

SOCIAL CONTROL, TRANSPARENCY, AND PARTICIPATION IN THE BIG DATA WORLD

By Alessandro Mantelero

The Big Data paradigm underlines the fundamental role of control over information, in terms of competitive advantages in the business world and in terms of an increased possibility of social control by states and groups of power.

Control over the information deriving from Big Data is not accessible to everyone, as it is based on the availability of large datasets, expensive technologies, and specific human skills to develop sophisticated systems of analysis and interpretation. For these reasons, big business and governments are in the best position to take advantage of Big Data.

In order to increase access to information and to boost the participatory and transformative potential of information it is necessary to increase the number of people able to create and manage large amounts of data.

From a legal perspective, specific regulations can contribute to achieving this goal. A mandatory notification of the creation of large databases to independent authorities, rights to access to information and open data policies are different ways to increase the transparency of the information society and data sharing. They reduce the level of control over information and limit the advantage of Big Data owners in terms of technical and cultural analysis skills.

Alessandro Mantelero is Aggregate Professor Polytechnic University of Turin (Italy), Director of Privacy and Faculty Fellow, Nexa Center for Internet and Society at Polytechnic University of Turin and Research Consultant of Sino-Italian Research Center for Internet Torts at Nanjing University of Information Science & Technology (China); and was admitted to the Italian bar in 2001.

The evolution of human societies has been characterized by the increasing capacity to collect and manage information. From the beginning, the relationship among data, information,¹ and knowledge also had a strong link with the power held by public and private entities, due to two different reasons: (1) the knowledge drawn from organized information is fundamental for adopting successful strategies in public policies, social control and competition; (2) access to the information necessary to achieve knowledge is limited.

With regard to access to the information, we can distinguish between cultural barriers, such as low level of education, technical language, linguistic barriers, and restrictions to access, which include restricted access to libraries, copyright,² trade secrecy, documents, and databases owned by governmental agencies not available to the citizens or to any citizen. These limits create an asymmetric distribution of information and knowledge in society, which implies an unequal distribution of the opportunity to use them to understand, predict, and manage the different aspects of social interaction, from business to government policies, from technological innovation to science.³

The present era, characterized by digital information, huge databases, global transborder data flows and petabytes of information freely available online, seems to offer a different scenario. Nevertheless, it veils the paradox of information abundance: More information available means neither more knowledge available, nor a more diffuse access to information. Moreover, this period is characterized by an increasing concentration of the control over the information in the hands of a limited number of private and public entities, which, in different cases, cooperate in sharing information with each other to increase their position as owners and gatekeepers of knowledge.

From this perspective, this article provides an overview of these phenomena and their effects on society at large in order to suggest some thoughts on the possible policies that, from a legal perspective, may be able to boost a more democratic access to information and offer the opportunity to create new bottom-up initiatives and increase the number of centers of knowledge. With regard to these policies, attention focuses more on the general remedies to control and reduce the increasing concentration of power over information, rather than on the subsequent policies

that can facilitate the possible access and reuse of information, which are extensively described and analyzed in literature already.⁴ The article points out some requirements that are necessary to reduce the risk that these forms of concentration of information are hidden and hard to put under democratic control.

The analysis of these different topics does not focus on a specific legal system but assumes a global perspective. Given the global nature of the big players, governments, and multinational corporations, it is necessary to adopt policies that work at a supranational level.

AN OVERVIEW OF THE SOCIAL AND POLITICAL IMPLICATION OF BIG DATA

Due to the value that information has assumed in all fields of human activities, our society has been named the information society,⁵ nevertheless this synthesis of the role of information represents only a provisional stage in a continuous evolution: Information changes over time, with regard to its distribution, quantity, availability, and capacity to be stored and analyzed.

