

VOL. 5, NO. 4, 2023, 93-125

UTOPIAN AND DYSTOPIAN SOCIOTECHNICAL IMAGINARIES OF BIG DATA

Hana Marčetić^a and Jan Nolin^b

ABSTRACT

Data feminism, a way of thinking about and “doing” data utilizing feminist tools and perspectives, has emerged in recent years as a part of a critical discourse surrounding datafication. The aim of this study is to analyze and identify shared perceptions of big data as expressed in a corpus of scholarly writings published in the domain of data studies and data feminism. We analyzed a set of 44 scholarly texts engaging in feminism concerned with the concept of big data. For the purpose of this article, we refer to this set of texts as data feminism and examine how authors frame and describe big data. We compare future visions in data feminist material with policies by the European Commission and explore what tensions arise among them. Furthermore, we explore and delineate social and political alternatives that emerge from data feminist texts. Both corpora describe futures inclusive of big data and imagine possible positive outcomes from different perspectives and with different ideas of the current role of big data. We found that sociotechnical imaginaries of big data within the data feminist corpus are considerably richer and more nuanced than those of the European Commission. In the data feminist corpus, big data is described as a multiplicity of things and often implicated in perpetuating power imbalances and large societal issues. The European Commission corpus employs the perspective of “data as a resource” to be exploited.

Keywords: big data; data; the European Commission; feminism; sociotechnical imaginaries; VOSViewer.

^a Swedish School of Library and Information Science, Sweden.

^b Swedish School of Library and Information Science, Sweden.

1 INTRODUCTION

The rise of the Internet and social media, as well as the accumulation of vast amounts of data have created a setting for practices and structures associated with what has come to be called big data. It is a curious concept that has served as the foundation for both optimistic and pessimistic visions of the future. Big data is a contested concept and there is no consensus on its definition. In lieu of the lack of an agreed-upon definition, the “three V’s” attributed to Gartner (Ward & Barker, 2013, p. 1) are often used to explain big data. The “three V’s” are Volume, Velocity, and Variety, and new words have been introduced since (such as Value and Veracity). Simultaneously, critical discussions are continuing to shed light on the array of systems, decisions, and processes that influence what we perceive as big data. Together with predominant narratives about big data, they help shape how we envision and prepare for the future with big data.

According to the theoretical approach of sociotechnical imaginaries (Jasanoff, 2015; Jasanoff et al., 2007), visions of the future are formed among groups in society and frequent conflicts can occur between the imaginaries of different groups. Sociotechnical imaginaries are self-fulfilling and agenda-setting: today’s dominating imaginaries set the boundaries for how the future will unfold. This has proven particularly relevant in the current decade, as long-term planning and investments by governments and corporations in the West are focused on images of smart cities, the Internet of Things, 5G telecommunications, augmented reality, virtual reality, and much more. The creation and application of big data involve practices that bind together many of these imaginaries.

The notion of big data as fostering both utopian and dystopian discussions goes back at least a decade, when boyd and Crawford (2012) stated that the application of big data at a scale triggers both utopian and dystopian rhetoric. According to them, utopian rhetoric describes big data as a helpful tool for simplifying and streamlining complex systems. In contrast, the dystopian rhetoric predominantly regards big data as capable of enabling privacy invasions and decreased civil freedoms. In this text, we use the notions of dystopian and utopian sociotechnical imaginaries to signal the utilization of the dystopian and utopian rhetoric described by boyd and Crawford. This is not to imply or feed into narratives describing feminists as being against technology or progress. In fact, women have long been at the forefront of technological development, even though their contributions were often unrecognized and erased.

With this in mind, *utopian sociotechnical imaginaries* encompass those imaginaries that depict big data as capable of improving social systems and economies. These sociotechnical imaginaries are promoted by two very different but highly influential groups. The first comprises IT companies, particularly those

that have their headquarters in Silicon Valley. These companies often appear unified in marketing positive perspectives about technologies in the making (Lindh & Nolin, 2017). The second group entails policymakers who aspire to boost their economies by implementing data-driven innovation. These two groups frame big data as a resource and a tool for empowerment and social change (Levina and Hasinoff, 2016). Big data is described as beneficial across domains, from commerce to health and government (Intel, 2021; Chen et al., 2012) and helpful for solving local and social problems (e.g., Guha, 2021). Zuboff (2019) argues that positive discourses are dominated by Big Tech through the power of declarations. They function by “impos[ing] new facts on the social world while their declarers devise ways to get others to agree to those facts” (p. 177). According to Zuboff, Big Tech companies move into uncharted territories to claim them, subsequently tailoring the direction of their development.

In dystopian rhetoric, big data is seen as problematic and capable of inflicting damage (Gregory & Halff, 2020; O’Neil, 2017). Considering this, discourses developed by scholars from various fields who choose a critical perspective toward big data fall under the scope of *dystopian sociotechnical imaginaries*. Critical data studies, surveillance studies, and feminist studies are examples of the critical lenses used to problematize and scrutinize various aspects of big data. Another way of describing the dystopian rhetoric would be to call it anti-utopian, as such accounts can provide constructive and positive ideas. In this article, we are particularly concerned with understanding the *dystopian* in the sociotechnical imaginaries within data feminism.

The term data feminism was popularized in recent years with the publication of the book “Data Feminism” in which it was defined as “a way of thinking about data, both their uses and their limits, that is informed by direct experience, by commitment to action, and by intersectional feminist thought” (D’Ignazio & Klein, 2020, p. 3). Feminist scholars have a tradition of developing tools and theories for studying structures of power and how they are subverted by and intertwined with different phenomena in society. Although data feminism is a relatively new concept associated with a specific program, for simplicity, we will use it broadly to cover an array of feminist critical approaches. We will contrast this by comparing data feminist sociotechnical imaginaries with those of the European Commission, as identified by Rieder (2018). We also want to better understand the relationship between the utopian and the dystopian within the sociotechnical imaginaries of big data. The questions that we explore in this article are therefore:

- What characterizes the dystopian sociotechnical imaginaries of big data within data feminism?

- In which ways do these dystopian sociotechnical imaginaries differ from the utopian sociotechnical imaginaries within European Commission policies?
- What are the particular contributions and added value to critical studies that data feminism provides in this area?

In order to perform a meta-analysis of the contributions of data feminism to the topics of this study, we aim to analyze the contributions of data feminism as researchers who are not active within it. We do this by conducting a text analysis of the data feminist corpus in addition to a comparison of official European Commission documents as analyzed by Rieder (2018).

2 SOCIOTECHNICAL IMAGINARIES

One of the most prolifically employed frameworks for studying the confluence of society and technology is the Actor Network Theory (ANT). However, the ANT has been criticized for flattening the thickness of social relationships, hierarchies, and power distributions (Jasanoff, 2015). The concept of sociotechnical imaginaries was developed as a response to this flattening and is used to describe the role of collective imagination in society. Social imaginaries upon which the concept is built have been predominantly defined as:

[T]he ways in which people imagine their social existence, how they fit together with others, how things go on between them and their fellows, the expectations that are normally met, and the deeper normative notions and images that underlie these expectations (Taylor, 2003, p. 106).

According to this theory, the role of collective imagination is to enable the development and legitimization of practices and routines through the creation of common understandings of what is or is not acceptable, desirable, or even thinkable. However, Taylor's theory does not account for the increasingly important role of technology in the collective imagination. Since technology mediates so many aspects and routines of the everyday and mundane, accounting for it in the collective imagination has been grounds for developing the concept of sociotechnical imaginaries.

In this paper, we employ the theoretical perspective of sociotechnical imaginaries as developed by Sheila Jasanoff. According to Jasanoff, sociotechnical imaginaries are “collectively held and performed visions of desirable futures” (2015, p. 28). In other words, the way we imagine the future impacts our practices through the ways we prepare for it. Moreover, visions of undesirable futures can signify the emergence of competing sociotechnical imaginaries between different groups. Jasanoff explains that sociotechnical imaginaries are “animated by forms of social life and social order attainable through, and supportive of, advances in science and

technology” (2015, p. 28). According to this view, dramatic technological changes, such as those in the 20th and 21st centuries, have an impact on social life and social order. This is not a technological determinist perspective; rather, it is situated in the tradition of Marshall McLuhan as articulated by Culkin (1967, p. 70): “we shape our tools, and thereafter our tools shape us”. In addition to being mutually influenced and mediated by technology, sociotechnical imaginaries are important for how practices and technologies take shape.

3 IMAGINING THE FUTURE OF BIG DATA

When attempting to define big data beyond the “three V’s”, different authors mention scalability, frequent or continuous generation of data as well as methods used to process and analyze it (boyd & Crawford, 2012; Kitchin, 2014; Secundo et al., 2017). One way to think about the phenomenon is suggested by Mayer-Schönberger and Cukier (2013) who refer to big data as:

[T]hings one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more (p. 6).

The changes in relationships they mention indicate how deeply big data is embedded in societal and cultural contexts. Women have played a central role in developing a critical understanding of big data, Women of Color and nonbinary/trans women in particular. Women have described the social consequences of big data their racial implications and how they reinforce existing inequalities. They pioneered research on the use of big data for algorithmic decision making that is often impossible to audit and correct when blunt mistakes are made, and inequalities perpetuated. Some examples of their contributions concern big data reinforcing social inequalities in education and work force (O’Neil, 2017) and over policing (Benjamin, 2020). Furthermore, women of color have described how training datasets tend to overwhelmingly represent lighter-skinned subjects (Boulamwini and Gebru, 2018) and how digital spaces can be sites of racialization (Nakamura, 2013). These and other contributions enriched the understanding of big data as a technical as well as a social phenomenon.

