1. Introduction – new legal challenges of big data

Joe Cannataci, Valeria Falce and Oreste Pollicino

Big data is no longer a new phenomenon, but its implications are still compelling.¹ The increasing spread of automated technologies involves all sectors and different areas of the world with different cultures and legal systems. These technologies can provide predictive answers, replacing the human role in decision-making. The fortunes of these technologies are linked to their ability to process data and extract value from them. Unlike oil, data do not lose their value after processing and use. This form of perpetuity leads to the question how to address the new challenges raised by big data analytics and the role of private and public law from a transnational perspective.

As with many other expressions in the field of new technologies, the term ‘big data’ has become a metaphor for the development of the information society. In 2011, the McKinsey Global Institute defined big data as a dataset whose size exceeds a database’s ability to acquire, store, manage and analyse data and information.² Taking a few steps backwards, in 2001, Laney outlined the three-dimensional growth model of the ‘3 Vs’: Volume, Velocity and Variety.³ These characteristics, outlined in the context of e-commerce, describe, first of all, an increase in the processing of data coming from new sources such as sensors (Volume). There is also a spread of the homogeneous and heterogeneous sources from which data are processed. If, on the one hand, the increase in quantity constitutes an index to be kept under careful consideration, on the other hand, the heterogeneity of sources and types of data comprise a fundamental element to fully understanding the phenomenon (Variety). While in the past, the processing of data was primarily based on information

¹ V. Mayer-Schönberger and K. Cukier, *Big Data. A Revolution that Will Transform How We Live, Work, and Think* (Houghton Mifflin Harcourt 2013).

² J. Manyika et al., ‘Big Data: The Next Frontier for Innovation, Competition, and Productivity’, McKinsey Global Institute, 2011.

³ D. Laney, ‘3D Data Management: Controlling Data Volume, Velocity and Variety’, META Group, 2001.

kept in rigid databases, the new possibilities of data analytics also enable the exploitation of unstructured data. The third element of growth involves the rapid production and exchange of vast amounts of data in limited time spans (Velocity).

Laney's model was then enriched by (at least) two other characteristics, namely Veracity and Value.⁴ The former explains how the accuracy of the data inevitably affects the quality of the final result. The latter expresses the ability of big data to create business opportunities by generating values from the processing of information. Big data analytics can define models or find correlations between structured and unstructured datasets. In other words, the big data analytics phase can be considered in a twofold guise: on the one hand, it constitutes the stage from which the value derives through the analysis of different categories of data; on the other, it constitutes the stage that produces higher risks from the predictability of the output.

This definition is particularly relevant since it allows appreciation of an evident paradigm shift compared with traditional forms of data processing based on deterministic rules. Big data analytical techniques allow new forms of processing using unstructured or semi-structured data such as multimedia content and social media accounts.⁵ Therefore, considering big data as a phenomenon that combines quantitative and qualitative data, it is possible to consider big data as a 'new generation of technologies and architectures, designed to economically separate value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery and analysis'.⁶

Whereas the big data definition is settled, in the data ecosystem, the developing mix of processing technologies and analysis of data and information raises unexpected opportunities as well as new legal challenges. First of all, from an intellectual property rights perspective, the current legal framework does not fit exactly with the new digital technologies. As big data are also increasingly strategic resources for private companies, software and database rights together with trade secret protection could either lead to an overprotection of the market players that process information thanks to powerful and self-learning algorithms or, under a different line of reasoning, leave totally unprotected the investments underlying the creation of industrial data, generated by machines (such as sensors gathering climate information, satellite imagery, digital pictures and videos, purchase transaction records, GPS signals, etc.).

⁴ C. Tsai et al., 'Big Data Analytics: A Survey' (2015) 2 *Journal of Big Data*.

⁵ R. Cumbley and P. Church, 'Is Big Data Creepy?' (2013) 29 *Computer Law and Security Review* 601.

⁶ P. Jain, M. Gyanchandani and N. Khare, 'Big Data Privacy: A Technological Perspective and Review' (2016) 3 *Journal of Big Data*.

