The social imaginaries on governance through data: Q-methodological analysis of data professionals' views in the field of mobility

Masso, Anu; Männiste, Maris; Calzati, Stefano

DOI
10.1016/j.futures.2023.103199

Publication date
2023

Document Version
Final published version

Published in
Futures

Futures 152 (2023) 103199

Contents lists available at ScienceDirect

Futures

journal homepage: www.elsevier.com/locate/futures

The social imaginaries on governance through data: Q-methodological analysis of data professionals' views in the field of mobility

Anu Masso [a,*], Maris Männiste [b], Stefano Calzati [c]

[a] Ragnar Nurkse Department of Innovation and Governance, Tallinn University of Technology, Estonia
[b] Institute of Social Studies, University of Tartu, Estonia
[c] Department of Urbanism, Delft University of Technology, the Netherlands

ARTICLE INFO

Keywords:
Big data
Social datafication
Spatial mobility
Social imaginaries
Q-methodology
Discourse

ABSTRACT

This study strives to contribute to the discussions on the increasing use of data in organisational settings within specific contexts like governance and control of spatial mobilities. Whereas legal, regulatory, and technical aspects are widely studied, there is less knowledge about public discourses on data, especially the data professionals' interpretations (re)constructing these understandings. This study conducted Q-methodological interviews among data professionals employed in the mobility domain in Estonia—a leading digital society—who work in public, private, and third-sector organisations. The findings show the emergence of four discourses on data: (I) data-based liberal internationalism and equal access; (II) integrative data activities through cooperation and disciplinary continuity; (III) (big)data self-reflections and organisational data culture; and (IV) data discrimination aware data practices and privacy protection. Thus, governance through data does not constitute a single imaginary future but creates various scenarios across the types of both data and mobility and the sector. We propose that governance through data emerges in these discourses not only as 'enculturation', where data are seen as an ultimate medium of knowledge, but as 'acculturation'—critically self-reflexive learning through data as a form of change in organisations and among data professionals.

1. Introduction

Restricted mobility during a pandemic, forced migration, or increasing digital migration requires rapid decision-making and governance solutions. These processes have accelerated the need to use and adapt datafied solutions to govern and control mobilities. However, there is still little evidence about data professionals as agents initiating, developing, using, and implementing datafied solutions nor how they contribute to constructing new imaginary futures.

In this article, we examine social imaginaries based on the use of data in decision-making, which has become the subject of public attention, starting with using big data for developing services for public administration (Ahas & Mark, 2005), accelerated with the launch of artificial intelligence strategies (e.g., Männiste & Masso, 2020) and intensified with the COVID-19 outbreak (Kummitha,

* Correspondence to: Ragnar Nurkse Department of Innovation and Governance, Tallinn University of Technology, Akadeemia tee 3, 12616 Tallinn, Estonia.
E-mail address: anu.masso@taltech.ee (A. Masso).

https://doi.org/10.1016/j.futures.2023.103199
Received 3 August 2022; Received in revised form 2 June 2023; Accepted 12 June 2023
Available online 15 June 2023

2020). Previous research has thoroughly investigated a broad variation of principles in how datafied solutions are implemented (Janssen et al., 2017; Latzer & Just, 2020; Taylor & Purtova, 2019), including discussions on not only the technical processes of data mining but also related ethical implications how the data are collected, used, processed and shared (Berendt et al., 2022). Research suggests that the understanding of data is framed and grounded institutionally, distributed, and (re)constructed publicly (Guenduez et al., 2020; Grosman & Reigeluth, 2019). However, there is less knowledge about how these publicly constructed social imaginaries or collectively shared understandings of sociotechnical progress (see, e.g., Jasanoff & Kim, 2015) are realised by communities, initiating, designing, and implementing data technologies, but also (re)constructing these public imaginaries. Although prior research emphasises the central role of data professionals in constructing the future through social imaginaries of data (Hepp, 2016; Meng & Disalvo, 2018; O'Neil, 2016; Schrock & Shaffer, 2017; Männiste & Masso, 2020), there is still a lack of knowledge about the comparison of the alternative futures as reflected in the social imaginaries. Besides, prior research has mainly used a qualitative approach that still has limitations in systematically comparing and testing the alternative imaginary futures that a combination of quantitative and qualitative analysis would enable.

Although there is much research on datafied governance in general (Amoore, 2019; Just & Latzer, 2017), there are fewer empirical studies on the sociotechnical imaginaries of datafication in organisations within a specific context like mobility data (Aradau & Blanke, 2022; Molnar, 2021; Männiste & Masso, 2020). Data on mobility is one of the fastest-growing types of data due to the increasing number of mobile devices used among the population globally (Berendt et al., 2022). Therefore, we still do not know how the data technologies used in governing a broad variety of mobile groups are understood and what the potential futures are embedded in these social imaginaries. In this article, we take mobility as a case study and a lens to explore the diverse discursive articulations by data professionals on the kind of imaginaries triggered by, and through, data. On a broader level, as van Lente (2000, 43) remarks, 'technological futures are forceful', meaning that discourses and expectations on technology's design, use, and implementation contribute to realising those same (imagined) scenarios they envision. Besides, as sociotechnical imaginaries are discourses (Foucault, 1980) producing power relations and constructing the social realm, exploring the discourses surrounding data-driven solutions for policing various forms of mobilities and subjects acquires a particularly performative relevance. Performativity must be intended here in a double-sided meaning as, on the one hand, the force of language in reshaping reality by intervening in it (Austin, 1962); on the other hand, as the technologisation of reality, that is, the ensemble of means and techniques for obtaining efficient and accountable outcomes (Lyotard, 1984). Therefore, based on this, we assume that the social imaginaries bear concrete (dis)enabling consequences for both those who utter them and also those towards whom they are directed.

We examine the meanings, practices, and implications surrounding data technologies as seen, understood and (re)constructed by data professionals in Estonia, working mainly in the mobility domain. This study focuses on Estonia, often characterised as a leading digital society (Männiste & Masso, 2020; Tammpuu & Masso, 2018), where the development of data technologies is often used as a tool for national branding. Therefore, data technologies' norms, values and understandings are publicly prominent. We analyse what certain data experts say about data governance rather than criticising directly how governance of mobility is specifically performed through data. To examine data professionals' understandings about using data in decision-making and formulate the challenges and futures of datafied governance, we apply a Q-methodological mixed-method approach to assess data professionals' subjective perspectives. For this purpose, we formulated the following research questions: (1) How do the data professionals who contribute to developing datafied governance solutions imagine datafied decision-making? (2) How are the emerging social imaginaries similar and different, considering the data professionals' various experiences with data, mobile groups and depending on the institution where they work? (3) What are the key characteristics of the shift towards datafied governance based on visions of the future expressed by data professionals?

2. Literature review

2.1. The shift towards governance through data

The research community has extensively studied the shifts towards implementing data technologies (e.g., Masso & Kasapoglu, 2020; Broeders & Dijstelbloem, 2016; Taylor, 2016). Prior research agrees (Jutel, 2021; Lupton, 2016; McQuillan, 2016; McStay and Rosner, 2021; Smith, 2018) that change towards using data technologies in governance entails significant shifts in data sharing and using habits. Some authors understand this shift as the 'enculturation' of individuals into data-sharing habits as a potential solution to societal problems, a medium of knowledge creation and an indication of progress (Smith, 2018). Other authors understand this change process as an 'acculturation', whether in the form of a shift in the ways how the ideas of transparency and surveillance are considered in society (McQuillan, 2016) or a change in the respective practices with data and relationships between individuals (McStay and Rosner, 2021; Lupton, et al. 2016). However, research does not agree on the processes, obstacles, and (un)expected outcomes of this change process of introducing big data in organisations.

