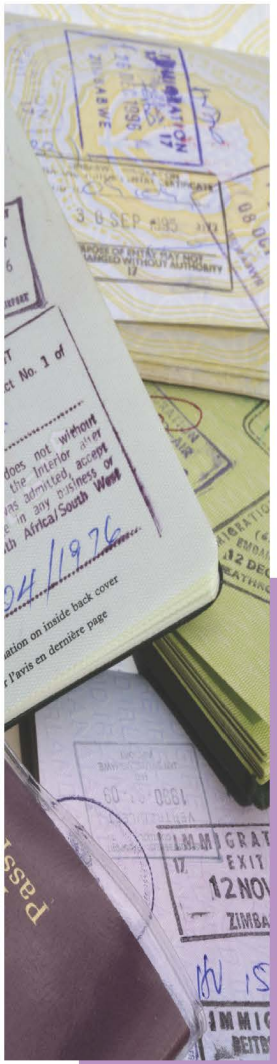
European Parliament

DIRECTORATE-GENERAL FOR INTERNAL POLICIES

POLICY DEPARTMENT C
CITIZENS' RIGHTS AND CONSTITUTIONAL AFFAIRS



Constitutional Affairs

Justice, Freedom and Security

Gender Equality

Legal and Parliamentary Affairs

Petitions

Big Data and Smart Devices and Their Impact on Privacy

Study for the LIBE Committee

EN

2015

DIRECTORATE GENERAL FOR INTERNAL POLICIES

POLICY DEPARTMENT C: CITIZENS' RIGHTS AND CONSTITUTIONAL AFFAIRS

CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS

Big Data and smart devices and their impact on privacy

STUDY

Abstract

The numerous debates triggered by the increased collection and processing of personal data for various - and often unaccountable - purposes are particularly vivid at the EU level. Two interlinked, and to some extent conflicting, initiatives are relevant here: the development of EU strategies promoting a data-driven economy and the current reform of the EU personal data protection legal framework in the context of the adoption of a General Data Protection Regulation (GDPR). In this context, and focusing on the development of Big Data practices, smart devices and the Internet of Things (IoT), this Study shows that the high degree of opacity of many contemporary data processing activities directly affects the right of the individuals to know what is being done with the data collected about them. This Study argues that the promotion of a data-driven economy should not underestimate the challenges raised for privacy and personal data protection and that strengthening the rights of digital citizens should be the main focus of the current debates around the GDPR.

DOCUMENT REQUESTED BY THE
COMMITTEE ON CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS (LIBE)

AUTHORS

Dr Gloria González Fuster, Research Professor at the Vrije Universiteit Brussel (VUB)

Dr Amandine Scherrer, European Studies Coordinator and Associate Researcher at the Centre d'Etudes sur les Conflits, Liberté et Sécurité (CCLS)

This Study was coordinated by the Centre d'Etudes sur les Conflits, Liberté et Sécurité (CCLS) and the Centre for European Policy Studies (CEPS) and conducted under the scientific supervision of Prof. Didier Bigo (Director of CCLS and Professor at Sciences Po Paris and King's College London).

The authors would like to express their gratitude to Prof. Dominique Boullier (Sciences Po Paris) and Prof. Evelyn Ruppert (Goldsmiths, University of London) who have provided their comments on an earlier draft. Any errors or omissions are the sole responsibility of the authors.

RESPONSIBLE ADMINISTRATOR

Mr Alessandro DAVOLI
Policy Department C - Citizens' Rights and Constitutional Affairs
European Parliament
B-1047 Brussels
E-mail: poldep-citizens@ep.europa.eu

LINGUISTIC VERSIONS

Original: EN

ABOUT THE EDITOR

Policy Departments provide in-house and external expertise to support EP committees and other parliamentary bodies in shaping legislation and exercising democratic scrutiny.

To contact the Policy Department or to subscribe to its monthly newsletter please write to:
poldep-citizens@ep.europa.eu

European Parliament, manuscript completed in September 2015.

This document is available on the Internet at:
http://www.europarl.europa.eu/studies

DISCLAIMER

CONTENTS

EXECUTIVE SUMMARY

EU citizens and residents and, more generally, all individuals deserving protection as 'data subjects' by EU law, are directly impacted by EU strategies in the field of Big Data. Indeed, the **data-driven economy** poses significant challenges to the EU Charter of Fundamental Rights, notably in the fields of **privacy and personal data protection**.

Big Data refers to the exponential growth both in the availability and automated use of information. Big Data comes from gigantic digital datasets held by corporations, governments and other large organisations; these are extensively analysed (hence the name 'data analytics') through computer algorithms. There are numerous applications of Big Data in various sectors, including healthcare, mobile communications, smart grids, traffic management, fraud detection, or marketing and retail (both on- and offline). The notion, primarily driven by economic concerns, has been largely promoted through market-led strategies and policies. Presented as an enabler of powerful analytical and predictive tools, the concept of Big Data has also raised numerous criticisms emphasising such risks as biased information, spurious correlations (associations that are statistically robust but happen only by chance), and statistical discrimination. **Moreover, the promotion of Big Data as an economic driver raises significant challenges for privacy and digital rights in general.** These challenges are even greater in a digital ecosystem with a proliferation of cheap sensors, numerous apps on mobile devices and **an increasingly connected world that sometimes does not even require human intervention** (as shown in the increasing development of the **Internet of Things** [IoT]). The flows of information on- and off line, shared and multiplied across computers, mobile devices, watches, SmartBands, glasses, etc., have dramatically increased the availability, storage, extraction and processing of data on a large scale. It has become increasingly difficult to track what is made of our data. This situation is complicated further by the wide variety of actors engaged in data collection and processing.

The numerous debates triggered by the increased collection and processing of personal data for various – and often unaccountable - purposes are particularly vivid at the EU level. Two interlinked, and to some extent conflicting, initiatives are relevant here: the development of EU strategies promoting a data-driven economy and the current reform of the EU personal data protection legal framework, in the context of the adoption of a General Data Protection Regulation (GDPR).

In order to address the issues at stake, the present Study provides an overview of Big Data and smart devices, outlining their technical components and uses (**section 2**). This section shows that **many contemporary data processing activities are characterised by a high degree of opacity**. This opacity directly affects the ability of individuals to know how data collected about them is used; it also hinders their capacity to assess and trust the manner in which choices are (automatically) made - whether, in other words, these choices are appropriate or fair. As regards smart devices, cheap sensors or the IoT, the pervasiveness of sensors and extensive routine data production might not be fully understood by individuals, who may be unaware of the presence of sensors and of the full spectrum of data they produce, as well as the data processing operations treating this diverse data. If Big Data, smart devices and IoT are often promoted as key enablers of market predictions and economic/social dynamics, data processing raises the question of who controls one's data.

In this perspective, Section 3 presents the different EU approaches on the digital economy and the questions raised in terms of privacy and personal data protection (**Section 3**). This section argues that in the current context of the development of a Digital Single Market for Europe (DSM), the European Commission's perspective is very much commercially and economically driven, **with little attention to the key legal and social challenges regarding privacy and personal data protection**. Even though the European Commission points out some of the key challenges of processing data for economic and market purposes (i.e., anonymisation, compatibility, minimisation), the complexity of these challenges is somehow under-estimated. These challenges can be grouped around the following questions any digital citizen may ask her/himself under EU law: which data about me are collected and for what purposes? Are data protected from unauthorised access and to what extent is control exercised upon the processing of my personal data?

Section 4 then considers these questions in the specific context of the **Data Protection Reform package**. Arguing that **the digital citizen's rights should be the main focus of the current debates around the GDPR, this Section underlines that Big Data, smart devices and the IoT reveal a series of potential gaps in the EU legal framework**, in the following areas in particular: transparency and information obligations of data controllers; consent (including consent in case of repurposing); the need to balance public interest and the interests of data subjects for legitimising personal data processing; the regulation of profiling; and proper safeguarding of digital rights in case of data transfers to third parties and third countries.

In light of these findings, the Study concludes with **key recommendations** for the European Parliament and, in particular, the **LIBE Committee** responsible for the protection of natural persons with regards to the processing of personal data. These recommendations aim at ensuring that negotiations around the GDPR promote a strong and sustainable framework of transparency and responsibility in which the data subject's rights are central.

In particular, **the guiding principle of any exploitation of personal data should be driven by the requirement of guaranteeing respect for the Fundamental Rights (privacy and personal data protection) laid down in EU primary and secondary law** (recommendations 1 & 2). The role of data controllers in this perspective is central as they are legally required to observe a number of principles when they process personal data, compliance of which must be reinforced. The degree of information and awareness of data subjects must be of prime concern whenever personal data processing takes places, and the responsibility for protecting Fundamental Rights should be promoted along the data production chain and gather various stakeholders. Furthermore, **the GDPR should ensure that individuals are granted complete and effective protection in the face of current and upcoming technological developments of Big Data and smart devices** (recommendation 3). The GDPR currently under discussion should in any case not offer less protection and guarantees than the 1995 Data Protection Directive, and users should remain in complete control of their personal data throughout the data lifecycle. Finally, **effective protection of individuals** cannot be guaranteed solely by the adoption of a sound GDPR. It will **also require a consistent review of the e-Privacy Directive** (recommendation 4), an instrument that not only pursues the safeguarding of personal data protection but, more generally, aims to ensure this right and the right to respect for private life.

1. INTRODUCTION

EU citizens and residents and, more generally, all individuals deserving protection as 'data subjects' by EU law, are directly impacted by EU strategies in the field of Big Data. Indeed, the data-driven economy poses significant challenges to the EU Charter of Fundamental Rights, notably in the fields of privacy and personal data protection.

Big Data refers to the exponential growth both in the availability and automated use of information. Big Data stems from gigantic digital datasets held by corporations, governments and other large organisations; these are extensively analysed (hence the name 'data analytics') through computer algorithms. There are numerous applications of Big Data in various sectors, including healthcare, mobile communications, smart grids, traffic management, fraud detection, or marketing and retail (both on- and offline). **The concept, primarily driven by economic concerns, has been largely promoted through market-led strategies and policies**.

The value of Big Data lies in the aggregation of personal data, which is how patterns and correlation are detected and become relevant and actionable (data mining, profiling, grouping, categorising, identifying outliers). Through mass aggregation and the development of computerised techniques making sense of this data, Big Data can be used to identify general trends and correlations but can also be processed in order to directly affect individuals.[1] Big Data can indeed target specific groups of people, sometimes first identifying and later focussing on groups not previously envisioned as such. In these processes, data related to concrete individuals might be used to ascribe other individuals to certain categories, influencing or determining decisions concerning the latter.

The rise of cheap sensors and mobile devices has led to an increasingly connected world, mobilising further data processing and fuelling the development of Big Data. This increasingly connected world is closely identified with the Internet of Things (IoT), a notion that relates to a cluster of objects that are readable and/or controllable via the Internet or other technologies, such as Radio-Frequency Identification (RFID), objects that can communicate with each other without human interference.

