

Big Data, Big Brother?

Striking the right balance with privacy

Stéphane Hurtaud
Partner
Governance, Risk
& Compliance
Deloitte

Laureline Senequier
Senior Manager
Governance, Risk
& Compliance
Deloitte

Aurelia Schwander
Consultant
Governance, Risk
& Compliance
Deloitte

“The availability and use of Big Data is crucial for maintaining the EU's competitiveness. Currently there is a lack of clarity about who owns these industrial data, and about how they may or may not be used. This reduces incentives to develop data-analytics services.”¹ Such was the assessment of *Günther Oettinger, Commissioner for Digital Economy and Society*, regarding the current state of Big Data use in the EU.

¹ Speech at Hannover Messe: "Europe's future is digital" https://ec.europa.eu/commission/2014-2019/oettinger/announcements/speech-hannover-messe-europes-future-digital_en



Indeed, while Big Data may lead to new opportunities, it also raises concerns as current legislation may fail to give precise directions for its development, and it is therefore important to understand what the current trends and good practices are surrounding Big Data today.

Welcome to the Big Data era

Over the past few years, technological innovations and device affordability have led to an explosion of “real-time digital data”². According to a paper published by the UN Global Pulse, the stock of digital data is expected to increase 44 times between 2007 and 2020, doubling every 20 months. In this context, **Big Data** acts as a response to market actors’ need for lower uncertainty and to create valuable information from the constantly produced new data.

Big Data can be defined as a “data collection, merging data from multiple sources into a single one”³. The types of data that are collected comprise various items,

such as call logs, mobile-banking transactions, blog posts, tweets, online searches, satellite images, or even videos. One purpose of the Big Data solutions is indeed to overcome traditional database systems, requiring very structured data to operate, and to identify unseen trends and patterns from chaotic collected datasets.

In addition to this variety, Big Data also aims at managing amounts of data so large that computational tools are required to extract data and obtain usable information in a very short time span. This may be achieved by methods such as cluster analysis (grouping similar elements or people) or association analysis (co-occurrence of items).

The ability of Big Data solutions to process and create valuable information from a massive amount of heterogeneous data therefore represents a tremendous opportunity for both the private and public sector in the EU to compete and take advantage of today’s digital environment.

² *Big Data for Development: Challenges & Opportunities*, UN Global Pulse <http://unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf>

³ *Big data: opportunities and privacy concerns*, Piotr Bąkowski, Members’ Research Service, European Parliamentary Research Service [http://www.europarl.europa.eu/RegData/bibliotheque/briefing/2014/140771/LDM_BRI\(2014\)140771_REV1_EN.pdf](http://www.europarl.europa.eu/RegData/bibliotheque/briefing/2014/140771/LDM_BRI(2014)140771_REV1_EN.pdf)



Big Data promises: a wealth of opportunities

Higher productivity, improved service offerings, and better client targeting are just some of the outcomes expected to be achieved through the adequate processing of these large volumes of data.

In the public sector, **decision-makers** in areas such as public health, national security, urban planning, food security, energy efficiency, and intelligent transport systems could benefit from Big Data. It would, for instance, “allow to track development progress, improve social protection, and understand where existing policies and programs require adjustment”⁴ by complementing official statistics and surveys. For this purpose, real-time feedback provided by Big Data analysis on the effectiveness of public policies would indeed narrow the time and knowledge gap between studies, allowing a targeted and adaptive approach by political leaders to be adopted in a timely manner.

By way of example, the UK government is currently supporting Big Data to monitor public transportation infrastructure and services and better understand users’ needs. Analysis of public transport data gives useful insights on journey patterns, which can then be used to support the planning of services⁴.

Another practical illustration of the use of Big Data in the public sector is the way in which some police enforcement agencies adopt new strategies that maximize the use of information produced by emerging technology solutions, such as gunshot sensors, surveillance video, social media, etc. to put officers in a position to more quickly and effectively prevent—or at least respond to—criminal activities.

Decision-making processes benefit from Big Data, as, unlike Business Intelligence which focuses on explaining current trends, the latter supports at forecasting consumer behavior and market trends through predictive models, thus enabling market players to be proactive.

In short, Big Data can help to take advantage of the ever-growing amount of digital data produced every day. However, concerns have been raised regarding the lack of legislation framing the use of Big Data and particularly the rules enforcing the protection of data **privacy**.

