ENTERING THE ERA OF DECENTRALIZATION: AN INVESTIGATION OF NEW TRENDS IN DATA MANAGEMENT

Erik Karger *, Marvin Jagals *

* Chair of Information Systems and Strategic IT Management, University of Duisburg-Essen, Essen, Germany

How to cite: Karger, E., & Jagals, M. (2023). Entering the era of decentralization: An investigation of new trends in data management. In E. Karger & A. Kostyuk (Eds.), *Corporate governance: An interdisciplinary outlook* (pp. 79–84). Virtus Interpress. https://doi.org/10.22495/cgaiop17

Received: 23.09.2022
Accepted: 03.10.2022
Keywords: Data Governance, Data Management, Data Quality, Governance, Blockchain
JEL Classification: L86, M10, M15
DOI: 10.22495/cgaiop17

Abstract

Data build the foundation for many modern technologies like artificial intelligence. Data are of increasing relevance for companies. Nowadays, however, the management of data is no longer only an intraorganizational phenomenon. To get access to more data, companies are increasingly willing to share their own data with other organizations. Furthermore, decentralized technologies like blockchain are supposed to have potentials for the management, storage, or sharing of data. This study aims to investigate three current trends of decentralization within data management: Interorganizational data governance, data ecosystems, and blockchain-based data management. We provide an initial overview of these developments and present avenues for future research.

1. INTRODUCTION

A lot of different sources increasingly generate data. According to a report published by the World Economic Forum in 2019, it is estimated that 200 exabytes of data will be produced every day in 2025

(Desjardins, 2019). One major reason for this boost in data generation speed in recent years is the emergence of smart devices and connected networks, resulting in a rising spread of the IoT (Khan et al., 2016; Silva et al., 2019, p. 14; Silva et al., 2018). For companies, data can have significant advantages when analyzed, leading to the metaphor of data being the new oil (Hartmann et al., 2016, p. 1382; Wiener et al., 2020).

In the last decade, technologies to analyze data, like artificial intelligence (AI) and big data analytics, experienced exponential growth (Günther et al., 2017, p. 191; Wiener et al., 2020, p. 66). Some researchers believe that big data "offer unprecedented opportunities" (Wiener et al., 2020, p. 66) for companies and that it is "one of the most significant technology disruptions for businesses since the meteoric rise of the Internet and the digital economy" (Chen et al., 2017, p. 19). By supporting, for example, the optimization of supply chain management or setting optimal prices for products and services, big data furthermore is believed to help achieve more efficient and effective operations (Chen et al., 2012; Davenport et al., 2012; McAfee & Brynjolfsson, 2012).

The increasing relevance of data for companies has also led to more trends that exhibit decentralization characteristics. To benefit from multiple sources of data, and to enable the exchange of data between several companies, organizations are increasingly willing to soften organizational boundaries. This makes it necessary to govern data not only within one company but across different organizations, leading to developments such as interorganizational data governance (Jagals & Karger, 2021) and data ecosystems. Furthermore, the blockchain technology is increasingly investigated in relation to data, for example, for the management of AI data (Karger et al., 2021) or for enhancing inter-organizational data governance (Jagals et al., 2021). We aim to shed light on these recent data management developments characterized by decentralization. Hereby, we aim to provide interested researchers and practitioners with an overview of these new trends and outline avenues for future research. Hence, we address the following research question:

RQ: What are relevant and recent decentralized developments in data management?

In total, this study deals with three decentralized developments of data. First, the potential of blockchain for data management will be investigated. Furthermore, data ecosystems will be presented as a recent manifestation of interorganizational networks and ecosystems. Finally, we introduce interorganizational data governance.

2. FOUNDATIONS — DATA IN COMPANIES

Companies have already tried to use and analyze the available data for a long time to gain benefits, insights, and competitive advantages.

In the beginning, data were first utilized for automated data processing for specialized operations in companies. Thus, data were seen from a functional viewpoint, with data management as a subset of data administration on certain database systems (Aiken et al., 2013). The second business data management phase can be dated to the late 1980s and 1990s. In these years, integrated information systems were introduced. Furthermore, data was progressively exchanged across end-to-end processes rather than being confined to discrete company activities (Grover & Teng, 1991). In the 2010s, internal and external data were increasingly used for digital business models or data-driven services (Buhl et al., 2013; Lycett, 2013; Wixom & Ross, 2017). Due to the increasing relevance of data, technological and organizational skills for handling and using the data became increasingly relevant. For organizations, it is not only important to be able to gather and store increasing volumes of diverse data. Also, the ability to utilize modern and advanced analytics technologies is nowadays crucial for companies (Abbasi et al., 2016; Müller et al., 2016).

