Towards a big data regulation based on social and ethical values. The guidelines of the Council of Europe

Hacia una regulación de los datos masivos basada en valores sociales y éticos. Las directrices del Consejo de Europa

Cap a una regulació de les dades massives basada en valors socials i ètics. Les directrius del Consell d’Europa

Alessandro Mantelero *
Abstract

This article discusses the main provisions of the Guidelines on big data and data protection recently adopted by the Consultative Committee of the Council of Europe. After an analysis of the changes in data processing caused by the use of the predictive analytics, the author outlines the impact assessment model suggested by the Guidelines to tackle the potential risks of big data applications. This procedure of risk-assessment represents a key element to address the challenges of Big Data, since it goes beyond the traditional data protection impact assessment encompassing the social and ethical consequences of the use of data, which are the most important and critical aspects of the future algorithmic society.

Keywords: big data; data protection; privacy impact assessment; Council of Europe.

Resumen

Este artículo ofrece un análisis de las principales disposiciones de las Directrices sobre datos masivos y protección de datos recientemente aprobadas por el Comité Consultivo del Consejo de Europa. Después de un examen de los cambios ocurridos en el procesamiento de datos por el uso de la analítica descriptiva, el autor describe el modelo de evaluación de impacto que se sugiere en las Directrices para encarar los riesgos potenciales de las aplicaciones que utilizan la analítica de datos masivos. Este procedimiento de evaluación de riesgos es un elemento clave para gestionar los datos masivos, ya que no se limita a la evaluación tradicional del impacto en la protección de datos sino que abarca también las consecuencias sociales y éticas de su uso, aspectos que son importantes y críticos de la futura sociedad algorítmica.

Palabras clave: datos masivos; protección de datos; evaluación de impacto de la protección de datos; Consejo de Europa.

Resum

Aquest article ofereix una anàlisi de les principals Directrius sobre dades massives i protecció de dades recentment aprovades pel Comitè Consultiu del Consell d’Europa. Després d’un examen dels canvis ocorreguts en el processament de dades mitjançant l’ús de l’analítica descriptiva, l’autor descriu el model d’avaluació d’impacte que se suggereix en les Directrius a fi d’encarar els riscos potencials de les aplicacions que utilitzen l’analítica de dades massives. Aquest procediment d’avaluació de riscos és un element clau per gestionar les dades massives, ja que no es limita a l’avaluació tradicional de l’impacte en la protecció de dades sinó que inclou també les conseqüències socials i ètiques del seu ús, aspectes crítics de la futura societat algorítmica.

Paraules clau: dades massives; protecció de dades; avaluació d’impacte de la protecció de dades; Consell d’Europa.
1. Introduction

The “Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data” (hereinafter Guidelines), adopted by the Consultative Committee of Convention 108 in January 2017 represents the first international guidance on the use of big data, which is a form of data processing that rises significant questions concerning the protection of fundamental rights.¹

The role of individual self-determination with regard to the use of data and the risk assessment of big data applications represent two of the main aspects of the Guidelines and, in this regard, the Council of Europe suggests novel solutions to address the challenges of the new data processing paradigm based on analytics.

In light of the above, this article is divided into two main parts: the second section describes the impact of the new model of predictive analysis on the main principles of data protection regulation, while the third section discusses the provisions of the Guidelines and focuses on the risk assessment procedure adopted by the Consultative Committee.

2. Big Data: a new paradigm of data processing

The advent of big data analytics² has suggested a new paradigm in portraying our societies, where the traditional approach adopted in statistical studies is complemented or replaced by predictive analysis. Data visualization has played a relevant role in this change, making it possible real-time analysis of streams of data and prediction of their future trends.³ Moreover, algorithms are used to discover hidden correlations between the variables that characterize large datasets.

---


² This term is used to identify computational technologies that analyse large amounts of data to uncover hidden patterns, trends and correlations. According to the European Union Agency for Network and Information Security (ENISA), the term Big Data analytics “refers to the whole data management lifecycle of collecting, organizing and analysing data to discover patterns, to infer situations or states, to predict and to understand behaviours”. See ENISA, 2014.

