

Data Publics: An Empirical and Comparative Approach to Datafication

FLORENCE MILLERAND¹

MÉLANIE MILLETTE

ALEXANDRE COUTANT

Université du Québec à Montréal, Canada

GUILLAUME LATZKO-TOTH

Université Laval, Canada

This article presents an empirical and comparative framework for understanding datafication through the lens of data publics. We define data publics as social groups that engage, or might be expected to engage, with data generated about, by, or for them. Through ethnographic case studies conducted in various contexts, including open data, health data, and data activism, this article explores how data initiatives are developed in relation to imagined publics, which ultimately confront actual publics. This study introduces an analytical framework that integrates diverse tensions between conceptions of the public (passive versus agentive, ideal versus empirical, cognitive versus affective) to facilitate a nuanced analysis of data publics. The article emphasizes the complexity inherent in these publics and the ways developers of data initiatives relate to them, and highlights the necessity of examining underlying power dynamics to gain a more in-depth understanding of publics' agency within the broader landscape of datafication.

Keywords: datafication, data publics, ethnographic case study, imagined public, empirical public, open data, health platform, data activism

Florence Millerand: millerand.florence@uqam.ca

Mélanie Millette: millette.melanie@uqam.ca

Alexandre Coutant: coutant.alexandre@uqam.ca

Guillaume Latzko-Toth: guillaume.latzko-toth@com.ulaval.ca

Date submitted: 2025-01-22

¹ This article is based on a research project funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) (Insight Grant No. 435-2018-1021). Both the research and this article's conception and production have been a fully collective effort by the coauthors. We particularly wish to thank Lucie Delias, Maxime Harvey, Sarah Meunier and Nadia Seraiocco for assisting us with the field work. We also thank the two anonymous reviewers for their thorough and constructive feedback on earlier versions of this paper.

Copyright © 2026 (Florence Millerand, Mélanie Millette, Alexandre Coutant, and Guillaume Latzko-Toth). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <https://ijoc.org>.

The phenomenon of datafication has generated considerable interest in the social sciences and humanities over the past decade. Coined by Mayer-Schönberger and Cukier in 2013, the term initially described a novel form of economic value production stemming from the collection, processing, and analysis of vast amounts of data (big data), particularly by digital platforms. Since then, the definition of datafication has expanded to include the increasingly pivotal role that data play in everyday life (Burgess et al., 2022; Flensburg & Lomborg, 2021).

Datafication has been theorized as a fundamental aspect of surveillance capitalism (Dijck, 2014; Zuboff, 2019) and as a contemporary form of colonialism, characterized by the extraction of data that adversely affects the populations involved (Couldry & Mejias, 2019). Some scholars contend that data can empower social actors who possess the means to interpret them effectively (Lycett, 2013; Powell, 2014). Conversely, others point to a disconnect between the potential applications of data and their actual utilization by the communities they impact (Andrejevic, 2014; Ruppert et al., 2017). In their literature review, Flensburg and Lomborg (2021) noted that early research on the datafication of society was still largely theoretical and not well-grounded in empirical evidence. Since then, several studies have been conducted to better understand how datafication processes shape social life, including everyday life (e.g., Pybus et al., 2025). Some of these studies have offered new perspectives for studying the publics affected by datafication (e.g., Millerand et al., 2025; Møller Hartley, Bengtsson, et al., 2023; Møller Hartley, Sørensen, & Mathieu, 2023). This article aims to contribute to this body of work.

Datafication is often associated with large-scale projects and infrastructure, such as corporate digital platforms, yet some research has shed light on medium- and small-scale initiatives, municipal open data portals, university dashboards tracking academic performance, open science projects, and local activist efforts to document public issues (e.g., Gabrys et al., 2016). This article focuses on such data initiatives and examines them through the lens of data publics, both the imagined publics targeted by these initiatives and the empirical publics observed in practice.

In this comparative analysis, we examine three ethnographic case studies that explore different data types and concern distinct public demographics: (1) the City of Montreal's Open Data Portal (ODP), a municipal open data initiative; (2) Pulsar, a Quebec City area-focused platform dedicated to managing health research data; and (3) Data for Good—Montreal (DFG), a collaborative network of volunteer data scientists working alongside community organizations. Each of these cases encounters unique challenges pertinent to their intended publics—city residents, health system users, and community organizations—particularly in promoting engagement with the data provided. These case studies are situated within the broader context of a research project investigating the phenomenon of datafication. Our comparative analysis allows us to establish a conceptual framework that enriches the empirical investigation of datafication in diverse contexts. The primary contribution of this article lies in the development of a “theory/methods package” (Clarke & Star, 2008, p. 113) designed to facilitate the empirical study of data publics.

More specifically, we are interested in how developers of data initiatives interact with publics, both imagined and actual. We propose an analytical framework combining different conceptions of the public divided into three axes (passive versus agentive, ideal versus empirical, cognitive versus affective) to analyze and compare how developers conceive of publics. Based on our case studies, we identify three

central dynamics—conciliation, exclusion, and embodiment—that describe how different understandings of publics shape the design and evolution of data initiatives. We argue that this analytical framework allows for the empirical examination of data publics and facilitates their comparison across different contexts while offering a better understanding of the logic behind publics' engagement or disengagement with data.

The article opens with a summary of the literature on publics, supplemented by recent work on digital publics, and then presents a three-axis analytical framework for the empirical study of data publics. The next section outlines our ethnographic methodology. The fourth section presents the results, detailing how each case intersects with the analytical axes. We show how different conceptions of the public are negotiated within each initiative. The final section discusses the added value of the analytical framework on data publics and its relevance for empirical studies of datafication.

Theoretical Approaches to Publics: A Synthesis and Framework for Analysis

From Media Publics to Data Publics

The concept of "public" eludes a strict definition, as it is intrinsically linked to the object around which it coalesces. This ambiguity is precisely what makes the concept analytically valuable (Esquenazi, 2009), particularly for interpreting empirical findings. The notion of the public emerged in early sociology, nearly simultaneously in France and the United States, as an analytical category developed in contrast to the idea of the crowd by Gabriel Tarde (1901/1989) and Robert Park (1904/1972). It is intertwined with the advent of mass media—especially the mass circulation press—and signifies a community of minds united through the mediation of news media. Subsequently, John Dewey (1927) and Walter Lippmann (1925) redefined this notion, emphasizing a more explicitly political rather than merely media-centric ontology. In their writings, the public is depicted as a collective of citizens embedded in modern societies where media plays a crucial role. However, their conceptions differ significantly: Lippmann (1925) views the public as a homogeneous and passive mass, largely acted on by institutions, whereas Dewey (1927) imagines plural, active publics that emerge within specific contexts, formed by citizens mobilized around particular issues.

