BRUSSELS PRIVACY HUB
WORKING PAPER
VOL. 4 • N° 13 • SEPTEMBER 2018

BRUSSELS
PRIVACY
HUB
Vrije Universiteit Brussel

BIG DATA ANALYTICS BY TELECOMMUNICATIONS OPERATORS AND THE DRAFT ePRIVACY REGULATION

by Vagelis Papakonstantinou & Paul de Hert

Abstract

Big data analytics has been defined by the EPDS, under a common denominator approach, as the practice of combining huge volumes of diversely sourced information and analysing them, using more sophisticated algorithms to inform decisions. Notwithstanding the discussion whether personal data constitute a new asset for companies, the fact remains that organisations find new value through constant re-processing of personal data either already in their possession or coming from different third parties. Indeed, over the past few years, big data analytics have forcefully entered the mainstream. However, despite public attention and high volumes of expert analyses, that customarily focus on challenges to personal data protection by similar operations, the vast majority of approaches remains purely theoretical; Aim of this Working Paper is to map and analyse a specific field, focusing on what is actually taking place and what big data analytics operations do take place within telecommunications organisations today.

Keywords: GDPR, ePrivacy Regulation, ePrivacy Directive

Contents

The Brussels Privacy Hub Working Papers are intended to circulate research in progress for comment and discussion. Available at www.brusselsprivacyhub.org/publications.html
ISSN N° 2565-9979. This version is for academic use only.

Disclaimer

Introduction

Big data analytics has been defined by the EPDS, under a common denominator approach, as the practice of combining huge volumes of diversely sourced information and analysing them, using more sophisticated algorithms to inform decisions. Notwithstanding the discussion whether personal data constitute a new asset for companies,[1] the fact remains that organisations find new value through constant re-processing of personal data either already in their possession or coming from different third parties. Indeed, over the past few years, big data analytics have forcefully entered the mainstream. Within the EU they are considered an important part of the European Digital Data Economy.

Despite public attention and high volumes of expert analyses, that customarily focus on challenges to personal data protection by similar operations, the vast majority of approaches remains purely theoretical. **What is actually taking place, what operations do take place within organisations today, is largely unexplored**. This is both on account of unwilling to disclose organisations, as well as, unwilling (or unable) to get involved DPAs. Accordingly, the EDPS speaks of the "*black box*" of big data analytics. Both the Article 29WP and the EDPS have asked for an "*innovative*" mind-set from all parties concerned, when implementing data protection principles on big data analytics operations.

Telecommunications operators ("*telecommunication network operators*", offering telephony and telephony-related services, to be distinguished from Internet Service Providers) are prime candidates, or even already prime users, of big data analytics. They sit on a wealth of subscriber information collected in their course of business. These data are, or can be, continuously re-examined in view of extracting added value from them.

On 7 December 2017 a by-invitation-only workshop was organised by the Brussels Privacy Hub, whereby views were exchanged on **(a)** current data analytics processing scenarios, as well as, **(b)** on difficulties encountered while pursuing compliance with the current regulatory framework (the General Data Protection Regulation and the draft ePrivacy Regulation, as proposed by the Commission).

On 22 February 2018 an open workshop was organised by the Brussels Privacy Hub, where participants from Data Protection Authorities, the industry and academia exchanged views on **(a)** the data analytics processing scenarios, as formulated above, and **(b)** compliance options and difficulties.

On 26 September 2018 a draft final report was discussed and validated in a by-invitation-only workshop organised by the Brussels Privacy Hub.

[1] For example, World Economic Forum, "Personal Data: The Emergence of a New Asset Class", 2011. This would however open up discussions on proprietary rights over personal data.

1. Personal data processing by telecommunications operators

1.1. A need for classification of services rendered / service providers

Consensus is not met on classification of the services (essentially, internet access and voice communication, the "**basic services**") provided by telecommunications operators. While the traditional approach considers them **essential services** to individuals,[2] the argument can be made that technological developments have by now enabled provision of the same services in a number of alternative ways for individuals (for example, WiFi networks or internet telephony). In other words, providers of the basic services are by not only telecommunications operators but also internet service providers (that may include large multi-national entities and also SMEs).

