percepito che di prezzo. Si tratta, a ben vedere, di una tendenza che ha già preso piede in altri contesti, tra cui in particolare quello del cibo, dove le certificazioni biologiche o biodinamiche hanno consentito agli attori del mercato che le hanno ottenute indubbi vantaggi competitivi e la possibilità di richiedere prezzi più alti. Resta, peraltro, da comprendere, in un sistema fondato su servizi apparentemente gratuiti, ma finanziati dalla pubblicità, quale potrà essere il successo e la diffusione effettiva di servizi che garantiscono il trattamento etico dei dati, ma che saranno necessariamente a pagamento.

5.7 Big Data e Regolamento Privacy

Alla luce dell'analisi condotta, un aspetto che apparentemente rimane al di fuori dalla panoramica sin qui svolta sulle nuove disposizioni introdotte dal GDPR è rappresentato dal trattamento dei big data.

Come definito dal WP29, i big data costituiscono un vasto insieme di operazioni di trattamento dati, un enorme patrimonio informativo che, tuttavia, a causa delle proprie peculiarità, può comportare dei rischi per la tutela dei dati trattati ed, in particolare, per la loro riservatezza[470]. I big data rappresentano un sottoinsieme potenzialmente molto rilevante di tutti i dati disponibili in natura: le operazioni possono essere condotte sia su dati personali che su dati che rimangono sostanzialmente estranei alla normativa privacy. In questo secondo caso si suole parlare di dato non-personale e

[468] Sugli aspetti legati alla condivisione e commercializzazione dei dati tra società, si veda *Information Commissioner's Office, Data sharing, code of practice*, 2016.

[469] Baldini G., Botterman M*., Ethical Design in the Internet of Things*, Springer, 2016.

[470] Bolognini L., Bistolfi C. - *Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation*; Nunziante E*., Big Data - Come proteggerli e come proteggerci* – Medialaws 5/2017.

classici esempi possono essere i dati sul traffico raccolti da sensori stradali, il numero di rotazioni di un tornello all'entrata della metropolitana in una grande metropoli, il dato statistico inerente al numero di piante coltivabili su un suolo agricolo, la loro dimensione, colore, struttura e così via. Anche senza conoscere cifre o statistiche si può capire quanto il volume totale dei dati non-personali presenti all'interno del mercato europeo sia maggiore rispetto ai dati personali identificabili in un soggetto (persona fisica).

Seguendo questa linea di pensiero e considerando la normazione dei flussi di dati come una vera e propria strategia, la Commissione Europea ha recentemente avviato tramite la comunicazione '*Costruire un'economia europea dei dati*', una proposta di politiche e soluzioni giuridiche per realizzare un'economia dei dati dell'UE.[471] Il concetto centrale - che non è sfuggito di vista al legislatore europeo - è che questi dati non-personali costituiscono un asset indispensabile allo sviluppo dell'economia Europea e del Digital Single Market. Secondo la già citata comunicazione '*istituendo per tempo un assetto programmatico e giuridico per l'economia dei dati, il suo valore* [dell'economia dei dati] *potrà raggiungere i 643 miliardi di euro nel 2020, pari al 3,17% del PIL complessivo dell'UE*'. I punti cruciali per poter liberare completamente l'appena citato potenziale sono molteplici. In *primis*, i diritti in capo al soggetto che produce il dato (*Data Producer's right*) e in particolare l'elemento della proprietà del dato. In breve: chi produce il dato per la prima volta ne è allo stesso tempo proprietario e detentore? In linea di massima sì. Si pensi ai dati prodotti da società attraverso l'utilizzo di propri macchinari o sistemi informativi. Un esempio: il dato afferente al numero finale di ordini che vengono effettuati da parte di clienti nei confronti di una società che produce scarpe. Il possedere questo dato è sicuramente un punto di forza per la società, ma è anche un dato che afferisce alla storia personale di un individuo e che potrebbe essere oggetto dell'esercizio di uno dei diritti previsti dal GDPR, quali il diritto di accesso o il diritto alla portabilità.

Il trattamento dei dati non-personali non è peraltro scevro da rischi. Aspetti critici relativi al rapporto tra normativa privacy e trattamento di dati (personali e non-personali) in operazioni di big data si possono riscontrare in tutte le fasi del trattamento di questi dati. Il primo momento critico è rappresentato dalla raccolta: che tipo di informativa, ed in rapporto a quale specifica finalità, l'interessato deve essere trasparentemente avvertito della raccolta del dato? Il principio della correttezza e trasparenza nei confronti dell'interessato, alla luce dei principi sanciti dalla normativa italiana e dalle raccomandazioni delle Autorità Europee (in particolare nella 36° Conferenza Internazionale di Protezione dei dati e privacy), impone di fornire una "adeguata" informativa all'interessato. Dovrebbe dunque essere esplicitamente riportata la finalità del compimento di attività di big data? Basta ancora, tutt'oggi, la menzione al semplice svolgere "attività di analisi e statistiche"? Quello che è certo è che tale informativa sarà il risultato di un bilanciamento del principio di trasparenza con l'esigenza del titolare di poter utilizzare il grande volume di dati raccolti per finalità e con modalità di trattamento che non sono sempre preventivamente conoscibili.

[471] Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Costruire un'economia dei dati europea*, adottata in data 10 gennaio 2017, http://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52017DC0009&from=EN (ultimo accesso 12 ottobre 2017).

