Perspectives on the value of Big Data sharing

Christopher L. Tucci, [orcid.org/0000-0001-8733-9530] Imperial College Business School, London, UK - email: c.tucci@imperial.ac.uk
Gianluigi Viscusi, [orcid.org/0000-0003-0770-7108] Imperial College Business School, London, UK - email: g.viscusi@imperial.ac.uk

Introduction

The concept of Big Data has gone through several cycles of awareness, with no small amount of industry hype (Davenport *et al.*, 2012) followed by a growing interest in academia (Abbasi *et al.*, 2016; Batini *et al.*, 2015; Buhl *et al.*, 2013; Goes, 2014; Günther *et al.*, 2017; Rai, 2016). The current common understanding of Big Data can be summarized by the following definition that appeared in 2013 in the first issue of *Big Data*, one of the first journals on the topic: *Big Data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or doesn't fit the structures of your database architectures. To gain value from this data, you must choose an alternative way to process it* (Dumbill, 2013). The hype (and phenomenon) followed and overlapped with public sector interest in open government data (Bertot *et al.*, 2014), symbolically reinforced at a global level by the memoranda and directives signed by Barack Obama in the early years of his first mandate (Chignard, 2013; Obama, 2009).

The overlap of interests between the private and public sectors raises the question of the different types of value (economic, public, and social value) that Big Data may have and the challenges related to having access to and sharing them, such as data quality and privacy (Batini *et al.*, 2015; Jain *et al.*, 2016; Menon and Sarkar, 2016). This Special Section aims to provide an outlook on research on these issues, considering specifically the connection on one hand between Big Data, public safety, security, and quality of life; and on the other hand, on the different paths of business model innovation (cf. Massa and Tucci, 2021) enabled by Big Data such as social innovation (Misuraca *et al.*, 2018) and crowd-driven innovation (Afuah and Tucci, 2012; Tucci *et al.*, 2018).

Inspired by the rise of Big Data platforms and infrastructure that handle both structured and unstructured data from a multitude of domains and data sources (ranging from environmental and weather data to wearables, passenger vehicle sensors, financial and insurance institutions' data streams, and social web data), this Special Section explores the benefits and advantages—as well as the challenges, limitations, and threats (at the data security and privacy level)—that emerge from the Big Data value chain (Curry, 2016; Miller and Mork, 2013), delivering "intelligence" to support operations that surround various aspects of human living.

It is worth noting that the Special Section also considers the social value impact of Big Data-driven innovation in terms of capabilities and *functionings* enabled by emerging Big Data ecosystems. In particular, we refer to the construct of the "capability approach" (Nussbaum, 2011; Sen, 1992) where the focus is on "human functionings," which are the various states of human beings and the doings or activities that a person can undertake; and *capabilities*, i.e., the opportunities to achieve functionings as outcomes. For example, traveling is a functioning, and the real opportunity to have a safe and secure trip is the corresponding capability. Taking these issues into account, Big Data and open linked data are a key resource for enabling capabilities, support decision-making on these issues, and develop appropriate policies and

services (e.g., Viscusi et al., 2014). Furthermore, the Big Data-related phenomenon of the *quantified self* (individuals' self-tracking of any kind of biological, physical, behavioral, or environmental information: Swan, 2013) has recently been associated with subjects other than human beings, such as cars. Vehicles in general are able to capture sensor data about themselves and about their environment, thus becoming *quantified vehicles* and creating value to be captured by emerging ecosystems enacted by data-driven services (Kaiser *et al.*, 2021; Stocker *et al.*, 2017). Accordingly, the emergence of different *quantified subjects* raises questions about the role of Big Data for public safety and security as well as the need for understanding the resulting infrastructural challenges and for designing new platforms and services.

In summary, the Special Section provides a multidisciplinary understanding of the impact of Big Data on personal safety, personal security, and well-being. In addition, the Special Section presents both solutions and case studies. In what follows we provide an overview of the articles in the Special Section, before proposing directions for future research on the value of Big Data and data sharing amongst individuals as well as private and public organizations.