Prior research reveals diverse forms of how the use of data in governance has been introduced and practised. This research includes both explorations in the use of data in government in general (Drechsler, 2019; Guenduez et al., 2020; Janssen et al., 2017; Molnar & Fletcher, 2021) or in the domain of mobility in a particular (Broeders & Dijstelbloem, 2016; Scheel, Ruppert, & Ustek-Spilda, 2019; Masso & Kasapoglu, 2020) like indicators, statistics, and automated tools. Often, the issue is framed in adopting big data in decision-making as governance driven *by* data. Concepts like data-driven, data-informed, techno and human-driven governance have often been used in parallel to emphasise the power of data in directing the meaning-making of datafied social processes. Research suggests (see Ranerup & Henriksen, 2019) that new data technologies have increased accountability, decreased costs and enhanced efficiency in governance. Therefore, such 'data revolution' is often discursively framed as a form of social progress, innovation, or even

a branding instrument, neglecting the embeddedness of data in society at large or the role of individuals as active agents in this process (Guenduez et al., 2020; Morozov, 2017).

However, other research proposes alternatives to the data-driven forms of governance, framed by effectivity, efficiency and progress, such as governance *through* or *with* the data (Saurwein et al., 2015). These studies suggest moving towards decisions informed by the data (Haardörfer, 2019; Männiste & Masso, 2020) to emphasise meaning-making based on theory-driven questions, where human agents participating in this process have an active role in formulating and revising these questions (Masso, Maris Männiste, & Andra, 2020). Therefore, some studies are sceptical about decision-making, considering any indicators other than efficiency, suggesting even modelling supra data and indicators and governing without them (Drechsler, 2019). These studies emphasise that when neglecting the role of the human participants, the change in implementing big data technologies in governance may have several negative (un)intended consequences. For example, research highlights that focusing on speed and efficiency but also lacking transparency may lead to legitimising social inequalities (Eubanks, 2018; Saurwein, Just, & Latzer, 2015; Masso & Kasapoglu, 2020), and neglecting the data sovereignty—fundamental rights of data subjects to collect and manage their data (Hummel et al., 2021)—as several studies warn.

Therefore, the literature increasingly emphasises the essential role of not only data agents (Milan & Treré, 2019) but also data professionals (Männiste & Masso, 2020) in initiating novel datafied initiatives). Scholars propose concepts like the 'pioneer communities' (Hepp, 2016) to signify that data professionals are leading the way in introducing the data-based cultural shift in the public and organisations they work for. These actors are presumed to have specific reflexivity skills (O'Neil, 2016), to respond to structural disruptions and possible dangers related to data, besides their favourable position in accessing, using and translating the data for the public (Schrock & Shaffer, 2017). Research claims that data pioneers' understandings and (counter)activities as grassroots initiatives (Meng & Disalvo, 2018) generate critical collective consciousness besides their effort to mobilise the data resources. Those active groups also express their faith or resistance in these collective understandings, shaping their use and design of the data technologies (Thomas et al., 2018).

In summary, the research agrees that there is a significant shift in organisations towards using datafied solutions. However, in which form this shift happens, whether *by*, *through,* or *with* data, is still unknown. Besides, we do not know how these alternative futures are imagined by the data experts as 'insiders', supposedly knowing the workings of data, and what the potential social dilemmas on data are that contribute to or work against developing the unified expert community. From here, we aim to examine how data professionals, as part of active communities, express their understanding of governance through data.

2.2. The social imaginaries on data

The increasing use of data by institutions produces a variety of discursive imaginaries based on hopes and fears, utopias and dystopias related to technological innovations. Therefore, the increasing use of data technologies has led scholars to analyse social imaginaries about different sets of data or data-based futures concerning approaches incorporating data in the decision-making process in the organisations (Bucher, 2016; Hummel et al., 2021).

Jasanoff and Kim (2015) propose the concept of 'sociotechnical imaginaries' to analyse technology-related imaginaries, inspiring a collective understanding of what is suitable or desirable in a particular society, i.e., how technologies can fulfil public desires. In this respect, sociotechnical imaginaries are discourses in the Foucault (1980) sense of the term: a *dispositif* of power. The idea that discourses are technologies producing power relations and contributing to shaping the social realm is well-rooted. Austin (1962) was the first to speak of the performativity of language— the pragmatic force of words in actualising certain states of realities and enabling certain conditions of existence. On Austin's path, Lyotard (1984) denounces the instrumentalisation of discourse to enact that same efficiency that machines were already putting forth in labour in the societal realm at large. Therefore, discourses and technologies are two faces of the same coin. Technology exists as much as it can be imagined, i.e., actualised first and foremost through language as an apparatus of power: 'technological futures are forceful', van Lente (2000) notes in this regard. Such technology-related futures often mean reconfiguring physical infrastructures, transforming social infrastructures, establishing new patterns of life and work, and reallocating societal benefits and burdens.

Scholars have thoroughly explored and agree that such social imaginaries on governance through data are rooted in public discourses (Guenduez et al., 2020; Pink et al., 2018; Rieder, 2016) in particular social contexts and historical narratives. For example, recent research substantiates the belief that human interaction with technology is mediated by the political and institutional contexts (Kummitha, 2020), where data, for example, on mobility, are created and implemented. Some studies estimate public understandings about data as significant drivers of accountability for governmental institutions (Ojala et al., 2019), while Carlson (2018) proves the reverse process where practised datafied judgments impact and legitimise data imaginaries. Consequently, we do not oppose discourses and imaginaries in this article but consider the two are entangled—as the existence of imaginaries depends on discourses and vice versa. How imaginaries and discourses are shaped through language is a double-sided process which sees language implicated as both a performative force able to define reality (imaginaries) and as a technologisation of reality, meaning that language *is* technology, responsible for producing a certain accountable reality (discourses).

However, these data imaginaries are not always seen as unified ideas and understandings but in the form of social dilemmas that contribute to developing data commons (Taylor & Purtova, 2019) as a social change towards sustainable and socially acceptable use of data technologies. Therefore socially constructed and culturally situated understandings about data form certain 'data ideologies' (Pink et al., 2018; Thornham & Gómez Cruz, 2016), which experts working with data and institutions often use to justify and legitimise their data practices. Research agrees that the discursively constructed understandings of data (Drechsler, 2019; Latour & Woolgar, 1979) are materialising in the everyday work of data experts when they manage, process and analyse data. Besides, individuals also

construct counter-imaginaries (Hummel et al., 2021; Pink et al., 2018) that highlight the social responses to the complex changes related to data technologies and the hidden potential risks.

In this article, we mean data futures in the sense that any technological imaginary is forceful, as van Lente emphasises (2000), that is, it creates and establishes its condition of existence (and, therefore, further development of governance through data). This article contributes to unpacking such imaginaries to delineate trajectories of the possible evolution of data governance. Although numerous studies on social imaginaries of data have been conducted, there is a lack of empirical research about social imaginaries on data as seen by data professionals working in specific domains, like human mobility. Therefore, the discursive strategies on how datafication is understood in organisations and by experts within a specific context, like spatial mobilities, are still unknown. Besides, we do not know the social imaginaries of these diverse forms of datafied governance, which the data experts see as 'insiders', supposedly knowing the workings of data.

3. Q-methodology

We operationalise the research on social imaginaries by applying a Q-methodological discourse analysis (Ramlo & Newman, 2011). This method enables us to systematically assess data experts' positions on public discourses related to datafication in organisations, using the field of mobility as a specific example. According to the principles of the Q-methodology (McKeown & Thomas, 2013; O'Neil, 2016; Watts & Stenner, 2012), we combine statistical analysis techniques to reveal factors and qualitative analysis to interpret patterns of subjective views on data. In Q-methodology, the discourses arise from correlations between statements and the groupings of individuals who agree with these statements, unlike the traditional quantitative factor (Finch, 2019) or qualitative discourse analysis (Fairclough, 2003; Wodak, 1996).

We rely on Stephenson's original approach to the Q-methodology (O'Neil, 2016; Stephenson, 1953), which suggests examining the interviewees' interpretations on a given topic rather than exclusively quantitatively striven initial approaches (Burt, 1937 as cited in Cattell, 1973). This article uses principles similar to qualitative discourse analysis, like the systematic approach used in the analysis process (Fairclough, 2003) and emphasises the context (Wodak, 1996) when interpreting the results of subjective positions.