Privacy at the age of Yottabyte: a continuous challenge

Every day, “individuals leave permanent traces across cyberspace, like silkworms depositing “silk” which is then “farmed” for different purposes by public and private organizations”⁵. As personal data such as photos, email addresses, bank details, medical data, or IP addresses may be collected for the purpose of a subsequent Big Data analysis, the risk of retaining personal or sensitive information about individuals without their consent is heightened. This is a major concern as the “right of individuals to respect for their private and family life, home and communications” is a **fundamental right** governed by **EU law**, and therefore a **legal obligation** for all market actors.

Furthermore, Big Data is, by its very nature, able to connect different data, which would not be considered as personal data if processed alone, but once aggregated, may lead to individuals being identifiable. This is a particularly challenging feature of Big Data as it may limit **traditional de-identification methods**; in this case, even anonymized data could be re-identified and attributed to an individual, which may create a data privacy breach.

⁴ *Big data: opportunities and privacy concerns*, Piotr Bąkowski, Members’ Research Service, European Parliamentary Research Service [http://www.europarl.europa.eu/RegData/bibliotheque/briefing/2014/140771/LDM_BRI\(2014\)140771_REV1_EN.pdf](http://www.europarl.europa.eu/RegData/bibliotheque/briefing/2014/140771/LDM_BRI(2014)140771_REV1_EN.pdf)

⁵ *Big and Open Data in Transport*, Houses of Parliament, July 2014, <http://researchbriefings.files.parliament.uk/documents/POST-PN-472/POST-PN-472.pdf>

Big Data may consequently pose a threat to privacy due to misuse or voluntary abuse, such as behavioral tracking, profiling, or discrimination systems using data and algorithms to include and exclude people from various programs and make decisions based on the scoring of individuals. The capacity to analyze and cross a various amount of data to detect patterns could also lead to mass surveillance, which is, for instance, the consequence of the United States National Security Agency's PRISM program. Finally, Big Data could also lead to the disclosure and loss of control of confidential information due to database sharing with third parties or a lack of proper data security.

By its nature, Big Data is a challenge for privacy, and **classic data protection principles** cannot be applied as such. For instance, Article 4 of the Luxembourg Data Protection law on data quality, requiring the principles of necessity and proportionality to be applied, is incompatible with the very nature of Big Data.

Big Data & Privacy: the possible reconciliation

Despite the concerns outlined above, Big Data is a growing trend in EU legislation seen as "an enabler for Big Data services in Europe"⁶. This tendency is embodied by various concrete actions such as the Directive on the re-use of public sector information promoting the availability of information produced, collected and commissioned by the public sector in the EU, also called open data policy.

The recent EU General Data Protection Reform stated that data protection should not prevent innovation enabled by Big Data but, on the contrary, be considered as a competitive edge and a condition for success.

The remaining question is then, how can I launch Big Data projects while abiding by the current legislation? For now, the following may be initiated to deal with the current data protection framework.

Big Data can help take advantage of the ever-growing amount of digital data produced every day

On the one hand, a Privacy Impact Assessment (PIA) should be performed when determining the processing of data, with a special focus on the specific purposes for using Big Data.

Moreover, data controllers should consider privacy throughout the Big Data processing's lifecycle meaning at the time of data collection and access, data retention and aggregation and data analysis and presentation.

In addition, transparency should be a key principle for improving the quality and accuracy of the data processed. It entails informing on how, by whom and under whose responsibility personal data will be processed. Early communication at the very beginning of the project with your national data protection authority can provide helpful insight.

Data security is another important topic to be addressed in accordance with the Privacy Impact Assessment. Personal and sensitive data classification should be defined and audit and traceability of data from the origin should be ensured.

Finally, data integrity should be guaranteed and any access or modifications should be restricted via means such as granular access control and encryption of the data-at-rest and in-transit.

6 Report of workshop on Privacy, Consumers, Competition and Big Data 2 June, [https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared-Documents/Consultation/Big percent20data/14-07_11_EDPS_Report_Workshop_Big_data_EN.pdf](https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared-Documents/Consultation/Big%20data/14-07_11_EDPS_Report_Workshop_Big_data_EN.pdf)

7 The EU Data Protection Reform and Big Data—Factsheet—April 2015 http://ec.europa.eu/justice/data-protection/files/data-protection-big-data_factsheet_web_en.pdf

