



## Puzzles in the big data revolution: an introduction

Alain Marciano<sup>1</sup> · Antonio Nicita<sup>2</sup> · Giovanni Battista Ramello<sup>3</sup>

Accepted: 21 October 2020 / Published online: 13 November 2020

© The Author(s) 2020

Our society is undergoing a genuine digital transformation that implies “a set of technological, cultural, organizational, social, creative and managerial changes associated with digital technology applications in all aspects of human society” (Stolterman and Fors 2004, 689). According to McDonald and Rowsell-Jones (2012), “the digital transformation goes beyond the simple adoption of new technologies and allows, on one side, to provide services, goods and experiences, and on the other side, to find, to process and to make accessible large quantities of contents, creating pervasive new connections between people, places and things”. Indeed, this “great digital transformation” is shaping a new society—a “digital market society”, in which all previous social, economic, and political dimensions of life have changed. There are now or there will very soon be new—and different compared to nowadays—norms of consumption, modes of production, institutions, business models and individual interactions. As many voices claim, we are just at the beginning of this co-evolutionary path leading toward an epochal revolution. This process is affecting all sectors and all countries. The disruptive innovations resulting from this digital transformation involves more and more individuals, businesses and objects, having a global rather than national dimension. It is thus not a surprise that the United Nations Sustainable Development Goals, which set access to information and communications technology and universal and affordable access to the Internet as keys to a future sustainable world.

Like each long journey, the digital transformation has proceeded step by step: at the beginning there were only telecommunications networks, then the internet came and afterward the Web and its evolutions. According to the *IDC Data Age*

---

✉ Giovanni Battista Ramello  
giovanni.ramello@uniupo.it

<sup>1</sup> University of Montpellier and MRE, Montpellier, France

<sup>2</sup> Lumsa University, Roma, Italy

<sup>3</sup> Università del Piemonte Orientale, Alessandria, Italy

2025 forecast, in the next 6 years the volume of data created in the “datasphere” is expected to increase by 5 times.

Volume, velocity and variety<sup>1</sup> are the main ‘v’ that characterize the data economy and that will shape global datasphere in the near future, changing the nature of products and services exchanged in the digital ecosystem, as well as the structure of digital markets. In turn, this process raises new trade-offs and dilemmas for law and economics scholars, such as for regulators and policymakers.

As the digital ecosystem becomes fundamentally driven by *datasphere* growth, traditional value chains and vertical organisation in the communications industry will be rapidly transformed by new modes of production and exchange of economic goods and services as well as by the emergence of new services and user needs. Standard economic transactions are thus redesigned in terms of producer and/or recipient of data, while algorithm-driven platforms take centre stage, with data exchange at the core of economic transactions involving multisided markets.

The 5G revolution, with the increase of IoT and M2M connections, will further reveal that data is king being the real ‘economic product’ which creates value in the new digital product cycle for “matchmaker” platforms, online advertisers, and so on. The digital market society is nowadays a networked multi-layered ecosystem, whose fundamental backbone is composed of very high capacity fixed and mobile communications physical infrastructures, which however are neither its brain nor its soul. Services, contents, applications and their providers permeate the digital society with sophisticated, pervasive, self-nurturing networks made of economic, social and operational interactions (human and non-human), taking place at various levels of the materiality-immateriality scale. This ecosystem is extremely complex, and it is populated, animated and influenced by a wide range of diverse actors:

- private and public organizations, running very different kind of business or providing public services (e.g., ICT manufacturers, network operators, communications service providers, digital platforms, creative content producers and aggregators, advertisers, applications producers, “vertical” companies);
- mid and end-users consuming different kind of digital and ICT products (e.g., private and public organizations, “digital champions”,<sup>2</sup> digital experts and unsophisticated consumers)

Among all the fundamental actors in the digital market society, nowadays digital platforms are crucial. They are also called Over-The-Top providers (OTTs), because they provide services to users via the public internet and telecoms infrastructures, but “over the top” of the traditional telecoms market value-chain. As a matter of

<sup>1</sup> The 3 Vs seems to be one of the leit motive of Big Data. Volume refers to the amount of data generated through websites, portals and online applications. Velocity refers to the speed with which data are generated and variety stands for the amount of different data that can be generated. See e.g. Gewirtz (2018).

<sup>2</sup> Establishing a group of EU “Digital Champions” is one of the policy within the Digital Single Market strategy put in place by the European Union: Each Eu member state appoint a Digital Champion to support the establishment of an inclusive digital society and to advise the European Commission. <https://ec.europa.eu/digital-single-market/en/digital-champions>.

fact, within the digital transformation, the linear product value chain of the industrial economy has been replaced with the non-linear platforms' ecosystems.