³ See Bollier, 2010.
Like in the past, with regard to the traditional statistical studies, this kind of analysis is not exclusively carried on for mere scientific purposes, but is mainly conducted to provide insights about individuals and society to decision-makers. This relationship between data processing and the adoption of strategic decisions—which affect individuals in different contexts (e.g. financial services, healthcare services, urban planning)—has become progressively stronger by reason of the increased availability of data as a result of the so-called datification process.

This is an ongoing process to “capture quantifiable information”\textsuperscript{4}, which aims to transform reality into data.\textsuperscript{5} In this sense, even human beings can be considered as aggregates of information,\textsuperscript{6} which represent their private or public identities. This relationship between individual nature and personal data has been recognised by the courts in various decisions concerning the right to privacy and—more recently—the right to be forgotten.\textsuperscript{7}

Nevertheless, the complexity of human beings cannot be reduced to a mere aggregate of data, since they are primarily persons. For this reason, the information referring to them are not neutral or raw data that, in the present digital economy, can be freely used and assimilated to mere goods, regardless they are qualified as private or common. Personal data, as well as the other forms of expression of a given person, are part of her identity and, therefore, should be safeguarded within the framework of personality rights and fundamental rights and freedoms.\textsuperscript{8}

According to this theoretical framework, individual name, image and social identity (i.e. reputation and honour) have been recognised as attributes of human beings and safeguarded by law over the centuries. Against this scenario, the right to the protection of personal data represents the most recent development of the category of personality rights, since its origin is strictly related to the early stages of the digitalisation of information, which is the foundational stone of the present process of datafication and of our (big) data-driven society.

\textsuperscript{4}See Mayer-Schönberger and Cukier, 2013, 78. See also Lycett, 2013; Ericsson, 2014.

\textsuperscript{5}An early step in this strategy of datafication was made by the Auto-ID Center which, set up in 1999, was a not-for-profit consortium created with the aim “to develop a system for using the Internet to identify goods anywhere in the world”.

\textsuperscript{6}See Floridi, 2014, 96-98 (“the fourth revolution has brought to light the intrinsically informational nature of human identity”) and Mayer-Schönberger and Cukier, 2013, 93 (“Datafication is not just about rendering attitudes and sentiments into an analysable form, but human behavior as well”).

\textsuperscript{7}See Court of Justice of the European Union, 13 May 2014, case C-131/12, Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González. Available at: http://curia.europa.eu (accessed 4 March 2017).

\textsuperscript{8}On the notion of personality right see, inter alia, Resta, 2014; Cannataci, 2008; Alpa and Resta, 2006.
Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero

Rev Bio y Der. 2017; 41: 67-84

www.bioeticayderecho.ub.edu - ISSN 1886-5887

Like the right to privacy was the answer to the assault to the private sphere conducted by the penny press at the end of the XIX century,\(^9\) the right to the protection of personal information is the answer given by legislators in the '70s to the rising citizens' concern about the risks of new forms of computer-based social control. Over the years, this risk has changed its nature and source, from the original concern about government surveillance and economic exploitation of personal information to the present public and private partnership in surveillance and the adoption of information-based predictive decision-making systems.

This increasing exploitation of personal information and the development of data processing technologies led legislators to adopt different procedural regulations on data protection and, since the so-called second generation of data protection laws,\(^10\) the right to the protection of personal information was placed in the wider context of fundamental rights. In this sense, the Convention 108 on the Protection of Individuals with regard to Automatic Processing of Personal Data, adopted by the Council of Europe in 1981, considers data protection as an expression of the broader right to privacy.\(^11\)

Although nowadays the right to data protection is—more correctly—considered as an autonomous right, which differs from the right to privacy, the qualification provided by the Convention 108 was consistent with the theoretical framework at that time. Anyway, the most important element of this qualification consists in the level of protection accorded to personal information. In this sense, the protection of personal information is put at the highest level, in the context of fundamental rights.

