DATA DRIVEN ECONOMY
Market Trends and Policy Perspectives

Executive Summary – Research Report

With the scientific contribution of

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

Augusto Preta  
Project Leader

Mariateresa Maggiolino  
Editorial coordinator Part II

Other author: Giuseppe Colangelo

Contributors: Aura Bertoni, Paolo Biedanelli, Carlo Rossi Chauvenet, Giulia Filosa, Carlo Salatino.

Data Driven Economy: Market Trends and Policy Perspectives is a 342 pages report carried out by ITMedia Consulting, with the scientific contribution of the Bocconi University Research Center ASK.

Its aim is to provide insights allowing antitrust and regulatory authorities to update interpreting models and develop more efficient tools fitting the complex and quick nature of the changes the data driven digital economy is going through.

The data included in the report, updated to October 2017, come from ITMedia Consulting internal sources, public data, papers, essays and other academic and scientific publications, reports from analysts, balance sheets and companies' internal documentation, annual reports.

The work is divided into two parts: the first part, consists of two chapters, related to markets trends in digital economy and data driven economy in particular; the second part concerns the policy perspectives focusing on competition, regulation and privacy profiles.
In order to carry out the study, prominent experts and companies where consulted and involved also in internal workshops. Namely: 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) and Andrea Stazi, Public Policy & Government Relations' Counsel of Google Italia, which provided the financial support to the project. We also wish to thank Emilio Calvano (Università di Bologna) and Federico Ghezzi (Università Bocconi) for their friendly precious suggestions and valuable insights.

In expressing appreciation and gratitude to all the contributors, it should also be noted that ITMedia Consulting, together with ASK and all the authors for their specific parts, are the sole responsible of the content and the opinions featuring this work.

Data Driven Economy: Market Trends and Policy Perspectives (http://www.itmedia-consulting.com/DOCUMENTI/economiaideidati.pdf) is published in Italian, in January 2018, and distributed under a creative commons license (non-commercial attribution – no derivative works, CC BY-NC-ND), in order to guarantee the widest possible distribution and circulation.
Part I – Market

Digital Economy

We are at the dawn of an era that has the potential to define a new phase of human existence. This era will be led by the digitization and connection of everything and everyone with the goal of automating much of life, effectively creating time by maximizing the efficiency of everything we do, augmenting our intelligence with knowledge that expedites and optimizes decision making and everyday routines and processes. That revolution is the result of a continuous, relentless process of innovation that has characterized the evolution of the Internet over the last two decades through the diffusion of digital economy in all sectors, which is not more exclusively linked to the ICT.

This disruptive transformation begins with the first spread of the world wide web (1995-2005) and proceeds with social networks and so-called web 2.0 (2005-2015). Just as these transformations have been driven by technological innovation that has introduced several paradigm shifts in connectivity, information processing, usage and exploiting, so the next transformation - the third generation of the Internet - will be driven by technological changes that will lead to changes in demand and consumption patterns, as well as in the whole economy and business models.

First of all, digital data exponentially grows: if, in the past, there were analogue representations of reality, business, public institutions, etc., that could be digitised, nowadays the "datafication" of every aspect has dramatically increased. Therefore, according to Moore's law, technology innovation continues to provide, at a greater speed and at lower costs, the improvement of performances as well as the ease access to the three basic information and data operations: processing, computing power of microprocessors; storage availability and capacity; transmission capacity and speed, thanks to the development of broadband and ultrabroadband networks.

Innovative technologies, platforms and systems like Cloud Computing, Internet of Things, Big Data & Analytics, Blockchain, Artificial Intelligence, Augmented Reality & Virtual Reality, Advanced robotics & 3D printing e 5G are the new enablers of the digital economy.

With its pervasive diffusion in all sectors, digital economy promises to establish a new era of economic and social development and, more generally, a new and more evolving phase of human existence.
Fourth Industrial Revolution

Innovative technologies, platforms and systems like *Cloud Computing, Internet of Things, Big Data & Analytics, Blockchain, Artificial Intelligence, Augmented Reality & Virtual Reality, Advanced robotics & 3D printing e 5G* are the new enablers of the digital economy. With its pervasive diffusion in all sectors, digital economy promises to establish a new era of economic and social development and, more generally, a new and more evolving phase of human existence.

