LGBTQI+ Migrants in a Datafied City: A Qualitative Study on the Use of (Geo)Data in Amsterdam

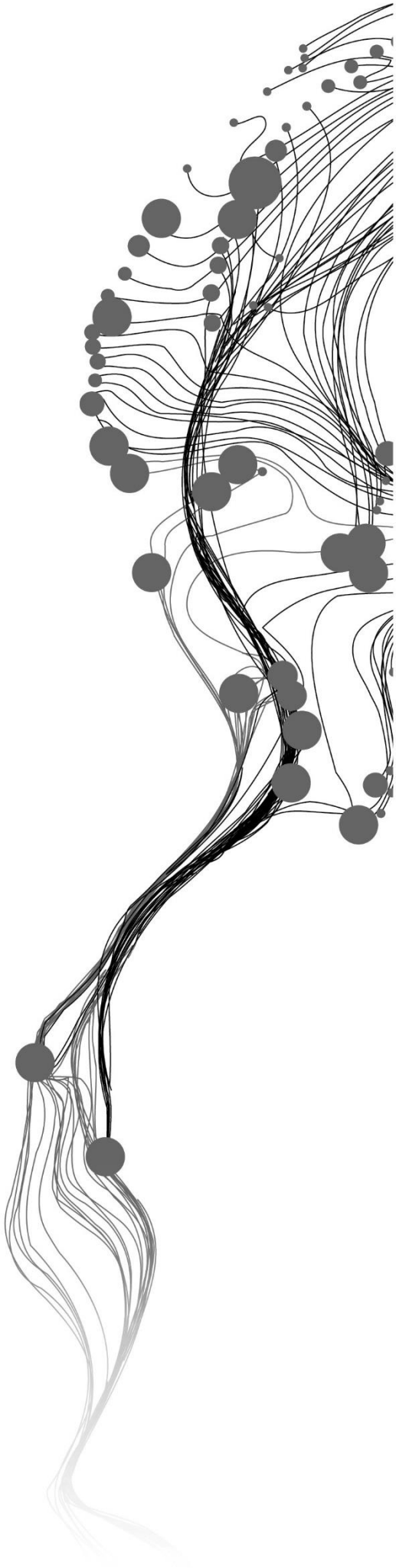
UDIPTA BORO

July 2021

SUPERVISORS:
Dr. A.M. (Ana) Bustamante Duarte
Dr. F.V.M. (Fran) Meissner

ADVISOR:
Prof. Dr. K. (Karin) Pfeffer

THESIS ASSESSMENT BOARD:
Dr. J.A. (Javier) Martinez (Chair)
Dr. K.H.A. (Koen) Leurs (External Examiner, Universiteit Utrecht)



LGBTQI+ Migrants in a Datafied City: A Qualitative Study on the Use of (Geo)Data in Amsterdam

UDIPTA BORO
Enschede, The Netherlands, July 2021

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Urban Planning and Management

SUPERVISORS:
Dr. A.M. (Ana) Bustamante Duarte
Dr. F.V.M. (Fran) Meissner

ADVISOR:
Prof. Dr. K. (Karin) Pfeffer

THESIS ASSESSMENT BOARD:
Dr. J.A. (Javier) Martinez (Chair)
Dr. K.H.A. (Koen) Leurs (External Examiner, Universiteit Utrecht)

ABSTRACT

Data is the most valuable resource of the 21st century. Cities such as Amsterdam are adopting a data-driven approach to develop people-oriented services and elevate city residents' quality of life. In the data-driven development of cities, the use of (geo)data also plays a prominent role. While the rapid datafication of cities can be viewed as a step towards futuristic urban development, cities must take into account the growing concerns on the invasion of people's privacy, hypervisibility, or biased profiling of the data subjects. Such issues are of particular importance when looked at from the perspectives of groups that have been historically discriminated against. Drawing on critical data studies literature, my research aims to understand such (geo)data concerns from the perspective of self-identified LGBTQI+ migrants living in the rapidly datafying city of Amsterdam. After identifying the concerns of my interlocutors, I ask how far-reaching the city's data policies are to account for those concerns. I conducted eight semi-structured online interviews with LGBTQI+ migrants living in Amsterdam and collected 32 reports from the Personal Data Commission, Amsterdam between the years 2017 to 2020, to gain insights on the latter. The data was analyzed using thematic analysis. From the interviews, three main themes emerged about the respondents' concerns and actions over the use of their (geo)data: 1) Safety and Convenience Trump Privacy; 2) Awareness of Datafication Shapes Perceived Risk; and 3) Consent is Subject to Power Dynamics in Urban Datafication. The city reports, however, have limited response to the concerns identified. I critically look at this asymmetry in concerns and response through the lens of data justice. Based on these findings, I argue that a city needs more than mere data collection to support its marginalized community; the city must acknowledge and address the unequal power dynamics to build a just and inclusive datafied city.

Keywords: *datafication, data-driven development, (geo)data, data concerns, LGBTQI+, migrants, marginalized communities, Amsterdam, data justice*

"Take away, for the moment, the identifiable markers of the gay and lesbian experience, and imagine a social protest movement that, throughout the twentieth century, has created an independent urban culture, suffered police harassment, been legally subject to housing and employment discrimination, and, in response, waged a campaign for social justice that has intensified over the past fifty years. Then imagine, that, as planning historians, we have overlooked these experiences. If nothing else, the implausibility of this occurrence marks the gay and lesbian experience as worthy of current attention."

~ Moira Rachel Kenney
1998
Remember, Stonewall was a Riot: Understanding Gay and Lesbian Experience in the City

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my supervisors Ana Maria Bustamante Duarte and Fran Meissner for without your guidance and constant support, I would never have been able to write this thesis. Your insightful yet friendly mentoring kept me going even when I was finding myself at sea. Thank you for being patient with me and inspiring me with your kind words. I could not have wished for any better supervisors. I am grateful to my advisor Karin Pfeffer for chiming in with valuable feedback and critical questions which immensely helped me structure the arguments I make in this thesis. I am indebted to my research participants. This thesis would not have come to life without your help. Thank you for investing your time with me.

I am thankful to ITC Excellence Scholarship Program for making it (financially) possible for me to study and conduct research at ITC. I am proud and grateful to be part of the diverse and talented ITC community.

My friends from ITC and beyond, what an amazing support system you have been! I send my gratitude to Deepak, Aparupa, Sadichchha, Priscilla, Liu, and Surakshya for checking on me and keeping me afloat the river called *Doing A Master's Degree*. Matthijs, every time I felt the blues, you were there to cheer me up. I owe you big time for keeping me safe and sane. A big thank you to Ken for always being there for me despite your busy schedule and the six-hours' time difference. Chandrama, I cannot thank you enough for being the person I could look up to since my bachelor's. You all have made this journey possible, thank you!

Now please excuse the shift in language as I thank my parents: মা, দেউতা, তোমালোকৰ মৰম আৰু সাহস অবিহনে আজি হয়তো মই এই স্থানত নাথাকিলোঁহেঁতেন। আজি মই যি, সেয়া কেৱল তোমালোকৰ কাৰণেই সম্ভৱ হৈছে। আৰু তাৰবাবে মই তোমালোকৰ ওচৰত চিৰঋণী।

And last but not least, I would like to thank the Almighty for giving me the courage to carry out this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CPA: Commissie Persoonsgegevens Amsterdam (Amsterdam Personal Data Commission)

FDP: Forcibly Displaced Population

FRA: European Union Agency for Fundamental Rights

GDPR: General Data Protection Regulation

LGBTQI+: Lesbian, Gay, Bisexual, Transgender, Queer, Intersex+

1. INTRODUCTION

Data has become one of the most valuable resources in the present times. While data is termed as the "new oil" by many (The Economist, 2017), recent voices have critiqued this analogy by pointing out that data is not precisely oil, but people who are being traded in the form of quantifiable/digital information in a complex market system (Martínez, 2019; Naughton, 2021). In the modern technology-dependent society, data is used and generated in almost every facet of our lives. In this thesis, I particularly focus on geodata. Geodata is a broad term that ranges from earth observation data to specific location data. For my research, I am defining geodata as spatially tagged digital data that can reveal someone's spatial movements or location (Taylor et al., 2016). Such geodata is produced by, for example, our use of the mobile phones that emit spatial information, the use of smart travel cards or the GPS, our digital interaction with the city to receive services and offer feedback, and the city authorities' attempt at crowd control through surveillance and monitoring (Taylor et al., 2016). It is noteworthy that geodata is intrinsically linked with other personal data[1]. This entanglement may create confusion regarding the meaning of the term "geodata" in some parts of the thesis. To avoid such confusion, I would use the term "geodata" to refer to those data that strictly fit under the previously stated definition and the term "(geo)data when geodata is entangled with other personal data.

In cities such as Amsterdam (Gemeente Amsterdam, 2019a), there is extensive use of (geo)data in the process of datafication – "the growing presence, use and impact of data in social processes" (Heeks & Shekhar, 2019; p. 992). That implies that (geo)data is extensively used in various processes and actions such as service provision, decision-making, and urban governance. (Geo)Data is also used to understand migration flows and inform migration-related policies as well as to monitor and control these migration flows (Gillespie, Osseiran, & Cheesman, 2018; Latonero & Kift, 2018). Urban datafication, informed by (geo)data, is claimed to improve urban quality of life, facilitate decision-making, and bring the city's marginalized communities (such as migrants) to the purview of urban policies/decisions (Heeks, Graham, Evans, & Taylor, 2020). However, datafication may also exasperate the urban fault-lines of the existing socio-economic inequalities inducing manifold discrimination leading to further marginalization of those already marginalized or at the risk of marginalization (Heeks et al., 2020; Redden, Brand, & Terzieva, 2020).

These issues are of particular interest if looked at from the perspective of groups that have been historically discriminated against. One such group is the LGBTQI+[2] migrant community who faces multiple discrimination and social inequalities (Carroll & Itaborahy, 2015; FRA[3], 2015; Gatehouse et al., 2018) and needs special support from the authorities in combating these issues. Support can be provided by collecting and analyzing (geo)data from the community in question and making policy decisions based on such data (e.g. assess their specific needs in specific neighborhoods). However, one must consider

[1] The GDPR (2016, p.33) defines personal data as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

[2] An umbrella term referring to those who identify as lesbian, gay, bisexual, trans, queer, intersex, asexual, etc. In this thesis, I use this term to refer to anyone who identify as non-heterosexual or non-cisgender.

[3] European Union Agency for Fundamental Rights

issues such as privacy and security before collecting and analyzing their (geo)data. These (geo)data can also be used by data collectors to track human behavior up to the individual level potentially breaching individual privacy (Armstrong & Ruggles, 2005). Moreover, such (geo)data collection practices must also meet the data collection regulations of that particular country/area.

In this qualitative research, I explore the (geo)data practices and concerns of Amsterdam's self-identified LGBTQI+ migrant community and ask if the city's policies are far-reaching enough to account for those concerns. I mainly draw from *data justice* (Heeks & Shekhar, 2020; Taylor, 2017) to critically comment on the LGBTQI+ migrant community's position in the discussion surrounding Amsterdam's data-driven development. The concept of data justice looks into how data and social justice intersects with each other to tackle the questions of fair representation, treatment, and visibility of people as an outcome of their creation of digital data; and how data is related to power and inequality – particularly focusing on those at the risk of marginalization – which determines who gains and who loses from the process of datafication in the physical city (Heeks & Shekhar, 2020; Taylor, 2017). As the LGBTQI+ migrants are prone to facing intersectional discrimination and oppression, the concept of data justice is useful for assigning meaning to the experiences of the former and situating them in the emerging discussion of social justice in the process of datafication.

1.1. Background and Justification

Migration is a complex and multi-faceted phenomenon and is crucial for economic development and/or to protect or prosper in life. Migration includes economic migration, forced migration, migration in search of a better lifestyle, or migration as a way of life for some nomadic groups (Boyle, Halfacree, & Robinson, 2013). These categorizations of migration types, however, are often difficult to disentangle in the lived experience of being a migrant. Migration streams also consist of the members of the LGBTQI+ community as Szulc (2020; p.220) points out, "Some migrants are queer[4]. Some queers are migrants.". Manalansan IV (2006) states that sexuality and sexual idenitities are one of the major factors for migration. And for many of the LGBTQI+ migrants, the LGBTQI+-friendly European countries are the preferred destination (Mole, 2021). As of May 2015, the members of the LGBTQI+ community risk being criminalized in a total of 76 countries (Carroll & Itaborahy, 2015). However, the reason for the LGBTQI+ migration is not merely confined to persecution or threats to life due to their sexual orientation or gender identity. Roshchupkin (2020; p.10) states that "46% of migrant MSM [men who have sex with men] and trans people move to other countries in search of work, 27% for education, and almost 16% are looking for a social environment where they, as trans people, gay or bisexual, can live in greater security than in their home countries". Moreover, a major part of migration flows towards cities owing to what Boyle et al. (2013) term as the "lure of the city" (e.g., safety, more opportunities, economic prosperity).