La tutela dell'Interessato, a parere di chi scrive, non potrà quindi che spostarsi a monte, garantendo allo stesso tempo la conoscenza dei dati raccolti e l'adozione di *policy* e misure organizzative che riducano i rischi di re-identificazione dell'interessato, ove si sia proceduto alla pseudo-anonimizzazione dei dati. E' infatti la pseudo-anonimizzazione (e non l'anonimizzazione, che da un punto di vista tecnico costituisce un processo irreversibile) che potrebbe mettere a rischio la tenuta dei presidi (in *primis* quello dell'informativa e del consenso) a tutela delle libertà fondamentali dell'individuo: si pensi ad esempio ai dati anonimi o pseudo-anonimizzati raccolti dai grandi operatori della telefonia mobile. Quali sono personali? Quanti sono anonimizzati o pseudo-anonimizzati? Quanti vengono aggregati con altri dati non-personali? Sono molte le incertezze che permangono circa il carattere "personale" o non personale di questi dati. E' infatti a tutti evidente come, a fronte di analisi ed incroci di dati non-personali, gli operatori del settore della telefonia potrebbero più facilmente di altri svolgere un'attività di "single-out" dei propri clienti persone fisiche. Quello che è certo è che il titolare del trattamento dei dati deve provvedere a pratiche di data minimisation, così come ricordate nella recente opinione del WP29 sui processi decisionali automatizzati e sulla profilazione delle persone fisiche per i fini di GDPR.[472]

Da ultimo, il trattamento dei big data impone di verificare l'esattezza dei dati raccolti e la loro veridicità in funzione del loro utilizzo all'interno di algoritmi automatizzati. Questi ultimi, infatti, grazie alla rielaborazione di grandi quantità di dati possono condurre alla definizione di *cluster* da utilizzare per lo sviluppo di nuovi prodotti e servizi, con effetti predittivi sulle esigenze di un certo numero di individui. L'utilizzo di dati non corretti o la loro manipolazione potrebbe ri-determinare le scelte di consumo degli individui, trasformando i big data da mere informazioni utili ad interpolare gli accadimenti futuri, in un pericoloso strumento di coartazione delle libertà dell'individuo attraverso l' induzione di scelte future.

Queste sono sfide che necessiteranno un intervento del legislatore europeo, ma anche e soprattutto il definirsi di un approccio tecnologico unitario che disciplini la raccolta, il trattamento e la conservazione dei big data.

[472] Gruppo di lavoro Articolo 29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, adottato in data 3 ottobre 2017.

Bibliografia

5GPP, (2015), *5G Vision*

Accenture, (2016), *Digital disruption: The growth multiplier*.

Accenture, (2016), *Exploring Next Generation Financial Services: The Big Data Revolution*.

Acquisti A., (2010), *The economics of personal data and the economics of privacy*, Joint Working Party for Information Security and Privacy and Working Party on the Information Economy, Roundtable, background paper 3. Paris: Organisation for Economic Co-operation and Development.

Acquisti A., Adjerid I., Balebako, R. H., Brandimarte L., Cranor L. F., Komanduri S., Leon P. G., Sadeh N., Schaub F., Sleeper M., Wang Y., Wilson S., (2017) *, Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online*.

Acquisti A., Taylor C., Wagman L., (2016), *The Economics of Privacy*.

Acquisti A., Varian H. R., (2005), *Conditioning Prices on Purchase History*, 24 Marketing Sci. 367, 368.

AgID, (2017), *Task Force sull'Intelligenza Artificiale*, White Paper IA.

Agree P.E., Rotenberg M., (2001), *Technology and Privacy. The New Landscape*, MIT Press, Cambridge; Massachusetts, 7.

Allcott H., Gentzkow M., (2017), *Social Media and Fake News in the 2016 Election*, Journal of Economic Perspectives 2017, p. 211.

Almuhimedi H., Schaub F., Sadeh N., Adjerid I., Acquisti A., Gluck J., Cranor L., Agarwal Y., (2015), *Your location has been shared 5,398 times!: A field study on mobile app privacy nudging*, in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 787–796.

Amato G., Schwitzer H., Denozza F., Stallibrass D., e Nicita A., (2009), *Tutela della concorrenza e tutela dei consumatori. Due fini confliggenti?*, in Mercato concorrenza regole, 2, pp. 381-400.

Anderson C., (2008), *The end of theory: The data deluge makes the scientific method obsolete*, Wired.

Areeda P., Hovenkamp H., (2007), *Antitrust law. An analysis of antitrust principles and their application, Boston*, Aspen Publishers, IIB, 517, p. 150.

Argentesi E., Filistrucchi L., (2007), *Estimating Market Power in a Two-Sided Market: the Case of Daily Newspapers in Italy*, Journal of Applied Econometrics, Vol. 22(7), pp.1247-1266.

Argenton C., Prüfer J., (2012), *Search Engine Competition With Network Externalities*, 8 J. Competition L. & Econ. 73.

Armstrong M., (2006), *Competition in two-sided markets*, Rand Journal of Economics Vol. 37(3): 668-691.

Armstrong M., Wright J., (2007), *Two-sided Markets, Competitive Bottlenecks and Exclusive Contracts*, Economic Theory , 32 (2), 353-380.

Arun K., Jabasheela L., (2014), *Big Data: Review, Classification and Analysis Survey, International Journal of Innovative Research in Information Security (IJIRIS)*

Auer D., Petit N., (2015), *Two-sided markets and the challenge of turning economic theory into antitrust policy*.

Autorité de la concurrence and Bundeskartellamt, (2016), *Competition Law and Data*.

Ayres I. e Schwartz A., (2014), *The no-reading problem in consumer contract law*.

Bakos Y., Marotta-Wurgler F., Trossen D. R., (2014*), Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts*, New York University Law and Economics Working Papers. Paper 195.

Baldini G., Botterman M., (2016), *Ethical Design in the Internet of Things*, Springer.

Barbaro M., Zeller T. Jr., (2006), *A Face Is Exposed for AOL Searcher*, No. 4417749, N.Y. TIMES, Aug. 9.

Bellamy & Child, (2013), *European Union Law of Competition*, Oxford, Oxford University Press, p. 800.

Bennett J., Lanning S., (2007), *The Netflix Prize*.

Ben-Shahar O. e Schneider C.E., (2014), *More Than You Wanted to Know: The Failure of Mandated Disclosure*, Princeton University Press.

Ben-Shahar O., (2009), *The myth of the 'opportunity to read'*.

Ben-Shahar O., Schneider C. E., (2010), *The failure of mandated discourse*, U. Pa. Law Rev. 159, 3.

Bollier D., (2010), *The promise and peril of big data*.

Bolognini L., Bistolfi C. (2017), *Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation*.