Through the introduction and implementation of these technologies and their widespread circulation, **human experience is undergoing deep changes**, characterized by a high degree of efficiency due to the use of data, to the return to local production paradigms and to integrated cloud systems which are capable of connecting people and objects, physical and virtual places.

This new phase, commonly referred to as the “*fourth industrial revolution*”, is characterized by a series of technological innovations designed to completely modify our habits, touching every aspect of our lives and improving their quality.

In this new and fast changing scenario, one of the key factors driving the development of the digital economy is **data**: a resource which can be fully exploited by means of appropriate analytical tools, and at the same time through a radical change in the organizational culture of businesses and public institutions.

While the third industrial revolution or Industry 3.0, following today’s classification, focused on the automation of the single machine and the single process, Industry 4.0 will focus on the **end-to-end digitization** of every physical asset and their integration within the digital ecosystem.

This process will go through several stages:

(i) **Digitization and integration of horizontal and vertical value chain.** Industry 4.0 will lead to the digitization and integration of vertical processes passing through the whole organization, from the development of a product to its acquisition thanks to production, logistics and sales systems.

(ii) **Digitizing the supply of products and services.** Product digitization will include the enlargement of the existing product line, for example by adding intelligent sensors or communication devices that can be used with data analytical tools, as well as through the creation of new digital products that will focus on fully integrated solutions. Thanks to new data collection and data analysis methods, companies will be able to obtain data about the product designed to improve it in order to better meet the
expectations of the final consumer, accordingly companies will associate their products with a range of services.

(iii) **Digital business models and direct “access” to consumer.** The main market leaders are already increasing their offer, being able to provide important digital solutions such as fully data-based services and solutions that are highly-integrated with platforms. Disruptive digital business models will be often focused on development of further digital revenues and on the optimization of interrelationships with the consumer.

**Data at the heart of the Digital Economy**

In this scenario, **data represents the engine of transformation of the digital economy** and it is considered the "lubricant" of the new economy. In particular, over the last few years, more than talking about simple or single data, it is common to refer to a much more complex and widespread concept, that could be explained with the term "**smart data**" or more commonly "**big data**". It is due to the huge amount of data which are available within the new digital ecosystem, high-speed produced and originated from a multitude of sources, whose management and analysis require new, more powerful and intelligent tools in terms of processors and algorithms.

The first element of data diversity compared to the past is represented by their quantity or **volume**: ordinary databases are no longer able to handle the ever-increasing number of data. Consequently, it has become necessary to develop new database models capable of storing, classifying and processing large amounts of data at supersonic speeds. The second aspect concerns the **ways of analyzing data** which, as already mentioned, definitely changed. Today we are talking about innovative and advanced analysis techniques: "**predictive analytics**", "**data mining**" and "**data science**".

The majority of doctrine focuses on the quantitative aspect concerning the volume of data which, for some authors, represents one of the main problem linked with this phenomenon. It is due to the fact that it is particularly complex to keep up with the pace of such exponential growth. According to estimates, in fact, every day about 2.5 exabytes are produced, which is a really huge number.

However, volume is not the only important feature. The **speed** at which data is generated and made accessible is likewise impressive. To provide a few examples, in 2016 from data which are generated on social networks, Facebook registers about 10 billion "like" every day from 1.09 billion users, while on Instagram more than 95 million images per day are shared with a daily average of 4.2 billion "like" and finally over 400 hours of uploaded content to YouTube are recorded every minute of every single day.
However, the speed of data development is a cause of concern particularly with regard to their management, especially when it far exceeds that of processing. Indeed, in order that data could guarantee best performances, they have to be analyzed at a speed which is comparable to the lightning speed which characterizes the information flow into the Enterprise Information Management systems. In microseconds it is necessary to decide whether it is worth acquiring a particular data, considering its relevance after having compared it or combined it with other available information.

In some cases, the big data phenomenon is also defined by the ability to analyse a variety of unstructured data sets from different sources such as: web logs, social media, smartphones, sensors, and financial transactions. This requires the ability to connect different sets of data, in addition to the ability to extract information from a combination of deconstructed dataset. The element of variety of data mainly refers to their structure: structured data enters into Enterprise Information Management systems when they are already marked with a tag and they are easy to sort; today, however, data are mostly unstructured, as random information, which are difficult to analyze and manage.