Therefore, a starting point for a sound analysis is that intellectual property rights are necessary but not sufficient to promote innovation and competition and to realize an efficient digital single market.⁷ As nicely put by **Sofia Oliveira Pais**, the flow of data (and the intellectual property rights insisting on it) is crucial to the well-being of companies in a data-driven economy, but needs, nevertheless, to be reconciled with other interests in order to maximize the social welfare. EU innovation policy, in fact, on the one hand promotes data production and collection as well as data disclosure (for instance the Directive on the protection of databases⁸ or the new Directives on trade secrets⁹ and copyright),¹⁰ and, on the other hand, explores several possible solutions

⁷ As Herbert Zech explained in ‘A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data’ [2016] *Journal of Intellectual Property Law & Practice* 460–470, ‘the task of the law is to ensure that data markets exist (since the exchange and use of data are desirable) [and] to minimize regulatory risks for market participants’. The existence of a market for data as well as property rights for personal information has also been discussed by M.M.S. Karki in ‘Personal Data Privacy and Intellectual Property’ [2005] 10 *Journal of Intellectual Property Rights* 59. Although the concept of data property right (or moral right or sui generis right) may conflict with the European view of privacy as a fundamental civil right, the author suggests that those solutions should be considered.

⁸ Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, [1996] OJ L 77/20. For a recent overview of this Directive, see Valeria Falce, ‘Copyrights on Data and Competition Policy in the Digital Single Market Strategy’ [2018] *Rivista Italiana di Antitrust* 1, 32–44.

⁹ Trade Secrets Directive, Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure (Text with EEA relevance) [2016] OJ L 157/1.

¹⁰ European Parliament and Council, ‘Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market’ COM/2016/0593 final – 2016/0280 (COD), which extends EU copyright law. The new Directive entered into force in June 2019, cf. Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (EU Copyright Directive) [2019] OJ L 130/92.

to import data protection in competition analysis via access rights,¹¹ liability¹² and contractual protection.¹³ From this enlarged innovation policy viewpoint, exploitative abuses¹⁴ and abuse of economic dependence have proven to be suitable means of qualifying competition law as another layer of data rights protection and should be more consistently considered at the European level.¹⁵

¹¹ See Inge Graef, SihYuliana Wahyuningtyas and Peggy Valcke, 'Assessing Data Access Issues in Online Platforms' [2015] 39 Telecommunications Policy 375, highlighting that EU authorities are more likely to accept antitrust liability for data access refusals than other jurisdictions like the US; see also Alberto Alemanno, 'Data for Good: Unlocking Privately-held Data to the Benefit of the Many' [2018] 9 European Journal of Risk Regulation, 2. See also Bertin Martens, 'The Importance of Data Access Regimes for Artificial Intelligence and Machine Learning' [2018] JRC Digital Economy Working Paper 09, arguing that, although the conclusions in this context are not clear, 'data access rights matter for the downstream economic value generated by the use of data, in particular for value generated in data aggregation and use in ML applications'.

¹² European Commission, 'Building a European Data Economy', Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 10 January 2017, COM (2017) 9 final, 13. See also, European Commission, 'Staff Working Document on the Free Flow of Data and Emerging Issues of the European Data Economy', Brussels, 10 January 2017, SWD(2017) 2 final, 33–38.

¹³ Staff Working Document, pp. 43–44.

¹⁴ Directive (EU) 2019/633 on unfair trading practices in business-to-business relationships in the agricultural and food supply chain [2019] OJ L111/59. Some Member-States (like Portugal or Germany for instance), have already introduced the concept of abuse of economic dependence in their competition law. For an overview of this problem, see ICN, 'Report on Abuse of Superior Bargaining Position' (2008), available at: <http://www.internationalcompetitionnetwork.org/uploads/library/doc386.pdf> (accessed 1.7.2019). Bundeskartellamt, Facebook Decision from 6.2.2019 B6-22/16, Case Summary, 15.2.2019, available at: https://www.bundeskartellamt.de/SharedDocs/Entscheidung/EN/Fallberichte/Missbrauchsaufsicht/2019/B6-22-16.pdf?__blob=publicationFile&v=4 (accessed 1 March 2019); Bundeskartellamt, 'Preliminary Assessment in Facebook Proceeding: Facebook's Collection and Use of Data from Third-party Sources is Abusive', 19 December 2017, available at: https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2017/19_12_2017_Facebook.html (accessed 1 March 2019); cf. 'Competition Law and Data', 10 May 2016, pp. 23–24 available at: https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2 (accessed 1 March 2019).