Since the start of the 2010s, big data as an imaginary has been at the center of numerous discussions of the future. Considering the lack of consensus regarding how the concept should be defined, the way big data is imagined plays a significant role in how it figures in society and what kinds of futures are implied. Utopian visions of big data emerged quickly in the early 2010s. Big data seemed to offer a promise of new forms of rational planning that could also provide solutions to societal challenges such as climate change. Resources could be used with much more

efficiency with the continued development of social media as well as smart homes and smart cities. All such developments depended on a combination of big data and artificial intelligence.

Within critical scholarship, discussions aligning with the dystopian rhetoric emerged within numerous disciplines and interdisciplinary fields. The commercialization of online space during the early 2000s resulted in a handful of companies seizing control of massive amounts of data that could be relatively freely combined into valuable prediction products (see for example McNamee, 2019; Zuboff, 2019; Cheney-Lippold, 2019). The sheer volume of available data signals the shaky promise of objectivity, truth and accuracy (boyd & Crawford, 2012). Nonetheless, authors such as McNamee (2019) and Zuboff (2019) explore how Big Tech and Silicon Valley shape narratives and thus the impact that they can have on how futures are shaped.

A key concept used within the discussions of big data is datafication. It is a concept that explains how social action is translated into data that, in turn, makes accessing, understanding, and monitoring human behavior possible (van Dijck, 2014; Mayer Schoenberger & Cukier, 2013). The concept of datafication became another buzzword within the IT industry, carrying positive connotations. The mathematical understanding of the concept was launched by Mayer Schoenberger and Cukier (2013), who emphasized the aspect of quantification of human behavior and contextual information. They were among the first within the big data context to suggest that human behavior can be treated mathematically and thus open the possibility of collection and analyses of various forms of data. Scholars have also critically discussed the potential of datafied mechanisms and automation to perpetuate inequality and discrimination (e.g., Benjamin, 2020; Eubanks, 2019; Noble, 2018; O’Neil, 2017). In this article, we position discussions about datafication as part of big data imaginaries.

In a sense, big data appears as the ultimate form of the sociotechnical imaginary. It conjures up a future in which planning of automated and human practices can be organized for optimal efficiency in basically any realm of society. Indeed, the sociotechnical imaginaries approach has previously been used in critical studies of unfolding technologies related to big data. Reviewing such studies, a remarkable pattern can be discerned in which the imaginaries of Big Tech have become so powerful and uncontested that they are appropriated by the very people that are, seemingly, exploited.

Such research shows how users have come to accept profiling as a natural part of the online experience. Lupton (2020) explores how people imagine and talk about personal data profiling on social media. She defines the profiles created by collecting data generated through user participation in online life as “data personas”.

Lupton uses this term to provoke a conversation about which data are collected, to whose benefit, as well as the perceived effect they have on her participants.

Corporations such as Google and Meta (formerly Facebook) have an advertisement-driven business model based on big data. The viability of this business model is dependent on minimizing the risks associated with public backlash and government regulation. To that end, it becomes necessary for these companies to position data profiling as a public good, perhaps even as a public service. One study found that even individuals with a sophisticated understanding of datafication and surveillance described the trade-off of control over personal data in positive terms (Sörum & Fuentes, 2022). The prioritization of receiving information in a seamless manner is closely connected with accepting datafication and framing it as positive. Furthermore, in a study where participants received guidance about developing critical consciousness toward digital futures, Markham (2020) found that it was difficult for the participants to envision alternatives. These results demonstrate how deeply ingrained such imaginaries can be, especially when established by big corporations that invest heavily in creating and establishing standards for the online practices of the future. These imaginaries can be called utopian as they identify diverse problems of contemporary life and present technological workarounds. In this way, Big Tech companies take on roles traditionally reserved for public institutions and governing bodies, increasingly dominating the imaginative power for various groups in society (Mager & Katzenbach, 2020).

In many other ways, the role of the policies that regulate Big Tech is heavily influenced by the utopian visions produced within Big Tech. Trying to understand how technology can be reimagined has prompted studies of smart cities (Deitz et al., 2021; Sadowski & Bendor, 2018; Sandeep, 2017); studies of agency in social media (Saariketo, 2020; 2014); search engines (Mager, 2016); and algorithms (Kazansky & Milan, 2021; Bucher, 2017). In the following section, we look at how big data can and has been reimagined within critical feminist perspectives.

3.1 Feminist critical discussions

Feminist thinkers and scholars have developed perspectives that place structures and dynamics of power at the center of their attention. A tradition of studying intricate and ubiquitous oppression equipped feminist scholars with tools and theories that inspect and illuminate power dynamics, thus making it possible to criticize systems that support discriminatory practices and inequality in power distribution.

The introduction of intersectionality to mainstream feminisms reflects aspirations to advocate for the importance of localities and minorities (Daniels,

2015; Carbin & Edenheim, 2013; Davis, 2008). Intersectionality was introduced by legal scholar and civil rights advocate Kimberlé Crenshaw. She described the specific situation of Black women facing discrimination both by race and gender (Crenshaw, 1991). These compounded experiences of discrimination were situated at the intersection of racial and gendered identities, and as such are particularly difficult to advocate for within a system that favors whiteness and maleness. Appropriation of the term intersectionality has been controversial, the main criticism being that the use of intersectionality among feminist scholars has become increasingly vague and wide in scope, reducing its potential as a critical tool (Carbin & Edenheim, 2013). Even so, intersectionality remains a significant focus among feminist scholars, as many established norms tie into perspectives that were previously overlooked in (white) feminisms (Lykke, 2020; hooks, 2000). In digital landscapes, intersectionality implies opening to, and actively drawing from, indigenous and postcolonial research (Odumosu, 2020), LGBTQ+ research (Baucom, 2018; Drabinski, 2013; Huffer, 2013), and critical race studies (Knight Steele, 2021; Love, 2018), among others.

Feminisms have long operated between utopian and dystopian visions of the meaning that emerging technologies can have for women (Wajcman, 2004). On one hand, digital technologies are seen to enable spaces of freedom from gender norms and old social relations. On the other, the dramatic gender differences in access to and control over new technologies reiterates old power disbalances at scale. Women as innovators of technology have historically been overlooked and had their contributions erased (D'Ignazio & Klein, 2020a; Perez, 2019; Onouha, 2016). Feminist discourses on these topics address the issues of seeking equal representation in data as well as concerns about excessive surveillance and threats against privacy caused by such representation (Dubrofsky & Magnet, 2015). Simultaneously, feminist interest in topics of visibility and representation within datafied systems – in a world where maleness is the default – predates the current upscaling of datafication.

In the context of contributing to critical perspectives on data, feminist approaches such as standpoint epistemology (Harding, 1986; 1991) and situated knowledge (Haraway, 1998) emerge as useful tools for scrutinizing the notions of data as inherently objective, reliable, and trustworthy. Haraway and Harding are influential in the feminist refutation of absolute objectivity, i.e., “a view from nowhere” that is often perpetuated by removing the contexts in which data are created and exist. This critical tradition frames data as situated and embodied, not isolated from the circumstances in which they were produced. Feminist critical perspectives also highlight the processes that data underwent since their creation and the circumstances under which they are employed.

The application of feminist discourse to a data-driven society can help interrogate and challenge power relations (D'Ignazio & Klein, 2020a) and broaden the understanding of the challenges they engender. According to boyd and Crawford (2012), the mythology of big data is a significant part of this socio-technical phenomenon, resulting in “the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy” (p. 663). Indeed, the intentional mystification of big data can hinder critical inquiry by obfuscating the understanding that data are “not an autonomous force or a unidimensional technical fix”; rather that they are shaped by “social and material factors, including social institutions and technologies” (Beaulieu & Leonelli, 2022, p. XV).

4 METHOD

In this article, we explore a collective set of dystopian imaginaries by comparing a corpus of articles that are here labeled data feminist with discourses developed in the European Union. In order to identify and highlight conceptions of big data and visions of the future in our corpora, we utilize the concept of sociotechnical imaginaries, or collectively held ideas about how things work, what is conceptualized as good or bad, and what constitutes desirable or undesirable futures.

We identified scholarly texts concerned with datafication that employ feminist tools and perspectives through searches in academic databases, discovery engines, and citation chaining. The initial searches were conducted in the Winter of 2020 and Spring 2021. At this stage, a set of 103 texts was collected. We utilized VOSViewer¹, a software for constructing and visualizing bibliometric networks, to identify clusters of related terms. The largest cluster formed around the term big data, drawing attention to the term as a central concept warranting further exploration. We then reviewed the 103 identified texts to assess their engagement with the concept of big data. By early 2022, when final texts were added, we narrowed down the corpus to 44 articles, commentaries and essays. A description of our search strategy is provided in Table 1.