In contrast to the sociopolitical approach, communication and media studies have adopted a cultural perspective on publics, focusing on the study of audiences and the reception of cultural and media products. This perspective has its roots in studies of reading practices, which examine the disparity between the abstract publics envisioned by authors and the real, embodied publics that engage with their work (Livingstone, 1998b). Drawing from Umberto Eco's distinction between the model reader and the empirical reader, Dayan (2000) and Livingstone (1998a) introduced the "text-reader" metaphor to differentiate between what they term the "empirical publics" of television programs and the "model publics" conceived by their producers. We apply a similar approach to data production. In data initiatives, developers frequently envision model or imagined publics that do not necessarily correspond to the actual publics that engage with the data.

The concept of the public serves as a rich theoretical framework and a multidimensional lens for exploring the diverse relationships (often absent or fragile) among data, the infrastructures that facilitate

their collection and dissemination, and the individuals who engage with them (or choose not to). Mobilized by Internet studies, it has led to new terminology such as “networked publics” (boyd, 2010), “ad hoc publics” (Bruns & Burgess, 2015, p. 1), “refracted publics” (Abidin, 2021), and “affective publics” (Papacharissi, 2014), among other variants. While these conceptualizations have particularly highlighted the affordances of social media, their predominant focus on individual platforms limits their relevance for understanding public formation processes in the era of datafication—a phenomenon that inherently transcends platform boundaries (Møller Hartley, Bengtsson, et al., 2023).

To move beyond a media- or platform-centric perspective, Møller Hartley, Bengtsson, and colleagues (2023) propose studying publics through the various interplays among four distinct analytical entry points: data infrastructures, journalistic media, algorithmic mechanisms, and audiences. For instance, this involves examining how the interplay between algorithmic and journalistic logics facilitates or constrains the emergence of audiences or how audiences are transformed through the combined influence of algorithmic operations and user practices.

In this perspective, the concept of “datapublics” is proposed by Møller Hartley, Sørensen, and Mathieu (2023) to shed light on how the logics of datafication transform media practices, and journalism in particular, as well as the formation of audiences. These authors critique the prevalent, oversimplified narrative about the interplay between datafication and publics in datafied democracies, where publics are frequently depicted as passive victims of algorithms, misinformation, and the commercial imperatives of big tech. While acknowledging this asymmetry, the authors underscore the scarcity of empirical research exploring how data infrastructures within small and medium-sized media organizations contribute to the development of publics. Similarly, Mathieu and Jorge (2020) and Skrubbeltrang Mahnke et al. (2024) advocate for a renewed focus on audience studies to gain a more in-depth understanding of how individuals’ situated experiences with data influence their roles in datafication processes. By concentrating on datapublics rather than datafied publics, this research creates opportunities to investigate the “reciprocal influences that one [the data] may have on the other [the public] without presuming hierarchy or directionality” (Møller Hartley, Sørensen, & Mathieu, 2023, p. 13). In this context, data publics are distinguished from datafied publics (Baack, 2015), calculated publics (Gillespie, 2014), and algorithmic publics (Christin, 2020), which all emphasize algorithmic objectification and directly link the formation of publics to technological infrastructures and the data they produce.

Building on this scholarship and drawing on Møller Hartley, Sørensen, and Mathieu’s work (2023), we define data publics as “social subjects who interact—or are expected to interact—with data generated about them, by them, or for them” (Millerand et al., 2025, p. 11). This notion serves as a heuristic for empirically analyzing how publics are constructed, imagined, and embodied within data infrastructures across different actors and contexts. Its analytical value lies in shifting the focus away from data per se and toward the publics that form around data production, circulation, use, or nonuse. For example, if we are studying the data sets published by a city’s open data portal, we might ask: Which publics are targeted by the release of municipal data? How are these imagined publics translated into the design of the portal? Who is excluded (either by default or by design) because of accessibility or intelligibility barriers? Which publics actually engage with the data? Finally, how do empirical publics interact with imagined publics? Through these various instantiations of data publics, we can explore the underlying biases that shape them, such as

the idea of a competent, data-literate public versus one composed of “ordinary” citizens who lack the skills to use data effectively (Millerand et al., 2023).

Three Analytical Axes

We propose to organize theoretical approaches to publics into a three-axis analytical framework for empirically studying data publics. Each axis juxtaposes different conceptions of the public: passive versus agentive, ideal versus empirical, and cognitive versus affective. We understand these axes as continuums, not mutually exclusive, and subject to fluctuation over time.

Axis 1: From Passive to Agentive Publics

The first axis spans a continuum between a passive view of the public—as a collective acted on and formed in response to an external initiative—and an agentive view, where the public emerges autonomously in response to specific situations. This tension is often linked to the distinction between political and commercial conceptions of publics, which leads Livingstone (2005) to differentiate the public (with its engaged, political dimension) from the audience (a marketing construct representing a target market to be reached). At its core, this axis helps us think about the varying degrees of autonomy or heteronomy involved in the constitution of publics.

This tension plays out between two normative visions of data publics: on one end, the “commercial and datafied profiling of audiences,” or “publics-as-data,” and on the other, publics as “collectives gathered around a debate on a given issue,” or “publics-as-citizens” (Møller Hartley, Sørensen, & Mathieu, 2023, pp. 13, 198). These contradictory conceptions have always been present in the ways media organizations view their audiences, and they have evolved as new audience measurement technologies have emerged, shifting from democratic collectives to consumer segments and, eventually, to aggregated data points (Møller Hartley & Schjøtt, 2023).

Axis 2: From Ideal to Empirical Publics

The second axis contrasts ideal publics—those imagined or anticipated in the design of a project—with empirical publics, those who engage with it. This distinction is rooted in a long-standing tradition in audience studies that highlights the often stark difference between the publics imagined by media producers and those who engage with media in real-world contexts (Livingstone, 1998a). These imagined publics—sometimes referred to as “constructed,” “targeted,” or “designed”—are often shaped by the strategic visions of institutions, funders, or developers. In contrast, empirical publics emerge from actual practices, interactions, and appropriations (Mathieu, 2023). Empirical publics are similar to the “material” publics examined by Marres (2015), which form around issues, objects, or material practices and often diverge from idealized conceptions.

In the context of data initiatives, this distinction is especially salient. Data infrastructures rely on ways of filtering and categorizing publics that are guided by worldviews (e.g., publics gathered around political issues on Twitter or around social connections on Facebook; Birkbak & Carlsen, 2016). Developers

may imagine certain types of users or beneficiaries (e.g., “informed citizens” or “data-literate professionals”), but empirical publics are sometimes more diverse and less predictable or interact with the data in unforeseen ways, which leads to a reassessment of how publics are affected by data (see Mathieu, 2023). At other times, the imagined publics never materialize, revealing a disconnect between institutional projections and situated realities.