3.1. Q and R samples

Inherent to Q-methodology, we designed the Q sample of the statements (Q = 36) and the R sample of the individual participants (R=24). The sample size chosen in this study follows the principles of Q-methodology so that the number of statements is greater than the size of the participants' sample (Watts & Stenner, 2012).

To form the Q sample, we first mapped the full range of public understandings and opinions on data, i.e., political and strategic documents, general discussions in the media (i.e., news articles about different examples of data-based governance), and research conducted on datafied governance to form a *concourse* in our study from where we extracted the final statements' Q-sample. Public discussions on data covered two main dimensions: 1) subjective views on data technologies, including individual experiences or activities perceived in their work or society in general; and (2) possible sociocultural consequences (i.e., data discrimination) of data technologies that participants consider to be negative, neutral, or positive. We introduced a conceptual matrix formed due to the intersections of these dimensions (Table 1). We extracted an equal number of statements in each matrix field (n = 4 from the total Q = 36) to balance the chosen statements across the complete discourse representing the public discourse on the data.

We followed the qualitative purposeful sampling principles (Suri, 2011) to design the sample of the individuals. Based on this principle, the sample includes both homogeneous and heterogeneous properties. The homogeneous characteristic of the sample is that all the participants use and design data technologies for decision-making purposes, including the specific field of human mobility in Estonia. The interviewed experts have been working with diverse data-based solutions in general, but also including work with data within a specific context of mobility, like analysing, managing or regulating everyday mobilities, forced migration or digital migration (e.g., Estonian e-residency). However, although the participants were chosen based on their prior experiences with data in the field of mobilities, the participants mostly had additional expertise beyond this particular domain. To ensure the heterogeneity of subjective views, the participants had distinct roles: analysts and developers of registry databases, analysis department managers, and software and algorithms designers (R = 24). These varied roles also ensured that the sample represents the experts acting as procurers, i.e.,

Table 1
The conceptual framework applied to extract statements from the *concourse*.

Focus of positions	The tone of the statements		
	Obstacles or negative implications to mobility data technologies	Advantages or positive implications to mobility data technologies	Neutral or normative implications to mobility data technologies
Mobility data technologies in the individual expert's work	*1, 2, 3, 4*	*13, 14, 15, 16*	*25, 26, 27, 28*
Mobility data technologies in the institutions' cooperation	*5, 6, 7, 8,*	*17, 18, 19, 20*	*29, 30, 31, 32*
Mobility data technologies in the context of society and state	*9, 10, 11, 12*	*21, 22, 23, 24*	*33, 34, 35, 36*

*Statement number (see statements in Discourses 1-4, Tables 3–7) – in the text, these statement numbers are prefixed with #.

ordering data-based analysis or governance solutions and executing procurements as third-sector analysts or developers of data technologies. Experts from the public, private, and third-sector institutions are proportionally represented in the sample. We engaged data experts dealing with various data, such as register, survey, or tracking data (i.e., social media, mobile phone, online transactions, and identity verification data). The sample is diverse regarding gender, educational discipline, and age.

To examine the subjective views of the experts, we (1) asked open-ended questions to encourage experts' spontaneous opinions on big data, 2) then asked them to sort the 36 Q-statements, and (3) finally asked them to justify their Q-sorting choices. We asked participants to sort the statements on an agreement scale, where -5 indicated complete disagreement, +5 showed complete agreement, and 0 showed neutrality (see Fig. 1). To encourage the expression of the subjective positions, the 0-point scale was conventional rather than 'forcing' participants to choose on a given symmetrical scale. Interviews were conducted in 2018 and lasted an average of ninety minutes. This study was conducted in the period when the public discussions at the start of the implementation of the data-based solutions were most vivid, and the imagined futures were constructed. Therefore, the Estonian case reveals the potential manifestations of the shifts towards datafied futures, focusing on the most important breaking moment in this long change process.

3.2. Centroid factor analysis

Following the main principle of Q-methodology, we combined statistical techniques with the qualitative interpretation of the patterns of the individual positions. First, we conducted a quantitative factor analysis with the centroid rotation method to extract the discourses on data technologies. Unlike the traditional varimax rotation in the factor analysis (Finch, 2019), we used the centroid rotation method (Ramlo, 2016) to extract non-orthogonal factors[1]. The manual, pair-wise centroid rotation method considers the participants' views and contextual information when choosing the final factor structure. We used Ken-Q software to analyse the discourses quantitatively. Second, we compared and interpreted the quantitatively found discourses with qualitative thematic analysis techniques. Similar to qualitative discourse analysis (Fairclough, 2003; Wodak, 1996), we systematically compared and interpreted differences within and between discourses, counting the background information of each interviewee and the topic context. Manual and computer-aided techniques were combined (Woolf & Silver, 2017) to organise the opinions on pre-defined statements and compare the participants' positions. We analysed open-ended interview questions with Maxqda software.

The analysis presents factor loadings and factor scores of distinguishing discourses[2]. We compare the factor loadings to define the distribution of the individual interviewees between the factors. We calculated the factor scores to assess the composition of each discourse based on the individuals' agreements. We then interpreted and illustrated the factor structure with short extracts from the qualitative material. In the quantitative and qualitative analysis, we compare the similarities and differences in the participants' views across the discourses. To assure the anonymity of the respondents, only the type of data and the field in which the interviewee works are presented at the end of each direct quote (i.e., tracking data, everyday mobility).

4. Results: Imaginaries of governance through data technologies

Four factors emerged in the analysis, highlighting how governance through data technologies is understood. The four-factor solution was chosen based on pair-wise comparisons of the factors and interpreting the qualitative contextual information. We compare factor scores of statistically significantly distinguishing statements[3] in each factor and provide an interpretation of discourses on governance through data technologies, considering both the distribution of the respondents across discourses (Table 2), the consensus statements (Table 3), the statements differing across discourses (Tables 4–7), and comparing participants' spontaneous clarifications within and between the discourses.

4.1. Discourse I: 'Data-based liberal internationalism and equal access'

We call Discourse I—'Data-based liberal internationalism and equal access'. Based on their agreement with two statements (#34, #36), three interviewees working in public, private or third-sector organisations represent this discourse (see Tables 2, 4). They are involved in analysing register, tracking or survey data and developing solutions for everyday mobilities and forced migration. Most of the interviewees in this discourse work in a managerial position, which is essential in analysing their experiences related to the practices with mobility data analyses in their organisations.

As the quantitative analysis reveals (see Table 4), the experts understand data as a 'human right' and expect equal access to any data that should be guaranteed to experts and all members of society (see statement #36). The global data transmission standard must be developed to use the full potential of the available digital data, as perceived by the interviewees in this discourse (#34). As such, Discourse I emphasises big data not just as a resource to grant fair solutions within national territories but as a global potentiality to

[1] Instead of maximising the sum of the squares of the factor loadings, as is usually the case when analysing the principal components, we maximised the sum of the factor loadings according to the principles of centroid factor analysis.

[2] In this study, we call patterns of subjective positions 'discourses' rather than 'factors', as previously suggested (Watts & Stenner, 2012) to emphasize the abductive nature in Q methodology.

[3] Interviewees expressed consensus regarding 6 statements. Of the 36 statements, 19 were distinguished by 4 factors and 9 did not distinguish the discourses either statistically or significantly. These 9 statements, which did not have a significant statistical effect on the composition and structure of the factors, have been excluded from the present analysis.

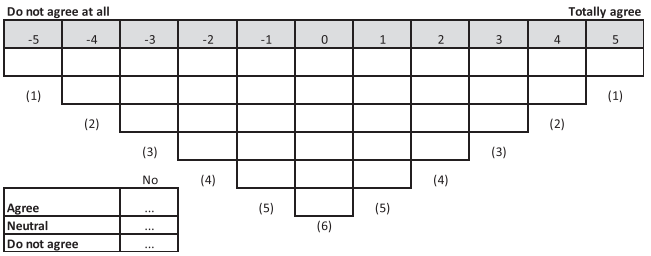

Fig. 1. Applied symmetrical distribution for the Q-sorting of 36 statements.