We conducted a close reading of the texts in the resulting corpus to determine how big data figures in these documents, identify the concerns raised about it, and analyze the implications of these concerns for understanding and integration of big data into future visions. Our analysis encompassed searching for contexts in which the term appeared in the corpus and examining the words and metaphors used to

¹ <https://www.vosviewer.com/>

discuss and describe it. For organizing our notes and creating mind maps of instances of the term "big data" in the texts, we utilized Miro software². Much like the widespread adoption of the concept of datafication has been widely adopted for critical discussions (see above), data feminist authors have also adopted big data. We refer to this corpus as data feminism, with the understanding that data feminism is a larger discourse and an umbrella term that surpasses the materials identified in this study.

Table 1. Collection of the data feminist corpus

Outcome	Details																																				
Goal	Finding literature engaging in a meaningful way with both data and feminism																																				
Databases and, discovery and search engines searched	WoS, Scopus, ProQuest, Taylor&Francis, Springer, SAGE journals, EbscoHost, Emerald Insight, Google Scholar																																				
Search terms clusters	<table border="0"> <tr> <td>data feminism</td> <td>or</td> <td>surveillance studies</td> </tr> <tr> <td>feminism</td> <td></td> <td>data</td> </tr> <tr> <td>critical data studies</td> <td></td> <td>feminism</td> </tr> <tr> <td>data studies</td> <td></td> <td>data science</td> </tr> <tr> <td>data science</td> <td></td> <td>gender</td> </tr> <tr> <td>data activism/t</td> <td></td> <td>critical race studies</td> </tr> <tr> <td>data ethics</td> <td></td> <td>data studies</td> </tr> <tr> <td>datafication</td> <td></td> <td></td> </tr> <tr> <td>algorithm(-s/-ic)</td> <td></td> <td></td> </tr> <tr> <td>big data</td> <td></td> <td></td> </tr> <tr> <td>artificial intelligence</td> <td></td> <td></td> </tr> <tr> <td>machine learning</td> <td></td> <td></td> </tr> </table>	data feminism	or	surveillance studies	feminism		data	critical data studies		feminism	data studies		data science	data science		gender	data activism/t		critical race studies	data ethics		data studies	datafication			algorithm(-s/-ic)			big data			artificial intelligence			machine learning		
data feminism	or	surveillance studies																																			
feminism		data																																			
critical data studies		feminism																																			
data studies		data science																																			
data science		gender																																			
data activism/t		critical race studies																																			
data ethics		data studies																																			
datafication																																					
algorithm(-s/-ic)																																					
big data																																					
artificial intelligence																																					
machine learning																																					
Additional strategies	Citation chaining																																				
Final corpus	103 texts collected, 44 engaging meaningfully with the concept of big data																																				

Following the construction of our data feminist corpus, we compared the results of our analysis of the corpus to the results of a 2018 study of sociotechnical imaginaries of big data in European Commission policies (Rieder, 2018). We created a table of Rieder's findings and complemented those with close readings of three additional texts published since the 2018 study. In the final stage, we surveyed and analyzed

² <https://miro.com/>

policy documents of the European Commission for mentions and considerations of minorities, gender, and other traces of intersectional and (data) feminist issues. Jasanoff and Kim (2009) have positioned textual criticism as a crucial method for the construction of imaginaries, and comparing these corpora aided us in teasing out sociotechnical imaginaries of big data. Overall, the study used a combination of bibliometrics, data visualization, and close reading techniques to analyze and understand the literature related to feminism and data studies.

4.1 Comparison as a method

Comparison is a common research method employed across disciplines, such as history, anthropology, and literary studies. In its most elementary form, comparison refers to a systematic process of contrasting two or more cases to each other, enabling the exploration of parallels and differences among them (Azarian, 2011). Comparison has the potential to elucidate nuances that are difficult to identify in conventional single case studies and to explore assumed knowledge and perspectives within the different cases. As a prevalent method in studies investigating sociotechnical imaginaries, comparison helps to identify the content and contours of those imaginaries by highlighting and contrasting the epistemic and ethical assumptions that are situated and particular (Jasanoff, 2009). Studies that implement comparison as a method to investigate sociotechnical imaginaries are numerous. One study conducted a comparison of cross-national science and technology policies (Jasanoff et al., 2007). Another (Levenda et al., 2019) focused on sociotechnical imaginaries of governance in energy innovations, comparing the cases of Portland, OR and Phoenix, AZ in the United States. Most recently, a comparative analysis of US and EU data governance and their respective sociotechnical imaginaries of personal digital data was published (Guay and Birch, 2022). Sociotechnical imaginaries are embedded with normative implications. Jasanoff and Kim (2009) state that descriptions of potentially attainable futures tend to become prescriptive and positioned as futures that *ought to* be attained. According to Jasanoff (2019), comparison helps tease out sociotechnical imaginaries and can reveal underlying assumptions and normative commitments by identifying commonalities and differences between them.

Comparison helps sketch out different sociotechnical imaginaries or visions of the future that depend on understanding existing technologies as well as future technological innovation. Comparing two corpora of texts can imply variations in terms of affiliations, types of authorship, and ways of achieving legitimacy. In our case, data feminist texts are usually affiliated with one or more universities and attributed directly to the individuals who wrote them, often by a single author. By contrast, the Commission's policies represent the work and reflect the values of the

European Commission. Keeping in mind instances of authorship transparency issues (Nelhans & Nolin, 2022), these policies are published and legitimized under the name of the European Commission. We chose these two corpora to compare in order to contrast assumptions in data feminist and European Commission corpora. This juxtaposition highlights the sociotechnical imaginaries within data feminism, and those that arguably reflect a significant influence on the formation of big data imaginaries in Europe, particularly stemming from the European Commission.

Comparing one corpus of texts to another, using analysis performed by another author, means that this article relies heavily on the analysis conducted by Rieder in 2018. Perhaps the most consequential difference between the two corpora is that of genre. Moreover, when comparing corpora, disparities in underlying assumptions and taken-for-granted aspects become apparent. For example, having been written for different audiences with different perspectives, the studied texts assume different levels of *a priori* knowledge of certain concepts. Furthermore, their perspectives and intended audience differ. The intended audiences further influence the actors that the texts address and that are deemed central.

5 RESULTS

We found that data feminist texts and European Commission policies diverge in the ways in which they frame, understand, and discuss big data. European Commission documents identify and explain interpretations of socioeconomic and political realities to provide guidelines for policy. Such documents are designed to create broad-scale changes across the European Union. These are, therefore, powerful utopian imaginaries, situated at the core of a centralized decision-making system. Being a policy-making body and attempting to work in the interest of citizens of the European Union, the perspective of the European Commission is decidedly top-down.

Contrarily, authors who write with a data feminist lens take on *bottom-up* perspectives, with a focus on local communities and knowledge, as well as the protection of vulnerable members. Local minority groups and localities are central within data feminist texts, whereas legal bodies, businesses, and governments are the actors in focus for the European Commission. If the utopian imaginaries are, therefore, all-inclusive and symmetrical, all citizens across the 27 member states are to be treated in the same way by emerging algorithms. Our own, as well as Rieder's (2018), readings of European Commission documents indicate that citizens are imagined as strong and healthy and flexible enough to adapt to change. Interestingly, class, race, age, and ability are barely mentioned in relation to big data. Data feminism focuses on the marginalized and the vulnerable in society,

those that are most likely to suffer when citizens are classified according to various big data plans.

The academic genre of data feminism presupposes a critical and investigative approach. Differences in these two contexts have implications for their perspectives and the way language is employed. Feminist writing is not intended to be prescriptive, unlike the policy documents of the European Commission. Furthermore, the utilization of feminist tools contributes to the critique and destabilization of dominant narratives, particularly where they reproduce problematic hierarchies and enable or recreate discriminatory practices.

Table 2 describes the result of the encoding of the differences between points of the dystopian and utopian sociotechnical imaginaries. It contains the results of close reading data feminist and European Commission texts, as well as our analysis of Rieder's (2018) findings.

Table 2. Encoding of main differences in data feminist texts and European Commission policy texts

	Data feminism	European Commission
<i>Main focus</i>	Groups pushed to the margin	Centralized steering and control
<i>Level of focus</i>	Micro, meso and macro perspectives	Macro perspective
<i>Argued effect of big data</i>	Increased exploitation	Improved economy
<i>View of citizens</i>	Citizens are vulnerable	Citizens are technology oriented and flexible
<i>View of data</i>	Data as process	Data as resource
<i>Ambition of texts</i>	Destabilizing dominating narratives	Stabilizing dominating narratives
<i>Effect on society</i>	Big data leads to marginalization	Big data as a tool for a better society
<i>Main characteristics of big data</i>	Big data as domination and violence	Big data as new oil, goldmine, and magic material
<i>View of classification systems</i>	Basic classification system is biased	Data is neutral and raw
<i>View of collection and processing of data</i>	Collection and processing of data is biased	Data is the lifeblood of economic development
<i>View of the paradox of exposure</i>	Dual problems of the paradox of exposure	There is no problem of exposure

<i>View of flexibility of classification systems</i>	Concepts of classification systems will age	Classification is made on raw data, need not be revised
<i>View of privileging of classification systems</i>	Classification systems privilege coding of the majority society	Classification is made on raw data, total objectivity
<i>Big data and gender</i>	Big data underprivileges women	Implicitly, technologies are masculine, but this has no consequences
<i>Public and private</i>	Categories of public and private need to be analyzed separately	Categories of public and private are collapsed
<i>Big data as power</i>	Instrumental, structural, and symbolic power	Economic power
<i>What big data strengthens</i>	Strengthened imperialism	Strengthened European Union
<i>Control of data flows</i>	Citizens have no control of data flows	Growing corporations have control of data flows
<i>Purpose of ordering people</i>	Ordering people for domination and exploitation	Ordering people for economic growth
<i>Policy addressed at</i>	Minorities and vulnerable groups	National Governments, Silicon Valley
<i>Timescale of development</i>	Need for slow developments to avoid multiple problems	Urgency to become early adopters
<i>Scale of problems identified</i>	Major problems	Minor problems
<i>Data</i>	Embodied	Separation of data and bodies

The results of the encoding will be discussed further below. We first present the motifs and phrases that occur within each of the corpora followed by each of the three prominent themes that occur across the data feminist texts. The first, titled *Visibility and representation*, focuses on discussions about big data and minorities and the relationship between the private and the public. The second is *Power and domination*, discussing the two in relation to datafication. The final focuses on the differences between those corpora employing the rhetoric of big data as a resource and those that challenge it. The final theme is *Exploitation and vulnerability*.