More recently, the Charter of Fundamental Rights of the European Union has expressly recognised the “right to the protection of personal data”\(^12\) as an autonomous right, different from the right to respect for private and family life, and has grant to everyone “the right to the protection of personal data concerning him or her”.

Against this scenario, the European model of data protection is based on the safeguard of the data subject’s individual right to control “his or her personal data and the processing of such data”.\(^13\) This is in line with the original notion of data protection as data control that was

---


\(^11\) See Article 1, Convention 108. See also Parliamentary Assembly of the Council of Europe, Recommendation 509 (1968) on Human rights and modern scientific and technological developments (1968).

\(^12\) See Charter of Fundamental Rights of the European Union, Article 8.

\(^13\) See the Preamble of the Draft Modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data (hereafter Draft Modernised Convention 108). Consolidated text of the modernisation proposals of
elaborated by legal scholars in the ’70s, which led legislators to adopt a model of protection primarily focused on the individual dimension.

Nevertheless, both this idea of control over personal information and the notion of data protection as an individual right show their limits in the context of the present forms of data processing based on analytics. In this sense, on the one hand, the traditional paradigm of “notice and consent” does not adequately address the complexity of data processing and, on the other hand, data collection and analysis are even more focused on the collective dimension, due to their attempt to understand, predict and orient the behaviour of groups of persons.

The use of big data analytics creates “a new truth regime”, in which general strategies are adopted on a large scale on the basis of descriptions of society generated by algorithms, which predict future collective behaviour. These strategies are then applied to specific individuals, given the fact that they are part of one or more groups generated by analytics.

Nevertheless, this “categorical” approach characterizing the use of analytics, leads decision-makers to adopt common solutions for individuals belonging to the same cluster generated by analytics, without considering each individual per se, her unique identity that may differ from the stereotypical models created by algorithms.

---

15 See Bygrave, 2004, 324-325; Solove, 2008, ch. 2. See also Bennett and Raab, 2003, ch. 1; Post, 1989; Cohen, 2000, 1426-1428. The collective dimension of this right has been recognised in the U.S. and Europe, but protected mainly indirectly, as an aggregation of individual privacy issues and not as an autonomous dimension. See Mantelero, 2016.
16 See Mantelero, 2014; Rubinstein, 2013.
21 See also Vedder, 1997, 215–226.
Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero

Rev Bio y Der. 2017; 41: 67-84

In this sense, the use of big data analytics to support decisions exceeds the boundaries of the individual dimension and assumes a collective dimension, with potential harmful consequences for some groups. Therefore, the potential prejudice is no longer circumscribed to the well-known privacy-related risks (e.g. illegitimate use of personal information, data security), but it also concerns the negative impact on other fundamental rights, such as the right to non-discrimination.

Against this background, the adoption of a fundamental rights impact assessment has been proposed by the United Nations Special Rapporteur on the right to privacy (Joe Cannataci), but it seems still far from being developed at global level. Nevertheless, a first step in this direction is the Privacy, Ethical and Social Impact Assessment (PESIA), which has been adopted by the Council of Europe in its Guidelines on Big Data. This model of assessment goes beyond the traditional impact assessment focused on data quality and data security, since it also encompasses the societal consequences of data uses and the analysis of their potential conflicts with ethical values.

Moreover, the intent of the Guidelines to take into account the collective dimension of the use of personal information is not only evident in the scope of PESIA, but also in the participatory model adopted in the assessment procedure, which aims to give voice to the different stakeholders potentially affected by data processing.

Although there are several provisions of these Guidelines that suggest novel approaches in protecting personal information and fundamental rights in the big data environment (e.g. the provisions concerning the role of the human intervention in Big Data-supported decisions) the

---

22 See Mantelero, 2017, 139-158. See also Vedder, 1997.
23 See also Crawford et al., 2013, 6-7; Boyd, Levy, and Marwick, 2014, 56.
25 Regarding the PESIA model, see also the H2020 project “VIRT-EU: Values and ethics in Innovation for Responsible Technology in Europe”. Available at: http://www.virteuproject.eu/ (accessed 16 March 2017). With regard to the ethical assessment in research and innovation, see Shelley-Egan et al., 2014, 42-44.
26 See Mantelero, 2016.
27 A similar participatory approach is also adopted by the UNESCO Universal Declaration on Bioethics and Human Rights (Article 18).
28 See Guidelines, Section IV, para 7.
Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero

Rev Bio y Der. 2017; 41: 67-84

www.bioeticayderecho.ub.edu - ISSN 1886-5887

risk assessment procedure represents the core of the Guidelines. In this sense, the risk assessment procedure\textsuperscript{29} plays a central role with regard to different elements concerning the social and ethical dimensions of data uses, data subject's self-determination, the relationships between the purposes of data collection and data uses,\textsuperscript{30} the by design approach\textsuperscript{31} and the use of anonymous data\textsuperscript{32} and open data.\textsuperscript{33}

3. The PESIA model in the context of the Guidelines of the Council of Europe

The Guidelines adopted by the Consultative Committee of Convention 108 are non-legally binding practical and operative instructions provided by the Council of Europe to the Parties of the Convention. This is in line with the regulatory model of the Council of Europe, which adopts a principle-based approach complemented by guidelines that provide a sector-specific interpretations of the principles of the Convention. Through the adoption of its guidelines, the Consultative Committee aims to facilitate an effective application of the principles of the Convention.\textsuperscript{34} In this sense, the Guidelines are primarily addressed to data controllers and data processors.

The main limit of these Guidelines regards their scope, since they concern a given technology in general (i.e. big data), rather than its application in a given sector (e.g. healthcare services). For this reason, the reached result is not completely satisfactory for various operators, who would like to have specific answers with regard to the applications of analytics in given fields.

This outcome is inevitable due to the wide range of big data applications, but the awareness of this limit led the Consultative Committee of the Convention to recognise in the Guidelines that "given the expanding breadth of Big Data in various sector-specific applications, the present

\textsuperscript{29} A central role to risk assessment and risk management is also recognised by the UNESCO Universal Declaration on Bioethics and Human Rights, see Article 20.

\textsuperscript{30} See Guidelines, Section IV, para 3.1 ("Exposing data subjects to different risks or greater risks than those contemplated by the initial purposes could be considered as a case of further processing of data in an unexpected manner"). On the purpose limitation principle, see also Article 9 of the UNESCO Universal Declaration on Bioethics and Human Rights.

\textsuperscript{31} See Guidelines, Section IV, para 4.

\textsuperscript{32} See Guidelines, Section IV, para 6.

\textsuperscript{33} See Guidelines, Section IV, para 8.

\textsuperscript{34} See Guidelines, Section II (Scope).
Guidelines provide a general guidance, which may be complemented by further guidance and tailored best practices on the protection of individuals within specific fields of application of Big Data (e.g. health sector, financial sector, public sector such as law enforcement). Apart from this limit, the Guidelines represent an important step in regulating big data use, since the issues concerning analytics are not specifically addressed by the most recent data protection regulations, such as Regulation (EU) 2016/679.

The main instrument to address the potential negative impact of big data on individuals and society is represented by risk management. In defining the key principles for risk management, the Guidelines suggest the adoption of a precautionary approach to regulating data protection in the field of big data.

The precautionary approach is adopted with regard to any new application of technology that may produce potential risks for individuals and society, which cannot be exactly calculated or quantified in advance. In this sense, the obscurity of big data uses, the uncertainty characterising the concrete applications of data science and the potential high impact of big data analytics on essential aspects of society may warrant the adoption of this approach as the default setting.

