

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Aslam, Bilal; Karjaluoto, Heikki; Varmavuo, Eevi

Title: Data obstacles and privacy concerns in artificial intelligence initiatives

Year: 2022

Version: Published version

Copyright: © Authors, 2022

Rights: CC BY-NC-ND 4.0

Rights url: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Please cite the original version:

Aslam, B., Karjaluoto, H., & Varmavuo, E. (2022). Data obstacles and privacy concerns in artificial intelligence initiatives. In O. Niininen (Ed.), *Contemporary Issues in Digital Marketing* (pp. 130-138). Routledge. <https://doi.org/10.4324/9781003093909-16>

12 Data obstacles and privacy concerns in artificial intelligence initiatives

Bilal Aslam, Heikki Karjaluoto and Eevi Varmavuo

Introduction

Artificial Intelligence (AI) is considered a revolutionary technology in the marketing industry. Although it has existed as a technology and a field of study for decades, it has only recently shown rapid growth among different markets. AI has disrupted a variety of industries (Campbell *et al.*, 2020), and the growth of AI adoption and its positive effects on businesses are clear; estimations predict that AI will increase the global economy by 14% (equivalent to 15.7 trillion USD) by 2030 (see Future Reading: AI Sheibani, Cheung, and Messon, 2018). By investing in AI, firms may increase their return on investment significantly, which will further accelerate AI adoption.

One factor that influences AI is data, which are the raw materials that are required to run AI machinery. Data serve as a new source of idea generation for product development, customer service, shelf location, distribution, dynamic pricing, etc. (Erevelles, Fukawa, and Swayne, 2016). Good data are the foundation of AI modelling, and the appropriate data are needed to optimise projects (eMarketer, 2019 – see Further reading). During the past few years, the use of data sciences, which facilitate decision-making and the extraction of actionable insights and knowledge from large datasets in the marketing environment, has remarkably increased. Despite these advances, strategies for improving the management of data sciences in DM remain scarce (Saura, 2020). AI both lives and dies by data; problematic data can introduce risks that can be economically devastating for any company (Harrison *et al.*, 2019)

Therefore, we must explore whether companies are preparing themselves in terms of data collection and management to fully capitalise on AI technology. Specifically, businesses reportedly waste considerable time organising, cleaning and structuring the databases of their users and customers (Kelleher and Tierney, 2018). There are problems in data acquisition, data labelling and the improvement of existing data (Roh, Heo, and Whang, 2019). It is important to determine whether companies have organisational chassis to successfully support AI initiatives. A strong dimension of data that is often ignored is ethical issues related to data privacy. For example, the two tech giants Google and Facebook were both sued for submitting consumers' photos for biometric scanning (Solove and Schwartz, 2014). This development might affect data-driven AI initiatives, especially since the implementation of the General Data Protection Regulation (GDPR) in Europe. Data privacy can also create problems because customers are becoming more aware of the collection of their personal data, and it is raising concerns (Davenport *et al.*, 2020).

This chapter aims to determine whether firms have sound structures for collecting and organising data to fully harness AI technology. We used a special data privacy angle to monitor the industry's current outlook regarding data privacy issues. For this purpose, we conducted in-depth interviews with relevant AI and data industry experts in five countries. The results showed that companies are currently lacking in structures and systems for collecting and managing data; thus, their ability to harness AI technology is weak. It was also evident that data privacy issues are extremely important, and measures like the GDPR will help the industry cultivate a more ethical use of technology while not blocking overall industry growth. Our study will assist managers who want to capitalise on data based on AI technology.

Literature review

The concept of AI in its contemporary sense is not new. It was firstly initiated during the Dartmouth Summer Research Project on AI in the 1950s. However, AI appears new to those looking to repackage Big Data (Elish and Boyd, 2017). Marketers can now generate more accurate results for a variety of marketing intelligence tasks, including customer segmentation and profiling, product reputation management, pricing strategy, competitor analysis, promotional marketing analysis, recommender systems, location-based advertising and community dynamic analysis (Fana *et al.*, 2015). These results have important implications for forecasting purchases, and they can affect sales forecasting (Liu, Xiao, and Ding, 2016). AI and Big Data are not only changing existing marketing tasks but also supporting innovative marketing solutions. For example, Liu, Xiao, and Ding (2016) propose an automated and scalable garment recommender system using real-time in-store videos that can improve the experiences of garment shoppers and increase product sales.