5.1 Motifs and phrases

In data feminist texts, big data is described in various ways, with the consensus being that it is an inevitable part of the future. As long as big data is tightly wrapped into dystopian imaginaries, a problematic development seems to be inevitable. That said, there are also data feminists who are attached to utopian imaginaries. Proximity to natural sciences and technology is correlated to positive visions of big data in some data feminist materials (e.g., Diaz Martinez et al., 2020; Vaitla et al., 2020; Larrondo et al., 2019).

Some authors in our data feminist corpus assert that big data projects can be transformed and therefore serve to create attractive future outcomes. Vaitla et al. (2020, p. 18) argue that big data “can have a profound influence on improving the lives of all”, adding that it can only happen if “it is intentionally managed as a vehicle for equity and empowerment”. Hong (2016) claims that big data “offer a chance to open up new opportunities to produce a more egalitarian society”. Other authors refrain from making such claims and contributing to similar discourses. They do not describe a future rid of, or with less, data extraction either. Instead, some authors explore the representational limits of data in the context of a societal system in which dominating societal norms already exist. They suggest *Queering data* (Giesecking, 2018) and *Black data* (Rueberg & Ruelos, 2020) as two respective methods of centering perspectives of minorities in datafied systems and data narratives. Foucault Welles (2014, p.1) suggests searching for outliers and minorities within big datasets and putting them in focus but points out that “a large dataset quickly becomes small when you focus on a minority population”.

Themes of marginalization and bias occur throughout the data feminist corpus that we surveyed. This is reflected in the phrases used in relation to big data, including *bias* (Diaz Martinez et al., 2020), *violence* (Hong, 2016), and *oppression* (Giesecking, 2018). Big data is described in the corpus as a part of a positivist, colonialist, imperialist, and capitalist paradigm (e.g., in Hoffman, 2020; Corple & Linabary, 2019; Leurs, 2017).

In his 2018 study of the sociotechnical imaginaries of the European Commission, Rieder finds the utilization of stable phrases and a limited set of discursive elements recurring across the European Commission’s policy documents. An additional reading of the policies published since 2018 that we conducted for this study corroborates that Rieder’s findings still hold true. The variety of phrases represents different ways of conceptualizing big data. The phrases that occurred in the European Commission documents included *new oil*, *gold mine*, *game changer*, and *magic material* (Rieder, 2018). These are all highly positive metaphors that link big data with economic wealth.

Additionally, this way of framing big data feeds into the discourses of data as *raw* and isolated technical artifacts, as opposed to the data feminist understanding of data as a *process* (Cruz, 2020) or as a set of ongoing negotiations. Data as a resource, such as oil or gold, makes data disembodied, removed from the subjects that produce data, or are targeted for data collection. On the other hand, data as a process implies that humans, together with the algorithms they design, make a multitude of decisions in shaping what is characterized as “raw data”. The implication is that data-driven decision-making, whether done by humans or machines, does not build upon neutral data. Rather, it is a matter of a sequence of (potentially biased) decisions being made about the various perspectives and ideologies that govern processes when creating the *appearance* of neutral data.

Other important concepts used by the European Commission are *key asset*, “motor and foundation of the future economy”, and the *lifeblood of digital markets* (Rieder, 2018). Similar rhetoric persists in the policies of the Commission: data is labeled as the *lifeblood of economic development* (EC, 2020) and as an *essential resource* (EC, 2022a). These phrases point in different directions, with data being portrayed either as a stepstone upon which prosperity can be reached, or an entity that enables and fuels successful systems and practices. However, all the phrases employed by the European Commission have a strong connection to capitalist vocabulary. Imaginaries of big data in the European Commission hinge upon three storylines: big data as the cornerstone of a thriving data economy; big data as a way to transform and improve public services; and big data as a tool for evidence-based decision-making (Rieder, 2018). Prominent themes in data feminism are *visibility and representation*, *power and domination*, and *vulnerability and exploitation*.

5.2 Visibility and representation

Topics of under- and misrepresentation, as well as data bias, have been extensively covered in the data feminist corpus. The topics of silencing voices of minorities based on gender, sexual orientation, and race are characteristic of intersectional data feminism and occur across the material (e.g., Cruz et al., 2020; Ruberg & Reilos, 2020; Giesecking, 2018). Ruberg and Ruelos (2020) state that data play a role in valuing or devaluing the identities and experiences of marginalized groups. This *paradox of exposure* (D’Ignazio & Klein, 2020b) is evidenced in simultaneous advocacy for better representation and arguing against collecting data on already vulnerable and marginalized populations. The argument is that classification systems precede data collection and processing. Therefore, critical scrutiny of the ideologies underpinning the classification system itself is needed, particularly as they concern marginalized groups. Ideally, these groups should be considered during the construction of classification systems to avoid structural bias within the

big data that is created. However, visibility increases exposure to over-policing, surveillance, and discrimination (Favaretto et al., 2019; Corple & Linabary, 2019). As Snapp et al. (2016, p. 135) noted, “[t]he right for participation, and thus representation in data, science, and policy, is often understood as conflicting with the right for protection, that is, safety from the disclosure of a marginalized orientation or identity”. However, O’Neil (2017) stresses that certain big data systems can be specifically designed to exploit vulnerable people, by steering them toward questionable universities and other financial scams. According to dystopian imaginaries, the complex problems of the paradox of exposure are likely to increase in the future, as today’s big data sets are unlikely to be redesigned and will only grow in size and scope.

Bogers et al. (2020) discuss online representations of pregnancy as featuring predominantly white, able-bodied, CIS women, while overlooking non-white, disabled and LGBTQ+ people. This is an example of a taken-for-granted classification system, implicitly constructed to ignore certain groups. For Hong, a woman who is “allowed” visibility must be “white, pleasant, and subversive only to the extent that she does not disturb the expectations of normal feminine behavior” (2016, p. 3). Extending this line of thought, classification systems can be built upon conservative notions of core concepts such as gender and will therefore be unable to change over time by, for instance, considering non-binary gender identities. Within dystopian imaginaries, big data resources age in size only, becoming increasingly larger, while the core ideas of classification remain stagnant.

Overrepresentation is also discussed by Bogers et al. (2020) in what they call *power shadows*, or the hypervisibility of a selected identity or group that casts a shadow rendering other identities or groups less visible. That numerous classification systems of big data do not typically highlight the experiences of minoritized groups is an issue with no simple solution. Big data does not account for all people, and as currently often utilized, is not good at representing the experiences of minorities (Giesecking, 2018). Categorization and datafication of fluid identities and minorities is therefore an issue highlighted by data feminism in relation to big data. As Giesecking argued, datasets representing queer people may never be *big enough* in scale (2018, p. 150). Moreover, analytical tools are not built to *see* data concerning queer lives, which means that existing data are not immediately legible when they are available (Ruberg and Ruelos, 2020). These dystopian imaginaries signal a future in which the traits of individuals and minorities are overpowered and disappear within power shadows.

Regarding the ways machines can both see and overlook people, Agostinho (2019) notes that optical visual metaphors enhance the role of the senses and heighten the conviction of knowability. According to Agostinho, sensory imaginaries of big data as microscopes or magnifiers perpetuate the misconceptions

of inherent objectivity and reliability. This perspective can be applied to both the European Commission policy documents and our data feminist corpus. Similarly, Giesecking (2018) points out the inherent masculine undertones of equating the *bigness* of data with legitimacy. Data feminist texts adopted these concepts to enter the discussion, but few have explicitly characterized the implications of using them. The concepts utilized by the European Commission, as analyzed by Rieder (2018), imply that the adoption of masculine terminology also draws from capitalist and economic narratives. Indeed, before programming became a male-dominated profession, the title *computer* had little pedigree, and was predominantly used to describe women who performed computations on early computers. In this instantiation, engineers were considered to do the *real* work. This changed as programming became more prestigious, and alongside the change in terminology, the number of women in computing dropped as its perceived importance increased. The concerted effort to align technology with masculinity ultimately resulted in the underrepresentation of women in computing (Van Oost, 2000; Terras & Nyhan, 2016). Several authors explore this phenomenon, both in literature on the history of women in computing and in analyses of the perceived pedigree of professions in relation to gender distribution within them.