Although in the current information age individuals have new and various means to access information, the distribution of information is asymmetric in terms of access to relevant, valued, and reliable data and in terms of ability to use it, given the concentration of power over information in the hands of few.⁶

The role played by specific subjects in the generation of data flows is the principal reason for this concentration. Governments and big private companies collect huge amounts of data while performing their daily activities. This bulk of information represents a strategic and economically relevant asset, because 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, by limiting access to the data, perhaps to specific subjects only or to circumscribed parts of the entire collection, or by keeping it completely closed.

Not only governments and big private companies acquire this power, but also the intermediaries in the information flow (*e.g.*, search engines,⁷ Internet providers, credit report agencies, marketing companies), which do not generate information, but play a key role in circulating it.

There are also different cases in which information is accessible to the public 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 content, which represent a relevant and increasing portion of the information available online.

At a first glance, the concurrent effect of all these different sources diminishes the concentration of power over information; however, access to information is not equivalent to knowledge.⁸ 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.⁹ Without these skills, data only produces confusion and less knowledge in the end, with information interpreted in an incomplete or biased way.

In the Big Data context,¹⁰ the availability of data is not sufficient. It also is necessary to have the adequate human and computing resources to manage it. For this reason, control over information does not only regard limited access data, but can also concern open data,¹¹ over which the information intermediaries create an added value by means of their instruments of analysis.

Because only a few entities are able to invest heavily in equipment and research, the dynamics described above enhance the concentration of power over information and are increased by the new expansion of Big Data and its global dimension.

Predictive ability is the fundamental element that distinguishes this new kind of power from the ordinary power to acquire a large amount of data or profile millions of people.¹² The creation of datasets of enormous dimension (Big Data) and new powerful analytics make it possible to draw inferences about unknown facts from statistical occurrence and correlation, with results that are relevant in sociopolitical, strategic, and commercial terms.¹³ Despite the weakness of this approach, more focused on correlation than on statistical evidence,¹⁴ it is useful to predict and perceive the birth and evolution of macro-trends, that can be later analyzed in a more traditional statistical way in order to identify their causes. These

results concerning the emerging tendencies give the owner of Big Data a competitive advantage in the business world and increase the possibility of social control by states and groups of power.

Limited access to this information has consequences in terms of increasing its economic value. In a society based on information that requires large amounts of data in order to make decisions, having large datasets and the instruments to extract predictive knowledge from them is a fundamental asset, which can be exploited with a big payback.

Another element characterizing the concentration of power over information is that of being geolocalized in terms of the nationality of the owners of the datasets and of the place in which these Big Data farms are placed. Geo-location has several relevant consequences, not only with regard to the applicable law and the level of data protection, but also in terms of political influence exerted by local governments. In many cases national laws give judicial and administrative bodies power to access databases, which can compromise data protection, trade secrets, and public interests (when the data stored is related to governmental activities). There is another important aspect from an European Union perspective, related to the geographical concentration of power over information: Information technology (IT) is the sector that is most involved in Big Data management and the big players in this market are US companies, therefore there is an asymmetric distribution of control over data between the United States and the European Union.¹⁵ This situation implies evident risks in terms of service continuity, competition, lock-in, and, moreover, can represent a possible threat to EU interests due to the influence of the US legislative and political context.

For these reasons, the recent EU draft of new data protection regulation has adopted the strategy to protect personal data of European citizens wherever they are, even outside EU borders.¹⁶ Because the European Union does not have an IT sector strongly involved in Big Data management and it cannot exert political influence over foreign companies, European legislators try to use legal instruments in order to limit the risks of unauthorized access and fraudulent use of data. The intent of EU regulators is to influence development and business strategies of the owner data in order to compel third-party countries to reach a compromise on the guarantee offered with regard to

data protection.¹⁷ This is an ambitious goal, based on the consciousness of the importance of international data flows and of the effects of the introduction of legal barriers.

Finally, it should be noted that various commentators consider that the privacy risks related to Big Data analytics are low, pointing out the large amount of data processed by analytics and the de-identified nature of most of this data. This conclusion is wrong. Anonymity by de-identification is a difficult goal to achieve, as demonstrated in a number of studies.¹⁸ The power of Big Data analytics to draw unpredictable inference from information undermines many strategies based on de-identification.¹⁹ In many cases a reverse process in order to identify individuals is possible; it also is possible to identify individuals using originally anonymous data.²⁰ 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.