Data are the foundation of AI-driven marketing; AI requires data to run, and most AI-powered solutions require existing datasets to run. If a project lacks the appropriate data, the results will be less than optimal (eMarketer, 2019 – see Further reading). AI makes data meaningful through cognitive computing because analysis of data by humans can be a time-consuming task; thus, the utilisation of AI techniques helps clarify data (Gupta *et al.*, 2018). More data are available today than ever before but companies cannot generate effective results if their data are compromised (Alshura, Zabadi, and Abughazaleh, 2018). Data offer greater insights into marketing performance than they did in the past, and these insights help marketers make successful decisions for optimising their marketing actions and improving their return on investment (Wedel and Kannan, 2016). While finding data used to be a challenge for marketers, they now face the challenge of transforming data into value. Data have attained a central role in marketing solutions. As data become larger, more complex and more inexplicable, the limited mental capacities of humans pose difficulties in deciphering and interpreting an unknown environment (Sammut and Sartawi, 2012). Data create a competitive advantage, and marketers are creating more personalised experiences through data (Wright *et al.*, 2019; Ozelik and Varnali, 2019; Jarek and Mazurek, 2019; Aguirre *et al.*, 2015). Companies should be using an infrastructure that can facilitate the adoption of AI, and this particular area of research needs further examination.

Data-based innovation and marketing can trigger consumers' privacy concerns from a contextual integrity perspective. Such concerns can, in turn, influence the future of these data-intensive fields. Firms must carefully evaluate their use of consumer data in

their innovation and marketing efforts (Bleier, Goldfarb, and Tucker, 2020; Davenport *et al.*, 2020). Addressing data ethics and politics is an integral task of data studies. Big Data and their meaning are socially constructed and influenced by evolving social, political and technological forces (Chen and Haase, 2020). Martin, Borah, and Palmatier (2017) state that growing efforts towards data collection and usage increase customers' concerns about their privacy. They might feel uncomfortable receiving personalised advertising and content when they realise how much of their data are being collected and analysed (Aguirre *et al.*, 2015). Per Martin and Murphy (2017), the more worried customers are about their data privacy, the more negative their responses are towards the brand. In addition, regulations, such as the GDPR, are considered revolutionary in terms of data privacy, but they create challenges for marketers (Kietzmann, Paschen, and Treen, 2018).

Research methodology

To reach a holistic understanding of what data issues challenge marketers, managers and other industry executives, semi-structured interviews were used for data collection in this research. The interviewed experts were working with AI technology and had knowledge of and/or experience in marketing. Given the complexity of the research problem and that the phenomenon under study may vary between the interviewees depending on their expertise, it was important to consider that new information could emerge during the interviews. This study used purposive sampling. A review of the interviewees' titles shows that they varied in their level of expertise. We conducted in depth-interviews (see Table 12.1 for details) with 14 relevant managers, CEOs, entrepreneurs and consultants.

We conducted telephonic interviews with the informants from Finland, the United Kingdom, the United States, Switzerland and Peru to ensure balanced and comprehensive results. Most of the interviews were conducted in Finland. The interviewees were selected because of their knowledge/practice/experience, particularly in the AI field and generally for their ability to reflect from marketing perspectives. The interviews, which lasted ~45–60 minutes, were recorded and transcribed.

Our data analysis was conducted via thematic analysis (the ethnographic imagination approach). Ethnography is a way to imagine 'social'. In other words, ethnography is a type of qualitative research that involves immersing yourself in a particular community or organisation to observe their behaviour and interactions up close. Per Watson (2011), organisational ethnography involves creating systematic generalisations about a topic that must be theoretically informed, informing and contributing to the broader body of knowledge that constitutes organisation and management studies. This approach enables theoretical rather than empirical generalisations. Sociological imagination allows individuals to rise above their everyday social context, making it possible to acquire the distance necessary for critical reflection and change (Mills, 1959).

