POLITECNICO DI TORINO
Repository ISTITUZIONALE

The future of consumer data protection in the E.U. Rethinking the “notice and consent” paradigm in the new era of predictive analytics

Original
The future of consumer data protection in the E.U. Rethinking the “notice and consent” paradigm in the new era of predictive analytics / Mantelero, Alessandro. - In: COMPUTER LAW & SECURITY REPORT. - ISSN 0267-3649. - STAMPA. - 30:6(2014), pp. 643-660. [10.1016/j.clsr.2014.09.004]

Availability:
This version is available at: 11583/2556555 since:

Publisher:
Elsevier BV: PO Box 211, 1000 AE Amsterdam Netherlands: 011 31 20 4853757, 4853642 011 31 20, 011

Published
DOI:10.1016/j.clsr.2014.09.004

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

01 September 2025

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

Alessandro Mantelero
Aggregate Professor, Politecnico di Torino
Director of Privacy and Faculty Fellow, Nexa Center for Internet and Society
alessandro.mantelero@polito.it

THE FUTURE OF CONSUMER DATA PROTECTION IN THE E.U. RETHINKING THE "NOTICE AND CONSENT" PARADIGM IN THE NEW ERA OF PREDICTIVE ANALYTICS.

The Computer Law and Security Review, 2014 (forthcoming)

Abstract

The new E.U. proposal for a general data protection regulation has been introduced to give an answer to the challenges of the evolving digital environment. In some cases, these expectations could be disappointed, since the proposal is still based on the traditional main pillars of the last generation of data protection laws.

In the field of consumer data protection, these pillars are the purpose specification principle, the use limitation principle and the "notice and consent" model. Nevertheless, the complexity of data processing, the power of modern analytics and the "transformative" use of personal information drastically limit the awareness of consumers, their capability to evaluate the various consequences of their choices and to give a free and informed consent.

To respond to the above, it is necessary to clarify the rationale of the "notice and consent" paradigm, looking back to its origins and assessing its effectiveness in a world of predictive analytics. From this perspective, the paper considers the historical evolution of data protection and how the fundamental issues coming from the technological and socio-economic contexts have been addressed by regulations.

On the basis of this analysis, the author suggests a revision of the "notice and consent" model focused on the opt-in and proposes the adoption of a different approach when, such as in Big Data collection, the data subject cannot be totally aware of the tools of analysis and their potential output.

For this reason, the author sustains the provision of a subset of rules for Big Data analytics, which is based on a multiple impact assessment of data processing, on a deeper level of control by data protection authorities, and on the different opt-out model.

Keywords: data protection, consent, data protection impact assessment, big data, data protection authorities

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

1. Introduction

In the last few years, the debate surrounding data protection and privacy has focused on the future wave of new regulations. Driven by the Web 2.0 environment and the economy of data,[1] private companies and governments have become even more data-centric. However, the high demand for personal information, the complexity of the new tools of analysis and the increasing numbers of sources of data collection,[2] have generated an environment in which the "data barons" (i.e. big companies, government agencies, intermediaries)[3] have a control over digital information which is no longer counterbalanced by the user's self-determination.[4]

Nevertheless, all the ongoing proposals for a reform of data protection regulations, both in the U.S.[5] and Europe,[6] are still focused on the traditional main

[1] On the economic value of personal information, see Ian Brown, 'The economics of privacy, data protection and surveillance' in J.M. Bauer and M. Latzer (eds.) Research Handbook on the Economics of the Internet (Edward Elgar 2014) (forthcoming) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2358392 accessed 27 February 2014; Joseph W. Jerome, 'Buying and selling privacy: Big Data's Different Burdens and benefits' (2013) 66 Stan. L. Rev. Online 47, 47-49; OECD, 'Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value' (2013). OECD Digital Economy Papers, No. 220, OECD Publishing http://www.oecd-ilibrary.org/docserver/download/5k486qtxldmq.pdf?expires=1403110041&id=id&accname=guest&checksum=1F20BE8EB6E36BA2F7D94A175C5FB089 accessed 27 February 2014. ENISA, 'Study on monetising privacy. An economic model for pricing personal information' (2012) http://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/monetising-privacy accessed 27 February 2014. See also Howard J. Beales and Jeffrey A. Eisenach, 'An Empirical Analysis of the Value of Information Sharing in the Market for Online Content' (2014) Navigant Economics http://images.politico.com/global/2014/02/09/beales_eisenach_daa_study.html accessed 27 February 2014.

[2] See Luciano Floridi, *The 4TH Revolution. How the Infosphere is Reshaping Human Reality* (OUP 2014) 96 ("the fourth revolution has brought to light the intrinsically informational nature of human identity").

[3] See Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data. A Revolution That Will Transform How We Live, Work and Think* (John Murray 2013) 182; Federal Trade Commission, 'Data brokers. A Call for Transparency and Accountability' (2014) http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf accessed 25 June 2014. See also Alessandro Mantelero and Giuseppe Vaciago, 'Social media and big data' in Babak Akhgar, Francesca M. Bosco, and Andrew Staniforth (eds), *Cyber Crime and Cyber Terrorism Investigator's Handbook* (Elsevier Science 2014).

[4] Since the articles focuses on consumer data protection, for the purposes of the article, consumer, user and data subject are use considered as synonyms. On the right to informational self-determination, it is worth mentioning the influential decision adopted by the Federal German Constitutional Court (Bundesverfassungsgericht), 15 December 1983, Neue Juristische Wochenschrift, 1984, 419 https://www.zensus2011.de/SharedDocs/Downloads/DE/Gesetze/Volkszaehlungsurteil_1983.pdf?__blob=publicationFile&v=9 accessed 25 June 2014. See also Michael A. Froomkin, 'The Death of Privacy?' (2000) 52(5) Stan. L. Rev. 1461, 1464; Anita L. Allen, 'Coercing Privacy' (1999) 40 (3) Wm. & Mary L. Rev. 723; Paul M. Schwartz, 'Privacy and Democracy in Cyberspace' (1999) 52 Vand. L. Rev. 1690, 1661; Jerry Kang, 'Information Privacy in Cyberspace Transactions' (1998) 50(4) Stan. L. Rev. 1193, 1203; Alan F. Westin, *Privacy and Freedom* (Atheneum 1967) 330-399; Charles Fried, 'Privacy' (1968) 77 (3) Yale L. J. 475, 482-483.

[5] See The White House, 'A Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy' (2012) http://www.whitehouse.gov/sites/default/files/privacy-final.pdf accessed 25 June 2014, 47-48; The

pillars of the so called "fourth generation" of data protection laws.[7] In the field of consumer data protection, these pillars are the purpose specification principle, the use limitation principle and the "notice and consent" model (i.e. an informed, freely given and specific consent).[8]

White House, 'Executive Office of the President. Big Data: Seizing Opportunities, Preserving Values' (2014) http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf accessed 25 June 2014, 17-21, 55-57, 60-61. See also Daniel J. Solove, 'Introduction: Privacy Self-Management and the Consent Dilemma' (2013) 126 Harv. L. Rev. 1880; Julie E. Cohen, 'Between Truth and Power' in Mireille Hildebrandt and Bibi van den Berg (eds), *Freedom and Property of Information: The Philosophy of Law Meets the Philosophy of Technology* (Routledge) (forthcoming) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=234645910-11 accessed 25 June 2014.

[6] See Proposal for a regulation of the European Parliament and 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, Brussels, 25 January 2012 (hereinafter abbreviated as PGDPR) http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf accessed 27 February 2014; Proposal for a regulation of the European Parliament and of the Council on the protection of individual with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), (COM(2012)0011 – C7 0025/2012 – 2012/0011(COD)), compromise amendments on Articles 1-29 and on Articles 30-91 (hereinafter abbreviated as PGDPR-LIBE) http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/dv/comp_am_art_01-29/comp_am_art_01-29en.pdf and http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/dv/comp_am_art_30-91/comp_am_art_30-91en.pdf accessed 27 February 2014. In March 2014 the LIBE text was voted and approved by the whole Parliament http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P7-TA-2014-0212&language=EN accessed 25 March 2014

[7] See Viktor Mayer-Schönberger, 'Generational development of data protection in Europe' in Philip E. Agre, Marc Rotenberg (eds), *Technology and privacy: The new landscape* (MIT Press 1997) 219-241.

[8] See art. 2 (h), Directive 95/46/EC and art. 4 (8) PGDPR-LIBE (" 'the data subject's consent' means any freely given specific, informed and explicit indication of his or her wishes by which the data subject, either by a statement or by a clear affirmative action, signifies agreement to personal data relating to them being processed;"). Although the Directive does not recognize the consent as the principal or preeminent legal ground for data processing, it should be noted that the five other grounds require a "necessity" test, which strictly limits the cases in which they can be applied. See Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent', adopted on 13 July 2011, 7 http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf, accessed 27 February 2014; see also Article 29 Data Protection Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC', adopted on 9 April 2014, 11, 23-32 http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf accessed 27 February 2014. See also above at n 4; Brendan Van Alsenoy, Eleni Kosta and Jos Dumortier, 'Privacy notices versus informational self-determination: Minding the gap' (2014) 28 (2) Int. Rev. Law, Comp. & Tech. 185, 188; Viktor Mayer-Schönberger (n 7) 229-234; European Commission, Directorate-General Justice, Freedom and Security, 'Comparative study on different approaches to new privacy challenges, in particular in the light of technological developments: Working Paper No.2: Data protection laws in the EU. The difficulties in meeting challenges posed by global social and technical developments' (2010) http://ec.europa.eu/justice/policies/privacy/docs/studies/new_privacy_challenges/final_report_working_paper_2_en.pdf accessed 5 July 2014; Roder Brownsword, 'Consent in Data Protection Law: Privacy, Fair Processing and Confidentiality' in Serge Gutwirth, Yves Poullet, Paul De Hert, Cécile de Terwangne and Sjaak Nouwt, *Reinventing data protection?* (Springer 2009) 83-110. Differently, in the U.S., the traditional approach based on various sectorial regulations has

As it will be explained below in Section 4, this kind of approach seems to be inadequate in the present context[9], where the "transformative"[10] use of Big Data makes often impossible to explain the description of all the possible uses of information at the time of its initial collection.

Moreover, the digital world is characterized by an asymmetric distribution of the control over information, in terms of access to relevant, valued, and reliable data and in terms of ability to use it. In this sense, the control over the information derived from predictive analytics is not accessible to everyone, as it is based on the availability of large data sets, expensive technologies, and specific human skills to develop sophisticated systems of analysis and interpretation.[11]

Finally, in our digital economy, consumers often accept not having an effective negotiation of their personal information, due to market concentration[12] and related social and technological lock-ins.[13] The social lock-in effect is one of the consequences of the dominant position held by some big players and is evident in the social networks market. It is the incentive to remain on a network, given the numbers of connections and social relationships created and managed by the user

---

underestimated the role played by user's choice, adopting a market-oriented strategy. Nevertheless the recent guidelines adopted by the U.S. administrations seem to suggest a different approach, reinforcing self-determination. See The White House, 'A Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy' (n 5) 47-48. On the U.S. "notice and choice" regime, see also Neil M. Richards and Jonathan H. King, 'Big Data Ethics' (forthcoming 2014), Wake Forest Law Review, draft version, January 2014, 25 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2384174 accessed 25 June 2014; Paul Ohm, 'Branding Privacy' (2013) 97 Minn. L. Rev. 907, 929-930; Lorrie F. Cranor, 'Necessary but not sufficient: standardized mechanisms for privacy and choice' (2012) 10 J. on Telecom & High Tech L. 273.