SOME POSSIBLE ACTIONS IN ORDER TO LIMIT THE RISKS RELATED TO BIG DATA

In order to limit the different risks related to Big Data, it seems necessary to adopt various remedies, without focusing the attention only on data protection regulations.²¹ The complexity of the phenomenon requires different approaches, in order to control and limit the information asymmetries and their consequences in terms of economic advantages and social control.

For these reasons, we need to identify different solutions in order to limit the power over information and obtain a better allocation of it. To achieve this goal it is important to adopt adequate measures to control those who have this power, in order to limit possible abuse and illegitimate advantages. At the same time, we need to increase access to the information, in order to increase the number of subjects able to create and manage large amounts of data, spreading the informational power currently in the hands of a few bodies.

The need to control these great aggregations of data is related to their political and strategic relevance and should lead to the introduction of a mandatory notification of the creation of a big and

important database—as happened at the beginning of the computer age when there was a similar concentration of power in the hands of a few due to the high cost of the first mainframes²²—and the creation of specific international independent authorities. These authorities will be able to control the invasive attitude of government with regard to large databases and the power of the owner of Big Data, but also have an important role in the definition of specific standards for data security.

This approach seems to be inline with the value assumed by Big Data, as happens with all the relevant and strategic assets. This will be a long and tortuous journey, as it is based on international cooperation. Therefore, it is important to start as soon as possible, using the existing international bodies (Council of Europe, OECD, and APEC) and multilateral dialogues among the European Union, the United States, and other countries, rather than introducing new regulations when it is late. At the same time, any solutions should be implemented in an appropriate manner, avoiding the involvement of every kind of data-farm built somewhere in the world, but considering only the data-farms with an absolutely remarkable dimension or considerable importance because of the data collected (*e.g.*, police or military databases).

Another important aspect of these policies concerns the transparency of the information society, in terms of knowing who holds great informational power due to Big Data management. Knowing this makes it possible to act on the other central aspects, namely access to data and data sharing, in order to limit the power of the owners of Big Data and give society the opportunity to have access to information, that fundamental resource. From this perspective, a key role is played by open data.²³

Opening public databases and potentially private archives²⁴ to citizens and giving them raw data not only reduces the power of the owners of information, in terms of the exclusive access to the data, but also limits their advantage in terms of technical and cultural analysis skills. The access to data does not mean that everyone will have new knowledge and predictive capacity immediately, because, as mentioned above, technical equipment is necessary. However, the availability of the data permits citizens to put together their economic and cultural resources,²⁵ even without a business-oriented action, in order to

constitute groups dedicated to the analysis and processing of the raw data. In this way, alternative and autonomous centers focused on Big Data management could be created. These centers will take a different and longer amount of time to achieve the same results that the owners of Big Data are able to obtain in a short time, thanks to their resources. However, for some research with relevance to the public the speed of obtaining the result is less important than the result itself (*e.g.*, Who's Lobbying,²⁶ Openpolis,²⁷ OpenSpending,²⁸ and Open Street Maps²⁹).³⁰

It also is possible to adopt open data policies that are more oriented to tackle the risks of the “data divide”³¹ and focus not only on the access to data, but also on the provision of adequate information and technical tools to the citizen. These contribute to an effective use of the data that reduces asymmetries in the distribution of the knowledge drawn from this public accessible information.

The final critical issue regarding Big Data concerns the geopolitical distribution of informational power, which represents an emerging problem for Europe. From this perspective the comparison with the United States assumes relevant importance, because in the US public sector, mainly with regard to federal agencies, not only is there a structural homogeneity still unknown to the European Union, fragmented into different national systems, but also many resources have been invested for modernizing the management of information through the use of cloud computing technologies, creating a limited number of big databases. Considering the US private sector, even if big European companies are able to collect and analyze a large amount of data, the excellence of the US companies in some key areas of the information and communication technology sector (search engines, cloud computing services, user-generated-content platforms, social networks) puts this nation in a better position to control the world's informational flows, because the largest stream of data is present in the areas most closely related to information and communications technology (ICT).