Borgesius F. Z., van Eechoud M., Gray J, (2015), *Open data, privacy, and fair information principles: towards a balancing framework*.

Bourreau M., De Streel A. Graef I., (2017), *Big Data and Competition Policy: Market Power, Personalised Pricing and Advertising*.

Boyd D., Crawford K., (2012), *Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon*, in Information, Communication & Society 2012, p. 662.

Brand Finance, (2017), *Brand Finance Global 500 2017*.

Brauneis R., Goodman E. P., (2017), *Algorithmic Transparency for the Smart City*.

Calandrino J., Kilzer A., Narayanan A., Felten E., and Shmatikov V., (2011), *You might also like: Privacy risks of collaborative filtering*.

Califano L., (2016), *Privacy: affermazione e pratica di un diritto fondamentale*, Editoriale Scientifica, Napoli, 69.

Calvano E. e Jullien B., (2011), *Issues in on-line advertising and competition policy: a two-sided market perspective*.

Calvano E., Polo M., (2015), *Strategic Differentiation by Business Models: Free-to-Air and Pay-TV*, CSEF Working Paper.

Cambridge University Judge Business School, (2015), *Study on Personal Data Stores*, commissioned by DG CONNECT.

Carbonneau S., McAfee P., Mialon S., (2004), *Price discrimination and market power*, Emory economics, p. 413.

Carlton D. W., (2016), *Roundtable on price discrimination*, Directorate for financial and enterprise affairs Competition Committee.

Cavoukian, A., (2010), *Privacy by Design: The 7 Foundational Principles*.

Centre for Media Pluralism and Media Freedom, (2013), *European Union competencies in respect of media pluralism and media freedom*.

Chakravorti S., Roson R., (2006), *Platform Competition in Two-Sided Markets: The Case of Payment Networks*.

Chun-Wei Tsai, Chin-Feng Lai, Han-Chieh Chao, Athanasios V. Vasilakos, (2015), *Big data analytics: a survey*.

Cisco, (2012), *The Cloud Value Chain Exposed*, White Paper

Clerides S., (2002), *Price discrimination with differentiated products: definition and identification*, 20 International journal of industrial organization, p. 1385.

Cohn, Barbara L., (2015), *Data Governance: A Quality Imperative in the Era of Big Data, Open Data and Beyond*, ISJLP 811.

Colangelo G., Maggiolino M., (2017), *Big Data as a Misleading Facility*, 2017.

Colangelo G., Maggiolino M., (2017), *Data Protection in Attention Markets: Protecting Privacy through Competition?*.

Commission Staff Working Document, (2017), *The free flow of data and emerging issues of the European data economy*.

Cooper J.C., (2013), *Privacy and Antitrust: Underpants Gnomes, the First Amendment, and Subjectivity*, 20 Geo. Mason L. Rev. 1129.

Cory N., (2017), *Cross-Border Data Flows: Where Are the Barriers, and What Do They Cost?*.

Culnan M. J., Armstrong P.K., (1999), *Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation*, School of Business, Georgetown University, Washington, D.C. 20057-1008.

Curry E. (2014), *The Big Data Value Chain, European Data Forum*

D'acquisto G., Naldi M., (2017), *Big data e Privacy by design. Anonimizzazione, pseudonimizzazione e sicurezza*, Giappicchelli, Torino, 41-171.

Damien G., Ioannis L., (2017), *Competition Policy: Between Equity and Efficiency*, Cambridge: Cambridge University Press.

Deloitte, (2015), *Inside the Internet of Things (IoT)*.

Deutscher E., (2017), *The role of personal data and privacy in merger analysis – a critical reassessment of the EU Commission's Facebook/WhatsApp decision.*

Di Porto F., (2017), *La regolazione degli obblighi informativi le sfide delle scienze cognitive e dei big data*, Editoriale Scientifica, ISBN 978-88-9391-095-8.

Dragos, D.C., Neamtu B., (2009), *Reusing public sector information — policy choices and experiences in some of the Member States with an emphasis on the case of Romania*, in European Integration Online Papers.

Drexl J., (2016), *Designing competitive markets for industrial data – between propertisation and access*.

Drexl J., (2016), *Economic Efficiency versus Democracy: On the Potential Role of Competition Policy in Regulating Digital Markets in Times of Post-Truth Politics*.

Dutton, W. H., Reisdorf B. C., Dubois E., Blank G., (2017), *Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States*.

Edelman B., (2015), *How to launch your digital platform*, Harv. Bus. Rev. 90.

Edward S., (2015), *Reflections on the Concept of Open Data*, SCRIPTed 154.

Einav L., Levin J. D., (2013), *The data revolution and economic analysis*.

Eisenmann T., Parker G., Alstyne M.W.V, (2006), *Strategies for Two-Sided Markets*.

Eldeman B. e Lockwood B., (2011), *Measuring Bias in Organic Web Search*.

Eli Pariser, (2011), *The Filter Bubble: What the Internet Is Hiding from You*.

ENISA – European Union Agency for Network and Information Security, (2015), *Big data security. Good practices and recommendations on the security of big data systems*.

ENISA – European Union Agency for Network and Information Security, (2015), *Privacy by design in big data*.

European Commission DG CONNECT, (2013), *A European strategy on the data value chain*.

European Commission DG CONNECT, (2016), *Definition of a research and innovation policy leveraging Cloud Computing and IoT combination*.

European Commission DG CONNECT, (2016), *Legal study on ownership and access to data*.

European Commission, (2011), *Open data: An engine for innovation, growth and transparent governance*,COM(2011)882.

European Commission, (2014), *HORIZON 2020*.

European Commission, (2014), *Towards a thriving data-driven economy*.

European Commission, (2015), *Benchmarking of the symbolic machine learning classifier with state of the art image classification methods*.

European Commission, (2015*), Collaborative research-grade software for crowd-sourced data exploration*.

European Commission, (2016), *Advancing the Internet of Things in Europe*, Commission Staff Working Document.

European Commission, (2016), *Connectivity for a Competitive Digital Single Market - Towards a European Gigabit Society*.