Table 2
The sample structure of the study across the four discourses (factor loadings).

No.	D I	D II	D III	D IV	Field of expertise	Sector	Data type	Data field
1	0,635*	0,439			Data analyst	Private sector	Tracing data	Everyday mobility
2	0,689*				Manager	Third sector	Register data	Forced migration
3	0,628*				Manager, adviser	Public sector	Survey data	Everyday mobility
4	0,553				Data analyst	Third sector	Tracing data	Forced migration
5	0,439				Data analyst	Private sector	Register data	Digital migration
6		0,683*			Adviser	Public sector	Register data	Forced migration
7		0,618*			Data analyst	Public sector	Tracing data	Forced migration
8		0,593*			Data analyst	Public sector	Register data	Forced migration
9		0,546*			Data analyst	Third sector	Survey data	Forced migration
10		0,398			Data analyst	Third sector	Tracing data	Forced migration
11		0,454	0,660*		Manager	Public sector	Tracing data	Everyday mobility
12			0,542*		Manager	Public sector	Tracing data	Everyday mobility
13			0,536*		Manager	Private sector	Survey data	Forced migration
14			0,522*		Manager	Third sector	Survey data	Forced migration
15		0,578	0,612		Manager	Public sector	Register data	Forced migration
16			0,577		Data analyst	Third sector	Tracing data	Digital migration
17			0,367		Manager	Private sector	Tracing data	Everyday mobility
18				0,599*	Manager, analyst	Private sector	Tracing data	Everyday mobility
19				0,598*	Manager, analyst	Private sector	Tracing data	Forced migration
20				0,568*	Data analyst	Private sector	Tracing data	Everyday mobility
21				0,479*	Data analyst	Private sector	Register data	Digital migration
22		0,482		0,555	Data analyst	Third sector	Survey data	Forced migration
23	0,441			0,53	Data analyst	Private sector	Tracing data	Forced migration
24				0,511	Manager, analyst	Private sector	Survey data	Forced migration

* Statistically significant factor loadings, distinguishing the factors ($p > 0.01$).

Table 3
Consensus statements that do not distinguish the factors (factor scores).

No	Statement	D I	D II	D III	D IV
2	*In today's societies, 'data-rich' and 'data-poor' are emerging; I often do not have access to big data as an 'elitist resource'.*	-3	-1	-1	-1
15	Thanks to the daily work with the data, I can influence processes, find solutions to problems (social cohesion).	1	2	2	-1
21	*Today, data can be considered the most critical resource, the new 'mineral', like physical assets and financial capital.*	1	0	1	1
22	Open big data gives new opportunities for 'active citizenship', people have better chances to participate (repeat analyses).	1	1	0	0
28	*Institutions do not violate privacy; people must be responsible for their own 'digital footprints,' develop the 'digital literacy.'*	0	-1	-1	-2
31	Institutions that hold data about people (their spatial mobility) have a responsibility to provide solutions to societal issues.	0	2	1	0

Final factor scores (integer values of z-scores) instead of initial z-scores are used for interpreting the results in Table 2–6. D is the abbreviation of Discourse.

Table 4
Data-based liberal internationalism and orientation to equal access (Discourse I).

No	Statement	D I	D II	D III	D IV
36	*In today's society, data are like a 'human right.' All members of society should have equal access to data and information (e.g., the right to use cross-border digital services).*	3*	0	-1	-1
34	The development of a global data transmission standard is vital so that open, liberal internationalism can work in the 21st century.	1*	-3	-2	-2

* Statistically significantly distinguishing statements, p<=.01. The discourse is explained by three core and two additional interviewees.

strive towards universal obligations and rights.

Therefore, the participants agree that it is necessary to introduce regional or global governance principles to assure equal access to data, particularly in the novel situation where data instead of humans are moving across national territories. However, although interviewees see data as a form of 'human right', their understanding of human rights is somewhat varied from the traditional views – data as a form of human rights means equal and open access to the data, besides avoiding discrimination through data.

The additional qualitative data enabled us to examine how this liberal data internationalism intertwined with orientation to equal access is practised and understood in detail. As the additional in-depth interviews revealed, the experts in this discourse emphasise the data as a form of 'human rights' since they have often faced difficulties accessing the data, even when it is essential for conducting contractual tasks in their everyday work. The interviewees involved in this discourse noted they depend on the expectations formulated in the procurements' or projects' calls. As they feel as being under pressure from the tight competition, they also rarely offer alternative methods or approaches to the terms of reference provided by the customer. Their work is mainly guided by customers' needs and expected results and formulated standards for using data and making decisions. However, one of the experts working with mobility tracing data emphasises their role as a 'translator' – through processing and analysing data, the expert knows the possibilities and limits of data and offer the solutions to customers:

> *"We deal with translation [of data] (…) In this sense, we are seeking value for the client (…) We must understand what the client needs (…)" (tracing data, everyday mobility).*

While the aim of providing data is to justify decisions, the objective of collecting various data is to create simple systems and to offer data to other interested parties. Thus, experts interpret data as goods that can be sold to others under suitable conditions for a reasonable price and based on an agreement. That was also visible in the case of a respondent from a third-sector organisation, who, similarly to others, faces the need to find the monetary value of data:

> *"We want to create a kind of query environment so that if someone is interested in a section of these data, they can buy these directly via an e-interface* (…) *It is priced so that it would be cheaper if the entire database were bought"* (register data, forced migration).

If the sale of data would include a contract, then sharing the data with interested parties should be based on previous good experiences and mutual trust that the person who requests the data is sufficiently competent to use it, as seen in the following extract:

> *"We will not publish data files* (…) *There have been opposition politicians who have demanded files. We always ask people to introduce themselves"* (survey data, everyday mobility)

Besides particular technical solutions, the interviewees also emphasised the need for conceptual innovations regarding open government data. Potential issues like the lack of uniform data standards were mentioned due to varying understandings:

> *"I have understood that everyone defines this* (open data; authors) *differently. Open data here in our city, we must have requirements that there must be open data"* (survey data, everyday mobility).

Although experts see themselves as data providers, they emphasise the content of data rather than its openness, and constant access to data is a form of a right that is crucial to developing data technologies in government agencies, as seen here:

> *"Data are like human rights in today's society, they are not* (just; authors) *data, they are information"* (tracing data, everyday mobility).

This discourse highlighted the need to develop common standards for open data, data exchange, and data subjects' rights, applicable nationally and globally. Emerging global 'data internationalism' potentially fosters traded data relations regulated through contract or mutual trust and strong social ties. The emergence requires data experts to have the 'translation' skills to turn technical data analysis into social value and to intermediate the rights and needs of data subjects with decision-makers. Global data internationalisation development requires standards that govern profit-oriented, traded data relationships and guarantee the data subjects' rights. The idea of open government data presumes combining these standards without conflict, emphasised by the consensus of third and public sector interviewees dealing with broad data types.

Table 5
Data activities through cooperation and disciplinary continuity (Discourse II).

No	Statement	D II	D I	D III	D IV
30	*Data should be combined, and cooperation between authorities strengthened, to streamline the organization of public authorities.*	5	4	2	2
27	The big data era is nothing new; it is the acceleration of the same processes, characteristic of 100 years ago.	1	3	1	1
1	*Big data is primarily a massive technical 'craft.' It's mainly the domain of computer scientists.*	0*	-4	-5	-3
25	(Big) data is something mystical to me; I have not fully understood what it is.	-1	4	3	4
13	*The institution where I work has opportunities for further development with analytical methods; I feel confident in the context of the 'data revolution.'*	-2*	3	3	1

* Statistically significantly distinguishing statements, p<=.01. The discourse is explained by four core and two additional interviewees.