Regarding the scope of risk assessment, while in the Regulation (EU) 2016/679 – as well as in Directive 95/46/EC – it mainly focuses on data security and data quality, in the Guidelines the Data Protection Impact Assessment evolves into a broader and more complex Privacy, Ethical and

35 See Guidelines, Section II (Scope).
36 See Mayer-Schönberger and Padova, 2016, 326, 332. See also Moerel, 2014.
37 See Guidelines, Section IV, para 2.1 (“Given the increasing complexity of data processing and the transformative use of Big Data, the Parties should adopt a precautionary approach in regulating data protection in this field”). On the distinction between the precautionary approach and the precautionary principle, see Peel, 2004.
38 See Guidelines, Section IV, para 2.1 (“Given the increasing complexity of data processing and the transformative use of Big Data, the Parties shall adopt a precautionary approach in regulating data protection in this field”). Only few contributions in law literature take into account the application of the precautionary approach in the field of data protection, see Costa, 2012; Gonçalves, 2017; Gellert, 2015. See also Council of Europe, 2005, para 10; Pieters, 2011, 455. On the precautionary approach in data protection, see also Narayanan, Joanna and Felten, 2016, 357-385; Böröcz, 2016, 476-477; Raab, and Wright, 2012, 364; Lysnskey, 2015, 83; Raab, 2004.
39 Moreover, in Section IV, Paragraph 2.2, the Guidelines require data controllers to adopt “preventive policies” to adequately address and mitigate the potential risks related to the use of big data analytics. This is consistent with Article 8bis (2) of Draft Modernised Convention 108, which focuses on risk analysis and requires that data processing is designed “in such a manner as to prevent or minimise the risk of interference with [...] rights and fundamental freedoms”. On the precautionary principle, see also Tosun, 2013; Aven, 2011; Stirling and Gee, 2002.
Social Impact Assessment (PESIA) to encompasses the societal consequences of data uses mentioned above.\textsuperscript{40}

Obviously, an assessment concerning the compliance of data use with ethical and social values is more complicated than the traditional data protection assessment, since social and ethical values are necessarily context-based and change from one community to another.\textsuperscript{41} In this sense, the Guidelines recognise the relative nature of social and ethical values.\textsuperscript{42}

To address this issue, the Guidelines urge both data controllers and data processors to use personal information in a manner that is not in conflict with the “ethical values commonly accepted in the relevant community or communities and should not prejudice societal interests, values and norms”.\textsuperscript{43}

Moreover, in order to provide a general benchmark of values to be taken into account in conducting the PESIA, the Guidelines identify “the common guiding ethical values” in the international charters of human rights and fundamental freedoms, such as the European Convention on Human Rights.\textsuperscript{44}

Nevertheless, international charters and ethical values commonly accepted in a community\textsuperscript{45} may only provide a high-level guidance. For this reason, the Guidelines combine this general suggestion with a more tailored option, represented by “ad hoc ethics committees”,\textsuperscript{46} which should identify the specific ethical values to be safeguarded with regard to a given use of data, providing more detailed and context-based guidance for risk assessment.

In conclusion, the PESIA model is based on a system of values which is organised on three different layers with a progressive granularity: the “common guiding ethical values” provided by the international charters of human rights, the values and social interests of given communities

\textsuperscript{40} Definition of the PESIA model is still in its infancy, see above fn. 25.
\textsuperscript{41} See Guidelines, Section IV, para 1.
\textsuperscript{42} On the different ethical values and their harmonization, with regard to ethics assessment of research and innovation, see Bray et al., 2015.
\textsuperscript{43} Guidelines, Section IV, para 1.2.
\textsuperscript{44} See also, in this sense, Wright, 2011, 201–202.
\textsuperscript{45} On the importance of cultural diversity and pluralism see also Article 12 of the UNESCO Universal Declaration on Bioethics and Human Rights.
\textsuperscript{46} See Guidelines, Section IV, para 1.3 (“the assessment of the likely impact of an intended data processing described in Section IV.2 highlights a high impact of the use of Big Data on ethical values, controllers could establish an ad hoc ethics committee, or rely on existing ones, to identify the specific ethical values to be safeguarded in the use of data”). With regard to the role played by ethics committees, see also Article 19 and 22.2 of the UNESCO Universal Declaration on Bioethics and Human Rights.
and the tailored application of these values provided by ethics committees, which focuses on a given use of data.