¹⁵ See ICN, 'Report on Abuse of Superior Bargaining Position' (n. 109); *Magill TV Guide/ITP, BBC and RTE*, Case IV/31.851, Commission Decision 89/205/EEC [1988] OJ L 78/43, para 22: The abuse of economic dependence is, therefore, 'characteristic of the existence of a dominant position'. See on this point Case C-226/84 *British Leyland Public Limited Company v Commission* [1986] ECR 1986-03263 and Case T-65/98 *Van den Bergh Foods Ltd v Commission of the European Communities* [2003] ECR II-04653. See the Case C-95/04 P *British Airways plc v Commission*

In addition to property rights, the accumulation and processing of big data as such may raise some new legal challenges. Moving from here, **Antonio Capobianco** and **Pedro Gonzaga** provide a thoughtful analysis of some of the most debated implications of big data, querying satisfactory policy response and verifying whether it would entail the use of competition law or any other regulatory tool, such as consumer protection or data protection. The conclusions they reach are straightforward: researchers and policy analysts should cautiously evaluate the implications of data in order to better distinguish the positive effects from the risks, as well as to identify the best policy tools to address existing concerns. Addressing the legal toolbox as it stands, competition law appears to be suitable to address algorithmic collusion, whereas consumer protection will probably play a more prominent role in addressing the risks of personalized pricing, and data protection will be devoted to tackling privacy concerns. Of course, it would be possible to use competition law, consumer protection and data protection as complements for a more effective enforcement, although this would probably require enhanced coordination between distinct regulatory agencies. **Renato Nazzini** further elaborates the topic. Under his analysis, whereas privacy standards are relevant to competition analysis as qualitative parameters of competition and can play, therefore, a role in the assessment of mergers, agreements and unilateral conduct, it is important to keep data protection regulation and competition policy clearly distinct. Competition law protects competition. When a breach of the data protection rules is either a form of exercise of market power or an element of an exclusionary strategy, competition law can and should intervene. However, while it is superficially attractive to expand the scope of competition policy to deal with new concerns with digital markets, this approach would ultimately

of the European Communities [2007] ECR I-02331 in which the dependence of the agent on British Airways was also relevant in finding dominance; see also the Case T-128/98 *Aéroports de Paris v Commission of the European Communities* [2000] ECR II-03929. In these judgments, abuse of economic dependence is considered a characteristic of dominance; it is a complementary factor in the finding of the dominant position. See the Japanese Act on Prohibition of Private Monopolization and Maintenance of Fair Trade (Act No. 54 of 14 April 1947), available at: https://www.jftc.go.jp/en/legislation_gls/amended_ama09/index.html (accessed 1 July 2019). See for instance the case *San'yō Marunaka, Edion, Ralse and Direx* analysed by Simon Vande Walle and Tadashi Shiraishi in 'Competition Law in Japan', *Comparative Competition Law* (John Duns, Arlen Duke, Brendan Sweeney eds), Edward Elgar Publishing, 2015, p. 14 Available at: SSRN: <https://ssrn.com/abstract=2636263> or <http://dx.doi.org/10.2139/ssrn.2636263> (accessed 1 July 2019). See also Japan Fair Trade Commission, 'Yūetsuteki chii no ranyō ni kansuru dokusen kinshi hō jō no kangaekata' [Guidelines Concerning Abuse of a Superior Bargaining Position Under the Antimonopoly Act] (30 November 2010) (under II.1), available at: https://www.jftc.go.jp/en/legislation_gls/imonopoly_guidelines_files/101130GL.pdf (accessed 1 July 2019).

weaken competition enforcement, turning it into a discretionary and populist tool that may be bent to the contingent political objectives of the day. If undertakings, as well as breaching the General Data Protection Regulation (GDPR), also collude with competitors on privacy standards, or reduce privacy standards in a way that amounts either to an exploitative or to an exclusionary abuse, then, subject to the principle of *ne bis in idem*,¹⁶ they would also breach competition law because they have brought about specific harm to competition that the rules prohibit.