It must be noted that migrants face a range of challenges in their destination countries or more specifically, in the cities they migrate to. These challenges rise above the linguistic or cultural barriers that migrants face in host cities. These barriers may already cause social exclusion by intensifying serious societal problems such as lack of access to housing or employment, a sense of isolation, or lack of trust among social groups (Bernardino & Santos, 2021). Europe is also witnessing a rise of xenophobic populist nationalism across much of the continent leading to migrants being increasingly seen as threats to local culture and traditions (French Brennan, 2017). Adding to these challenges, FRA (2020) reports that discrimination and violence against people based on sexual orientation or gender identity still remain a problem across the EU. The members of the LGBTQI+ community, in general, are also vulnerable to

[4] Many scholars use the term 'queer' to refer to the whole LGBTQI+ community or part of it. While citing the work of those scholars, I am keeping the term unchanged.

hate crimes in the form of physical/verbal abuse, or social discrimination (Gatehouse et al., 2018). Situated at the intersection of being migrants and LGBTQI+, the LGBTQI+ migrants, therefore, may face exacerbating intersectional discrimination and inequality in Europe. To combat such challenges, further informed governance arrangements such as policies, programs, and strategies to support communities at the risk of marginalization are of particular importance. Such efforts aimed at evidence-based policymaking may help recognize how LGBTQI+ migrants can more effectively be supported and receive specific services.

One way for the government to make informed decisions and deliver people-oriented services is to collect and analyze people's geodata (Traunmueller et al., 2018; Witanto et al., 2018). Today, almost all our actions produce geodata. By tracking our movements through geodata, it becomes possible to paint the most detailed picture of our activities (Taylor et al., 2016). The traces of our lives we leave digitally can be used to create our digital selves – often inaccurate and distorted compared to the real selves – termed as 'data doubles' (Jones, 2018; Moe-Pryce, Bellanova, & Bergersen, 2016). Haggerty & Ericson (2000; p.606) note that today's world of an emergent surveillant assemblage "operates by abstracting human bodies from their territorial settings and separating them into a series of discrete flows. These flows are then reassembled into distinct 'data doubles' which can be scrutinized and targeted for intervention". There are various ways of tracking people's digital traces in a datafied city. Surveillance (CCTV) cameras (Jameson et al., 2019), Wi-Fi or mobile phone signals (Taylor, 2018; Traunmueller et al., 2018), geotagged social media data (Witanto et al., 2018), or even digital payments through bank cards (Lerman, 2013; Taylor et al., 2016) or smart transport cards (Gutiérre et al., 2020) – all these have the potential to record our every move which can be used by city authorities to analyze our behavior and improve their services and policies. For example, Enschede, a city in the eastern part of the Netherlands, is using smart traffic sensors for traffic management in the city. A phone's Wi-Fi signal is picked up by these smart traffic sensors and the MAC[5] address of the smartphone is recorded by the register. By analyzing this data, the city wants to understand the travel behavior (preferred routes and favorite spots) of people who visit Enschede (Naafs, 2018). In Stratumseind, Eindhoven, on the other hand, the city is installing cameras, microphones, and Wi-Fi trackers in lampposts as a means to control crowd behavior (Naafs, 2018). Technological projects such as this give rise to several socio-political, legal, and ethical concerns including data security and privacy issues and increased surveillance in cities (Galič, 2018).

However, there have been growing concerns about the increasing datafication of cities. Concerns have been raised about breaching people's privacy, tracking individual or group behavior, and indirect or biased profiling among many others. For instance, Traunmueller et al. (2018) point out that digital technologies such as facial recognition methods in a datafied city are breaching individual privacy. While biometrics are important in the process of datafication, they do not necessarily have the "geo" component. Geo-tracking and surveillance technologies such as tracking of mobile devices, GPS or Wi-Fi signals, IP addresses, and CCTV cameras may violate individual and/or group privacy by collecting and processing people's location information or geodata (Bridges, 2019; Taylor, Floridi, & Van Der Sloot, 2017). Tracking location information may reveal a lot about a person's life. For example, knowing a person's home address may tell us about their income level based on the area they live in. Moreover, tracking a person's daily commute may inform us about their office location or even their job. There is a risk of such data being accessed by other parties for malicious use. This type of data can be used by businesses to target the right customers or by governments to make targeted interventions. Such data may also be used as a proxy for income which might inadvertently disadvantage an individual or a group of individuals (e.g., creditworthiness). In a datafied city, these risks are even higher as Wadhwa (2015; p.130) states "to a hacker, the smart city is a playground unlike any other". In addition to that, Taylor (2018) in her work paints an illustrative case of

[5] MAC stands for Media Access Control, a unique identifier for every device

how a digital city might collect data about a person that can be used to create profiles and how these profiles might be biased towards a specific group of people. Such cases put the LGBTQI+ migrant community at serious risk. As previously mentioned, this group of people may face multiple discrimination and social exclusion and is subject to abuse or hate crime. That is why it is very important to engage with this particular group and their data concerns in a datafied city.

While many scholars have worked on datafication of cities and its potential downsides, no work, to the best of my knowledge, has been carried out that has investigated the LGBTQI+ migrants' position in a datafied city. It is very important to ask what data concerns the LGBTQI+ migrants may have, how they want their data to be handled, or if and how they want to trade off data for services. With the growing discussion surrounding smart yet inclusive urban development, it is time that we included communities at the risk of marginalization in our vision for urban development. Kitchin (2019) argues that a 'genuinely humanized' smart city is built around the concepts of democracy, ethics, equity, and fairness and that it addresses the issues of inequality and discrimination bringing more inclusivity to the smart urban development discourse. This notion of a datafied city, he further states, is reinforced by social justice, public-led intervention, and citizen participation.

1.2. Amsterdam – a haven for innovation and inclusion through technology?

I selected Amsterdam, the capital city of the Netherlands, as the field site for my research project. Amsterdam actively advertises itself as a progressive city in its data-driven development practices. The city's innovative approaches have made it a pioneer in data-driven urban development. Amsterdam is, in fact, the first municipality in Europe to have initiated a smart city program extending its previous digital city program (Dameri, 2014). The city is working on becoming more and more digital in order to provide its residents with better and effective services (Gemeente Amsterdam, 2019a). The European Digital City Index (EDCi, 2016) has ranked Amsterdam in the top three among other European cities and termed the city as a living lab for innovative experiments. Following Amsterdam's collaborative and creative technology-driven development against complex urban challenges, the European Commission in 2016 also named the city as the European Capital of Innovation (European Commission, 2016). With the rapid datafication of the city, important questions about data collection and processing and privacy protection come into play. Apart from the GDPR (2016), the city has its own urban framework to deal with data-related issues. For example, the Urban Framework for the Processing of Personal Data (Gemeente Amsterdam, 2018). This framework builds on the principles of anonymity of data subjects, control over and access to data, and addressing data-led discrimination in the city. Amsterdam also has a Personal Data Commission (Commissie Persoonsgegevens Amsterdam, CPA) (Gemeente Amsterdam, 2021). This commission advises the municipality on the implementation of privacy policies and guides official organizations in case of complex and politically sensitive matters related to personal data.

The city also boasts of its progressive tradition of being accepting of the LGBTQI+ community. The Netherlands is the first country in the world to have legalized same-sex marriage in 2001 (Dunbar, 2012). Amsterdam, too, is termed as the 'global gay capital' where the Dutch place gender and sexuality at the center of the city's rich and progressive LGBTQI+ history (French Brennan, 2017). The city had decriminalized homosexuality back in 1811 and has been an epitome of LGBTQI+ rights since then (IAmsterdam, 2021). In 2015, the city adopted the Pink Agenda (Gemeente Amsterdam, 2015). This document laid out the city's policies for LGBTQI+ inclusion and protection in the city. These policies are aimed at reducing intolerance and discrimination towards the LGBTQI+ community. Apart from the efforts to make Amsterdam LGBTQI+-inclusive, the Pink Agenda also sought to bring an inter-city collaboration for promoting LGBTQI+ social acceptance by being part of external networks such as the

Rainbow City Network. The Pink Agenda was based on the motto that Amsterdam belongs to everyone and through its inclusive policies, the city did strive to make that happen. After the Pink Agenda which was effective until 2018, the city adopted another set of policies under the name Nota Regenboogbeleid (Rainbow Policy Memorandum) for the period 2019-2022 (Gemeente Amsterdam, 2019b). As an extension of the Pink Agenda, this policy puts emphasis on acceptance, equality, and emancipation of Amsterdam's LGBTQI+ community. The policy aims to create inclusive spaces in the city for people of all gender identities or sexual orientations. The vision of the Rainbow Policy is to empower Amsterdam's LGBTQI+ community to fully participate in society without being confronted with violence, discrimination, insecurity, and social exclusion. Through these policies, the city seeks to go beyond just merely offering support, it wants to take an extra step to engage with the communities. That includes not just respecting the differences among Amsterdam dwellers, but to celebrate and make them visible.

Amsterdam also has a long migration history. Large-scale migration to the Netherlands dates back to the 16th century (Hoekstra, 2014). As the capital of the Netherlands, Amsterdam's population consisted of around thirty percent non-western people in the 17th and 18th century which was three times more than the national average (Lucassen & Penninx, 1994). In 2019, more than fifty percent of the population in Amsterdam had a migration background[6] (Gemeente Amsterdam, 2020a). In a multicultural and diverse city like this, it is not surprising to find ample manifestation of integration policies. Hoekstra (2014) notes that Amsterdam's policies are more diversity- and inclusion-oriented as compared to those of the national government. She also highlights the city's approach *de boel bij elkaar houden* (keep everything together) which emphasizes on concepts such as diversity (that encompasses ethnicity, age, gender as well as sexual orientation), connection, citizenship, and the undivided city.

However, the city is facing several criticisms in terms of datafication and the acceptance of the LGBTQI+ or migrant community. Jameson et al. (2019) point out that Amsterdam's 'smart city project' that seeks to bolster growth and digital connectivity in the city frames datafication as a "purely economic and technical phenomenon" (p. 1469) and ignores the social impacts of datafication or questions related to inequality and diversity. Taylor et al. (2016) also found that in datafied Amsterdam, "individuals became objectified and were seen as incidental to data flows – rather than living parts of the city's operations and dynamics, they easily became problems to be solved (public safety or public health risks), or groups to be influenced and controlled – users of the city, rather than its living infrastructure." (p. 5). Moreover, as previously stated, inequality and discrimination against the migrant and LGBTQI+ community are still present in Amsterdam (HRW, 2021; Selm, 2019) which is why the LGBTQI+ migrants living in Amsterdam risk facing intersectional discrimination and inequalities. Plus, in a datafied city such as Amsterdam where the social impacts of datafication are sometimes paid less attention to, data-led discrimination and inequality could intensify the intersectional discrimination they are already facing. These commitments and problems of the city discussed above make Amsterdam an excellent case study area for my study. The city's commitments towards datafication, diversity, inclusion, and equality and the problems such as privacy issues, intolerance present Amsterdam with novel challenges that need scientific exploration. Moreover, conducting my study in Amsterdam helps me provide useful insights to the discussion surrounding marginalized communities in a datafied Amsterdam.

In the following section, I lay out the specific objectives and the associated research questions of this research.

[6] The CBS defines a person with a migration background as "a person of whom at least one parent was born abroad". https://www.cbs.nl/en-gb/onze-diensten/methods/definitions/person-with-a-migration-background

1.3. Research Objectives and Research Questions

My research focuses on the city of Amsterdam and the adoption of data-driven technologies in its urban policies. Taking Amsterdam as my case study area, I aim to investigate the (geo)data practices and concerns of the self-identified LGBTQI+ migrants living in Amsterdam and how far-reaching the city's data policies are to account for those concerns. Below, I outline the research objectives and objective-specific research questions that this thesis addresses.

1. To understand the (geo)data practices of the LGBTQI+ migrant community in Amsterdam.
   a) How are the LGBTQI+ migrants in Amsterdam producing (geo)data in everyday life?
   b) Why are they producing this (geo)data?
   c) With whom are they sharing this (geo)data?
   d) Why are they sharing this (geo)data?

2. To understand the (geo)data concerns of the LGBTQI+ migrant community in Amsterdam.
   a) What potential concerns of datafication of Amsterdam do the city's LGBTQI+ migrants identify?
   b) How does, according to the city's LGBTQI+ migrants, Amsterdam's datafication affect their location privacy?
   c) How important is location privacy to Amsterdam's LGBTQI+ migrants?

3. To understand Amsterdam's LGBTQI+ migrant community's position on the city's datafication discourse from the city government perspective.
   a) How is the LGBTQI+ migrant community represented in the city's datafication policies?
   b) How does the city address the concerns identified by the LGBTQI+ migrants?

1.4. Thesis Structure

This thesis is divided into six chapters. Each chapter is further divided into subsections. In the Introduction chapter, I present the background to my study detailing the process and effects of urban datafication – informed by (geo)data – specifically in a migrant city. Following that, I justify my choice to focus particularly on the LGBTQI+ migrants for this study. Then, I outline the overarching research question and the specific research objectives. Following the Introduction chapter, the chapter on related works introduces readers to the current discussions and the state-of-the-art in the field of critical data studies, urban datafication, and migration and queer studies. In this chapter, I collate the interdisciplinary discussions and identify the gap to be addressed. The methodology chapter follows Related Work where I explain and justify my methods of data collection and analysis. This chapter ends with a note on the ethical considerations associated with the study. Next comes the chapter on results and discussions. In this chapter, I report and describe the findings and answer the research questions. In the discussion section of this chapter, I employ the framework I use from the current literature on data justice and data feminism to critically comment on the findings. The final chapter summarizes the study, gives the readers an overview of the limitations of the study, draws conclusions based on the findings, and points to future research directions to explore the nexus between (geo)data, urban datafication, migration, and data justice.