4.2. Discourse II: 'Integrative data activities through cooperation and disciplinary continuity'

We call Discourse II—'Integrative data activities through cooperation and disciplinary continuity'. The title was based on four interviewees from public and third sector institutions and their agreement with two statements (see #1, #13 in Table 5). All the interviewees in Discourse II are involved in the analysing register or tracking data related to forced mobilities. Most of the interviewees in this group work as analysts. Therefore, unlike other discourses, they reflect relatively high data literacy skills when describing their experiences with data analysis and management processes in their organisations.

Experts in this discourse emphasise an organisational perspective, unlike the international focus in Discourse I, as the quantitative analysis of the agreements with the presented statements reveals. For example, the experts in this discourse emphasise and expect clear work division between disciplines (as the agreements with statement #1 indicate) when working with data in their organisations. Besides, the interviewees expressed high expectations regarding training opportunities in their organisations (#13), which may have resulted from their high analytical skills, in which they saw a significant competitive advantage.

Additional qualitative analysis revealed that interviewees in this group valued high disciplinary expertise but expected collaboration between different parties to enable their work and datafied decision-making. Therefore, unlike other discourses, emphasis on certain conservative approaches with disciplinary continuity was characteristic of this discourse. For example, the interviewees highlighted that the key to turning data into value is the clear division of the work and cooperation between experts in different fields. In the case of some experts, the division of labour means they did not feel forced or obliged to improve their knowledge in working with different types of data. However, in the case of vacant analyst positions, decision-makers and advisers may be under pressure to develop skills to work with data. In those cases, there is no time for developing new data solutions, like merging existing data, enabling the use of the potential of the available data:

> *"The everyday work of advisers is very intense, and they just do not have time, and we also, unfortunately, do not have an analyst's position"* (register data, forced migration).

> *"We do not use [algorithms]* (…) *We have separate people who are directly involved with data"* (tracing data, forced migration).

However, the emphasis on disciplinary continuity also meant the experts were active in proposing novel solutions when doing their everyday work with data and negotiating the chosen approach within their organisation and with the institutions ordering analysis projects from them. Indeed, the interviewees in Discourse II introduced original solutions, as alternatives to restricted access to the data, unlike in Discourse I, where experts relied on prescribed norms and standards when working with data. For example, when preparing their analyses and tenders, the experts were aware of the register data they could obtain for the necessary analyses. They emphasise the problems of access to data considerably more often, compared to Discourse I, due to data protection rules. Interviewees invited various parties to cooperate to find joint solutions to these restrictions and to turn the data into value for the parties, for example, in the form of the merged database:

> *"We were trying to create a research database between various ministries to concentrate them all* (…) *We discovered that even finding an environment was difficult* (…) *One that would be sufficiently secure and universal"* (register data, forced migration).

The motivation to develop original data solutions was reasoned in Discourse II with the inappropriateness of the universal solutions, which, taken from another social context, would not work in specific contexts like mobility. In the case of governance through data, the experts also indicated that a severe issue is the lack of necessary data and access to indicators that correspond to actual needs. Besides the technical obstacles, the interviewees emphasised potential issues of premature generalisations and interpretations people tend to make when using mobility data. Consequently, the datafied services developed and offered by public institutions may not meet the needs of any particular group:

> "*How this topic of refugees is dealt with in the registers proceeds from a completely different logic than when looking at things from their viewpoint* (…) *The way they are coded in some information system* (…) *it may become a sort of violence"* (survey data, forced migration).

A characteristic feature of Discourse II is the expression of the experts' agency. The agency of the experts consists of both expressing criticism towards limited access to necessary data or indicators and implementing universal data solutions in the mobility field. Although expressing active support in the shift in the experts' organisations towards governance through data, the lack of those organisations' support may become the main obstacle. Abiding by traditional disciplinary boundaries, the experts of Discourse II are clearly distinguished from Discourse III by not introducing a novel open data culture. This discourse indicates that governing forced migration as a complex and sensitive topic assumes cooperation from the public and private sector organisations and crossing disciplinary boundaries, regardless of the form of the data used.

4.3. Discourse III: '(Big) data self-reflections and organisational data culture'

We call Discourse III—'(Big) data self-reflections and organisational data culture'. This title was based on the (dis)agreement of four interviewees working in public, private and third sector organisations regarding five statements (see Table 2, 6). In their work, primarily as managers or division leaders, the interviewees in this group mainly analysed and worked with data related to everyday mobilities and forced migration. This discourse comprised interviewees who had introduced a distinctive organisational decision-making culture based on extensive analytical skills and a self-critical, reflexive approach to governance through data. They, as

managers, developed a unique data culture in their organisation based on their passion for working with various data that the quantitative analysis of the agreements revealed a significant statement (#14). Interviewees also believe the use of data gives an organisation a significant competitive advantage (see statement #18). Therefore, the experts in this group are impatient regarding the slow development in the use of digital technologies and big data (#24). Participants in Discourse III also criticised the focus on speed in the development of data technologies without assessing the social consequences (#4, see Table 6).

In addition, the quantitative data analysis results are supported by the qualitative analysis of freely expressed explanations during the interviews, which show that the experts in this group expressed enthusiasm for developing a data culture in their organisation. However, they were also aware of the potential risks, so especially in the context of minority groups any possible biased decisions based on datafied solutions developed in their organisation get full public attention. Therefore, as the interviewees indicated, being 'data philanthropists' who actively develop data culture in their organisation comes with a responsibility to the public:

> *"I would call myself an evangelist; I just went and told everyone what data analysis is, why it's used, where it's useful"* (tracing data, everyday mobility).

Interviewees expressed a need to develop a data culture in their organisation due to the competitive advantage and new opportunities that big data opens up for both their organisations and the state. They also emphasise the speed necessary to introduce the data culture for future success in their field. Thus, the use of data technologies in governance offers possibilities to communicate the progress and the reputation of the state, as an interviewee emphasised:

> *"Using big data gives an essential advantage* (…) *There are so many examples from history that whoever leads the way is later copied by everyone"* (tracing data, everyday mobility).

Introducing novel data analysis tools entails a shift toward a data-oriented culture, as the interviewees emphasised. The data culture experts' proposal included decision-making using real-time analysis principles, interactive analysis tools, and the ability to estimate analysis results and make management choices accordingly and critically. Even when experts have felt some difficulties like resistance, they still actively strove towards the ideal of the data culture they proposed. For example, as an expert emphasised:

> *"I used data for reorganising management. I gave up work plans and traditional plans.* (…) *Managers must immediately get an online overview of how things are* (…) *I taught managers to ask which of the trends is better?"* (tracing data, everyday mobility).

Besides introducing data analysis tools in their organisations, the experts in this group are passionate about explaining the benefits of a data-based approach to their clients. These experts not only strove to use and implement novel data-based approaches but also spread their 'faith in data' within and outside their institutions through organising joint seminars with other interested parties, as one of the group pointed out:

> *"Spreading this culture, so that there would be this data-based organisation. At first, we had to look for work for ourselves, but now we get lots of orders"* (tracing data, everyday mobility).

However, as the interviewees highlighted, the analyst's choices may be restricted to just disseminating the data culture. Like the experts in other discourses, those in Discourse III saw legal restrictions, like limited access to data, as a significant obstacle to introducing innovative data solutions. The "traded" data relations, where an organisation is financially dependent on the customer, proved to be another obstacle to novel data solutions, as in Discourse I. The dependence on customers' expectations was evident in the case of third-sector organisations, as an interviewee working in that sector explained:

> *"On the other hand, we do the things we're paid to do* (…) *That sometimes it's sort of boring and simple"* (survey data, forced migration).

Unlike the previous two discourses, the interviewees in Discourse III, were mainly managers of their organisation or division, who introduced a data culture into each organisation as an alternative approach to data use. Instead of blind 'faith' in datafied solutions, where data were seen exclusively as an opportunity to disseminate social progress and innovation, Discourse III emphasises the responsibilities in governance through data. Despite possible obstacles, like resistance within their organisation, or financial dependence

Table 6
(Big) data self-reflections and organisational data culture (Discourse III).