In the majority of the European Commission policies and the data feminist corpus we examined, the notion of privacy and what constitutes personal data is implied, but not explicitly discussed. In her study of privacy in the context of the digital economy, Weinberg (2017, p.16) writes: “[w]ithout critically engaging the underlying assumptions of the categories of public and private, the forms of political resistance against data exploitation remain tethered to presuppositions about the liberal democratic subject”. Weinberg (2017, p. 11) also argues that corporations and governments have found ways to aggregate data by using the “fiction of the sovereign subject to resist surveillance” in ways that are technically legal, and by fragmenting the information about individuals and combining them into mass collections of data. This imaginary warns of a future where the boundary between the private and the public has been renegotiated, blurred, and provided with false legitimation.

5.3 Power and domination

Critical data perspectives discuss data as a form of power. The data feminist materials in our corpus actively contribute to that discussion. Hoffman (2020) investigates instrumental, structural, and symbolic power and how big data can underwrite and manipulate the informational bases of other forms of power. Access to data is one form of power, alongside the power to decide what is or is not included in datasets, and the extent to which certain topics are represented. Bogers et al. (2020) and

Cooky et al. (2020) are among those authors that identify the absence of women in data streams as an issue of power. Leurs (2017) equates power differences grounded in gender and race with the imperialistic and sexist design of technological systems. Thompson (2020) in particular reflects on the kinds of power given to “factual” numerical data that is intended to represent all people, despite its underrepresentation of women, and how this affects governmental, educational, and legal systems. Foucault (1980) famously discussed the inextricable connection between knowledge and structures of power. Many feminists have been inspired by Foucault and go further than associating power with knowledge. Rather, data feminists conceptualize data, the bedrock of knowledge, as associated with power structures.

Situating data used across various systems as representations of real-world conditions of power means that data bias has material consequences within and outside the digital. Cooky et al. (2020) pinpoint access to data which privileges powerful corporate bodies as an example of inequality in power dynamics. Moreover, Suarez and Gonzalo (2019) note that companies and corporations have control over data flows and benefit from them, while citizens whose data are collected know little of what is collected and to which end. These imaginaries warn of a future (and a present) where big data is used to shift power from citizens to corporations. Classification systems utilizing big data are developed so quickly and with such complex consequences that governments simply cannot design appropriate legislation fast enough. As such, the imaginaries of Big Tech drive the transformation of society.

Domination of and through data is a recurring motif in the data feminist corpus (e.g., McQuillan, 2016; Luka & Millette, 2018). Suarez and Gonzalo (2019) call it *data domination*; conceptualizing data as a tool for those already in power to increase or solidify their dominance over vulnerable groups. It is also implied that big data serves to reinforce existing structures of dominance and plays a role in building new structures of oppression. In our corpus, data feminists are concerned with power and disruption of solidified power structures, especially where they (re)inforce discriminative practices. Indeed, big data is repeatedly related to exploitation, violence and power that are exerted over groups of people (e.g., Weinberg, 2017; Giesecking, 2018; Leurs, 2017). Some data feminist authors also claim that big data is capable of inflicting and facilitating harm and violence (e.g., Hoffmann, 2020; Cooky et al., 2018; Luka & Millette, 2018). According to Hoffman (2020, p.4), these harms reproduce “racist, sexist, and other norms and stereotypes that position some people as subordinate, inferior, or irredeemably ‘other’”. Some of the explicit forms in which harm is inflicted through big data include surveillance, harassment, commodification, privacy issues and abuse, most notably of minorities (e.g., in Cooky et al., 2020; Corple & Linabary, 2020;

Hoffmann, 2020). For Hoffman (2020), *data violence* occurs when people are labeled and classified, and can be material, symbolic, and representational.

Imbedded in the topics of power, violence, and domination through technology are discussions about capitalism and colonialism in relation to datafication. Whereas data feminist texts in our corpus direct attention to complex social issues that would require significant time for key actors to negotiate and address, by contrast, European Commission policy texts denote a sense of urgency to capitalize on big data (Rieder, 2018). These policy texts argue that the EU needs to collectively act as early adopters of new innovations developed through big data practices. Otherwise, the next wave of dominating digital corporations will be US-based, just like Apple, Google, Amazon, Microsoft, and Meta. Framing big data as a resource reiterates the capitalist agenda and a narrative of data as a raw material that needs to be exploited. It should be noted that the European Union has made significant attempts at regulating security, privacy, and surveillance online. Such inroads are made through legislation, primarily the General Data Protection Regulation (GDPR) from 2018, the Digital Markets Act (DMA) and the Digital Service Act (DSA) from 2022.

Our data feminist corpus also draws parallels between the corporate capture and selling of user data and the capture of territory to extract value and exert control (e.g., in Corple & Linabary, 2020; McQuillan, 2016). Giesecking (2018) notes that big data cannot be disconnected from historical examples of domination and exploitation, without replicating them in the digital context. For the same author, the perceived objectivity and authority of big data is derived from “masculinist, racist, colonialist, ableist, and heteronormative structural oppressions” (2018, p. 150). Denying the context in which data are collected and organized and the oppressions that they reflect allows for the perpetuation of the same destructive practices.

5.4 Exploitation and vulnerability

Both our and Rieder’s analyses suggest that within the European Commission imaginaries, big data is described as fostering a large economic potential that needs to be “tapped into” (EC, 2020, p. 1) and “exploited” (Rieder, 2018, p. 5; EC, 2022b, p. 1; p. 47). Rieder states that the broader European imaginary of a technological race with the United States has a large influence on the formation of big data imaginaries of the European Commission. This broad imaginary further influenced the perspective of the Commission and the actors that figure in policies. There is a sense of urgency, a race to exploit data before someone else does it. For the Commission, balancing power on an international scale is of great importance, and seizing opportunities as well as protecting European citizens is a priority. A sense of missing out on the first

wave of digital innovation dominated by Big Tech companies based in Silicon Valley fuels the urge to jump on the figurative big data train (Rieder, 2018; Burgelman et al., 2010). Because of the urge to tap into the resource that is big data, European Commission policies situate companies in Silicon Valley and national governments as key players. This is in stark contrast to data feminist texts, which place minority groups and local communities at the center of discussion about big data. Rieder found that in the policies of the European Commission, the benefits are “believed to outweigh any potential harm”, and detrimental effects are considered exceptions to that rule (2018, p. 6). In the sociotechnical imaginaries of the European Commission, high velocity of data collection is equated to increased technological advances, economic prosperity, as well as the identification and mitigation of societal issues. Data feminist authors in our corpus are particularly concerned with how big data mining renders people vulnerable (Favaretto et al., 2019; Leurs, 2017) and further oppresses the already marginalized (Giesecking, 2018). The bottom-up approach of data feminism highlights detaching data from bodies as a structural feature. Envisioning data as embodied and situated challenges the discourses about the need for exploitation that are identified as part of a “gold rush” mentality (Luka & Millette, 2018). Centering the bodies and locations from which knowledge arises (Corple & Linabary, 2020) helps to counter big data neo-positivism. Across the corpus, data are envisioned as introducing or reinforcing various types of harm to vulnerable groups. As such, there is a notion of resistance against (D’Ignazio & Klein 2020b; Giesecking, 2018) and refusal of (Linabary et al., 2020) big data oppression.

The difference in perspectives between the European Commission and data feminist corpora are visible in the actors that are considered, the language that is employed, and how power is perceived and seen to be enacted. Within data feminism, exploitation through big data is repeatedly connected to colonialism (e.g., Cooky et al., 2020; Linabary et al., 2020; Dionne, 2019; McQuillan, 2017). Likewise, exploitation often denotes economic exploitation through datafication (e.g., Weinberg, 2017) but in several instances refers to the exploitation of free labor on social media (e.g., Corple & Linabary, 2020; Cooky et al., 2018). The focus on workplace optimization through big data has been shown to have negative consequences on the workloads for health workers (Dionne, 2019), women (Michailidou, 2018), and people of color (Cooky et al., 2018; Corple & Linabary et al., 2020; Michailidou, 2018). Some authors explore how the use of big data for research, particularly when utilizing social media, can make researchers complicit in inflicting harm (Corple & Linabary, 2020; Luka & Milette, 2018).

The struggle for the power to capture and shape political and social imaginaries influence the directions of technology development and common understandings of what constitutes attainable futures. Subsequently, the tension

between dominant and alternative imaginaries is ultimately the struggle to sketch out boundaries of current understandings of the possible and future (technological) developments.

6 DISCUSSION: UTOPIAN AND DYSTOPIAN IMAGINARIES OF BIG DATA

In our data feminist corpus, big data is understood in a variety of messy and fuzzy ways. It is described as susceptible to change and imagined as a process. Big data is not depicted as a fixed technical artifact; rather it is understood as determined by social contexts and decisions made during stages of collection and processing. The data feminist corpus highlights big data as created by someone, for someone and for a specific purpose, indicating an awareness of big data as a sociotechnical phenomenon. In these texts, the conceptualization of big data entails multiplicity, and as such, allows for different perspectives and interpretations. This multiplicity introduces nuance into sociotechnical imaginaries and decentering of the Big Tech narratives. Conversely, the European Commission corpus engages with big data as a thing that simply is and tends to overlook the implications big data for the society beyond its alleged potential to improve the economy. Describing data as a raw material overlooks the choices during data collection, the decision-making processes during integration into various systems, and the social consequences after integration. In other words, the social aspects of big data.