Axis 3: From Cognitive to Affective Publics

The third axis articulates a continuum between cognitive publics—defined by rational, informational engagement—and affective publics, mobilized through emotions, values, or shared experiences. This distinction originates in communication theory, particularly from classical works on crowds, audiences, and publics (Park, 1904/1972; Tarde, 1901/1989), as well as from recent studies on affective publics (Papacharissi, 2022) that show how public formation increasingly involves emotional registers, particularly in digital environments.

Cognitive publics are typically envisioned as rational actors who engage with information to make informed decisions, participate in democratic processes, or enhance their lives through knowledge. This sociopolitical conception of the public (Habermas, 1989) aligns with technocratic ideals of data transparency, informed citizenship, and evidence-based decision making. It is often reflected in the language of public policies (e.g., open data), platform interfaces, and institutional discourses, where data are framed as a neutral tool for enlightenment or efficiency.

In contrast, affective publics form around shared feelings such as anger, fear, hope, or empathy and emerge in response to collective experiences or injustices. They may be motivated by embodied experiences or moral indignation and form collectives around an “affective problematic,” such as air pollution (Pritchard & Gabrys, 2016). This conception aligns with that of Dewey, who identifies affective sensations and practical concerns as the primary drivers of public formation around issues. Although often viewed with suspicion in policymaking circles (as irrational or unpredictable), affective publics play a key role in mobilizing action and contesting dominant discourses, as well as in the production of data and (counter-) narratives around data (Gabrys et al., 2016).

By identifying and grouping these three axes into a unified framework, this analytical model functions as a theory/methods package (Clarke & Star, 2008), combining “both epistemological and ontological assumptions along with concrete practices” (Clarke et al., 2018, p. 24). The framework draws on existing theories to operationalize them into a methodological model to guide the analysis of results and examine data publics while facilitating their comparison in various empirical contexts.

It is important to note that the axes are continuums and that their extremes do not mutually exclude each other. Publics often combine affective and cognitive dimensions, particularly when interpreting data on topics that directly affect them (public health, safety, or inequalities). These continuums aim less to hierarchize than to observe how the publics imagined by data infrastructure designers interact, coexist, or clash with empirical publics in specific data arrangements. They also invite

reflection on who is excluded from either vision and which mechanisms enable or hinder the emergence of empirical publics.

Methodology

To analyze how data publics are envisioned and defined across three distinct cases, we adopted a comprehensive and critical qualitative methodology. Our approach followed an ethnographically oriented case study method (Hammersley et al., 2000), with the originality of combining three different sites corresponding to three “particular” cases in Becker’s (2014) sense of the term. For Becker (2014), linking cases that share specific characteristics about a given phenomenon allows for reflection on the underlying mechanisms of that phenomenon and, ultimately, on how society functions. This involves conducting “in-depth studies of situations, organizations, and singular types of events” (Becker, 2014, p. 11), using “reasoning by cases” to understand the processes that characterize a phenomenon. Through our three cases, we aimed to describe and analyze different contexts of datafication to identify the concrete logics behind how data sets are produced and used and how these logics relate to their various publics.

Because the actors and practices we examined operate across both online and offline spaces, we adopted a “connective” ethnographic approach (Hine, 2015) based on the premise that these spaces are deeply intertwined. Our observations, conducted both online and in person, generated field notes that we triangulated with three additional data sources: semi-structured interviews, document analysis, and platform analysis (including interface and functionalities; see Table 1). Each case had its particularities, so we adapted our methods and data collection tools accordingly.

Table 1. Empirical Material for Each of the Three Cases.

Case 1: City of Montreal’s ODP	Case 2: Pulsar Platform	Case 3: DFG—Montreal
<i>In situ</i> observation: public events (2020–2021)[1]	<i>In situ</i> observation: public events, internal meetings (2021)	<i>In situ</i> observation: monthly meetings (2019)[2]
Online observation: regular consultation of the portal (2020–2021)	Online observation: two-month participant observation in 2020: involvement in two health data collection projects	Online observation: monthly meetings and ongoing activity on platforms (MeetUp, Slack, Trello, GitHub, GitLab) from 2019 to 2021
Interviews: 10 interviews (site management team, developers, advisers, municipal employees)	Interviews: 12 interviews (platform management team, public representatives’ committee, communications leads)	Interviews: 12 interviews (members of DFG—Montreal and other chapters across Canada)
Document corpus (81): official documents (e.g., Open Data Policy), technical documentation (e.g., data forms), online publications (blog posts, social media accounts, press articles)	Document corpus (125): internal documents, data catalog, online publications (e.g., social media	Document corpus: internal documents (e.g., data sets, slide decks), online publications (e.g.,

Website analysis: versions of the portal from 2011 to 2021 (via Internet Archive)	accounts, newsletters, press articles)	social media accounts, website texts)
	Platform analysis: versions of the platform from 2018 to 2021	

Notes. [1] *In situ* observation of City of Montreal employees could not be conducted because of the COVID-19 health crisis. [2] Monthly in-person meetings were moved online starting in April 2020 because of the COVID-19 health crisis.

Results

We describe how each data initiative was developed by focusing on the relationships between data and publics.

The City of Montreal's ODP

The City of Montreal's ODP is a local expression of a broader movement to open public data across Canada. Launched in 2011, the portal now offers over 350 data sets across a wide range of domains (finance, environment, education, society, culture, etc.). It includes features such as data set downloads and data visualization tools.²

The development of the portal coincided with the wave of enthusiasm for open data at the end of the 2000s, which carried with it a heterogeneous set of imaginaries ranging from citizen empowerment to commercial innovation and public service optimization (Kitchin, 2014). The ODP reflects these dynamics. In our ethnographic observations, we quickly identified a gap among (a) the imagined publics invoked by the portal's designers, often found in municipal promotional discourse and generally referring to a vague "ordinary citizen"; (b) the publics configured through the design and functionalities of the website; and (c) the publics constructed through specific engagement strategies initiated by the portal team (e.g., hackathons). This finding is not unique to Montreal: Many municipal ODPs have struggled for years to build and sustain an active user base (Broomfield & Reutter, 2022).

Interactions with actual users prompted the portal team to recognize the mismatch between intended and real publics. This led them to reassess the portal's goals: Should it serve citizens by providing access to information, help them organize their services, stimulate commercial innovation, or optimize existing public services? These questions led to significant changes in the portal's design and orientation. We identify three major phases that illustrate the evolution of relationships between data and publics.