Articles in the Special Section

The articles in the Special Section tackle several of the original aims as articulated above. The article by Cho et al. (2021), "Creating value using public Big Data: Comparison of driving factors from the provider's perspective," investigates the use of open government data (OGD) from the perspective of data providers in the public sector. To this end, the authors use the analytic hierarchy process and analytic network process methodologies to rank the importance of the factors driving the use of OGD, and a series of interviews with experts to compare the characteristics of OGD-related tasks among the departments in a public agency. Their findings suggest that the factors driving the use of OGD are not only related to the political or technological environment, but also depend on the institutional culture within the organization of the data provider. Moreover, considering the focus on the different perceptions of the factors valued by public officials in charge of data in institutions, the study suggests the need to explore officials' perceptions of value creation in Big Data domains such as OGD.

Moving to the article by Hsiang et al. (2021), "Predicting popular contributors in innovation crowds: The case of My Starbucks Idea," here the purpose of the study is to predict popular contributors in crowdsourcing through the analysis of textual representations of user-generated content in open crowds. The results of the experiments presented in the article show that popular contributor offerings in open crowds can actually be predicted through user-generated content. Also, this research suggests that the learning effect that can benefit the companies that pay attention to popular contributors might lead to an improvement in their agility in satisfying customer needs.

Finally, the article by Halwani et al. (2021), "Job qualifications study for data science and Big Data professions," provides an interesting analysis of the different skills required for data science and other professional profiles related to the exploration and exploitation of Big Data analytics. Besides providing an empirical perspective on the existing portfolio of data science skills and capabilities, the authors' findings show that a good alignment exists between the industry view and the academic view on what is needed for performing data science, blending statistical and programming skills. The authors also note that the industry view has remained relatively stable during the fours years covered by the study.

Future research directions

The articles in the Special Section develop the imperative for a systemic perspective on Big Data value that covers the spectrum from public value (exemplified by the paper of Cho et al., 2021) to the marketing value whose driving factors have been explored in the study by Hsiang et al. (2021). Big Data's variety of formats and content, spanning texts to videos to sensors, embed different aspects and actually multidimensional fragments of daily life of people and operations of businesses (e.g., one could not only check the code of an order, but have actually the video and the environmental data of the outlets crossed during the shipping and delivery as well as the opinions and sentiments of the various actors involved or the news from the local contexts touched), thus requiring not simply data engineering or analytical competences but also interpretive lenses acquired through training in social sciences, psychology, and humanities. Otherwise, one runs the risk of ignoring the interdependencies, thus distorting the learning opportunities of the full system, especially considering their increased sharing amongst individuals as well as private or public organizations through the adoption of open innovation or crowdsourcing platforms. Future research could also therefore further explore the role Big Data have in business model innovation, acknowledging business models as complex systems, characterized by interdependencies, nested hierarchies, and information processing (Massa *et al.*, 2018).

The skills required for data science and for any activity aiming to create and capture value from Big Data would require further empirical research, elaborating on the elements and sources considered in the paper by Halwani et al. (2021) in this Special Section. Clearly, from the point of view of organizations in both the public and private sectors, legacy skills are not immediately substitutable by digital skills, nor should they necessarily be (cf. Lanzolla et al., 2020). However, empirical work that attempts to find a balance between digital and legacy skills and the conditions under which digital skills can be developed internally to the benefit of the organization would help shed light on better transitions to Big Data exploitation.

Nowadays, there is a vast corpus of research on Big Data, analytics, and how to create and capture value from them (Fosso Wamba *et al.*, 2015; Parvinen *et al.*, 2020; Schüritz *et al.*, 2017; Spiekermann, 2019). Nevertheless, we argue that future research should further question the different facets and dimensions of the value of Big Data and what they have in common with their use in areas such as robotics, autonomous vehicles, and the increasing adoption of artificial intelligence (A.I.) in the public sector (Kaiser *et al.*, 2021; Viscusi *et al.*, 2020) and the challenges of digital transformation for public and private organizations (Lanzolla *et al.*, 2020). This would imply increasing, for example, the capacity of modeling vast amounts of different distributed data (cf. Viscusi and Batini, 2016) and to introduce specific social science competences coupled with statistical and computer science skills to harness information production by integrating a multifaceted valuation of both Big Data sources and A.I. or analytics outcomes.