From a geo-political and industrial perspective, this situation may be a weakness for the European Union, in terms of the loss of control over the data of its citizens due to the need to entrust the management of strategic information to foreign entities. In order to reduce this risk, the European Union has strengthened the protection of personal data by

applying the European data protection law when goods or services are offered to “data subjects in the Union,”³² or when their behavior is monitored. At the same time, European industry is being urged by the European Commission to assume a more important role in the implementation of new computer architectures, such as cloud computing.³³

Stimulating competition in the development of new technologies in the ICT sector and strengthening the regulatory framework for the protection of information are the only possible responses in order to limit the potentially negative effects due to the use of services provided by foreign companies and the risk to give the control of the (big) data concerning European citizens to companies that potentially may be influenced by legal and political foreign interests.

CONCLUSION

Governments and big companies are increasing their control over information. This concentration is reducing the transparent and democratic use of information in our societies, facilitating social control, and producing an asymmetric distribution of knowledge in society.

These negative effects of the above can be tackled by the adoption of a multilayered strategy, which includes data protection and privacy regulations, open access to information and the development of a European competitive industrial policy in the ICT sector.

NOTES

- On the differences and relationship between data and information, see Boland, R. J. and Hirschheim, R.A., *Critical issues in information systems research*, Chichester, England: John Wiley & Sons (1987); McKinney, E. H. and Yoos, C. J., “Information about information: A taxonomy of views,” *MIS Quarterly* 34 (2): 329–344 (2010).
- See Lessig, L., *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*, New York: The Penguin Press (2004), <http://www.free-culture.cc/freecontent/>.
- On open access to scholarly publications and scientific data, see Nicol, A., Caruso, J. and Archambault, E., *Open Data Access Policies and Strategies in the European Research Area and Beyond*, Montréal: Science-Metrix (2013), http://www.science-metrix.com/pdf/SM_EC_OA_Data.pdf.
- A bibliography on access and reuse of (public sector) information is available at <http://www.evpsi.org/biblio> (accessed Sept. 15, 2013).
- See Bell, D., *The Coming of Post-Industrial Society: A Venture in Social Forecasting*, New York, NY: Basic Books (1973); Masuda, Y., *The Information Society as Post-Industrial Society*, Tokyo: World Future Society (1980); Toffler, A., *The Third Wave*. New York: Bantam Books (1980); Beniger, R. J., *The Control Revolution, Technological and Economic Origins of the Information Society*, Cambridge (MA)-London: Harvard University Press (1986); Webster, F., *The Information Society Reader*, London, New York: Routledge (ed. 2004); Webster, F., *Theories of the information society*, London: Routledge (2006); Castells, M., *The Information Age: Economy, Society, and Culture*, Vol. 2 of *The Rise of the Network Society* of M. Castells, Oxford: Wiley-Blackwell (2009).
- See The Aspen Institute, Bollier, D. Rapporteur, *The Promise and Peril of Big Data*, Washington, D.C. (2010), http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf. See also OECD, Joint WPISP-WPIE Roundtable, *The Economics of Personal Data and Privacy: 30 Years after the OECD Privacy Guidelines* (2010), accessed August 28, 2013, <http://www.oecd.org>; McKinsey Global Institute, “Big data: The next frontier for innovation, competition, and productivity,” (2011), <http://www.mckinsey.com>; Boyd, D. and Crawford, K., “Six Provocations for Big Data,” paper presented at Oxford Internet Institute’s “A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society,” Oxford, September 21, 2011, <http://ssrn.com/abstract=1926431>; Boyd, D. and Crawford, K., “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon,” *Information, Communication, & Society* 15 (5): 662–679 (2012), doi: 10.1080/1369118X.2012.678878. See also Tene, O. and Polonetsky, J., “Privacy in the Age of Big Data. A Time for Big Decisions,” *Stan. L. Rev. Online* 64: 63–69 (2012), http://www.stanfordlaureview.org/sites/default/files/online/topics/64-SLRO-63_1.pdf.
- See also Sparrow, B., Liu, J. and Wegner, D.M., “Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips,” *Science*, August 5, 2011:776–778, Published online July 14, 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 Marton, A., Avital, M. and Blegind Jensen, T., “Reframing Open Big Data,” in *ECIS 2013 Proceedings*, Atlanta, GA (2013).
- See Gurstein, M., “Open data: Empowering the empowered of effective data use for everyone?” *First Monday*, 16 (2) (2011), <http://firstmonday.org/ojs/index.php/fm/article/view/3316/2764> (accessed Sept. 4, 2013).
- See The Aspen Institute, 13 (“As a large mass of raw information, Big Data is not self-explanatory”); Boyd, and Crawford, *supra* n.6 at 666–668.
- See Mayer-Schonberger, V. and Cukier, K. 2013. *Big Data. A Revolution That Will Transform How We Live, Work and Think*. London: John Murray. See also Kallinikos, J. 2012. “The Allure of Big Data.” *ParisTech Rev.*, November 16. <http://www.paristechreview.com>.
- See also Benjamin, S., Bhuvaneshwari, R. and Rajan, P. 2007. “Bhoomi: ‘E-governance’, or, an anti-politics machine necessary to globalize Bangalore?” CASUM–m Working Paper. <http://casumm.files.wordpress.com/2008/09/bhoomi-e-governance.pdf>.
- The Aspen Institute, *supra* n.6 at 12.
- See Boyd and Crawford, *supra* n.6.
- See The Aspen Institute, *supra* n. 6. See also Marton, Avital, and Blegind Jensen, *supra* n.7.
- This asymmetry has become moreover evident in the N.S.A. Internet surveillance case; see European Parliament resolution of July 4, 2013, on the US National Security Agency surveillance program, surveillance bodies in various Member States and their impact on EU citizens’ privacy, available at <http://www.europarl>.

- europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2013-0322+0+DOC+XML+V0//EN*. See the various articles published by *The Guardian*, available at <http://www.guardian.co.uk>; Auerbach, D., Mayer, J. and Eckersley, P., "What We Need to Know About PRISM, June 12, 2013." <https://www.eff.org/deeplinks/2013/06/what-we-need-to-know-about-prism>. See also the various documents available at <https://www.cdt.org>; Perloth, N., Larson, J. and Shane, S., "N.S.A. Able to Foil Basic Safeguards of Privacy on Web," *The New York Times*, Sept. 5, 2013, http://mobile.nytimes.com/2013/09/06/us/nsa-foils-much-internet-encryption.html?nl=today'sheadlines&emc=ed_th_20130906. This case is just the latest in a series of programs adopted by intelligence services in different nations to pursue a massive social control; nevertheless, it has caused a wide debate about privacy and data protection. Many opinions have been voiced on the comparison with the United States. See European Parliament 2001, "Report on the existence of a global system for the interception of private and commercial communications (EHELON interception system)," <http://www.fas.org>; DARPA 2002, "Total Information Awareness Program (TIA), System Description Document (SDD), Version 1.1," <http://epic.org/privacy/profiling/tia/systemdescription.pdf>. See also National Research Council, Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment, Appendix 1 e Appendix J, Washington, D.C.: The National Academies Press; Congressional Research Service 2008, CRS Report for Congress, "Data Mining and Homeland Security: An Overview," www.fas.org/sgp/crs/homesec/RL31798.pdf.
16. See Article 3(2) of the Proposal for a Regulation 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, http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf. See also Article 3 (2), 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, available at http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libel/dv/comp_am_art_01-29/comp_am_art_01-29en.pdf.
 17. See Greenleaf, G., "Sheherazade and the 101 data privacy laws: Origins, significance and global trajectories," Submitted to the *Journal of Law & Information Science*, 2013, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2280877.
 18. See Ohm, P., "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization," *UCLA L. Rev.* 57: 1701-1777 (2010); United States General Accounting Office 2001, Record Linkage and Privacy, Issues in creating New Federal Research and Statistical Information 68-72, <http://www.gao.gov/assets/210/201699.pdf>. See also Zang, H. and Bolot, J., "Anonymization of location data does not work: A large-scale measurement study," in *MobiCom '11 Proceedings of the 17th annual international conference on Mobile computing and networking*, New York: ACM (2011); Golle, P., "Revisiting the uniqueness of simple demographics in the US population," in *Proceedings of the 5th ACM workshop on Privacy in electronic society*, New York: ACM (2006); Sweeney, L., "Simple Demographics Often Identify People Uniquely," *Data Privacy Working Paper 3* (2000), Pittsburgh: Carnegie Mellon University; Sweeney, L., "Foundations of Privacy Protection from a Computer Science Perspective," in *Proceedings, Joint Statistical Meeting, AAAS, Indianapolis* (2000). But see Tene, O. and Polonetsky, J., "Big Data for All: Privacy and User Control in the Age of Analytics," *Northwestern Journal of Technology and Intellectual Property*, 11 (5): 239-273 (2013).
 19. See Mayer-Schönberger, V. and Cukier, K., *Big Data: A Revolution That Will Transform How We Live, Work and Think* (London: John Murray Publishers), 154-156 (2013). See also Schwartz, P. M. and Solove, D. J., "The PII Problem: Privacy and a New Concept of Personally Identifiable Information," 86 *New York Univ. Law Review*, 1841-1845 (2011).
 20. See *supra* n.18.