2. RELATED WORK

Various scholars have been contributing to discussions on the different aspects of geodata, urban datafication, migration, and the LGBTQI+ community. In this chapter, I connect these critical discussions and give an overview of the state-of-the-art of research in these fields. I group the literature into four categories - (1) geodata and its societal implications, (2) (geo)data in city datafication, (3) migration and (geo)datafication, (4) LGBTQI+ population, (geo)data, and urban spaces - contextualizing debates as they are relevant to the present study.

2.1. Geodata and its Societal Implications

Geodata is spatially tagged digital data that can reveal someone's spatial movements or location (Taylor et al., 2016). People in present-day society produce a large range of spatially tagged digital data (or geodata) in day-to-day activities. For example, using an ATM or a credit/debit card that registers the location of transactions, using mobile phones that can be tracked using the phone signal or the location-based apps on that phone (Höhnle, Michel, Glasze, & Uphues, 2013). The importance of geodata has been rising in our increasingly datafied society. The German Federal Ministry of Education and Research (BMBF[7], 2018) has referred to geodata as the "resource of the 21st century". Savinykh & Tsvetkov (2014) state that geodata is a unique and strategic information resource for national-level development. They further emphasize geodata's importance by describing its different characteristics:

> "The technological characteristic of geodata is that they are not obtained from direct measurements but rather result from the postprocessing of the information measured. The systemic characteristic is that, upon being formed, they become a system that coordinates and unites data of different types into a single complex. The informational characteristic is determined by the fact that geodata are a new information resource that makes it possible to solve tasks from diverse topical areas" (p. 365).

Geodata is now extensively used in many aspects of our society. Joh (2015) describes the police's use of geodata where they rely on surveillance and movement tracking for effective patrolling and preventing potential disruption of peace. But she also expresses concerns that data-led policing might lose its democratic values. Jansen (2018), while also highlighting the importance of location data in data-driven policing, is concerned over the very base on which such data-driven technologies have developed. She raises questions on such data practices and the unclear processes and criteria of feeding data of individuals to the police databases. This research is a call for more scientific exploration from social, political, and economic perspectives on how such data-led practices may adversely affect specific communities or neighborhoods.

There are other examples of geodata use in scenarios with large societal implications as well. For instance, disaster management (Barker & Macleod, 2019; Lingad, Karimi, & Yin, 2013; Wu & Cui, 2018), transport (Neisse, Baldini, Steri, & Mahieu, 2016), and healthcare (Christensen, Kjeldskov, & Rasmussen, 2007), among many others. While mainly discussing the importance and versatility of geodata, some of the

[7] Bundesministerium für Bildung und Forschung

literature have also focused on the possibility of potential privacy breaches and called for a stronger and effective implementation of informed consent.

In this section, I discussed the use and potential of geospatial technologies in various fields. In the following section, I turn to the literature on geodata and urban datafication which tends to be more critical about the adoption of geospatial technologies.

2.2. (Geo)data in urban datafication

Heeks et al. (2020; p.7) define urban datafication as the "growing velocity, volume and variety of data used in urban decision-making; and expanding presence for the city's 'data twin': a virtual but skewed simulacrum through which the city is increasingly planned and even experienced." These authors claim that the datafication of cities has aided in facilitating urban decision-making and has thrown light on the hitherto invisible issues or groups to include them in formal decision-making. (Geo)Data and geo-spatial technologies (such as location-based services, mobile phone data, CCTV surveillance) play a significant role in urban datafication leading to the emergence of geo-information infrastructures[8] in cities. However, the increasing use of (geo)data in a datafied city has seen critical voices express concerns about the associated risks that may affect the urban society, especially those at the risk of marginalization.

The introduction of urban datafication with the help of urban big data has transformed the way in which cities are governed, seen, and function. Such transformation has enabled the cities to incorporate real-time data into the fabric of urban governance and create a longitudinal, highly granular understanding of the urban system that can be managed and governed in real-time (Kitchin, 2016). Kitchin also notes that the rise of datafication is transforming urban development from being data-informed to being data-driven tightly. As such, it is tightly interlinking city-systems with infrastructures and helping urban dashboards and urban operating systems generate synoptic city intelligence.

In spite of the allure of urban datafication, there have been strong criticisms against the cities' failure to uphold their citizens' rights and freedom. Pasquale (2015) notes that big data tried to order our social and financial institutions in an asymmetrical manner using hidden algorithms that may exasperate existing inequalities. Taylor (2018) points out that the data collected by various actors in an urban environment, by getting aggregated with other data, may adversely influence certain communities. In cities like Amsterdam where datafication and surveillance are actively used, people may get more informed about the city and its (digital) services but, on the other hand, expose themselves to a new form of hypervisibility and get known to public and private data collectors at the individual level (Jameson et al., 2019). These authors have also identified that being 'extremely visible' leads to uncertainty and concerns among people about the use of their data. They also claim that the feeling of insecurity among people is increasing with the increasing dependence on technology and datafication as personal data is now more prone to getting leaked. Taylor et al. (2016) have also identified people having little control over how their personal data is shared. They also identified that in a datafied environment, the shared data may lose its context during its cross-sectoral flow between public and private data vendors. Such public-private data flow results in individuals being denied certain services. For example, business registration data flow from the Kamer van Koophandel (Chamber of Commerce) to other sectors resulted in sex workers in Amsterdam being denied housing.

[8] Taylor et al. (2016) define geo-information infrastructure as "the technologies and organisational channels used to collect, process and analyse information about us that relates to the way we use or occupy space" (p. 12)

Datafication and the adoption of geospatial technologies are not solely confined to urban development. Datafication, including geospatial technologies, is also widely used in understanding, monitoring, and controlling migration. In the next section, I outline the current discussions surrounding datafication, migration, and geo-technologies.

2.3. (Geo)data, datafication, and migration

Various research works are contributing to the ongoing discussion on datafication and its implications on migration. Datafication of migration can be viewed from two angles: (i) from the migrants' perspective, (ii) from the authorities'/governments' perspective.

The use of data-driven technology begins as soon as migration starts at the point of origin. Technologies such as smartphones and location-based services are used by migrants to facilitate their migration journeys (Gillespie et al., 2018). Their research highlights that for migrants, it is important to stay connected with friends, family, and other information providers – all activities are facilitated by smartphones. Zijlstra & Liempt (2017) identify that accessing online information through smartphones aids migrants' mobility, decision-making, and even financing their journey. Adding to this discussion, Dekker et al. (2018) state that social media applications on smartphones have become a popular source of information among migrants which influences their migration decision to a great extent. These authors notice that triangulating online information sources and cross-checking the available information with trustworthy social ties are some strategies developed by the migrants to avoid spreading rumors about migration-related information. Bustamante Duarte, Degbelo, & Kray (2018) show that smartphones and geospatial services are useful for migrants to navigate through daily life in the destination cities. Diminescu (2020) has termed them the connected migrants – those who have access to "at least one digitalized device which enables him/her to instantaneously switch between several lifestyles. This device gives migrants access and allows them to navigate in a connected digitalized environment." (p. 74). Monachesi (2020) has shown that in cities like Amsterdam, creative migrants use social media to interact with the city and contribute to urban development by highlighting "social activities and projects for the common good" (p. 2). These mobile devices, social media, or other similar technologies constitute the digital passage infrastructure that facilitates digital interaction among different actors such as migrants and/or refugees, governments, and multinational corporations (Latonero & Kift, 2018).

Data-driven technologies have also been adopted by governments to monitor and control migration. For example, Latonero & Kift (2018) mention the European regulations 'Eurodac' and 'Eurosur'. Eurodac focuses on the monitoring and controlling of migration through the collection of biometric data to ease identification and enact border control and monitoring while in the European territory. Eurosur entails the drones and satellite imagery surveillance of the Mediterranean Sea which is targeted at groups' movements disregarding their individual identities. These authors argue that while these practices might be sometimes "legitimate and even helpful" for these groups, it is relevant to explore all instances of how data is captured, collected, managed, the purposes and values embedded in this, as the current conditions present clear risks to migrants and asylum seeker groups.

Several other authors have also identified potential risks of datafication facing migrants. Gillespie et al. (2018) observe that by constantly interacting with the digital environment, migrants risk being trapped in a web of misinformation or revealing their geolocation data to third parties who may, then, use such data maliciously. Taylor's (2016) work shows that while using mobile phone data to understand migration or

human mobility facilitates authorities' responses to conflict and (forced) migration, it is likely to push certain groups to invisibility. These data are also prone to get misinterpreted and hence create a distorted picture of what is visible. Taylor (2018) describes how a *connected* migrant may become a victim of biased algorithms in a datafied urban society. These algorithms are meant to facilitate the governance of migrants and non-migrants alike and are not inherently biased. But when fed with data that is biased towards particular groups or communities, they produce results that are also biased. This is likely because the process of datafication has mainly been a technical one and requires reevaluation from a social justice perspective (Taylor, 2017). Leurs & Smets (2018) also support the argument and state that the current development in digital migration studies should focus more on the lived experiences of the (forced) migrants shaped by the intersections of culture, history, politics, and power rather than on the technological spectacle.

This section shows that datafication and geo-spatial technologies have changed the way we view migration and approach controlling and monitoring of the same. Despite the manifold use of (geo)data and datafication for migrants and authorities alike, there is a need for incorporating social justice principles in migration studies to better understand the potential exacerbation of sociopolitical discrimination rising out of datafication. In order to incorporate the ideas of social justice in the discussion surrounding migration and equality, it is important to move beyond techno-centric discourse and address the societal problems first before engaging with technology (Gregory, McMillan Cottom, & Daniels, 2017; Leurs & Smets, 2018). In the following section, I situate the LGBTQI+ population in the discussion surrounding urban spaces and datafication.

2.4. LGBTQI+ Population, Migration, and Urban Spaces

There have been discussions by various authors about urban datafication and the position of the LGBTQI+ community in the datafication debate. Considering the available space and time, only a few are pointed here. Goh (2018) has brought into light the issue of safe queer spaces in the urban politics discussion. Goh states that the LGBTQI+ population remains marginalized in contemporary urban spaces and faces multiple socio-spatial oppression rising from the imbricated identities entangled in race, gender, class, and sexuality. Goh further notes the need to assert LGBTQI+ rights in the heterosexually produced urban spaces where systemic oppression against the LGBTQ+ community exasperates marginalization. Shield (2019) also shows that migrants who identify as LGBTQI+ face intersectional challenges in Europe – both online and offline. Shield observes that even within the LGBTQI+ community, immigrants, specifically the non-Whites and non-Europeans, go through discrimination in digital spaces. The structure of our society is also largely influenced by heteronormativity marginalizing those who do not conform to the traditional gender norms or sexual expressions (Pertzel, 2020). Heteronormativity in society restricts the ways in which the LGBTQI+ community can express themselves and interact with others (Pertzel, 2020). Such marginalization and heterosexual constraints lead to the creation of what Foucault (1967) has termed as heterotopias. Heterotopias are spaces that exist outside the traditional heteronormative society and facilitate interaction and free expression of sexual orientation and/or gender identity among the LGBTQI+ community. Complementing the heterotopias, as noted by Gieseking (2016), the LGBTQI+ community creates urban territories. These territories offer a sense of belongingness and safety to the LGBTQI+ community in the otherwise heteronormative society. Using a queer-feminist lens, Gieseking comments on the importance of queer bodies in shaping urban territories for lesbians and queer women in a heteronormative world.

There is also an emergence of LGBTQI+ discussion in migration studies. Carrillo (2004) has coined the term "sexual migration" which is defined as "international relocation that is motivated, directly or indirectly, by the sexuality of those who migrate." (p. 58). Carrillo claims that focusing on sexual migration has the potential to enrich policy and strategy formulation for, for example, promoting awareness regarding sexual health issues among migrants as well as challenges the predominance of heteronormative practices in migration studies. Grewal & Kaplan (2001) also point out that LGBTQI+ migrants are frequently and wrongly placed within a western-centric narrative that attempts to show the LGBTQI+ migration as a journey from oppression to freedom. Such narratives discard the embodied experiences of LGBTQI+ migrants.

To counter the predominance of heteronormativity and western centrism in migration studies, scholars draw from, for example, queer and feminist theories. At the intersection of the LGBTQI+ community, migrants, technology, and datafication, it is important that we use critical theories to understand the emerging challenges and the resulting resistances. McKenna & Chughtai (2020) have found that the online digital world can function as a place of both safety and resistance against harassment and discrimination for marginalized communities such as the LGBTQI+ community. There is also a call for "queering" GIS and new spatial media to enforce social change, democratization, and justice in terms of sexualities and gender identities in contemporary discussions (Gieseking, 2018; Leszczynski & Elwood, 2015).

The discussion above groups the literature into four main categories – (i) geodata and its societal implications (ii) (geo)data in urban datafication (iii)(geo)data, datafication, and migration (iv) LGBTQI+ population, migration, and urban spaces. While space is limited, it is clear that the relevant literature goes beyond those pointed to here. At the same time, it also notable that no studies discuss how datafication may affect the LGBTQI+ migrant communities living in cities. It is important to establish the link between these four types of literature if we want to better understand how all these are related. As this thesis will show, there is ample scope for discussing how urban datafication discourses should and can talk about the LGBTQI+ migrant community.