Moving to the actors of the big data economy, platform roles and intermediary responsibility are fully explored by **Damiano Canepa** and **Vicente Bagnoli**. Whereas the data-driven economy relies on data¹⁷, some antitrust-related topics, such as market definition, market power, vertical assessment, assessment of anticompetitive effects and data as an input for platforms, shall be addressed in order to understand how digital platforms and mega-platforms operate and to what extent they may be entitled to intermediary liability under competition law. For Damiano, in particular, the current framework of analysis

¹⁶ Which may be relevant in certain circumstances, although, if the approach suggested in this article is adopted, data protection infringements and competition infringements should be quite distinct both in law and in fact, as the harm addressed by competition law and the ingredients of a competition infringement would be clearly different from the harm and the ingredients of a data protection breach: see R. Nazzini, 'Parallel Proceedings in EU Competition Law: *Ne Bis in Idem* as a Limiting Principle' in B van Bockel, *Ne Bis in Idem in EU Law* (Cambridge, Cambridge University Press 2016), pp. 131–166 and R. Nazzini, 'Fundamental Rights beyond Legal Positivism: Rethinking the *Ne Bis in Idem* Principle in EU Competition Law' (2014) 2 Journal of Antitrust Enforcement 1–35.

¹⁷ OECD, 'Big Data: Bringing Competition Policy to the Digital Era' (2016) <[https://one.oecd.org/document/DAF/COMP\(2016\)14/fr/pdf](https://one.oecd.org/document/DAF/COMP(2016)14/fr/pdf)> accessed 1 August 2019, para. 22; Jacques Crémer, Yves-Alexandre de Montjoye and Heike Schweitzer, 'Competition Policy for the Digital Era' (2019) <<http://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf>> accessed 1 August 2019, 24. The importance of data has been recognized by the European Commission, which explained in *Apple/Shazam* that user data plays an important role in the music industry. In that case, the device data, behavioural data and demographic data gathered by automatic content recognition software such as *Shazam* allowed the party in their possession to develop certain competitive advantages based on data analytics such as identifying upcoming music trends, understanding users' music tastes in order to offer customised playlists and performing target advertising; *Apple/Shazam* (M.8788) Commission Decision [2018] OJ C417, paras 64 ff. On the motives that may explain the motivation a company has to acquire more data, see Nathan Newman, 'Search, Antitrust and the Economics of the Control of User Data' 2014 31(2) Yale Journal of Regulation, art. 5. See also Maurice E. Stucke and Allen P. Grunes, 'Big Data and Competition Policy' (2016) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2612562> accessed 1 August 2019, para. 1.06.

of merger control provides a good and flexible basis to guarantee that competition is protected in digital markets where large amounts of data, which may provide market power to their owner, are at stake.¹⁸ Data may, according to the circumstances, play a role in the various stages of a merger control procedure. The transfer of a database may firstly constitute a concentration within the meaning of the EU Merger Regulation (EUMR); data may then influence the definition of the relevant markets and may be an important element of the substantive assessment of horizontal and non-horizontal mergers (data protection rules have to be taken into account and data may be linked to gains in efficiency); data may constitute an essential element of structural and behavioural commitments taken by the merging undertakings in order to remove a competitive concern. Some further efforts could, however, have a positive effect. The general assessment would however gain in quality if the Commission, in the future, were to better define the framework under which the competitive effects of data and big data must be assessed.¹⁹ In addition, the turnover thresholds should be revised. Further research will have to take into account the different categories of data that may be distinguished depending on their source, nature and use.²⁰ The angle chosen by Vicente is complementary. He suggests instead that the traditional regulatory framework is mainly focused on ‘bipolar’ transactions between businesses and consumers. Hence, an appropriate solution for the increasing number of ‘triangular’ transactions arising from the use of a platform has not been sufficiently provided. The behavioural change in the market structure caused by the rise of digital platforms has brought together duties of the platform operator and the potential liability of the platform operator for non-performance according to law, which can be understood as the platform role and intermediary responsibility. Dominant platforms are and shall be subject to special obligations as with market power comes a special responsibility to the intermediary.