The data feminist corpus we analyzed is often aligned with what boyd and Crawford (2012) have called the dystopian rhetoric, problematizing power and highlighting the potential for bias and perpetuating discrimination. Dystopian rhetoric is therefore often collapsed with critical perspectives stemming from the feminist tradition. While data feminist texts do offer visions of optimistic futurities with big data, the focus is on critical inspection and questioning of big data-related systems and practices. The European Commission corpus is more consistent with the utopian rhetoric, visible in the use of metaphors like oil and raw material. Characterizing sociotechnical imaginaries of big data as dystopian and utopian sociotechnical imaginaries based on the type of rhetoric they predominantly employ helps contrast and compare the differences in corpora. However, it also collapses critical engagement with big data behind the dystopian rhetoric. In fact, critical perspectives are often lacking in the European Commission corpus, while sociotechnical imaginaries in the data feminist corpus are richer and more nuanced. Four themes separate these particular sets of sociotechnical imaginaries, addressed in the following text.

6.1 The concept of data

Discussions of data in the data feminist corpus are very rich. Numerous sub-themes can be noted regarding how data is collected, processed, and owned. In this aspect, the themes align with those most prominently described in critical data studies. Critical data studies cover topics of how datafied systems serve and privilege certain groups (O’Neil, 2017; Zuboff, 2019) while overlooking or discriminating against others (e.g., Benjamin, 2020; Eubanks, 2019, Noble, 2018). Data feminist texts add to the critical discourse by conceptualizing data as a process, embedded in politics, enabling discussions of datasets as created within a context and never fully representative of reality. Moreover, various strategies for teaching computers how to collect and categorize data are scrutinized within a firmly established feminist tradition.

The policy documents of the European Commission portray a far less nuanced notion of data. As argued by Rieder (2018), the concept is described as a resource that needs to be mined and thereafter can produce value as *the new oil*. Our own additional reading reaffirms this portrayal in more recent policies. Through this rhetoric, the European Commission upholds and perpetuates decidedly positive imaginaries of big data (Rieder, 2018). The language used in the policies is aimed at creating excitement by using phrases like *game-changer* and *magic material*. This mythology of big data (boyd & Crawford, 2012) creates a sense of urgency to exploit and capitalize on big data. Sociotechnical imaginaries of the European Commission rest upon the notion of a utopian future that will unfold with the mining and exploitation of big data.

Our study highlights the differences between the conceptualization of the relationship between personal data and big data in the corpora. An understanding of knowledge as situated gives feminist conceptualizations of big data a multiplicity in a way that is not present in the policy documents of the European Commission. However, it should be noted that in both data feminist and in European Commission documents, big data is often used as a shorthand to denote personal data. It is important to note that not all big data refers to personal information but avoiding explicit discussions of personal data altogether obscures the instances when it does have certain implications. While the problematic distinction between public and private spheres has been a longstanding discussion in feminist and data feminist discourse, these complexities are collapsed and simplified in the sociotechnical imaginaries of the European Commission. In the case of the European Commission and its characterization of big data as a resource, understanding personal information collected through datafication as a commodity would require consideration of compensation for citizens. In the imaginaries identified in the European Commission’s texts, citizens are technology-oriented and flexible,

separate from and wholly unaffected by the presence, ordering, and classification of data. Moreover, in the Commission's view, big data is a positive and stable fixture of European society. Thus, whatever data are available today, they will remain so tomorrow, making them ideal for exploitation and economic gains.

6.2 Micro-, meso- and macro perspectives

The first theme concerns how the perspective levels differ in the data feminism and European Commission corpora. We have demonstrated that the top-down perspective of the Commission differs in implications from the bottom-up perspective of data feminists. Building on this, we introduce different ways that macro-, meso-, and micro-perspectives are present within these texts.

Firstly, our analysis discloses that the European Commission works with the top-down macro perspective, while data feminism engages simultaneously with macro-, meso-, and micro-perspectives. This is particularly visible in the way that data feminists position access to and utilization of data as a power negotiation, and how they build upon the tradition of discussing patriarchal macrostructures. The macro perspective is present in the strong tradition of talking about patriarchal macrostructures.

Secondly, in keeping with a feminist tradition, there is a strong trend of investigating local experiences, which aligns with the micro perspective. Data feminist interest in how classification biases privilege the powerful, while underserving marginalized groups such as women, people of color, and LGBTQ+ communities. Even though texts in the data feminist corpus define individuals and groups as resilient and holding agency, they perceive big data as capable of inflicting harm and rendering some people vulnerable to exploitation and oppression. However, optimistic visions of big data futures occur in data feminist texts as well. While they argue that big data is inherently tied up with masculine rhetoric and needs to be reframed and redefined, data feminists also see big data as a persistent element in the imaginaries of the future.

Finally, meso perspectives are visible through an emphasis on how institutions engage with and provide support for vulnerable people. In feminist imaginaries of data futures, big data often figure as a facilitator and a vehicle for capitalist narratives and aspirations. The feminist focus on localities and identifying vulnerable groups within established structures gives rise to imaginaries of domination, violence, and exploitation through data. Dystopian data feminist imaginaries build on different categories of data allowing for rich, critical discussions of future outcomes. By comparison, the top-down perspectives of the European Commission are tried and found wanting.

6.3 Addressing majorities and minorities

Dystopian imaginaries build upon the ways in which minorities and vulnerable groups in society may fare as big data technologies are increasingly incorporated into everyday experiences and interactions. Because, as we argue, the European Commission builds policy with dominant communities in mind, citizens are positioned as robust and flexible. However, this focus on an imaginary majority arguably creates vulnerable groups in society whose little power or voice within institutions are likely to suffer the consequences of becoming underserved. The commodification of data takes on different shapes for dominant groups compared to minority groups. However, much of data feminism builds upon intersectionality. This allows for critical engagement with imaginaries of exploitation that raise questions of what big data does *to* people, as well as who is prioritized as a citizen and how that prioritization perpetuates the privilege and discrimination described by critical race and intersectional feminist theory. The topic of hypervisibility and subsequent over-policing of marginalized groups that is reiterated and exponentialized in datafied systems is additionally complicated in the context of feminist traditions that seek to enhance the visibility and representation of marginalized identities. Such engagement with dystopian imaginaries also provides rich perspectives that are missing in utopian imaginaries.

6.4 Economy and human values

Following an intersectional feminist tradition, data feminisms center people as the focus of big data narratives. This is demonstrated by the language used in our data feminist corpus, which highlights that women, nonbinary people, people of color, and LGBTQ+ people are particularly exposed to discrimination, exploitation, and violence through data and classification. From this perspective, inequalities, bias, and discrimination come to the forefront.

Conversely, the European Commission is primarily concerned with continuously molding the European Union into an IT superpower that can rival the many successes of the United States since the 1990s. Betting on the development of big data within the EU is therefore deemed desirable and even necessary. This is not contextualized as a political choice fraught with underlying discourses of power and domination in the policy documents, it is instead framed as a logical imperative to stimulate European economic growth.

Macroeconomic aspects are not the focus of the data feminist corpus to the same extent that they are for the European Commission. Instead, there is an emphasis on the private sphere of experience which privileges the importance of basic human values. Several texts in the corpus equate mining data as a resource for profit with colonial tendencies. By contrast, the language used in the European

Commission documents implies much less critical consideration of how, exactly, infiltrating data markets will result in universal benefits and prosperity.

7 CONCLUSION: SOCIOTECHNICAL IMAGINARIES OF BIG DATA

Data feminism entails a bottom-up perspective with a focus on local communities and experiences. As a result, it positions vulnerable groups and minorities as key actors in discussions of big data. In their position as policymakers, members of the European Commission employ a top-down perspective of power that prioritizes the protection of European Union citizens. For the Commission, Big Tech companies in Silicon Valley, national groups, and governments are central agents of interest, while data feminist texts often do not afford them the same degree of importance.

In data feminism, big data is a multiplicity of things, always fluid and under development, and understanding big data requires an understanding of the diverse underlying perspectives through which data are generated. For the European Commission, big data is something that can be used and reused, which implies a fixed entity and focus on openness and exploitation over variability and understanding data contexts.

Imaginarities of both data feminism and the European Commission feature futures inclusive of big data and the systems built upon it. However, while imaginaries of the European Commission reflect economic narratives and a top-down perspective, data feminist bottom-up perspectives highlight new forms of power and structural imbalance that signify potential for harm of distinct groups of people. For the Commission and its economic imperative, the focus is on sharing, generating, and exploiting data. That perspective reflects the position of power to govern and aims to increase the financial gain and leadership role of the European Union with the help of big data. Situating big data in this way supports the idea that citizens of the EU will only benefit from its increased presence in society. However, in our data feminist corpus, questioning who benefits from datafication and whose livelihoods are jeopardized figures at the forefront.

Imaginarities within data feminist texts are considerably richer than those identified within the European Commission policy documents. Imaginaries present in data feminism build upon a broad and inclusive view of different intersectional social perspectives (macro, meso, micro) that allow for recognition of a multitude of vulnerable groups, a highly sophisticated understanding of data, as well as a focus on human values over economic growth. Without conscientious and intentional integration of (data) feminist and other perspectives, dominant Big Tech imaginaries will prevail in the narratives of European policymakers. In this study, we focused on the sociotechnical imaginaries present in data feminism research. Future studies

could identify existing big data practices and projects that inhabit and act within data feminist perspectives, and how they work with, use, and resist datafication when necessary.