Regarding the procedure of assessment, the Guidelines adopt the traditional circular scheme that characterises the risk-assessment,\textsuperscript{47} which is divided into four stages:\textsuperscript{48} 1) identification of risks, 2) analysis of the potential impact of these risks, 3) selection and adoption of the measures to prevent or mitigate the risks, 4) regular review of the effectiveness of the measures.\textsuperscript{49}

With regard to the measures to prevent or mitigate the risks, the Guidelines also make an explicit reference to by-design and by-default solutions.\textsuperscript{50} The existing strict relationship between risk assessment and solutions by design implies that any change in the nature of the assessment affects the architectural solutions adopted. Thus, the multiple impact assessment suggested by the Council of Europe necessarily leads data controller to consider a broader range of by-design solutions to mitigate the additional ethical and social concerns.\textsuperscript{51}

Given the complexity of this assessment and the various aspects that should be taken into account, it cannot be conducted only by data protection experts, but “should be carried out by persons with adequate professional qualifications and knowledge to evaluate the different impacts, including the legal, social, ethical and technical dimensions”.\textsuperscript{52}

Finally, the collective dimension of the potential impact of the use of data leads the Consultative Committee to encourage the involvement of all the relevant stakeholders, giving voice to the different groups of persons potentially affected by the use of data.\textsuperscript{53}

\textsuperscript{47}See, e.g., CNIL, 2015.
\textsuperscript{48}See Guidelines, Section IV, para 2.5.
\textsuperscript{49}See Guidelines, Section IV, para 2.9. Moreover, to enable subsequent control of the effective level of compliance, data controllers should document both the risk assessment and the measures adopted, see Guidelines, Section IV, para 2.10. See also Koivisto and Douglas, 2015.
\textsuperscript{50}These two kinds of solutions, which are also mentioned in Regulation (EU) 2016/679, represent a key component of the modern risk-based approach to data protection.
\textsuperscript{51}See Wright, 2011.
\textsuperscript{52}See Guidelines, Section IV, para 2.6. See also, with regard to the Regulation (EU) 2016/679, Article 29 Data protection Working Party, 2017.
\textsuperscript{53}See Guidelines, Section IV, para 2.7 (“With regard to the use of Big Data which may affect fundamental rights, the Parties should encourage the involvement of the different stakeholders (e.g. individuals or groups potentially affected by the use of Big Data) in this assessment process and in the design of data processing”). See also Wright, and De Hert, 2012, 467. On ethical values and stakeholder analysis, see Shelley-Egan et al., SATORI Deliverable D2.2, 2015; Shelley-Egan et al. Ethical Assessment of Research and Innovation, 2015, 28-29.
4. Conclusions

The Guidelines on the protection of individuals with regard to the processing of personal data in the big data context represent the first attempt to provide practical guidance to address the issues related to the use of big data and to reduce their potential negative impacts on society.

These Guidelines, as well as the Privacy, Ethical and Social Impact Assessment that they outline, confirm the importance of going beyond the mere declarations of fundamental rights and to provide practical instructions and operative methodologies to put them into practice.

In light of the above, these Guidelines confirm the attention of part of our society to the potential implications of the use of data, adopting a viewpoint that refuses vague notions (such as “citizen empowerment” or “digital sovereignty”) often used to provide a mere formal protection to personal data, but looks ahead to concrete and robust forms of assessment of the compliance of data use with the ethical and social values accepted in a given community.54

Bibliography


Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero

Rev Bio y Der. 2017; 41: 67-84


COUNCIL OF EUROPE. Progress report on the application of the principles of Convention 108 to the collection and processing of biometric data. 2005, para 10. Available at: https://rm.coe.int/16806840ba (accessed 4 May 2017).


Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero

Rev Bio y Der. 2017; 41: 67-84


- HILDEBRANDT, M. Profiling: "From Data to Knowledge. The challenges of a crucial technology", Datenschutz und Datensicherheit 30(9), 2006:548-552.


Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero


Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero
Rev Bio y Der. 2017; 41: 67-84


Towards a big data regulation based on social and ethical values. The guidelines of the council of Europe - Alessandro Mantelero

Rev Bio y Der. 2017; 41: 67-84


Fecha de recepción: 5 de junio de 2017
Fecha de aceptación: 30 de junio de 2017