The first phase (2011–2014) reflects a somewhat naive stance on open data, driven mainly by a desire for institutional transparency. The underlying assumption was that making data sets available would naturally lead to their appropriation. One member of the portal team summarized this mindset: "It feels

² For an in-depth analysis of this case, see Millerand et al. (2023).

like, for the past 10 years, we haven't really paid attention—we've been hoarding data compulsively, and then we don't know what it's for."

The second phase of the analyzed initiative (2015–2020) evidences a dual transformation. First, the emergence of the smart city paradigm has engendered heightened expectations concerning the utilization of data sourced from the ODP. Stakeholders have articulated a belief in the potential of these data as an asset that can enhance public service optimization and drive economic development. As one participant from the portal team articulated, "We know data is valuable, . . . and we think there's real potential to use it more in decision making. We also believe there's a data culture to be developed—among both citizens and city staff." This perspective reflects an evolving recognition of the importance of fostering a data-centric culture within the urban environment.

Concurrently, this phase marks a growing emphasis on enabling "ordinary" citizens to effectively appropriate and engage with the data provided by the portal. This commitment is manifest in the evolution of the portal's design, which has incorporated new visualizations, the introduction of "data stories" that illustrate practical applications of data sets, and features that allow users to request the release of specific data sets.

The subsequent phase, commencing in 2021, is characterized by a continuation of these efforts, with an explicit focus on reconciling the needs of expert users with the demand for data accessibility among nonexperts. This period is also shaped by increasing concerns about data security and digital sustainability, reflecting broader societal debates surrounding technology and its implications.

The first analytical axis of passive versus agentive engagement offers a valuable framework for understanding the dynamics at play within the ODP case. On one hand, the city's rhetoric underscores the empowerment of citizens through data reuse, positing that individuals can actively engage in civic affairs and democratic processes. Conversely, the portal's design appears to configure participation predominantly as the passive consumption of preformatted data sets by citizens. In the former conception, the data public is envisioned as an autonomous entity that reuses data sets pertinent to their needs, exemplified by a citizen group that develops an application to assist drivers in navigating the city's complex parking regulations. In contrast, the latter perspective identifies user groups in advance, treating them as a relatively passive audience for a curated selection of data sets. Empirical observations indicate that these two conceptions coexist, despite their inherently contradictory perspectives on public agency.

The second axis (ideal versus empirical) highlights an important tension. Portal managers often refer to "the citizen" as the primary public for open data. However, ongoing interactions with actual users have gradually dismantled this singular, abstract figure. Over time, particularly through the three phases of the portal's development, the idealized vision of "the" citizen has evolved into a more nuanced understanding of diverse user groups—such as "experts," "citizens," and "corporate citizens"—each with distinct data needs and expectations.

The third axis (cognitive versus affective) underscores a rational perspective of the public as articulated by the portal team. The appeal of open data is portrayed as self-evident; it provides vital

information about the city, benefiting citizens both individually and collectively, for instance, by providing access to snow removal schedules, construction permits, and air quality data. This perspective assumes that public engagement emerges from shared rationality and mutual interests. However, tensions arise within this framework. For example, one team member expressed concerns about “sensitive” data that might elicit emotional or disruptive responses, such as geolocated data on bed bug infestations. The apprehension surrounding potential public “overreaction” to such data sets highlights a preference for rational discourse and a skepticism toward affective responses. Consequently, the portal aims to cultivate a rational/reasonable data public rather than an emotionally reactive crowd.

The Pulsar Platform

Launched in 2019 by a consortium of public and private health-sector actors in the Quebec City region under the leadership of Université Laval, Pulsar presents itself as a digital platform, a network of actors, and a data infrastructure. One of the project’s objectives is to provide access to longitudinal data on a large population of Quebec City residents for research and public policy improvement. Pulsar positions itself as a “sustainable health” initiative. Citizens who register on the platform can participate in various research projects and access their data to support the self-management of their health. Their anonymized data are then made available to researchers through the Sustainable Health Data Bank (*Banque de données en santé durable*).

From the project’s initial phase, three distinct groups were identified as users of the platform: “citizens,” “researchers,” and “decision-makers.” The citizen group was actively engaged in various ways throughout the project’s development. During the design phase of the platform, citizens participated in consultations alongside researchers and public health institutions. In the development phase, they were mobilized through an engagement strategy aimed at establishing a “public representatives’ committee.” At that stage, they served as an envisioned public (i.e., the anticipated users of the platform). The other two groups identified during the platform’s development were researchers and decision-makers. Researchers include individuals leading or collaborating on health research projects utilizing Pulsar and the Sustainable Health Data Bank. “Decision-makers” refers to institutional partners—primarily governmental or quasigovernmental entities, as well as some professionals within the health sector.

Our attention was particularly drawn to how citizens were conceptualized as a public, revealing the ambiguities and contradictions that frequently characterize the perceptions of data initiative designers about their target audience. Citizens were referred to in three distinct ways: as “participants” in research projects facilitated by the platform (adhering to the “patient-partner” model), “public representatives” (an abstract, preexisting general public), and “users” of the platform—also termed “members” of Pulsar—who are invited to utilize its functionalities, such as discovering new projects to join.

The “public representatives’ committee” was initially envisioned as a group of 10 citizens who would actively engage in co-constructing and governing the platform. They were intended to play a role in selecting research projects that could significantly affect population health. However, the recruitment process revealed several contradictory imperatives. On one hand, members were expected to be aligned with the goals of “sustainable health” and to be broadly representative of the population, in addition to being

individuals already involved in community organizations. On the other hand, they were instructed not to advocate for their organizations' causes and to fully align their interests with those of the platform. These conflicting demands ultimately led to the dissolution of the committee.

The figure of the citizen as a "user" or "member" of the platform presents another paradox. While citizens are invited to consider themselves as part of a "community"³ (brought together via newsletters, a Facebook page, etc.), they are also instructed not to interact with one another. This is meant to prevent mutual influence or coordinated action, which the platform's promoters believe could bias health data. The contrast becomes especially striking when compared with the trajectories of the researcher and decision-maker publics. Although both of those groups also diverged somewhat from their original imagined forms, they navigated the development process in ways that allowed them to express their expectations within formal decision-making structures.

The examination of the Pulsar case through the first analytical axis (passive versus agentive) reveals a persistent tension between the prevailing passive conceptualization of the citizen public and the citizens' self-perception as active agents. The challenges associated with the establishment of a public representatives' committee serve as a poignant illustration of the rejection of a genuinely agentive public in favor of a more extractive paradigm, wherein the public is predominantly viewed as a "resource."