These technological developments have had a fundamental impact on our everyday life, from reinventing society to transforming notions of identity. Government policy-making has been influenced, a radical change in information production has been mobilized, and new economies have been created and formatted.[2] Presented as an enabler of powerful analytical and predictive tools, **the notion of Big Data has also raised numerous criticisms.** These criticisms have emphasised such risks as biased information and spurious correlations (associations that are statistically robust but happen only by chance).[3] **Moreover, the promotion of Big Data as an economic driver raises significant challenges for privacy and digital rights in general.**

Big Data has already sparked much interest and prompted debate across different disciplines on both sides of the Atlantic. In the United States Big Data has been depicted as

[1] WP29 Opinion 03/2013 on purpose limitation.
[2] E. Ruppert et al. (2015), "Socialising Big Data: From Concept to Practice", CRESC Working Paper Series, Working Paper no. 138.
[3] See "The Backlash against Big Data", *The Economist*, 20 April 2014.

one of the greatest public policy challenges of our time.[4] In 2013, Viktor Mayer-Schönberger and Kenneth Cukier famously described Big Data as "A Revolution That Will Transform How We Live, Work and Think".[5] Some scholars have coined the concept of *habitèle*, suggesting that mobile technologies reshape our everyday interactions (from coordination skills to mood changes) and offer us the opportunity of switching between social worlds.[6] This global environment of digital identities constitutes the *habitèle*, which is seen as nothing less than an anthropological mutation.[7] Other scholars speak of a conflict between (traditional) functional rationality and the 'digital reason' of contemporary 'algorithmic lives'.[8] Others still have described a shift towards a society of hyper-control based on mobile equipment.[9]

The value of data in general has undoubtedly increased due to the flows of information on- and offline, shared and multiplied across computers, mobile devices, watches, SmartBands, glasses, etc. By exploiting large data sets through advanced predictive analytics, **the processing of data enables the generation of new insights about how individuals live, work, travel, study, eat, or sleep, and how and what they consume**. The data we produce are now part of a digital 'ecosystem', opening up numerous avenues for corporations (market predictions, targeted advertising, etc.) and governments (e-Health, smart cities-related developments such as waste management and traffic predictions). **This digital ecosystem poses significant challenges when it comes to respecting such Fundamental Rights recognised by the European Union as the rights to privacy and to personal data protection.**

In a keynote speech given at "Recent Developments in Data Protection Law", a conference sponsored by the Academy of European Law (ERA), Giovanni Buttarelli, the European Data Protection Supervisor (EDPS), declared:

> In both the business and government spheres, there is a worrying drift towards thinking that, with regards to personal information, whatever is possible is also desirable: if personal data are available, they should be collected and stored indefinitely and exploited for any expedient purpose.
>
> … We need to find new ways for applying data protection principles to the latest technologies, be they Big Data, the internet of things, cloud computing, artificial intelligence, drones or robotics.
>
> This means placing the individual more firmly at the heart of technological development, through transparency, user control and accountability.[10]

The debates triggered by the **increased processing of personal data for various – and often unaccountable - purposes** are particularly vivid at the EU level. Two interlinked,

[4] C. Wolf (2015), "Envisioning Privacy in the World of Big Data" in Rotenberg, M., Horwitz, J. and Scott, J. (eds.), *Privacy in the Modern Age: The Search for Solutions*, New York, NY: The New Press, p. 204.
[5] V. Mayer-Schönberger and K. Cukier (2013), *Big Data: A Revolution That Will Transform How We Live, Work and Think*, London: John Murray.
[6] D. Boullier (2011), "Habitèle virtuelle: une nouvelle enveloppe pour commuter, notre téléphone portable", *Urbanisme* 376, pp. 42-44.
[7] D. Boullier (2014), "Habitele: mobile technologies reshaping urban life", *Urbe* 6(1), pp. 13-16.
[8] E. Sadin (2015), *La Vie algorithmique: Critique de la raison numérique*, Paris: Éditions L'Échappée.
[9] B. Stiegler (2015), *La Société automatique 1: L'avenir du travail*, Paris: Fayard.
[10] G. Buttarelli (2015), "Big data, big data protection: challenges and innovative solutions", keynote speech, ERA Conference, "Recent Developments in Data Protection Law", 11 May, at https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/EDPS/Publications/Speeches/2015/15-05-11_ERA_speech_EN.pdf.

and to some extent conflicting, initiatives are relevant here: the development of EU strategies promoting a data-driven economy and the current negotiations of the General Data Protection Regulation (GDPR). These tensions are particularly highlighted in a recent interview given by Robert Madelin, former director general of Directorate General (DG) CONNECT and now policy advisor to the Commission on Innovation.[11]

In this context, this Study argues that **the digital citizen and his or her rights should be the main focus of the current debates around the GDPR**. Concerns over Big Data, smart devices and IoT and their impacts on privacy are indeed critical aspects of our digital rights. **Big Data and smart devices do not represent the end of privacy and personal data protection. On the contrary, they call for a reworking of the EU privacy and personal data protection framework to ensure it is up-to-date and operational**. The digital citizen, a fundamental figure for conceiving politics and rights in relation to digital technologies, should be central.[12] In this sense, the concept of the digital citizen is broader than the notion of the 'data subject'. The digital citizen is more generally the subject of all rights that are relevant in the digital realm.[13] As this Study later describes, **the economic aspects of Big Data and the promotion of a data-driven economy have too often prevailed over social concerns or Fundamental Rights.**

A strong and sustainable framework of transparency and responsibility, in which the data subject's rights are central, is needed. In this perspective, one key aspect of the GDPR proposal of the European Commission currently under negotiation as part of the trilogue is **the requirement for the valid consent of data subjects for data processing to be legitimate.** Securing the digital citizen's rights is essential at the EU level, and the **GDPR can demonstrate and stress the EU's necessity when it comes to protecting its citizens' digital rights**.

In order to address the issues at stake, the present Study provides an overview of Big Data and smart devices, outlining their technical components and uses (**section 2**). This is followed by a presentation of the different EU perspectives on the digital economy and the questions raised in terms of privacy and personal data protection (**section 3**). These perspectives are then considered in the specific context of the Data Protection Reform Package (**section 4**). The concluding section (**section 5**) finally presents some key recommendations for the European Parliament (EP) in light of these findings.

[11] Interview given to Euractiv on September 1st, 2015. Available at: http://www.euractiv.com/sections/innovation-industry/madelin-i-have-been-appointed-job-be-creative-317192.
[12] E. Isin and E. Ruppert (2015), *Being Digital Citizens*, London: Rowman & Littlefield.
[13] The reference to digital citizens in this context is to be understood as broader than EU citizenship: just as both EU citizens and third-country nationals can be 'data subjects', they are all entitled to enjoy, under EU law, their digital rights.

2. UNDERSTANDING BIG DATA AND SMART DEVICES

KEY FINDINGS

- Big Data can be broadly depicted as the massive and rapid processing of data (through modern data analytics) in the search for information (including unforeseen information). The practice of data mining poses a significant challenge due to the degree of opacity characterising many contemporary data processing activities.
- Envisioned through the lens of Big Data, smart devices are singled out for their ability to further extend data mining practices. The production of data by smart devices can be quite varied (such as sensors planned for data capture); the pervasive and extensive routine data production of smart devices might not be fully grasped by individuals.
- Data mining practices may result in 'behavioural targeting' and further encourage a 'datafication' of society that poses significant challenges for privacy and digital rights in general. Due to such risks as statistical discrimination, there are calls for up-to-date regulations.

This Section introduces Big Data and smart devices, outlining their technical components and uses. The main challenge of Big Data from the perspective of privacy and personal data protection lies in the degree of opacity that characterises many contemporary data collection and processing activities (2.1). Simultaneously, the development of smart devices and IoT further feeds data mining practices and allows enhanced automated decision-making and 'behavioural targeting' that is generally not fully grasped by individuals (2.2). Within this context, practices and trends (such as cloud computing and the move towards 'quantified societies) are developed further, raising a set of concerns (2.3).

2.1. Big Data

The history of the term 'Big Data' can be traced back to the beginning of the 2000s. Initially popular with American companies, the term 'Big Data' reached scientific circles by the end of the decade. It eventually reached policy makers and the general public, which progressively envisioned Big Data as an economic and social driver.[14] Although in its original definition Big Data was commonly seen through '3 Vs' (volume, velocity and variety), over the years Big Data has expanded – primarily via the marketing strategies of private companies – to include '7 Vs' (the original three, plus viscosity, variability, veracity, and volatility).[15]

Big Data relies on data analytics that can process massive quantities of data in the search for information, including unforeseen information, which can *potentially* generate unexpected insights. Big Data is characterised by two basic features: first, the possibility of

[14] P. Delort (2015), *Le Big Data, Que Sais-Je ?*, Paris: Presses Universitaires de France.
[15] See on this subject L. K. Stapleton (2011), "Taming big data", *IBM Data Management Magazine*, 16(2), pp. 12-18; and K. C. Desouza, L. Kendra and K. L. (2014), "Big Data for Social Innovation" *Stanford Social Innovation Review*, 12(3), pp. 39–43.

accessing and using large quantities of data (meeting the conditions of the '3 Vs', namely huge size [volume], created in near real-time [velocity], and diverse [variety]), and, second, the use of data processing techniques that allow for the recognition of previously unidentified patterns. Such patterns can entail identifying correlations or detecting anomalies; on the basis of past and current data, these patterns *might* have a predictive quality in the sense that they aim to forecast what may still happen. Recent newspaper articles have, however, underlined the limits of these predictive analytics.[16]

Considering these two basic features together reveals Big Data's fundamental underlying postulate: the more data available to be processed, regardless of its apparent interest or value, the higher the chances that unexpected, and potentially valuable, information can be obtained. Because any data could lead to interesting information, all data is potentially salient. The progress of Big Data in recent years can be explained by the removal of two of the traditional obstacles to the development of data mining practices, namely the inability to store large quantities of data and the cost of computer power.[17] **Big Data's utility and relevance come from algorithms and advanced data processing techniques** that exploit computer power and vast quantities of data.

However, despite the apparent success of Big Data in winning over marketing and economic discourses, **the various drawbacks of Big Data-driven strategy and policy** include biases inherent to data and the risk of spurious correlation.[18] Furthermore, as later described, data mining poses a significant challenge to privacy and digital rights in general, including the risk of statistical discrimination.

In any case, the notion of Big Data is best understood as part of a wider movement of what has been called an on-going 'data revolution' embracing such closely related trends as digitisation and the linking and scaling-up of data into networked data infrastructures.[19] **The advance of Big Data goes hand in hand with the 'datafication' of society, or the increasing transformation into data of multiple aspects of the lives of individuals.[20]** The importance of the Internet in our societies and the related and widespread use of social media and data production by Internet users results in 'datafication' as online activities give rise to a far-reaching collection of data. If many Internet services originate massive datasets that feed Big Data practices, these services often rely themselves on modern data mining practices to function.

The result of data mining practices predominantly manifests itself online through advertising or so-called 'behavioural targeting'. **This practice, which illustrates the degree of opacity characterising many contemporary data processing activities, reveals some of the challenges posed by Big Data**. Indeed, even though behavioural targeting is widespread, the logic behind the particular choice of adverts shown to each individual is generally unknown to the user of online services. **This opacity directly affects the ability of individuals to know how data collected about them is used; it also hinders their capacity to assess and trust the manner in which**

[16] For the health sector, see K. Leetaru (2014) "Why Big Data Missed the Early Warning Signs of Ebola", *Foreign Policy* (26 September), http://foreignpolicy.com/2014/09/26/why-big-data-missed-the-early-warning-signs-of-ebola; for the field of security, see A. Edwards (2015), "Big Data, Predictive Machines and Security: Enthusiasts, Critics and Sceptics", *Discover Society* 23, at http://discoversociety.org/2015/07/28/big-data-predictive-machines-and-security-enthusiasts-critics-and-sceptics/

[17] T. Craig and M. E. Ludloff (2011), *Privacy and Big Data,* Sebastopol, CA: O'Reilly, p. 5.

[18] See "The Backlash against Big Data," *The Economist*, 20 April 2014.