A distinction may also be made between the above basic services and added-value electronic communication services ("**added-value services**"). These include, for example, geolocation services, video (television) services, handset-related services (upgrade/replacement), contract-related services (bundle sale, friends/family/company packages), etc. Added-value services are related but are not necessary to the provision of the basic services.

Finally, a series of services (for example, fraud detection, signal improvement, customer care, technical support) are services related and necessary for the provision of the basic services ("**necessary services**")

1.2. Categories of personal data collected by telecommunications operators

Telecommunications operators In their course of business typically collect, and are in possession of, the following categories of personal data:

1. **Subscriber personal data**.
   These are the data provided by subscribers themselves during the application process;
2. **Metadata**.
   These are the use of services data, potentially including and or all of the following categories:
   - Numbers called;
   - Websites visited, URL data;
   - Location data / data related to the connection point of the terminal equipment (antenna, Wi-Fi connection point, home router; **however, no GPS data is collected**);
   - Use of services personal data;
   - TV channels viewed;
   - Date, time, duration of the communication;
   - Type of communication (voice, data).

Various qualifies may be attached to metadata: They can be necessary for the provision of the service, or useful for the improvement of its quality. They can also be of low or of high privacy risk. Certain categories can also reveal behavioural patterns by subscribers (behavioural data).

[2] See also Recital 18 of the ePrivacy draft Regulation (Commission proposal).

3. **Content of the communication (voice call, SMS, voice mail message, email content, messaging).**

Content of the communication is normally not accessed by telecommunications operators (the cases of access/processing exceeding the purposes of this Working Paper).

1.3. Types of personal data processing

The basic assumptions on big data analytics operations already carried out or planned for the future to take place within a typical telecommunications operator refer to the following:

1. **Billing – Customer support (most likely, to be classified under "necessary services")**
    - Personal data collected during application process;
    - For all new subscribers;
    - Necessary for the provision of the service;
    - Required to answer questions / complaints from customers;
    - Invoice related: For example, to explain data volume invoiced;
    - Any roaming information.

2. **Simple internal marketing (most likely, to be classified under "added-value services")**
    - Examples: Offers on contract renewals, upselling, cross selling;
    - Sometimes based on legal obligation; better offer, BIPT tariff simulator;
    - Using call details (e.g. family and friends only) / volume of data or voice (metadata).

3. **Customer insight models (most likely, to be classified under "necessary services")**
    - Examples: Reasons why people call customer service / bad device detection / detection of family circles;
    - Using various metadata;
    - Processing run periodically (e.g. once every two months);
    - Automated-decision making: subscribers may (or may not) be made offers on the basis of findings;
    - Opt-out option for subscribers only on whether to be contacted, not on running the processing itself;
    - Fraud detection/mitigation.

4. **Data analytics for third parties (most likely, to be classified under "added-value services")**
    - Examples: Smart cities applications (counting visitors in a city or recurring visitors);
    - Upon third parties' request (and contract);
    - Data are aggregated / pseudonymized / anonymized;
    - No access to the raw data by these third parties / only findings in reports, or customer profiles.

2. The applicable legal framework

From a data protection perspective, the General Data Protection Regulation and the ePrivacy provisions form the applicable legal framework. The ePrivacy legal instruments being currently under processing, attention will be paid to the current (Commission-released) draft for an ePrivacy Regulation. As regards their relationship, for the moment the two instruments are complementary, in the sense that the ePrivacy Regulation "*complements and particularises*" the GDPR.[3] In addition, the ePrivacy Regulation does not intend to "*lower the level of protection*" afforded by the GDPR[4].

What is important to the purposes of this Working Paper is whether big data analytics carried out by telecommunications operators are regulated by the ePrivacy Regulation or the GDPR. **For an intermediate period, until the ePrivacy Regulation is released and comes into effect, it would appear that the two instruments will have to co-exist: In particular, although metadata and the lawful bases for the processing are set in the ePrivacy Regulation,[5] reference will ultimately be made to the provisions of the GDPR, as applicable to relevant processing**.