The three above mentioned features - volume, velocity, and variety - are considered the three major characteristics of the big data and are commonly referred to as the three V. They are "technical" properties that depend on data storage capacity and on processing technology.

The development of the analysis of the phenomenon has led to consider, in addition to the three above mentioned V, a fourth V namely the value, which strongly depends on the growing economic potential and on social value which, in turn, attributes to data the quality of "new production factors".

Concerning the existing classification modes, it is possible to distinguish different approaches to catalogue the data, with reference to:

(i) **Type of information**: data can be classified according to the different kinds of information they are able to provide and that, in turn, may have a different degree of utility for a certain enterprise. For example, data can provide information about individuals, their behaviors, their preferences, and their geographic position; about economic subjects regarding their turnover and the number of commercial transactions carried out; or even about objects, such as the real-time position or the speed of a car.

(ii) **Data Structure**: data can be differentiated depending on whether they are structured or not. From this peculiarity depends on the possibility to extract an economic value and the choice of the modality needed to do it. For structured data, that are numeric or financial data (so tables, records, office automation documentation, M2M data, etc.), is required an extraction model that defines a set of parameters such as: the reference sector; the typology of data concerning that particular sector; and finally the
relationship existing between the data itself. An example of structured data may be represented by a consumer database that contains information about his name, surname, address, age, telephone number, etc. Structured data can be more easily processed and used for commercial purposes than deconstructed data, especially with reference to traditional methods of processing.

(iii) **Collection Mode:** data can be collected in different ways and, depending on the method adopted, there is a different impact on their economic value. Data are mostly provided voluntarily and actively by users. Data can also be collected simply through freely available sources on the Internet (open source) or through the study of user behavior, even though there is no personal information about the user. Examples of the first type are the so-called generic search engines based on systematic collection and processing of each web page at their disposal - a technique known as "crawling". The second mode is also the most common one. In fact, a large number of enterprises trace the path made by the user from one web page to another - and sometimes even within a single page (technically it is possible to monitor which part of a web page is actually viewed by the user) - through different techniques. Nowadays, these techniques have been further developed in order to increase the study of user behavior on multiple mobile devices and not just on the web. Yet, some of the techniques used for this purpose can be easily overcome by users without any special effort, while some others are difficult to avoid. The data generated in that way can also (technically) be easily combined with data released voluntarily to a company in order to create highly differentiated user profiles. However, privacy laws can limit the field of application of this kind of approach. Lastly, data can be generated by deducing new information using existing data.

**A more and more “data driven” world**

Pervasive digitization sets, therefore, data as a catalyst for innovation. So it is becoming necessary that more and more organizations become data driven, adopting a holistic information management strategy, which should include and integrate new types of data and of management and analysis techniques. Data can influence positively a company's products and services improvement; they can enable companies to take advantage of new business opportunities; **finally, they can also be used to better serve users demand by providing them with customized services or products.**

The overall economic effects of the data-driven phenomenon are not entirely clear yet, however current studies show an evident positive effect on companies and public administrations productivity. Altogether, these results predict an increase in productivity of approximately 5-10% for economic subjects deciding to adopt a data-driven approach.
Nevertheless benefits are not limited to companies: an efficient use of Big Data Analytics by European governments and public administrations, could reduce administrative costs of approximately 15%-20% thanks to a higher efficiency resulting from greater transparency, greater tax revenues (due, for example, the provision of personalized services) and a lower risk of fraud or errors, tied more directly to the data analysis capabilities.

However according to McKinsey Global Institute, it is necessary to make human capital investments in order to train appropriate professionals and, more generally, raise awareness about the value of data in organizations. One of the biggest challenges to the spread of big data is represented, in fact, by the skill shortage, so that one of the main big data challenge depends on human resources and on the one hand, in spreading the "data culture" throughout the organization, and on the other, in developing more and more the ability to ask the right questions, recognize descriptive patterns of certain phenomena, predict behaviors and make assumptions based on facts and evidence supported by data.