[9] See Fred H. Cate, Mayer-Schönberger, Viktor, 'Data Use and Impact. Global Workshop' (2013) iii http://cacr.iu.edu/sites/cacr.iu.edu/files/Use_Workshop_Report.pdf accessed 27 February 2014 ("The technologies and data applications of the 21st century are rapidly combining to make data protection based on notice and choice irrelevant"); Ira S. Rubinstein, 'Big Data: The End of Privacy or a New Beginning?' (2013) Int'l Data Privacy L., 3 (2), 74; Marc Rotenberg, 'Fair Information Practices and the Architecture of Privacy (What Larry Doesn't Get)' (2001) Stan. Tech. L. Rev. 1, paras. 29-32 https://journals.law.stanford.edu/sites/default/files/stanford-technology-law-review/online/rotenberg-fair-info-practices.pdf accessed 20 December 2013.

[10] See Omer Tene and Jules Polonetsky, 'Privacy in the Age of Big Data: A Time for Big Decisions' (2012) Stan. L. Rev. Online 64. Big Data analytics make possible to collect a large amount of information from different sources and to analyse it in order to identify new trends and correlations in data sets. This analysis can be conducted to pursue purposes not defined in advance, related to the emerging correlations and different from the purposes of the initial collection.

[11] See Alessandro Mantelero, 'Social Control, Transparency, and Participation in the Big Data World' (2014) Journal of Internet Law, April, 23-29. On privacy and control over information see Westin (n 4) 7; Arthur R. Miller, *The Assault on Privacy Computers, Data Banks, Dossiers* (University of Michigan Press 1971) 25; Daniel J. Solove, *Understanding Privacy* (Harvard University Press 2008) 24-29; Cohen, 'Between Truth and Power' (n 5) 1, 5.

[12] See Science and Technology Options Assessment, 'Potential and Impacts of Cloud Computing Services and Social Network Websites' (2014) 94-99, 116-121 http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/513546/IPOL-JOIN_ET%282014%29513546_EN.pdf accessed 27 February 2014.

[13] See also Spiros Simitis, 'Reviewing privacy in an information society' (1987) 135(3) Pen. L. Rev., 707, 737 ("the value of a regulatory doctrine such as "informed consent" depends entirely on the social and economic context of the individual activity"); Schwartz, 'Privacy and Democracy in Cyberspace' (n 4) 1661-1662.

of a social networking platform. This lock-in intrinsically limits the user's possibility to recreate the same network elsewhere. The different technological lock-in is related to technological standards and data formats that are adopted by service providers. This lock-in effect limits data portability and migration from one service to another that offers the same functions.

For these reasons, it is necessary to re-consider the existing data protection legal framework with regard to the "notice and consent" model and define new models, which better address the various issues of the present and future digital environment.

Different proposals have been advanced by legal scholars and computer scientists, which focus on privacy by design,[14] contextual privacy,[15] differential privacy, [16] data uses[17] and other combined solutions. Nevertheless, many of these

[14] See Ann Cavoukian, 'Privacy by design. From rhetoric to reality' (2014) 12-18, 65-100 http://www.ipc.on.ca/images/Resources/PbDBook-From-Rhetoric-to-Reality.pdf accessed 27 February 2014; Ann Cavoukian, 'Privacy by Design: Leadership, Methods, and Results' in Serge Gutwirth, Ronald Leenes, Paul De Hert, Yves Poullet (eds), *European Data Protection: Coming of Age* (Springer 2013) 175-202; Ann Cavoukian, 'Privacy by Design: Origins, Meaning, and Prospects for Assuring Privacy and Trust in the Information Era' in Yee, G O M (ed), *Privacy Protection Measures and Technologies in Business Organizations: Aspects and Standards* (IGI Global 2012) 170-208; Ira S. Rubenstein, 'Regulating Privacy By Design' (2011) 26 Berkeley Tech. L. J. 1409-1456; Peter Schaar, 'Privacy by Design' (2010) 3(2) Identity in the Information Society 267-274. See also Article 29 Data Protection Working Party, 'Opinion 07/2013 on the Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems ('DPIA Template') prepared by Expert Group 2 of the Commission's Smart Grid Task Force' (2013) http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp209_en.pdf accessed 27 February 2014; Article 29 Data Protection Working Party, 'Opinion 9/2011 on the revised Industry Proposal for a Privacy and Data Protection Impact Assessment Framework for RFID Applications' (2011) http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp180_en.pdf accessed 27 February 2014; Woodrow Hartzog and Frederic Stutzman, 'Obscurity by Design' (2013) 88 Wash. L. Rev. 385, 397; Federal Trade Commission, 'Protecting Consumer Privacy in an Era of Rapid Change. Recommendations for Business and Policymakers' (2012) 22-24 http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf accessed 25 June 2014.

[15] See Helene Nissenbaum, *Privacy in Context. Technology, Policy, and the Integrity of Social Life* (Stanford University Press 2010) (it should be noted that the main aspects of the context-based approach suggested by the author are already existing in the European legal framework on data protection); Solove, *Understanding Privacy* (n 11) 69-70; Robert C. Post, 'The Social Foundations of Privacy: Community and Self in the Common Law Tort' (1989) 77 (5) Cal. L. Rev. 957, 980-981.

[16] See Cynthia Dwork, 'The Promise of Differential Privacy. A Tutorial on Algorithmic Techniques' in *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2011)* http://research.microsoft.com/apps/pubs/default.aspx?id=155617 accessed 27 February 2014; Cynthia Dwork, Guy N. Rothblumy and Salil Vadhanz, 'Boosting and Differential Privacy' in *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS 2010)* http://research.microsoft.com/pubs/155170/dworkrv10.pdf accessed 27 February 2014; Cynthia Dwork, 'Differential Privacy' in *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming, part II (ICALP 2006)* (Springer Verlag 2006), 1-12 http://research.microsoft.com/apps/pubs/default.aspx?id=64346 accessed 27 February 2014; Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil Vadhan, 'Computational Differential Privacy' in Thomas Beth, Norbert Cot, and Ingemar Ingemarsson (eds), *Advances in Cryptology. CRYPTO `09'* (Springer-Verlag 2009) 126-142 http://people.seas.harvard.edu/~salil/research/CompDiffPriv-crypto.pdf accessed 25 June 2014.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

proposals adopt a holistic approach to the problem. In contrast, this article suggests the adoption of different solutions for situations in which the role of the consent-based model is outdated (e.g. Big Data). These situations should be distinguished from the different contexts in which the traditional model based on self-determination can be preserved.

In doing so, the experience from the past should not be forgotten. In many cases, the first answer given by the legal system to new technological and social revolutions[18] is represented by the introduction of new *ad hoc* rules. Nevertheless, the lack of knowledge of past experiences makes it difficult to find adequate answers to the new questions that technology poses.

From this perspective, this article reconsiders the evolution of data protection and the role played by the data subject from mainframe to Big Data, in order to identify and clarify the rationale of the "notice and consent" paradigm and to give an answer to the contemporary problems of data protection.

In the light of the above, the first part of this article[19] deals with the rationale of the first generations of data protection regulations, in which there was no place for the notice and consent model, and then focuses on the changing of paradigm adopted by the following generations of regulations, in which the "notice and consent" model play an important role. The analysis points out the relationship between awareness, concentration of power over information and enhancement of data subjects' self-determination.

The second part of the article[20] focuses on the present Big Data era and considers the different aspects, briefly mentioned above, that caused the crisis of the traditional model.

The analysis of the past experiences and the existing similarities between the context of the 50's-60's and the present can offer elements to address the new challenges and to reconsider the data protection framework.

Nevertheless, a complete picture of the future legal framework of consumer data protection is difficult to produce for a number of factors: the complexity of the topic; the fact that Big Data applications are in their infancy; the impact that policy makers will have in defining regulations and their differing approaches.[21] Moreover, the use of Big Data analytics covers a wide range of different

---

[17] See Cate and Mayer-Schönberger, 'Data Use and Impact. Global Workshop' (n 9).

[18] See Mayer-Schönberger and Cukier, (n 3).

[19] See below at paras. 2 and 3.

[20] See below at paras. 4 and 5.

[21] See PGDPR; PGDPR-LIBE; The White House, 'A Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy' (n 5); The White House, 'Executive Office of the President. Big Data: Seizing Opportunities, Preserving Values' (n 5) 17-21, 55-57, 60-61; OECD, 'The OECD Privacy Framework' (2013) http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf accessed 27 February 2014; Asian-Pacific Economic Cooperation, 'Privacy Framework' (2005) https://cbprs.blob.core.windows.net/files/APEC%20Privacy%20Framework.pdf accessed 27 February 2014. See also the Cross Border Privacy Rules (CBPR) system adopted by the Asian-Pacific Economic Cooperation, available at http://www.cbprs.org/ accessed 27 February 2014.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

situations, characterized by conflicting interests, which require *ad hoc* guidelines and solutions.

For the above reasons, the definition of this new legal framework is outside the scope of an academic article. Furthermore this definition should involve policy makers and stakeholders in a wide debate. Nevertheless, it is possible and necessary to clarify the interests that should be taken into consideration and provide an initial outline of the limits and general principles for the re-definition of the legal framework on consumer data protection.

In light of the above, the following paragraphs provide further considerations for the debate rather than a definitive solution to the problem.

2. The reasons of data protection. The first generations of regulations

Before considering the different reasons that induce the law to protect personal information, it should be noted that European legal systems do not recognize the same broad notion of the right to privacy that exists in U.S. case laws. At the same time, data protection laws in the European countries do not draw their origins from the European idea of privacy and its related case law.

With regard to the notion of right to privacy (and in brief), in the U.S. the right to privacy covers a broad area that goes from informational privacy to the right of self-determination in private life decisions.[22] On the other hand, in European countries this right mainly focuses on the first aspect and is related to the activities of the media.

With regard to the origins of data protection in Europe, it is worth pointing out that the European data protection regulations, since their origins in the late 60's, have focused on the information regarding individuals, without distinguishing between their public or private nature.[23] The right to privacy and data protection do not concern the same aspects, even if they are entangled and connected in many senses. There is only a partial overlapping, given that private facts are also referred to individuals. At the same time, a lot of personal information is publicly available and, for this reason, it does not fall into the field of the right to privacy. However, the legal issues related to the protection of personal information had a more recent recognition in law, both in the U.S. and Europe.[24] This dates from the

[22] See Richard S. Murphy, 'Property Rights in Personal Information: An Economic Defense of Privacy' (1996) 84 Geo.L.J. 2381; William A Parent, 'A New Definition of Privacy for the Law' (1983) 2(3) Law & Phil. 305; Diane L. Zimmerman, 'Requiem for a Heavyweight: A Farewell to Warren and Brandeis's Privacy Tort' (1983) 68 Cornell L. Rev. 296, 299; Raymond Wacks, 'The Poverty of "Privacy"' (1980) 96 L.Q.R. 73, 77-78; Raymond Wacks, *The Protection of Privacy* (Sweet & Maxwell, 1980) 10; Luis Henkin, 'Privacy and Autonomy' (1974) 74(8) Colum. L. Rev. 1419.

[23] See Luiz Costa and Yves Poullet, 'Privacy and the regulation of 2012' in this Review (2012), vol. 28, issue 3, 255.

[24] See fn. 3. See also Paul M. Schwartz, 'The E.U.-US Privacy Collision: A Turn to Institutions and Procedures' (2013) 126 Harv. L. Rev. 1966, 1969-1992.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

60's, whereas the primitive era of the right to privacy was at the end of 19th century, when the penny press assumed a significant role in limiting the privacy of the people of upper classes.[25] For these reasons, our analysis should start from the computer revolution of the late 50's and not one century before, when the first decision on informational privacy were adopted in Europe,[26] independently from the U.S. legal doctrine and before the milestone article of Warren and Brandeis.[27]

The first generations of data protection regulations were characterized by a national approach. Regulations were adopted at different times and were different in the extension of the protection they provided and the remedies they offered.