On the second axis (ideal versus empirical), the case elucidates two contrasting perspectives. First, the strictly ideal public is envisioned as an abstract "general public" comprising the health service population. This idealized construct was deemed so intangible that the designers felt compelled to represent it via a committee of "real" citizens. Conversely, the empirical public of citizens ultimately experienced a curtailment of agency, as many of the mechanisms intended for public engagement were neglected or abandoned. Although the creation of a user account on the platform remains feasible, the anticipated functionalities, such as personalized dashboards and health tracking tools, were never realized. This outcome underscores the earlier analytical observation that the researcher publics were prioritized, relegating citizens to the status of research subjects.

Lastly, the third axis (cognitive versus affective) addresses a public imagined through a lens of rationality and altruism. Citizens are inherently expected to contribute to health research for the collective good and to place their data in the hands of researchers and platform designers. However, one platform designer astutely noted the reluctance among researchers to share their data for subsequent inquiries: "Researchers panic when you use their data." Furthermore, members of the public representatives committee were instructed to temporarily set aside their personal motivations, often inherent in patient advocacy. The rationale offered was that such affective dimensions might introduce bias. Nonetheless, the degree to which these affective expressions were acknowledged varied according to the perceived value of the citizen public and the ethical constraints guiding the designers. This situation further accentuates the reduction of the citizen public to the role of "research subject."

³ Term used on the platform.

The DFG Collective

DFG is a volunteer-based citizen initiative launched in 2013. Following a bottom-up logic, DFG operates through a dozen chapters across Canada, offering free data science expertise to local community organizations. The initiative relies entirely on volunteer commitment and supports organizations working on the ground for social justice. For our ethnographic study, we focused on the Montreal chapter. We observed the project that was occupying volunteers at the time: working with data sets from a network of shelters for women experiencing intimate partner violence. In this case, we identified three distinct data publics: the DFG volunteers, the individuals within local organizations who collect the data (i.e., DFG's "clients"), and the people who benefit from these organizations' services.

What makes this case distinctive is its relationship to both data and publics. DFG produces little amounts of original data; instead, it specializes in analyzing existing data sets provided by community organizations. The collective acts as a mediator of sorts. Its "client" organizations are typically nongovernmental organizations working in health, safety, or access to justice, supporting individuals in vulnerable situations. These organizations approach DFG to optimize their services and increase their capacity to secure funding, in both cases by leveraging their internal data. As the Montreal chapter founder explained, representatives from these organizations often show up at the first meeting saying, "We've been collecting data for 10 years, but we don't really know what to do with it, and we don't have the expertise to do so."

The volunteers of DFG constitute a particularly engaged public. Relying on their investment and skills—many are data science professionals—they work to "make the data speak." Among the volunteers, some described themselves as activists, while others explicitly rejected any political label. However, all emphasized that their motivation to volunteer stemmed from a belief in the cause pursued by the "client" organization.

The community organization itself, as the "client," also represents a public for the data. Its frontline workers (e.g., social workers, outreach staff) generate data through their day-to-day activities, yet typically have limited data literacy. Nonetheless, this public plays a critical role by contextualizing the data. For example, one volunteer raised issues related to cleaning nonstandardized Excel tables, where some figures appeared to be errors. Representatives from the organization explained that these "anomalies" reflected crucial information: The numbers exceeding the official shelter bed capacity indicated the number of women and children admitted even after the shelter was at full capacity. On such nights, staff would prepare makeshift bedding on the floor. Removing those data points would erase a compelling argument for funding: The documented bed capacity is insufficient to meet demand.

Finally, we identified a third data public: those most directly impacted by datafication. This group consisted of women utilizing the shelter network. This public occupies a unique position, serving as the source of the data while having little to no interaction with them. Their agency is limited by urgent safety concerns for themselves and their children, placing them in a state of precarity that must be recognized. Frequently, those data publics most at risk of adverse effects from datafication are either underrepresented

in data or extensively tracked (D'Ignazio & Klein, 2020), yet they are often unlikely to engage directly with the data initiatives that affect them.

The investigation into the DFG initiative reveals nuances relevant to the first analytical axis—the political conceptualization of an active public. The public of volunteers is characterized by a collective intent to support community organizations. Initially, the assembly of data volunteers appears to be committed and even activist in nature (Milan, 2017). However, many volunteers reported their initial motivation as the desire to improve their employability. This complexity aligns with Dewey's (1995) broader interpretation of the political, positing that politics is inherently interwoven with quotidian choices and actions that empower individuals to define themselves as publics. Furthermore, community organization staff exhibit a degree of agency by strategically recruiting volunteers to address deficiencies in internal data expertise. One should also note the elusive, yet significant, empirical public represented by the datafied women. Evaluating their agency solely through the framework of data publics would be overly reductive. Nonetheless, their inclusion in the organization's records signifies a form of agency that positions them as agentive data publics.

Analyzing DFG through the lens of the second axis, it becomes apparent that the presence of an ideal public is largely absent; rather, all publics are empirical in nature. The collective's work is initiated by diagnosing an explicit need articulated by a community organization, which serves as both the data provider and the primary beneficiary of the outcomes. The data publics in this scenario are not only situated but also embodied. A notable strength of this case lies in its demonstration of the rarity of such fully embodied representations within most data initiatives.

Finally, regarding the third analytical axis, the DFG case tends to adopt a rational rather than affective interpretation of publics. Data publics are conceptualized as collectives engaged in civic endeavors for the common good. However, volunteers often exhibit affective motivations; they engage in laborious tasks (e.g., cleaning disordered data sets) because of an emotional connection to the cause—namely, supporting women affected by violence. While some individuals initially join the initiative with the aim of enhancing their CVs, those who remain over the long term are often driven by an alignment with the mission of the client organization, which they perceive as enabling them to effectuate meaningful change.

Discussion

Historically, the concept of public has provided the analytical flexibility needed to describe the mutual constitution of cultural objects and social groups (Esquenazi, 2009). Data sets can be seen as informational and cultural objects (Couldry & Powell, 2014) that need to be grasped by the very publics they contribute to shape (Mathieu, 2023). Viewing data through the lens of their publics enables us to adopt an epistemological perspective that distances us from what we might call the "cult of data" (the data doxa), for which the centrality of data in our society is naturalized (Smith, 2018). Millerand et al. (2025) note that

interrogating data publics requires examining what datafication "counts" and then "renders public"—that is, the processes of publicization it entails—as well as how datafication "makes a public"—that is, what kinds of collective entities are constituted (or not) by these processes. (p. 18)

The heuristic value of the concept of public takes on its full significance as the analysis shows how ways of understanding publics shape the design and evolution of data initiatives.