 21. It also should be noted that the rules concerning personal data protection only impact within a limited area of the entire amount of information managed by the owners of Big Data, because in many cases they hold databases that do not refer to identified or identifiable subjects, even if they have a predictive relevance. However, when open data sets contain personal data, it is necessary to make a prior data protection assessment to analyse the effects of open data strategies on the access to personal information. See Article 29 Data Protection Working Party, "Opinion 06/2013 on open data and public sector information (PSI) reuse," adopted on June 5, 2013, http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp207_en.pdf. See also Article 29 Data Protection Working Party, "Opinion 03/2013 on purpose limitation," adopted on April 2, 2013, Annex 2, 45-50, http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf; European Data Protection Supervisor, "Opinion 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," Brussels, April 18, 2012, https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2012/12-04-18_Open_data_EN.pdf; Australian Government-Departure of Finance and Deregulation-Australian Government Information Management Office, 2013, "The Australian Public Service Big Data Strategy. Improved understanding through enhanced data-analytics capability," 21, <http://agict.gov.au/sites/default/files/Big%20Data%20Strategy.pdf>.
 22. See Article 8 (a) of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, opened for signature in Strasbourg on January 28, 1981, recital 48 in the preamble to Directive 1995/46 and Articles 18-21 of Directive. See also Article 29 Data Protection Working Party, "Article 29 Working Party report on the obligation to notify the national supervisory authorities, the best use of exceptions and simplification and the role of the data protection officers in the European Union," adopted on January 18, 2005, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2005/wp106_en.pdf; Article 29 Data Protection Working Party, "Working Document: Notification," adopted on December 3, 1997, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/1997/wp8_en.pdf; Bygrave, L. A., "Data Protection Law, Approaching Its Rationale, Logic and Limits," *The Hague, London, New York: Kluwer Law International*. 63, 75-76 (2002).
 23. On the differences and interactions between Big Data and open data see Marton, Avital, and Blegind Jensen, *supra* n.7, which point out that "while big data is about distributed computation and infrastructures, open data is about standards on how to make data machine-readable, and hence linkable." From a European perspective, see the recent approved Directive 2013/37/EU of the European Parliament and of the Council of June 26, 2013, amending Directive 2003/98/EC on the re-use of public sector information, available at <http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2013:175:SOM:EN:HTML>. See also Veenswijk, M., Koerten, H. and Poot, J., "Unravelling Organizational Consequences of PSI Reform – An In-depth Study of the Organizational Impact of the Reuse of Public Sector Data," Helsinki: ETLA (2012), <http://www.eta.fi/en/julkaisut/dp1275-en/>. With regard to the US open data policies, see Executive Office of the President, National Science and Technology Council, "Smart Disclosure and Consumer Decision Making: Report of the Task Force on Smart Disclosure," http://www.whitehouse.gov/sites/default/files/microsites/ostp/report_of_the_task_force_on_smart_disclosure.pdf (2013).
 24. See Deloitte, "Open data, Driving growth, ingenuity and innovation," London, 16-20 (2012), <http://www.deloitte.com/assets/dcom-uk/uk/local%20assets/documents/market%20insights/deloitte%20analytics/uk-insights-deloitte-analytics-open-data-june-2012.pdf>, and <http://data.enel.com/> (accessed Sept. 29, 2013) project that shares data sets regarding Enel, an Italian multinational group active in the power and gas sectors. See also the data sets made available by Nike Inc., which are available at <http://www.nikeresponsibility.com/report/>