European Commission, (2016), *Digitalizzazione dell'industria europea. Cogliere appieno i vantaggi di un mercato unico digitale*, COM(2016) 180.

European Commission, (2016), *EU eGovernment Action Plan 2016-2020 Accelerating the digital transformation of government*.

European Commission, (2016), *European Cloud Initiative - Building a competitive data and knowledge economy in Europe*, COM(2016) 17.

European Commission, (2016), *ICT Standardisation Priorities for the Digital Single Market*.

European Commission, (2016), *Il 5G per l'Europa: un piano d'azione*, COM(2016) 588.

European Commission, (2016), *Legal study on ownership and access to data*.

European Commission, (2016), *Online Platforms and the Digital Single Market*.

European Commission, (2016), *Priorità per la normazione delle TIC per il mercato unico digitale*, COM(2016) 176.

European Commission, (2016), *Proceedings of the 2016 conference on Big Data from Space.*

European Commission, (2016), *Study on Big Data in Public Health, Telemedicine and Healthcare*.

European Commission, (2017), *European Data Market study*.

European Commission, (2017), *Facilitating cross border data flow in the Digital Single Market*.

European Commission, (2017), *Next steps against fake news: Commission sets up High-Level Expert Group and launches public consultation*.

European Commission, staff working document, (2014), *Accompanying the document Towards a thriving data-driven economy*.

European Parliament, Directorate-General for Internal Policies of the Union, (2015), *Big Data and smart devices and their impact on privacy*.

European Parliament, Directorate-General for Internal Policies of the Union, (2016), *Industry 4.0*.

Evans D., (2011), *Platform Economics: Essays on Multi-Sided Businesses*.

Evans D., (2011), *The antitrust economics of free*, 7 Competition Policy Int. l 71.

Evans D., Schmalensee R., (2008), *Markets with Two-sided Platforms*, Issues in Competition Law and Policy, Vol. 1:667-693.

Evans D.S., (2002), *The antitrust economics of two-sided markets*.

Evans D.S., (2011), *The Antitrust Economics of Free*.

Evans D.S., (2016), *Multisided platforms, dynamic competition, and the assessment of market power for internet-based firms*.

Evans D.S., Schmalensee R., (2013), *The antitrust analysis of multi-sided platform business*, in RD Blair e DD Sokol, Oxford Handbook on International Antitrust Economics.

Faull J., Nikpay A., (2014), *The EU law of competition*, Oxford, Oxford University Press, p. 524 e 533.

FCA, (2016), *Call for Inputs on Big Data in retail general insurance*.

Filistrucchi L., (2010), *How many markets are two-sided?*, Competition Policy International-Antitrust Chronicle, Vol.7.

Filistrucchi L., Geradin D., Damme E.V., Affeldt P., (2014), *Market definition in two-sided markets: theory and practice*, Journal of Competition Law and Economics, Volume 10 (2), pp. 293-339.

Frignani A., (2013), *Intese discriminatorie*, sez. 5, in Disciplina della concorrenza nella UE, p. 242.

FTC, (2014), *Data broker report.*

FTC, (2016), *Big Data. A Tool for Inclusion or Exclusion?*.

Gal M. S., Rubinfeld D. L., (2016), *Access Barriers to Big Data*.

Gal M.S., Rubinfeld D.L., (2015), *The Hidden Costs of Free Goods: Implications for Antitrust Enforcement*.

Garante per la protezione dei dati personali, (2016), *Guida al nuovo regolamento europeo in materia di protezione dei dati personali*.

Gartner (2017), Gartner, Forecast: Public Cloud Services, Worldwide, 2015-2021

Gaudeul A., Jullien B., (2007), *E-commerce, two-sided markets and info-mediation*, in Internet and Digital Economics (Brousseau E. and Curien N. eds.), 2007, 268.

Gerhardt B., Griffin K., Klemann R., (2012), *Come estrarre valore nell'universo frammentato dell'analisi dei Big Data*, Cisco.

Girolami M., (2014), *Etichettatura, informazioni e rimedi privatistici nella vendita di prodotti alimentari ai consumatori*, in Nuove Leggi Civili Commerciali.

Goldfarb A., Tucker C., (2011), *Privacy Regulation and Online Advertising*, 57 MGMT.SCI.57, 61.

Gomez C., Hunt N., (2015), *The Netflix recommender system: Algorithms, business value, and innovation. ACM Trans. Manage. Inf. Syst. 6, 4, Article 13*

Graef I., (2015), *Market Definition and Market Power in Data: The Case of Online Platforms*.

Graef I., (2016), *EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility*.

Graef I., Verschakelen J., Valcke P., (2013), *Putting the Right to Data Portability into a Competition Law Perspective*, Law: The Journal of the Higher School of Economics, Annual Review.

Grunes A., (2013), *Another Look at Privacy*, 20 Geo. Mason L. Rev. 1107.

Grunes A., Stucke M., (2015), *No Mistake About It: The Important Role of Antitrust in the Era of Big Data*, 14 Antitrust Source 1.

GSMA Intelligence, (2016), *Global mobile trends*.

GSMA, (2016), *Resetting competition policy frameworks for the digital ecosystem*.

Gurin J., (2014-2015), *Big Data and Open Data: How Open Will the Future Be*, ISJLP 691.

Hacker F., (2017), *Personal data, exploitative contracts, and algorithmic fairness: autonomous vehicles meet the internet of things*, International Data Privacy Law, 2017, Vol. 0, No. 0, 1.

Hadfield G.K., Howse R., Trebilcock M. J., (1998), *Information-Based Principles for Rethinking Consumer Protection Policy*, 21 Journal of Consumer Policy, p. 150.

Hagiu A., Rothman S., (2016), *Network Effects Aren't Enough*, Harv. Bus. Rev. 64.

Hagiu, A., B. Jullien (2011), *Why do intermediaries divert search?*, The RAND Journal of Economics, 42 (2), 337-362.

Haigiu A., (2006), *Pricing and commitment by two-sided platforms*, Rand Journal of Economics, Vol.37(3): 720-737.