The thematic analysis included five steps. (1) Braun and Clark (2006) stress that the researcher should familiarise themselves with the data. When transcribing the interviews for this study, strict attention to detail ensured that no information of value was overlooked. The interview recordings were firstly transcribed in writing, and filler words (e.g. so, um and like) were left out. Both the interviews and the transcriptions were conducted in English. (2) The second step included generating initial codes. Coding is defined as 'the process of assigning meaningful numerical values or names that reduce data from a

Table 12.1 Interview participant details

<i>Interviewee</i>	<i>Country</i>	<i>Title</i>	<i>Type</i>	<i>Duration</i>
I-1	Finland	Head of data	Face-to-face	28 min
I-2	Finland	Chief Growth Officer	Telephone	52 min
I-3	Finland	Lead Data Scientist	Telephone (with I4)	46 min
I-4	Finland	Business Lead AI	Telephone (with I3)	46 min
I-5	Finland	CEO	Telephone	37 min
I-6	Finland	CEO and Co-Founder	Telephone	49 min
I-7	Finland	Director, Analytics	Telephone	36 min
I-8	Finland	CEO	Telephone	24 min
I-9	Finland	Founder and Chairman	Telephone	38 min
I-10	Bulgaria	CEO and Co-Founder	Telephone	51 min
I-11	Switzerland	AI expert	Telephone	55 min
I-12	The United Kingdom	CEO	Telephone	39 min
I-13	The United States	Previous Digital Director/ Adjunct Professor	Telephone	1 h 2 min
I-14	Peru	CEO	Telephone	60 min

large amount of undifferentiated text' (Hair *et al.*, 2015, p. 302). Thus, coding helps the researcher focus on valuable key characteristics of the data. During this step, the entire dataset was read and coded with precision. (3) The third step involved searching for themes within the data (Braun and Clarke, 2006). Similar codes were put together in potential themes; the codes that demonstrated patterns throughout the dataset were considered themes. The themes were then colour coded to help analyse the data. Braun and Clarke (2006) suggest (4) reviewing the themes and (5) naming them. Here, the themes were reviewed against emerging similar themes to avoid duplicates, and they were named to provide a clear visual map of the data.

Findings

Figure 12.1 outlines the framework for factors affecting the successful implementation of AI and required actions.

The analysis consists of themes that emerged from the semi-structured interviews regarding data issues. Most of the interviewees stressed that the current data collection and management methods that are used in most organisations are insufficient.

The biggest issue with AI is the data quality and the storing of the data. And I see there are a lot of things to be done to get AI really – Because 80% of our time, when we work with customers and AI-related projects, it is to just start collecting the data, cleaning the data and managing the data, and only after that can we start running the AI.

(I-3, Lead Data Scientist)

I don't think they know what data is there. They never made sense of it. It's not been tuned, not been managed, I doubt. And there's no kind of data lineage. What is the purpose and why was it collected historically vs its purpose now? Is it labelled? Is it normalised?

(I-13, Digital Director and Adjacent Professor)

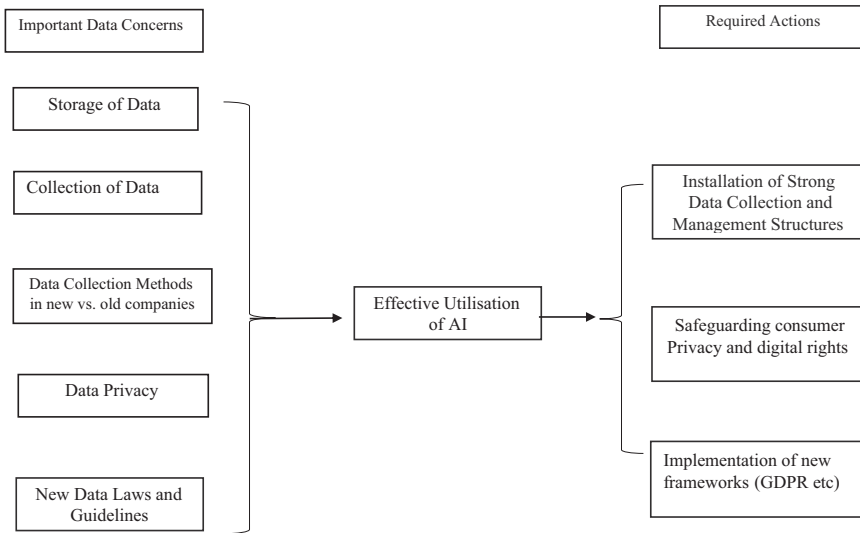


Figure 12.1 Result framework: factors affecting successful implementation of AI and required actions.

The findings also showed that, in some cases, companies are collecting data, but they have no idea how to use it.

I believe that companies have quite a lot of data, of which a big part is not used. Nobody cares, and they do not know what they have. So, I would start with what data is there and make conclusions about whether it is in the right format or not.

