

Big Data Discourses: Introduction to the Special Section

Introduction

CHARLOTTE KNORR¹
Leipzig University, Germany
LMU Munich, Germany

CHRISTIAN PENTZOLD
Leipzig University, Germany

Big data discourses are integral to how we come to understand and engage with datafication. This IJoC Special Section explores the semantics of big data in contemporary society. By weaving together three central themes—big data’s materiality, the role of emerging technologies, and the sensemaking around datafication—it interrogates how communication, imagination, and social practices shape and are shaped by the expanding landscape of data-driven innovation. The contributions offer interdisciplinary perspectives on the semantics, imaginaries, and social implications of big data, drawing on empirical research and critical theory to illuminate the evolving relationship between technology, discourse, and meaning-making in the digital age.

Keywords: big data, datafication, sensemaking, discourse, digital society

Starting from the observation that data and its technologies have become an integral part of everyday life, this Special Section brings together contributions that take issue with public deliberation and imaginaries around datafication. The relationship between technology and meaning-making is neither linear nor unidirectional. As Feenberg (2002) reminds us, “while social institutions adapt to technological development, the process of adaptation is reciprocal, and technology changes in response to the conditions in which it finds itself as much as it influences them” (p. 143). In that respect, the Special Section foregrounds the mutual constitution of technological artifacts and social realities.

Charlotte Knorr: charlotte.knorr@ifkw.lmu.de
Christian Pentzold: christian.pentzold@uni-leipzig.de
Date submitted: 2025-11-20

¹ The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant 447465824/PE2436/3-1.

Copyright © 2026 (Charlotte Knorr and Christian Pentzold). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <https://ijoc.org>.

The turn to the discursive layer of big data is important, we argue, and does not mean to detract us from its concrete technological force. Quite the contrary, the Special Section orients our attention to discourses whose programs of thought actively shape the social constitution of big data and translate into distinctive practices, organizational forms, policies, and social network institutions (Ellison & boyd, 2013). Also, big data invite us to look at what datafication is or should be for a variety of publics and speakers and how they discuss, criticize, and envision the collection and use of data at different places, speaking from different situations, and at different times. In that vein, the contributions in this Special Section are grounded in three central themes: big data's materiality, the role of emerging technologies, and the sensemaking around datafication. These three themes help us to critically engage with how datafication's social imaginations are shaped and constructed—namely, what data are, what data become, and how data are understood. Next to dominant positions and views on big data, the contributions also point us to alternative readings that challenge established narratives and practices that shape our relation to big data.

Take, for example, Emma Kaylee Graves-Sandriman's article, where she presents an analytical model to examine media communication on emerging technologies. Graves-Sandriman demonstrates its utility with two case studies on the reporting of the British media on big data and on generative artificial intelligence (Gen-AI). Her proposed model, Frame Categories for Emerging Technologies (FCET), comprises four categories: conceptualization, novelty, user experience, and evaluation. They afford the inductive identification of media frames and the comparison between different studies and technologies.

The Special Section further includes a contribution by Maria Cristina Paganoni and Gaston Becerra that looks at the discursive evolution of the data debate in the news media from 2019 to 2024, with a focus on the transition from big data to artificial intelligence (AI). Here, the authors consider power relations and ethical issues that go beyond data protection, such as data accessibility, job losses, and social justice. The analysis is based on two extensive news corpora in English and Spanish and combines quantitative methods such as topic modeling with qualitative discourse analysis.

With a focus on AI, Jascha Bareis describes how the release of chatbots has sparked social euphoria. He highlights the actors, dynamics, and motives that ignited and spread the hype, and shows how different social spheres interact to create hype as a powerful social phenomenon. In his analysis, the perspective on emerging technologies and actors in public discourse is interlinked—the public hype involves both material innovations (data, models) and sensemaking efforts alike.

Andreas Hepp puts forth the idea of "curating AI into being," with a focus on journalism and technological performances. In his contribution, he analyzes the role of the Hacks/Hackers community as a pioneering network and shows how it has transformed traditional journalistic practices in the wake of digitalization. The focus is on the way Hacks/Hackers shape and reinforce the idea of AI as the future of journalism through curatorial processes. Through this lens, emerging data technologies are understood as forces of social transformation that have a performative impact on journalistic and technological practices.

A further central concern of this Special Section is the sensemaking around datafication in society and, in that regard, the significance attributed to it. Investigating “data feelings” for instance, Ash Watson and Deborah Lupton examine the diverse effects of data streams on everyday life. Their study follows the concept of “data feelings”—the affective and physical connections that people form with their personal data and the technologies that generate it, including how they perceive their data, the feelings tied to it, and the everyday activities through which they engage with datafication. That way, their analysis deciphers the materiality, meaning, and multisensory impressions of personal data for the participants and thus also provides insights into their speculative ideas and sensemaking about data. In addition to the affective-physical dimension, data feelings are shaped by new technologies to collect and generate data (e.g., wearables, apps, and platforms).

With respect to cinematic aesthetics, Magdalena Krysztoforska and Oliver Kenny examine how cinema contributes to the “construction of visions of data,” particularly through the cinematic representation of scale. They analyze key scenes from films such as *Moneyball*, *Minority Report*, *Snowden*, *Anon*, and *Heart of Stone*, all of which have played and continue to play a prominent role in the data debate. In these cultural products, the materiality of big data (scaling) is linked to cinematic sensemaking on big data (imaginaries). The authors argue that the cinematic aesthetic, which they describe as an “aesthetic of boundless insight,” has a politically performative effect by presenting big data as an all-encompassing, context-free source of knowledge that is itself unlimited because it has no predetermined agenda.