3. METHODOLOGY

In this chapter, I introduce and explain Data Justice as the theoretical stance where my thesis is positioned. That is followed by a section on research design where I describe and justify the data collection and analysis methods chosen. The chapter ends with a note on the ethical considerations taken in this study.

3.1. Conceptualizing the Research Problem through the Lens of Data Justice

"Data is a source of power" (Heeks et al., 2020; p. 7). Data, specifically digital data produced by people's interaction with technological devices and services, determines how the State and private actors see and treat people including the political and practical implications that may have (Taylor, 2017). As such, data is a means to challenge existing socio-political injustices and unequal power structures (D'Ignazio & Klein, 2020). With the world witnessing a data revolution[9], certain population groups hitherto invisible are becoming (digitally) visible which is engendering new research and discussions (Taylor, 2017). However, the data revolution is also raising several issues and concerns. Those in power are claimed to determine how the benefits of data availability and datafication are distributed. Especially in urban areas, the fault-line of socio-political inequalities may get exacerbated due to the availability, or unavailability, of new datasets and faster data flow. As these issues impact deeply our current socio-political, economic, and cultural structures and dynamics, they require to be examined through the lens of data justice.

Data justice, as stated earlier, revolves around the idea of fairness – "fairness in the way people are made visible, represented and treated as a result of their production of digital data" (Taylor, 2017; p. 1). To incorporate data justice in the discussion surrounding datafication, it is important to have a better understanding of the instances of data injustices first. Taylor (2017) gives two examples through which she identifies and provides a comprehensive account of the data injustices happening in society. She presents the case of Aadhaar – the biometric population database of India. This program aimed at facilitating service provision (both governmental and non-governmental) specifically to the poor and marginalized which was expected, in turn, to expedite the distribution of the benefits of datafication among the marginalized. However, the Aadhaar program failed to acknowledge and contextualize the specific situation of the marginalized population it had initially planned to serve. As a result, the physically marginalized groups were made to experience virtual marginalization along with facing new challenges and barriers in the process of datafication. This example shows how the poor and marginalized often have to unfairly bear the negative implications of datafication.

Taylor's (2017) second example explores further the incidences of data injustices and brings the "geo" component into the discussion. By triangulating the data gathered from satellite images, social media, and local online reporting, the EU Space Agency proposed to monitor and predict the movements of migrant groups moving towards Europe's southern border. The aim of this project was to visualize and predict the flow of migrants attempting to reach Europe. These predictions, after being sold to migration authorities, were to be used in algorithmic sorting where the authorities would filter out the "undesirable" migrants

[9] According to the United Nations (2014), data revolution is when "new technologies are leading to an exponential increase in the volume and types of data available, creating unprecedented possibilities for informing and transforming society and protecting the environment. Governments, companies, researchers and citizen groups are in a ferment of experimentation, innovation and adaptation to the new world of data, a world in which data are bigger, faster and more detailed than ever before." (p. 2)

and prevent them, by introducing new measures, from reaching Europe. This was done with a view to controlling the number of asylum claims made in the European countries. The problem with this project is the categorization of migrants based on remotely observed behavior. The migrants' spatial information (their movement, their location) – some of which are produced voluntarily (e.g., geo-tagged social media data, sharing of location with friends and family, use of Google Maps) and some involuntarily (e.g., information recorded by satellites) – were used by the authorities to intercept the migrants' movements in order to meet the former's interests.

Both the examples show how datafication may induce discrimination and bring injustices to certain social groups. The two examples also give rise to questions regarding power imbalances and fairness on the visibility, representation, and treatment of people. Questions that rise from this that act as meta-questions crosscutting different aspects of my research are for example, being the data subjects in a datafied world, 1) how does the current underserved or marginalized population groups (such as migrants, LGBTQI+ people) experience the implications of datafication? and, 2) how can they control and decide on their inclusion and exclusion from the process of datafication? These two questions can be seen embedding the aspects of the RQs 2.(a), 2.(b), and 3.(a) of my research. Also, they have informed the definition of the codes of analysis for the interviews with LGBTQI+ migrants participating as well as from the documents. Some examples of codes, and the afterwards discussion, directly emerging from there are, for instance, data sovereignty, conditions to a digital society, and rights of the citizens.

Drawing on the emerging discussion on data and justice, Taylor (2017) proposes a data justice framework – a framework to guide our analyses and understanding of data-induced injustices. She terms it the "three pillars of data justice": visibility, engagement with technology, and non-discrimination (see Fig. 1). This is the framework I am going to use in this thesis to critically comment on my findings as well as to position them in the contemporary critical data studies literature. The framework is depicted in the following diagram.

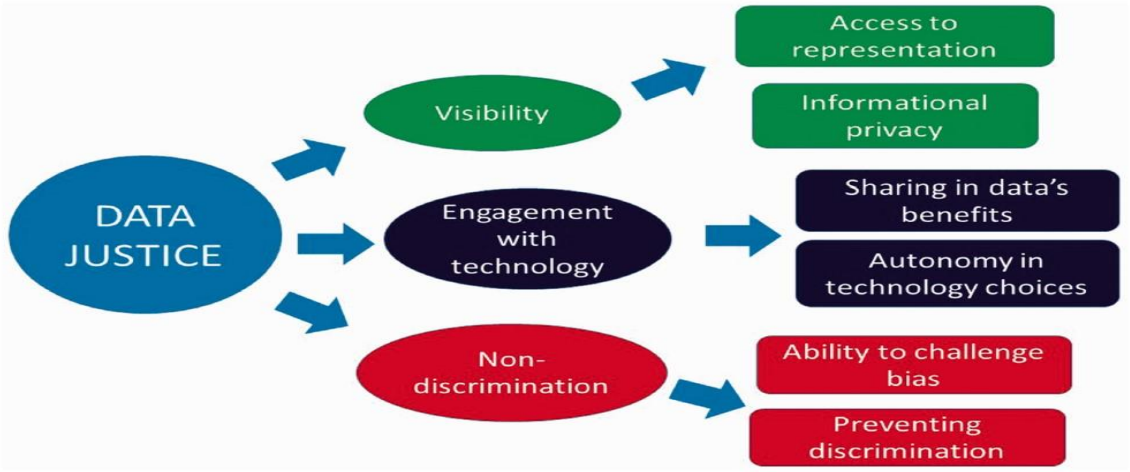


Figure 1: The Three Pillars of Justice: A Framework for Data Justice

(Source: Taylor, 2017)

This framework guides our thinking to consider data "at a level that goes beyond particular domains and applications, and instead to address data technologies primarily as they relate to human needs." (Taylor, 2017; p. 9) rather than telling us what we should or should not do with data. Each of the three pillars are as follows:

1. Visibility: This pillar deals with the interplay between representation and privacy. Visibility (of certain groups) in a datafied world needs proper representation while also respecting group privacy.
2. Engagement with technology: This pillar revolves around one's freedom to determine when and how to engage with what type of technologies and the ability to resist arbitrary inclusion of oneself in commercial databases. In other words, this pillar looks at one's power to control and determine their own visibility.
3. Non-discrimination: This pillar questions the power one possesses to identify and challenge those biases induced by datafication. It also explores one's freedom to resist potential discrimination through datafication.

This framework primarily focuses on the power dynamics in the datafied world. It proposes the questions of balancing and integrating the need for representation and visibility without interrupting the autonomy or integrity of certain groups. Using this framework, I guide my analysis in this thesis. I contextualize the three pillars to my own study and explore, in the Discussion section, visibility, engagement with technology, and non-discrimination from the perspective of the LGBTQI+ migrants living in Amsterdam.

3.2. Research design and research methods

This researcher adopts a qualitative approach to achieve the objectives of this study and to address the overarching research question which is: what are the (geo)data practices and concerns of the self-identified LGBTQI+ migrants living in Amsterdam and how far-reaching are the city's data policies to account for those concerns? A qualitative research approach explores and aims to understand social relations (de Sousa Santos, 2018). Moreover, a qualitative approach proves to be more fruitful if the research demands interaction between the researcher(s) and the participants to explore experiences and perspectives (Gray, 2014). In my research's context, a qualitative approach would, therefore, be well suited.

With the goal to commence my research with the community in question, the first step was to identify potential respondents. It has to be noted that the initial focus on the target research population has shifted from the LGBTQIA+ forcibly displaced population (FDP) to LGBTQIA+ migrants. LGBTQI+ FDP proved to be a hard-to-find research population as anticipated. I designed this project from start to allow for this shift from LGBTQI+ FDP to LGBTQI+ migrants. As a result, I had to adapt the participant recruitment process to match my new target population. The following sections detail the participant recruitment process.

3.2.1. Recruitment process

In this section, I explain the steps followed in the participant recruitment process. The section is divided into two subgroups.

3.2.1.1. Initial considerations

First, I had identified various public and private organizations in Amsterdam that were working for either the LGBTQIA+ community or refugees/asylum seekers in Amsterdam, or both. I contacted them with a view to getting in touch with possible participants through them (Appendix 1). However, this process did not result in participant recruitment from the target group. Therefore, I broadened my research focus from LGBTQI+ FDP to LGBTQI+ migrants.

3.2.1.2. Adaptations to recruitment process

After broadening my research focus, I reached out to several other organizations and social media and messaging groups (via Facebook, WhatsApp) (Appendix 2) of migrants but not particularly of refugees/asylum seekers. I created a registration form using Maptionnaire[10] where interested participants could register. I was successful in recruiting four participants in this phase. As I failed at recruiting more participants through this method, I decided to employ snowball sampling to recruit more participants. Finally, I recruited eight participants (Appendix 5). As the goal of my research is only to bring into light the LGBTQIA+ migrants' data concerns and set the ground for more critical work in the future, I deemed recruiting eight participants was enough. The participants were anonymized and their personal data were stored securely in an encrypted folder created using VeraCrypt (see DMP). The participants were assigned random numbers, e.g., PN3, PN11 to ensure anonymity.

The snowball sampling process that was followed had some limitations such as community bias or non-randomness which may influence my research results. However, the first respondents of the study came from different backgrounds and were gathered randomly. That helped me minimize community bias and non-randomness by getting me in touch with a somewhat diverse group of new respondents.

3.2.2. Data Collection

In this subsection, I describe and justify the data collection methods namely semi-structured interviews and policy document collection. I conducted pilot interviews with three other persons in advance who were not related to or aware of my research content to ensure that no leading questions were asked. Based on the pilot interviews, the interview script was modified, and I went on to using it with my study participants.

3.2.2.1. First Phase – semi-structured interviews

During this phase, I conducted a semi-structured online interview with each of the eight recruited participants. The interviews lasted for 45 minutes on an average. In the interviews, I asked them about their general data practices and concerns (not specifically geodata). I, then, extracted the geodata elements, based on the definition I used, from their responses during my analysis. I did so because they might not have sufficient knowledge about geodata and explaining the geodata concept to them might be time-consuming or might even confuse them. Therefore, I reckoned asking them about their general data practices instead of geodata was important.

I chose semi-structured interviews as the preferred method as my research requires to handle sensitive and complex issues of the LGBTQIA+ migrants' data practices and concerns. Semi-structured interviews, as (Barriball & While, 1994) state, "are well suited for the exploration of the perceptions and opinions of respondents regarding complex and sometimes sensitive issues and enable probing for more information and clarification of answers" (pp. 330).

[10] Maptionnaire was used as it is GDPR-compliant and approved by the ITC Faculty of the University of Twente

The interviews were online due to the current COVID-19 restrictions on traveling and social distancing. I used Skype or Duo (as preferred by the participants). Both offered end-to-end encryption on calls and text messages. All the participants were fully informed about the study purpose and potential risks through an information sheet and were encouraged to ask questions about the study before beginning the interview. They were, then, asked to sign a consent form allowing me to collect and use their data as described in the information sheet. The participants were also compensated with a gift voucher of their choice to compensate them for the time they had afforded me. Compensation was done using the funding available for MSc. students from the ITC Faculty.

The interviews were recorded with the participants' consent (either by audio-recording or noting down responses using pen and paper based on the participant's consent) and saved in a secure manner (see DMP in Annex.). Seven participants consented to being audio-recorded while one wanted the responses to be noted down using pen and paper. Following the interviews, I transcribed the audio-recorded interviews using AmberScript[11]. The transcripts of all eight responses were then checked for any possible errors and were corrected. After that, I prepared the interview transcripts for analysis which is discussed later in this document.

3.2.2.2. Second Phase – collecting policy documents

Parallel to conducting interviews, I also explored policy documents which helped me to understand the city's data authorities take on the concerns identified by my participants. I aimed for policy documents that detailed the city's data practices. I decided on collecting *Verslagen* (reports) published by the Personal Data Commission (CPA)[12] of the Municipality of Amsterdam (Gemeente Amsterdam, 2021). I collected the verslagen between the years 2017 and 2020 as the CPA website contained verslagen for only that time period at the time of data collection for this study. These verslagen are the official reports detailing the public meetings of the CPA on Amsterdam's data practices. "The CPA advises on the privacy policy of the municipality and its implementation. In addition, the committee advises the official organization on complex and/or politically sensitive issues concerning personal data. The CPA identifies what is going on within the municipality in the field of personal data and brings this to the attention of the municipal council and the management of the official organization." (Gemeente Amsterdam, 2021; p. 1). I selected these documents from the CPA because they contain a rich details of the city's data-related activities in a chronological order. These were helpful for me to follow Amsterdam's development through the lens of an agency in terms of datafication throughout these years. Analyzing merely the verslagen is not enough to get a full account of what Amsterdam is doing in its process of datafication. But for a small-scale research such as this, the verslagen provided enough data to address , on an exploratory form, the defined research questions. It should be noted that the verslagen are only available in Dutch. Since my level of proficiency in the Dutch language is not enough for document analysis, I translated them using DeepL and Microsoft Translation. This may create the issue of "lost in translation" and may mislead my understanding. In order to minimize this effect, I took help from a native Dutch speaker if sometimes some translation did not make sense.