The notion of data protection was originally based on the idea of control over information, as demonstrated by the literature of that period.[28] At that time, the migration from dusty paper archives to computer memories was a Copernican revolution which, for the first time in history, permitted the aggregation of information about every citizen previously spread over different archives.[29] For this reason, the first regulations represented the answers given by legislators to the rising concern of citizens about social control as the introduction of big mainframe computers gave governments[30] and large corporations the opportunity to collect and manage large amount of personal information.[31]

In that period, people were afraid of being visible like a goldfish in a glass bowl:[32] a concentration of information, which was massive for the time, was in the hands of few entities, which were able to support the investments required by the new mainframe equipment. This concentration was also induced by the centralized architecture of mainframes. They had a single central processing unit and a main memory in which all the computational power was placed and made available to other specialized terminals, which were connected to the central unit by cables.

---

[25] See Michael Schudson, *Discovering the News. A Social History of American Newspaper* (Basic Books 1978) 12-60. See also below at n. 26 and 27.

[26] See Trib. civ. Seine, 16 June 1858, in *D.P.*, 1858.3.62.

[27] See Samuel D. Warren and Luis D. Brandeis, 'The Right to Privacy' (1890) 4(5) Harv. L. Rev. 193-220.

[28] See Westin (n 4), 158-168, 298-326; Adam C. Breckenridge. *The Right to Privacy* (University of Nebraska Press 1970) 1-3. See also Solove, *Understanding Privacy* (n 11) 4-5. See also above at n. 4 and below at n. 32.

[29] See Secretary's Advisory Committee on Automated Personal Data Systems, 'Records, Computers and the Rights of Citizens' (1973) http://epic.org/privacy/hew1973report/ accessed 27 February 2014 ("A persistent source of public concern is that the Social Security number will be used to assemble dossiers on individuals from fragments of data in widely dispersed systems").

[30] Miller (n 11) 54-67; Mayer-Schönberger (n 7) 221-225.

[31] See Colin J. Bennett, 'Regulating Privacy: Data Protection and Public Policy in Europe and the United States' (Cornell University Press 1992) 29-33, 47; Mayer-Schönberger (n 7) 221-222.

[32] See Myron Brenton, *The Privacy Invaders* (Coward-McCann 1964); Vance Packard, *The Naked Society* (David McKay 1964); Miller (n 11), chs 1 and 2. See Secretary's Advisory Committee on Automated Personal Data Systems (n 29) ("In more than one opinion survey, worries and anxieties about computers and personal privacy show up in the replies of about one third of those interviewed. More specific concerns are usually voiced by an even larger proportion"). See also Mayer-Schönberger (n 7) 223. See also Lee A. Bygrave, *Data Protection Law. Approaching Its Rationale, Logic and Limits* (Kluwer Law International 2002), 107-112.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

The solution given by the legal systems was the opportunity to have a sort of counter-control over the collected data. [33] The purpose of the regulations was not to spread and democratize power over information, but to increase the level of transparency about data processing and guarantee the right to access to information. Citizens felt they were monitored and the law gave them the opportunity to know who controlled them, which kind of data were collected and for which purposes.

Technically speaking, the mandatory notification of new databases to independent authorities, registration and licensing procedures[34] were the fundamental elements of these new regulations. They were necessary in order to know who had control over information and to monitor data processing. Another key component of the first legal frameworks was the rights to access, which allowed citizens to ask the data owners about the way in which the information was used and, consequently, about their exercise of power over information. Finally, the entire picture was completed by the creation of *ad hoc* public authorities, to guarantee the respect and enforcement of citizen's rights, control over the data owners and reaction against abuses.

In this model there was no space for individual consent, due to the economic context of that period.

The collection of information was mainly made by public entities for purposes related to public interests, so it was mandatory and there was no space of autonomy in terms of negotiation about personal information. At the same time, personal information did not have an economic value for the private sector. The data about clients and suppliers were mainly used for operational functions regarding the execution of the activities of the company.

Nevertheless, there was also another element that contributed to exclude the role of self-determination: the lack of knowledge, the extreme difficulty for ordinary people to understand how the mainframes worked. The computer mainframes were a sort of modern god, with sacral attendants, a selected number of technicians that was able to use this new equipment. In this scenario, it did not make sense to give citizens the chance to choose, since they were unable to understand the way in which the date was processed.

Finally, it is worth pointing out that all these aspects (concentration of information, centralised architecture, complexity of data processing) are now present again in the Big Data context, hence the practical relevance of this past experience.[35]

[33] See Secretary's Advisory Committee on Automated Personal Data Systems (n 29). See also Viktor Mayer-Schönberger (n 7) 223.

[34] See Mayer-Schönberger (n 7) 223 and, on the extent of the licensing rules, Lee A. Bygrave, *Data Privacy Law. An International Perspective* (OUP 2014), 183-184.

[35] See below at paras. 4 and 5.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

3. The economic value of personal information: the new generations of regulations

The following period – from the mid 70's to the 90's – can be considered as the era of distributed computers, in which a lot of people bought a personal computer to collect and process information. The big mainframe computers "became" the small desktop personal computers, with a relatively low cost. Consequently, the computational capacity was no longer an exclusive privilege of governments and big companies, but became accessible to many other entities and individual consumers.

This period witnessed another transformation involving direct marketing, which was no longer based on the concept of mail order and moved towards computerized direct marketing solutions.[36] The new forms of marketing were based on customer profiling and required extensive data collection to apply data mining software. The main purpose of profiling was to suggest a suitable commercial proposal to any single consumer. This was an innovative application of data processing driven by new purposes. Information was no longer collected to support supply chains, logistics and orders, but to sell the best product to single users. As a result, the data subject became the focus of the process and personal information acquired an economic and business value, given its role in sales.

These changes in the technological and business frameworks created new requests from society to legislators since citizens wanted to have the chance to negotiate their personal data and gain something in return.

Although the new generations of the European data protection laws placed personal information in the context of fundamental rights[37], the main goal of these regulations was to pursue economic interests related to the free flow of personal data. This is also affirmed by the Directive 95/46/EC,[38] which represents both the general framework and the synthesis of this second wave of data protection

[36] Although direct marketing has its roots in mail order services, which were based on personalized letter (e.g. using the name and surname of addressees) and general group profiling (e.g. using census information to group addressees in social and economic classes), the use of computer equipment increased the level of manipulation of consumer information and generated detailed consumer's profiles. See Lisa A. Petrison, Robert C. Blattberg and Paul Wang, 'Database Marketing. Past, Present, and Future' (1997) 11 (4) J. Direct Marketing 109, 115-119 ("During the decade, companies not only learned their customer's names and addresses, they also began to collect detailed personal and purchasing information, thereby beginning to understand them as individuals rather than as part of a traditional mass audience"); Daniel J. Solove, 'Privacy and Power: Computer Databases and Metaphors for Information Privacy' (2001) 53(6) Stan. L. Rev. 1393, 1405-1407.

[37] See Council of Europe, Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, opened for signature on 28 January 1981 and entered into force on 1st October 1985 http://conventions.coe.int/Treaty/Commun/QueVoulezVous.asp?NT=108&CL=ENG accessed 27 February 2014; OECD, Annex to the Recommendation of the Council of 23rd September 1980: Guidelines on the Protection of Privacy and Transborder Flows of Personal Data http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm#preface accessed 27 February 2014.

[38] 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 [1995] OJ L281/31.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

laws[39]. Nevertheless the roots of data protection still remained in the context of personality rights and, for this reason, the European approach is less market-oriented [40] than in other legal systems. The Directive also recognizes the fundamental role of public authorities in protecting data subjects against unwilled or unfair exploitation of their personal information for marketing purposes.

Both the theoretical model of fundamental rights, based on self-determination, and the rising data-driven economy highlighted the importance of user consent in consumer data processing[41]. Consent does not only represent an expression of choice with regard to the use of personality rights by third parties, but is also an instrument to negotiate the economic value of personal information.[42]

In this new data-driven economy, personal data cannot be exploited for business purposes without any involvement of the data subject. It is necessary that data subjects become part of the negotiation, since data is no longer used mainly by government agencies for public purposes, but also by private companies with monetary revenues.[43]

Nevertheless, effective self-determination in data processing, both in terms of protection and economic exploitation of personality rights, cannot be obtained without adequate and prior notice[44].

For these reasons, the "notice and consent" model[45] has added a new layer to the existing paradigm based on transparency and access.

---

[39] The EU Directive 95/46/EC has a dual nature, since it was written on the basis of the existing national data protection laws, in order to harmonize them, but at the same time it also provided a new set of rules. See the recitals in the preamble to the Directive 95/46/EC. See also Yves Poullet, 'EU data protection policy. The Directive 95/46/EC: Ten years after' in this Review (2006), vol. 22, issue 3, 206, 207; Spiros Simitis, 'From the Market to the Polis: The EU Directive on the Protection of Personal Data' (1995) 80 Iowa L. Rev. 445.

[40] On the different approach based on granting individuals property rights in personal information, see Paul M. Schwartz, 'Property, Privacy and Personal Data' (2004) 117(7) Harv. L. Rev. 2055; Pamela Samuelson, 'Privacy as Intellectual Property?' (2000) 52(5) Stan. L. Rev. 1125; Lawrence Lessig, *Code and Other Laws of Cyberspace* (Basic Books 1999). For criticism, see Julie E. Cohen, 'Examined Lives: Informational Privacy and the Subject as an Object' (2000) 52 Stan. L. Rev. 1373.

[41] See Charter of Fundamental Rights of the European Union (2010/C 83/02), art. 8 [2010] C83/389. See also Productores de Música de España (Promusicae) v Telefónica de España SAU, C-275/06, paras. 63-64 http://curia.europa.eu/juris/liste.jsf?language=en&jur=C,T,F&num=C-275/06&td=ALL accessed 27 February 2014; Federal German Constitutional Court (Bundesverfassungsgericht), 15 December 1983 (n 4). Among the legal scholars, see also Schwartz, 'The E.U.-US Privacy Collision: A Turn to Institutions and Procedures' (n 24); Maria Tzanou, 'Data protection as a fundamental right next to privacy? 'Reconstructing' a not so new right' (2013) 3 (2) Int'l Data Privacy L. 88 ss.; Daniel J. Solove, 'Introduction: Privacy Self-Management and the Consent Dilemma' (n 5). See also above at n 4.

[42] But see Alessandro Acquisti and Jens Grossklags, 'Privacy and rationality in individual decision making' (2005) 3(1) Security & Privacy, IEEE, 26–33.

[43] See OECD, 'Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value' (n 1); European Data Protection Supervisor, 'Preliminary Opinion of the European Data Protection Supervisor. Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy' (2014) https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2014/14-03-26_competitition_law_big_data_EN.pdf accessed 27 February 2014.

[44] The notice describes how the data is processed and the detailed purposes of data processing.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

Finally, it is important to highlight that during the 80's and 90's data analysis increased in quality, but its level of complexity was still limited. Consequently, consumers were able to understand the general correlation between data collection and related purposes of data processing (e.g. profiling users, offering customized services or goods). Clearly, at that time, informed consent and self-determination were largely considered as synonyms. This changed in the Big Data era.

4. The future generation of regulations in the Big Data era

The present Big Data era is different from the previous period both in terms of economic and technological context, with direct consequences on the adequacy of the legal framework adopted to protect personal information.

The new environment is mainly digital and characterized by an increasing concentration of information in the hands of a few entities, both public and private. The role played by specific subjects in the generation of data flows is the main reason for this concentration. Governments and big private companies (e.g. large retailers, telecommunication companies, etc.) collect huge amounts of data while performing their daily activities. This bulk of information represents a strategic and economically relevant asset, since the management of large databases enables these entities to assume the role of gatekeepers with regard to the information that can be extracted from the datasets. They are able to keep information completely closed or to limit access to the data, perhaps to specific subjects only or to circumscribed parts of the entire collection.