No	Statement	D III	D I	D II	D IV
14	*Everything that's related to data fascinates me. I enjoy my daily work with data irrespective of the topic and the types of the data.*	5*	3	1	1
18	The big data gives private and public institutions an essential competitive advantage, they are considerably more efficient.	4*	2	3	2
4	*Information technologists only think about the speed of code. They often do not have the skills and knowledge to assess the consequences of algorithms in a 'socially critical' manner.*	2*	-2	-1	-3
33	The whole universe can be reduced to data flows. Organisms are like datasets, and the mission of humankind is to create a comprehensive data processing system.	0*	-3	-4	-5
19	The use of big data allows data to be collected cheaper and faster and thus to find solutions to societal issues more quickly.	0	2	2	4
24	*Estonia, with its digital society, is one of the leaders in the world; the big data opportunities for finding solutions to social problems.*	-3*	2	1	3
8	The blind faith of public and private institutions in digital technology can endanger national and individual security.	-4	1	2	0

* Statistically significantly distinguishing statements, p<=.01. The discourse is explained by four core and three additional interviewees.

on the traditionally-oriented customer, the experts in Discourse III actively disseminated their ideas on organisational data culture among partners and in public.

4.4. Discourse IV: Data discrimination aware data practices and privacy protection

We call Discourse IV—'Data discrimination aware data practices and privacy protection'. This title was based on the agreements of interviewees all working in private sector organisations as analysts. Most interviewees here worked with tracking data, distinguishing this discourse from the other three.

The interviewees agreed that privacy and individuals' low awareness of the use of their data by institutions are the most significant obstacles related to data technologies used for governance (statement #11, see Table 7), as the quantitative analysis revealed. The interviewees expressed scepticism regarding the opportunities for efficient anonymisation, especially in merging databases (#6). This discourse also entails self-evaluated analytical confidence regarding competencies with data analytics techniques (#13) and awareness of the possible power relations in data-based governance (#9). Their extensive experiences in private companies, and access to detailed data, could explain the awareness of privacy concerns (three of four interviewees). Unlike the other discourses, the interviewees here expressed that they were relatively independent in their decisions, having no obligation to provide input to governing bodies.

The qualitative analysis confirmed and illustrated the quantitative analysis results, in that enabling a more detailed explanation of what exactly were these power relations that quantitative analysis revealed. The qualitative analysis revealed that interviewees emphasised potential issues with the use of data, such as privacy concerns, discrimination, or depersonalisation challenges; they were also disenchanted with the usability of data, as there was little public awareness of the services provided through these technologies. Minority groups may be unaware of the data used and the services provided. Informing disadvantaged groups about these services may, however, be economically harmful to the state, as an interviewee expressed:

> *"If the state informs all the citizens entitled to a benefit or allowance, then the fear is that we don't have money. Today, our life is based on the assumption that some citizens don't come to exercise their rights"* (tracing data, everyday mobility).

The interviewees in Discourse IV believed the organisations that manage and use data are responsible for protecting the rights of the data subjects to avoid harm, for example, by protecting privacy. For example, an expert working with tracking data pointed out that people's competence in the field of data was too weak to allow them to adequately protect themselves against potential risks. The expert also raised the issue that individuals cannot protect their privacy since they did not control the data collected, especially when these data were given to third parties:

> *"How can I be responsible for my digital footprint when my mobile operator gives my data to others? I also don't know how the police are using my data"* (tracing data, everyday mobility).

The interviewees in this group also condemned the construction of exclusive data solutions, where the public did not have many opportunities to understand and evaluate the design process. Furthermore, they felt the selective approach led to biased solutions, which were mostly a threat in public sector organisations. The interviewees emphasised that the biases in data do not come from the decision-makers but the standards and opinions prevailing in society, carried into system development via digital technology workers. For example, one expert stressed the responsibility to the public was emphasised when designing and using data technologies for governance purposes in a specific context like mobility:

> *"The police should not create such an exclusive algorithm because the state's competency is what it is. The state could develop something good that would not have unwanted consequences. It should be possible to audit it publicly"* (tracing data, everyday mobility).

These claims regarding the development of "closed" and "exclusive" solutions were based on a mediated experience – the interviewees were aware of the possible negative consequences of data technologies communicated in the media and about global companies. Their claims about potential biases related to data relied on personal experiences, like one data expert, who considered it essential that the use of data in organisations within a context of governing human mobility must be transparent:

> "*Everything that isn't reasonably closed could be open* (…) *The restriction exists since nobody wants to say publicly that there is rubbish going on* (…)" (tracing data, forced migration).

Table 7
Discrimination-aware data practices and privacy protection (Discourse IV).

No	Statement	D IV	D I	D II	D IV
3	*One of the biggest problems is the lack of experts and analysts who know their field and have the experience to analyse big data.*	4	0	2	2
11	The primary threats with big data are personal data protection; people are not aware of who uses their 'digital footprints'.	3*	-2	0	-1
6	*Efficient anonymization may be challenging to achieve; privacy and personal data protection are the biggest threat with big data.*	3	0	0	-2
17	Private and public sector institutions have excellent opportunities to use big data to regulate cross-border mobility, control borders.	2	0	-1	0
9	*Data have become the most crucial source of power at present – data make it possible to control people's daily movements.*	1*	-2	-4	-2
13	In the institution where I work, there are ample opportunities to develop myself; I feel confident with the 'data revolution'.	1*	3	-2	3
29	*To strengthen data protection and digital opportunities, we need strong leadership, quick and bold political decisions.*	0	2	3	3

* Statistically significantly distinguishing statements, p<=.01. The discourse is explained by four core and three additional interviewees.

In Discourse IV, the interviewees perceived the data as a possible power source when used commercially and implemented through closed and exclusively developed data solutions. They saw open data as an opportunity to move towards a more transparent state and governance and believed organisations and the state were obliged to guarantee system audits to prevent any possible issues of data-related bias, which are often visible in the use of datafied solutions in organisations within mobility as a specific context. Whereas the experts in the other Discourses also emphasised the awareness of the potential negative consequences, those in Discourse IV were most critical regarding data-driven governance. Experiences as designers in the private sector have raised their awareness of the potential problems with governance through data in the context of human mobilities.

5. Conclusions

Our study highlighted the composite scenario of sociotechnical imaginaries—i.e., discourses on specific technological innovations and their use—that the data experts involved in governance and controlled through data in the specific context of spatial mobilities expressed and put forth. The experts' visions were particularly relevant as data experts can be regarded as 'insiders'; they have a cognizant understanding of how data-driven technologies worked, manifesting in their everyday work with data.

The analysis identified four sociotechnical imaginaries on data, as seen by the experts working in the field of spatial mobility and related data: (I) protection of rights to equal access and the development of universal data internationalism across state borders; (II) balancing between disciplinary resources and continuity and cross-sectorial cooperation in data relations; (III) introducing innovation through data and developing a self-reflective data culture in the organisations, and (IV) ensuring data discrimination aware data practices and appropriating understandings on privacy protection. This study demonstrated that there are parallel and often contradictory imaginaries where public values, social norms and routine practices among data experts may affect the specific process of datafied governance.

This research revealed that the imaginaries on the datafied governance were differently understood by experts active in the diverse areas of mobility and working with various types of data. Our study revealed that social and ethical aspects (Discourses II, III, and IV) were more expressed by experts working in public, private and third-sector organisations, who are predominantly involved in working with data related to everyday mobility and forced migration. In these four discourses, the limitations the interviewees considered affecting individual agencies were emphasised through the institutional and organisational responsibilities to develop self-reflective data culture in their organisations. The results also indicated that management often tend to have more concrete expectations related to data-based futures and the means for materialising these in an organisational context. Also, experts working as analysts were much more affected by cooperation across levels, producing various imaginaries centred around the ideals of openness and universalism.