Studying the discursive dimension of the sensemaking around big data, there is not one universal dataist imaginary, but a variety of different and also conflicting data imaginaries championed by Silicon Valley conglomerates, intelligence agencies, data activists, and others (Lehtiniemi & Ruckenstein, 2019; van Dijck, 2014). In order to challenge the objective facticity and abstract neutrality often ascribed to data, all contributors to this Section draw on critical data studies to argue that data presuppose interpretation. As Gitelman and Jackson (2013) put it, “data need to be imagined as data to exist and to function as such, and the imagination of data entails an interpretative base” (p. 3).

These sensibilities also resonate with Jun Yu’s contribution on how end-users of social media platforms make sense of, and respond to, the data processes of those platforms in everyday situations. Yu’s article, “A New Source of the Self? A Critical View on the Domestication of Data,” critically examines how data functions not only as a tool but also as an epistemic and moral infrastructure that profoundly influences users’ self-image, social practices, and moral thinking. Also highlighting moral discourses and the treatment of sensitive data, Alison B. Powell argues in the contribution “Deceptive Stories of Scale” that the digitization of healthcare promises efficiency gains and cost savings but also leads to tensions and challenges. Powell’s case study of Babylon Health, a start-up that introduced software and AI for rapid scaling (“blitzscaling”) in the UK’s National Health Service (NHS), illustrates how technological discourses and practices can create friction between time, knowledge, and ethics. It underscores that assuming technology to be inherently efficient and fast obscures normative aspects such as justice and epistemic fairness, which can have problematic consequences.

Taking a longitudinal approach, Preeti Raghunath examines the development of the Digital Public Infrastructure (DPI) in India over nearly two decades, beginning with digital identification for low-income families in 2006 and ending with the official recognition of the term DPI at the 2023 G20 summit. Hereby, Raghunath uses the theoretical framework of the Deliberative Policy Ecology Approach to examine how data is discussed and conceptualized in the context of DPI nationally and internationally, and concludes by emphasizing the need to strengthen the deliberative nature of DPI policy to promote a sustainable and collaboratively supported data future. Both contributions, Powell's and Raghunath's, reflect the diversity of actors and interests involved in the shaping of big data across years.

Victor Kuansong Zhuang and Gerard Goggin also direct their attention to the issue of diversity. In their article, "Toward Disability Data Justice: A Critical Discussion of Disability and Big Data Discourses," the authors discuss the significance of the increasing collection of disability data in national censuses, as exemplified by the first instance in Singapore in 2020. They argue that this development marks a crucial step toward enhancing inclusivity. The analysis of data facilitates the development of more effective service designs for specific population groups. However, it also imposes heightened demands for greater levels of detail and granularity in data collection and analysis.

Turning to media discourses again, Charlotte Knorr, Andreas Niekler, and Christian Pentzold reconstruct the media framing of big data in User Generated Content (Twitter/X, Facebook, Reddit) over the course of 10 years, 2011–2020. It is in this respect that big data give rise to its own mythology, that is, "the widespread belief that large data sets offer a higher form of intelligence and knowledge . . . with the aura of truth, objectivity, and accuracy" (boyd & Crawford, 2012, p. 663). The article builds on this line of scholarship and looks at the "interpretative work" (Bowker, 2013, p. 170) involved in making sense of data and the sociotechnological assemblage that shapes its production and understanding (Kitchin, 2014).

All contributions form an intervention aimed at fostering dialogue and debate about the discourses surrounding datafication today, thereby encouraging alternative and critical reflections on the trajectories of data-driven futures. They provide empirically rich conceptualizations and examinations of imaginaries around datafication and cognate technologies. The contributors and their respective studies cover a global range of countries and use cases, from Argentina to Italy, Germany, Spain, the United Kingdom, the United States, Australia, Singapore, and South Africa. Taken together, the Special Section turns its attention to the discourses and imaginaries that actively shape the social constitution of big data and translate into practices, organizational forms, policies, and institutions. The rise of AI has further intensified these dynamics, as AI systems increasingly rely on big data for training and inference. As such, the discursive dimension of datafication is still pressing because the ideas, affects, visions, and beliefs encapsulated in a discourse have tangible effects on the course of technological development, policymaking, and public expenditure. Approaching datafication through discourse means to engage with the eminent reality-making power of communication, deliberation, and imagination. This foregrounds the materiality, emerging technologies, and sensemaking. They not only mark three analytical perspectives, but in their interaction, they make datafication a socially relevant phenomenon and problem: Big data's materiality prompts discourse and is shaped by it, emerging technologies afford innovations and transformation in all sorts of fields, whilst cultural sensemaking renders them meaningful and politically

effective. Thus, the contributions of this Special Section do not merely interrogate the status quo of big data. Rather, the discourses they study commonly also involve prospective ambitions and normative stances about potential, desirable, or unwanted innovations.

References

- Bowker, G. (2013) Data flakes. In L. Gitelman (Ed.), *"Raw data" is an oxymoron* (pp. 167–172). Cambridge, MA: MIT Press.
- boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Ellison, N. B., & boyd, d. (2013). Sociality through social network sites. In W. H. Dutton (Ed.), *The Oxford Handbook of Internet studies* (pp. 151–172). Oxford, UK: Oxford University Press.
- Feenberg, A. (2002). *Transforming technology: A critical theory revisited*. New York, NY: Oxford University Press.
- Gitelman, L., & Jackson, V. (2013). Introduction. In L. Gitelman (Ed.), *"Raw data" is an oxymoron* (pp. 1–14). Cambridge, MA: MIT Press.
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. London, UK: SAGE Publications.
- Lehtiniemi, T., & Ruckenstein, M. (2019). The social imaginaries of data activism. *Big Data & Society*, 6(1). <https://doi.org/10.1177/2053951718821146>
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208. <https://doi.org/10.24908/ss.v12i2.4776>