The three case studies elucidate the tensions that arise between varying conceptualizations of publics and how data initiatives operationalize these tensions in practice. The DFG initiative embraces the empirical diversity of data publics, while the City of Montreal's ODP seeks a balance between an idealized vision of the public and more embodied representations. In contrast, the Pulsar platform largely maintains an idealized conception of its publics. Publics, even imaginary ones, influence action. Rather than viewing datafication strictly in terms of technological progress or as an instrument of control, our analysis of the three cases reveals various dynamics. Comparatively speaking, our ethnographic studies, examined through the lens of the three-axis framework, reveal three different dynamics in terms of how designers of data initiatives relate to their publics, each reflecting specific tensions. The first is conciliation, which occurs when actors with disparate expectations receive adequate recognition, as exemplified by the City of Montreal's ODP. The second is exclusion, which transpires when such actors operate within unequal power relations, as illustrated by the Pulsar platform. The third is embodiment, where designers consciously cede portions of their normative framing authority to facilitate the emergence of a multiplicity of data publics. In the end, our research shows a constant in the three case studies: The less attention data developers pay to data publics in the conception of a data initiative, the more decisions are made on their behalf (passive-agentive axis); the more data publics remain abstract figures (ideal-empirical axis); and the more they are reduced to rational agents, overlooking other dimensions of their experience, such as emotions (cognitive-affective axis).

The three-axis analytical framework allows us to operationalize the existing body of theoretical work on publics along three continuums. Organizing them as coexisting continuums allows for a more nuanced understanding of the mechanisms by which data infrastructures (whether a website, a platform, or data services) evolve in response to gaps between designers' expectations and users' actual practices. The results of such an analysis can be effectively visualized as a polygon on a radar chart, which takes the form of a triangle or prism, where each axis of tension is represented as a line going from the center of the triangle (one end of the axis) to one of its corners, figuring the other end of the axis (see Figure 1). This provides a synthesized depiction of a given data public's positioning along each axis, forming a triangular shape.

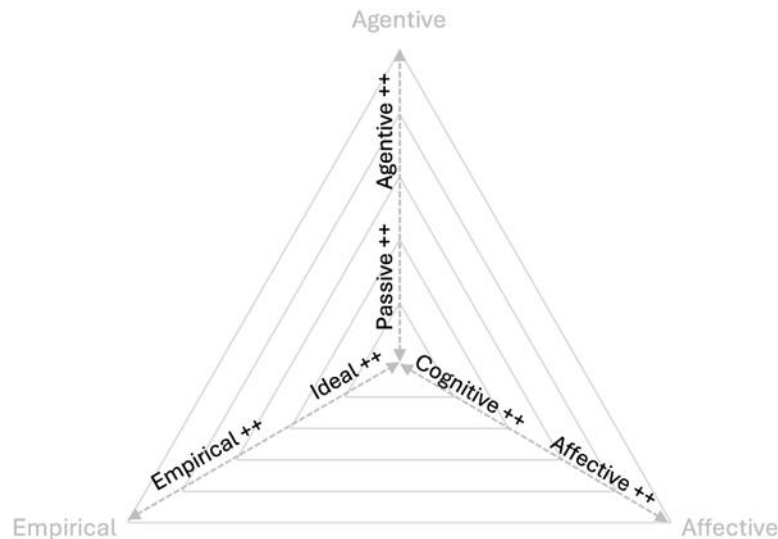


Figure 1. Analytical prism for data publics.

The location of each triangle's point illustrates how a data public is understood in relation to each axis (see Figure 2). Beyond its synoptic appeal, this graphic representation of the three-axis model also proves helpful in the context of a comparative analysis of data publics. The complexity of the object leads researchers to rely on in-depth case studies, which are often difficult to compare, making it challenging to generalize across findings. The prism thus serves to identify recurring patterns between the continuums. In this regard, Figure 2 illustrates across our three case studies that publics conceived of as passive (e.g., Pulsar members) tend to also be idealized and seen as rational. In contrast, more agentive publics (e.g., DFG—Montreal beneficiaries and ODP users) tend to be more empirically grounded and affectively engaged. The affective dimension appears to be underestimated by the designers of data initiatives, yet recognizing affect as a legitimate mode of public engagement (Papacharissi, 2022) allows for a better understanding of how data circulate and acquire (or fail to acquire) meaning in various contexts. It also paves the way for analyzing the emotional dimensions of data governance, such as the fear of misuse, trust in institutions, or fatigue from technocratic discourse.

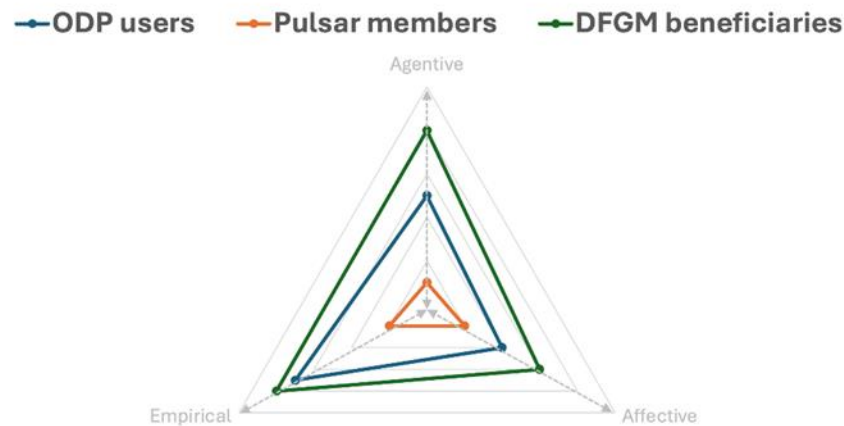


Figure 2. Data publics for the three case studies.

This approach also opens the door to comparing field sites that are otherwise extremely diverse—something the literature has so far struggled to do given the fragmentation of datafication processes across domains (Millerand et al., forthcoming; Møller Hartley, Sørensen, & Mathieu, 2023)—thereby contributing to the development of richer and more grounded empirical investigations (Pybus et al., 2025). This analytical framework and its visualization enable the integration of initiatives that generate data publics at vastly different scales, from large platforms shaping today’s digital landscape to highly localized, community-based initiatives. It allows for measuring the gaps between the idealized local publics imagined by media companies (e.g., those cultivated by design through imaginaries, data logics, and technical infrastructures; Møller Hartley, Schjøtt, & Sørensen, 2023) and the publics that actually materialize in practice.