Haigiu A., (2014), *Strategic decisions for multisided platforms*, MIT Sloan Management Review.

Haigiu A., Spulber D., (2013), *First-party content and coordination in two-sided markets*, Management Science, Vol.59(4) 933-949.

Harbour P. J., T. I. Koslov, (2010), *Section 2 in a Web 2.0 World: an expanded vision of relevant product markets*, 76 Antitrust L J 769, 794.

Harbour P.J., (2007), *Dissenting statement*, In the matter of Google/DoubleClick, p. 4.

Harbour, T. I. Koslov, (2011), *Privacy in Antitrust: A Relevant Product Market Factor*.

iCom, (2016), *Internet of Things & 5G revolution*.

IDC, (2015), *Key success Factors for digital transformation in the Banking Industry*.

IDC, (2016), *FutureScape: Worldwide IT Industry 2017 Predictions*.

Il-Horn Hann, Kai-Lung Hui, Tom Lee, I. Png, (2002), *Online Information Privacy: Measuring the Cost-Benefit Trade-off*, 23 Int'l Conf. On Info. Systems 1.

ITMedia Consulting, LUISS DREAM, (2016), *Il video e la terza rivoluzione di internet*.

James F. Ponsoldt J.F., David C. D., (2007), *A Comparison Between U.S. and E.U.Antitrust Treatment of Tying Claims Against Microsoft: When Should the Bundling of Computer Software Be Permitted?*, 27 NW. J. INT'L L. & BUS. 421, 448.

Jan Whittington & Chris Jay Hoofnagle (2012), *Unpacking Privacy's Price*, 90 N.C. L. REV. 1327.

Jansenn K., Hugelier S., (2013), *Open Data: A New Battle in an Old War Between Access and Privacy?*, in Digital Enlightenment Yearbook 2013 190, 199.

Jullien B., (2015), *Competition policy for on-line platforms*, intervento al Convegno 'On Line Intermediation and Competition in E-Commerce', Università Roma Tre.

Jullien, B. (2011), *Competition in multi-sided markets: Divide-and-Conquer*, American Economic Journal: Microeconomics, No 3, 1-35.

Kaiser U., Wright J., (2006), *Price structure in two-sided markets: Evidence from the magazine industry*.

Kalimo H., Majcher K., (2017), *The concept of fairness: linking EU competition and data protection law in the digital marketplace*, European Law Review, p. 210.

Keim Daniel A., (2017), *Data Drive Innovation open summit*.

Keim Daniel A., Kohlhammer J., Ellis G., (2010), *Mastering the Information Age Solving Problems with Visual Analytics*.

Kerber W., (2016), *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis*.

Kerber W., (2016), *Digital Markets, Data, and Privacy: Competition Law, Consumer Law, and Data Protection*, 11 Journal of Intellectual Property Law & Practice 856.

King S.P., (2013), *Two-Sided Markets*.

Kitchin R., (2014), *Big Data, new epistemologies and paradigm shifts*, Big Data & Society, p. 1.

Kitchin R., (2014), *The Data Revolution*.

Klitou D., (2014), *Privacy-Invading Technologies and Privacy by Design: Safeguarding Privacy, Liberty and Security in the 21st Century*, Springer, 260-268.

Kuner, C., (2017) *Reality and Illusion in EU Data Transfer Regulation Post Schrems*, German Law Journal, Vol. 18, 881.

Lambrecht A., Tucker C. E., (2015), *Can Big Data Protect a Firm from Competition?*.

Lasica J.D., (2009*), Identity in the Age of Cloud Computing: The Next-Generation Internet's Impact on Business, Governance and Social Interaction*, The Aspen Institute, Washington DC.

Laudon, K. C. (1996), *Markets and Privacy*, Communications of the ACM (39:9), pp. 92-104.

Lerner A. V., (2014), *The Role of 'Big Data' in Online Platform Competition*.

Leucci S., (2014), *Preliminary Notes on Open Data Licensing*, J. Open Access L. 1.

Levine M. E., (2002), *Price discrimination without market power*, 19 Yale J. on Reg. (2002) p. 1.

Libertini M., (2014), *Diritto della concorrenza dell'unione europea*, Giuffré, Milano, p. 326.

Lundqvist B., Forsberg Y., De Vries M., Maggiolino M., (2015), *Open Data and Competition Law: Some Issues regarding Access and Pricing of Raw Data*, Masaryk U. J.L. & Tech. 95.

Lynskey O., (2014), *Deconstructing data protection: The 'Added-Value' of a Right to Data Protection in the EU Legal Order*, International and Comparative Law Quarterly 63(3), 569–597.

Maggiolino M., (2012), *Il riutilizzo dell'informazione detenuta dal settore pubblico: alcune riflessioni di politica e diritto della concorrenza*, Concorrenza e mercato, pp. 765 – 802.

Maggiolino M., (2016), *Big data e prezzi personalizzati*, in Concorrenza e Mercato, pp. 95-138.

Manne G. A., Ben Sperry R., (2014), *The law and economics of data and privacy in antitrust analysis*.

Manne G.A., Sperry R.B., (2015), *The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework*, 5 CPI Antitrust Chronicle.

Mantelero A., (2017), *From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era*, in Taylor L., Floridi L., van der Sloot B. (a cura di), Group Privacy. New Challenges of Data Technologies, Springer International Publishing, 139-158.

McDonald A. M., Cranor L. F., (2008), *The Cost of Reading Privacy Policies,* I/S: A Journal of Law and Privacy for the Information Society 4:540–65.

McKinsey Global Institute, (2011), *Big data: The next frontier for innovation, competition, and productivity*.

McKinsey Global Institute, (2011), *The challenge and opportunity of Big data*.

McKinsey Global Institute, (2016), *The Age Of Analytics: Competing In A Data-Driven World*.

Meeker, M., 2015. "*Internet Trends 2015"* KPCB website

Michael Luca et al., (2015), *Does Google Content Degrade Google Search? Experimental Evidence*, Harvard Business School Working Paper.