The digital market society comprises a multitude of OTTs, which have a broad scope of activities, comprising for example online advertising, marketplace services, internet search engines, social media, creative contents aggregations and distribution, video-sharing, communications services, products price comparison, application distribution, payment system services, collaborative activities, etc.<sup>3</sup> The main OTTs, such as Amazon, Apple, Facebook, Google and Microsoft have been renamed “Big Techs”, being the few global platforms pushing the economic interactions over the internet from the mere trading from on-line shops needing many additional off-line actions, to several diverse forms of digital capitalism, acting as “matchmakers” among different groups of consumers and/or businesses in the digital market society (Schiller, 2000).

Moreover, with the widespread development of this immersive, enveloping new economies, and the increasing use of the internet, particularly via mobile devices, all interactions, transactions, consumption and production within the digital market society became an unlimited source of data, which started to become an economic and social key-asset and productive factor. The “data economy” is composed of different types of market players—such as manufacturers, researchers and infrastructure providers—collaborating to ensure that data is accessible and usable. This enables the market players to extract value from this data, by creating a variety of applications with a great potential to improve daily life (e.g. traffic management, optimisation of harvests or remote health care).

As with any revolution, the emergence of the datasphere society generates important challenges growing both at the level of the law and economics theory of institutions and at level of policy design. Studying these main challenges is what the papers gathered in this special issue aim at doing. Most of the papers were presented during a workshop organized in Rome, on December 3, 2018, and sponsored by the *Autorità per le Garanzie nelle Comunicazioni* (AGCOM, Authority for Communications Guarantees) which is the Italian regulatory authority for the communication industries.

First, Marciano et al. (2020) propose a paper about one of the major challenges that result from the digital revolution occurring in our societies bears on the organization and regulation of the economy. This leads to a creation of a new form of capitalism—platform capitalism, that rests on a fundamental dilemma between ‘decentralization’ on the one side and ‘concentration’ on the other is the main puzzle that the emergence of a big data driven economy is actually offering to law and economics scholars and to policy makers. This paper analyzes the major aspects of this dilemma.

One of them clearly relates to the role of platforms and their importance in the “new digital society”. Their position, that allows them to gather data about individual behaviors—the so-called Big Data –, gives them a power that has probably never been achieved in any economy previously. As Arpetti and Iovanella (2020) explain

<sup>3</sup> See EU Commission Communication on Online Platforms and the Digital Single Market Opportunities and Challenges for Europe—COM/2016/0288 final. See also BEREC Report on OTT services - BoR (16) 35.

in their article, “Towards more effective consumer steering via network analysis”, there is now an “an unprecedented concentration of information related to the individuals’ preferences in the hands of a few gatekeepers.” Platforms could nonetheless be even more performant, so to speak, than they are in their use of the information consumers provide them when they buy online. This is what Arpetti and Iovanella show: platforms could extract even more information than they do—and even, “more information than it is reasonable to imagine”, as they conclude—and therefore “boost their predictive capabilities in terms of individuals’ purchasing capacity and potentially consumers’ steering practices.” To evaluate this advantage they propose a measure called “Network Information Patrimony”.

Big data do not only allow platforms to know what consumers want to buy and their reservation price as well. Thanks to technology, platforms are thus able to perform first-degree price discrimination and capture the consumer’s surplus. Yet, as Botta and Wiedemanb (2020) emphasize in “To discriminate or not to discriminate? Personalised pricing in online markets as exploitative abuse of dominance”, “there is limited evidence that online platforms are already enforcing this kind of discriminatory strategy.” This might be because firms are reluctant to use a strategy that consumers do not like. Or, also, because it is prohibited by the law, even though, as Botta and Wiedermann explain, it is not sure how the law can actually be effectively implemented—the proof would indeed be difficult to establish. They suggest that personalized pricing should not be banned a priori but assessed on a case-by-case basis and that European authorities should negotiate “behavioural commitments” with online platforms. That would be more efficient than fines. In any case, the situation seems to be unclear, for the legislator and for platforms—that “are waiting for further guidance from courts and competition law enforcers.”

In a way that nicely complements Botta’s and Wiedermann paper, “AI algorithms, price discrimination and collusion: a technological, economic and legal perspective”, by Gautier et al. (2020) studies how technology and, more precisely, artificial intelligence can indeed give birth to price discrimination and tacit collusion. They show that price discrimination and tacit collusion do not require the same kind of data and do not use the same kind algorithms. It even appears that the technology allowing price discrimination is maturing more rapidly than the one allowing tacit collusion. This analysis confirms that personalized pricing could be a real threat in the future and that, as Botta and Wiedermann have shown, enforcement agencies have to find legal solutions to face the possible use of this strategy by platforms. The analysis developed by Gautier, Ittoo and van Cleynenbreugel also shows that lawyers and policy makers have to pay attention to how technologies develop before envisaging new legal rules. There is real need to understand the nature of the technological challenges we are facing before starting to regulate and produce new laws.