**Distruptive technologies: Blockchain and Artificial Intelligence**

In this new data driven world, a particularly important role, especially due to their wide disruptive application potential, is covered by Blockchain and Artificial Intelligence which will affect more and more sectors and industries.

The Blockchain technology, originally used to manage Bitcoin virtual currency, is a distributed database that leverages peer-to-peer technology, enabling currency exchange, traceability and, above all, allows to the identity of subjects who carry out their operations. In other words, it is the accounting book in which all transactions made in Bitcoin have been recorded since 2009, and where the currency holders identity is “logged”.

Specifically, the blockchain protocol is characterized by a spread chain of information between multiple nodes within a network (public or private), composed by machines of varying power and connected to each other. Each node holds a copy of the log, consisting of a series of chained blocks, including information about the exchanges between those who participate in the network. The fact that this information is distributed simultaneously, synchronously on all network nodes, makes it impossible to delete data. Any transaction carried out in this environment is subject to a ‘confirmation’, or ‘consensus’, expressed through validation by solving a mathematical problem. This confirmation is then performed by a shared and accepted algorithm in the network’s ecosystem. Individual transactions can be made up in terms of smart contracts, IT objects whose conditions are self-evident and allow automatic executions between two or more shares, making it applicable to every market segment or sector such as health, retail, media, energy, and public administration.
Differently from banks platforms, that are owned by banks and can be managed only at centralized level, *blockchain* is an “open” and “free” technology that can be used also by privates and organized as a giant shared database. Everyone can download a “copy”, that is to access the system: identities are protected by intelligent encryption.

*Artificial Intelligence* refers to the ability of a computer to do tasks and reasonings typical of the human mind. This technological approach is emerging as the main catalyst that will push the development of innovative sectors and applications such as *connected car and smart home*, branding itself among the various technology enablers as a sort of super enabler of digital services. In its purely technological aspect, it includes theories and techniques for developing algorithms that allow machines to show ability and / or intelligent activities, at least in specific domains. Among the current and major application areas there are *virtual assistants* (such as Alexa of Amazon, Cortana for Microsoft, Siri for Apple, Home for Google, Google Assistant and Google Now as well as the latest Bixby of Samsung) providing assistance on generic requests, requests contextualized to obtain travel forecasts, business planning, etc. Another rising area of usage is *Conversational Interfaces*, especially *Chatbots*, whose applications embedded in messaging services, such as Messenger, are capable of self-guiding users through a series of questions in the solution of different tasks.

The rising interest in Artificial Intelligence is linked to the prospective of substituting humans in various roles and tasks. This perspective is based on the conviction that, as they will be used in resolving complex problems, these applications will evolve more and more, learning automatically from the data and inputs involved. However, there are many doubts about the real chances that these applications, oriented to the creation of the so-called "perfect knowledge" could really replace humans. For this reason, in literature is mentioned a new approach, called *Augmented Intelligence (AugI)*, aimed to increase the intelligence and the knowledge of an individual in order to support him in the decision-making process and in the acquisition of further knowledge, without an artificial system replacing him. The goal of these techniques is to provide potential answers or possible pathways to explore an issue, rather than merely show a list of responses sorted by relevance that do not allow to solve the problem raised.

AugI applications work as a personalized partner that provides support offering relevant answers, feedbacks and suggestions based on access to external data and information, and, above all by, it operates under the guide of the user, who will dynamically evaluate the relevance of what is suggested though sophisticated interaction modes that are simple and designed to "educate" the machine / application. Answers and suggestions are based on algorithms that calculate and display the probability associated with different potential responses, helping to understand why a given response was suggested or not. In addition, through user-supervised learning, the more these applications are used, the more they become accurate and reliable, tending to the ideal behaviour expected for this type of system.
Part II – Policy

Antitrust

Big data are used to produce value through knowledge. They are inputs that firms may use to improve their decision-making processes, create new goods or services, or produce other information-based products such as recommendations, reviews, or search results. This makes case-by-case analysis the most suitable for understanding the competitive impact of the big data held by each firm.