Michael Luca, Reviews, (2011), *Reputation and Revenues: The Case of Yelp.com*, Harvard Business School.

Mik E., (2016), *The Erosion of Autonomy in Online Consumer Transactions*, 8 Law, Innovation and Technology 1.

Miller G., (2014), *The huge, unseen operation behind the accuracy of Google Maps*.

Min Chen-Shiwen Mao- Yunhao Liu, (2014), *Big data: a survey*.

Monopolkommission, (2015), *Competition policy: The challenge of digital markets, Special Report*.

Morozov E., (2011), *L'ingenuità della rete*, Codice Edizioni, Torino.

Motta M., (2009), *Competition Policy*, Cambridge University Press, Cambridge.

Narayanan A., Huey J., Felten E.W., (2016), *A Precautionary Approach to Big Data Privacy*, Data Protection on the Move. Law, Governance and Technology Series, vol 24. Springer, Dordrecht.

Narayanan A., Shmatikov V., (2008), *Robust de-anonymization of large sparse datasets*.

Nelson P., (2015), *Monopoly Power, Market Definition and the Cellophane Fallac*y, U.S. Department of Justice.

Newman J. H., (2015), *Antitrust in zero price markets: Foundations*, 164 University of Pennsylvania law review 149.

Newman J. H., (2015), *Antitrust in Zero-Price Markets: Application*, Washington University Law Review, Vol. 94, Issue 1, 49

Newman N., (2013-2014), *The Costs of Lost Privacy: Consumer Harm and Rising Economic Inequality in the Age of Google*, 40 Wm. Mitchell L. Rev. 849.

Newman N., (2014), *Feature: Search, Antitrust, and the Economics of the Control of User Data*, 31 Yale J. on Reg. 401.

Newman N., (2014), *Search, Antitrust, and the Economics of the Control of User Data*, 31 Yale J. on Reg. 401.

Newman N., (2015), *Data Justice- Taking on Big Data as an Economic Justice Issue*.

Nicita A., (2017), *L'insostenibile leggerezza della post-verità,* prefazione al volume Extra Media a cura di Enzo Argante, Egea.

Nicita A., (2017), *Pluralism in the digital age: public service media and technological changes*, Relazione al Convegno Eurovisione.

Nokia, (2016), *The Future X Network: A Bell Labs Perspective*.

Noto La Diega G., Walden I.(2016), *Contracting for the 'Internet of Things': Looking into the Nest*.

Nunziante E., (2017), *Big Data - Come proteggerli e come proteggerci*, Medialaws.

Nunziante E., (2017), *Big Data. Come proteggerli e come proteggerci. Profili di tutela tra proprietà intellettuale e protezione dei dati personali*, in Law and Media WorkingPaper Series.

OCSE, (2015), *Data Driven Innovation Big Data for Growth and Well-Being*.

OCSE, (2015), *Data-Driven Innovation Big Data for Growth and Well-Being*.

OECD, (1980), *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*.

OECD, (2008), *The Interface Between Competition And Consumer Policies*, DAF/COMP/GF(2008)10.

OECD, (2013), *Supporting Investment in Knowledge Capital Growth and Innovation*.

OECD, (2013), *The role and measurement of quality in competition analysis*.

OECD, (2016), *Price Discrimination*.

OECD, (2017), *Algorithms and Collusion - Background Note by the Secretariat*.

Office of Government Commerce (2011), *IT Infrastructure Library (ITIL) v.3*

Ogus A, (1994), *Regulation. Legal Theory and Economic Forms*, Oxford University Press, Oxford.

Ohlhausen M., Okuliar A., (2015), *Competition, Consumer Protection, And The Right [Approach] To Privacy*.

Oracle (2014), *Information Management and Big Data, A Reference Architecture*

Oracle (2016), *An Enterprise Architect's Guide to Big Data, Reference Architecture Overview*

P. Belleflamme, Peitz M., (2015), *Industrial Organization. Market and Strategies*, Cambridge, Cambridge University Press, p. 198.

Pagnatelli V., (2016), *Access, Accessibility, Open Data. The Italian Model of Public Open Data in the European Context*, Giornale di Storia Costituzionale 205.

Pappalardo A., (2007), *Il diritto comunitario della concorrenza*, p. 177.

Pasquale F., (2013), *Privacy, Antitrust, and Power*, 20 Geo. Mason L. Rev. 1009.

Patterson M., (2012), *Manipulation of product ratings: Credit-Rating Agencies, Google, and Antitrust*, CPI Antitrust Chronicles.

Patterson M., (2017), *Antitrust Law in the New Economy*.

Peng P., (2016), *Implications of Big Data for Competition Policy*.

Perez C., (2002), *Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages*, Edward Elgar, Cheltenham

Pigou A.C., (1920), *The economics of welfare*, London, Macmillan and Co..

Pino G., (2010), *L'identità personale*, in Trattato di biodiritto, diretto da S. Rodotà e P. Zatti, vol. I, Ambito e fonti del biodiritto, a cura di S. Rodotà e M. Tallacchini, Giuffrè, Milano, 297-321.

Pizzetti F., (2016), *Privacy e il diritto europeo alla protezione dei dati personali. Il Regolamento europeo 2016/679,* Giappichelli.

Polo M., (2010), *Concentration and Competition in Media Markets*, EUI.

Preibusch S., (2015), *How to Explore Consumers' Privacy Choices with Behavioral Economics*.

Presidenza del Consiglio dei Ministri, (2015), *Strategia per la crescita digitale 2014-2020*.

Preta A., (2012), *Televisione e mercati rilevanti*, Vita e Pensiero.

Preta A., Peng P., (2016), *Competition and innovation in digital markets*, Working paper, EALE Poland 2017.

Prufer J., Schottmüller C., (2017), *Competing with Big Data*, Tilburg Law School Legal Studies Reseach Paper Series No. 06/2017.

Przybylski A. K., Murayama K., DeHaan C. R., & Gladwell, V., (2013), *Motivational, emotional, and behavioral correlates of fear of missing out*, Computers in Human Behavior, Vol. 29.