In conclusion, we hope you enjoy these thought-provoking articles, and that they stimulate you to explore and expand our knowledge base in this emerging domain. Academic research in the organizational and policy implications of Big Data and related digital topics is in its infancy with many exciting studies to come!

Acknowledgements
The Special Section dissemination and organization has been partially supported by the AEGIS (Advanced Big Data Value Chain for Public Safety and Personal Security) project, which received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 732189.

References

Abbasi, A., Sarker, S. and Chiang, R.H.L. (2016), "Big data research in information systems: Toward an inclusive research agenda", *Journal of the Association for Information Systems*, Vol. 17 No. 2, p. 3.

Afuah, A. and Tucci, C.L. (2012), "Crowdsourcing as a solution to distant search", *Academy of Management Review*, Vol. 37 No. 3, pp. 355–375.

Batini, C., Rula, A., Scannapieco, M. and Viscusi, G. (2015), "From data quality to big data quality", *Journal of Database Management*, Vol. 26 No. 1, pp. 60–82.

Bertot, J.C., Gorham, U., Jaeger, P.T., Sarin, L.C. and Choi, H. (2014), "Big data, open government and e-government: Issues, policies and recommendations", *Information Polity*, Vol. 19, pp. 5–16.

Buhl, H.U., Röglinger, M., Moser, F. and Heidemann, J. (2013), "Big Data - A Fashionable Topic with(out) Sustainable Relevance for Research and Practice?", *Business & Information Systems Engineering*, Vol. 5 No. 2, pp. 65–69.

Chignard, S. (2013), "A brief history of Open Data", *Paris Innovation Review*, available at: http://parisinnovationreview.com/2013/03/29/brief-history-open-data/ (accessed 17 March 2017).

Cho, J.Y. and Lee, B.G. (2022), "Creating value using public big data: comparison of driving factors from the provider's perspective", *Information Technology & People*, Vol. 35 No. 2, pp. 467-493.

Curry, E. (2016), "The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches BT - New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe", in Cavanillas, J.M., Curry, E. and Wahlster, W. (Eds.), , Springer International Publishing, Cham, pp. 29–37.

Davenport, T.H., Barth, P. and Bean, R. (2012), "How 'Big Data' Is Different", *MIT Sloan Management Review*, Vol. 54 No. 1, pp. 43–46.

Dumbill, E. (2013), "Making Sense of Big Data (Editorial)", *Big Data*, Vol. 1 No. 1, pp. 1–2.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2015), "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study", *International Journal of Production Economics*, Vol. 165, pp. 234–246.

Goes, P. (2014), "Editor's Comments: Big Data and IS Research", *Management Information Systems Quarterly*, Vol. 38 No. 3, pp. iii–viii.

Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M. and Feldberg, F. (2017), "Debating big data: A literature review on realizing value from big data", *The Journal of Strategic Information Systems*, Vol. 26 No. 3, pp. 191–209.

Halwani, M.A., Amirkiaee, S.Y., Evangelopoulos, N. and Prybutok, V. (2022), "Job

qualifications study for data science and big data professions", *Information Technology & People*, Vol. 35 No. 2, pp. 510-525..

Hsiang, C.-Y. and Rayz, J.T. (2022), "Predicting popular contributors in innovation crowds: the case of My Starbucks Ideas", *Information Technology & People*, Vol. 35 No. 2, pp. 494-509.

Jain, P., Gyanchandani, M. and Khare, N. (2016), "Big data privacy: a technological perspective and review", *Journal of Big Data*, Vol. 3 No. 1, p. 25.

Kaiser, C., Stocker, A., Viscusi, G., Fellmann, M. and Richter, A. (2021), "Conceptualising value creation in data-driven services: The case of vehicle data", *International Journal of Information Management*, Vol. 59, pp. 102335–102335.