2.1. The GDPR and big data analytics

Personal data and anonymisation

The GDPR sets the EU data protection scene, but also made forays into ePrivacy area, for example when setting that location data are personal data (Art 4.1), despite its Article 95. Other than that, the issue of **data anonymization** is central to this Working Paper. The GDPR excludes anonymous information from its scope, and introduces a proportionality criterion thereof, "*on reasonably likely to be used means to identify a natural person*".[6]

However, the Art29WP has raised the bar considerably with regard to anonymous data: "*Anonymisation of data cannot be achieved by just stripping a dataset of some directly identifying attributes. The bigger and the more comprehensive a collection of data becomes, the more possibilities exist to identify the individuals whom the data relates to, especially when data is retained for longer periods of time and/or shared*". The notion of "**linkability**" is used by it in this regard: *"[consent would still be required] if the data are simply hashed, aggregated on an event level or otherwise pseudonymized, but there remains a possibility to single out users, or linkability of individual events in the aggregated data to the original data, also if future collection of traffic or location data creates linkability to events in the aggregated data set*".

In view of the above, it has been established that "full" anonymisation is not at all possible within the above processing scenarios, particularly given the "linkability" criterion.

Taking into account supervisory authorities' extremely restrictive approach, it appears that any data analytics operation cannot apply an "anonymisation" approach, that would otherwise exempt it from data protection law requirements, in order for it to be lawfully performed. They thus need to apply data protection legislation and related criteria.

[3] Article 1.3.
[4] Recital 5.
[5] In Articles 4 and 6.
[6] Recital 26. See also Article 11.

Regarding the **value of anonymised data**, it was found that they **lose substantially, if not entirely,** in value for all parties involved in the processing (for example, municipalities in smart cities scenarios, telecoms operators in client management). If data cannot even be aggregated per event or other categories or pseudonymized, then they lose their intrinsic value for the Data Driven / Digital Data Economy. On the other hand, the Art29WP, for example, states that *"most legitimate processing of location data and other metadata does not require a unique identifier"*. It also suggested in the past that *partial anonymization* or **partial de-identification** (key-coding, keyed-hashing, replacing unique IDs) may be the appropriate solution in some situations when complete anonymisation is not practically feasible. None of the above, however, exempt data protection law from relevant personal data processing.

Lawful bases for the processing – Consent

The GDPR outlines the acceptable, lawful bases for personal data processing.[7] The ePrivacy Regulation complements them further.[8] Attention has been brought to **consent**, and on whether it should be the only acceptable legal basis for the processing in big data analytics.

While this would be the preferred approach of supervisory authorities, substantial difficulties need to be surpassed for this to become the case.

On the one hand, the Art29WP places high standards on what constitutes **lawful consent**: "*Mere silence or inaction, such as in the case of default settings of online social networks or web browsers, is not valid. Consent should be requested prior to the data processing and only after the data controller has given notice to the data subject of the processing operations in clear and understandable language*". Similarly, "*bundled consent for processing for multiple purposes*" should not be allowed, and therefore "*consent should be granular*".

On the other hand, telecommunications operators face a series of difficulties in their attempt to acquire such, *granular*, consent:

- ***Existing*** **subscribers who may need to renew or update their consent to processing of their personal data are likely to be indifferent/not care to take any action (form submission etc.);**
- ***New*** **subscribers, while they can be asked to consent while entering their new contracts, may also soon need to renew their consent, each time a new data analytics operation is planned;**
- **Efficient management requires the planning and running of new types of processing at a constant pace;**
- **Granularity in consent means that, normally, subscribers have to enter multiple consents within a short period of time.**

The above situation is bound to:

- **Lead to *"consent fatigue"*, as is the case today with cookies, whereby internet users are most likely to apply a mechanistic acceptance approach, undermining thus the data protection purposes;**
- Lead to substantial compliance difficulties, requiring disproportionate effort, for telecommunications' operators.

[7] Article 6.
[8] Article 6.

The basic data protection principles: Data specification and further processing

The basic data protection principles are outlined in the GDPR.[9] Of particular relevance to big data analytics is the principle of purpose specification. Some flexibility is provided in the GDPR itself.[10] Further processing is permitted, even without consent, when the new purpose is compatible with the initial. To do this, "any link" to the original purposes needs to be assessed, or the context/relationship between data subjects and the controller.