Purtova N., (2011), *Property in Personal Data: a European Perspective on the Instrumentalist Theory of Propertisation*.

PWC, (2016), *Catalyst or threat? The strategic implications of PSD2 for Europe's banks*.

Pwc, (2016), *Industry 4.0: Building the digital enterprise.*

Quattrociocchi W., Scala A., Sunstein C. R., (2016), *Echo Chambers on Facebook*.

Reding V., (2012), *Outdoing Huxley: Forging a High Level of Data Protection for Europe in the Brave new Digital World*, Speech at Digital Enlightenment Forum.

Reidenberg J. R., Bhatia J., Breaux T.D., Norton T. B., (2016), *Ambiguity in Privacy Policies and the Impact of Regulation*, Journal of Legal Studies 45:S163–S190.

Rochet J. C., Tirole J., (2003), *Platform Competition in Two-Sided Markets*, Journal of the European Economic Association, 1:4, pp. 990-1029.

Rochet J. C., Tirole J., (2006), *Two-sided markets: a progress report*.

Rodotà S., (2012), *Il diritto di avere diritti*, Editori Laterza, 404.

Roman, Charles, (2016), *Open Data*, ConLawNOW 19.

Rouvroy A., Poullet Y., (2009), *The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy*, in Reinventing Data Protection, Springer, Dordrecht. 45-76.

Salerno F., (2014), *Consenso informato in medicina e qualità soggettive del paziente*, in Giurisprudenza italiana.

Santander InnoVentures, (2015), *Rebooting financial services*, in collaboration with its partners Oliver Wyman and Anthemis Group.

Saurwein F., (2011), *Regulatory Choice for Alternative Modes of Regulation: How Context Matters*, in Law & Policy, 33(3):334-366.

Savage S., Waldman D. M., (2013), *The Value of Online Privacy*.

Schatsky D., Muraskin C., Deloitte University Press, (2015), *Beyond bitcoin, Blockchain is coming to disrupt your industry*.

Schepp N. P., Wambach A., (2016), *On Big Data and Its Relevance for Market Power Assessment*, pp. 121-122.

Schumpeter J. A., (1994), *Capitalism, Socialism and Democracy*, 1994 [1942], London: Routledge. pp. 82–83.

Schwartz A., Wilde L., (1979), *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis,* University of Pennsylvania Law Review 127:630.

Shaked A., Sutton J., (1982), *Relaxing Price Competition through Product Differentiation*, Review of Economic Studies, Vol. 49:3-13.

Shapiro C., Varian H.R., (1999), *Information rules: a strategic guide to the network economy*, Boston, Harvard Business School Press.

Sica S., D'Antonio V., (2016), *Verso il Privacy Shield: il tramonto dei Safe Harbour Privacy Principles*, 137-168, in La Protezione Transnazionale dei Dati Personali (Resta G., Zeno-Zencovich V., a cura di), Roma Tre Press.

Sensortower estimates (2017), *Data that Drive App Growth Report*

Sidak J. G. & Teece D. J., (2009), *Dynamic competition in antitrust law*.

Sokol D. D., Comerford R. E., (2016), *Antitrust and regulating big data*.

Stahl F., Schomm F., Vossen G., Vomfell L., (2016), *A classification framework for data marketplaces*, 3 Vietnam J. Comput. Sci. 137.

Steadman I., (2013), *Big data and the death of the theorist*, Wired.

Stigler G. J., (1966), *The theory of price*, New York, The Macmillan Company, p. 209.

Stole M. A., (2003), *Price discrimination and imperfect competition*.

Stone E., e Stone D., (1990), *Privacy in organizations: Theoretical issues research findings, and protection mechanisms*, Research in personnel and human resources management 8 (3), 349-411.

Strahilevitz L., Kugler M., (2016), *Is Privacy Policy Language Irrelevant to Consumers?,* Coase-Sandor Working Paper Series in Law and Economics, No. 776.

Strozzi G., Mastroianni R., (2011), *Diritto dell'Unione Europea. Parte istituzionale. Dal trattato di Roma al trattato di Lisbona*, Giappichelli.

Stucke M. , Grunes A., (2015), *The Curious Case of Competition and Quality*.

Stucke M. E., Ezrachi A., (2016), *Virtual Competition - The Promise and Perils of the Algorithm-Driven Economy*.

Stucke M., Ezrachi A., (2016), *When Competition Fails to Optimise Quality: A Look at Search Engines*, 18 Yale Journal of Law and Technology 70.

Stucke M., Grunes A., (2016), *Big Data and Competition Policy*, 337.

Stucke M.,Ezrachi A., (2015), *Artificial Intelligence & Collusion: When Computers Inhibit Competition*.

Sutton J., (1991), *Sunk costs and market structure: price competition, advertising, and the evolution of concentration*, MIT Press, Cambridge, Mass. ISBN 0-26-219305-1.

Swanson D.G., Baumol W.J., (2005), *Reasonable and Nondiscriminatory (RAND) Royalties, Standards Selection, and Control of Market Power*, 73 Antitrust L.J. 1.

Swire P., (2007), *Protecting Consumers: Privacy Matters in Antitrust Analysis*, Center for American Progress.

Swire, P., Lagos Y., (2013), *Why the Right to Data Portability Likely Reduces Consumer Welfare: Antitrust and Privacy Critique*, 72 Maryland Law Review 335 (2013); Ohio State Public Law Working Paper 204.

Thaler R. H., (1988), *Anomalies: The Winner's Curse*, Journal of Economic Perspectives 2 (1): 191–202.

Thaler R.H., Sunstein C. R., (2008), *Nudge: Improving decisions about health, wealth, and happiness*, TC Leonard - Constitutional Political Economy.

The Boston Consulting Group, (2016), *Inside OPS.*

The Future Group, (2008), *Freedom, Security, Privacy: European Home Affairs in an Open World.*

Thierer A., (2013), *Relax and Learn to Love Big Data*.

Thomas H. Davenport, (2006), *Competing on Analytics.*

Thomas L.D.W., Leiponen A., (2016), *Big Data Commercialization*, 44 IEEE Eng. Manag. Rev. 74.