(I-5, CEO)

Building a large data warehouse was not considered the optimal approach for companies, as explained by one of the interviewees (I-6). Instead, the interviewees found starting with the business goals and the use case, followed by determining what data and tools were required as well as how to label the data the correct approach.

Davenport *et al.* (2020) stress the importance of ensuring that these processes do not hinder gathering relevant data insights from customers and innovating. This conclusion can also be drawn from the interviews; the interviewees agreed that privacy requirements and regulations guide the operations of companies. They considered the GDPR the most significant change in data privacy regulations in recent years and stressed its role in marketing practices today. The interviewees saw regulations as having a positive effect on the incorporation of AI into marketing practices.

I am 100% convinced that GDPR will affect it in a good way. Data privacy is a huge issue nowadays already, and it will be more so in the future. And now when we have GDPR in place, it will force the companies to concentrate on these privacy issues already in the beginning. It is not possible to if something is already up and running; it is very difficult to take into account certain privacy things. But if we have those rules in place already in the beginning, it will be a great thing.

(I-6, CEO and Co-Founder)

I think GDPR was the best thing that could happen in the data business . . . I would say that these regulations help create fair and sustainable AI business. Fair is kind of the same as being ethical but an easier concept.

(I-9, Founder and Chairman)

As seen earlier, the GDPR guides companies in the right direction regarding consumer data privacy from the outset, and it is a major factor when it comes to incorporating AI into marketing practices. It only has a negative impact when a lack of understanding is involved.

Another interviewee (I-1) explained that, thanks to the GDPR, consumers are now more aware of what data are collected about them, and they also understand their responsibility regarding the data they provide, which can be considered positive. However, customers also expect to gain value from giving out their data via better services. This connects to the findings regarding the expectations that consumers have from younger companies, especially when considering marketing activities that are based on collected data, such as personalisation. Interviewee I-1 also highlighted that compliance with the GDPR or other regulations allows companies to create better services and increase their overall brand image as well as the confidence of certain partners.

While the positive role of the GDPR was emphasised by the interviewees, few challenges were mentioned concerning related limitations in AI-driven marketing practises. One challenge described by an interviewee (I-4) was data collection and management processes.

You have to understand what data you are using and how you are using it and whether the usage of that data is purposeful for why it was created.

(I-4, Business Lead AI)

Another interviewee (I-11) agreed that the GDPR has complicated things for marketing practices. However, he emphasised the role of trust between the company and its customers correspondingly to the rising privacy concerns among consumers that were noted by Davenport *et al.* (2020) and Martin, Borah, and Palmatier (2017). Regarding the opinion presented earlier, Martin and Murphy (2017) also stress the role of providing customers some control over their data to build trust with the company and increase the likelihood of customers providing their data. Thus, building trust between a company and its customers is critical for receiving customer data while complying with certain regulations, such as the GDPR.

Discussion

This chapter aimed to determine whether companies are structurally prepared to adopt AI in terms of data collection and management. This step is important to ensure the optimisation of AI benefits. We sought to determine how privacy issues and new reforms, such as the GDPR, can affect a company's AI initiatives and processes. Our main findings indicated that companies might be neglecting the expertise required for sound data collection and management practices. This is consistent with Saura's (2020) claim that relevant measures for improving data management remain scarce. Henke *et al.* (2016– see Further reading) emphasise the importance of having proper data ecosystems in place to deliver successful AI campaigns (Chui, 2017 – see Further reading). In some cases, even companies that are

collecting large datasets are not utilising them in their favour. We also found that privacy should be at the centre of data policy when implementing any data initiatives. Martin, Borah, and Palmatier (2017) state that growing efforts in data collection and usage are increasing customers' privacy concerns. Firms must carefully evaluate their usage of consumer data in their innovation and marketing efforts (Bleier, Goldfarb, and Tucker, 2020; Davenport *et al.*, 2020). New initiatives, such as the GDPR, are helping industries safeguard consumer privacy without causing problems for AI initiatives. Figure 12.1 extracts the findings from this paper. Many data issues must be fixed if companies want to harness AI technology. Currently, most companies' data collection and management systems are inadequate; they are not striving to collect all possible forms of internal data. External data can be useful but only after internal data are fully understood. While data collection methods in older companies are more inadequate than those in newer companies, older companies can survive without data because they have long-standing, strong relationships with their customers, and the market recognises their products and/or services. However, to be more competitive, they must introduce new data collection and management policies. There is no simple answer regarding which data are the most effective in any campaign; instead, the desired results need to be identified before beginning a project to know which data will help achieve effective results.