For example, when it comes to the relationship between big data and market power, a firm controlling big data does not necessarily hold (significant) market power in the markets for the generation and collection of digital data. Likewise, a firm controlling big data does not necessarily hold (significant) market power in the downstream markets for the production and distribution of information-based products and services. In other words, no generalization is possible. Similarly, whereas no extraordinary barriers to entry shelter the markets for the generation and collection of digital data, in some cases it could be true that big data represent barriers to entry protecting the markets for the production and distribution of information-based products and services. Yet, we should collect more empirical data to check whether, in each scenario, factors such as the first mover advantage, the feedback loop, or the network effects are stronger than other factors, such as non rival good, product differentiation, multi-homing, or decreasing data returns.

Then, when it comes to the relevant markets and the market power of multisided media platforms collecting big data - that is, of firms that gather digital data while they offer zero price products and use multi-sided business models - some issues must be distinguished. Firstly, not every firm amassing digital data is also a media multi-sided platform. Secondly, as a matter of practice, it is true that the identification of the markets where these platforms work becomes more complicated, because: (a)
zero-price products neutralize the use of the SSNIP test; and (b) multi-sidedness business models bring together two or more demands. Thirdly, as a matter of theory, zero prices and multi-sidedness do not prevent antitrust authorities from assessing the market power of multisided media platforms. Likewise, where big data are involved, market power can be appreciated in terms of quality and attention: namely, an increase in market power can result from a reduction in quality, and/or from an increase in the amount of the attention, also due to the amount and quality of the available data.

As to access to digital data, in the era of the Fourth Industrial Revolution they are ubiquitous, cheap to produce and distribute, non-excludable and non-rival in consumption. Still, it might be true that a firm happens to be the only one to witness/perceive an event registered into its digital data. Nevertheless, as the foreclosure theory suggests, such a de facto exclusivity (or even a legal exclusivity over some digital data) does not mean – or does not prove – that competitors are prevented from developing rival products and services by using fungible data. More in general, up to now, no antitrust authority has ever showed that firms holding big data leave their rivals without access to (fungible) digital data. In other words, no unilateral conduct or multilateral practice, merger included, has ever be found illegal because of its foreclosing effects. For example, in Facebook/WhatsApp the EU Commission cleared the merger because «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».

Then, one may discuss whether big data can be used to harm competition. In this regard, as to the algorithmic collusion run over big data, the main concern is related to the case of algorithms that reproduce oligopolistic interdependence. And this because antitrust provisions about agreements - that is, Section 1 of the Sherman Act and Article 101 of the TFEU - require that firms meet their minds, while oligopolistic interdependence does not amount to a multilateral practice. In addition, as to personalized prices calculated on the bases of big data, the main concern regards the case of dominant firms practicing perfect discrimination, because these discriminatory prices neutralize consumer welfare, but maximize total welfare. Thus, antitrust authorities should come up with a solution as to the role of consumer welfare and total welfare. Finally, as to the refusal to share big data, it is very hard to drive such a practice under the essential facility doctrine. In addition, any duty to share big data would be problematic and cumbersome to apply, because authorities would struggle to indicate factors such as the data to be shared, the timing of the sharing, and the privacy issues connected to this sharing.

Not by chance, indeed, one could analyze also the antitrust-privacy interface. It can be squared according to four different perspectives: a) data portability increases competition by reducing switching costs; b) no wonder if antitrust law intervenes to fine practices reducing the quantity and quality of products and services that are privacy-friendly or privacy-
enhancing; c) as the German Facebook case shows, dominant firms' practices aimed at collecting digital (personal) data may also be unfair within the meaning of EU competition law; d) privacy concerns may affect antitrust theories of harm when it is assumed that a privacy reduction entails a reduction in goods' quality which, in turn, implies a consumer welfare reduction.

Regulation

When markets (are prone to) fail ex ante economic regulations may work better that ex post antitrust actions. The Italian Communication Authority (AGCOM) is one of the regulatory authorities that looks with particular interest to the big data phenomenon for two main reasons. Firstly, according to the Electronic Communications Code (Legislative Decree No 259 of 1 August), the AGCOM must ensure market access to communications services which are the backbone of big data; and this access must occur in accordance with criteria of objectivity, transparency, non-discrimination and proportionality.