The Article 29 Working Party recommended that compatibility should be assessed in the light of the context in which the data were collected, of reasonable expectations of the data subjects, of the nature of the personal data in question, of the impact of further processing, and of safeguards to protect the data subject.

In view of the substantial difficulties encountered while assessing the feasibility of individual consent as the legal basis for data analytics processing by telecommunications operators, it appears that the notions of "*further processing*" and "*legitimate interest of the controller*" can be the applicable legal grounds for such types of processing.

2.2. The draft ePrivacy Regulation and big data analytics

A first point to be noted is that there is a difference in scope between the GDPR and the ePrivacy legislation. Since its introduction, ePrivacy EU legislation was aimed at protecting privacy and the private life of individuals, as well as, confidentiality of communications. On the other hand, the GDPR protects the individual right to data protection. The right to privacy has in theory a wider scope than the right to data protection.

The second point refers to the relationship between the two instruments. As said, the ePrivacy regulation will "particularise" the GDPR. Unless expressly regulated by its provisions, any processing of personal data matter will fall under the GDPR scope.

With regard to the subject-matter of this Working Paper, the proposed ePrivacy Regulation most notably:

- Treats "metadata" in bulk, as "data processed in an electronic communications network for the purposes of transmitting, distributing or exchanging electronic communications content";[11]

- Acknowledges two types of "location data", the first one referring to data generated in the context of providing electronic communication services, and the second on data generated other than in the context of providing such;[12]

[9] Article 5.
[10] Article 6.4.
[11] Article 4.
[12] Recital 17.

- Acknowledges three (3) lawful bases for the processing of metadata: mandatory quality of service requirements, billing purposes or against fraudulent and abusive use, and user consent, "provided that the purpose or purposes concerned could not be fulfilled by processing information that is made anonymous".[13]

Both the EDPS and the Art29WP, in contrast with the telecommunications sector, wish for the same level of protection for both metadata and content, and for both to be placed under a consent-only legal basis for processing (admittedly with some flexibility as regards better technical provision of services, and the provision of services requested by the user).

It remains questionable if, under the current wording, other lawful bases of the GDPR are applicable for the processing of metadata. It also remains to be seen whether, even under the consent legal basis, further processing may take place under the ePrivacy Regulation.

The EDPS, on October 2017, issued recommendations against including the "legitimate interests" of the operators as an additional legal basis for the processing of metadata, nor "further processing".

Nevertheless, telecommunications operators raise the following with regard to metadata:

- "Metadata" are connected to technological developments, and therefore their exact content may vary substantially from time to time; consequently, no blanket regulation may cover them;

- The same is the case with "location data", that too is a dynamic term difficult to analyse in a constantly changing processing environment: for example, WiFi networks collect all data within their reach and not only those of customers who consent, for example, these of passers-by;

- "Location data" are not sensitive data ***per se:*** The level of geographical identification in the telecommunications context is not as accurate, as to permit exact location of an individual (ie. a mobile set is placed within a wider region);

- Location data are customarily held in separate databases by telecommunications operators than metadata; only through their possible connection with metadata is identification of an individual possible.

In view of the above, telecommunications operators consider that the request to treat content of communication and metadata in the same manner is not reasonable and justified. A preferable approach would be for content of the communications to be regulated by the ePrivacy Regulation and metadata by the GDPR.

[13] Article 6.2.

Findings – Towards A New Guideline (Checklist)

Difficulties in implementation of applicable legal framework and compliance are caused by a certain lack of specificity and common understanding, mostly as follows:

- **Blanket treatment of "services":** Electronic communication services can be essential, necessary or added-value services; Each type merits different legal treatment;

- **Blanket treatment of "service providers":** Service providers can be anything from traditional telecommunications operators, global internet service providers and SMEs, each one providing one or more of the above services (including access); each one merits different legal treatment;

- **Blanket treatment of the term "metadata":** Current treatment of metadata lacks flexibility and technical/business understanding; Most pressingly, "metadata" needs to be distinguished from "location data.