To regulate platforms is not only a matter of economic welfare, of consumer’s surplus—that is of economic freedom—but has also to do with non-economic freedom such as free speech. This is what Frank Fagan’s paper (2020), “Optimal social media content moderation and platform immunities”, studies. Speech is certainly one of the most important area that has to be protected from coercion but speech has to be regulated. Especially on online platforms where fake news can spread easily, at a very low cost, and damage the “institutional health” of a society. Would

lawmakers, “concerned with platform civility and its impact on institutional health”, choose “platform immunity” or “implementing a platform liability regime”? Fagan proposes a model in which he shows that lawmakers prefer immunity “if the costs of implementing a platform liability regime are greater than the costs of enforcing status quo law” and also “inasmuch as implementation of a platform liability regime is coupled with new speech restrictions that are unconstitutional or prohibitively costly”. But, Fagan concludes, “platforms are free to set strong content moderation policies consistent with existing law.”

Technology is not only costly and a matter of concern. It can also have positive consequences. This is what shows the next paper, “Automated fact-value distinction in court opinions” by Cao et al. (2020). Cao, Ash and Chen start from “a major theme of legal discourse in Common Law countries”, namely that legal decisions are a mix of statements about facts, which are descriptive, and statements about values, which are normative. Do judges rely mainly on facts or on values? Or the reverse? Answering the question is crucial to improve our understanding of legal decision making and judges reasoning. Cao, Ash and Chen thus propose a “machine learning model” that automatically distinguishes between fact and value statements in written legal opinions. The paper also proposes “two learning experiments, one to evaluate this model by comparing its performance with those of the methods proposed in the previous literature, and the other to understand how its behavior differs from its precedents by analyzing the texts on which their judgments differ.” They show that “the value segments of opinions are more informative than fact segments of the ideological direction of U.S. circuit court opinions.”

Finally, the special issue ends with a paper with broad and particularly important consequences. “Infrastructure and technologies: a technology flow framework”, written Hogendorn and Frischmann (2020), raises the fundamental question of the nature of “technology”. Most of the time, Hogendorn and Frischmann argue, economists carelessly put technologies such as electricity, railroads, the Internet or the steam engine and the computer are put together in the same category—general purpose technologies. They thus neglect a major difference that exist between the first three—they can be defined as “infrastructures”—and the last two—they are not infrastructures. To understand the difference and to understand whether or not a facility qualifies as “infrastructure” is crucial since, for various reasons Hogendorn and Frischmann present, infrastructural resources “are key to the development and growth”. But, on the other hand, these resources have specific features that raise challenges—partial nonrivalry or tethering—that must necessarily be taken into account by policy makers.

**Funding** Open access funding provided by Università degli Studi del Piemonte Orientale Amedeo Avogadro within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article

are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arpetti, J., & Iovanella, A. (2020). Towards more effective consumer steering via network analysis. *European Journal of Law and Economics* (this issue).
- Botta, M., & Wiedemanb, K. (2020). To discriminate or not to discriminate? Personalised pricing in online markets as exploitative abuse of dominance. *European Journal of Law and Economics* (this issue).
- Cao, Y., Ash, E., & Chen, D. L. (2020). Automated fact-value distinction in court opinions. *European Journal of Law and Economics* (this issue).
- Fagan, F. (2020). Optimal social media content moderation and platform immunities. *European Journal of Law and Economics* (this issue).
- Gautier, A., Ittoo, A., & Van Cleynenbreugel, P. (2020). AI algorithms, price discrimination and collusion: a technological, economic and legal perspective. *European Journal of Law and Economics* (this issue).
- Gewirtz D. (2018). *Volume, velocity, and variety: Understanding the three V's of big data*. ZDnet available at <https://www.zdnet.com/article/volume-velocity-and-variety-understanding-the-three-vs-of-big-data/>
- Hogendorn, C., & Frischmann, B. (2020). Infrastructure and general purpose technologies: A technology flow framework. *European Journal of Law and Economics* (this issue).
- Marciano, A., Nicita, A., & Ramello, G. B. (2020). Big Data and Big Techs: Understanding the Value of Information in Platform Capitalism. *European Journal of Law and Economics* (this issue).
- McDonald, M. P., & Rowsell-Jones, A. (2012). *The Digital Edge, Exploiting Information and Technology for Business Advantage*. Stamford: Gartner Inc.
- Schiller, D. (2000). *Digital Capitalism. Networking the Global Market System*. Cambridge: MIT Press.
- Stolterman, Erik, & Fors, Anna Croon. (2004). Information Technology and the Good Life. In B. Kaplan, D. P. Truex, D. Wastell, A. T. Wood-Harper, & J. DeGross (Eds.), *Information Systems Research. Relevant Theory and Informed Practice* (pp. 687–692). New York: Springer.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.