The prism can also be employed in a diachronic approach, serving as a concise visual tool for tracking the evolution of conceptions of the public within a specific case. Figure 3 illustrates the evolution of the “citizen” data public as it is conceived of throughout the various phases of the City of Montreal’s ODP, as detailed in the previous section. Initially, this public was perceived as competent, autonomous, and motivated by instrumental considerations, in a conception akin to that of the sociopolitical actor: a rational city resident who would develop a use of data to potentially participate in civic affairs. This conception accounts for ODP designers’ decision to publish raw data sets. However, over time, the ODP team began to consider the actual conditions affecting public access (e.g., hardware and software), the necessity for support in utilizing the data, and the diverse motivations that drive users to engage with the portal. In other words, the absence of empirical publics—or rather, the inability to expand them beyond expert users—has highlighted the importance of material aspects in the formation of data publics (Marres, 2015).



Figure 3. Evolution of the Montreal ODP developers' conception of the "citizen" data public.

Finally, the observable evolution of public engagement underscores the dominant paradigms that inform the conceptualizations of publics—frameworks that are profoundly normative in nature (Møller Hartley, Sørensen, & Mathieu, 2023)—as well as the power dynamics that influence the trajectories of these conceptualizations. Notably, the phenomenon of datafication does not extend the same advantages to all publics; indeed, certain groups are conspicuously marginalized. Questioning ways of understanding data publics, including investigating their emergence and actualization, presupposes considering the imaginaries, discourses, practices, and sociotechnical mediations that help shape them.

Conclusion

Drawing on three ethnographic case studies, we have observed how data publics developed as they took shape within different data initiatives. We proposed an analytical framework operationalizing different theoretical conceptions of publics along three axes. The results of our study suggest that this framework can be used in future studies; it allows for the empirical examination of data publics while facilitating their comparison in various contexts. This analytical prism aims to facilitate the development of comparative case studies to understand datafication from the perspective of the publics affected by the data.

While our three case studies shed light on the intricacies of data publics, they equally underscore the limitations inherent in this conceptual framework. Although our framework facilitates the identification of power dynamics that enhance or restrict the agency of individuals involved in datafication processes, it falls short of fully elucidating the foundational mechanisms that drive these interactions. Future research endeavors may address these gaps by incorporating theoretical constructs that enable the exploration of power relations at both micro- and macrosocial dimensions as they affect the data public. Nevertheless, we contend that the concept of data public remains a useful and generative framework, as it offers a nuanced perspective on the collectives that emerge around, for, and through data within a multifaceted sociotechnical landscape.

References

- Abidin, C. (2021). From "networked publics" to "refracted publics": A companion framework for researching "below the radar" studies. *Social Media + Society*, 7(1), 1–13. <https://doi.org/10.1177/2056305120984458>

- Andrejevic, M. (2014). The big data divide. *International Journal of Communication*, 8(17), 1673–1689. <https://ijoc.org/index.php/ijoc/article/view/2161>
- Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, 2(2), 1–11. <https://doi.org/10.1177/2053951715594634>
- Becker, H. S. (2014). *What about Mozart? What about murder?: Reasoning from cases*. Chicago, IL: University of Chicago Press.
- Birkbak, A., & Carlsen, H. (2016). The public and its algorithms. Comparing and experimenting with calculated publics. In L. Amoore & V. Piotukh (Eds.), *Algorithmic life: Calculative devices in the age of big data* (pp. 21–34). New York, NY: Routledge.
- boyd, d. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In Z. Papacharissi (Ed.), *Networked self: Identity, community, and culture on social network sites* (pp. 39–58). New York, NY: Routledge.
- Broomfield, H., & Reutter, L. (2022). In search of the citizen in the datafication of public administration. *Big Data & Society*, 9(1), 1–14. <https://doi.org/10.1177/20539517221089302>
- Bruns, A., & Burgess, J. (2015). Twitter hashtags from ad hoc to calculated publics. In N. Rambukkana (Ed.), *Hashtag publics: The power and politics of discursive networks* (pp. 13–27). New York, NY: Peter Lang Publishing.
- Burgess, J., Albury, K., McCosker, A., & Wilken, R. (2022). *Everyday data cultures*. Cambridge, UK: Polity.
- Christin, A. (2020). *Metrics at work. Journalism and the contested meaning of algorithms*. Princeton, NJ: Princeton University Press.
- Clarke, A. E., Friese, C., & Washburn, R. S. (2018). *Situational analysis: Grounded theory after the interpretive turn* (2nd ed.). Los Angeles, CA: SAGE.
- Clarke, A. E., & Star, S. L. (2008). The social worlds framework: A theory/methods package. In E. J. Hackett, O. Amsterdamska, M. Lynch, & J. Wajcman (Eds.), *The handbook of science and technology studies* (3rd ed., pp. 113–137). Cambridge, MA: The MIT Press.
- Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4), 336–349. <https://doi.org/10.1177/1527476418796632>
- Couldry, N., & Powell, A. (2014). Big data from the bottom up. *Big Data & Society*, 1(2), 1–5. <https://doi.org/10.1177/2053951714539277>

- Dayan, D. (2000). Télévision: Le presque-public [Television: The almost-public]. *Réseaux*, 100(2), 427–456. <https://doi.org/10.3406/reso.2000.2232>
- Dewey, J. (1927). *The public and its problems*. New York, NY: H. Holt and Company.
- Dewey, J. (1995). La démocratie créatrice: La tâche qui nous attend [Creative democracy: The task that awaits us]. *Horizons Philosophiques*, 5(2), 41–48. <https://doi.org/10.7202/800979ar>
- D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. Cambridge, MA: MIT Press.
- Dijck, J. van. (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>
- Esquenazi, J.-P. (2009). *Sociologie des publics* [Sociology of audiences]. Paris, France: La Découverte.
- Flensburg, S., & Lomborg, S. (2021). Datafication research: Mapping the field for a future agenda. *New Media & Society*, 25(6), 1451–1469. <https://doi.org/10.1177/14614448211046616>
- Gabrys, J., Pritchard, H., & Barratt, B. (2016). Just good enough data: Figuring data citizenships through air pollution sensing and data stories. *Big Data & Society*, 3(2), 1–14. <https://doi.org/10.1177/2053951716679677>
- Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Media technologies* (pp. 167–194). Cambridge, MA: The MIT Press.
- Habermas, J. (1989). *The structural transformation of the public sphere. An inquiry into a category of bourgeois society*. Cambridge, MA: Polity.
- Hammersley, M., Foster, P., & Gomm, R. (2000). Case study and theory. In R. Gomm, M. Hammersley, & P. Foster (Eds.), *Case study method: Key issues, key texts* (pp. 234–258). London, UK: SAGE.
- Hine, C. (2015). *Ethnography for the Internet: Embedded, embodied and everyday*. London, UK: Bloomsbury.
- Kitchin, R. (2014). *The data revolution. Big data, open data, data infrastructures and their consequences*. London, UK: SAGE.
- Lippmann, W. (1925). *The phantom public*. New York, NY: Harcourt, Brace and Cie.
- Livingstone, S. (1998a). *Making sense of television: The psychology of audience interpretation*. London, UK: Routledge.