[19] In this sense, see R. Kitchin (2014), The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences, Los Angeles, CA: SAGE, p. xv.

[20] For instance, see V. Mayer-Schönberg and K. Cukier (2013), *Big Data: A Revolution That Will Transform How We Live, Work and Think*, London: John Murray, p. 73.

choices are (automatically) made - whether, in other words, these choices are appropriate or fair. In this sense, an April 2015 paper on the interaction between user behaviours and Google advertising claimed that female users were shown fewer instances of an ad for high-paying jobs.[21] Another case exemplifying some of the limitations of Big Data is the July 2015 revelation that the service Google Photo Tags, which uses facial recognition software to automatically tag pictures, labelled the portraits of some African-Americans as 'gorillas'.[22] A White House report on Big Data and Privacy released in May 2014 underscored the risk of algorithms being biased against certain groups due to the fact that their input data explicitly or implicitly encodes for a protected characteristic like gender or race.[23] The report refers to a study that found that web searches with black-identifying names (e.g., Jermaine) were more likely to display ads with the word 'arrest' than those using white-identifying names (e.g., Geoffrey). Social scientists have warned against these discriminatory risks, on numerous occasions arguing that statistical discrimination helps reproduce and legitimise social inequalities.[24]

While Big Data is advertised as a massive opportunity for corporations and governments, it raises significant issues not only in terms of efficiency when it comes to predictive analytics but also in terms of social discrimination and, more generally, EU Fundamental Rights (notably privacy and data protection). Smart devices and the IoT, which feed data mining practices further, undoubtedly amplify these risks for digital citizens.

2.2. Smart devices

Smart devices are electronic tools capable of operating interactively and autonomously; they are usually networked. The term may actually refer to many different types of pieces of electronic equipment, ranging from devices that are principally manipulated by individuals, such as smartphones, to the constitutive elements of so-called 'ubiquitous computing', that is, an environment with pervasive sensors and information-processing capability.

In this context, smart devices are also closely related to the IoT, a notion which had already originated in the 1980s. IoT relates to a cluster of objects that are readable and/or controllable via the Internet or other technologies such as RFID; these objects sometimes communicate with each other without human interference. In some cases, smart devices can be everyday objects, simple 'things' that beforehand lacked electronic components but that are now embedded with sensors and microprocessors. The adjective 'smart' is sometimes understood as referring to the fact that a device enjoys 'machine learning' and adaptive capabilities, allowing it to programme its activity on the basis of gathered data.

When considered through the lens of Big Data, the prime interest of smart devices is that they can feed data mining practices and make use of the information

[21] A. Datta, M. C. Tschantz and A. Datta (2015), "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination", *Proceedings on Privacy Enhancing Technologies*, 1, pp. 92–112, DOI: 10.1515/popets-2015-0007.

[22] See "Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software", *Forbes*, 1 July 2015, at http://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/.

[23] Executive Office of the President (2014), "Big Data: seizing opportunities, preserving values", at https://www.eff.org/files/2014/05/01/big_data_privacy_report_may_1_2014.pdf

[24] O. Gandy (2009), *Coming to Terms With Chance: Engaging Rational Discrimination and Cumulative Disadvantage*, Farnham: Ashgate; D. Bigo (2013), "Sécurité maximale et prévention? La matrice du futur Antérieur et ses grilles", in B. Cassin (ed.), *Derrière les grilles: sortir du tout évaluation*, Paris: Fayard/Mille et Une Nuits; and D. Wright and R. Kreissl (eds.) (2015), *Surveillance in Europe*, Abindon: Routledge.

obtained through them; most smart devices actually do both, simultaneously. They are certainly one of the key enablers of the current data explosion.

The production of data by smart devices can occur in quite varied ways. Smart devices often include sensors designed for data capture. These generate continuous streams of data such as temperature, waves, movements or other variables. Further data might also be created through information processing. In any case, smart devices generate data that inevitably includes indexical data, that is, data allowing for the identification of the produced data sets and for the linking of data sets with other data sets. **The pervasiveness of sensors and extensive routine data production might not be fully understood by individuals, who may be unaware of the presence of sensors (which are often low-cost and miniscule) and of the full spectrum of data they produce, as well as the data processing operations treating this diverse data**. Smart devices are sometimes labelled 'everyware', alluding to the fact they embody the colonisation of everyday life by information technology.[25]

The previous decades already witnessed the rise of personal devices relying on the processing of vast quantities of data. In the 1990s the first mobile telephones with Internet connectivity and the first email-enabled mobile phone systems appeared. In the 2000s, smartphones opened the door for the widespread use of mobile web applications, both producing and using great quantities of data, including data about the device's location. Since 2010, tablet devices have also contributed to the proliferation of said applications.

Wearable devices, that is, accessories or clothing incorporating advanced electronic technologies, are special types of smart devices. Although their origins could be found in the 1980s calculator watch, 'wearables' today most often integrate connectivity features. Modern smartwatches, for instance, typically run mobile applications and might also function as mobile phones. The paradigmatic example of a smart wearable device is the Google Glass prototype (marketed from April 2013 to January 2015), an optical head-mounted display developed by Google that allows wearers to communicate with the Internet through voice commands.

2.3. Practices and trends

A series of practices and trends, developing alongside the evolution of Big Data and smart devices, are particularly relevant in assessing the impact of Big Data and smart devices on privacy and personal data protection.

2.3.1. Cloud Computing

Collecting and processing data involved in Big Data benefits from, and supports, the progress of cloud computing, enabling ubiquitous network access to a shared pool of computing resources.

Cloud computing and storage solutions provide the basic supporting infrastructure for the data processing necessary for large-scale data analytics. They can thus be considered as **key enablers of Big Data.** The use of cloud computing, however, also results in specific privacy and security challenges. Such challenges are directly related to the quantities and quality of the data processed, the potential involvement of many different actors in their processing, the connected proliferation of copies of data generated in their processing, and

[25] A. Greenfield (2006), *Everyware: The Dawning Age of Ubiquitous Computing*, Berkeley, CA: New Riders, p. 33.

the possible multiplication of relevant jurisdictions.[26] These parameters put additional pressure on ensuring the security of the data processed, which requires both organisational and technical measures.[27]

2.3.2. A variety of (expanding) domains

Big Data today appears to interest, in one way or another, most large corporations and governments. This interest occurs in a myriad of sectors, including scientific research. Big Data attracts 'big industry' but also many other actors are drawn to it by its 'big benefits' in many areas, from healthcare management to the detection of fraudulent payments.[28] Companies particularly active in this field include Google, Facebook, Oracle, eBay and Amazon, that is, companies primarily dealing with information processing, but also business devoted to retail, banking or real estate. Additionally, Big Data sometimes crosses the path of Open Data, a notion calling on the free use and re-use of data by anyone, including ordinary citizens.

In the last few years, the IoT has developed mainly in three sectors: homes and buildings (monitoring and controlling); automobiles and transport (the 'smart car' and other applications); and health (including self-tracking, clinical remote monitoring and personal environment monitoring).[29] Because these relatively established sectors are not isolated from one another, practices often build on different uses. In this sense, many care applications can be regarded as falling under the paradigm of both the smart home and e-health.

An example of a recent development in the field of smart devices is 'intelligent toys', which in some cases can be spoken to and in others allow homeowners to turn on or off appliances or lights. Google Inc. filed for a patent in the US for a teddy bear and a toy rabbit equipped with cameras in the eyes and microphones in the ears; both toy animals have integrated speakers and are connected to the Internet, which allows for individuals to interact with home equipment.[30] **The nature of the objects, however, raises the question of whether they could be used to surreptitiously monitor people, and more concretely children.** The vulnerability of minors to smart devices was illustrated by the reactions in March 2015 to the announcement of a doll that uses voice-recognition technology and sends private recordings of children to third parties.[31]

A concept that tellingly brings together Big Data and smart devices is the notion of the 'smart city', or the use of digital technologies to enhance urban services. The services that could be improved in smart cities, normally by data-driven systems, include transport and traffic management, energy, health care, water or waste management, but also law enforcement.

---

[26] D. Bigo et al. (2012), "Fighting cyber crime and protecting privacy in the cloud", Study for the European Parliament's Committee on Civil Liberties, Justice and Home Affairs (LIBE), PE 462.509, Brussels.
[27] The EU Agency for Network and Information Security (ENISA) has been particularly active in mapping security and defining security approaches for 'going cloud'. See, for instance, ENISA (2015), *Security Framework for Governmental Clouds*,
[28] As pointed out in O. Tene and J. Polonetsky (2013), "Big Data for All: Privacy and User Control in the Age of Analytics", *Northwestern Journal of Technology and Intellectual Property*, 11(5), pp. 239-273.
[29] M. Swan, (2012), "Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0," *Journal of Sensor and Actuator Networks*, 1, pp. 217–53, at p. 218.
[30] Patent# US20150138333, United States Patent and Trademark Office.
[31] "Privacy Fears over 'Smart' Barbie that Can Listen to Your Kids", *The Guardian*, 13 March 2015, at http://www.theguardian.com/technology/2015/mar/13/smart-barbie-that-can-listen-to-your-kids-privacy-fears-mattel.

2.3.3. Quantified societies

The deployment of smart devices is partly indebted to individual self-tracking practices, whereby individuals decide to adopt self-measuring gadgets, often for health or leisure purposes. In this context, **the Quantified Self movement aims to incorporate technology, combining sensors and computing technology, for data acquisition in multiple aspects of everyday life**. Taking the idea even further, life-loggers aim to fully capture their lives.

The popularisation of self-quantifying practices has led to the use of such expressions as 'quantified communities' or even 'quantified societies'.[32] Such expressions, however, might be misleading to the extent that self-tracking is, predictably, only practiced by a portion of the relevant community enjoying the material conditions that allow for them to engage in digital self-measurement, leaving the others 'un-quantified' and potentially unaccounted. Although these tensions are not unique to Big Data, **they raise the question of the data (re)distribution and the inclusion/exclusion of individuals and groups in Big Data practices**. If data mining analytics can reach or include subjects that other data practices might have missed, they could also sideline or discriminate against groups and individuals. Furthermore, as stated earlier, the manner in which data is extracted and processed might not be fully grasped by individuals, who may not be necessarily aware of what is done with their data, thus raising the question of data ownership.

The techniques, uses and trends described above thus pose significant challenges to privacy and digital rights in general. In particular, the practice of data mining and extraction raises the question of the degree of **opacity that characterises many contemporary data processing activities**. If Big Data is often promoted as a key enabler for market predictions and economic/social dynamics, **data processing raises the question of who controls one's data**. In this perspective, and as detailed below, transparency is central in a digital ecosystem for both businesses and citizens.

[32] See, for instance, A. Croll (2012), "Big Data Is Our Generation's Civil Rights Issue, and We Don't Know It: What the Data Is Must Be Linked to How It Can Be Used', *O'Reilly Radar*, 2 August, at http://radar.oreilly.com/2012/08/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it.html.

3. DIGITAL ECONOMY AND PRIVACY: EU PERSPECTIVES

KEY FINDINGS

- While advocating the economic benefits of a data-driven economy, the European Commission has reiterated its concerns over the need for adequate data protection and infrastructure security, underlining the need for 'a high level of trust'.
- However, in the current context of the development of the Digital Single Market for Europe (DSM), the European Commission perspective is very much commercially and economically driven, with little attention to the key legal and social challenges regarding privacy and personal data protection.
- As discussed in the last few years by EU data protection authorities (especially the EDPS and the Article 29 Data Protection Working Party [WP29]), these challenges will prove hard to overcome in the current Data Protection Reform Package.