Not only governments and big private companies acquire this power, but also the intermediaries in information flows (e.g. search engines,[46] Internet providers, data brokers, [47] marketing companies), which do not generate information, but play a key role in circulating it.

---

[45] See arts 2 (h), 7 (a) and 10, Directive 95/46/EC. See also Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 8) 5-6. With regard to personal information collected by public entities, the Directive 95/45/EC permits the data collection without the consent of data subject in various cases; however, the notice to data subjects is necessary in these cases. See arts 7, 8 and 10, Directive 95/46/EC. See also Christopher Kuner, 'The European Commission's Proposed Data Protection Regulation: A Copernican Revolution in European Data Protection Law' (2012) 11 Privacy & Sec. L. Rep., 1, 5.

[46] See also Betsy Sparrow, Jenny Liu, Daniel M. Wegner, 'Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips' (2011) Science 776-778. Published online 14 July 2011. doi: 10.1126/science.1207745. This study suggests that when people expect to have future access to information, they have lower rates of recall of the information itself and enhanced recall instead for where to access it. In a world in which the main part of collective information and knowledge are migrating to the online environment, this adaptation of the processes of human memory to the advent of new computing and communication technology increases the power of the gatekeeper of the information in the ICT context and mainly the role of search engines. See also Attila Marton, Michel Avital, and Tina Blegind Jensen, 'Reframing Open Big Data' in *ECIS 2013 Proceedings* (AISeL 2013).

[47] See Federal Trade Commission, 'Data brokers. A Call for Transparency and Accountability' (n 3);

There are also different cases in which information is accessible to the public, both in raw and processed form. This happens with regard to open data sets made available by government agencies, information held in public registries, data contained in reports, studies and other communications made by private companies and, finally, online user-generated contents, which represent a relevant and increasing portion of the information available online.

The concurrent effect of all these different sources only apparently diminishes the concentration of power over information, since access to information is not equivalent to knowledge.[48] A large amount of data creates knowledge if the holders have the adequate interpretation tools to select relevant information, to reorganize it, to place the data in a systematic context and if there are people with the skills to define the design of the research and give an interpretation to the results generated by Big Data analytics.[49]

Without these skills, data only produces confusion and less knowledge in the end, with information interpreted in an incomplete or biased way.

For these reasons, the availability of data is not sufficient in the Big Data context.[50] It is also necessary to have the adequate human[51] and computing resources to manage it. In this scenario, control over information does not only regard limited access data, but can also concern open data,[52] over which the information intermediaries create an added value by means of their instruments of analysis.

Given that only few entities are able to invest heavily in equipment and research, the dynamics described above enhance the concentration of power over information, which increases due to the new expansion of Big Data.

Committee on Commerce, Science, and Transportation, 'A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes' (2013) http://consumercal.org/wp-content/uploads/2013/12/senate_2013_data_broker_report.pdf accessed 20 February 2014; Nissenbaum (n 15) 79.

[48] See Michael Gurstein, 'Open data: Empowering the empowered of effective data use for everyone?' (2011) 16(2)First Monday http://firstmonday.org/ojs/index.php/fm/article/view/3316/2764 accessed 4 September 2013.

[49] See The Aspen Institute, *The Promise and Perils of Big Data* (2010) 13 http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf accessed 27 February 2014 ("As a large mass of raw information, Big Data is not self-explanatory"); danah boyd and Kate Crawford, 'Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly' (2012) 15(5) Inf., Comm. & Soc. 666-668. See also Julie E. Cohen, 'What Privacy is For' (2013) 126 Harv. L. Rev. 1904, 1924-1925; The White House, 'Executive Office of the President. Big Data: Seizing Opportunities, Preserving Values' (n 5) 7.

[50] See Mayer-Schonberger and Cukier (n 3); Cynthia Dwork and Deirdre K. Mulligan, 'It's not Privacy and It's not Fair' (2013) 66 Stan. L. Rev. Online 35. See also Jannis Kallinikos, 'The Allure of Big Data' (2012) ParisTech Rev., November 16 http://www.paristechreview.com/2012/11/16/allure-big-data/ accessed 27 February 2014.

[51] See Science and Technology Options Assessment (n 12) 95; Cohen, 'What Privacy is For' (n 49) 1922-1923.

[52] See Federal Trade Commission, 'Data brokers. A Call for Transparency and Accountability' (n 3), 13. See also S. Benjamin, R. Bhuvaneswari, and P. Rajan, 'Bhoomi: 'E–governance', or, an anti–politics machine necessary to globalize Bangalore?" (2007) http://casumm.files.wordpress.com/2008/09/bhoomi-e-governance.pdf accessed 27 February 2014.

Under many aspects, this new environment resembles the origins of data processing, when, in the mainframe era, technologies were held by a few entities and data processing was too complex to be understood by data subjects.

Could this suggest that, in the future, the scenario will change again in a sort of "distributed Big Data analytics", as it happened in the mid 70's?[53] I believe not.[54]

The new "data barons" do not base their position only on expensive hardware and software, which may become cheaper in the future. Neither is their position based on the growing number of staff with specific skills and knowledge, able to give an interpretation to the results of data analytics. The fundamental element of the power of "data barons" is represented by the large databases they have. These data silos, considered the goldmine of the 21st century, do not have free access, as they represent the main or the side-effect of the activities realized by their owners, due to the role they play in creating, collecting or managing information.

For this reason, with regard to Big Data, it seems quite difficult to imagine the same process of "democratization" that happened concerning computer equipment during the 80's. The access to the above-mentioned large databases is not only protected by legal rights, but it is also strictly related to the peculiar positions held by the data holders in their market and to the presence of entry barriers.

Another aspect that characterizes and distinguishes this new form of concentration of control over information is given by the nature of the purposes of data collection: data processing is no longer focused on single users (profiling), but it increased by scale and it is trying to investigate attitudes and behaviours of large groups[55] and communities, up to entire countries. The consequence of this large scale approach is the return of the fears about social surveillance, which characterized the mainframe era.

It is important to highlight that this new potentially extensive and pervasive social surveillance differs from the past, since the modern surveillance is no longer realized mainly by intelligence apparatus, which autonomously collects a huge amount of information through pervasive monitoring systems. It is the result of the interplay between private and public sectors,[56] based on a collaborative model

---

[53] See above at para. 3

[54] On the risks related to "democratized big data", see Woodrow Hartzog and Evan Selinger, 'Big Data in Small Hands' (2013) 66 Stan. L. Rev. Online 81, 84-85.

[55] On group privacy see Luciano Floridi, 'Open Data, Data Protection, and Group Privacy' (2014) 27(1) Philos. Technol. 1–3; danah boyd, 'Networked Privacy' (2012) 10(3/4) Surv. & Soc. 348-350 ("We need to develop models that position networks, groups, and communities at the center of our discussion"); Edward J. Bloustein, 'Group Privacy: The Right to Huddle' (1977) 8 Rut.-Cam. L. J. 219.

[56] See Colin J. Bennett, Kevin D. Haggerty, David Lyon, Valerie Steeves (eds.) *Transparent Lives Surveillance in Canada* (Athabasca University Press 2014) 55-69 http://www.aupress.ca/books/120237/ebook/99Z_Bennett_et_al_2014-Transparent_Lives.pdf accessed 27 February 2014; Neil M. Richards, 'The Dangers of surveillance' (2013) 126 Harv. L. Rev. 1934, 1940-41; Jon D. Michaels, 'All the President's Spies: Private-Public Intelligence Partnerships in the War on Terror' (2008) 96(4) California Law Review 901-966; Chris Hoofnagle, 'Big Brother's Little Helpers: How Choicepoint and Other Commercial Data Brokers Collect, Process, and Package Your Data for Law Enforcement' (2003) 29 N.C.J. Int'l L. & Com. Reg. 595, 595-597; Simitis, 'Reviewing privacy in an information society' (n 13) 707, 726. See also Alessandro Mantelero and Giuseppe Vaciago, 'The "Dark Side" of Big Data: Private and

made possible by mandatory disclosure orders, which are issued by courts or administrative bodies, and extended to an undefined pool of voluntary or proactive collaborations from big companies.[57]

In this way, governments obtain information with the indirect "co-operation" of the consumers who probably would not have given the same information to public entities if requested. Service providers, for example, collect personal data on the base of private agreements (privacy policies) with the consent of the user and for specific purposes[58], but governments exploit this practice by using mandatory orders to obtain the disclosure of this information.[59] This dual mechanism hides from citizens the risk and the dimension of the social control that can be realised by monitoring social networks or other services and using Big Data analytics technologies.[60]

Public Interaction in Social Surveillance, How data collections by private entities affect governmental social control and how the EU reform on data protection responds' (2013) Comp. L. Rev. Int'l 161-169.

[57] See also Council of Europe, 'Guidelines for the cooperation between law enforcement and internet service providers against cybercrime' (2008) http://www.coe.int/t/informationsociety/documents/Guidelines_cooplaw_ISP_en.pdf accessed 27 February 2014.

[58] On the existing relationship between data retention and access to personal information by government agencies or law enforcement authorities, see Joel Reidenberg, 'The Data Surveillance State in the US and Europe' (forthcoming) Wake Forest L. Rev. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2349269 accessed 10 December 2013.

[59] See Ira S. Rubinstein, Gregory T. Nojeim, and Ronald D. Lee, 'Systematic government access to personal data: a comparative analysis' (2014) 4 (2) Int'l Data Privacy L. 96-119; Christopher Kuner, Fred H. Cate, Christopher Millard, and Dan Jerker B. Svantesson, 'Systematic Government Access to Private-Sector Data Redux' (2014) 4 (1) Int'l Data Privacy L. 1-3; Fred H. Cate, James X. Dempsey, and Ira S. Rubinstein, 'Systematic government access to private-sector data' (2012) 2 (4) Int'l Data Privacy L. 195-199; Peter Swire, 'From real-time intercepts to stored records: why encryption drives the government to seek access to the cloud', (2012) 2 (4) Int'l Data Privacy L. 200-206; Ian Brown, 'Government access to private-sector data in the United Kingdom' (2012) 2 (4) Int'l Data Privacy L. 230-238; Stephanie K. Pell, 'Systematic government access to private-sector data in the United States' (2012) 2 (4) Int'l Data Privacy L. 245-254; see also the other contributions on the systematic government access to private-sector in different countries published in Int'l Data Privacy (2014) 4 (1) and (2012) 2 (4). See also Ian Brown, 'Lawful Interception Capability Requirements' (2013) Computers & Law http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2309413 accessed 27 February 2014; European Parliament, Directorate General for Internal Policies, Policy Department C: Citizens' Rights and Constitutional Affairs, Civil Liberties, Justice and Home Affairs, 'National Programmes for Mass Surveillance of Personal data in EU Member States and Their Compatibility with EU Law' (2013) http://www.europarl.europa.eu/committees/it/libe/studiesdownload.html?languageDocument=EN&file=98290 accessed 27 February 2014.

[60] See European Parliament (2013), 'Resolution of 4 July 2013 on the US National Security Agency surveillance programme, surveillance bodies in various Member States and their impact on EU citizens' privacy' http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2013-0322+0+DOC+XML+V0//EN accessed 27 February 2014; European Parliament, Directorate General for Internal Policies, Policy Department C: Citizens' Rights and Constitutional Affairs, Civil Liberties, Justice and Home Affairs, 'The US National Security Agency (NSA) surveillance programmes (PRISM) and Foreign Intelligence Surveillance Act (FISA) activities and their impact on EU citizens' (2013) 14-16 http://info.publicintelligence.net/EU-NSA-Surveillance.pdf accessed 14 December 2013; European Parliament, Directorate General for Internal Policies, Policy Department C: Citizens' Rights and Constitutional Affairs, Civil Liberties, Justice and Home Affairs, 'National

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

4.1 The crisis of the traditional data protection framework

In this scenario, the traditional data protection framework defined in the 90's[61] goes to crisis, since the new technological and economic contexts (i.e. market concentration, social and technological lock-ins) undermined its fundamental pillars: [62] the purpose specification principle, the use limitation principle,[63] and the "notice and consent" model.