Secondly, one of AGCOM tasks is to promote information pluralism, that is a hotly debate issue also for its non-economic aspects. Nowadays, due to the development of new access modes to information (online content, social media), regulating pluralism requires policymakers' intervention, on the supply side as well as on the demand side. If the economic dimension (competition and dominant positions) are effectively tackled through antitrust analysis, the individual and social dimension (right to information) requires a specific regulation. The use of technologies and big data can help in order to better regulate information pluralism in traditional media as well as in the new media, taking into account all the aspects involved and not only the increasing phenomenon of fake news and post truth.

From an economic perspective, the lack of competence of “data supplier” consumers and “recipient of information and goods” consumers would create information asymmetries. In this regard, it is possible to suppose regulatory interventions aimed to define requirements for the exploitation of users’ personal data. This would be coherent with the traditional approach applied to privacy issues. However, this approach has proved not to be particularly effective over the years: interventions that increased duties to inform ended up being largely formal, inadequate to face the problem related to the proper consumers' information and conscious choices. In the digital environment, the overall picture is even more depressing. The belief that consent can be the main element of legitimacy also in the digital environment exposes any regulatory intervention to the tragic outcome of Sisyphus's effort, especially if we consider that, in comparison to business models exploiting personal data in order to provide behavioral advertising, no effective alternatives seem available.
Different expectations are related to measures that have been recently implemented in Europe to promote users’ mobility among platforms by reducing switching costs through the guarantee of data portability. Such a measure is intended to foster competition among platforms and may be useful to encourage competition on services quality. Furthermore, also to overcome the practical difficulties implied by the enforcement of the essential facility doctrine, other national regulators, such as the French one, has enacted new rules obliging private firms to open their data when they are of public interest.

**Interventions on pricing policies of platforms would be less desirable.** Indeed, it is worth recalling that the markets at issue are characterized primarily by the interrelations and the interdependence between the two (or more) sides of the platforms. Indeed, platforms must carry “both sides on board”, that is to intercept a sufficient number of economic agents on each side so that they can ensure a sufficient critical mass to feed the indirect network effects. For example, in the case of platforms providing search services or social networking, while services are offered free of charge to users’ side, in order to engage consumers, a tariff for the advertising spaces made available for behavioural targeting will be charged to advertisers’ side. This objective function often requires platforms to set asymmetrical prices to groups on different sides (skewed pricing), prices that do not reflect the actual cost of the service offered but take into account differences in demand elasticity. The non-neutrality of the price structure applied by the platform is functional to the need to feed both sides, so any regulatory intervention aimed to prohibit null prices could undermine not only the business model but the very existence of a platform.

**Increasing consumers’ awareness** about data value is a preliminary condition in order to set rules to balance the trade-off between disclosure and data protection. Several studies have raised questions, deriving mixed conclusions. Consumer education is mostly considered not sufficient to create a desirable balance between the need to protect data and the need to share information. Nevertheless consumer awareness is still a goal to be pursued, even if it involves costs and risks.

**Privacy**

Due to our increasingly digital reality, where data represent the lymph of the so-called knowledge economy, it is unavoidable to examine what role is assuming the specific category of personal data. Today, collecting and processing data related to identified or identifiable users of online services is a common practice in a growing number of sectors, both private and public ones. From insurance to healthcare, from banking and finance to advertising, to name a few, personal data processing is becoming more and more consistent. Meanwhile, this same digital reality is determining profound social changes, which can be defined in terms of knowledge society, where the representation
of ourselves is more and more linked to our virtual selves, that is, to the images derived from
the information pertaining to us, and that are available on the Internet. The virtual image of
each of us can variably represent the stratification of personal data revealed to build up the
image of who we are over time, together with personal data released without awareness, or
without a meaningful consent when it is the only option for accessing the services of data
controller.