There are certain limitations to this research. For example, the data collection was conducted via semi-structured expert interviews on a rather broad scale instead of focusing on a specific factor. Although the semi-structured interviews offered the possibility of finding new aspects, this thesis offers a general overview of the factors that make data appropriate for practising effective AI campaigns in marketing versus a detailed description.

Key lessons for future research

- Future research should seek an exact framework for the deployment of data management and collection structures as well as a general privacy framework model.
- A clear/new framework is required to explain how to establish an AI strategy and the how company's data should be managed to make the strategy work
- Exploration is needed into how data collections and management practices can be tailored in different organisational settings (e.g. large organisations vs. small organisations; old organisations vs. new organisations).

Disclaimer

The research presented in this chapter was remodified from the University of Jyväskylä Master's thesis '*Factors affecting the success of AI campaigns in marketing: data perspective (2020)*'. The copyright for this JYU thesis belongs to Varmavuo, Eevi as the Author. Research presented here has not been otherwise previously published.

Further reading

- Al Sheibani, S., Cheung, Y., and Messon, C. (2018). 'Artificial Intelligence adoption: AI-readiness at *on Information Systems*, Japan 2018. Available at: <https://aisel.aisnet.org/pacis2018/>
- Chui, M. (2017). *Artificial Intelligence: The Next Digital Frontier?* London: McKinsey and Company Global Institute. Available at: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world> (accessed 15 Oct 2020).

- Emarketer. (2019). *Getting Smart about Artificial Intelligence*. Available at: www.emarketer.com/content/getting-smart-about-artificial-intelligence
- Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B and Sethupathy, G. (2016). 'Big data's potential just keeps growing. Taking full advantage means companies must incorporate analytics into their strategic vision and use it to make better, faster decisions'. *McKinsey and Company*. Available at: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world> (accessed 15 Oct 2020).
- Jacques Bughin and James Manyika are directors of the McKinsey Global Institute, and Michael Chui is an MGI partner; Nicolaus Henke and Tamim Saleh are senior partners in McKinsey's London office, Bill Wiseman is a senior partner in the Taipei office, and Guru Sethupathy is a consultant in the Washington, DC, office.

References

- Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., and Wetzels, M. (2015). 'Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness'. *Journal of Retailing*, 91(1), pp. 34–49. <https://doi.org/10.1016/j.jretai.2014.09.005>
- Alshura, M. S., Zabadi, A., and Abughazaleh, M. (2018). 'Big Data in marketing arena. Big opportunity, big challenge, and research trends: An integrated view'. *Management and Economics Review*, 3(1), pp. 75–84. <https://doi.org/10.24818/mer/2018.06-06>
- Bleier, A., Goldfarb, A., and Tucker, C. (2020). 'Consumer privacy and the future of data-based innovation and marketing'. *International Journal of Research and Marketing* <https://doi.org/10.1016/j.ijresmar.2020.03.006>
- Braun, V., and Clarke, V. (2006). 'Using thematic analysis in psychology'. *Qualitative Research in Psychology*, 3(2), pp. 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Campbell, C., Sands, S., Ferraro, C., Tsao, H.-Y., and Mavrommatis, A. (2020). 'From data to action: How marketers can leverage AI'. *Business Horizons*, 63, pp. 227–243. <https://doi.org/10.1016/j.bushor.2019.12.002>
- Chen, W. and Haase, A. Q. (2020). 'Big Data Ethics and Politics: Toward New Understandings'. *Social Science Computer Review*, 38(1), pp. 3–9. <https://doi.org/10.1177/0894439318810734>
- Davenport, T., Guha, A., Grewal, D., and Bessgott, T. (2020). 'How Artificial Intelligence will change the future of marketing'. *Journal of the Academy of Marketing Science*, 48, pp. 24–42. <https://doi.org/10.1007/s11747-019-00696-0>
- Elish, M. C. and Boyd, D. (2018). 'Situating methods in the magic of Big Data and AI'. *Communication Monographs*, 85(1), pp. 57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- Erevelles, S., Fukawa, N., and Swayne, L. (2016). 'Big Data consumer analytics and the transformation of marketing'. *Journal of Business Research*, 69(2, February), pp. 897–904. <https://doi.org/10.1016/j.jbusres.2015.07.001>
- Fana, S., Raymond, Y. K., Laub, Zhaob, J. L. (2015). 'Demystifying Big Data analytics for business intelligence through the lens of marketing mix'. *Big Data Research*, 2(1, March), pp. 28–32. <https://doi.org/10.1016/j.bdr.2015.02.006>
- Gupta, S., Kar, A. K., Baabdullah, A., and Al-Khowaiter, W. A. A. (2018). 'Big Data with cognitive computing: A review for the future'. *International Journal of Information Management*, 42(October), pp. 78–89. <https://doi.org/10.1016/j.ijinfomgt.2018.06.005>
- Hair, J. F., Wolfinbarger, M. E., Money, A. H., Samouel, P., and Page, M. J. (2015). *The Essentials of Business Research Methods* (3rd ed.). London: Routledge. <https://doi.org/10.4324/9781315716862>
- Harrison, T. M., Reyes, L. F. L., Pardo, T. A., DePaula, N., Najafabadi, M. M., and Palmer, J. M. (2019). 'The data firehose and AI in government: Why data management is a key to value and ethics'. In *Proceedings of the 20th Annual International Conference on Digital Government Research* (June). pp. 171–176. New York: Association for Computing Machinery. <https://doi.org/10.1145/3325112.3325245>
- Jarek, K., and Mazurek, G. (2019). 'Marketing and Artificial Intelligence'. *Central European Business Review, Prague University of Economics and Business*, 2, pp. 46–55. DOI: 10.18267/j.cebr.213