FUNDING STATEMENT AND ACKNOWLEDGMENTS

We want to thank Merisa Martinez for helpful comments and for bringing to our attention Melissa Terras and Julianne Nyhan's work on Father Busa's female punch card operatives. A big thank you to Maja Krtalić, Pieta Eklund and Marija Marčetić for taking the time to read this text, offer helpful advice and brainstorming session.

REFERENCES

- Agostinho, D. (2019). The optical unconscious of Big Data. *Big Data and Society* 6(1).
<https://doi.org/10.1177/2053951719826859>
- Azarian, R. (2011). Potentials and Limitations of Comparative Method in Social Science. *International Journal of Humanities and Social Science* 1(4), 113–125.
https://www.ijhssnet.com/journals/Vol._1_No._4;_April_2011/15.pdf
- Baucom, E. (2018). An exploration into archival descriptions of LGBTQ materials. *The American Archivist* 81(1), 65–83. <https://doi.org/10.17723/0360-9081-81.1.65>
- Beaulieu, A. & Leonelli, S. (2022). *Data and Society: A Critical Introduction*. London: Sage.
- Benjamin, R., (2020). *Race after technology: Abolitionist tools for the New Jim Code*. Cambridge, UK: MA Polity.
- Bogers, L., Niederer, S., Bardelli, F., & De Gaetano, C. (2020). Confronting bias in the online representation of pregnancy. *Journal of Research into New Media Technologies*. 26(5- 6), 1037–1059. <https://doi.org/10.1177/1354856520938606>
- Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.
- 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>
- Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society* 20(1), 30–44.
<https://doi.org/10.1080/1369118X.2016.1154086>
- Burgelman, J. C., Osimo, D., & Bogdanowicz, M. (2010). Science 2.0 (change will happen...). *First Monday* 15(7). <https://doi.org/10.5210/fm.v15i7.2961>
- Carbin, M. and Edenheim S. (2013). The intersectional turn in feminist theory: A dream of a common language? *European Journal of Women's Studies* 20(3), 233–248.
<https://doi.org/10.1177/1350506813484723>

- 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>
- Cheney-Lippold, J. (2019). *We are data: algorithms and the making of our digital selves*. Paperback edition. New York: New York University Press.
- Cooky, C., Linabary, J. R., & Corple, D. J. (2018). Navigating Big Data dilemmas: Feminist holistic reflexivity in social media research. *Big Data and Society* 5(2).
<https://doi.org/10.1177/2053951718807731>
- Corple, D. J., & Linabary, J. R. (2020). From data points to people: feminist situated ethics in online big data research. *International Journal of Social Research Methodology* 23(2), 155–168. <https://doi.org/10.1080/13645579.2019.1649832>
- Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review* 43(6), 1241–1299.
<https://doi.org/10.2307/1229039>
- Cruz, T. M. (2020). Perils of data-driven equity: Safety-net care and big data's elusive grasp on health inequality. *Big Data and Society* 7(1).
<https://doi.org/10.1177/2053951720928097>
- Culkin, S.J. (1968). A Churchman's Guide to Marshall McLuhan. *Religious Education* 63(6), 457–462. <https://doi.org/10.1080/0034408680630609>
- Daniels, J. (2015). The trouble with white feminism: Whiteness, Digital Feminism and the Intersectional Internet. In Umoja Noble, S., & Tynes, S.M. (eds) *The Intersectional Internet*. New York: Peter Lang. 41–60. Retrieved on 19 February 2023 from <https://archive.org/details/intersectionalin0000unse/page/n3/mode/2up>
- Davis, K. (2008). Intersectionality as buzzword: A sociology of science perspective on what makes a feminist theory successful. *Feminist Theory* 9(1), 67–85.
<https://doi.org/10.1177/1464700108086364>
- Diaz Martinez, C., Diaz Garcia, P. and Navarro Sustaeta, P. (2020). Hidden Gender Bias in Big Data as Revealed by Neural Networks: Man is to Woman as Work is to Mother? *Revista Española de Investigaciones Sociológicas* 172, 41–60.
<https://doi.org/10.5477/cis/reis.172.41>
- Dionne, E. (2019). Algorithmic Mediation, the Digital Era, and Healthcare Practices: A Feminist New Materialist Analysis. *Global Media Communication, Canadian Edition* 11(2), 49–65. http://gmj-canadianedition.ca/wp-content/uploads/2020/06/05_Dionne-Volume-11-issue-2-Final-.pdf
- D'Ignazio, C., & Klein, L. F. (2020a). *Data feminism*. Cambridge, MA: MIT Press.
- D'Ignazio, C., & Klein, L. F. (2020b). Seven intersectional feminist principles for equitable and actionable COVID-19 data. *Big Data and Society* 7(2).
<https://doi.org/10.1177/2053951720942544>
- Drabinski, E. (2013). *Queering the Catalog: Queer Theory and the Politics of Correction*. *Library Quarterly* 83(2), 94–111. <https://doi.org/10.1086/669547>

- Dubrofsky, R. E., & Magnet, S. E. (2015). *Feminist Surveillance Studies*. Durham, NC: Duke University Press.
- Eubanks, V. (2019). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: Picador.
- European Commission (2020). Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the regions: A European strategy for data. COM/2020/66. Retrieved on 19 February 2023 from <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>
- European Commission (2022a). Shaping Europe's digital future: A European Strategy for data. Retrieved on 19 February 2023 from <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>
- European Commission (2022b). Proposal for a regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act). COM/2022/68. Retrieved on 19 February 2023 from <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0068>
- Favaretto, M. F., De Clercq, E., & Elger, S. E. (2019). Big Data and discrimination: perils, promises and solutions. A systematic review. *Journal of Big Data* 6, 12. <https://doi.org/10.1186/s40537-019-0177-4>
- Foucault, M. (1988). *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*. New York: Random House.
- Foucault Welles, M. (2014). On minorities and outliers: The case for making Big Data small. *Big Data & Society* 1(1). <https://doi.org/10.1177/2053951714540613>
- Gieseking, J. J. (2018). Size Matters to Lesbians, Too: Queer Feminist Interventions into the Scale of Big Data. *The Professional Geographer* 70(1), 150–156. <https://doi.org/10.1080/00330124.2017.1326084>
- Gregory, A., & Halff, G. (2020). The damage done by big data-driven public relations. *Public Relations Review* 46(2). <https://doi.org/10.1016/j.pubrev.2020.101902>
- Guay, R. and Birch, K. (2022). A comparative analysis of data governance: Socio-technical imaginaries of digital personal data in the USA and EU (2008–2016). *Big Data and Society* 9(2). <https://doi.org/10.1177/20539517221112925>
- Guha, R. V. (2021, April 21). Data Commons: Making sustainability data accessible. [Blog Post] Google. Retrieved on 19 February 2023 from <https://blog.google/outreach-initiatives/sustainability/data-commons-sustainability/>
- Harding, S. (1986). *The Science Question in Feminism*. Ithaca: Cornell University Press.
- Harding, S. (1991). *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Ithaca: Cornell University Press.
- Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of the partial. *Feminist Studies* 14(3), 575–599. <https://doi.org/10.2307/3178066>

- Hoffmann, A. L. (2020). Terms of inclusion: Data, discourse, violence. *New Media & Society* 23(12), 3539–3556. <https://doi.org/10.1177/1461444820958725>
- Hong, R. (2016). Soft skills and hard numbers: Gender discourse in human resources. *Big Data & Society* 3(2), 1–13. <https://doi.org/10.1177/2053951716674237>
- hooks, b. (2000). *Feminism is for Everybody: Passionate Politics*. London: Pluto Press.
- Huffer, L. (2013). *Are the Lips a Grave? A Queer Feminist on the Ethics of Sex*. New York: Columbia University Press.
- Intel. (2021, February 2). Intel and Google Cloud Team with All of Us Research Program. [Press Release] Retrieved on 19 February 2023 from <https://www.intel.com/content/www/us/en/newsroom/news/all-us-research-program.html#gs.ypgaqv>
- Jasanoff, S. (2015). Future imperfect: Science, technology, and the imaginations of modernity. In Jasanoff, S., & Kim, S.-H. (eds). *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. Chicago: The University of Chicago Press.
- Jasanoff, S., & Kim, S.-H. (2015). *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. Chicago: The University of Chicago Press.
- S., Kim, S. -H., & Sperling, S. (2007). Sociotechnical imaginaries and science and technology policy: a cross-national comparison. [National Science Foundation Research Project Proposal]. Cambridge, MA: Harvard University. Retrieved on 19 February 2023 from <https://stsprogram.org/admin/files/imaginaries/NSF-imaginaries-proposal.pdf>
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1). <https://doi.org/10.1177/2053951714528481>
- Knight Steele, C. (2021). *Digital Black Feminism*. New York: NYU Press.
- Larrondo, A., Morales-i-Gras, J., & Orbegozo-Terradillos, J. (2019). Feminist hashtag activism in Spain: measuring the degree of politicisation of online discourse on #YoSíTeCreo, #HermanaYoSíTeCreo, #Cuéntalo y #NoEstásSola. *Communication and Society* 32(4), 207–221. <https://doi.org/10.1558/003.32.4.207-221>
- Leurs, K. (2017). Feminist Data Studies: Using Digital Methods for Ethical, Reflexive and Situated Socio-Cultural Research. *Feminist Review* 115(1), 130–154. <https://doi.org/10.1057/s41305-017-0043-1>
- Levenda, A. M., Ritcher, J., Miller, T., & Fisher, E. (2019). Regional sociotechnical imaginaries and the governance of energy innovations. *Futures* 109, 181–191. <https://doi.org/10.1016/j.futures.2018.03.001>
- Levina, M., & Hasinoff, A. A. (2016). The Silicon Valley Ethos: Tech Industry Products, Discourses, and Practices. *Television and New Media* 18(6), 489–495. <https://doi.org/10.1177/1527476416680454>
- Lindh, M., & Nolin, J. (2017). GAFAspeak: Metaphors in the promotion of cloud technology. *Journal of Documentation* 73(1), 160–180. <https://doi.org/10.1108/JD-03-2016-0039>