3.2.3. Data Analysis

In this section, I detail the data analysis methods and justify them. I analyzed all data using thematic analysis (TA). TA is "a method for identifying, analyzing, and reporting patterns (themes) within data. It minimally organizes and describes your data set in (rich)detail" (Braun & Clarke, 2006; p. 6). I have

[11] The University approves AmberScript as a GDPR-compliant software that can be used to transcribe personal/sensitive interview data.

[12] Commissie Persoonsgegevens Amsterdam (https://www.amsterdam.nl/bestuur-organisatie/organisatie/overige/adviesraden/commissie-persoonsgegevens-amsterdam/)

chosen TA for my qualitative analysis as "a rigorous thematic analysis can produce trustworthy and insightful findings" (Nowell, Norris, White, & Moules, 2017; p. 2). Atlas.ti was used to perform the analysis.

I adopted a quasi-inductive approach as proposed (Perry & Jensen, 2001) to initiate the data analysis phase. This approach allows the researcher to develop themes before the coding process and makes the researcher aware of the possible dimensions to be studied in the analysis (Perry & Jensen, 2001). By adopting this approach, I have identified themes by consulting literature (Table 1). I primarily drew on highly cited critical data studies and data justice literature[13]. The following table states the themes and refers to the literature that they emerged from. These themes were used to guide my analysis of the interview data. My coding scheme was partly inspired by these themes. I started coding deductively at first based on these themes. Then, as I went deeper into the data, new concepts came up for which some new codes emerged. These new codes were not inspired by or connected to these literature-born themes.

Themes	Literature
Data and the rights over data	(Taylor, 2016b)
Consent, identifiability, and data sharing	(Taylor, 2016b)
Geo-surveillance and privacy in a data-driven city	(Kitchin, 2016)
Privacy, responsibility, and accountability in a datafied city	(Taylor et al., 2016) (Kitchin, 2016)
Security, efficiency, and data maximization	(Taylor et al., 2016)
Representation, visibility, and inequality of marginalized communities in urban datasphere	(D'Ignazio & Klein, 2020) (Heeks & Shekhar, 2020)

Table 1: Themes to proceed with the analysis

3.2.3.1. Interview Data

After identifying the six themes present across the selected critical data studies literature, I analyzed the interview responses to find (dis)similarities in the new themes emerging from my research participants' narratives. For the analysis, I coded the data using Atlas.Ti in three iterations (Appendix 3). After the coding, the codes were categorized into different groups. These categories were, then, grouped together to form new themes. These new themes covered the data practices and concerns of my research participants. The bew themes are (1) safety and convenience trump privacy, (2) awareness of datafication shapes perceived risk, (3) consent is subject to power dynamics in urban datafication. As shown in the result chapter, these new themes went beyond the immediate scope of the literature-born themes.

3.2.3.2. Data from the *verslagen*

To analyze the verslagen, I followed the same procedure as described in section 3.2.3.1. I coded the data in three iterations. Following that, each code was assigned to a theme. These themes, discussed in the result

[13] I did not include all the highly cited literature available in the field of data justice and critical data studies. I only included those that seemed most relevant to my thesis.

chapter, represented the city's data collection/processing practices and how it addresses the concerns identified by my research participants. These themes are (1) data transparency and privacy protection are important in datafication, (2) datafication to improve city life?

3.3. Ethical Considerations, Risks, and Contingencies

My research deals with a social group that faces intersectional marginalization. Therefore, it is important to make sure the research is conducted ethically. For my research, I followed the ethical guidelines from the ITC Ethics Committee and the Netherlands Code of Conduct (ITC, 2021; KNAW et al., 2018).. Below, I will discuss the underlying framework I followed to address the ethical dimensions involved in my research. (To see how I treated the personal and sensitive[14] data collected, refer to the DMP in the appendices).

I viewed the ethical issues from two dimensions in this research – procedural and situational ethics. Procedural ethics involve institutional requirements to "adequately deal with informed consent, confidentiality, rights to privacy, deception, and protecting human subjects from harm" (Ellis, 2007; p4). The second dimension, situational ethics, "deals with the unpredictable, often subtle, yet ethically important moments that come up in the field" (Ellis, 2007; p4).

For procedural ethics, I considered the following aspects:

Informed Consent: It is very important that the participants decide for themselves about their participation in the study. For this, they were informed about what this study entails, what the objectives were, the collection and use of their data, the potential results, and that they have the right to withdraw their participation anytime they want. I used the ITC Consent Form template (ITC, 2021) for this purpose.

Confidentiality and anonymity: I ensured and informed the participants that their data would be treated confidentially. Their responses and personal data would not be accessible by anyone else other than myself. Confidentiality was maintained by not publishing or sharing the participants' personal data and also by storing their data in a secure way which is described in the previous section. The potential risks of this research were also made clear to the participants through the information sheet. Due to the very small number of participants, it is a high risk for the participants to be re-identified especially by people that know them personally. In order to minimize the risk, I assigned each respondent with a code name so that their personal identifiers could not be linked to them in the future.

Rights to Privacy: All their data were collected and processed with the participant's prior consent. I did not ask for any private information from them that was not needed in my study. However, since I am particularly focusing on the LGBTQI+ migrants, I needed to have data on their sexual orientation or gender identity. The participants might not want to provide this information to anyone else which is why I had to treat the information very carefully. I expected I might also come across other sensitive data not necessary for my research such as health-related data or political/religious opinion. That is why I, before the interview, informed the participants that I would not publish or share any such information with anyone else; also, that I would securely store their personal identifier which could be linked to the individual participants and might reveal their private information and would destroy any other sensitive data not needed for my research.

[14] Sensitive data are those revealing for example political/religious opinion, health-related data, genetic/biometric data, or data on racial/ethnic origin (European Commission, 2019).

Rewarding the participants: It was another dilemma I faced about how to offer appropriate compensation to the participants for their investment of time and knowledge on the research. While it seems logical to compensate for their time with a reward, it can also be said that rewarding them may bias their participation or responses in the research. I discussed with my supervisors and decided to compensate my participants with €25 gift vouchers of their choice. The compensation was made as a form of encouragement and not as the main incentive to take part in my study.

And for situational ethics, I expected that I might come across sensitive personal information during the data collection phase, which is not necessary for my research. In such a case, I had planned to rely on the flexible nature of the semi-structured interview which would help keep the conversation going. Concerning sensitive information, I had accounted to treat it in a confidential manner and destroy the sensitive data as soon as possible.

4. RESULTS

In this chapter, I describe the findings from the analysis to answer the RQs. To recall, I am specifically interested in investigating the (geo)data practices and concerns of the self-identified LGBTQI+ migrants living in Amsterdam and if the city's data policies are far-reaching enough to account for those concerns. Below, I detail the themes that emerged from the data.

4.1. Results from the Interviews

My analysis of the interviews revealed three themes that were not literature-born but rather emerging from the interviews themselves. These themes were: (1) safety and convenience trump privacy, (2) awareness of datafication shapes perceived risk, (3) consent is subject to power dynamics in urban datafication. When commencing the interviews, I anticipated that some data concerns emphasized in the critical data studies literature would surface. These three themes, as I will show in the following section, indicate that the concerns from the literature are intermingled with the participants' ways of digital interaction with the city. It should be noted that the themes are not mutually exclusive but are complementary to each other.

During the interviews, the participants mentioned a range of devices/technologies with which they interact with the city or receive services digitally. For example, one participant (PN3) uses the OV-Chipkaart for traveling purposes which records her movements. Another participant (PN7) uses social media where he puts his information including location data to receive recommendations for personalized services. There were also references to surveillance and monitoring technologies such as CCTV cameras in the city and tracking of digital movements via, for example, website cookies. All the participants, when asked, mentioned that location data (such as addresses, movement) is something they consider as part of personal data. However, none of them classifies location data as sensitive and as I show below, has varying concerns regarding such data.

4.1.1. Safety and Convenience Trump Privacy

This theme revolves around the tendency of the participants to trade privacy for the safety and convenience that urban datafication provides. There were several points raised by the participants that helped form this theme. The first point is about the feeling of safety. About half of the participants feel that there is safety in surveillance and that people tend to behave in a more orderly way under surveillance. The feeling of safety is also strengthened by the high level of trust the participants have in the city authority/government. The majority of the participants believe that the government acts responsibly when it comes to handling (geo)data. Therefore, they are not concerned about the latter having access to their (geo)data. However, almost all the participants were skeptical about private data vendors who collect and process their (geo)data. One participant (PN12) was fine with even private data vendors having access to his information:

> "I posted on Facebook about gay rights and no, here, I wouldn't mind, I feel safe here being gay and that sharing that I'm gay actively online or just people when I'm talking and they ask about it, even companies, I wouldn't mind."

But the majority of the participants assumed that the city government and the GDPR[15] would protect any potential data mishandling by the private data vendors. The participants from outside Europe also mentioned the importance of consent in data collection in the European context which, they said, gives them the impression of their (geo)data being handled safely.

Another point raised during the interviews was the convenience of living in a datafied city. One participant (PN7) pointed out that life is much easier when everything is just one click away. He further added about government services,

> "So anything, any governmental institution, I mean, you just have simply have like one login. Yeah… And, basically, all your data is stored somewhere online, which I personally find really good because if I go to Germany, you have this old-fashioned bureaucracy where, erm, you need to go there, they need to see you and you need to register; and all that stuff and everything is happening online here, which for me works fine."

Another participant (PN11) stated,

> "Well, I do all the registrations online, I never go to the physical office here in Amsterdam and I find it's really comfortable and I don't mind."

When asked about the concerns over sharing sensitive information with the government online, the same participant said,

> "There are kind of sensitive data about my financial situation, which I was a bit concerned in the beginning to share. But at the end, I did it and it turns out I benefited from that. So, it was OK."

These findings on the trade-off between privacy for safety and convenience among the LGBTQI+ migrant community in Amsterdam are similar to what other scholars have been debating about. For example, a similar discussion about "Have nothing to hide" (Cofone, 2020; Solove, 2011) resonate with these themes. These scholars have argued that the balance between privacy and security is shifting towards security and that privacy is an abstract and vague concept for many while the concept of security is readily understood. That results in people believing that one must trade privacy for more security. My findings also highlighted that privacy is a fluid concept for my participants. What I mean by "fluid" is that the importance of privacy changes with the level of safety/security and convenience gained from datafication. While the "nothing to hide" argument revolves around the trade-off between privacy and security, my participants' responses highlighted that it is more the ease of living that comes with datafication that they want to trade privacy for. While they recognize that datafication may provide safety in society, the word "safety" was also used by them to refer to the assumed safety of their (geo)data in the hands of city authorities. This assumed safety reinforces the arguments made by several scholars (e.g., Kitchin, 2019; Leurs & Smets, 2018; Taylor, 2017) on making data policies more citizen-oriented and placing human needs first. Without such citizen-oriented policies, the city may risk losing the trust of its citizens. Nevertheless, the tendency to trade off privacy for safety and convenience is a dangerous one. Solove (2011) argues that while security is important, it should not be at the cost of privacy. Without ensuring the privacy of its (resident and non-resident) citizens, a city cannot guarantee proper safety and security.

[15] The majority of the participants were aware of the GDPR's existence. Some of them had a basic understanding of what the GDPR does (for one participant, it was "some kind of a data protection document") and a few others had a deeper understanding of the rules and regulations of the GDPR.

4.1.2. Awareness of Datafication Shapes Perceived Risk

I found that the level of personal awareness of how the city collects and uses data determines how one perceives the potential risks of datafication. The majority of the participants frequently mentioned that they have a low level of understanding of why their (geo)data is collected and what happens to their data once it is collected. All of them have a high perception of their awareness about who they are willingly sharing their data with. But the level of understanding of how else their data is being collected and used is limited. Datafication itself is a relatively new and complex process. Even though all the participants were conversant with the increasingly digital lifestyle in Amsterdam, it is no surprise that many were not aware of the potential of datafication. Similarly, those participants were unaware of the probable risks and the level of vulnerability that comes with datafication. Such a finding is in line with what Shaw (2017; p1) points out, "The increased interconnectivity of the world we are living in [has led to] a level of vulnerability that we don't truly understand." Because of the different levels of understanding, the perceived risks mentioned by the participants were diverse. For example, one participant (PN3) is worried that an unwanted third-party may get their hands on their email address which was used in the registration for an online service:

> "Now, if I was just bombarded with spam mail and advertisements, that would be worse[…]".

On the other hand, another participant (PN9) expressed concerns about the possibility of the surveillance footage being misused:

> "But just the fact that there are security cameras, that's not a problem. It's just the possibility of mishandled video material that is concerning. However, if we don't know exactly how they are processing this dump, the footage might go into the wrong hands and might be a tool for people to analyze my habits, find out where I get and possibly commit hate crimes."