From a different but intertwined perspective, **Mateja Durovic** and **Franciszek Lech** evaluate the main big data concerns in consumer relations

¹⁸ Same opinion: Ben Holles de Peyer, ‘EU Merger Control and Big Data’ (2018) 13(4) *Journal of Competition Law & Economics* 767, 776 ff.

¹⁹ Jay Modrall, ‘Big Data and Merger Control in the EU’ (2018) 9(9) *Journal of European Competition Law & Practice* 569, 571.

²⁰ Jacques Crmer, Yves-Alexandre de Montjoye and Heike Schweitzer, ‘Competition Policy for the Digital Era’ (2019) <<http://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf>> accessed 1 August 2019, 24; Autorit de la concurrence/Bundeskartellamt, ‘Competition Law and Data’ (2016) <<http://www.autoritedelaconcurrence.fr/doc/reportcompetitionlawanddatafinal.pdf>> accessed 1 August 2019, 6, 12: for example, the distinction between online and offline data, between data generated by mobile applications and data collected by fixed devices, between data retrieved from search queries and from social networks.

(security breaches,²¹ companies' inability to faithfully rely on consumer's consent²² and discriminatory treatment), reconciling them to a breach of the fundamental right to privacy when interpreted as the right to informational self-determination²³ and the right to control personal data. Consequently, the fragility of genuine, free and informed consumer consent in the age of big data is another breach of the right to privacy. Since both consumer law and privacy protection aim to promote autonomous decision-making by data subjects or consumers²⁴ (that data protection and consumer laws share, at least to some extent, the same purposes seems evident), a coordinated response shall be advanced in order to avoid an overlap of mutually incoherent legal actions.

Rupprecht Podszun and **Sarah Langenstein** analyse data as an input in competition law cases under an institutional approach. In their opinion, collection of data and its use for solving competition cases are necessary. A major improvement in competition law in recent years has been to make active use of economic evidence, integrating the possibilities of big data computing into the toolbox of modern competition law. However, a tougher stance in checking the accuracy of information provided is required, also because the sheer mass of data offers ample opportunity for 'cherry picking' and confirming hypotheses instead of a neutral assessment. The Court of Justice paved the way in the UPS/TNT merger case where the Court confirmed that the right to defence of the parties had been violated when changes to the econometric model after the Statement of Objections had not been communicated to the merging parties. The Court of Justice held that the parties need to be able to comment 'on the accuracy and relevance of all the factors that the Commission intends to base

²¹ David C. Vladeck, 'Consumer Protection in an Era of Big Data' (2016) 42(2) Ohio N.U.L. Rev. 493, 500, 506–507; Inge Graef, Damian Clifford and Peggy Valcke, 'Fairness and Enforcement: Bridging Competition, Data Protection and Consumer Law' (2018) 8(3) IDPL 200, 202.

²² On consumer consent for data collection see GDPR, art. 6(1)(a); Max N. Helveston, 'Consumer Protection in the Age of Big Data' (2016) 93(4) Wash. U.L. Rev. 859, 864; Wolfgang Kerber, 'Digital Markets, Data and Privacy: Competition Law, Consumer Law and Data Protection' (2016) Joint Discussion Paper Series in Economics, No. 14-2016, Philipps-University Marburg, School of Business and Economics, Marburg, 4.

²³ Stefano Rodotà, 'Data Protection as Fundamental Right', Reinventing Data Protection, International Conference, Bruxelles, 12–13 October 2007, 2; Stefano Rodotà has argued that privacy must be understood as a right to 'keep control over one's own information and determine the manner of building up one's own private sphere'.

²⁴ Inge Graef, Damian Clifford and Peggy Valcke, 'Fairness and Enforcement: Bridging Competition, Data Protection and Consumer Law' (2018) 8(3) IDPL 200, 202.

its decision on'.²⁵ The conclusion they reach is sharable: data scientists in competition law should look not only into the effects of data in markets, but also the use of data by competition agencies and courts.

Whereas competition law, intellectual property rights and consumer law tend to face the new challenges of big data from the perspective of the market, nevertheless, the potentialities of big data go even beyond the private dimension involving individuals' fundamental rights and freedoms. In other words, the role of big data is relevant from a public law standpoint and democratic values.