- Kelleher, J. D., and Tierney, B. (2018). *Data Science*. Cambridge, MA: MIT Press.
- Kietzmann, J., Paschen, J., and Treen, E. (2018). 'Artificial Intelligence in advertising. How marketers can leverage Artificial Intelligence along the consumer journey'. *Journal of Advertising Research*, 58(3), pp. 263–267. <https://doi.org/10.2501/JAR-2018-035>
- Liu, S., Xiao, L., and Ding, M. (2016). 'Video-Based Automated Recommender (VAR) system for garments'. *Marketing Science*, 35(3), pp. 484–510. <https://doi.org/10.1287/mksc.2016.0984>
- Martin, K. D., Borah, A., and Palmatier, R. W. (2017). 'Data privacy: Effects on customer and firm performance'. *Journal of Marketing*, 81(1), pp. 36–58. <https://doi.org/10.1509%2Fjm.15.0497>
- Martin, K. D., and Murphy, P. E. (2017). 'The role of data privacy in marketing'. *Journal of the Academy of Marketing Science*, 45(2), pp. 135–155. <https://doi.org/10.1007/s11747-016-0495-4>
- Mills, C. W. (1959). *The Sociological Imagination*. London: Oxford University Press.
- Ozcelik, A. B., and Varnali, K. (2019). 'Effectiveness of online behavioral targeting: A psychological perspective'. *Electronic Commerce Research and Applications*, 33(January–February), pp. 1–11. <https://doi.org/10.1016/j.elerap.2018.11.006>
- Roh, Y., Heo, G., and Whang, S. E. (2019). 'A survey on data collection for machine learning: A Big Data–AI integration perspective'. *IEEE Transactions on Knowledge and Data Engineering*, pp. 1041–4347. <https://doi.org/10.1109/TKDE.2019.2946162>
- Sammur, G., and Sartawi, M. (2012). 'Perspective-taking and the attribution of ignorance'. *Journal for the Theory of Social Behaviour*, 42(2), pp. 181–200. <https://doi.org/10.1111/j.1468-5914.2011.00485.x>
- Saura, J. R. (2020). 'Using data sciences in digital marketing: Framework, methods, and performance metrics'. *Journal of Innovation and Knowledge* <https://doi.org/10.1016/j.jik.2020.08.001>
- Solove, D. J., and Schwartz, P. (2014). *Information Privacy Law*. New York: Wolters Kluwer Law and Business.
- Watson, T. J. (2011). 'Ethnography, reality and truth: The vital needs for studies of 'how things work' in organizations and management'. *Journal of Management Studies*, 48(1), pp. 202–216. <https://doi.org/10.1111/j.1467-6486.2010.00979.x>
- Wright, L. T., Robin, R., Stone, M., and Aravopoulou, E. (2019). 'Adoption of Big Data technology for innovative B2B marketing'. *Journal of Business-to-Business Marketing*, 26(3–4), pp. 281–293. <https://doi.org/10.1080/1051712X.2019.1611082>