3. BLOCKCHAIN FOR DATA MANAGEMENT

To the best of our knowledge, no common definition for the term "blockchain" exists because blockchain is nowadays a complex technology with many different facets and manifestations. In this article, we follow the definition proposed by Treiblmaier (2018) and define blockchain as "[…] a digital, decentralized and distributed ledger in which transactions are logged and added in chronological order to create permanent and tamper-proof records" (p. 547). This definition already contains the essential criteria and characteristics from which the blockchain obtains its uniqueness. First, a blockchain network is decentralized and distributed. No central authority or instance is required to maintain or run the blockchain network (Nakamoto, 2008; Raval, 2016). Furthermore, the blockchain's blocks are cryptographically linked to each other with hash functions. When data in one of the blocks are changed, this is immediately recognized when these blocks are distributed within the network. This fact leads to resistance to data modification (Narayanan et al., 2016). A blockchain's basic units are transactions stored together in one block. These transactions can be linked to payment processes or simply contain information (Treiblmaier, 2018, p. 546). According to Iansiti and Lakhani (2017, p. 125):

- *Distributed database*: The whole database and its complete history are available for each blockchain party. No single party controls the data and information, and all parties can verify the transactions without using an intermediary.
- *Peer-to-peer transmission*: The network peers can communicate without a central node. This includes the storing and forwarding of information to all other peers.

• *Transparency with pseudonymity*: Transactions and their associated values are visible to everyone within a blockchain network. The single nodes on a blockchain have an alphanumeric address as a clear identifier, and transactions occur between these blockchain addresses.

• *Irreversibility of records*: Once a transaction is part of the blockchain, it cannot be altered or changed, because a transaction is linked to each transaction carried out beforehand. Various approaches and algorithms ensure transactions' permanence and correct order.

• *Computational logic*: The blockchain's computational logic makes it possible to program on the blockchain. Users can set up algorithms and rules that automatically trigger transactions between nodes.

There are different ideas and proposals on how blockchain might be used in data and data governance. For example, blockchain might serve as a foundation for interorganizational data governance (Jagals et al., 2021). Apart from data governance, blockchain is also discussed for the management or storage of AI data (Karger et al., 2021). The potential advantages and benefits of using blockchain for AI data management are long. One problem is that data can be easily hacked or tampered with when stored centralized (Salah et al., 2019). As it allows for distributed storage, the decentralized blockchain might be an advantage against this threat (Karger et al., 2021). Furthermore, by using smart contracts that enable an access control that allows making the data accessible to authorized users only, blockchain can help to ensure the data owner's privacy (Johnson et al., 2019; Karger et al., 2021; Passerat-Palmbach et al., 2019).

REFERENCES

1. Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems, 17*(2), 1–32. https://doi.org/10.17705/1jais.00423
2. Aiken, P., Gillenson, M., Zhang, X., & Rafner, D. (2013). Data management and data administration: Assessing 25 years of practice. In K. Siau (Ed.), *Innovations in database design, web applications, and information systems management* (pp. 289–309). IGI Global. https://doi.org/10.4018/978-1-4666-2044-5.ch011
3. Buhl, H. U., Röglinger, M., Moser, F., & Heidemann, J. (2013). Big data. *Business & Information Systems Engineering, 5*, 65–69. https://doi.org/10.1007/s12599-013-0249-5
4. Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188. https://doi.org/10.2307/41703503
5. Chen, H.-M., Schutz, R., Kazman, R., & Matthes, F. (2017). How Lufthansa capitalized on big data for business model renovation. *MIS Quarterly Executive, 16*(1), 19–34. https://aisel.aisnet.org/misqe/vol16/iss1/4/