There was also the feeling that sometimes, their (geo)data may be used to discriminate against them, racially or otherwise. The participants also identified no or very few risks associated with their location information or sexual orientation or gender identity. Indeed, a varying level of awareness of the type of data being collected was evident – but more of the comments were about actively generating data than of data as a byproduct of everyday activities (e.g., beyond the smartphone - sensors, commercial geo-fencing). However, all the participants mentioned that more awareness about how their data is collected and used would have given them a clearer picture of the associated risks. They also agreed that a higher level of awareness would give them the power to opt out of processes they deemed risky.

These results support the arguments made by Mejias & Couldry (2019) where they argue that awareness is the key to empowering individuals in a datafied society. Awareness of the data type collected and used by authorities facilitates resistance and challenges datafication. Even though it is unclear whether such resistances and challenges can avert the arbitrary growth of datafication (Mejias & Couldry, 2019), the responses of the participants of this research clearly highlight that a high level of awareness about datafication is desirable on their part. However, with a high level of awareness about the city's data practices may arise what Brayne (2014) has called "system avoidance". This means that with increasing awareness and knowledge about data collection and processing practices, individuals may avoid being captured through data by choosing channels and services that are unofficial and not monitored. This may, in turn, expose certain groups to new forms of violence and exploitation which are not regulated by legal systems (Latonero & Kift, 2018).

4.1.3. Consent is Subject to Power Dynamics in Urban Datafication

Another finding from the interview responses is that power dynamics play an important role in how the participants interact with the datafied city. It was mentioned several times during the interviews that a layman, nowadays, does not possess enough power to prevent the government or the companies from accessing any of their data.

> "I don't know, I don't think I have too much power to decide what I want to give away and what not" – PN16 said.

This can also be linked to the discussion on awareness above. This instance shows that not only is there a lack of awareness of the city's data practices, but also of the options for resisting the unwanted practices. The sense of power comes from a person's position in society. One participant (PN11) mentioned that as migrants, they do not want to risk being sent back home. That is why they consent to giving away their data even if they did not initially want to. Although there is a high level of trust in the government among the participants, I found that being data subjects with limited power over datafication processes leads to discomfort among the participants. Another important finding from the interviews is that most of the participants feel that the power dynamics between public and private authorities are changing. They think that private authorities are gaining more power and that the existing regulations sometimes become ineffective in preventing arbitrary data collection by these authorities.

While many scholars (e.g., D'Ignazio & Klein, 2020; Gieseking, 2018) have argued that the imbalanced power dynamics tend to mask certain communities and thus they require more visibility, I found that the imbalanced power structure takes away the resistance power from these communities which leads to hypervisibility. PN13 said:

> "We're always seen [within air quotes] in one way or the other. If someone wants to follow you, it is very easy for him to do so now. You are always there digitally."

PN7, while talking about the growing dependence on technology in today's society, stated:

> "It's scary if you think about it. I don't know who is probably watching me right now. But what can I do?"

Haggerty & Ericson (2000) has termed such incidences as the *disappearance of disappearance*. This finding aligns with what Jameson, Richter, & Taylor (2019) have found in their research. These authors state that hypervisibility is often followed by uncertainty and fear around cybersecurity. However, the important factor to note here is that the resulting hypervisibility is not challenging the need for more visibility. The imbalanced power structure is making the LGBTQI+ migrants visible in ways that are not used for understanding or representing their needs in the datafication discourse of the city (I elaborate on this in the next section).

4.2. Policy Document Analysis

In this section, I describe the results of the thematic analysis of the policy document. I detail what the city of Amsterdam is doing in its data-related policies. As noted in the methods section this analysis is limited

to a corpus of *verslagen*[16] of the Amsterdam Personal Data Protection Authority (CPA). The commitment of the city to the LGBTQI+ community might lead us to conclude that this specific community or other more marginalized groups are a central concern of data policies. However, there is a lack of an explicit discussion in the data policies about the people belonging to this community. The analysis resulted in two themes: (1) data transparency and privacy protection are important in datafication, (2) datafication to improve city life? I relate these results with the findings from the interviews.

4.2.1. Data Transparency and Privacy Protection are Important in Datafication

Amsterdam is a forerunner in actively utilizing (geo)data for development. (Geo)Data is becoming part of how the city is run with a citizen focus in mind. For example, the Responsible Sensing Lab and the Algorithm Register (AMS, 2020; Gemeente Amsterdam, 2020b). These initiatives have been or will be trying to improve the city services and solve urban challenges by collecting and analyzing data and aim to involve "public values" (such as privacy, transparency, autonomy, empowerment, and inclusiveness) in the process. These commitments are evident in the verslagen as well.

The CPA puts emphasis on being transparent about its (geo)data collection and processing practices. It recognizes that transparency creates the basis for trust. That is why it aims to inform the public about its (geo)data collection and processing practices. It wants to do so by, for example, establishing general information web pages, placing colored stickers with a reference to the website in municipal devices so that people are informed of the purpose of the data collected and used. The CPA also wants to distinguish between government devices and privately owned devices such as private CCTV cameras in public spaces. The CPA wants to make it mandatory to provide private cameras in public spaces with a code. This code will enable citizens to check for themselves the purposes of these privately owned cameras. In many cases, the CPA wants to make the use of private cameras in public spaces punishable unless the purpose is announced. Although the CPA places importance on being transparent, it also recognizes some potential risks of transparency. For example, the CPA feels that being too transparent about the data collection processes may help criminals avoid screening through data – leading to a case of system avoidance (Brayne, 2014) which I have already discussed in the previous section. That is why the CPA wants to find an optimal level of transparency by assessing what the right level of transparency could be. In order to do so, the CPA needs to involve in the discussion different actors of datafication as well as the citizens. While the city's commitments to transparency do address the issue of the lack of awareness about datafication among the participants of this research to some extent, it does not fully inform them about the possibilities of their (geo)data usage or the potential risks associated with the collection and processing of such (geo)data.

The CPA also highlights the importance of privacy protection in the process of datafication. Examples of datafication in the city include the digital management of bicycle parking garages to reduce nuisances from parked bicycles (Verslag July 6, 2017), the use of panoramic images as a means to manage public space (Verslag August 23, 2018), and the Responsible Sensing Lab that uses data for various purposes such as crowd control or parking management (Verslag November 5, 2020). The CPA aims to ensure privacy protection of its data subjects by measures like data anonymization and secure data storage which is evident by the following quote from Verslag August 23 (2018) about privacy during the use of panoramic images:

[16] These verslagen (or reports in English) are the official reports detailing the public meetings of the CPA on Amsterdam's data practices.

> "The images are only used by employees of the municipality, but because they are paid for with tax money, they also want to be made available to citizens. The images are now made in-house, with their own equipment, at a quarter of the original cost. Photos were not anonymized in the past, but it was decided that this was no longer possible because of the privacy of the people in the photos: all civil servants have access to the photos and eventually they will also go to the City Archives. The privacy aspect will become even more important if the photos become publicly accessible. Anonymizing, 'blurring' photos requires special software. DataPunt has software, but the target of making 85% of the faces unrecognizable was not achieved. There is a feedback function: when viewing a photo, a button appears on the screen. Using that button, people can indicate that they are recognizable in the picture. Then, without discussion, the photo is anonymized within three working days. This has happened twice in three years. About 2 million photos are taken each year. In the meantime, the search for better software for anonymization has continued. Rijkswaterstaat indicated that they have software that makes up to 98% of the faces unrecognizable. A contract was concluded with Rijkswaterstaat: The Department of Public Works will fit the software, after which the municipality will own it."

As opposed to the participants' willingness to trade privacy for safety and convenience, the CPA emphasizes protecting privacy through its policies. Verslag July 6 (2017) states:

> "Privacy must be ensured in design and in the working principles that are also enshrined in the urban policy in the field of personal data[…]".

They are of the view that safety and security are inseparable from privacy and that a strong feeling of privacy creates a feeling of safety and security. It follows that one of the aims of the CPA is to create stronger privacy protection measures in the process of datafication.

4.2.2. Datafication to Improve City Life?

The CPA, through datafication, is providing and managing different services such as health monitoring of Amsterdam residents through the GGD (Municipal Health Organization) (Verslag June 8, 2017) and the use of camera surveillance in tunnel and passageways to ensure public order on the road (Verslag March 9, 2017). Through strategies such as monitoring and surveillance by, for instance, tracking of smart devices, the CPA wants its service provision to be more effective and efficient. The CPA aims to achieve responsible datafication. Responsible datafication is when the CPA prioritizes addressing the citizens' concerns associated with datafication.

The CPA is aiming to make its interaction with its citizens more and more digital. In this process, they want to give the citizens more control over their data. For example, people can set up their privacy profiles in order to have more control over what data to share with the municipality. This strategy could perhaps address the issues presented by the research participants during the interviews where they stated they currently not being aware of all the provisions of data collection and processing in the city. They also mentioned feeling less powerful regarding having control over sharing their (geo)data. It is, however, not clear from the city policies how much control the citizens will have when it comes to (geo)data sharing with private authorities. The CPA largely depends on private companies when it comes to collecting and using (geo)data for datafication. The findings from the interview data showed that the participants have a low level of trust in the private companies when it comes to (geo)data collection and processing. The city authority's increasing dependence on private companies to realize its datafication plans may lead to accountability issues. This dependency was evident from all the verslagen where it is seen that the CPA is

collaborating with several companies for the datafication of the city. For example, the city authorities launched the Goochem chatbot as the cultural guide for people in the Amsterdam region and was specifically aimed to help young people make a choice from Amsterdam's large cultural offerings. For this, the city authorities had to rely on the company Axendo and the company itself had to collaborate with Facebook in order to create an effective chatbot:

> "The creators indicate that it is now only about the use of Facebook Messenger and that there is not yet a link with Facebook itself. One is, however, in the process with Facebook for a connection that Facebook would then have to first authorize. If that continues, the chatbot will only have access to public data. Goochem will itself not send push messages. The choice of Facebook as a platform was made because the target group mostly already uses that medium and has also already registered there as a user." (Verslag August 24, 2017)

As Naafs (2018) identified, the Dutch city governments are gradually losing control over such private companies' ways of data collection and processing. These instances raise public accountability questions particularly for their potential impact on marginalized urban population groups.

The CPA policies do not specifically mention the LGBTQI+ migrant community. These data policies do, however, include other social groups such as children, senior citizens, or even migrants (verslagen March 5, 2020; March 9, 2017). But these data policies leave out the LGBTQI+ migrant community from its purview even after its growing commitments towards the LGBTQI+ community (e.g., Gemeente Amsterdam, 2015, 2019). This is where the issue of visibility comes into play. By leaving out the LGBTQI+ migrant community from its datafication praxis, Amsterdam is reinforcing the long-standing claim of marginalized people being further marginalized in the process of datafication. This can be linked back to the concerns identified by the research participants during the interviews. For example, it was found during the interviews that the participants felt their ability to consent to or opt out of some data collection practices was influenced by the imbalanced power dynamics. Such imbalances might actually be widened when certain communities are left out of the city's data policy purview.

The city's plan to improve city life through datafication is indeed ambitious. However, without proper consideration of such issues, the life of whom will be improved is questionable. This reinforces the arguments made for inclusion (of different groups of people) through technology in contemporary urban societies (Wahba, 2019).

5. DISCUSSION AND CONCLUSION

This chapter presents the discussions on the findings of the study. I position the findings within the data justice framework. I, then, describe the limitations of the study. I conclude the chapter with a summary of the study, by looking back at the specific objectives, and presenting future research directions.

5.1. Discussion

In this study, I found that the privacy concerns among the respondents were influenced by several factors. The level of data awareness influences how the respondents perceive risks associated with datafication and the power structures between the authorities and the respondents shape their consent and resistances in the datafication process. The level of trust in the authorities also plays an important role in determining the concerns the participants identified. The city, while being focused on data transparency and privacy protection in general, misses the opportunity to be more inclusive of the LGBTQI+ migrant community in its data-related practices.

My findings can be understood better with the data justice framework proposed by Taylor (2017) which I describe in the theoretical framework section. Below I discuss what my results mean from the perspective of the three data justice pillars proposed by Taylor.

Visibility

This pillar deals with representation and privacy. I found that even though data is collected from the LGBTQI+ migrants for various purposes in the city, they are not discussed in the city data policies. As D'Ignazio & Klein (2020) notice, masking certain groups from the mainstream datafication discourse may strengthen the existing power differentials and historical inequalities. The lack of visibility in the mainstream discourse often takes away the power of those invisible groups to make their voices heard and claim their rightful place in the debate. As stated before, the data policies do have a migrant focus; the only thing it is missing is the LGBTQI+ representation. The city is, based on the development of specific policies such as the city's Pink Agenda, inclusive of its LGBTQI+ communities. But to achieve a just process of urban datafication, the city's data policies need to be more inclusive.

From the exploratory results of interviews conducted, it could be seen that in terms of visibility: LGBTQI+ migrant participants did not mention particular concerns on the sharing of data in Amsterdam (as a host city) concerning their gender identity or their sexual orientation. They also had no concerns in terms of geographical visibility. The sharing of (geo)data was not directly stated or addressed during the interviews as a matter of concern in their current data practices in Amsterdam. Plus, monitoring and surveillance as practices of geographical visibility were not explicitly mentioned. As stated earlier, the high level of trust in the city authorities helps shape this narrative. Potential implications of this narrative could be the exploitation of these individuals in terms of the commercialization of their (geo)data. While the CPA does put strong emphasis, for instance, on geographical aspects such as CCTVs attached to public spaces and how the recorded data would be stored and processed (e.g., Verslagen April 19, 2018; August 29, 2019), there, are no initiatives from the CPA in regard to the possible commercialization of the LGBTQI+ migrant community's (geo)data. This could be expected given the CPA did not have any specific mention of the LGBTQI+ community in their policies.