The purpose specification and use limitation principles have their roots in the first generations of data protection regulations, since they are strictly related to the intention of avoiding extensive data collections, which may imply risks in terms of social surveillance and control.

With the advent of the new generation of data protection regulations – during the 80's and 90's –, these principles not only put a limit to data processing, but also became key elements of the "notice and consent" model. They define the use of personal data made by data controllers, which represents important information impacting users' choice. Nevertheless, the advent of Big Data analytics makes it difficult to provide detailed information about the purposes of data processing and the expected outputs.

Since Big Data analytics are designed to extract hidden or unpredictable inferences and correlations from datasets, the description of these purposes is becoming more and more "evanescent". This is a consequence of the

---

Programmes for Mass Surveillance of Personal data in EU Member States and Their Compatibility with EU Law' (n 59) 12-16. See also DARPA, 'Total Information Awareness Program (TIA). System Description Document (SDD), Version 1.1' (2002) http://epic.org/privacy/profiling/tia/tiasystemdescription.pdf accessed 14 December 2013; National Research Council, *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment*. (National Academies Press 2008) Appendix I and Appendix J; Congressional Research Service, 'CRS Report for Congress. Data Mining and Homeland Security: An Overview' (2008) www.fas.org/sgp/crs/homesec/RL31798.pdf accessed 14 December 2013.

[61] See the previous paragraph.

[62] See Fred H. Cate, 'The Failure of Fair Information Practice Principles', in Jane K. Winn (ed.) *Consumer Protection in the Age of the 'Information Economy'* (Ashgate 2006) 343–345, also available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1156972 accessed 27 February 2014. See also Fred H. Cate and Viktor Mayer-Schönberger, 'Notice and Consent in a World of Big Data. Microsoft Global Privacy Summit Summary Report and Outcomes' (2012) http://www.microsoft.com/en-au/download/details.aspx?id=35596 accessed 27 February 2014; Fred H. Cate and Viktor Mayer- Schönberger, 'Notice and consent in a world of Big Data' (2013) 3 (2) Int'l Data Privacy L., 67; Rubinstein (n 9) 2; Solove, 'Introduction: Privacy Self-Management and the Consent Dilemma' (n 5) 1880-1903; Kate Crawford and Jason Schultz, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms' (2014) 55 B.C.L. Rev. 93, 108 http://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4 accessed 27 June 2014.

[63] See also Paul M. Schwartz, 'Data Protection Law and the Ethical Use of Analytics' (2011) 19-21 http://www.huntonfiles.com/files/webupload/CIPL_Ethical_Undperinnings_of_Analytics_Paper.pdf accessed 27 February 2014; Mireille Hildebrandt, 'Slaves to Big Data. Or Are We?' (2013) http://works.bepress.com/mireille_hildebrandt/52 accessed 27 February 2014.

"transformative"[64] use of Big Data, which makes it often impossible to explain all the possible uses of data at the time of its initial collection.[65]

These critical aspects concerning the purpose specification limitation have a negative impact on the effectiveness of the "notice and consent" model.

First, the difficulty in defining the expected results of data processing induces introducing generic and vague statements in the notices about the purposes of data collection. Second, also in the hypothesis of the adoption of long and detailed notices, the complexity of data processing in the Big Data environment does not offer to users a real chance to understand it and to make their choice.[66]

Moreover, this scenario is made worse by the economic, social and technological constraints, which definitively undermine the idea of self-determination with regard to personal information that represented the core principle of the generation of data protection regulations approved during the 80's and 90's.[67]

As mentioned before, we assisted to an increasing concentration of the informational assets, due to the multinational or global nature of some big players of the new economy, but also due to merger and acquisition processes, which created big companies both in the online and offline markets. In various cases, mainly with regard to online services, these large scale trends drastically limit the number of the companies that provide specific kind of services, which consequently have hundreds of millions of users. This dimension of the dominant players also produces social and technological lock-in effects that increase data

---

[64] See Omer Tene and Jules Polonetsky, (n 11).

[65] See also Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation' (2013) 23-27, 45-47 http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf accessed 27 February 2014; Article 29 Data Protection Working Party, 'Opinion 06/2013 on open data and public sector information ('PSI') reuse' (2013) 19-20 http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp207_en.pdf accessed 27 February 2014; European Data Protection Supervisor, 'Opinion of the European Data Protection Supervisor on the 'Open-Data Package' of the European Commission including a Proposal for a Directive amending Directive 2003/98/EC on re-use of public sector information (PSI), a Communication on Open Data and Commission Decision 2011/833/EU on the reuse of Commission documents' (2012) 4-5, 7, 10 https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2012/12-04-18_Open_data_EN.pdf accessed 27 February 2014.

[66] See Laura Brandimarte, Alessandro Acquisti, and George Loewenstein, 'Misplaced Confidences: Privacy and the Control Paradox' (2010), Ninth Annual Workshop on the Economics of Information Security http://www.heinz.cmu.edu/~acquisti/papers/acquisti-SPPS.pdf accessed 27 February 2014; Joseph Turow, Chris Jay Hoofnagle, Deirdre K. Mulligan, and Nathaniel Good, 'The Federal Trade Commission and Consumer Privacy in the Coming Decade' (2007), ISJLP 3, 723-749 http://scholarship.law.berkeley.edu/facpubs/935 accessed 27 February 2014; Federal Trade Commission, 'Data brokers. A Call for Transparency and Accountability' (n 3), 42. On the limits of the traditional notices, see also Rayan M. Calo, 'Against Notice Skepticism in Privacy (and Elsewhere)' (2013) 87(3) Notre Dame L. Rev. 1027, 1050-1055 http://scholarship.law.nd.edu/ndlr/vol87/iss3/3 accessed 27 February 2014; Daniel J. Solove, 'Introduction: Privacy Self-management and The Consent Dilemma' (n 5) 1883-1888; World Economic Forum, 'Unlocking the Value of Personal Data: From Collection to Usage' (2013) 18 http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf accessed 27 February 2014.

[67] See the previous paragraph.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

concentration and represents further direct and indirect limitations to consumer's self-determination and choice.[68]

4.2 Reconsidering the role of user's self-determination

In the above-described scenario, characterized by complex data processing and concentration of control over information, the decision to maintain a model mainly focused on "notice and consent" represents a risk. It is easy for companies to give notice and require the consent without an effective self-determination of users, given the above-mentioned reasons.[69]

This leads us to reconsider the role of user's self-determination in the situations in which consumers are not able to understand deeply data processing and its purposes,[70] or are not in the position to decide[71]. There it seems to be an analogy between the characters of data processing in the Big Data era and what it happened in the mainframe age. Today, data is collected by a limited number of entities and consumers are not able to understand purposes and methods of data processing, like at the beginnings of computer age.

In these cases the focus cannot be maintained mainly on the user and his or her self-determination: the role played by users should be restricted and conversely the role of independent authorities should be increased.[72]

Data protection authorities, rather than consumers, have the technological knowledge to evaluate the risks associated to data processing and can adopt legal remedies to tackle them[73]. Furthermore, they are also in the best position to balance all the different interests of the various stakeholders with regard to extensive projects of data collection and data mining.[74]

The suggestion is not to change the entire traditional model of data protection, but to reshape it with regard to the Big Data context and the other contexts in which

---

[68] See above at para 1.

[69] See Daniel J. Solove, 'Privacy Self-Management and the Consent Dilemma' (n 5) 1899.

[70] See The Boston Consulting Group, 'The value of our digital identity' (2012) 4, http://www.libertyglobal.com/PDF/public-policy/The-Value-of-Our-Digital-Identity.pdf accessed 27 February 2014.

[71] See also art. 7 (4), PGDPR ("Consent shall not provide a legal basis for the processing, where there is a significant imbalance between the position of the data subject and the controller"). In 2013, The Committee on Civil Liberties, Justice and Home Affairs (LIBE) of the European Parliament has dropped the former art. 7 (4), see art. 7 PGDPR-LIBE.

[72] See Bygrave, *Data Protection Law. Approaching Its Rationale, Logic and Limits* (n 32) 86 ("the monitoring and enforcement regimes set up by data protection laws are also a mixture of paternalistic and participatory control forms").

[73] See arts. 22, 23, 24 Directive 95/46/EC; Art. 53 and ch. VIII PGDPR-LIBE. See also n. 74.

[74] See, e.g., Article 29 Data Protection Working Party, 'Letter from the Article 29 Data Protection Working Party addressed to Google regarding the upcoming change in their privacy policy' (2012) http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2012/20120202_letter_google_privacy_policy_en.pdf accessed 27 February 2014.

asymmetries in data negotiation drastically reduce users' self-determination.[75] In the remaining cases, the "notice and consent" model, as traditionally designed, can still be effective, although it needs to be reinforced by increasing transparency,[76] service provider's accountability[77] and data protection-oriented architectures.[78]

It may be argued that an *ad hoc* regime is not necessary, considering that Big Data analytics can be applied to anonymized datasets. Nevertheless, it is worth pointing out that anonymity by de-identification is a difficult goal to achieve,[79] as demonstrated in a number of studies.[80]

---

[75] See below at para 5. See also Ryan Calo, 'Digital Market Manipulation' (forthcoming 2014), 82 Geo. Wash. L. Rev. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2309703 accessed 4 July 2014.

[76] See art. 13a, PGDPR-LIBE. See also Lorrie F. Cranor, 'Necessary but not sufficient: standardized mechanisms for privacy and choice' (n 8) 286-295, 304-307.

[77] See 32a, 33, 33a, 34, 35, 39 PGDPR-LIBE. See also Article 29 Data Protection Working Party, 'Opinion 3/2010 on the principle of accountability' (2010) http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp173_en.pdf accessed 27 February 2014; World Economic Forum, 'Rethinking Personal Data: A New Lens for Strengthening Trust' (2014) 9 http://www3.weforum.org/docs/WEF_RethinkingPersonalData_ANewLens_Report_2014.pdf accessed 27 February 2014.

[78] See arts. 23, 32a, 33, PGDPR-LIBE. See also Calo, 'Against Notice Skepticism in Privacy (and Elsewhere)' (n 66). The article suggests innovative ways to deliver privacy notice, based on consumer's experience of a product or service to warn or inform rather than entirely on words or symbols ("visceral notice" in the words of the author, since these notices are drawn upon consumer psychology to achieve greater salience). But see Solove, 'Introduction: Privacy Self-management and The Consent Dilemma' (n 5) 1885.

[79] See MIT- CSAIL, 'Exploring the Future Role of Technology in Protecting Privacy. Workshop report' (2013) 5 http://bigdata.csail.mit.edu/sites/bigdata/files/u9/MITBigDataPrivacy_WKSHP_2013_finalvWEB.pdf accessed 15 June 2014. See also Paul M. Schwartz, 'Data Protection Law and the Ethical Use of Analytics' (n 63) 7 ("dividing line between personally identified information (PII) and non-PII can be difficult to trace").