(known as the General Data Protection Regulation, GDPR) on the protection of natural
persons with regard to the processing of personal data and on the free movement of such
data, substitutes the previous Directive 95/46 / EC. The Regulation presents both elements of
continuity and elements of discontinuity with respect to the Directive, the main difference
being legal and systemic: whereas the Directive introduced a mere harmonization system
requiring Member States to achieve a particular result without dictating the means of
achieving that result, the new Regulation, which is already in force and will be applied from
25 May 2018, is immediately enforceable in all its elements and for all the Member States.
This choice finds its reason in the transnational feature both of the economic transformations
related to the knowledge economy and of the public interest of the knowledge society in the
protection of personal data. The legal differences between the two legislations also result in
different criteria for establishing the applicable law in the framework of personal data
processing. The Directive had to identify the applicable law within the European Union as
well as outside its borders given the variances in the national implementations. The
Regulation, on the contrary, overrides all the national laws dealing with data protection, and
introduces a single law and common standards for all the Member States. In this context, it is
important to underline the broad scope of the territorial scope of the new Regulation: it
applies when data processing takes place “in the context of the activities of an establishment”
in the EU, irrespective of the fact that, as currently envisaged, the use of tools located in the
territory of the Union, such as servers, are located outside the EU. In particular, there is a real
legal innovation in the provision of Article 3, para 2, which introduces the applicability of EU
rules on data protection in the case of data controllers and processors being outside the EU
whose processing activities relate to the offering of goods or services (even if for free) to, or
monitoring the behavior of EU data subjects (within the Union) holders or persons responsible
for processing not established in the territory of the Union, bases the applicability of the new
rules on the basis of two different criteria: 1) the supply of goods or the provision of
services to persons concerned in the Union; and 2) monitoring the behavior of these
individuals within the Union for 'profiling' and therefore to make decisions about them,
analyze them or predict their preferences and behaviors.

The thorny issue of data transfer outside the Union is a fundamental part of the new
Regulation. The process has been improved by the removal of the need for prior national
authorization for transfers based on approved safeguards such as Commission or DPA approved contracts (Binding Corporate Rules). Nevertheless the DPA’s authorization will remain required in the event a Data Controller intends to use ad hoc contractual clauses (not expressly recognized as adequate by the European Commission) or administrative controls concluded with administrative public authorities. The European Legislator has confirmed within the new regulation the requirement of consent, as well as other rights connected to the acknowledgement of his/her rights (right of information and access) and control (processing limitation, withdrawal of consent, opposition, erasure and portability). The data controller, according to the new European regulation, can process personal data without a prior consent of the data subject, only in the event the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defense of legal claims, or in the event the processing of personal data may be considered compatible with the original purposes (in accordance with a compatibility assessment carried out by the controller considering a set of specific requirements set forth in Article 6.4).

In the GDPR the relationship between consensus, contractual obligation and purpose constitutes the paradigm to measure the lawfulness of the data processing carried out by the controller, but the new regulation leaves the doors open to further processing of personal data provided that there is an adequate legal ground and that the risks for the safeguard of the rights of the data subject are limited. One of the problems that remains unsolved, even after the application of the GDPR, is the difficult balance between the transparency and the minimization principle, set forth in Article 4 and 5 of the GDPR, and the difficult task of the controller in the formalization of a privacy policy that shall clearly define the purposes of the data processing and the duration of the data retention. The continuous evolution of the business models and collection of users’ personal data to respond, or better to predetermine, their needs make very complex the compliance of the controllers to the regulations included in the GDPR, both for the data already processed and for the design of new products and services, that more and more are based on the processing of Big Data.

Issues of compliance with the privacy regulation remain in all the processing phases of this type of data. Which kind of privacy policy and purpose shall be submitted to the data subject to be granted the consent for the collection of data considered non-personal data? The simple mention of the processing’s purpose as “analysis and statistical purpose” could be considered too general in light of the new regulation? Only courts and legal scholars may provide a solution to these questions and efficiently balance the inviolable rights of the data subject with the companies’ need of processing larger and larger amount of data, for purposes that are difficult to be prior defined.
Bibliography

5GPP, (2015), 5G Vision


AgID, (2017), Task Force sull’Intelligenza Artificiale, White Paper IA.


Anderson C., (2008), The end of theory: The data deluge makes the scientific method obsolete, Wired.


Argenton C., Prüfer J., (2012), Search Engine Competition With Network Externalities, 8 J. Competition L. & Econ. 73.


Autorité de la concurrence and Bundeskartellamt, (2016), Competition Law and Data.