Engagement with technology

This pillar deals with people's engagement with technology and how its benefits are shared as well as how much autonomy people have in choosing their engagement with technology. This pillar advocates for freedom to opt out of certain technologies and resist being part of commercial databases. From this thesis results, I have found from the LGBTQI+ migrants in Amsterdam who participated in this research that they feel having very low autonomy in choosing how to engage with the data market. The freedom of controlling one's own engagement with data-driven technologies leads to controlling and managing one's own visibility. In Amsterdam, although the city government has plans and has implemented some actions to give people more ownership over their data, these have not yet been realized in their full potential. The city needs a more holistic approach to datafication to achieve this pillar of data justice.

Non-discrimination

Non-discrimination focuses on people's ability to challenge biases and preventing discrimination through datafication. It is important to note that to challenge biases and prevent data-related discrimination, one must also have proper representation through data (D'Ignazio & Klein, 2020). While the participants have enough interaction with the datafied city to produce a large amount of data, they do not interact as an LGBTQI+ migrant but as any other citizen. The city also does not have a particular focus on this specific group. On one hand, this, in turn, may result in this group's weak representation in the data debate which also will result in them having less power to stand up to the challenges and discrimination posed through the datafication of the society. On the other hand, not having such data might be a less risk factor for these population groups who are already marginalized or are at the risk of marginalization in the sense that they cannot be discriminated against based on their sexual orientation and gender identity and/or their migration background.

From the discussion above, it is seen that the interplay between visibility, invisibility, and hypervisibility affects the LGBTQI+ migrant community's position in the datafication discourse. This is also interlinked with their perceived power to assert their inclusion in and opt out from certain (geo)data technologies in the city's datafication development. In order to have a more inclusive datafied city, the city should include the narratives of the LGBTQI+ migrant community by acknowledging the emotions and embodiment to enrich the current debate surrounding datafication by contextualizing these experiences (D'Ignazio & Klein, 2020). Also, the absence of queer-specific and location data concerns among the participants should be taken seriously by the city as that may imply the city needs more efforts to educate the residents about the implications of datafication. The city should also involve its citizens and denizens in its planning for data acquisition and develop protection measures to safeguard its residents against commercial data exploitation. Even though the city emphasizes informed consent, the unbalanced power structures do not always allow the residents to resist such data collection and interferes with their liberty. There should be the implementation of rule of law and so that data (including (geo)data) can be used for the benefits of both parties. Then, such data can also be used to create an LGBTQI+ migrant narrative and will minimize their marginalization in the data-related discussion.

5.1.1. Limitations of the study

The main limitation of this study is the lack of diversity among the research participants in terms of gender identity and sexual orientation. Among the eight participants, only one identified as cisgender bisexual woman and another one as cisgender bisexual man. The rest of them identified as cisgender homosexual men. Also, the number of participants was very limited. Therefore, the data collected during

the interview is not a proper representation of the LGBTQI+ migrant community living in Amsterdam. While I managed to collect rich data from the eight participants, a more representative group would have provided me with diversified and more detailed data.

In terms of country of origin, half of the research participants were European migrants and the other half were from other parts of the world. Only one of them experienced forced migration. The rest of the participants identified as economic or highly skilled migrants. Based on the country of origin and the type of migration, the experiences migrants come across may vary greatly. In many instances, the highly skilled migrants especially from the Global North are portrayed as desirable while forced migrants are problematized and shown as undesirable in contemporary migration discussion (Leurs & Smets, 2018). Because of such cases and the demography of the participants I had for my research, the collective narrative I present in this thesis might be biased towards highly skilled European migrants. Future studies should try to have a more balanced representation of the different types of migrants in order to have a more representative story.

Another limitation is that I have based my policy analysis only on the reports provided by the CPA. They indeed provided me with a rich set of data since they discuss the city's official data policies and practices. However, including other policy documents such as the Pink Agenda, Nota Regenboogbeleid in the analysis would give a better picture of the city's commitment towards inclusive development. The study also did not explore in depth, due to the reasons of time and space, the role of different actors – public, private, or public-private partnership – in shaping the city's data practices. A more comprehensive account of the city's data practice would have been provided in this thesis if the role and responsibilities of the different actors were examined.

Lastly, I mainly cited the work of Western authors in this thesis. Indeed, my study is set up in the West and so the work currently cited are very relevant. Nonetheless, my reading of the data feminism literature made me realize the importance of citing a diverse range of work. In my future research work, I will try to keep this principle in mind.

5.2. Conclusion

In this study, I explored the LGBTQI+ migrant community's position in the datafication debate in Amsterdam. The overarching objective of the study was to find out the (geo)data concerns of the LGBTQI+ migrants living in Amsterdam and if the city's data policies are far-reaching enough to account for those concerns. By taking a qualitative approach, I asked what data concerns (including those related to location or location privacy) Amsterdam's LGBTQI+ migrants identify. I, then, investigated how the city addressed those concerns if they did at all. I found that the participants' concerns were shaped by the trade-off between privacy and perceived safety and convenience, level of awareness regarding the data practices of the city. The power dynamics also play an important role in the ways in which the participants interact with the datafied city. These findings were then linked to the findings from the policy document analysis. The results from the analysis of the policy documents showed that there is a growing emphasis on protecting privacy and increasing transparency in the city's data policies which can be seen as a step towards addressing the issue of data unawareness identified by the participants. The city policies also focused on improving city life through datafication but there arise the issues of consent, power, and visibility that were identified by the research participants. And this is where the city policies should pay more attention to – getting the most out of datafication while also empowering citizens and denizens by giving them autonomy to include and exclude themselves at their own will from any kind of data technologies.

My research had three specific objectives. I briefly present below the conclusions for each of them.

5.2.1. To understand the (geo)data practices of the LGBTQI+ migrant community in Amsterdam

This specific objective explored what kind of data technologies are used by the LGBTQI+ migrants in Amsterdam. A number of data technologies were mentioned by the research participants ranging from smart travel cards to surveillance technologies through CCTV, social media, and also website cookies. Knowingly or unknowingly, the participants produced (geo)data by using these technologies to interact with the city or other private actors. The participants seemed aware of whom they are sharing their (geo)data with. They mentioned the city authorities and a number of private actors that they know have access to their (geo)data. The participants shared their (geo)data with them in order to get services in return. Sometimes, the sharing of (geo)data was not voluntary but was done as the ability to resist was perceived to be low by the participants. I have explained and reflected on this in sub-section 4.1.3.

5.2.2. To understand the (geo)data concerns of the LGBTQI+ migrant community in Amsterdam

This objective explored the (geo)data concerns that the LGBTQI+ migrant community in Amsterdam linked with the datafication of the city. I found minimal privacy concerns among the participants. The participants also only had a vague idea about how datafication may affect their location privacy. Privacy seemed to be a fluid concept for the participants which depended on what they get in return in exchange for privacy. This finding comes down from the varying level of awareness that the participants had regarding the city's data practices – how data flows, is collected and processed. There was also a perceived lack of power which gave rise to the issues of visibility, invisibility, and hypervisibility. These findings are explained in the sub-sections 4.1.1 and 4.1.2.

5.2.3. To understand Amsterdam's LGBTQI+ migrant community's position on the city's datafication discourse from the city government perspective.

The third objective explored how far-reaching the city policies are while accounting for the concerns identified by the LGBTQI+ migrant community in Amsterdam. Even though the city of Amsterdam is a forerunner in terms of LGBTQI+ rights, the data policies of the city missed having a specific focus on this community. But issues such as data unawareness can still be addressed by the city's focus on transparency. However, to address the issues of visibility and representation, the city needs a more citizen-oriented inclusive approach. I explain these findings in subsections 4.2.1 and 4.2.2.

5.3. Future research

For future work, there is ample scope to explore the nexus between urban datafication, (geo)data, and LGBTQI+ migrants or migrants in general. My research incorporated a data justice angle to examine this nexus. By doing so, this research contributed to the urban data justice research agenda (Heeks et al., 2020) by exploring the notions of (in)visibility and (in)equality in the contemporary datafication discourse. Future work may examine this nexus by exploring the other issues of this urban data justice research agenda such as data uptake or data-enabled authoritarianism. Moreover, the majority of the work exploring datafication and marginalization is focused on the Global North (Lokanathan, 2017). Therefore, future research should focus more on the datafication experiences of the cities in the Global South.

Furthermore, future researchers can also take a stance from the angle of data feminism (D'Ignazio & Klein, 2020) to investigate the issues of power and fairness. Data feminism seeks to examine the power

dynamics. It challenges the imbalances in power structures. When we challenge the imbalances, we challenge the resulting injustices, and in turn, reinforce the need for justice rather than equity in the processes of the current datafication dynamics. Data feminism seeks to expand the narratives around big data from white-male and techno-heroic to a much larger and diverse group. The concept examines how unequal power structures play a crucial role in shaping the narratives around data and how intersectionality intensifies the structural oppression faced by some groups. It should be noted, however, that data justice and data feminism are not two mutually exclusive concepts. They inform and complement each other. While data justice remained the main theoretical point of view for this thesis, future research can contribute novel insights to the critical data studies literature by crosscutting these themes of data justice and data feminism.

LIST OF REFERENCES

AMS. (2020). Responsible Sensing Lab. Retrieved July 2, 2021, from Amsterdam Institute for Advanced Metropolitan Solutions website: https://www.ams-institute.org/urban-challenges/urban-data-intelligence/responsible-sensing-lab/

Armstrong, M. P., & Ruggles, A. J. (2005). Geographic information technologies and personal privacy. *Cartographica*, *40*(4), 63–73. https://doi.org/10.3138/RU65-81R3-0W75-8V21

Barker, J. L. P., & Macleod, C. J. A. (2019). Development of a national-scale real-time Twitter data mining pipeline for social geodata on the potential impacts of flooding on communities. *Environmental Modelling & Software*, *115*, 213–227. https://doi.org/10.1016/j.envsoft.2018.11.013

Barriball, K. L., & While, A. (1994). Collecting data using a semi-structured interview: a discussion paper. *Journal of Advanced Nursing*, *19*(3), 328–335. Retrieved from https://d1wqtxts1xzle7.cloudfront.net/34291860/Barriball___While_1994_Collecting_data_using_a_semi-structured_interview_JAN.pdf?1406324952=&response-content-disposition=inline%3B+filename%3DCollecting_data_using_a_semi_structured.pdf&Expires=1603026443&Si

Bernardino, S., & Santos, J. F. (2021). Integration Challenges of Immigrants in Smart Cities. *International Journal of Entrepreneurship and Governance in Cognitive Cities*, *1*(2), 39–56. https://doi.org/10.4018/ijegcc.2020070104

BMBF. (2018). Geodaten: Der Rohstoff des 21. Jahrhunderts. Retrieved April 27, 2021, from https://www.innovation-strukturwandel.de/de/geodaten-der-rohstoff-des-21-jahrhunderts-2315.html

Boyle, P., Halfacree, K., & Robinson, V. (2013). *Exploring Contemporary Migration*. New York: Routledge.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brayne, S. (2014). Surveillance and System Avoidance: Criminal Justice Contact and Institutional Attachment. *American Sociological Review*, *79*(3), 367–391. https://doi.org/10.1177/0003122414530398

Bridges, J. (2019). Why geotracking is a growing threat to online privacy. Retrieved May 15, 2021, from Reputation Defender website: https://www.reputationdefender.com/blog/privacy/why-geotracking-is-a-growing-threat-to-online-privacy

Bustamante Duarte, A. M., Degbelo, A., & Kray, C. (2018). Exploring Forced Migrants (Re)settlement & the Role of Digital Services. *Proceedings of 16th European Conference on ComputerSupported Cooperative Work - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies*. https://doi.org/10.18420/ecscw2018_7

Carrillo, H. (2004). Sexual Migration, Cross-Cultural Sexual Encounters, and Sexual Health. *Sexuality Research & Social Policy Journal of NSRC*, *1*(3), 58.

Carroll, A., & Itaborahy, L. P. (2015). *State-Sponsored Homophobia - A world survey of laws: Criminalisation, protection and recognition of same-sex love*.

Christensen, C. M., Kjeldskov, J., & Rasmussen, K. K. (2007). GeoHealth: A location-based service for nomadic home healthcare workers. *Australasian Computer-Human Interaction Conference, OZCHI'07*, (November), 273–281. https://doi.org/10.1145/1324892.1324951

Cofone, I. N. (2020). Nothing to hide, but something to lose. *University of Toronto Law Journal*, *70*(1), 64–90. https://doi.org/10.3138/utlj.2018-0118

D'Ignazio, C., & Klein, L. F. (2020). Data Feminism. In *MIT Press* (1st ed.). Cambridge, MA: MIT Press.