In May 2015, the European Commission set out a series of initiatives in the context of the Digital Single Market for Europe (DSM), laying the groundwork for Europe's digital future.[33] The strategy follows the Communication on the data-driven economy issued in 2014 in response to the European Council's conclusions of 24-25 October 2013 (EUCO 169/13) which called for EU action to provide the proper framework conditions for a single market for Big Data and cloud computing.[34] The overarching argument for a renewed EU strategy in the field is to 'release the digital economy's full potential' across the EU and globally. While the issue of impact for privacy and personal data protection will be addressed more specifically in Section 4, in the context of the current negotiations of the upcoming GDPR, this section takes stock of EU initiatives in the last decade in the fields of Big Data, the IoT and cloud computing, assessing to what extent EU strategy adequately addresses the challenges. In particular, this section argues that **the economic aspects of Big Data have too often prevailed over social aspects.**

3.1. From a data-driven economy to a digital single market: the EC perspective

While advocating the economic benefits of a data-driven economy, in its 2014 Communication the European Commission reiterated its concerns over the need for adequate data protection and infrastructure security, underlining the need for 'a high level of trust'. However, a careful analysis of the 2014 and 2015 Communications shows that the European Commission position is very much driven by commerce and economics, paying little attention to key legal and social challenges. While Big Data is presented as a market opportunity not to be missed, privacy and data protection, as well as the previously mentioned risks about Big Data, are addressed only marginally.

[33] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A Digital Single Market for Europe, Brussels, COM(2015) 192 final.
[34] Communication from the Commission to the European Parliament, the Council, the European Economic and social committee and the committee of the regions: Towards a thriving data-driven economy, Brussels, COM(2014) 442 final.

3.1.1. Big Data: an enabler for productivity and better services

From a market perspective, Big Data offers countless potential benefits, from an increased provision and efficiency of services to monitoring climate change, health trends and disease epidemics, as well as preventing government fraud and waste.[35] The European Commission view is that collecting, analysing and processing data lead to 'better results, processes and decisions', helping generate new ideas or solutions or more accurately predicting future events. 'Data-driven innovation' (DDI) refers to the "capacity of businesses and public sector bodies to make use of information from improved data analytics to develop improved services and goods that facilitate everyday life of individuals and of organisations, including SMEs [small and medium-sized enterprises]".[36] The 'data value chain' is thus given high priority.

This given priority can be traced back to the 2012 restructuring of the DG Information Society, which became DG CONNECT. Under this new banner, the activities of the DG were refocussed, reorganised and regrouped with the establishment of a new unit group: Data Value Chain (Unit G3). This Unit's mission is to foster commercial and social added value based on the intelligent use, management and re-use of data sources in Europe. This would involve a combination of research and innovation and legislative and deployment actions. An analysis of DG CONNECT's 2013 data value chain strategy helps to better understand the Commission's approach outlined in its 2014 and 2015 Communications.[37]

At the core of the data value chain is the process of extracting value from data, which presupposes that huge amounts of different types of data from a high number of various types of sources are processed efficiently. By building on the intelligent use of data sources across the EU member states, the data value chain strategy aims at extracting the maximum value from data to provide benefits for the economy and citizens. Three guiding principles filter through all the various segments and dimensions of the proposed data value chain strategy for Europe:

- a wide availability of good-quality data, including the free availability of publicly-funded data;
- the free flow of data across the EU as part of the digital single market;
- finding the right balance between potential privacy concerns for individuals and exploiting the potential reuse of their data.

The aggregated value of data is thus at the core of the data value chain concept. The EC 2014 communication clearly refers to the concept and takes on board the DG CONNECT strategy.

Ensuring the proper infrastructure for a data-driven economy is another key priority of the 2014 and 2015 European Commission Communications. Adopting a European Cloud Computing Strategy and establishing a European Cloud Partnership (ECP) are perceived as critical economic boosts and effective mechanisms to ensure European 'data sovereignty' in

[35] K. C. Li, H. Jiang, L. T. Yang, and A. Cuzzocrea (eds.) (2015), *Big Data: Algorithms, Analytics and Applications*, Boca Raton, FL: Chapman & Hall/CRC Big Data Series.
[36] Communication from the Commission to the European Parliament, the Council, the European Economic and social committee and the committee of the regions: Towards a thriving data-driven economy, Brussels, COM (2014) 442 final.
[37] DG CONNECT/Unit G3 - Data Value Chain, 7 November 2013, at https://ec.europa.eu/dgs/connect/en/%20content/data-value-chain-european-strategy.

the cloud.[38] Concerning the IoT, the 2014 EC Communication briefly states: "A series of large-scale projects will be funded to tackle the emerging questions of availability, quality and interoperability related to data gathered through smart connected objects and other IoT technology".

The subsequent DSM is built around three pillars: better access for consumers and businesses to online goods and services across Europe; creating the right conditions for digital networks and services to flourish; and maximising the growth potential of the European Digital Economy. The second pillar adresses the issues of cloud computing, Big Data and IoT, which are defined as 'central to the EU's competitiveness'. The EC highlights here restrictions to the free movement of personal data within the EU and promotes the removal of any unnecessary restrictions regarding the location of data within the EU.

3.1.2. Big Data opportunities and personal data protection

As regards personal data and consumer protection, the 2014 European Commission Communication makes clear that the fundamental right to personal data protection applies to Big Data when the data processed can be qualified as personal. Referring to the Commission's Data Protection Reform package (which will be analysed in the following section), the Commission underlines that it will work with member states and stakeholders to ensure that business, and in particular SMEs, receive adequate guidance, notably on issues such as data anonymisation and pseudonymisation, data minimisation, personal data risk analysis, as well as tools and initiatives enhancing consumer awareness. The European Commission also announces its support to projects aiming to regulate personal data breaches and to ensure that data is used in a manner compatible with its initial collection, recognising that "these measures will build the trust that is necessary to exploit the full potential of the data-driven economy".

Even though the Commission points out some of the key challenges of processing data for economic and market purposes (anonymisation, compatibility, minimisation), the complexity of these challenges is somehow under-estimated. In particular, **the successive European Commission Communications fail to clarify what is personal data and how the personal data life cycle may contain conflicting purposes and priorities**, opening up thorny debates for the future and the effective implementation of the digital single market.

The 1995 Data Protection Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data defines 'personal data' as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity".[39]

However, as the EDPS noted in its 2014 opinion on the subject, **it is now rare for data generated by user activity to be completely and irreversibly anonymised**.[40] **Therefore, the EC position according to which non-personal data, once recorded,**

[38] See Trusted Cloud Europe final report, 2014, at http://ec.europa.eu/digital-agenda/sites/digital-agenda/files/discussions/TrustedCloudEurope_3.pdf.
[39] Directive 95/46/EC, Article 2(a).
[40] EDPS Preliminary Opinion (March 2014): "Privacy and Competitiveness in the Age of Big Data: The Interplay between Data Protection, Competition Law and Consumer Protection in the Digital Economy".

can be re-used many times and escape the levels of protection of personal data is not as clear-cut as it seems. In its Opinion on anonymisation techniques on the Web, WP29 considers it a misconception for many data controllers to equate pseudonymisation with anonymisation.[41] This is because pseudonymised data still allows an individual data subject to be singled out and linkable across different data sets. Therefore, it can be concluded that in most instances pseudonymised data remains subject to data protection rules. Accordingly, all privacy and data protection principles fully apply. This aspect will be addressed further in Section 4.

Furthermore, personal data can be collected in a variety of ways, each of which raises specific challenges and entails potential abuses that undermine the trust of the citizen/consumer. As described in an OECD report:

- Data can be *volunteered* or *surrendered* by individuals when they explicitly share information about themselves or about third parties (e.g., when someone creates a social network profile, enters credit card information for online purchases, provides his/her personal information as a condition of registration to a given on-line service, or posts information about a friend, colleague, family member, etc.).
- Data can be legally *observed*, captured by recording the activities of users – in contrast to the data users volunteer (e.g., Internet browsing preferences, location data when using cellular mobile phones or telephone usage behaviour).
- Data can *inferred*, based on the analysis of personal data (e.g., credit scores can be calculated based on a number of factors relevant to an individual's financial history).[42]

Each of the steps of the personal data 'lifecycle' raises, moreover, distinct and significant challenges and involves different stakeholders from the collection/access stage, the process of storage and aggregation, the stage of analysis and distribution, up until the stage of data usage.[43]

As underlined by the European Parliament's Committee on Industry, Research and Energy (ITRE) in its draft motion for the resolution "Towards a thriving data-driven economy", while there are evident social and economic benefits associated with Big Data, there is an urgent need to tackle the challenges it raises.[44] ITRE importantly points out that **data protection and Big Data opportunities are not mutually exclusive**. In light of the above findings, it is unclear if the digital economy strategies devised by the European Commission have sufficiently taken into account these challenges.

3.2. Digital rights and welfare in a data-driven economy

EU citizens and residents and, more generally, all individuals protected as 'data subjects' by EU law, are directly impacted by EU strategies in the fields of the data-driven economy and

[41] WP29 Opinion 05/2014 on anonymisation techniques onto the web.
[42] OECD (2013), Working Party on the Information Economy & Working Party on Information Security and Privacy, "Exploring the economics of personal data: A survey of methodologies for measuring monetary value", DSTI/ICCP/IE/REG(2011)2/FINAL, p. 10, at http://http://www.oecd.org/officialdocuments/publicdisplaydocuemntpdf/?cote=DSTI/ICCP/IE/REG%282011%292/FINAL&docLanguage=EN.
[43] The personal data lifecycle is clearly described in OECD, Working Party on the Information Economy & Working Party on Information Security and Privacy, "Exploring the economics of personal data: A survey of methodologies for measuring monetary value", DSTI/ICCP/IE/REG(2011)2/FINAL.
[44] Committee on Industry, Research and Energy (ITRE), Draft motion for a resolution on Towards a thriving data-driven economy (2015/2612(RSP)), 10.4.2015, RE\1057329EN.

Big Data. According to EU primary laws, consumer welfare is an ultimate goal and "consumer protection requirements are to be taken into account in defining and implementing other Union policies and activities" (Treaty on the Functioning of the European Union [TFEU], Article 12). In parallel, the Charter of Fundamental Rights of the European Union enshrines respect for private and family life (Article 7) and personal data protection (Article 8) as Fundamental Rights. On the other hand, the Committee on Civil Liberties, Justice and Home Affairs (LIBE) is the European Parliament's Committee responsible for protecting natural persons with regards to the processing of personal data.[45] **These principles and priorities should lead to repositioning the safeguarding of rights, and in particular digital rights, at the centre of the analysis.** This sub-section aims at exploring the challenges derived from Big Data and a data-driven economy in general. Despite numerous opinions and recommendations from the EDPS and WP29, these challenges have been addressed only marginally by the European Commission.[46] These challenges can be grouped around the following questions any digital citizen may ask under EU law: what data about me is collected and for what purposes? Is my personal data protected from unauthorised access and to what extent is control exercised upon its processing?

3.2.1. The knowledge asymmetry: what data is collected and for what purposes?

The scale of data collected for commercial and economic purposes allows for tracking and profiling. **The process of aggregation implies that data is often combined from many different sources and that it is used and/or shared by many actors and for a wide range of purposes**.