Ben-Shahar O., (2009), The myth of the ‘opportunity to read’.


Bollier D., (2010), The promise and peril of big data.


Cambridge University Judge Business School, (2015), *Study on Personal Data Stores*, commissioned by DG CONNECT.


Centre for Media Pluralism and Media Freedom, (2013), *European Union competencies in respect of media pluralism and media freedom*.

Chun-Wei Tsai, Chin-Feng Lai, Han-Chieh Chao, Athanasios V. Vasilakos, (2015), Big data analytics: a survey.


Colangelo G., Maggiolino M., (2017), Data Protection in Attention Markets: Protecting Privacy through Competition?


Curry E. (2014), The Big Data Value Chain, European Data Forum


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.

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*.


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*.


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, (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), *Legal study on ownership and access to data.*


Evans D., (2011), The antitrust economics of free, 7 Competition Policy Int. 171.


FCA, (2016), Call for Inputs on Big Data in retail general insurance.


FTC, (2016), Big Data. A Tool for Inclusion or Exclusion?.


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


Graef I., (2016), EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility.


GSMA Intelligence, (2016), Global mobile trends.

GSMA, (2016), Resetting competition policy frameworks for the digital ecosystem.


iCom, (2016), *Internet of Things & 5G revolution*.


Maggiolino M., (2016), Big data e prezzi personalizzati, in Concorrenza e Mercato, pp. 95-138.


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.


Mik E., (2016), The Erosion of Autonomy in Online Consumer Transactions, 8 Law, Innovation and Technology 1.


Min Chen-Shiwen Mao- Yunhao Liu, (2014), Big data: a survey.


Narayanan A., Shmatikov V., (2008), *Robust de-anonymization of large sparse datasets*.


Noto La Diega G., Walden I.(2016), *Contracting for the ‘Internet of Things’: Looking into the Nest*.


OECD, (2013), *Supporting Investment in Knowledge Capital Growth and Innovation*.
OECD, (2013), *The role and measurement of quality in competition analysis.*


OECD, (2017), *Algorithms and Collusion - Background Note by the Secretariat.*

Office of Government Commerce (2011), *IT Infrastructure Library (ITIL) v.3*


Oracle (2014), *Information Management and Big Data, A Reference Architecture*


Polo M., (2010), *Concentration and Competition in Media Markets*, EUI.
Preibusch S., (2015), How to Explore Consumers’ Privacy Choices with Behavioral Economics.


PWC, (2016), Catalyst or threat? The strategic implications of PSD2 for Europe’s banks.

Pwc, (2016), Industry 4.0: Building the digital enterprise.


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.


Schatsky D., Muraskin C., Deloitte University Press, (2015), Beyond bitcoin, Blockchain is coming to disrupt your industry.


Sensortower estimates (2017), Data that Drive App Growth Report


Steadman I., (2013), Big data and the death of the theorist, Wired.


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.


The Boston Consulting Group, (2016), Inside OPS.


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)*.


Walker R., (2015), From Big Data To Big Profits: Success With Data And Analytics.


White House, (2015), Big data and differential pricing.


Williamson B., (2014), The death of the theorist and the emergence of data and algorithms in digital social research.

**About us**

**ITMedia Consulting**

ITMedia Consulting ([www.itmedia-consulting.com](http://www.itmedia-consulting.com)) is a research and consulting company working in the field of digital content and media, focused on the digital transition, content accessibility and convergence between media - internet - telecommunications. The company has improved its expertise by working with operators - incumbents, new entrants - and regulators on international markets and proposing itself as point of reference to face challenges of the digital environment and convergence.


**ASK - Bocconi University**

ASK, Centre for Research on Management and Economics of Arts and Culture Institutions, has been founded with the aim of carrying out researches in this field and promoting cultural planning and cultural institutions management activities.

The Centre is focused on activities and sectors significantly characterized by the relationship between Arts and Economics, and in particular:

- Economics, management, conservation and valorization of cultural goods and heritage, artistic and cultural traditions and productions.
- Economics, management and social promotion of cultural industries, such as publishing, media, and the film industry.
- Economics, management and valorization of leisure/entertainment and tourism activities.
- Urban and territorial valorization (Art cities)