6. Discussion

This study strove to contribute to the discussions on the social imaginaries of the data (see e.g., Amoore, 2019; Hepp, 2016; Jasanoff & Kim, 2015; Just & Latzer, 2017; Meng & Disalvo, 2018; O'Neil, 2016; Schrock & Shaffer, 2017) and the potential social and ethical implications of mobility data mining (Berendt et al., 2022). This article strove to reveal and test the social imaginaries of the data using the field of mobility as a specific case study and relying on the positions of data professionals. Whereas some of the discourses that emerged in this empirical research have received more academic attention, others have remained somewhat less explored.

Some social imaginaries on data have received broader academic attention, especially Discourses II and IV. For example, Discourse IV—on developing data discrimination aware practices, ensuring trust in data authorities and appropriating understandings of fairness and privacy protection—emphasised the emerging and ongoing tensions between the interests of individual data agencies and those of authorities. Therefore, this argument confirms research on constant negotiations about fairness and trust principles in governance (Guenduez et al., 2020), which is needed to avoid potential negative consequences of implementing data technologies (Eubanks, 2018; Saurwein, Just, & Latzer, 2015; Masso & Kasapoglu, 2020). Discourse II balances disciplinary resources, continuity, and cross-sectorial cooperation. Consequently, this imagination introduces alternatives to the publicly spread understanding of data as a form of social progress, which previous research has emphasised (Guenduez et al., 2020; Morozov, 2017).

The other two imaginaries (Discourses I and III) on datafied governance have received less attention in previous research. The social imaginary on assuring equal access and developing universal data internationalism (Discourse I) highlights the challenges in sharing, exchanging, and using data across state borders, besides previously emphasising data subjects' rights to collect and manage their data (Hummel et al., 2021). Moreover, the emerged discourse suggests the necessity to introduce governance principles in situations where data instead of people move across state borders and challenge the understanding of social diversities. The last imaginary (Discourse III) considers introducing innovation through data and developing self-reflective data culture, revealing the necessity to introduce standardised requirements for individuals' critical reflexivity in data work. Shifts in reflexive decisions are decisive, not only to avoid unintended risks with data technologies but to consider the diversities in data moving across state borders, besides technical data literacy skills in governance data work. Using data technologies in specific fields like mobility entails innovative location-based services with context and diversities aware services based on the data produced by individuals, as this empirical study proved.

Therefore, this research highlights that social imaginaries on data technologies not only reflect the state-of-the-art of implementation and practices with data analytics in a particular organisation, at a certain time point, being visible through the meanings and understandings the data experts have expressed (Guenduez et al., 2020; Hepp, 2016; Meng & Disalvo, 2018; Morozov, 2017; O'Neil, 2016). Instead, as this research revealed, the social imaginaries on data technologies enable us to trace the social shifts in the

organisations through the perspective of the agents developing and using datafied solutions. The analysis of four discourses indicated several ways in which the cultural shift towards the data-based organisation can happen. As Discourse III revealed, such a datafied shift may entail experts taking a responsible and active role in disseminating their data ideals within their organisation and the public. The changes towards data culture may also be extended beyond organisational and national borders, where technical aspects of data are turned into social value (Discourse IV). However, as Discourses II and IV revealed, the shift toward data culture necessarily entails developing a self-reflexive approach toward the data-driven forms of governance. The extent to which discourses are technologies of power and, in turn, enforce the realisation of technological futures, it is crucial to put forth self-reflexive practices (O'Neil, 2016; Schrock & Shaffer, 2017; Thomas et al., 2018) aimed at decoupling the 'technological' and the 'discursive' to reinstate control over data governance as a societal dynamic.

Therefore, this analysis of these data imaginaries indicates a variety of evolutions into data culture in institutions, which may not only occur in the form of previously stated 'enculturation' (Smith, 2018) with the view of data as a form of social progress and ultimate medium of knowledge. Instead, we suggest that these cultural shifts may also occur in the form of 'data acculturation' – data-informed, critically self-reflexive forms of governance. Acculturation through data entails engagement in organisational data culture and the larger society and adaptation to data technologies in governance with necessary critical and self-reflexive competencies. Through data acculturation, actors dealing with data as means for decision-making prevent themselves from being turned into (datafied) means, thus avoiding that instrumentalization that language has already undergone in its foregoing entanglement as a technology. The socio-technical imaginaries of datafication in organisations, relying on a specific example like mobility, reflect why 'data acculturation' as socialisation into data culture has been painful, entailing negative consequences. The 'data acculturation' concept enables us to strive to open discussions about the shifts towards awareness of, and orientation to, social (counter)imaginaries arising from using datafied technologies in governance. Based on empirical analysis, we assume that organisations within a specific context, like mobility, face a double process. Besides shifts towards datafied solutions, the social and cultural adaptation and change processes related to data subjects should be considered. However, we still do not know if these cultural and organisational shifts are inherent to using data-based solutions in organisations in general or only within a specific context like mobility.

This article is not free of limitations. This research has taken a retrospective perspective by examining the social imaginaries of data experts in a period when the most rapid changes in organisations using data-based technologies were happening. While the study of social imaginaries provides a mapping of possible understandings of data that may manifest in everyday practices in the future, we still do not know which discursive strategies, how and why have led to different outcomes. Therefore, we suggest that future research on the social imaginaries of data can take one step further in empirical mapping and examining these social, cultural and structural shifts in the data literacy skills of both organizations and individuals.

Funding

This study was financed by the Development Program ASTRA of Tallinn University of Technology for 2016-2022 (2014-2020.4.01.16-0032), European Commission through the H2020 project Finest Twins (grant No. 856602).

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication.

Acknowledgements

We would like to thank the anonymous rerviewers for their valuable comments that improved the quality of the paper. We are most appreciative of the help of all the interviewees and their readiness to participate in this study.

References

Ahas, R., & Mark, Ü. (2005). Location based services—New challenges for planning and public administration? *Futures, 37*(6), 547–561. https://doi.org/10.1016/j.futures.2004.10.012

Amoore, L. (2019). Doubt and the algorithm: On the partial accounts of machine learning. *Theory, Culture & Society, 36*(6), 147–169. https://doi.org/10.1177/0263276419851846

Aradau, C., & Blanke, T. (2022). *Algorithmic reason: The new government of self and other*. Oxford University Press.

Austin, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.

Berendt, B., Matwin, S., Renso, C., Meissner, F., Pratesi, F., Raffaetà, A., & Rockwell, G. (2022). *Mobility Data Mining. From Technical to Ethical, 12* pp. 35–66). Dagstuhl Reports. https://doi.org/10.4230/DagRep.12.1.35

Broeders, D., & Dijstelbloem, H. (2016). The datafication of mobility and migration management: The mediating state and its consequences. In I. Van der Ploeg, & J. Pridmore (Eds.), *Digitizing identities: Doing identity in a networked world* (pp. 242–260). Routledge.

Bucher, T. (2016). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 1–15. https://doi.org/10.1080/1369118X.2016.1154086

Burt, C. (1937). Correlations between persons. *British Journal of Psychology, 28*, 59–96.

Carlson, M. (2018). Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism. *New Media & Society, 20*(5), 1755–1772. https://doi.org/10.1177/1461444817706684

Cattell, R. (1973). *Personality and mood by questionnaire*. Jossey-Bass.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St Martin's Press.

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. Routledge.

Finch, W. H. (2019). *Exploratory factor analysis*. SAGE Publications.