An illustrative example comes from the smart metering system on which the EDPS issued a detailed recommendation in 2012.[47] The EU aims to replace by 2020 at least 80% of electricity meters with smart meters in order to foster smart grids that automatically monitor energy flows and adjust to changes in energy supply and demand accordingly. According to the European Commission, smart gas and electricity meters installed in the homes of energy consumers could reduce carbon emissions in the EU by up to 9% and annual household energy consumption by a similar amount. As the EDPS stresses, a key feature of smart meters is that they provide data via remote communications from the meter to energy suppliers, network operators and other third parties. As a result, there is a significant increase in the amount of energy-consumption data available to the consumer but also to third parties. Given the amount of information smart meters can amass, the EDPS notes that the potential for extensive data mining is thus very high: patterns can be tracked at the level of individual households but also for many households taken together, aggregated and sorted by area, demographics, and so on. Furthermore, if these patterns can be useful in analysing energy use to improve energy conservation, patterns and profiles can be used for many other purposes, including marketing and advertising. From the consumer perspective, what the smart metering system illustrates is **the need for a clear description of the key data processing operations, their purposes and the**

[45] Rules of procedure of the European Parliament, 8th Legislature, April 2015, Annex XI, Section XVII.
[46] See EDPS Recommendation (2012) on Smart metering system; EDPS Preliminary Opinion (2014) on Privacy and competitiveness in the age of Big Data: The interplay between data protection, competition law and consumer protection in the Digital Economy; EDPS Report of Workshop on Privacy, Consumers, Competition and Big Data (June 2014); WP29 Opinion 02/2013 on apps on smart devices; Opinion 03/2013 on purpose limitation; Opinion 05/2014 on anonymisation techniques onto the web; Opinion 06/2014 on the "Notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC"; Opinion 8/2014 on the Recent Developments on the Internet of Things; Statement of the WP29 on the impact of the development of Big Data on the protection of individuals with regard to the processing of their personal data in the EU (09.2014).
[47] EDPS 2012 Recommendation on smart metering system.

categories of data needed to achieve those purposes. This need for transparency is vital to ensuring well-informed consumer choice and consent.

In its 2014 Preliminary Opinion on Privacy and Competitiveness in the Age of Big Data, the EDPS further specifies on this aspect that choice depends on the availability of competing services and a consumer's ability to understand the information provided about those services.[48] The EDPS notes that several **obstacles can undermine the quality of choice, such as the difficulty in predicting what exactly will be done with one's personal data or lengthy and user-unfriendly 'privacy policies'**. The EDPS's Preliminary Opinion concludes that "this creates an asymmetry of knowledge which invokes the obligations of traders to provide clear and unambiguous information under EU consumer protection law, and calls into question whether data subjects have sufficient information to give informed consent to data processing".[49] However, if further transparency would be a welcome step in the field of data processing, **the guarantee that the data subject's rights are protected is equally critical in promoting a sound and sustainable business model**.

The EDPS highlights the fact that **the problem is likely to be compounded by the growth of the IoT, which will include many technical or embedded devices collecting personal data**. In many cases users will be 'unable to consult the privacy policy on the device itself, but would have to find paper documentation or more likely browse from another device to the relevant web sites'.[50] The challenges raised by the IoT have been similarly addressed by WP29, which has in particular underlined how the IoT undermines further **"the quality of the user's consent"** as classical mechanisms used to obtain individual consent may be difficult to apply in the IoT, resulting in "low-quality consent".[51] Moreover, given the increase of the amount of data generated in combination with modern techniques related to data analysis and cross-matching, possible inferences derived from data and the re-purposing of original processing are higher in IoT. As WP29 notes, this may lend this data to secondary uses that may not be related to the purpose assigned to the original processing. The WP29 thus recommends that **users must remain in complete control of their personal data throughout the product lifecycle**; when organisations rely on consent as a basis for processing, this consent should be fully informed, freely given and specific. In this context, it is also important to take into account that citizens can find it particularly difficult to recognise the connection between the different steps of Big Data processing practices that in some circumstances, which affects their ability to balance advantages (often short-term) vs. disadvantages (often long-term).[52]

Similar concerns were previously raised concerning applications on smart devices.[53] If apps are able to collect large quantities of data from the device in order to provide new and innovative services to the end user, these same data sources can be further processed (typically to provide a revenue stream) in a manner which may be unknown or unwanted by the end user. [54]

[48] EDPS Preliminary Opinion (March 2014), "Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy".
[49] Ibid.
[50] Ibid.
[51] WP29 Opinion 8/2014 on the Recent Developments on the Internet of Things.
[52] B. Custers (2004), The Power of Knowledge: Ethical, Legal and Technological Aspects of Data Mining and Group Profiling in Epidemiology, Nijmegen: Wolf Legal Publishers, p. 209.
[53] WP29 Opinion 02/2013 on apps on smart devices.
[54] WP29 Opinion 03/2013 on purpose limitation.

In 2014, the United Nations High Commissioner for Human Rights (UNHCR) focussed on controlling the collection and retention of personal data when it warned of the difficulties of safeguarding individual rights and freedoms when Big Data allows for non-obvious, unexpected uses of data.[55] Additionally, it can be said that **Big Data's inherent rapidity (or 'velocity') puts special pressure on the capacity of individuals to follow on-going data processing practices**. These are indeed based on reactive principles that appear to be in friction with regulatory efforts and, more generally, reflexive approaches to data processing.

As strongly underlined by both the EDPS and the WP29, **empowering individuals by keeping them informed, free and safe is key to trust and innovation**. As the WP29 notes, "when we share personal data with others, we usually have an expectation about the purposes for which the data will be used. There is a value in honouring these expectations and preserving trust and legal certainty".[56]

3.2.2. Data control and digital rights: are personal data protected from unauthorised access? Is data processing under effective control?

Given the amount of data processed in a context of Big Data, IoT or cloud computing, the security of data is clearly at risk. Given the potential sensitiveness of such data, many questions arise around the vulnerability of data while it is collected and processed, which often happens outside a traditional IT structure and lacks sufficient security.[57] These vulnerabilities include data losses and infection by malware but also unauthorized access to personal data, intrusive use of wearable devices, or unlawful surveillance. Any data collection and processing is potentially intrusive, opening up a wide range of possibilities for data mining and profiling, but also triggering the need for appropriate security measures.

The case of the smart meter reader is again illustrative. Access to this data may track what members of a household do within the privacy of the home. As noted by the EDPS in its opinion, while this also raises immediate security concerns (smart meter readings could be used by criminals to assess when a house is unoccupied), **it also raises the question of accessing and processing data by unauthorized third parties, such as corporations enhancing their marketing capabilities or security agencies engaging in unlawful surveillance. The latter aspect is particularly worrying in the context of the post-Snowden revelations**. As David Lyon notes, Big Data intensifies certain surveillance trends associated with information technology and networks, and its capacities (including metadata) intensify surveillance by expanding interconnected datasets and analytical tools.[58]

As previously underlined, **it is increasingly difficult to track the multiple uses of data. This situation is complicated further by wide variety of actors engaged in data collection and processing**. In the case of apps, for instance, the WP29 details the fragmented nature of the app ecosystem, which includes app developers, app owners, app stores, manufacturers of Operating Systems (OS) and devices, and other third parties involved in the collection and processing of personal data from smart devices, such as

[55] Office of the United Nations High Commissioner for Human Rights (2014), "The Right to Privacy in the Digital Age", June, p. 6.
[56] WP29 Opinion 03/2013 on purpose limitation.
[57] WP29 Opinion 8/2014 on the Recent Developments on the Internet of Things.
[58] D. Lyon (2014), "Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique", *Big Data & Society*, 1(2), pp. 1-13.

analytics and advertising providers.[59] **This considerably increases the risk that personal data is insufficiently protected from unauthorised access**.

The control of one's personal information is key when considering a data-driven economy. As noted in the EDPS's 2014 Preliminary Opinion on Privacy and Competitiveness in the Age of Big Data, consumers should be able to withdraw and to transfer data which record their activities and are stored in the cloud, whether in the context of social networks, search engines, online banking, energy consumption, or medical or fitness tracking applications. This raises the issue of data portability (the ability to move data among different application programs, computing environments or cloud services) and interoperability (understood as the ability for people to reuse their data across various applications and devices and, most importantly, to keep control over their personal data). Data portability is a typical field at the crossroads of data protection and competition policy. Indeed, the right to data portability not only has a privacy aspect (it would give individuals more control over their personal data) but also has a competition law aspect, as such a right may also reduce lock-in effects by enabling users to switch easily between services. [60] If data portability is often perceived to be a positive force for privacy as well as competitiveness, it opens up a wide range of complex issues, identified in a 2014 workshop organised by the EDPS on Privacy, Consumers, Competition and Big Data:

- it remain[s] unclear how it could work in practice and whether it could be effective without dominant networks being compelled to interconnect;
- users would need to know what data is held about them and for what purpose;
- they would also need to be able to retrieve 'dead data' lurking in defunct 'zombie' accounts of former users of existing networks, or zombie networks which were once popular but have since been driven from the market.[61]

In light of the above-mentioned challenges, **the current negotiations around the DPGR**, undertaken alongside the development of EC strategies in the field of a data-driven economy, appear to be **a critical overarching tool to ensure that consumers' and citizens' rights are protected**. These negotiations are also a valuable opportunity to address the above-mentioned tensions between strategies for a data-driven economy and guaranteeing citizens' rights.

[59] WP29 Opinion 02/2013 on apps on smart devices.
[60] See K. X. Zhu and Z. Z. Zhou (2011), "Lock-In Strategy in Software Competition: Open-Source Software vs. Proprietary Software", *Information Systems Research*, 23(2), pp. 536-545.
[61] EDPS Report of Workshop on Privacy, Consumers, Competition and Big Data, 2 June 2014, pp. 4-5.

4. IMPACT ON PRIVACY AND PERSONAL DATA PROTECTION

KEY FINDINGS

- **Privacy and personal data protection are granted EU protection through EU primary and secondary law**. The centrepiece of EU legislation on personal data protection is Directive 95/46/EC, known as the Data Protection Directive, whose revision is currently under discussion as part of the Data Protection Reform Package.
- Big Data and smart devices give rise to **a number of challenges and reveal a series of potential gaps in the EU legal framework**, in particular in the following areas: transparency and information obligations of data controllers; consent (including consent in case of repurposing); the need to balance public interest and the interests of data subjects for legitimising personal data processing; the regulation of profiling; and proper safeguarding of digital rights in case of data transfers to third parties and third countries and access to EU data.
- In the context of the trilogue opened up in June 2015, **different tensions and controversies can be foreseen**.

As outlined in the previous section, the development of Big Data and the spread of smart devices can be envisioned as making the rights to privacy and personal data protection more relevant and necessary than ever. The increased significance of Big Data and smart devices in our societies is indeed expected to have a direct impact on the lives of individuals, crucially intersecting with their right to respecting private life. Moreover, the multiplication of smart devices and the advance of Big Data rely on the continuous increase in the processing of personal data, triggering the need to ensure strict compliance with personal data protection safeguards. Big Data and smart devices, however, can also represent a challenge to the effective implementation of existing legal rules, revealing problematic gaps in the ways they have been designed. This Section thus describes applicable EU standards for privacy and personal data protection (4.1) and reviews key gaps and challenges generated by the deployment of Big Data and smart devices (4.2).

4.1. Applicable EU standards

EU applicable standards for privacy and personal data protection are provided by primary (4.1.1) and secondary laws (4.1.2).