The protection of privacy and personal data is one of the leading examples.²⁶ There are specific concerns strictly linked to the automated processing of huge amounts of data which could be used, in particular, for surveillance or profiling purposes by governments and private actors.²⁷ In the European framework, privacy and data protection are recognized as fundamental rights by the Charter of Fundamental Rights of the European Union ('Charter'),²⁸ and the European Convention on Human rights ('Convention'),²⁹ respectively by arts 7 and 8. This constitutional framework constitutes the primary safeguard against disproportionate interferences with the individual's rights and freedoms coming from the challenges raised by new data processing techniques.³⁰ The EU approach is not a global standard, but it can be considered a model of protection to follow when regulating data protection.

It is not by chance if the GDPR deals indirectly with big data analytics.³¹ In general, the GDPR adopts a comprehensive approach by requiring data controllers to ensure that the processing of personal data complies with general principles such as transparency, fairness, lawfulness, purpose limitation and data minimization.³² The processing of vast and heterogeneous amounts of data for multiple purposes challenges these principles, requiring data control-

²⁵ Case C-265/17 P *UPS/TNT* [2017] OJ C231/25, para. 31.

²⁶ Christopher Kuner et al., 'The Challenge of 'Big Data' for Data Protection' (2012) 2 *International Data Privacy Law* 47.

²⁷ Ira S. Rubinstein, 'Big Data: The End of Privacy or a New Beginning?' (2013) 3(2) *International Data Privacy Law* 74.

²⁸ Charter of Fundamental Rights of the European Union [2012], OJ C 326/391.

²⁹ European Convention on Human Rights [1950].

³⁰ Serge Gutwirth and Paul De Hert, 'Regulating Profiling in a Democratic Constitutional States', in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen* (Springer 2006), 271.

³¹ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ L 119/1.

³² *Ibid.*, art. 5.

lers to take into account the peculiarities of each data processing involving big data analytics. Furthermore, the principles of privacy by design and privacy by default constitute the other two pillars of EU data protection law.³³ Indeed, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organizational measures designed to implement data protection principles. Moreover, the same measures have to be implemented by default in order to ensure that only personal data which are necessary for each specific purpose of the processing are processed.

The GDPR also provides individuals with a right not to be subjected to automated decisions. In particular, it provides that the data subject shall have the right not to be subjected to a decision based solely on automated processing, including profiling which produces legal effects concerning the data subject or similarly significantly affects him or her.³⁴ Recital 71 provides an example of automatic decisions such as those related to ‘refusal of an online credit application or e-recruiting practices’ where a decision is ‘automatic’ or ‘without human intervention’. Moreover, it is also necessary to take into consideration other provisions of the GDPR requiring businesses using profiling techniques to provide individuals with ‘meaningful information about the logic involved, as well as the significance and the envisaged consequences’.³⁵ In particular, art. 15 requires data controllers to disclose ‘the existence of automated decision-making, including profiling’ upon request. These provisions are functional to make effective the provision of art. 22. However, Recital 63 refers to ‘trade secrets or intellectual property and, in particular, the copyright protecting the software’ as a sort of limit to the disclosure of information about profiling with the consequence that such disclosure should not involve the technical specifications of algorithms that are the basis for their functioning.

In this complex scenario, the chapter of **Giovanni De Gregorio** and **Sofia Ranchordás** enriches the debate on artificial intelligence and data protection by analysing how big data analytics can disrupt information silos allowing the different roles of individuals and the respective data to converge. In practice, this happens in several countries with data sharing arrangements or ad-hoc data

³³ Ibid, art. 25.

³⁴ Ibid, art. 22.