- Love, A. (2018). Digital black feminism. In Parry, D.C., Johnson, C.W., & Fullagar, S. (eds) *Digital Dilemmas: Transforming Gender Identities and Power Relations in Everyday Life*. London: Palgrave MacMillan. 53–71.
- Luka, M. E. and Millette, M. (2018). (Re)framing Big Data: Activating Situated Knowledges and a Feminist Ethics of Care in Social Media Research. *Social Media + Society* 4(2), 1–10. <https://doi.org/10.1177/2056305118768297>
- Lupton, D. (2020). ‘Not the real me’: Social imaginaries of personal data profiling. *Cultural Sociology* 15(1), 3–21. <https://doi.org/10.1177/1749975520939779>
- Lykke, N. (2020). Transversal Dialogues on Intersectionality, Social Feminism and Epistemologies of Ignorance. *NORA: Nordic Journal of Feminist and Gender Research* 28(3), 197–210. <https://doi.org/10.1080/08038740.2019.1708786>
- Mager, A. (2016). Search engine imaginary: Visions and values in the co-production of search technology and Europe. *Social Studies of Science* 47(2), 240–262. <https://doi.org/10.1177/0306312716671433>
- Mager, A., & Katzenbach, C. (2020). Future Imaginaries in the Making and Governing of Digital Technology: Multiple, Contested, Commodified. *New Media & Society* 23(2), 223–236. <https://doi.org/10.1177/1461444820929321>
- Markham, A. (2020). The limits of the imaginary: Challenges to intervening in future speculations of memory, data, and algorithms. *New Media & Society* 23(2), 382–405. <https://doi.org/10.1177/1461444820929322>
- Mayer-Schoenberger, V., & Cukier, K. (2013). *Big Data. A Revolution That Will Transform How We Live, Work, and Think*. London, UK: John Murray Press.
- McNamee, R. (2019). *Zucked: Waking Up to the Facebook Catastrophe*. New York: Penguin.
- McQuillan, D. (2016). Algorithmic paranoia and the convivial alternative. *Big Data & Society* 3(2), 1–12. <https://doi.org/10.1177/2053951716671340>
- Michailidou, M. (2018). Feminist methodologies for the study of digital worlds. *International Journal of Media & Cultural Politics* 14(1), 19–33. https://doi.org/10.1386/macp.14.1.19_1
- Milan, S., & Trere, E. (2019). Big Data from the South(s): Beyond Data Universalism. *Television & New Media* 20(4), 319–335. <https://doi.org/10.1177/152747641983773>
- Nakamura, L. (2013). *Cybertypes: Race, ethnicity, and identity on the Internet*. Routledge.
- Nelhans, G., & Nolin, J. (2022). Anti-transparency within the EU shift towards Open Science. *Science and Public Policy* 49(6), 928–941. <https://doi.org/10.1093/scipol/scac039>
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Odumosu, T. (2020). The Crying Child: On Colonial Archives, Digitization, and Ethics of Care in the Cultural Commons. *Current Anthropology* 61(22), 289–302. <https://doi.org/10.1086/710062>

- Onouha, M. (2016) The library of missing datasets. [Mixed media installation]. Retrieved 19 February 2023 from: <https://mimionuoha.com/the-library-of-missing-datasets>
- O’Neil, C. (2017). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin.
- Perez, C. C. (2019). *Invisible Women: Data Bias in a World Designed for Men*. New York: Abrams Press.
- Rieder, G. (2018). Tracing Big Data imaginaries through public policy: The case of the European Commission. In Rudinow Sætnan, A., Schneider, I., & Green, N. (eds) *The Politics and Policies of Big Data*. London, UK: Routledge. 89–109. <https://doi.org/10.4324/9781315231938>
- Ruberg, B., & Reilos, S. (2020). Data for queer lives: How LGBTQ gender and sexuality identities challenge norms of demographics. *Big Data & Society* 7(1). <https://doi.org/10.1177/2053951720933286>
- Saariketo, M. (2014). Imagining Alternative Agency in Techno-Society: Outlining the Basis of Critical Technology Education. In Kramp, L., Carpentier, N., Hepp, A., Tomanić Trivundža, I., Niemenin, H., Kunelius, R., Olsson, T., Sundin, E., & Kilborn, R. (eds) *Media Practice and Everyday Agency in Europe*. Bremen, DE: edition lumière. 129–137. http://www.researchingcommunication.eu/book9chapters/C06_Krotz1314.pdf
- Saariketo, M. (2020). Subversive imagination and possibilities of critical agency in the landscapes of code? Selected Papers of #AoIR2020: The 21st Annual Conference of the Association of Internet Researchers. Dublin, Ireland, 28–31 October 2020. <https://doi.org/10.5210/spir.v2020i0.11323>
- Sadowski, J., & Bendor, R. (2018). Selling Smartness: Corporate Narratives and the Smart City as a Sociotechnical Imaginary. *Science, Technology, & Human Values* 44(3), 540–563. <https://doi.org/10.1177/0162243918806061>
- Sandeep, M. (2017). Socio-technical imaginaries of a data-driven city: Ethnographic vignettes from Delhi. *The Fiberculture Journal* 29. <https://doi.org/10.15307/fcj.29.217.2017>
- Secundo, G., Del Vecchio, P., Dumay, J., & Passiante, G. (2017). Intellectual capital in the age of Big Data: Establishing a research agenda. *Journal of Intellectual Capital*, 18(2), 242–261. <https://doi.org/10.1108/JIC-10-2016-0097>
- Snapp, S. D., Russel, S. T., Arredondo, M., & Skiba, R. (2016). A Right to Disclose: LGBTQ Youth Representation in Data, Science, and Policy. In Horn, S.S., Ruck, M.D., & L.S.Liben (eds) *Equity and Justice in Developmental Science: Theoretical and Methodological Issues* 50, 135–159. <https://doi.org/10.1016/bs.acdb.2015.11.005>
- Suárez-Gonzalo, S. (2019). Personal data are political: A feminist view on privacy and big data. *Recerca: Revista de Pensament I Anàlisi* 24(2), 173–192. <https://doi.org/10.6035/Recerca.2019.24.2.9>

- Sörum, N., & Fuentes, C. (2022). How sociotechnical imaginaries shape consumers' experiences of and responses to commercial data collection practices. *Consumption Markets & Culture* 26(1), 24-46. <https://doi.org/10.1080/10253866.2022.2124977>
- Van Oost, E. (2000). Making the computer masculine: The historical roots of gendered representations. In Balka, E. & R. Smith (eds) *Women, Work and Computerization: Charting a Course for the Future*. The International Federation for Information Processing 44. Boston, MA: Springer. https://doi.org/10.1007/978-0-387-35509-2_2
- Taylor, C. (2004). *Modern Social Imaginaries*. Durham, NC: Duke University Press.
- Terras, M. & Nyhan, J. (2016). Father Busa's Female Punch Card Operatives. In *Debates in the Digital Humanities 2016*. Matthew K. Gold and Lauren F. Klein, Editors. Berkeley: University of California Press, Journals & Digital Publishing Division.
- Thompson, T. L. (2020). Data-bodies and data activism: Presencing women in digital heritage research. *Big Data & Society* 7(2). <https://doi.org/10.1177/2053951720965613>
- Vaitla, B., Verhulst, S., Bengtsson, L., Gonzalez, M. C., Furst-Nichols, R., & Pryor, E. C. (2020). The promise and perils of big gender data. *Nature Medicine* 26(19), 17-18. <https://doi.org/10.1038/s41591-019-0712-z>
- 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>
- Van Eck, N., & Waltman, J. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter* 7(3), 50-54. <https://www.issi-society.org/media/1131/newsletter27.pdf>
- Wajcman, J. (2004). *TechnoFeminism*. Cambridge, UK: Polity Press.
- Ward, J. S., & Baker, A. (2019). Undefined by data: A survey of Big Data definitions. *ArXiv*, abs/1309.5821. <https://arxiv.org/pdf/1309.5821.pdf>
- Weinberg, L. (2017). Rethinking Privacy: A Feminist Approach to Privacy Rights after Snowden. *Westminster Papers in Communication and Culture* 12(3), 5-20. <https://doi.org/10.16997/wpcc.258>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Profile Books.