4.1.1. Fundamental Rights

Privacy and personal data protection are both recognised as Fundamental Rights in the Charter of Fundamental Rights of the EU. Article 7 of the EU Charter on the right to respect for private life sets out that "[e]veryone has the right to respect for his or her private and family life, home and communications". This provision has to be interpreted as corresponding to Article 8 of the European Convention on Human Rights (ECHR) and in line with the case law thereof by the European Court of Human Rights (ECtHR). This implies that the notion of 'respect for private life' needs to be understood broadly, as generally referring to the right of individuals to live their own lives, covering their right to have social relations with others, and potentially applying both in private and in public spaces.

Article 8 of the EU Charter is specifically devoted to the right to the protection of personal data.[62] After establishing that "[e]veryone has the right to the protection of personal data concerning him or her", it notes that "[s]uch data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law", which means that personal data can only be processed on the basis of a legitimate ground, which in certain cases might be the consent of the individual. The second paragraph of Article 8 further indicates that "[e]veryone has the right of access to data which has been collected concerning him or her, and the right to have it rectified", referring to the need to ensure minimum rights of data subjects whenever personal data about them is processed. Finally, the third paragraph of Article 8 of the EU Charter sets out that "[c]ompliance with these rules shall be subject to control by an independent authority".

4.1.2. EU secondary law

The centrepiece of EU legislation on personal data protection is Directive 95/46/EC, known as the Data Protection Directive.[63] This instrument enshrines the two main objectives of EU personal data protection law: firstly, the protection of Fundamental Rights and freedoms of individuals; secondly, the achievement of the internal market, in the specific form of the free flow of personal data. It thus obliges member states to guarantee a high level of protection while forbidding the creation of any obstacles to the free flow of personal data between member states in the name of privacy and personal data protection (Art. 1).

Directive 95/46/EC generally applies to the processing of personal data wholly or partly by automatic means and to the processing otherwise than by automatic means of personal data which forms part of a filing system or is intended to form part of a filing system [Art. 3(1)]. 'Personal data' is defined as "any information relating to an identified or identifiable natural person ('data subject')", an identifiable person being "one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity" [Art. 2(a)]. 'Processing of personal data' is "any operation or set of operations which is performed upon personal data", including mere data collection [Art. 2(b)].

The Data Protection Directive describes as principles of 'data quality' the main principles to be complied with whenever personal data are processed: fair processing (data must be "processed fairly and lawfully"); purpose specification (data can only be "collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes"); accuracy (data must be "accurate and, where necessary, kept up to date"); and proportionality (data must be "adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed"; additionally, data cannot be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed") (Art. 6). The principles of confidentiality and security of data processing must as well be ensured (Arts. 16 and 17).

Directive 95/46/EC also details the grounds on which the processing of personal data can be legitimately grounded, such as, for instance, the consent of the data subject or the need

[62] On the appearance of this right in EU law, see G. González Fuster (2014), *The Emergence of Personal Data Protection as a Fundamental Right of the EU*, Dordrecht: Springer.
[63] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L281, 23/11/1995, pp. 31-50.

to process the data for the performance of a task carried out in the public interest (Art. 7). The Directive also specifies the information that must be provided to data subjects when personal data about them is collected directly from them (Art. 10) and when personal data that had been collected in another context is processed Art. 11). In addition to this general right to be informed, data subjects are granted a right to access personal data about them in the hands of data controllers and to request "as appropriate the rectification, erasure or blocking of data the processing of which does not comply with the provisions of this Directive, in particular because of the incomplete or inaccurate nature of the data" (Art. 12), as well as the right to object to certain uses of the data concerning them (Art. 14).

The processing of data "revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life" is, due to its sensitive nature, generally prohibited, except under particularly strict conditions (Art. 8).

Automated individual decisions - decisions producing legal effects concerning individuals or significantly affecting them and which are based solely on the automated processing of data intended to evaluate certain personal aspects relating to them, such as their performance at work, creditworthiness, reliability, or conduct [Art. 15(1)] - are also generally prohibited. These type of decisions might however be allowed when taken in the course of entering into or the performance of a contract, or when authorised by law [Art. 15(2)].

In relation to data transfers to third countries, the Data Protection Directive puts forward as a default rule that personal data can only be transferred to third countries which are recognised to ensure an 'adequate level of protection' (Art. 25). Derogations are, however, possible, for instance, on the basis of the unambiguous consent of the data subject (Art. 26). Taking into account that only a limited number of third countries have actually been recognised as providing an adequate level of protection, this approach raises many questions. Different solutions have been sought to nevertheless allow for the transfer of data across borders while attempting to condition the legitimacy of such transfers to the provision of a minimum level of protection, for instance, through the negotiation (and re-negotiation) of a Safe Harbour agreement with the United States[64].

The 1995 Data Protection Directive constitutes a horizontal instrument that has been complemented with other norms. Particularly relevant is Directive 2002/58/EC, known as the e-Privacy Directive, specifically concerned with the electronic communications sector.[65] The e-Privacy Directive puts forward the categories of 'traffic' and 'location' data, which are respectively defined as "any data processed for the purpose of the conveyance of a communication on an electronic communications network or for the billing thereof" and "any data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service" (Art. 2); both are subject to reinforced protection (Arts. 6 and 9).

4.1.3. Review of the EU personal data protection framework

The entry into force of the Lisbon Treaty in December 2009 represented a key moment for the EU's privacy and personal data protection framework. Not only did the Treaty endow

[64] D. Bigo et al. (2012), "Fighting cyber crime and protecting privacy in the cloud", Study for the European Parliament's Committee on Civil Liberties, Justice and Home Affairs (LIBE), PE 462.509, Brussels.
[65] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), OJ L201, 31.7.2002, p. 37–47.

the EU Charter of Fundamental Rights with legally binding force, it also introduced the right for protection of personal data in the EU treaties. [66] The Treaty thus introduced a new legal basis allowing for the establishment of comprehensive and coherent EU legislation on privacy and personal data protection.

In 2010, the European Commission issued a Communication stating that although the principles enshrined in the Data Protection Directive remained valid, rapid technological developments and globalisation had brought new challenges to the area.[67] The European Commission Communication did not refer to Big Data or smart devices but alluded to technological developments rendering the collection of personal data less easily detectable.[68] The 2010 Communication announced the European Commission's intention to consider policies that would ensure a coherent application of data protection rules taking into account the impact of new technologies on individual rights and freedoms as well as policies to meet the objective of ensuring the free circulation of personal data within the internal market.

In January 2012 the European Commission finally presented a legislative package aiming to reform the EU personal data protection legal framework. The package included the Communication *Safeguarding Privacy in a Connected World: A European Data Protection Framework for the 21st Century*, a proposal for a General Data Protection Regulation prepared to replace Directive 95/46/EC, and a proposal for a Directive on the protection of personal data processed for the purposes of prevention, detection, investigation or prosecution of criminal offences and related judicial activities to replace an existing Framework Decision.[69]

The 2012 Communication *Safeguarding Privacy in a Connected World* alludes to Big Data to support the argument that personal data has become an asset for many corporations and that collecting, aggregating and analysing the data of potential customers is often an important part of corporate economic activity.[70] This is indeed illustrated with a reference to the 2011 McKinsey Global Institute report *Big data: The next frontier for innovation, competition, and productivity*, which claims that Big Data will become a key basis of competition and underpin new waves of productivity growth and innovation, although such potential may only be achieved, the report notes, if supported by ad-hoc privacy policies.[71]

[66] TFEU, Article 16.
[67] European Commission (2010), Communication to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions: A Comprehensive Approach on Personal Data Protection in the European Union, COM(2010) 609 final, 4.11.2010, Brussels.
[68] Ibid., p. 2.
[69] European Commission (2012), Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Safeguarding Privacy in a Connected World: A European Data Protection Framework for the 21st Century, COM(2012) 9 final, 25.1.2012, Brussels; European Commission (2012), Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation), COM(2012) 11 final, 25.1.2012, Brussels; European Commission (2012), Proposal for a Directive of the European Parliament and of the Council 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, COM(2012) 10 final, 25.1.2012, Brussels; and Council Framework Decision 2008/977/JHA of 27 November 2008 on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters, OJ L350, 30.12.2008, pp. 60–71.
[70] Ibid., p. 2.
[71] On this report, see http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

The explanatory memorandum accompanying the proposed General Data Protection Regulation refers to the need to update Directive 95/46/EC in light of technological developments that have sparked a dramatic increase in the scale of data sharing and collecting.[72] While holding that the existing framework remains sound as far as its objectives and principles are concerned, the European Commission contends in the memorandum that the time has come to build a stronger framework, one that would put individuals in control of their own data.[73]

The text proposed by the European Commission aims, inter alia, to clarify the requirements for the valid consent of data subjects as a legitimate ground for data processing; puts forward a 'transparency principle', understood as requiring that clear information be provided to data subjects; amplifies information obligations of data controllers towards data subjects; and introduces compulsory notification of some personal data breaches.[74]

A significant innovation of the General Data Protection Regulation as drafted by the European Commission is the introduction of special provisions for the protection of children's personal data, in line with the explicit mention of the child's right to privacy in the United Nations Convention on the Rights of the Child.[75] In this sense, the European Commission has proposed to set out special conditions for the lawfulness of processing the personal data of children in relation to information society services offered directly to them.[76] As stated in the proposed Preamble, children are recognised as deserving specific protection of their personal data "as they may be less aware of risks, consequences, safeguards and their rights in relation to the processing of personal data".[77]

In June 2015 the European Parliament, the Council and the Commission initiated co-decision negotiations on the proposed General Data Protection Regulation - on the basis of the 2012 proposal by the Commission, a parliamentary legislative resolution adopted on 12 March 2014, and a General Approach of the Council of 15 June 2015 - with the ambition of concluding the discussions at the end of 2015.

As noted by the EDPS, the resulting text is not only supposed to provide a solid framework for protecting Fundamental Rights in the era of Big Data but also beyond it, as data-driven technologies progressively converge with artificial intelligence.[78] Once the future General Data Protection Regulation is adopted, EU institutions will most likely review the e-Privacy Directive in order to bring it in line with the upcoming text.

4.2. Gaps and challenges

Big Data, smart devices and, more generally, the 'data revolution' they signal, raise a number of challenges to safeguarding Fundamental Rights. These challenges must be discussed so that the complete and effective protection of these rights can be ensured.

[72] COM(2012) 11 final, p. 1.
[73] Ibid., p. 2.
[74] Ibid., pp. 8, 9, 10.
[75] United Nations Convention on the Rights of the Child, adopted and opened for signature, ratification and accession by General Assembly resolution 44/25 of 20 November 1989, entered into force 2 September 1990. See in particular Art. 16. This step was already taken by the United States almost two decades ago, with the adoption of the Children's Online Privacy Protection Act of 1998, 15 U.S.C. 6501–6505.
[76] COM(2012) 11 final, p. 8.
[77] Recital 29, ibid., p. 22.
[78] EDPS Opinion (July 2015), Europe's Big Opportunity - EDPS Recommendations on the EU's Options for Data Protection Reform, Brussels, p. 7.

4.2.1. Complete protection

As noted, EU law at the highest level currently protects two different Fundamental Rights, namely the right to respect for private life and the protection of personal data. EU secondary law, however, is primarily concerned with the regulation of the processing of personal data. Thus **the applicability of many EU norms is conditioned to data being qualified as personal**. It is extremely important, therefore

- to ensure that any personal data processing that falls under EU personal data protection law does not elude relevant safeguards on the basis of an incorrect claim according to which the data is not 'personal';

- to remember that the processing of data that is not qualified as personal might also constitute an unlawful infringement of the right to respect for private life.