Foucault, M. (1980). In Colin Gordon (Ed.), *Power/knowledge: Selected interviews and other writings, 1972–1977*. New York: Pantheon Books.
Grosman, J., & Reigeluth, T. (2019). Perspectives on algorithmic normativities: Engineers, objects, activities. *Big Data & Society, 6*(2). https://doi.org/10.1177/2053951719858742
Guenduez, A. A., Mettler, T., & Schedler, K. (2020). Technological frames in public administration: What do public managers think of big data? *Government Information Quarterly, 37*(1). https://doi.org/10.1016/j.giq.2019.101406
Haardörfer, R. (2019). Taking quantitative data analysis out of the positivist era: Calling for theory-driven data-informed analysis. *Health Education & Behavior, 46*(4), 537–540. https://doi.org/10.1177/1090198119853536
Hepp, A. (2016). Pioneer communities: Collective actors in deep mediatisation. *Media, Culture & Society, 38*(6), 918–933. https://doi.org/10.1177/0163443716664484
Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society, 8*(1). https://doi.org/10.1177/2053951720982012
Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research, 70*, 338–345. https://doi.org/10.1016/j.jbusres.2016.08.007
Jasanoff, S., & Kim, S.-H. (2015). *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. Chicago, IL: University of Chicago Press.
Just, N., & Latzer, M. (2017). Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture & Society, 39*(2), 238–258. https://doi.org/10.1177/0163443716643157
Jutel, O. (2021). Blockchain Imperialism in the Pacific, Big Data & Society, 8: 1, https://doi.org/10.1177/2053951720985249.
Kummitha, R. K. R. (2020). Smart technologies for fighting pandemics: The techno- and human- driven approaches in controlling the virus transmission. *Government Information Quarterly*. , Article 101481. https://doi.org/10.1016/j.giq.2020.101481
Latour, B., & Woolgar, S. (1979). Laboratory life: The social construction of scientific facts. *Sage, 80*.
Latzer, M., and Just, N. (2020). Governance by and of Algorithms on the Internet: Impact and Consequences, Oxford Research Encyclopedia of Communication, https://doi.org/10.1093/acrefore/9780190228613.013.904.
Lupton, D. (2016). The Quantified Self: A Sociology of Self-Tracking. Cambridge: Polity.
Lyotard, J. (1984). *The postmodern condition: A report on knowledge*, G. Bennington and B. Massumi (Trans), University of Minnesota Press, Minneapolis.
Männiste, M., and Masso, A. (2020). 'Three Drops of Blood for the Devil': Data Pioneers as Intermediaries of Algorithmic Governance Ideals." *Mediální Studia | Media Studies* 14: 1, 55–74. https://www.medialnistudia.fsv.cuni.cz/front.file/download?file=medialni_studia_1_2020%2004%20manniste_masso.pdf.
Masso, A., & Kasapoglu, T. (2020). Understanding Power Positions in a New Digital Landscape: Perceptions of Syrian Refugees and Data Experts on Relocation Algorithm. *Information, Communication & Society, 23*(8), 1203–1219. https://doi.org/10.1080/1369118X.2020.1739731
Masso, Anu, Maris Männiste, & Andra, Siibak (2020). End of Theory' in the Area of Big Data: Methodological Practices and Challenges in the Social Media Studies. *Acta Baltica Historiae et Philosophiae Scientiarum, 8*(1), 33–61. https://doi.org/10.11590/abhps.2020.1.02
McKeown, B., & Thomas, D. (2013). *Q Methodology* (2nd ed.). Sage, 7-066, Ed. 2.
Mcquillan, D. (2016). Algorithmic Paranoia and the Convivial Alternative, Big Data & Society, 3: 2. https://doi.org/10.1177/2053951716671340.
McStay, A., and Rosner, G. (2021). 'Emotional Artificial Intelligence in Children's Toys and Devices: Ethics, Governance and Practical Remedies, Big Data & Society, 8: 1, https://doi.org/10.1177/2053951721994877.
Meng, A., & Disalvo, C. (2018). Grassroots resource mobilization through counter-data action. *Big Data & Society, 5*(2). https://doi.org/10.1177/2053951718796862
Milan, S., & Treré, E. (2019). Big data from the south(s): Beyond data universalism. *Television & New Media, 20*(4), 319–335. https://doi.org/10.1177/1527476419837739
Molnar, P. (2021). Technological testing grounds and surveillance sandboxes: Migration and border technology at the frontiers. *The Fletcher Forum of World Affairs, 45*(2), 109–118.
Morozov, E. (2017). Opposing the exceptionalism of the algorithm. In Mirko Tobias Schäfer, & Karin van Es (Eds.), *The datafied society: Studying culture through data* (pp. 245–248). Amsterdam University Press.
O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy* (First ed.). New York: Crown.
Ojala, M., Pantti, M., & Laaksonen, S.-M. (2019). Networked publics as agents of accountability: Online interactions between citizens, the media and immigration officials during the European refugee crisis. *New Media & Society, 21*(2), 279–297. https://doi.org/10.1177/1461444818794592
Pink, S., Ruckenstein, M., Willim, R., & Duque, M. (2018). Broken data: Conceptualising data in an emerging world. *Big Data & Society, 5*(1). https://doi.org/10.1177/2053951717753228
Ramlo, S. (2016). Centroid and theoretical rotation: Justification for their use in Q Methodology Research. *Mid-Western Educational Researcher, 28*(1), 73–92, 73-Western Educational Researcher, 2016, Vol.28(1).
Ramlo, S. E., & Newman, I. (2011). Q methodology and its position in the mixed-methods continuum. *International Society for the Scientific Study of Subjectivity ISSSS*. https://doi.org/10.15133/j.os.2010.009
Ranerup, A., & Henriksen, H. Z. (2019). Value positions viewed through the lens of automated decision-making: The case of social services. *Government Information Quarterly, 36*(4), Article 101377. https://doi.org/10.1016/j.giq.2019.05.004
Rieder, B. (2016). Big data and the paradox of diversity. *Digital Culture & Society, 2*(2). https://doi.org/10.14361/dcs-2016-0204
Saurwein, F., Just, N., & Latzer, M. (2015). Governance of algorithms: Options and limitations. *Info, 17*(6), 35–49. https://doi.org/10.1108/info-05-2015-0025
Scheel, S., Ruppert, E., & Ustek-Spilda, F. (2019). Enacting migration through data practices. *Environment and Planning D: Society and Space, 37*(4), 579–588. https://doi.org/10.1177/0263775819865791
Schrock, A., & Shaffer, G. (2017). Data ideologies of an interested public: A study of grassroots open government data intermediaries. *Big Data & Society, 4*(1). https://doi.org/10.1177/2053951717690750
Smith, G. J. D. (2018). Data doxa: The affective consequences of data practices. *Big Data & Society, 5*(1). https://doi.org/10.1177/2053951717751551
Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology*. University of Chicago Press.
Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative Research Journal, 11*(2), 63–75. https://doi.org/10.3316/QRJ1102063
Tammpuu, P., and Masso, A. (2018). 'Welcome to the Virtual State': Estonian e-Residency and the Digitalised State as a Commodity." *European Journal of Cultural Studies,* 21: 5, 543–60. https://doi.org/10.1177/1367549417751148.
Taylor, L. (2016). No place to hide? The ethics and analytics of tracking mobility using mobile phone data. *Environment and Planning D: Society and Space, 34*(2), 319–336. https://doi.org/10.1177/0263775815608851
Taylor, L., & Purtova, N. (2019). What is responsible and sustainable data science? *Big Data & Society, 6*(2). https://doi.org/10.1177/2053951719858114
Thomas, S. L., Nafus, D., & Sherman, J. (2018). Algorithms as fetish: Faith and possibility in algorithmic work. *Big Data & Society, 5*(1). https://doi.org/10.1177/2053951717751552
Thornham, H., & Gómez Cruz, E. (2016). Hackathons, data and discourse: Convolutions of the data (logical). *Big Data & Society, 3*(2). https://doi.org/10.1177/2053951716679675
van Lente, H. (2000). Forceful futures: From promise to requirement. In Nik Brown, & Brian Rappert (Eds.), *Contested futures: A sociology of prospective techno-science* (pp. 43–64). New York: Ashgate.
Watts, S., & Stenner, P. (2012). *Doing Q methodological research: Theory, Method & Interpretation*. SAGE.
Wodak, R. (1996). *Disorders of discourse*. Longman.
Woolf, N., & Silver, C. (2017). *Qualitative analysis using MAXQDA: The five-level QDA™ method*. Routledge.