³⁵ Margot Kaminski, ‘The Right to Explanation, Explained’ (2019) 34(1) Berkeley Technology Law Journal 189; Sandra Wachter et al., ‘Why a Right to Explanation of Automated Decision-making Does Not Exist in the General Data Protection Regulation’ (2017) 7(2) International Data Privacy Law 76; Gianclaudio Malgieri and Giovanni Comandè, ‘Why a Right to Legibility of Automated Decision-making Exists in the General Data Protection Regulation’ (2017) 7(3) International Data Privacy Law 243.

requests. However, breaking down the existing structure of information silos in the public sector remains problematic. While big data disrupt artificial silos that may not make sense in the digital society and promote a truly efficient digitalization of data, removing information out of its original context may alter its meaning and violate individuals' fundamental rights like privacy and data protection. This chapter discusses how big data and data analytics are changing information silos and how digital technology is challenging citizens' autonomy. This chapter also explores the need for a more integrated approach to the study of information, particularly in the public sector.

The field of privacy and data protection does not exhaust the concerns at stake. Big data analytics indeed raise challenges for the protection of other fundamental rights and democratic values, especially freedom of expression.³⁶ In this framework, the role of big data and automated decision-making technologies is becoming predominant in the online environment. Such technologies are used by online platforms in order to select and index online contents, driving how information appears and is disseminated across society.³⁷ Focusing on freedom of expression, it is possible to observe that the role of big data and automated decision-making technologies is becoming predominant in the online environment with a strong impact on the ability of users to freely express their thoughts. This interference clearly hinders the freedom of speech which in Europe is protected as a fundamental right recognized by the Charter and the Convention respectively by arts 11 and 10. In particular, the fear of being subjected to general surveillance leads to chilling effects on freedom of expression consisting of self-censorship. Expressing opinions is firmly subject to the possibility of exercising this right freely. The increasing involvement of these technologies in daily life can have strong effects on free speech and, therefore, democratic values.

Without being exhaustive, there are at least other two main concerns: the first is related to the use of indexing and profiling techniques, which lead to the creation of 'filtering bubbles'.³⁸ In these spaces, users exchange and enjoy information that covers only their specific interests as defined by algorithms creating the right environment for the dissemination of disinformation.³⁹ The

³⁶ Jack M. Balkin, 'Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation' (2018) 51 University of California, Davis 1149.

³⁷ Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (2018) 131 Harvard Law Review 1598.

³⁸ Eli Pariser, *The Filter Bubble: What the Internet is Hiding from You* (Viking 2011).

³⁹ Cass Sunstein, *#Republic: Divided Democracy in the Age of Social Media* (Princeton University Press 2017).

second is related to the influence of such technologies on the management of those online contents considered as illicit in order to flag or remove them under the request of the users. One of the open debates in this field concerns the liability regime of internet service providers with regard to the spread of illicit contents uploaded by the users. The use of automated technologies raises specific issues in relation to the passive role of online intermediaries and their awareness. In September 2017, the EU Commission issued a Communication on ‘Tackling Illegal Content Online. Towards an Enhanced Responsibility of Online Platforms’ in which it stressed the risks for freedom of expression in relation to the removal of illicit content carried out by online platforms.⁴⁰ This issue is strictly related not only to the exemption of liability of internet service providers as recognized by the e-Commerce Directive but also to the lack of procedural safeguards.⁴¹

In this regard, the chapter of **Oleg Soldatov** addresses questions relating to how big data influence freedom of expression. To set the stage, it introduces the reimagined definition of freedom of expression in the age of big data, exploring both the nature of the Internet as a medium and the principal challenges presented by the technology. Next, the author explores the context in which this concept is applied, before moving on to provide concrete examples of the interplay between big data and freedom of expression. Before reaching a conclusion, the author discusses ways – both legislative and non-legislative – to move forward in this new reality.

The risks for freedom of expression, privacy and data protection are relevant but not exclusive. Indeed, they are not the only fundamental rights at stake. Indeed, big data analytics also affect human dignity owing to the discriminatory outcome which could result from the processing of data and its impact on minorities and minors.⁴² In particular, discriminatory results originate from biased evidence and inflexible decision-making processes. The right to non-discrimination is one of the fundamental principles enshrined in many constitutions and international charters of fundamental rights. It is based on the general principle of equality according to which similar situations have to be treated in the same way and different situations have to be treated differently. In particular, in the European framework, the principle of non-discrimination

⁴⁰ Communication on Tackling Illegal Content Online, Towards an Enhanced Responsibility of Online Platforms, COM(2017) 555 final. See, also, Recommendation of 1 March 2018 on measures to effectively tackle illegal content online (C(2018) 1177 final).