Making sure that all data that could be used to identify an individual remains under the scope of personal data protection law requires an acknowledgment that 'pseudonymised' data must be qualified as personal data.[79] This task also asks for considering the possibility that data that at a certain point is not considered personal data (in the sense that it cannot be related to any identified or identifiable individual) might later on actually be linked up with a concrete person precisely through data analytics and the processing of large unstructured sets of data.[80] In a way, understanding identifiability and the meaning of 'personal data' in the Big Data era means that it is necessary to take Big Data capabilities seriously.

4.2.2. Effective protection

When EU personal data protection applies, Big Data and smart devices undoubtedly put pressure on some of its principles. The scale of personal data being processed, along with the sensitivity of some of the data collected and exhausted, summon forward a particularly reinforced implementation of applicable and upcoming norms, for instance in relation to data security, privacy by design, or privacy impact assessments.

Some challenges deserve special mention:

- **Preserving transparency and information obligations of data controllers**: Data controllers must provide to the data subjects information on the processing they are going to undertake, the rights certain data subjects hold, and such information must be provided in a clear manner. These information and transparency obligations apply generally (although derogations are possible), not just when the data processing is based on the consent of the data subject. If data subjects are not informed in a clear manner about the processing of their personal data, they could be deprived de facto of the possibility of exercising their rights to access and rectification,

[79] Ibid., p. 5.
[80] G. Buttarelli (2015), "Big data, big data protection: challenges and innovative solutions", keynote speech, ERA Conference, "Recent Developments in Data Protection Law", 11 May, at https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/EDPS/Publications/Speeches/2015/15-05-11_ERA_speech_EN.pdf, p. 9.

which are explicitly mentioned in the EU Charter.[81] It is thus crucial that these transparency and information obligations are preserved also considering the deployment of smart devices. Beyond the preservation of existing transparency and information obligations, it might be necessary to assess whether increasingly rapid and possibly unpredictable data processing practices need to be accompanied by other protection mechanisms that acknowledge that asymmetries in knowledge cannot be (easily) overcome.

- **Repurposing and transparency**: Big Data analytics often engage in processing data for purposes that had not been initially scheduled, or could even, in theory, process data for purposes still to be discovered.[82] Data subjects, however, cannot be left uninformed about such repurposing. Transparency towards them must be preserved as data enters new purposes, encompassing not only the aim of the processing but also the manner in which it takes place.

- **Balancing public interest and the interests of data subjects for legitimising personal data processing**: As established in the 1995 Data Protection Directive, the need to process personal data for the performance of a task carried out in the public interest can constitute a legitimate basis for such processing. The possibility should, however, not be over-stretched so as to encompass any possible third-party interest. The concept of data minimisation is relevant in this case, acting as a reminder that data processing shall always be as limited as possible.

- **Consent**: Consent can only operate as valid legitimate ground for the processing of personal data when data subjects are not coerced into consenting and when they are informed about what they are 'consenting' to. In the case of repurposing, data subjects might need to be given again the opportunity to consent or reject a determined data processing practice. Data subjects need also to be able to express their consent in a clear manner, when applicable. The negotiation of consent and its limits raise also a specific set of challenges in IoT environments, especially as data subject's interactions are increasingly mediated by or delegated to (smart) devices and applications.

- **Special categories of data**: Even if all personal data must be protected under general EU personal data protection safeguards, some categories of data deserve special, reinforced guarantees. This notably affects 'sensitive' and location data. In the context of Big Data, it is crucial to note that the processing of non-sensitive data can lead, through data mining, to the generation of data that reveals sensitive information.

- **Profiling/automated decisions:** The General Data Protection Regulation currently under negotiation is expected to echo (although possibly with minor changes) the existing provision, in the Data Protection Directive, generally prohibiting automated individual decisions that significantly affect individuals.

[81] On the right to be informed in current and upcoming EU personal data protection law see G. González Fuster (2014), "How Uninformed is the Average Data Subject? A Quest for Benchmarks in EU Personal Data Protection", *IDP Revista de Internet, Derecho y Política*, 19, pp. 92-104.

[82] In this sense, see Information Commissioner's Office (ICO), *Big data and data protection*, at https://ico.org.uk/media/for-organisations/documents/1541/big-data-and-data-protection.pdf.

These decisions can now be officially referred to as 'profiling'. A significant change could be the reference to the fact that individuals can be submitted to these types of automated decisions when they consent to such practices, which brings to the fore, again, the challenges of ensuring informed consent. The Regulation's provision on 'profiling', however, only addresses part of the privacy issues generated by decision-making in Big Data scenarios because it focuses on a concrete (final) moment when individuals are confronted with automated decisions that 'significantly' affect them. Issues that are left unaddressed are, for instance, the need to inform individuals about the ways in which data about them might be use to take decisions about other individuals, as well as the manner in which such decisions are taken. These pending issues might require further transparency requirements.[83]

- **Transfers to third countries and third country access to EU data**: Reliance on cloud computing for Big Data practices demands an acceptance that massive quantities of personal data, including sensitive data, are extremely likely to move quickly across geographical and jurisdictional borders. As a matter of fact, even when data does not leave EU territory (and thus remain within EU jurisdiction), it might be at risk of being apprehended by third-country authorities for different purposes, including law enforcement and security purposes.[84] Therefore, thinking about the privacy and personal data protection implications of Big Data requires coordination with cloud computing policy.

- **Protecting children:** The recognition that minors are particularly vulnerable in the face of data processing practices, and thus require specific provisions to ensure their full protection, is one of the key novelties of the proposed General Data Protection Regulation. Smart devices and Big Data make this issue particularly significant, especially when devices aim to capture personal data from children and/or interact with them on the basis of automated data processing practices. These developments can require devising specially protected digital spaces for children as well as ad-hoc data protection provisions.[85]

4.2.3. The context of the trilogue and the issues at stake

The opening of the trilogue in June 2015 already announces possible tensions. If the 'package approach' has been welcomed by the EP and perceived as the adequate way to unify personal data protection laws, concerns have been raised regarding the agreement reached by the Council and the Commission on the proposal for a GDPR. The EP Rapporteur on the GDPR, Jan Philipp Albrecht, outlined the following issues opened up for negotiations:

- Users must be informed about what happens with their data, and they must in principle be able to consciously agree to data processing – or reject it. On that matter, the Rapporteur notes that while the Parliament insists on 'explicit' consent

[83] See also EDPS, Opinion 3/2015, p. 8.
[84] On law enforcement, see S. Carrera, G. González Fuster, E. Guild and V. Mitsilegas (2015), "Access to Electronic Data by Third-Country Law Enforcement Authorities: Challenges to EU Rule of Law and Fundamental Rights", Centre for European and Policy Studies (CEPS), http://www.ceps.eu/publications/access-electronic-data-third-country-law-enforcement-authorities-challenges-eu-rule-law. On security, see C. Bowden (2013), The US surveillance programmes and their impact on EU citizens' Fundamental Rights, European Parliament, Directorate General for Internal Policies, Policy Department C: Citizens' Rights and Constitutional Affairs, Civil Liberties, Justice and Home Affairs.

as proposed by the Commission, the Council's version of the draft law foresees the much more vague requirement of 'unambiguous' consent.

- The Rapporteur advocates that users should receive understandable information on how their data is processed or if the provider has transferred data to public authorities or intelligence services.
- The Rapporteur notes that companies should not be allowed to hand over data from Europe directly to the authorities of third countries.
- The Rapporteur underlines that all information that can be directly or indirectly linked to a person is defined as personal information and needs to be protected as such.
- In the case of illegal data processing and in severe cases, companies should face tough sanctions; according to the Rapporteur, the Parliament wants to raise the possible sanctions to up to five per cent of the global annual turnover, or 100 Million Euros (against up to two percent of global annual turnover proposed by the Commission).
- Privacy by Design/Privacy by Default should be encouraged: data processors, as well as producers of IT systems, should design their services in a data-minimising way and with the most data protection-friendly pre-settings.
- The mandatory appointment of a data protection officer (DPO) should depend on the amount and relevance of data processing, not on the size of a company. On that matter, the Rapporteur notes that the Council has suggested leaving it up to the member states if the data protection officer should be mandatory at all.
- The Rapporteur advocates a strong role for the future European Data Protection Board to ensure that it effectively gives voice to independent national data protection authorities; he notes that there are still open discussions on the details of the "one-stop shop" mechanism.[86]

These issues for negotiation echo those outlined in this Study. Additional controversial aspects of the proposed GDPR text as supported by the Council have been highlighted by digital rights organisations:

- The current text of the GDPR allows for the further processing of personal data "for archiving purposes in the public interest or scientific, statistical or historical purposes." However, it is unclear what those purposes are, and there is a risk that companies might assert that the processing pursues one of these targets even when, strictly speaking, this is not the case.
- The proposed Article 5(c) removes the obligation to keep processing to a minimum and weakens it to "non-excessive" processing.
- The "legitimate interest" justification for data processing without consent is the vaguest ground for processing, offering a lot of scope for industry to process data if they can claim a "legitimate interest" in doing so.
- Under the Council version, organisations defending the interests of citizens and consumers can no longer complain to authorities or take judicial actions on behalf of individuals whose privacy rights have been breached.
- As regards data transfers outside the EU, excessive significance is given to privacy seals/trust-marks (called "certification mechanisms") and codes of conduct.

[85] In this sense, see also D. Wright et al. (ed.) (2008), *Safeguards in a World of Ambient Intelligence*, Dordrecht: Springer, p. 210.
[86] J. Albrecht (January 2015), EU General Data Protection Regulation State of play and 10 main issues, at http: http://www.janalbrecht.eu/fileadmin/material/Dokumente/Data_protection_state_of_play_10_points_010715.pdf

- The Council proposals would allow further processing of health data, including genetic data on a massive scale and sharing of this data with third parties, including companies such as Google, without people's knowledge or consent.[87]

These concerns are all the more legitimate given the results of a May 2015 Eurobarometer survey asking 28,000 EU citizens what they think about the protection of their personal data.[88] The overall conclusion of the survey shows that the protection of personal data remains a very important concern for citizens:

- More than eight out of ten respondents feel that they do not have complete control over their personal data;
- A majority of people (53%) are uncomfortable with Internet companies using their personal information to tailor advertisements;
- 7 out of 10 people think the collection of their data should require their explicit approval;
- Around seven out of ten people are concerned about their information being used for a different purpose from the one it was collected for.

In light of the above-mentioned challenges related to Big Data from a privacy and personal data protection perspective, the confidence displayed after the first trilogue that the package will be adopted by the end of the year appears somehow unfounded.[89]

[87] See EDRi and Privacy International Press Release (15 June 2015), "Privacy and Data Protection under threat from EU Council agreement".
[88] Special Eurobarometer 431 on data protection, June 2015, at http://ec.europa.eu/public_opinion/archives/ebs/ebs_431_sum_en.pdf.
[89] See "Data Protection reform, first trilogue – next steps", Press Conference, 24 June 2015, at http://ec.europa.eu/avservices/video/player.cfm?ref=I105123.

5. GENERAL CONCLUSIONS AND RECOMMENDATIONS

As developed throughout this Study, Big Data, perceived as both an economic and social driver, raises significant challenges in terms of privacy and personal data protection. Furthermore, in a context of rapid developments in smart devices and IoT, it is increasingly difficult to track what is made of our data. This is complicated further by the fact that a wide variety of actors are engaged in data collection and processing. Individuals are too often left ignorant of the full spectrum of data they produce (or they are made to produce) and of the data processing operations that are undertaken on this data.

The Study has argued for more transparency in the field of data processing as well as a legal framework that fully respects and guarantees the digital citizen's rights. Such a **framework is critical to ensuring a sustained and fair digital economic model that would benefit both the private and public sectors and consumers and citizens**.