6. Davenport, T. H., Barth, P., & Bean, R. (2012). How 'big data' is different. *MIT Sloan Management Review, 54*(1), 43–46. https://sloanreview.mit.edu/article/how-big-data-is-different/
7. Desjardins, J. (2019, April 17). How much data is generated each day? *World Economic Forum*. https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/
8. Grover, V., & Teng, J. T. C. (1991). How effective is data resource management? Reassessing strategic objectives. *Information System Management, 8*(3), 16–20. https://doi.org/10.1080/07399019108964992
9. Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems, 26*(3), 191–209. https://doi.org/10.1016/j.jsis.2017.07.003
10. Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data: A taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management, 36*(10), 1382–1406. https://doi.org/10.1108/IJOPM-02-2014-0098
11. Iansiti, M., & Lakhani, K. R. (2017). The truth about blockchain. *Harvard Business Review, 95*(1), 118–127. https://hbr.org/webinar/2017/02/the-truth-about-blockchain
12. Jagals, M., & Karger, E. (2021). Inter-organizational data governance: A literature review. *ECIS 2021 Research Papers, 57*. https://aisel.aisnet.org/ecis2021_rp/57
13. Jagals, M., Karger, E., Ahlemann, F., & Brée, T. (2021). Enhancing inter-organizational data governance via blockchain-shaping scopes and research avenues. In *Proceedings of the International Conference on Information Systems (ICIS 2021), 7*. https://aisel.aisnet.org/icis2021/governance/governance/7
14. Johnson, M., Jones, M., Shervey, M., Dudley, J. T., & Zimmerman, N. (2019). Building a secure biomedical data sharing decentralized app (DApp): Tutorial. *Journal of Medical Internet Research, 21*(10), Article e13601. https://doi.org/10.2196/13601
15. Karger, E., Jagals, M., & Ahlemann, F. (2021). Blockchain for AI data — State of the art and open research. In *Proceedings of the International Conference on Information Systems (ICIS 2021), 2*. https://aisel.aisnet.org/icis2021/fintech/fintech/2
16. Khan, M., Silva, B. N., & Han, K. (2016). Internet of things based energy aware smart home control system. *IEEE Access, 4*, 7556–7566. https://doi.org/10.1109/ACCESS.2016.2621752
17. Lycett, M. (2013). 'Datafication': Making sense of (big) data in a complex world. *European Journal of Information Systems, 22*(4), 281–386. https://doi.org/10.1057/ejis.2013.10
18. McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60–68. https://hbr.org/2012/10/big-data-the-management-revolution
19. Müller, O., Junglas, I., vom Brocke, J., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal of Information Systems, 25*(4), 289–302. https://doi.org/10.1057/ejis.2016.2

20. Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system*. https://bitcoin.org/bitcoin.pdf
21. Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). *Bitcoin and cryptocurrency technologies: A comprehensive introduction*. Princeton University Press.
22. Passerat-Palmbach, J., Farnan, T., Miller, R., Gross, M. S., Flannery, H. L., & Gleim, B. (2019). *A blockchain-orchestrated federated learning architecture for healthcare consortia*. arXiv. https://arxiv.org/pdf/1910.12603
23. Raval, S. (2016). *Decentralized applications: Harnessing bitcoin's blockchain technology* (1st ed.). O'Reilly Media. https://dl.acm.org/doi/10.5555/3074145
24. Salah, K., Rehman, M. H. U., Nizamuddin, N., & Al-Fuqaha, A. (2019). Blockchain for AI: Review and open research challenges. *IEEE Access, 7*, 10127–10149. https://doi.org/10.1109/ACCESS.2018.2890507
25. Silva, B. N., Diyan, M., & Han, K. (2019). Big data analytics. In M. Khan, B. Jan, & H. Farman (Eds.), *Deep learning: Convergence to big data analytics* (pp. 13–30). Springer. https://doi.org/10.1007/978-981-13-3459-7_2
26. Silva, B. N., Khan, M., & Han, K. (2018). Internet of things: A comprehensive review of enabling technologies, architecture, and challenges. *IETE Technical Review, 35*(2), 205–220. https://doi.org/10.1080/02564602.2016.1276416
27. Treiblmaier, H. (2018). The impact of the blockchain on the supply chain: A theory-based research framework and a call for action. *Supply Chain Management, 23*(6), 545–559. https://doi.org/10.1108/SCM-01-2018-0029
28. Wiener, M., Saunders, C., & Marabelli, M. (2020). Big-data business models: A critical literature review and multiperspective research framework. *Journal of Information Technology, 35*(1), 66–91. https://doi.org/10.1177/0268396219896811
29. Wixom, B. H., & Ross, J. W. (2017). How to monetize your data. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/how-to-monetize-your-data/