⁴¹ Giovanni De Gregorio, ‘Democratising Online Content Moderation: A Constitutional Framework’ (2019) 36 Computer Law & Security Review.

⁴² Tal Zarsky, ‘Understanding Discrimination in the Scored Society’ (2014) 89 Washington Law Review 1375.

is enshrined both in the Convention and in the Charter.⁴³ The same principle is provided for by other more specific anti-discrimination regulations. For instance, EU law requiring the implementation of the principle of equal treatment between persons irrespective of racial or ethnic origin distinguishes between direct and indirect discrimination.⁴⁴ The first occurs in the case in which one person is treated less favourably than another has been or would be treated in a comparable situation on specific grounds such as racial or ethnic origin. Indirect discrimination is involved when an apparently neutral provision, criterion or practice would put individuals who belong to a specific category at a particular disadvantage compared with other individuals, unless that provision, criterion or practice is objectively justified by a legitimate aim and the means of achieving that aim are appropriate and necessary.

Usually, algorithms are not used directly to discriminate people on the grounds of racial origin, but it cannot be excluded that some automated results could lead indirectly to discriminatory results. This point is shown, for example, by profiling algorithms that produce discriminatory results against marginalized populations, as in the case of delivery of online advertisements according to perceived ethnicity or decisions about granting financing to specific subjects.

In this field the chapter of **Shulamit Almog** and **Liat Franco** focuses on the issue of children's digital rights and cyberbullying, underlining how big data analytics can be used to protect children or to harm them. They explore whether big data might be employed as a tool to retract data on behavioural mechanisms in children's communication and used to design new legal and non-legal frameworks to enhance children's rights in general and to deal with cyberbullying in particular.

The last chapter of this book addresses the regulation of big data analytics. After the book has analysed the new legal challenges of big data, this chapter provides perspectives on how to deal with this phenomenon. Similarly to other new technologies, big data shows a two-fold nature. On the one hand, these new processing techniques can foster fundamental rights and freedoms while, at the same time, promoting new paths of growth. On the other hand, they raise new legal, economic and ethical challenges. The dichotomy between risk and opportunity leads to a question about the potential regulatory model for big data analytics. Furthermore, one of the primary concerns also involves the (national or international) level from which these rules should be adopted.

⁴³ Convention, art. 14; Charter, art. 21.

⁴⁴ Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin OJ L 180/22.

The chapter of **Francesca Lagioia and Giovanni Sartor** addresses the risks and opportunities of big data analytics under a regulatory perspective, especially the need to understand how to ensure that artificial intelligence benefits citizens and communities, respecting individual rights and social values. After introducing the concept and methods of artificial intelligence, the authors consider the prospects of machine learning technologies, as applied to big data, as well as the opportunities and risks of these technological developments. In particular, the authors stress how artificial intelligence, on the one hand, further expands the persistence and pervasiveness of citizens' assessment and control, but, on the other hand, can also support individuals.

Within this framework, this book aims to detect and understand first the legal and economic implications of the new knowledge-based economy before coming to regulatory solutions and providing lawmakers and authorities with sound interpretative tools to implement in order to ensure balance between fundamental rights and competition and innovation. New digital technologies have raised challenges that require new skills and resources to assess all of the legal and economic implications. Indeed, the relevant issues are not merely restricted to a specific law field but involve fundamental rights of individuals and economic planning choices.

The book moves from the assumption that the deep understanding of the new challenges triggered by the data economy shall influence the institutional choices that are required at European as well as at the international level in order to balance innovation, competition and fundamental rights in the data era. For these reasons, this book firstly aims to investigate the drivers of the data-driven economy, focusing on data value and big data analysis and exploring the implications of the new challenges from a fundamental rights, intellectual property and competition policy perspective, with the aim of verifying the adequacy of the legal remedies that have been implemented so far. Secondly, this book provides a constitutional and public law standpoint which shows how big data analytics impact on fundamental rights and democratic values. The mix between the private and the public law perspectives is the added value of this editorial project which outlines how the boundaries between different areas of the world are increasingly blurred and going transnational.