[80] See Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA L. Rev. 1701-1777; United States General Accounting Office, 'Record Linkage and Privacy. Issues in creating New Federal Research and Statistical Information' (2011) 68-72 http://www.gao.gov/assets/210/201699.pdf accessed 14 December 2013. See also Hui Zang and Jean Bolot, 'Anonymization of location data does not work: a large-scale measurement study' in *Proc. MobiCom '11 Proceedings of the 17th annual international conference on Mobile computing and networking* (ACM 2011) 145-156; Philippe Golle, 'Revisiting the uniqueness of simple demographics in the US population' in *Proc. 5th ACM workshop on Privacy in electronic society*, (ACM 2006) 77-80; Latanya Sweeney, 'Simple Demographics Often Identify People Uniquely' (Carnegie Mellon University 2000); Latanya Sweeney, 'Foundations of Privacy Protection from a Computer Science Perspective' in *Proc. Joint Statistical Meeting, AAAS, Indianapolis*, (AAAS 2000). But cf. Omer Tene and Jules Polonetsky, 'Big Data for All: Privacy and User Control in the Age of Analytics' (2013) 11 Nw. J. Tech. & Intell. Prop. 239-274. Contra Ann Cavoukian and Khaled El Emam, 'Dispelling the Myths Surrounding De-identification: Anonymization Remains a Strong Tool for Protecting Privacy' in Ann Cavoukian, 'Privacy by design. From rhetoric to reality' (n 14) 227-245 ("As long as proper de-identification techniques, combined with reidentification risk measurement procedures, are used, de-identification remains a crucial tool in the protection of privacy."); Ann Cavoukian and Daniel Castro, 'Big Data and Innovation, Setting the Record Straight: Deidentification Does Work' (2014) http://www2.itif.org/2014-big-data-deidentification.pdf accessed 15 July 2014, contra Arvind Narayanan and Edward W. Felten, 'No silver bullet: De-identification still doesn't

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

The power of Big Data analytics to draw unpredictable inferences from information undermines many strategies based on de-identification[81]. In many cases a reverse process in order to identify individuals is possible and it is also possible to identify those using originally anonymous data.[82] Here, it is closer to the truth to affirm that each data is a piece of personal information than to assert that it is possible to manage data in a de-identified way.[83]

A potential solution may be represented by the introduction of specific provisions, by law or by contract,[84] which forbids any form of re-identification and provides *ad hoc* sanction.[85] Nevertheless, this approach solves only partially the problems related to Big Data analytics.

First, in many cases Big Data analytics are used for specific purposes that require non anonymous datasets (e.g. profiling, etc.). Second, the exclusion of any prior assessment of the risks related to Big Data projects makes difficult to monitor the effectiveness of the prohibition of re-identification. A case by case assessment, similar to the one suggested above, should be required anyhow, in order to provide an adequate enforcement of this prohibition and to assess the adequacy of the anonymising processes.

Finally, anonymity is usually related to single users and in this case it excludes the potential negative effect related to data processing, except the risk of re-identification. Nevertheless, anonymity does not affect group profiling and the related issues of non-discrimination or social control. In this sense, Big Data analytics make possible to identify patterns in the behaviours of groups[86] without identifying single individuals and these results can be used by data gatherers in order to define potentially invasive and discriminative solutions and policies.

---

work' (2014) http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf accessed 15 July 2014.

[81] See Mayer-Schönberger and Cukier (n 3) 154-156. See also Paul M. Schwartz and Daniel J. Solove, 'The PII Problem: Privacy and a New Concept of Personally Identifiable Information' (2011) 86 N.Y.U. L. Rev. 1841-1845.

[82] See above at n. 80.

[83] See World Economic Forum, 'Unlocking the Value of Personal Data: From Collection to Usage' (n 66) 12; Rubinstein (n 9) 74, 77-78.

[84] See Ann Cavoukian and Drummond Reed, 'Big Privacy: Bridging Big Data and the Personal Data Ecosystem through Privacy by Design' in Ann Cavoukian, 'Privacy by design. From rhetoric to reality' (n 14) 82; Yianni Lagos and Jules Polonetsky, 'Public vs. Nonpublic Data: The Benefits of Administrative Controls' (2013) 66 Stan. L. Rev. Online 103-109.

[85] See Cate and Mayer-Schönberger, 'Data Use and Impact. Global Workshop' (n 9) 13. See also Federal Trade Commission, 'Protecting Consumer Privacy in an Era of Rapid Change. Recommendations for Business and Policymakers' (n 14) 21.

[86] See Federal Trade Commission, 'Data brokers. A Call for Transparency and Accountability' (n 3), 19-20.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

5. Defining a subset of rules for Big Data analytics and the situations characterized by asymmetries in data negotiation

The context described above and the related observations suggest defining specific rules for Big Data uses and the situations characterized by asymmetries in data negotiation. [87]

The necessity to distinguish this area seems not to be felt neither by the E.U. legislator, in the proposal for a new data protection regulation, nor by the U.S. administration, in the Consumer Privacy Bill of Rights.[88]

Although the E.U. proposal provides various rules that can be useful, it still adopts a mainly holistic approach that does not considered autonomously the above-mentioned situations. This is an approach in which the consent is still "purpose-limited"[89], based on notice and on the opt-in model.

Conversely, legal scholars and companies propose a different approach, which focuses on the use of the data,[90] on the risks (of benefits and harms) associated with the proposed use and on accountability.[91]

This last solution has the undoubted merit to underline the crisis of the traditional model[92] and to suggest a solution more suitable to address the issues of the existing and future context of data processing.

Nevertheless, it offers a holistic solution and this "one solution fits all" approach does not seem to be consistent with the different existing contexts, since there are cases in which the traditional model can still be effective.[93] Moreover, applying this user-centred model in the Big Data context, it seems to address only partially the emerging issues: it adopts a risk analysis mainly focused on data protection assessment and security[94], rather than on a multi-criteria risk-analysis that considers also the ethical and social consequences of data processing.[95]

---

[87] For this reason, the following paragraphs will leave aside the cases in which users are able to understand the purposes of data collection and data processing, where the existing "notice and choice" and "purpose limited" consent can be kept valid.

[88] See The White House, 'A Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy' (n 5) 47-48. See also The White House, 'Executive Office of the President. Big Data: Seizing Opportunities, Preserving Values' (n 5) 61.

[89] See art. 7 (4) PGDPR-LIBE.

[90] See Cate and Mayer-Schönberger, 'Data Use and Impact. Global Workshop' (n 9) 6, 8-9; Edith Ramirez, 'The Privacy Challenges of Big Data: A View From the Lifeguard's Chair' Technology Policy Institute Aspen Forum Aspen, Colorado (keynote) (2013) 5-6 http://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard%E2%80%99s-chair/130819bigdataaspen.pdf accessed 27 February 2014.

[91] See Cate and Mayer-Schönberger , 'Data Use and Impact. Global Workshop' (n 9) 5, 17-18. On privacy harms, see also Rayan M. Calo, 'The Boundaries of Privacy Harm' (2011) 86 Ind. L. J. 1131.

[92] See Cate and Mayer-Schönberger , 'Data Use and Impact. Global Workshop' (n 9) 3-4, 7.

[93] See above at para. 4.

[94] See Cate and Mayer-Schönberger , 'Data Use and Impact. Global Workshop' (n 9) 12-13.

[95] On the contrary, this multi-criteria approach is adopted in the present article, see below in the text. See also Article 29 Data Protection Working Party, 'Statement on the role of a risk-based approach in data protection legal frameworks' (2014) http://ec.europa.eu/justice/data-

Finally, in this model the assessment is mainly conducted without a significant involvement of data protection authorities,[96] which conversely play a central role in the model here proposed.[97]

It is worth observing here that the new issues related to Big Data should not be necessarily addressed by making a choice between a "consent based" model and a "corporate accountability" model.

Although Big Data and lock-in effects drastically limit self-determination at the moment in which the data is collected, the fundamental right of any person to decide about his or her own information cannot be erased and users should have the right to be informed about data processing and not to take part of it.

In this sense, the model here suggested is the result of the past experiences: like in the first data protection regulations, the decision about data processing cannot be left to users, but at the same time user's rights to oppose to data processing and not to have personal data collected – codified in data protection laws during the 90's – should be preserved.

The fundamental pillars of this model are the definition of a rigorous multiple impact assessment of data processing, which should be widely adopted and publicly available,[98] and the adoption of the "opt-out" scheme.

It should be noted that the data protection impact assessment does not represent a completely new approach to data protection, as the privacy impact assessment exists in different experiences around the world and has represented an important tool since mid-1990s.[99] Nevertheless, it is necessary to distinguish the data

---

protection/article-29/documentation/opinion-recommendation/files/2014/wp218_en.pdf accessed 27 February 2014 ("The risk-based approach goes beyond a narrow "harm-based-approach" that concentrates only on damage and should take into consideration every potential as well as actual adverse effect, assessed on a very wide scale ranging from an impact on the person concerned by the processing in question to a general societal impact (e.g. loss of social trust)").

[96] See Cate and Mayer-Schönberger, 'Data Use and Impact. Global Workshop' (n 9), p. 17 ("after some initial transitional burdens, such an approach is likely to shift more of the burden for being accountable for data processing activities away from both regulators and individuals toward data users"). See the criticisms expressed by Information Commissioner's Office, 'Big data and data protection', Version: 1.0, p. 41 (2014) http://ico.org.uk/news/latest_news/2014/~/media/documents/library/Data_Protection/Practical_application/big-data-and-data-protection.pdf accessed 29 June 2014.

[97] See below in text.

[98] This is a critical aspect due to the need to balance the information about data processing provided to users and the security and competitive issues of enterprises. In order to balance these opposite issues, it is possible to provide business sensitive information in a separate annex to the impact assessment report, which will not be made public, or it is possible to publish a short version of the report without the sensitive contents. See Alessandro Mantelero, 'Competitive value of data protection: the impact of data protection regulation on online behaviour' (2013) 3(4) Int'l Data Privacy L. 234; Neil M. Richards and Jonathan H. King, 'Three Paradoxes of Big Data' (2013) 66 Stan. L. Rev. Online 41, 43; David Wright, 'A framework for the ethical impact assessment of information technology' (2011) 13(3) Ethics Inf Technol. 222. See also Recital 51 in the preamble to the PGDPR-LIBE.

[99] On the origins of the notion of privacy impact assessment, see Roger Clarke, 'Privacy impact assessment: Its origins and development' in this Review (2009), vol. 25, issue 2, 123-127; David Flaherty, 'Privacy impact assessments: an essential tool for data protection' (2000) 7(5) Privacy Law and Policy Reporter 85, fn. 3 http://www.austlii.edu.au/au/journals/PLPR/2000/45.html#fn3

protection impact assessment from the similar notion of privacy impact assessment. Both these assessments are based on the model of risk analysis and evaluate *ex ante* the future impact that a specific services or product could have on privacy or data protection, but the concepts of right to privacy and data protection are different,[100] consequently these assessments investigate different fields that are not completely overlapped.

A significant element of these assessments is the continuity of the evaluation, which follows the product and the service during their entire life-cycle, redefining the assessment when new features or modifications are introduced. This approach reduces the need for the legislator to follow technological developments and induces preventive solutions to ensure compliance with the principles of data protection.