- downloads* (accessed Sept. 29, 2013), the ASOS API project at <http://developer.asos.com/page> (accessed Sept. 29, 2013) and the Canadian Goldcorp Inc. case at <http://www.ideaconnection.com/open-innovation-success/Open-Innovation-Goldcorp-Challenge-00031.html> (accessed Sept. 29, 2013).
25. See the projects and activities of the Open Knowledge Foundation (OKF), which is a nonprofit organization founded in 2004 and dedicated “to promoting open data and open content in all their forms—including government data, publicly funded research and public domain cultural content”; see <http://okfn.org> (accessed Sept. 10, 2013).
 26. “Who’s Lobbying provides access to data about lobbying activity in the UK,” see <http://whoslobbying.com/about> (accessed Sept. 10, 2013).
 27. Openpolis develops and implements projects “to enable free access to public information on political candidates, elected representatives, and legislative activity thus promoting transparency and the democratic participation of Italian citizens,” see <http://www.openpolis.it/eng/> (accessed Sept. 4, 2013).
 28. The aim of Openspending “is to help track every (public) government and corporate financial transaction across the world and present it in useful and engaging forms for everyone from a school-child to a data geek,” see <http://openspending.org> (accessed Sept. 4, 2013).
 29. OpenStreetMap is a collaborative project to create a free editable map of the world, see <http://www.openstreetmap.org> (accessed Sept. 4, 2013).
 30. See also <http://www.openingparliament.org/> (accessed Sept. 4, 2013), a forum to help connect various world civic organizations that are engaged in monitoring, supporting, and opening up parliaments and legislative institutions; <http://opencorporates.com/> (accessed Sept. 4, 2013), a Web site that shares data on corporate entities. Another relevant bottom-up initiative is represented by Dbpedia, which is a project to extract structured content from Wikipedia and make this available on the Web in an open format, see <http://dbpedia.org> (accessed Sept. 4, 2013) (“Dbpedia allows you to ask sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data”). For this reason, Dbpedia represents a driving force for further bottom-up initiatives based on the use and reuse of information. See Cyganiak, R. and Jentzsch, A., “The Linking Open Data cloud diagram,” (2011), accessed Sept. 4, 2013, <http://lod-cloud.net/>.
 31. See Gurstein, M., *supra* n.8 (“Efforts to extend access to “data” will perhaps inevitably create a “data divide” parallel to the oft-discussed “digital divide” between those who have access to data which could have significance in their daily lives and those who don’t.”).
 32. Article 3(2)(a) of the Proposal for a Regulation 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, available at http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf.
 33. See Kroes, N., “The Digital Agenda: Europe’s key driver of growth and innovation,” SPEECH/11/629 (Brussels, October 4, 2011), http://europa.eu/rapid/press-release_SPEECH-11-629_en.htm.