Thomas, J., Cook, K., (2005), *Illuminating the Path: Research and Development Agenda for Visual Analytics*, IEEE-Press.

Timmer J., (2008), *Why Cloud cannot Obscure the Scientific Method?*, Ars Technica.

Trovato C. A., (2017), *La valutazione d'impatto privacy. Obblighi per le imprese e casi concreti*.

Tucker C. E., (2011), *Social Networks, Personalized Advertising, and Privacy Controls*.

Tucker C., (2013), *The Implications of Improved Attribution and Measurability for Antitrust and Privacy in Online Advertising Markets*, 20 Geo. Mason L. Rev. 1025, 1030.

Tucker D. S., Wellford Hill B., (2014), *Big Mistakes Regarding Big Data*.

U.S. Department of Justice and the Federal Trade Commission, (2010), *Horizontal Merger Guidelines.*

UK Competition & Markets Authority, (2015), *The commercial use of consumer data*, par. 3.56.

Vaidhyanathan S., (2011), *The googlization of everything (And why we should worry)*.

van der Sloot B., (2014), *Do data protection rules protect the individual and should they? An assessment of the proposed General Data Protection Regulation*, International Data Privacy Law 4(4), 307, 318-320.

Varian H. R., Tucker D. S., Wellford H. B., (2014), *Big Mistakes Regarding Big Data*.

Varian H.R., (1989), *Price Discrimination*, in Schmalensee R., Willig R.D., Handbook of industrial organization, Vol. 1, p. 600.

Varian H.R., (1996), *Economic Aspects of Personal Privacy, Privacy and Self-regulation in the Information Age*, National Telecommunications and Information Administration.

Varian H.R., (2013), *Beyond Big Data*.

Varian, H. R. (1997), *Economic Aspects of Personal Privacy*, in U.S. Department of Commerce, Privacy and Self-Regulation in the Information Age.

Walden I., Laise Da Correggio L., (2011), *Ensuring Competition in the Clouds: The Role of Competition Law?*.

Walker R., (2015), *From Big Data To Big Profits: Success With Data And Analytics*.

Weyl, E., (2010), *A price theory of multi-sided platforms*, The American Economic Review , 100 (4), 1642-1672.

White House, (2015), *Big data and differential pricing*.

Wiebe A., (2016), *Protection of industrial data—a new property right for the digital economy*, Gewerblicher Rechtsschutz und Urheberrecht Internationaler Teil (GRUR Int) 877 a 881.

Williamson B., (2014), *The death of the theorist and the emergence of data and algorithms in digital social research*.

Zevenbergen B., Brown I., Wright J., Erdos D., (2013), *Ethical Privacy Guidelines for Mobile Connectivity Measurements*, Oxford Internet Inst.

ITMedia Consulting

ITMedia Consulting (www.itmedia-consulting.com) è una società di ricerca e consulenza nel settore dell'economia digitale, operante in particolare nell'ambito dei contenuti e dei media digitali. La sua caratteristica è la focalizzazione sui temi della transizione al digitale, dell'accesso ai contenuti e della convergenza tra media, Internet e tlc.

Tra le principali aree d'attività si segnalano: analisi dei mercati e profili concorrenziali; accesso ai contenuti (es. gestione dei diritti, tematiche di pricing, barriere all'ingresso); nuovi modelli distributivi dei contenuti; offerte innovative (VOD, downloading e video streaming) e analisi dei modelli di business; accesso alle reti (assistenza ai fornitori di contenuti per l'accesso alle reti degli operatori di comunicazioni elettroniche e alle reti digitale terrestri); assistenza in operazioni antitrust, quali concentrazioni, joint-venture e fusioni tra aziende del settore televisivo e delle comunicazioni elettroniche.

I principali clienti: 3 Italia, AC Nielsen, Alcatel Lucent, Autorità per le Garanzie nelle Comunicazioni, Banca Nazionale del Lavoro, Cairo Communications, Canal Plus, Carat TV, Confindustria Servizi Innovativi, Council of Europe, Discovery, European Commission, Fastweb, Fox, France Télécom, Google, Infront, Mediaset, MTV, Paribas, Publikompass, Publitalia '80, Radio France, Rai, RaiSat, Raiway, RCS, SAT2000, Seat, Sky, Sipra, Sirti, Telecom Italia, Telecom Italia Media, Teleçinco, Telepiù, Time Warner, Tiscali, Viacom, Vivendi, Vodafone, Wind.

ASK Università Bocconi

Il Centro di Ricerca ASK è stato fondato nel 2004 all'interno dell'Università Bocconi con l'obiettivo di condurre ricerca teorica ed applicata in una continua dialettica tra arte, cultura e scienze sociali. Approfondisce la conoscenza dei settori artistici e delle industrie culturali attraverso la promozione e lo studio delle politiche culturali, del management delle istituzioni e dell'economia della cultura. Collabora attivamente con docenti e ricercatori dell'Università Bocconi ed è di riferimento e supporto per studenti e laureandi del corso di laurea triennale CLEACC e per il corso di laurea specialistica ACME.

Il Centro ASK è impegnato a studiare la gamma di attività e di settori caratterizzati in modo significativo dal rapporto tra arti ed economia. Le quattro principali aree tematiche di ricerca e lavoro sono arte contemporanea, patrimonio culturale e istituzioni, industrie culturali e proprietà intellettuale. Nel campo del patrimonio culturale, ASK vanta una lunga tradizione di lavoro ed ha svolto diverse ricerche per il MiBACT, accompagnando le istituzioni pubbliche nei momenti di riforma. Dopo un ampio studio per la Soprintendenza Speciale per i Beni Archeologici di Roma (con particolare riferimento alla struttura organizzativa e ai processi di pianificazione e controllo), il Centro sta svolgendo attività di ricerca per il Polo Museale della Lombardia e la Galleria Nazionale di Arte Moderna di Roma. Allo stesso modo, nel capo della proprietà intellettuale ASK è stato – e tuttora è partner di progetti europei, FP7, H2020, IPI e network tematici.