- Livingstone, S. (1998b). Relationships between media and audiences: Prospects for audience reception studies. In T. Liebes & J. Curran (Eds.), *Media, ritual and identity: Essays in honor of Elihu Katz* (pp. 237–255). London, UK: Routledge.
- Livingstone, S. (2005). On the relation between audiences and publics. In S. Livingstone (Ed.), *Audiences and publics: When cultural engagement matters for the public sphere. Changing media—Changing Europe series (2)* (pp. 17–41). Bristol, UK: Intellect Books.
- Lycett, M. (2013). "Datafication": Making sense of (big) data in a complex world. *European Journal of Information Systems*, 22(4), 381–386. <https://doi.org/10.1057/ejis.2013.10>
- Marres, N. (2015). *Material participation: Technology, the environment and everyday publics*. London, UK: Palgrave Macmillan.
- Mathieu, D. (2023). Deconstructing the notion of algorithmic control over datapublics. In J. Møller Hartley, J. K. Sørensen, & D. Mathieu (Eds.), *Datapublics. The construction of publics in datafied democracies* (pp. 27–48). Bristol, UK: Bristol University Press.
- Mathieu, D., & Jorge, A. (2020). The datafication of media (and) audiences: An introduction. *MedieKultur: Journal of Media and Communication Research*, 36(69), 1–10. <https://doi.org/10.7146/mediekultur.v36i69.122585>
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Boston, MA: Houghton Mifflin Harcourt.
- Milan, S. (2017). Data activism as the new frontier of media activism. In V. Pickard & G. Yang (Eds.), *Media activism in the digital age* (pp. 151–163). London, UK: Routledge. <https://doi.org/10.4324/9781315393940-13>
- Millerand, F., Coutant, A., Latzko-Toth, G., & Millette, M. (2025). *Les publics de données. Penser la datafication de la société* [Data publics. Rethinking the datafication of society]. Montréal, QC: Presses de l'Université de Montréal.
- Millerand, F., Delias, L., Coutant, A., & Fortier, M.-S. (2023). À la recherche du citoyen «ordinaire», les publics imaginés de l'ouverture des données publiques au niveau municipal [In search of the "ordinary" citizen: Imagined publics of municipal open data initiatives]. *Les Enjeux de l'Information et de la Communication*, 23(4), 109–129. <https://lesenjeux.univ-grenoble-alpes.fr/2023/dossier/07-a-la-recherche-du-citoyen-ordinaire-les-publics-imagines-de-louverture-des-donnees-publiques-au-niveau-municipal/>
- Millerand, F., Latzko-Toth, G., Millette, M., & Coutant, A. (forthcoming). Data publics: Viewing datafication through the lens of media studies. In U. A. Mejias, J. McNealy, & M. Miceli (Eds.), *Handbook of critical data studies*. Berlin, Germany: De Gruyter.

- Møller Hartley, J., Bengtsson, M., Schjøtt Hansen, A., & Sivertsen, M. F. (2023). Researching publics in datafied societies: Insights from four approaches to the concept of "publics" and a (hybrid) research agenda. *New Media & Society, 25*(7), 1668–1686.
<https://doi.org/10.1177/14614448211021045>
- Møller Hartley, J., & Schjøtt, A. (2023). Imagining publics through emerging technologies. In J. Møller Hartley, D. Mathieu, & J. K. Sørensen (Eds.), *Datapublics. The construction of publics in datafied democracies* (pp. 99–120). Bristol, UK: Bristol University Press.
- Møller Hartley, J., Schjøtt, A., & Sørensen, J. K. (2023). Personalization logics and publics by design. In J. Møller Hartley, D. Mathieu, & J. K. Sørensen (Eds.), *Datapublics. The construction of publics in datafied democracies* (pp. 121–141). Bristol, UK: Bristol University Press.
- Møller Hartley, J., Sørensen, J., & Mathieu, D. (2023). *Datapublics: The construction of publics in datafied democracies*. Bristol, UK: Bristol University Press.
- Papacharissi, Z. (2022). Affective publics. Solidarity and distance. In D. A. Rohlinger & S. Sobieraj (Eds.), *The Oxford handbook of digital media sociology* (pp. 61–72). New York, NY: Oxford University Press.
- Papacharissi, Z. (2014). *Affective publics: Sentiment, technology, and politics*. Oxford, UK: Oxford University Press.
- Park, R. E. (1972). *The crowd and the public, and other essays*. Chicago, IL: University of Chicago Press. (Original work published 1904)
- Powell, A. (2014). "Datafication," transparency, and good governance of the data city. In K. O'Hara, C. Nguyen, & P. Haynes (Eds.), *Digital enlightenment yearbook* (pp. 215–224). Amsterdam, The Netherlands: ISO Press Ebooks.
- Pritchard, H., & Gabrys, J. (2016). From citizen sensing to collective monitoring: Working through the perceptive and affective problematics of environmental pollution. *GeoHumanities, 2*(2), 354–371.
<https://doi.org/10.1080/2373566X.2016.1234355>
- Pybus, J., Lomborg, S., Gandini, A., & Lai, S. S. (2025). Empirical approaches to infrastructures for datafication: Introduction to the special issue. *New Media & Society, 27*(4), 1851–1867.
<https://doi.org/10.1177/14614448251314396>
- Ruppert, E., Isin, E., & Bigo, D. (2017). Data politics. *Big Data & Society, 4*(2), 1–7.
<https://doi.org/10.1177/2053951717717749>

Skrubbeltrang Mahnke, M., Swart, J., Mathieu, D., & Pruulmann-Vengerfeldt, P. (2024). Data reflectivity and user reflexivity: New conceptual pathways for connecting structural approaches with user perspectives. *Convergence: The International Journal of Research into New Media Technologies*, 30(6), 1859–1870. <https://doi.org/10.1177/13548565241301459>

Smith, G. J. (2018). Data doxa: The affective consequences of data practices. *Big Data & Society*, 5(1), 1–15. <https://doi.org/10.1177/2053951717751551>

Tarde, G. (1989). *L'opinion et la foule* [The opinion and the crowd]. Paris, France: Presses universitaires de France. (Original work published 1901).

Zuboff, S. (2019). *The age of surveillance capitalism. The fight for a human future at the new frontier of power*. New York, NY: PublicAffairs.