With regard to the impact assessment, the same approach that is used in the field of product safety and liability (e.g. drugs authorization)[101] should be extended to data processing: in presence of complex data processing systems or data collections influenced by lock-in effects, the risk and benefit assessment should not be made by consumers, but it should be made by companies, under the supervision of data protection authorities. Consumers should only decide to exercise or not their right to opt-out. [102]

These situations, characterized by asymmetries in data negotiation, seem not to be considered in the last amended version of the EU proposal[103], as demonstrated by the erasure of the provision of the proposal of the Commission which stated that "consent shall not provide a legal basis for the processing, where there is a significant imbalance between the position of the data subject and the controller".[104]

On the contrary, the EU Proposal indirectly addresses the issues related to Big Data analytics and data protection risk analysis. In this sense, article 32a does not mention Big Data, but, in listing the cases in which data protection impact

---

accessed 27 February 2014. See also David Wright, 'Should privacy impact assessments be mandatory?' (2011) 54(8) Communications of the ACM 121-131; David H. Flaherty, *Protecting Privacy in Surveillance Societies: The Federal Republic of Germany, Sweden, France, Canada, and the United States* (University of North Carolina Press, 1989) 277ff, 405. For a comparative analysis of the different regulations on privacy impact assessment, see Roger Clarke, 'Privacy impact assessment: Its origins and development', 127-129; David Wright,Kush Wadhwa, Paul De Hert, and Dariusz Kloza, 'PIAF A Privacy Impact Assessment Framework for data protection and privacy rights' (2011) 19-184 www.piafproject.eu/ref/PIAF_D1_21_Sept_2011.pdf accessed 27 February 2014. More in general on privacy impact assessment, see David Wright, 'The state of the art in privacy impact assessment' in this Review (2012), vol. 28, issue 1, 54; David Wright and Paul De Hert (eds), *Privacy Impact Assessment* (Springer 2012); Roger Clarke, 'An evaluation of privacy impact assessment guidance documents' (2011) 1(2) Int'l Data Privacy Law 111-120.

[100] See above at para 2.

[101] See Cohen, 'What Privacy is For' (n 49) 1925 ("Big Data represents the de facto privatization of human subjects research, without the procedural and ethical safeguards that traditionally have been required").

[102] See also Article 29 Data Protection Working Party, 'Opinion 06/2014' (n 8) 45.

[103] See Art. 7 (4) PGDPR-LIBE. See also above at n. 71.

[104] See Art. 7 (4) PGDPR. If the provision had been maintained, the above-suggested approach based on prior assessment and opt-out would have represented a possible solution for the provision of a legal basis for processing.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

assessment is required, introduces a new criterion that is represented by the "large scale" of data collection.[105]

More specifically, the EU Proposal requires a prior assessment when large quantities of data is collected for a long period ("more than 5000 data subjects during any consecutive 12-month period") [106] or when there is a "large scale" data collection that involves special categories of information or of data subjects.

Although the Proposal adopts a solution based on risk assessment, this approach diverges from that adopted in this article, which is not technologically neutral and focuses on the use of Big Data analytics. Moreover, the model here suggested grounds on the complexity of data processing rather than on the dimension of datasets, the purposes of data processing and the nature of information or of data subjects.[107]

The solution proposed by the European legislator is in line with the traditional approach, which is based on the elements of data processing (nature of the data, purposes of data processing, categories of data subjects, period of data processing) rather than on the adopted technologies and on the difficulty for the user to understand the implications of data processing.

Nevertheless, given the predictive nature of Big Data and the impossibility to define *ex ante* the "specified" purposes of data processing,[108] it seems to be more adequate to require a mandatory data protection impact assessment in any cases in which these analytics are applied to data sets containing personal information. In this sense, the focus is on the tools used to manage the information and the dimension of data collection merely represents a necessary corollary. Finally, with regard to the dimension of databases, it should also be noted that the border between Big Data archives and normal databases is difficult to define.

The criterion of the threshold value of 5.000 data subjects, which has been adopted by the EU Proposal,[109] seems to be questionable when considering the nature of data or data subjects. The use of Big Data analytics with regard to information concerning relatively small groups of subjects with particular characteristics (e.g. geographical distribution, occupation, similar preferences or behaviours) may have relevant predictive effects and also consequences in terms of discrimination or diversity of approaches that can be adopted by data processors.[110]

---

[105] See Art. 32a (2) PGDPR-LIBE. This parameter is added to the traditional criteria of the nature of the information processed (e.g. sensitive) and the nature of data subjects (e.g. children, employees).

[106] See art. 32a (2) (a), PGDPR-LIBE.

[107] See arts. 32a (2) (a), (b), (c), (d), (e), (f), (g) and (h); 32a (3) point c; 33 (1), PGDPR-LIBE.

[108] See art. 6 (1) (b), Directive 95/46/EC and art. 5 (b) PGDPR-LIBE(data shall be "collected for specified, explicit and legitimate purposes").

[109] See art. 32a (2) (a) PGDPR-LIBE.

[110] See Danielle Keats Citron, Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions' (2014) 89(1) Wash. L. Rev. 1, 5, 14-15; Lior J. Strahilevitz, 'Toward a Positive Theory of Privacy Law' (2013) 126 Harv. L. Rev. 2010, 2021-2022, 2027-2028; Nissenbaum (n 15) 208-210; Richards and King, 'Three Paradoxes of Big Data' (n 98), 44; Mayer-Schönberger and Cukier (n 3) 144; Article 29 Data Protection Working Party, 'Opinion 03/2013'

For this reason, the *a priori* definition of a threshold value of data subjects is inadequate and the adoption of the notion of "large scale" (which is also provided by the EU Proposal) would appear to be more suitable.[111] Although it is a less definite notion, it induces a case by case analysis of the existing relationship between the information contained in a database and the population to which it relates.

In the suggested model, companies intend to adopt a strategy based on Big Data analytics should conduct a prior assessment of the different impacts on data protection, social surveillance and discrimination,[112] in order to adopt all the adequate measures and standards to reduce them.[113]

This multiple assessment, as in clinical trials, should be conducted by third parties and supervised by data protection authorities, which should also define the professional requirements of these third parties.[114] Once the assessment is approved by data protection authorities, the process should be considered secure in terms of protection of personal information and social consequences; for this reason, companies can enlist all their users in the specific data processing, without any prior consent, but giving them a previous notice that mentions the results of the assessment[115] and providing them the opt-out option.

It is worth pointing that a prior consultation process is also provided by the EU Proposal;[116] nevertheless it differs from the model here suggested, since it is restricted to specific cases of data processing, in which the data protection assessment is required.[117] These cases do not cover all the potential hypotheses in

---

(n 65); The White House, 'Executive Office of the President. Big Data: Seizing Opportunities, Preserving Values' (n 5) 7, 45-47, 51-53, 59-60, 64-65; Federal Trade Commission, 'Data brokers. A Call for Transparency and Accountability' (n 3) 55-56.

[111] See Art. 32a (2) PGDPR-LIBE.

[112] Discrimination should also be considered in terms of exclusion of individuals or underrepresented groups or categories. See Kate Crawford, 'The Hidden Biases in Big Data' (2013) http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/ accessed 27 February 2014; Jonas Lerman, 'Big Data and Its Exclusion' (2013) 66 Stan. L. Rev. Online 55, 56-63.

[113] Following the traditional risk-assessment model, the impact assessment should consider each of the different stages of the use of analytics. See also Paul M. Schwartz, 'Data Protection Law and the Ethical Use of Analytics' (n 63) 16-17 ("Each period of analytics raises different kinds of issues for privacy, and, as I will argue below, Fair Information Practices (FIPs) should be tailored to make an effective contribution to promoting privacy at all four stages."); Commission nationale de l'informatique et des libertés, 'Methodology for privacy risk management. How to implement the Data Protection Act' (2012) http://www.cnil.fr/fileadmin/documents/en/CNIL-ManagingPrivacyRisks-Methodology.pdf accessed 27 February 2014.

[114] See also David Wright, 'A framework for the ethical impact assessment of information technology' (n 98) 221.

[115] The notice should also describe how to access to the impact assessment report. This report is a short version of the documentation related to the assessment and it does not contain corporate sensitive information, in order to balance trade secrets and publicity of the assessment. Nevertheless, in presence of litigations, courts or data protection authorities may have access to the complete documentation and may disclose it to the plaintiff.

[116] See also art. 20, directive 95/46/EC.

[117] As above-mentioned, the data protection assessment is required by the EU Proposal only in the cases mentioned in art. 32a (2) (a), (b), (c), (d), (e), (f), (g) and (h), PGDPR-LIBE and in art. 32a (3) point c; 33 (1), PGDPR-LIBE.

This is a pre-print version of the article appearing in "The Computer Law and Security Review". This version is distributed for non-commercial purposes (http://www.elsevier.com/journals/computer-law-and-security-review/0267-3649/guide-for-authors).

which Big Data analytics may be used.[118] Moreover, in the EU Proposal, the prior consultation is mainly addressed to data protection officers and data protection authorities are involved in the process only "in case a data protection officer has not been appointed".[119]

The complexity of the assessments related to the use of Big Data analytics, which also involve social and ethical aspects, it is evident. For this reason, these assessments cannot be conducted only by experts in data protection law, but requires external auditors with specific and multi-disciplinary skills.[120]

For this reason, in the suggested model, both the third parties that realize the assessment and the data protection authorities, which approve the assessment, play a fundamental role in balancing all the different implications of data processing.

Since the balancing test does not focus only on data security, but is a multiple assessment that considers also the social impact and ethical use of data[121], data protection authorities are in the best position to evaluate all the different aspects.

If it is necessary and in the interest coming from society at large, they may also suggest the adequate solutions to make the data processing proposed by companies compliant with the above-mentioned issues.[122] In this sense, data

[118] Arts. 34 (2) (b), (4), PGDPR-LIBE also provides that the supervisory authorities may define the cases in which "it necessary to carry out a prior consultation on processing operations that are likely to present specific risks to the rights and freedoms of data subjects by virtue of their nature, their scope and/or their purposes, and specified according to paragraph 4". The following paragraph 4 states that the European Data Protection Board "shall establish and make public a list of the processing operations which are subject to prior consultation". In this sense, both the data protection authorities and the European Data protection Board may consider the use of Big data analytics as a class of operation that require the prior consultation. On the European Data protection Board, see ch. VII, s. 3 of the EU Proposal (PGDPR-LIBE).

[119] See Art. 34, PGDPR-LIBE. On the data protection officer, see arts. 35, 36 and 37, PGDPR-LIBE.

[120] It is worth pointing out that the social and ethical assessments are similar to the data protection impact assessment in their nature, since they are prior assessments based on risk analysis. Nevertheless, in these cases, the wide range of interests that should be considered requires the involvement of different stakeholders and experts.

[121] See Paul M. Schwartz, 'Data Protection Law and the Ethical Use of Analytics' (n 63) 22-26; David Wright, 'A framework for the ethical impact assessment of information technology' (n 98) 199–226, which pointed out that "a prescriptive ethical guidance is problematic because contextual factors influence the ethics" and, consequently, an ethical impact assessment would be more appropriate than prescriptive rules. In order to define the values that serve as an ethical guidance, the author considers the Lisbon Treaty, Charter of Fundamental Rights of the European Union. See also Floridi, *The 4TH Revolution. How the Infosphere is Reshaping Human Reality* (n 2) 189-190; Nissenbaum (n 15) 231; Rayan M. Calo, 'Consumer Subject Review Boards: A Thought Experiment' (2013) 66 Stan. L. Rev. Online 97, 101-102. September 3, 2013; Cynthia Dwork and Deirdre K. Mulligan, 'It's not Privacy and It's not Fair' (n 50) 38; Bjørn Hofmann, 'On value-judgments and ethics in health technology assessment' (2005) 3 Poiesis & Prax. 277, 288-292. See also Barbara Skorupinski and Konrad Ott, 'Technology assessment and ethics Determining a relationship in theory and practice' (2002) 1 Poiesis Prax. 95, 102-107; Neil M. Richards and Jonathan H. King, 'Three Paradoxes of Big Data' (n 98) 46. On the "social acceptability" of data processing, see also Bygrave, *Data Protection Law. Approaching Its Rationale, Logic and Limits* (n 32) 61-62, 339; Cohen, 'What Privacy is For' (n 49) 1925-26.