Lanzolla, G., Pesce, D. and Tucci, C.L. (2020), "The Digital Transformation of Search and Recombination in the Innovation Function: Tensions and an Integrative Framework*", *Journal of Product Innovation Management*, John Wiley & Sons, Ltd, Vol. n/a No. n/a, available at:https://doi.org/https://doi.org/10.1111/jpim.12546.

Massa, L., and Tucci, C.L.. "Innovation and Business Models." Oxford Research Encyclopedia of Business and Management. 26 Apr. 2021; Accessed 2 Jul. 2021. https://oxfordre.com/business/view/10.1093/acrefore/9780190224851.001.0001/acrefore-9780190224851-e-296.

Massa, L., Viscusi, G., and Tucci, C. L. (2018), "Business Models and Complexity, Journal of Business Models", Vol. 6, No. 1, pp. 70-82.

Menon, S. and Sarkar, S. (2016), "Privacy and Big Data: Scalable Approaches to Sanitize Large Transactional Databases for Sharing", *MIS Quarterly*, Vol. 40 No. 4, pp. 963–981.

Miller, H.G. and Mork, P. (2013), "From Data to Decisions: A Value Chain for Big Data", *IT Professional*, Vol. 15 No. 1, pp. 57–59.

Misuraca, G., Pasi, G. and Viscusi, G. (2018), "Understanding the Social Implications of the Digital Transformation: Insights from Four Case Studies on the Role of Social Innovation to Foster Resilience of Society. BT - Electronic Participation - 10th IFIP WG 8.5 International Conference, ePart 2018, K".

Nussbaum, M.C. (2011), *Creating Capabilities - The Human Development Approach*, The Belknap Press of Harvard University Press, Cambridge (MA).

Obama, B. (2009), "Transparency and open government. Memorandum for the heads of executive departments and agencies", available at: https://www.whitehouse.gov/open/documents/open-government-directive (accessed 7 April 2016).

Parvinen, P., Pöyry, E., Gustafsson, R., Laitila, M. and Rossi, M. (2020), "Advancing data monetization and the creation of data-based business models", *Communications of the Association for Information Systems*, Vol. 47 No. 1, pp. 25–49.

Rai, A. (2016), "Synergies between big data and theory", *MIS Q.*, Vol. 40 No. 2, pp. iii–ix.

Schüritz, R., Seebacher, S. and Dorner, R. (2017), "Capturing Value from Data: Revenue Models for Data-Driven Services", *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Sen, A. (1992), *Inequality Re-Examined*, Clarendon Press, Oxford.

Spiekermann, M. (2019), “Data Marketplaces: Trends and Monetisation of Data Goods”, *Intereconomics*, Springer, Vol. 54 No. 4, pp. 208–216.

Stocker, A., Kaiser, C. and Fellmann, M. (2017), “Quantified Vehicles”, *Business & Information Systems Engineering*, Springer, Vol. 59 No. 2, pp. 125–130.

Swan, M. (2013), “The quantified self: Fundamental disruption in big data science and biological discovery”, *Big Data*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, Vol. 1 No. 2, pp. 85–99.

Tucci, C.L., Afuah, A. and Viscusi, G. (2018), *Creating and Capturing Value Through Crowdsourcing*, Oxford University Press.

Viscusi, G. and Batini, C. (2016), “Information Production and Social Value for Public Policy: A Conceptual Modeling Perspective”, *Policy & Internet*, Vol. 8 No. 3, pp. 334–353.

Viscusi, G., Castelli, M. and Batini, C. (2014), “Assessing social value in open data initiatives: a framework”, *Future Internet*, Vol. 6 No. 3, pp. 498–517.

Viscusi, G., Rusu, A. and Florin, M.-V. (2020), “Public Strategies for Artificial Intelligence: Which Value Drivers?”, *Computer*, Vol. 53 No. 10, available at:https://doi.org/10.1109/MC.2020.2995517.