In view of the above, the draft ePrivacy Regulation appears to lack the specificity to warrant effective regulation of data analytics operations carried out by telecommunications operators today. A more nuanced approach would be necessary, acknowledging the particularities of relevant processing. The required level of specificity could either be achieved, if not in the legislative text per se, through secondary guidance by competent supervisory (EU or Member State) authorities.

Basic bibliography (official reports only)

Article 29 Working Party, Opinion 1/2017 on the proposed Regulation for the ePrivacy Regulation

Article 29 Working Party, Statement 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, 2014

Article 29 Working Party, Opinion 05/2014 on Anonymisation techniques

Article 29 Working Party, Opinion 3/2013 on the purpose specification principle

CPVP, Big Data Rapport, AH-2016-0154

European Commission Communication, Building A European Data Economy, 2017

European Commission Communication, Towards a thriving data-driven economy, 2014

EDPS, Opinion 6/2017 on the proposal for a Regulation on Privacy and Electronic Communications (ePrivacy Regulation)

EDPS, Opinion 8/2016 on coherent enforcement of fundamental rights in the age of big data

EDPS, Opinion 7/2015, Meeting the challenges of big data

EDPS, Preliminary Opinion on privacy and competitiveness in the age of big data, 2014.

The Brussels Privacy Hub Working Papers series

N°1 "The data protection regime applying to the inter-agency cooperation and future architecture of the EU criminal justice and law enforcement area" (November 2014) by Paul De Hert and Vagelis Papakonstantinou (35 pages)

N°2 "The new cloud computing ISO/IEC 27018 standard through the lens of the EU legislation on data protection" (November 2014) by Paul de Hert, Vagelis Papakonstantinou, Irene Kamara (25 pages)

N°3 "Towards efficient cooperation between supervisory authorities in the area of data privacy law" (October 2015) by Dariusz Kloza, Antonella Galetta (24 pages)

N°4 "The data protection regime in China" (November 2015) by Paul De Hert and Vagelis Papakonstantinou (30 pages)

N°5 "The right to privacy and personal data protection in Brazil: time for internet privacy rights?" (February 2016) by Vinícius Borges Fortes (23 pages)

N°6 "Permissions and Prohibitions in Data Protection Jurisdiction" (May 2016) by Mistale Taylor (25 pages)

N°7 "Structure and Enforcement of Data Privacy Law in South Korea" (October 2016) by Haksoo Ko, John Leitner, Eunsoo Kim and Jong-Gu Jung (20 pages)

N°8 "The "Right to be Forgotten" and Search Engine Liability" (December 2016) by Hiroshi Miyashita (15 pages)

N°9 "European Human Rights, Criminal Surveillance, and Intelligence Surveillance: Towards "Good Enough" Oversight, Preferably but Not Necessarily by Judges" (March 2017) by Gianclaudio Malgieri & Paul De Hert (25 pages)

N°10 "Does Technology Drive Law? The Dilemma of Technological Exceptionalism in Cyberlaw" (July 2017) by Meg Leta Jones, JD, PhD (31 pages)

N°11 "The Microsoft Ireland case and the cyberspace sovereignty trilemma. Post-territorial technologies and companies question territorial state sovereignty and regulatory state monopolies" (July 2018) by Paul De Hert and Johannes Thumfart (27 pages)

N°12 "Understanding the balancing act behind the legitimate interest of the controller ground: a pragmatic approach" (August 2018) by Irene Kamara and Paul De Hert (35 pages)

N°13 "Big data analytics by telecommunications operators and the draft ePrivacy Regulation" (September 2018) by Vagelis Papakonstantinou & Paul de Hert (13 pages)

The Brussels Privacy Hub Working Papers series

The Brussels Privacy Hub Working Papers are intended to circulate research in progress for comment and discussion. The Working Papers focus on all areas of data protection and privacy research and can contain empirical research on privacy issues and analytical work on privacy governance and regulation in the EU; global flows of data; reconciling law enforcement and privacy interests; privacy challenges posed by new technologies; comparative research on privacy in different regions; jurisprudential issues of privacy protection; and many others

Available at www.brusselsprivacyhub.org/publications.html

Contact: info@brusselsprivacyhub.org

See list on previous page

VUB

BRUSSELS
PRIVACY
HUB