[122] See Paul Ohm, 'The Underwhelming Benefits of Big Data. In response to Paul M. Schwartz, Information Privacy in the Cloud' (2013) 161 U. Pa. L. Rev. Online 339, 345-346.

protection authorities can also involve in the assessment process the different stakeholders interested to the specific project of data processing and can audit them.[123]

A different assessment exclusively based on the adoption of security standards or corporate self-regulation would not have the same extent and independency. This does not mean that forms of standardization or co-regulation cannot be adopted;[124] nevertheless, the proposed reduction of the role of user's self-determination (opt-out) should have a necessary counterbalance in the active role of public and independent authorities acting in the interest of the whole society

The adequacy of this model is also empirically demonstrated by the most important cases in which data processing projects had social and ethical impacts. With regard to innovative products, services and business solutions, the initiative to evaluate their impact on data protection and on society did not come from data subjects, but from data protection authorities, which understood the potential risks related to these innovations. [125] Only these authorities were in the position to suggest the measures to be adopted by companies to reduce these risks, on the basis of balancing tests that placed data protection in the more general framework of the rights of the individual, as a single and as a member of a democratic society.

A significant element of these assessments is the continuity of the evaluations, which follow the product and the service during their entire life-cycle. For this reason, it is necessary to update each assessment (data protection assessment, social and ethical assessment) when new features or modifications are introduced. Moreover, changes with a significant impact on the existing data processing should require a specific authorization by data protection authorities.

Obviously the entire system works only if the political and financial autonomy of data protection authorities, both from governments and corporations, is guaranteed.[126]

For this reason, a model based on mandatory fees, paid by companies when they submit their requests of authorization to data protection authorities, would be

---

[123] See also David Wright, 'A framework for the ethical impact assessment of information technology' (n 98) 201-202, 215-220; Danielle Keats Citron, 'Technological Due Process' (2008) 85(6) Wash. U. L. Rev. 1249, 1312 http://digitalcommons.law.wustl.edu/lawreview/vol85/iss6/2 accessed 27 February 2014.

[124] See Calo, 'Consumer Subject Review Boards: A Thought Experiment' (n 121).

[125] See inter alia Article 29 Data Protection Working Party (2013). Letter to Mr. Larry Page, Chief Executive Officer http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2013/20130618_letter_to_google_glass_en.pdf accessed 27 February 2014; Irish Data Protection Commissioner (2012). Facebook Ireland Ltd. Report of Re-Audit http://dataprotection.ie/documents/press/Facebook_Ireland_Audit_Review_Report_21_Sept_2012.pdf accessed 27 February 2014; Italian Data Protection Authority (2013). Injunction and Order Issued Against Google Inc. http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/3133945 accessed 27 February 2014.

[126] See also FRA – European Union Agency for Fundamental Rights, 'Access to data protection remedies in EU Member States' (2013) 53 http://fra.europa.eu/sites/default/files/fra-2014-access-data-protection-remedies_en_0.pdf accessed 27 February 2014. See also Simitis, 'Reviewing privacy in an information society' (n 13), 707, 743.

preferable.[127] This solution will provide autonomous resources to authorities, proportionate to the increased activities of assessment, without being influenced by the companies under their surveillance.

The proposed model should offer clear and public procedures for assessment. These procedures undoubtedly represent an economic burden for companies. Nevertheless, in case of positive evaluation of data processing projects, these procedures allow companies to use data for complex and multiple purposes, without the inconvenience of acquiring a specific opt-in choice when data is used for new purposes. Companies should only inform users about any changes and give them the chance to opt-out.

From the user's point of view, on one hand the assessment conducted by the data protection authorities gives a guarantee of an effective evaluation of the risks related to data processing and, on the other hand, the opt-out allows users to receive information about data processing and to decide if they do not want to be part of the data collection.

With regard to the opt-out model, it might be noted that the suggested approach undermines the chances for users to negotiate their consent and to earn adequate revenues from data controllers. Nevertheless, the strength of this objection is reduced by the above-described existing limits to self-determination.

In the majority of the cases the negotiation is reduced to the alternative "take it or leave it". For these reasons, a prior assessment conducted by independent authorities and an opt-out model seem to offer more guarantees to users than an apparent, but inconsistent, self-determination based on "notice and consent"[128] and on the opt-in model.

Finally, in order to facilitate the implementation of the suggested approach, technical solutions based on user-oriented infomediaries could be adopted.[129]

---

[127] This self-financing model, based on licensing or notification fees, in the past was adopted in Sweden and United Kingdom, see Philip Schütz, 'Comparing formal independence of data protection authorities in selected EU Member States. Conference Paper for the 4th Biennial ECPR Standing Group for Regulatory Governance Conference 2012' (2012) 17, fn. 73, and 18 http://regulation.upf.edu/exeter-12-papers/Paper%20265%20-%20Schuetz%202012%20-%20Comparing%20formal%20independence%20of%20data%20protection%20authorities%20in%20selected%20EU%20Member%20States.pdf accessed 27 February 2014. In United Kingdom, the Information Commissioner's Office is still partially funded by notification fees paid by data controllers; see Information Commissioner's Office, 'Budget 2011-12. Spending plans 2012-13 to 2014-15' (2011) 2 http://ico.org.uk/about_us/boards_committees_and_minutes/~/media/documents/library/Corporate/Detailed_specialist_guides/ico_budget_2011-12.ashx accessed 27 February 2014. See also the fee-based model adopted by the European Medicines Agency http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000130.jsp&mid=WC0b01ac0580029336 accessed 27 February 2014.

[128] See Cate and Mayer-Schönberger, 'Data Use and Impact. Global Workshop' (n 9) 9 ("The existence of a privacy notice and the user's instant

And uninformed consent may give the illusion of privacy").

[129] On the original notion of infomediaries, see John Hagel III and Jeffrey F. Rayport, 'The Coming Battle for Customer Information' (1997) Harv. Bus. Rev. 53, Reprint 97104 http://thoughtleaderpedia.com/Marketing-Library/Quotes/The_Coming_Battle_for_Customer_Information.pdf accessed 27 February 2014

These agents are able to mediate between users and third parties companies interested in collecting consumers' personal information: on the basis of pre-definite consumer's preferences, infomediaries match single user's interest, and related propensity to share information, with the demand coming from companies gathering data.[130]

Infomediaries could be involved in negotiations, both in the opt-in and in the opt-out models, to communicate user's preferences. In the first case, they will act as representative of the data subject and will express his or her consent to data processing, notifying immediately to user the identity of data controller, as well as the nature and the purposes of data processing. In the second case (opt-out model), infomediaries will notify information about data processing to the user and his or her right to opt-out; if the data subject exercises this right, these agents communicate his or her decision to the data gatherer, making a machine-to-machine negotiation.[131]

6. Conclusion

The analysis of the evolutions of data protection, from mainframe era to the present new era of predictive analytics, points out that the new environment resembles the origins of data processing. For this reason it is difficult to maintain the holistic model of informed consent, adopted in the last decades.[132]

Like at the beginnings of computer age, data is collected by a limited number of entities and users are not able to understand the purposes and methods of data

---

("by connecting information supply with information demand and by helping both parties involved determine the value of that information, infomediaries would be building a new kind of information supply chain"); John Hangel III and Marc Singer, 'Unbundling the Corporation' (1999) Harv. Bus. Rev., 133-141; Paul M. Schwartz, 'Privacy and Democracy in Cyberspace' (n 4) 1685-1687. See also Jacques Bus and Carolyn M.-H. Nguyen, 'Personal Data Management – A Structured Discussion' in Mireille Hildebrandt, Kieron O'Hara and Michael Waidner (eds.), *Digital Enlightenment Yearbook 2013. The Value of Personal Data* (IOS Press, 2013) 270-287; Rubinstein (n 9) 81-87; Jerry Kang, Katie Shilton, Deborah Estrin, Jeff Burke, and Mark Hansen, 'Self-Surveillance Privacy' (2012) 97 Iowa L. Rev. 809; Doc Searls, 'The Intention Economy: When Customers Take Charge' (2012) Harvard Business Review Press, 177–179; World Economic Forum, 'Rethinking Personal Data: A New Lens for Strengthening Trust' (n 77) and World Economic Forum, 'Personal Data: The Emergence of a New Asset Class' (2011), both available at http://www.weforum.org/issues/rethinking-personal-data accessed 27 February 2014; Lee A. Bygrave, 'Electronic Agents and Privacy: A Cyberspace Odyssey 2001' (2001) 9 (3) Int'l J. L. Info. Technol., 275–294. See also the ProjectVRM created by the Berkman Center for Internet & Society, available at http://cyber.law.harvard.edu/research/projectvrm# accessed 27 February 2014; Mydex, 'The case for personal information empowerment: The rise of the personal data store' (2010) https://mydex.org/wp-content/uploads/2010/09/The-Case-for-Personal-Information-Empowerment-The-rise-of-the-personal-data-store-A-Mydex-White-paper-September-2010-Final-web.pdf accessed 27 February 2014; Hildebrandt (n 63).

[130] See World Economic Forum, 'Unlocking the Value of Personal Data: From Collection to Usage' (n 66) 13, 18.

[131] See also Carolyn M.-H.Nguyen, Peter Haynes, Sean Maguire, Jeffrey Friedberg, 'A User-Centred Approach to the Data Dilemma: Context, Architecture, and Policy' in Mireille Hildebrandt, Kieron O'Hara and Michael Waidner (eds.), *Digital Enlightenment Yearbook 2013* (n 129) 227-228, 233-238.

[132] See above at para 3.

processing. This leads us to reconsider the role of user's self-determination in data processing, as defined in during the 80's and 90's, when users are not able to understand deeply data processing and its purposes, or are not in the position to decide. In these cases, the role of users should be reduced and conversely the role of independent authorities should be increased.

In the Big Data era, data protection authorities, rather than users, have the technological knowledge to evaluate the risks associated to data processing[133] and can adopt adequate measures to reduce them. Furthermore, they are also in the best position to balance all the different interests of the various stakeholders with regard to extensive projects of data collection and data mining.

In the light of the above, the article suggests to adopt a subset of rules for Big Data and lock-in situations, based on the "opt-out" model and the definition of a rigorous multiple risk assessment, which should not only consider data processing, but also the social impact and ethical issues related to the use of personal information. These assessments should be conducted by third parties and supervised by data protection authorities, which can involve in the assessment process different stakeholders interested by the project of data processing.

Acknowledgements

I am indebted to all who provided feedback at the Oxford Internet Institute (OII) during the first presentation of my thoughts on this topic, during the seminar on "Big Data and the EU proposal on data protection: the crisis of the European paradigm?", which I held at the Institute in June 5, 2013. I am also grateful to the anonymous reviewers for their helpful comments. I am deeply grateful to Ian Brown, Luciano Floridi and Viktor Mayer-Schönberger for inspiring conversations at the OII, where I was visiting fellow in 2013 and 2014. I am also in debt to Sarah Spiekermann for the criticisms expressed during our lunch breaks in Oxford, which influenced the final part of this work.

I am grateful to the Network of Excellence in Internet Science project (http://www.internet-science.eu/network-excellence-internet-science) and to the Nexa Center for Internet & Society at Politecnico di Torino for having provided the financial support for visiting fellowships at the OII.

[133] Another aspect concerning risk assessment in the Big Data Era is represented by the transparency of the algorithms used by companies. About that, although algorithms are guarded by trade secrets, law should allow access to them by data Protection Authorities. See Danielle Keats Citron, Frank Pasquale (n 110) 1, 5, 10-11, 25. See also Mayer-Schönberger and Cukier (n 3) 179-182, which suggest a model based on independent internal and external audits. A wider access to the logic of the algorithms is required by Article 29 Data Protection Working Party, 'Opinion 03/2013' (n 65) 47 ("For the consent to be informed, and to ensure transparency, data subjects/consumers should be given access to their 'profiles', as well as to the logic of the decision-making (algorithm) that led to the development of the profile"). See also Tarleton Gillespie, 'The Relevance of Algorithms' in Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (eds.), *Media Technologies. Essays on Communication, Materiality, and Society* (MIT Press 2014) 167-194.