The key issues for reflection that have been underlined to promote such a framework for digital citizen's rights include:

- effective protection of personal data requires further transparency from data controllers;
- clearer information on the purpose and mechanism of data processing is necessary to ensure the quality of consent to such processing and that discriminatory practices are not taking place;
- guarantee a strong level of protection in the transfer of data to third parties and third countries.

The following recommendations aim at addressing these challenges:

Recommendation 1: The guarantee of the digital citizen's rights should be at the core of the data production chain

Within the data value chain strategy developed by the European Commission and the DG CONNECT in particular, the guiding principle of finding the right balance between individual privacy concerns and the exploitation of the potential of the reuse of data should be **driven by the requirement of guaranteeing respect for the Fundamental Rights (privacy and data protection) laid down in EU primary and secondary law**. The role of data controllers for this is central, and they are accordingly legally required to observe a number of principles when processing personal data, compliance of which must be reinforced (as detailed in recommendation 2).

Such requirements would entail going much further than the promotion of Privacy by Design/Privacy by Default or Privacy Enhancing Technologies (PETs). If these arrangements would undoubtedly enhance privacy settings, they would be insufficient in guaranteeing digital rights. **PETs are not the ultimate solution to safeguarding privacy**. These aspects should be central in the current trilogue on the GDPR.

Within this perspective, if the "one-stop shop" mechanism (which would facilitate reaching single supervisory decision for any issues related to EU personal data processing affecting multiple member states) is a welcome step in promoting the digital citizen's rights, the **responsibility for protecting Fundamental Rights should be promoted along the**

data production chain and gather various stakeholders, thus promoting a sound and fair data-driven economy. The European Parliament, for instance, could promote regular dialogue between companies (developing and distributing apps, processing data, providing analytics, etc.), regulators and data protection experts across policy boundaries. As underlined by the EDPS in relation to competition law, privacy and the protection of personal data should not be considered peripheral concerns but rather central factors in the appraisal of the activities of corporations and their impact on competiveness, market efficiency and consumer welfare. Labels could be devised and delivered to corporations that comply with privacy regulations.

Moreover, the quality of consent as a key instrument as the legitimate basis for personal data processing needs to be supported by quality information practices. The **degree of information and awareness of data subjects must be of prime concern whenever personal data processing takes places**, and this is even more the case when consent is used as a basis to legitimate data processing. The advent of Big Data and smart devices requires further progress in information practices and obligations.

Recommendation 2: Impact on Fundamental Rights should be systematically integrated in the definition of Big Data policy

The EU regulatory framework of privacy and personal data protection takes as its starting point that data processing activities can only take place under certain conditions, which include compliance with some obligations by data controllers, control by data subjects through a set of subjective rights, and monitoring by data protection authorities. Over the years, the framework has integrated notions and mechanisms that support the taking-into-account of privacy concerns, for instance through **the definition of privacy and data protection impact assessment obligations**. This approach should be sustained and reinforced to design a Big Data policy compliant with Fundamental Rights, where impact on these rights is evaluated and considered at all key stages.

Recommendation 3: The GDPR should grant individuals complete and effective protection in the face of current and upcoming technological developments of Big Data and smart devices

Big Data and smart devices do not represent the end of privacy and personal data protection. On the contrary, they call for a reworking of the EU privacy and personal data protection framework to ensure it is up-to-date and operational. A key step in this reinforcement is the adoption of a sound GDPR that grants individuals complete and effective protection in the face of current and upcoming technological developments. In particular:

- For individuals to enjoy full protection of their rights, it is necessary to guarantee that **personal data protection rules do not inappropriately exclude any data sets that relate to or could relate to identified or identifiable individuals**.
- For individuals to enjoy full protection of their rights, it is also crucial to consider that the right to respect for private life can be relevant when Big Data involves the processing of data that does not qualify as personal. In this sense, **the data processing practices of both private companies and public authorities must take into the account the possible impact on individual rights**, even when these data processing activities do not technically rely on, or generate, personal data, as they could nevertheless infringe the right to respect for private life or other rights and individual freedoms.

The GDPR currently under discussion should not offer less protection and guarantees than the 1995 Directive, and users should remain in complete control of their personal data throughout the data lifecycle.

Recommendation 4: The e-Privacy Directive should be reviewed in light of the upcoming GDPR and the challenges posed by the development of Big Data and Smart Devices, including the IoT.

Effective protection of individuals cannot be guaranteed solely by the adoption of sound GDPR but will require also a consistent review of the e-Privacy Directive. This instrument not only pursues the safeguarding of personal data protection but, more generally, aims to ensure this right and the right to respect for private life.

Crucially, the strong linkage between, on the one hand, the deployment of Big Data and smart devices, and, on the other, cloud computing, calls for **a clear EU policy on how to ensure privacy and personal data protection in cloud computing scenarios**. This requires addressing the issue of cross-border data transfers and, more generally, the legal conflicts that can occur in cloud computing environments.

6. REFERENCES

Bigo, D. (2013), "Sécurité maximale et prévention? La matrice du futur Antérieur et ses grilles", in B. Cassin (ed.), *Derrière les grilles: Sortir du tout évaluation*, Paris: Fayard/Mille Et Une Nuits.

Bigo, D. et al (2012), "Fighting cyber crime and protecting privacy in the cloud", Study for the European Parliament's Committee on Civil Liberties, Justice and Home Affairs (LIBE), PE 462.509, Brussels.

Boullier, D. (2011), "Habitèle virtuelle: une nouvelle enveloppe pour commuter, notre téléphone portable", *Revue Urbanisme*, no. 376: pp. 42-44.

Boullier, D. (2014), "Habitele: mobile technologies reshaping urban life', Urbe, 6(1), pp. 13-16.

Carrera, S., G. González Fuster, E. Guild and V. Mitsilegas (2015), "Access to Electronic Data by Third-Country Law Enforcement Authorities: Challenges to EU Rule of Law and Fundamental Rights", Centre for European and Policy Studies (CEPS).

Craig, T. and M. E. Ludloff M.E. (2011), *Privacy and Big Data*, Sebastopol, CA: O'Reilly.

Custers, B. (2004), The Power of Knowledge: Ethical, Legal and Technological Aspects of Data Mining and Group Profiling in Epidemiology, Nijmegen: Wolf Legal Publishers.

Datta, A., M. C. Tschantz and A. Datta (2015), "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination", *Proceedings on Privacy Enhancing Technologies*, 1, pp. 92-112.

Delort, P. (2015), *Le Big Data, Que Sais-Je ?*, Paris: Presses Universitaires de France.

Desouza, K.C., L. Kendra and K. L. Smith (2014), "Big Data for Social Innovation" *Stanford Social Innovation Review*, 12(3), pp. 39–43.

Gandy, O. (2009), *Coming to Terms With Chance: Engaging Rational Discrimination and Cumulative Disadvantage*, Farnham: Ashgate.

González Fuster, G. (2014), "How Uninformed is the Average Data Subject? A Quest for Benchmarks in EU Personal Data Protection", *IDP Revista de Internet, Derecho y Política*, 19, pp. 92-104.

González Fuster, G. (2014), The Emergence of Personal Data Protection as a Fundamental Right of the EU, Dordrecht: Springer.

Greenfield A. (2006), Everyware: The Dawning Age of Ubiquitous Computing, Berkeley, CA: New Riders.

Isin, E. and E. Ruppert (2015), *Being Digital Citizens*, London: Rowman & Littlefield.

Kitchin, R. (2014), The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences, Los Angeles, CA: SAGE.

Li, K. C., H. Jiang, L. T. Yang, and A. Cuzzocrea (eds.) (2015), *Big Data: Algorithms, Analytics and Applications*, Boca Raton, FL: Chapman & Hall/CRC Big Data Series.

Lyon, D. (2014), "Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique", *Big Data & Society*, 1(2), pp. 1-13.

Mayer-Schönberg, V. and K. Cukier (2013), *Big Data: A Revolution That Will Transform How We Live, Work and Think,* London: John Murray.

Ruppert, E., P. Harvey, C. Lury, A. Mackenzie, R. McNally, S. Baker, Y. Kallianos, and C. Lewis (2015), "Socialising Big Data: From Concept to Practice", CRESC Working Paper Series, Working Paper no. 138.

Sadin, E. (2015), *La vie algorithmique: Critique de la raison numérique*, Paris: Éditions L'Échappée.

Stapleton, L.K. (2011), "Taming big data", *IBM Data Management Magazine*, 16(2), pp. 12-18

Stiegler, B. (2015), *La Société automatique 1: L'avenir du travail*, Paris: Fayard.

Swan, M. (2012), "Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0," *Journal of Sensor and Actuator Networks*, 1, pp. 217–53.

Tene, O. and J. Polonetsky (2013), "Big Data for All: Privacy and User Control in the Age of Analytics", *Northwestern Journal of Technology and Intellectual Property*, 11(5), 239-273.

Wolf, C. (2015), "Envisioning Privacy in the World of Big Data" in M. Rotenberg, J. Horwitz and J. Scott (eds.), *Privacy in the Modern Age: The Search for Solutions*, New York, NY: The New Press.

Wright, D. et al. (eds.) (2008), *Safeguards in a World of Ambient Intelligence*, Dordrecht: Springer.

Wright, D. and R. Kreissl (eds.) (2015), *Surveillance in Europe*, Abingdon: Routledge.

Zhu, K. X. and Z. Z. Zhou (2011), "Lock-In Strategy in Software Competition: Open-Source Software vs. Proprietary Software", *Information Systems Research*, 23(2), pp. 536-545.

CAT: QA-02-15-737-EN-C (paper)
CAT: QA-02-15-737-EN-N (pdf)

DIRECTORATE-GENERAL FOR INTERNAL POLICIES

POLICY DEPARTMENT C
CITIZENS' RIGHTS AND CONSTITUTIONAL AFFAIRS

Role

Policy departments are research units that provide specialised advice to committees, inter-parliamentary delegations and other parliamentary bodies.

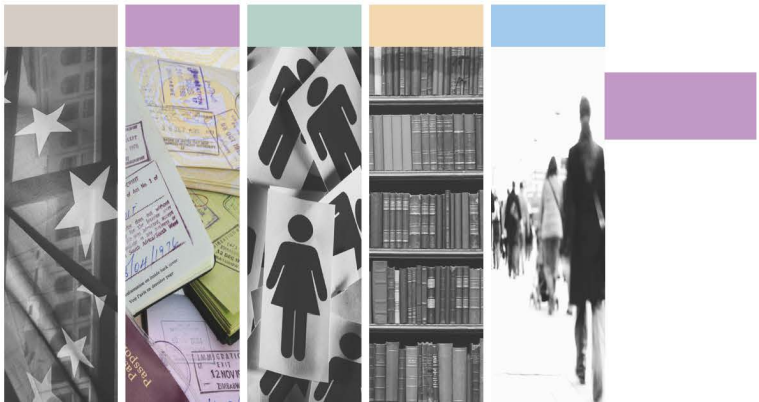
Policy Areas

- Constitutional Affairs
- Justice, Freedom and Security
- Gender Equality
- Legal and Parliamentary Affairs
- Petitions

Documents

Visit the European Parliament website:
http://www.europarl.europa.eu/supporting-analyses

PHOTO CREDIT: iStock International Inc.



ISBN 978-92-823-7957-8 (paper)
ISBN 978-92-823-7958-5 (pdf)

doi: 10.2861/787388 (paper)
doi: 10.2861/989586 (pdf)

Publications Office