Dameri, R. P. (2014). Comparing Smart and Digital City: Initiatives and Strategies in Amsterdam and Genoa. Are They Digital and/or Smart? In R. P. Dameri & C. Rosenthal-Sabroux (Eds.), *Smart City: How to Create Public and Economic Value with High Technology in Urban Space* (pp. 45–88). https://doi.org/10.1007/978-3-319-06160-3_3

de Sousa Santos, B. (2018). *The End of the Cognitive Empire: The coming of age of epistemologies of the south*. Retrieved from https://www.dukeupress.edu/Assets/PubMaterials/978-1-4780-0015-0_601.pdf?utm_source=twitter&utm_medium=social&utm_content=b-SantosIntro_Jul18

Dekker, R., Engbersen, G., Klaver, J., & Vonk, H. (2018). Smart Refugees: How Syrian Asylum Migrants Use Social Media Information in Migration Decision-Making. *Social Media and Society*, *4*(1). https://doi.org/10.1177/2056305118764439

Diminescu, D. (2020). Researching the Connected Migrant. In K. Smets, K. Leurs, M. Georgiou, S. Witteborn, & R. Gajjala (Eds.), *The SAGE Handbook of Media and Migration* (1st ed., pp. 74–78). SAGE Publications Inc.

Dunbar, W. (2012). Equal marriage around the world. Retrieved May 20, 2021, from Independent website: https://www.independent.co.uk/voices/comment/equal-marriage-around-world-8153205.html

EDCi. (2016). European Digital City Index. Retrieved May 18, 2021, from https://digitalcityindex.eu/city/1

Ellis, C. (2007). Telling secrets, revealing lives: Relational ethics in research with intimate others. *Qualitative Inquiry*, *13*(1), 3–29. https://doi.org/10.1177/1077800406294947

European Commission. (2016). European Capital of Innovation (iCapital) 2016. Retrieved May 20, 2021, from https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/prizes/icapital/icapital2016_en

European Commission. (2019). What personal data is considered sensitive? | European Commission. Retrieved November 5, 2020, from https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en

Foucault, M. (1967). Of Other Spaces: Heterotopias. Retrieved July 2, 2021, from foucault.info website: https://foucault.info/documents/heterotopia/foucault.heteroTopia.en/

FRA. (2020). *A long way to go for LGBTI equality EU-LGBTI II*. https://doi.org/10.2811/7746

FRA - European Union Agency for Fundamental Rights. (2015). *Protection against discrimination on grounds of sexual orientation, gender identity and sex characteristics in the EU – Comparative legal analysis – Update 2015*. https://doi.org/10.2811/556190

French Brennan, S. (2017). So much for Dutch tolerance: life as an LGBT asylum seeker in the Netherlands. Retrieved October 13, 2020, from The Conversation website: https://theconversation.com/so-much-for-dutch-tolerance-life-as-an-lgbt-asylum-seeker-in-the-netherlands-80332

Galič, M. (2018). Living labs and big data in practice: Stratumseind 2.0: Discussion of a living lab in the Netherlands. In A. Završnik & L. Selinšek (Eds.), *Pravo v dobi velikega podatkovja*.

Gatehouse, C., Wood, M., Briggs, J., Pickles, J., & Lawson, S. (2018). Troubling vulnerability: Designing with LGBT young people's ambivalence towards hate crime reporting. *Conference on Human Factors in Computing Systems - Proceedings*, *2018-April*, 1–13. https://doi.org/10.1145/3173574.3173683

GDPR. (2016). General Data Protection Regulation. Retrieved September 19, 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Gemeente Amsterdam. (2015). *Pink Agenda*. Retrieved from https://www.ohchr.org/Documents/Issues/LocalGvt/States/Amsterdam_4.pdf

Gemeente Amsterdam. (2018). *Stedelijk kader verwerken persoonsgegevens door de gemeente Amsterdam*. Retrieved from https://assets.amsterdam.nl/publish/pages/871672/stedelijk_kader_verwerken_persoonsgegevens_def_september_2018.pdf

Gemeente Amsterdam. (2019a). Ambitions and Implementation plan: Participatory and digital - City of Amsterdam. Retrieved October 20, 2020, from https://www.amsterdam.nl/en/policy/ambitions/participatory/

Gemeente Amsterdam. (2019b). *Nota Regenboogbeleid 2019-2022*. Retrieved from https://www.amsterdam.nl/sociaaldomein/diversiteit/amsterdam-regenboogstad/

Gemeente Amsterdam. (2020a). Allemaal Amsterdammers. Retrieved May 22, 2021, from Data en informatie website: https://data.amsterdam.nl/artikelen/artikel/allemaal-amsterdammers/cf965b3f-7570-41bf-b628-557b16351082/

Gemeente Amsterdam. (2020b). Amsterdam Algoritmeregister. Retrieved July 2, 2021, from City of Amsterdam Algorithm Register Beta website: https://algoritmeregister.amsterdam.nl/en/more-information/

Gemeente Amsterdam. (2021). Commissie Persoonsgegevens Amsterdam. Retrieved March 2, 2021, from https://www.amsterdam.nl/bestuur-organisatie/organisatie/overige/adviesraden/commissie-persoonsgegevens-amsterdam/

Gieseking, J. J. (2016). Crossing over into neighbourhoods of the body: urban territories, borders and lesbian-queer bodies in New York City. *Area*, *48*(3), 262–270. https://doi.org/10.1111/area.12147

Gieseking, J. J. (2017). Size Matters to Lesbians, Too: Queer Feminist Interventions into the Scale of Big Data. *Professional Geographer*, *70*(1), 150–156. https://doi.org/10.1080/00330124.2017.1326084

Gieseking, J. J. (2018). Operating anew: Queering GIS with good enough software. *The Canadian Geographer / Le Géographe Canadien*, *62*(1), 55–66. https://doi.org/10.1111/CAG.12397

Gillespie, M., Osseiran, S., & Cheesman, M. (2018). Syrian Refugees and the Digital Passage to Europe: Smartphone Infrastructures and Affordances. *Social Media and Society*, *4*(1). https://doi.org/10.1177/2056305118764440

Goh, K. (2018). Safe Cities and Queer Spaces: The Urban Politics of Radical LGBT Activism. *Annals of the American Association of Geographers*, *108*(2), 463–477. https://doi.org/10.1080/24694452.2017.1392286

Gray, D. E. (2014). *Doing Research in the Real World* (3rd ed.). Retrieved from https://www.academia.edu/29567720/Doing_Research_in_the_Real_Worl_David_E_Gray

Gregory, K., McMillan Cottom, T., & Daniels, J. (2017). Introduction. In J. Daniels, K. Gregory, & T. Cottom McMillan (Eds.), *Digital Sociologies* (pp. xvii–1xxx). Bristol, UK: Policy Press.

Grewal, I., & Kaplan, C. (2001). Global identities: Theorizing transnational studies of sexuality. *GLQ A Journal of Lesbian and Gay Studies*, *7*(4), 663–679. https://doi.org/10.1215/10642684-7-4-663

Gutiérrez, A., Domènech, A., Zaragozí, B., & Miravet, D. (2020). Profiling tourists' use of public transport through smart travel card data. *Journal of Transport Geography*, *88*(March), 102820. https://doi.org/10.1016/j.jtrangeo.2020.102820

Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *British Journal of Sociology*, *51*(4), 605–622. https://doi.org/10.1080/00071310020015280

Heeks, R., Graham, M., Evans, J., & Taylor, L. (2020). Global South Case Study Collection Introduction. In *Digital Development Working Paper Series The Urban Data Justice Case Study Collection*. https://doi.org/10.2139/ssrn.3705563

Heeks, R., & Shekhar, S. (2019). Datafication, development and marginalised urban communities: an applied data justice framework. *Information, Communication &Society*, *22*(7), 992–1011. https://doi.org/10.1080/1369118X.2019.1599039

Heeks, R., & Shekhar, S. (2020). An Applied Data Justice Framework Analysing Datafication and Marginalised Communities in Cities of the Global South. In *Digital Development Working Paper Series: The Urban Data Justice Case Study Collection*. https://doi.org/10.2139/ssrn.3705563

Hoekstra, M. (2014). Super-diversity and urban policies in Amsterdam. In *Interethnic Coexistence in European Cities*. https://doi.org/10.1553/icec_diversity_amsterdam2014

Höhnle, S., Michel, B., Glasze, G., & Uphues, R. (2013). Digital geodata traces - New challenges for geographic education. *International Research in Geographical and Environmental Education*, *22*(2), 97–108. https://doi.org/10.1080/10382046.2013.778713

HRW. (2021). The Netherlands Needs to Stay Vigilant Against Homophobic Violence. Retrieved July 2, 2021, from Human Rights Watch website: https://www.hrw.org/news/2017/04/04/netherlands-needs-stay-vigilant-against-homophobic-violence

IAmsterdam. (2021). LGBT History. Retrieved May 21, 2021, from https://www.iamsterdam.com/en/see-and-do/whats-on/lgbt/history

ITC. (2021). ITC Ethics Committee. Retrieved April 18, 2021, from https://www.itc.nl/about-itc/organization/boards-councils/ethics-committee/

Jameson, S., Richter, C., & Taylor, L. (2019). People's strategies for perceived surveillance in Amsterdam Smart City. *Urban Geography*, *40*(10), 1467–1484. https://doi.org/10.1080/02723638.2019.1614369

Jansen, F. (2018). Data Driven Policing in the Context of Europe. Retrieved June 1, 2021, from Data Justice Project website: https://datajusticeproject.net/wp-content/uploads/sites/30/2019/05/Report-Data-Driven-Policing-EU.pdf

Joh, E. E. (2015). The New Surveillance Discretion : Automated Suspicion , Big Data , and Policing. Retrieved May 30, 2021, from Harvard Law & Policy Review website: https://www.law.berkeley.edu/wp-content/uploads/2015/04/Joh-.-The-New-Surveillance-Discretion.pdf

Jones, K. M. L. (2018). What is a data double. Retrieved May 15, 2021, from https://datadoubles.org/2018/05/01/what-is-a-data-double/#markerref-53-1

Kitchin, R. (2016). The ethics of smart cities and urban science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083). https://doi.org/10.1098/rsta.2016.0115

Kitchin, R. (2019). Toward a Genuinely Humanizing Smart Urbanism. In P. Cardullo, C. Di Feliciantonio, & R. Kitchin (Eds.), *The Right to the Smart City* (pp. 193–204). https://doi.org/10.1108/978-1-78769-139-120191014

KNAW, NFU, NWO, TO2-federatie, Vereniging Hogescholen, & VSNU. (2018). Netherlands Code of Conduct for Research Integrity. In *DANS*. https://doi.org/10.17026/dans-2cj-nvwu

Latonero, M., & Kift, P. (2018). On Digital Passages and Borders: Refugees and the New Infrastructure for Movement and Control. *Social Media and Society*, *4*(1). https://doi.org/10.1177/2056305118764432

Lerman, J. (2013). BIG DATA AND ITS EXCLUSIONS. *66 STAN. L. REV. ONLINE*, *55*, 55–63. Retrieved from http://www.nytimes.com/2013/02/05/opinion/brooks-the-philosophy-of-data.html.

Leszczynski, A., & Elwood, S. (2015). Feminist geographies of new spatial media. *The Canadian Geographer / Le Géographe Canadien*, *59*(1), 12–28. https://doi.org/10.1111/CAG.12093

Leurs, K., & Smets, K. (2018). Five Questions for Digital Migration Studies: Learning From Digital Connectivity and Forced Migration In(to) Europe. *Social Media and Society*, *4*(1). https://doi.org/10.1177/2056305118764425

Lingad, J., Karimi, S., & Yin, J. (2013). Location extraction from disaster-related microblogs. *WWW 2013 Companion - Proceedings of the 22nd International Conference on World Wide Web*, 1017–1020. https://doi.org/10.1145/2487788.2488108

Lokanathan, S. (2017). *Mapping big data for development and the global goals : final technical report*. Retrieved from https://idl-bnc-idrc.dspacedirect.org/handle/10625/56905

Lucassen, J., & Penninx, R. (1994). *Nieuwkomers, nakomelingen, Nederlanders: Immigranten in Nederland 1550-1993*. Amsterdam: Het Spinhuis.

Manalansan IV, M. F. (2006). Queer intersections: Sexuality and gender in migration studies. *International Migration Review*, *40*(1), 224–249. https://doi.org/10.1111/j.1747-7379.2006.00009.x

Martínez, A. G. (2019, February 26). No, Data Is Not the New Oil. Retrieved June 25, 2021, from Wired website: https://www.wired.com/story/no-data-is-not-the-new-oil/?utm_source=WIR_REG_GATE

McKenna, B., & Chughtai, H. (2020). Resistance and sexuality in virtual worlds: An LGBT perspective. *Computers in Human Behavior*, *105*(July 2019), 106199. https://doi.org/10.1016/j.chb.2019.106199

Mejias, U. A., & Couldry, N. (2019). Datafication. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1428

Moe-Pryce, M., Bellanova, R., & Bergersen, S. (2016). Get to Know Your Data Double! . Retrieved June 26, 2021, from Law, and Ethics, Security website: https://blogs.prio.org/2016/11/get-to-know-your-data-double/

Mole, R. C. M. (2021). Queer Migration and Asylum in Europe. In R. C. M. Mole (Ed.), *Fringe* (1st ed.). https://doi.org/10.14324/ 111.9781787355811

Monachesi, P. (2020). Shaping an alternative smart city discourse through Twitter: Amsterdam and the role of creative migrants. *Cities*, *100*, 102664. https://doi.org/https://doi.org/10.1016/j.cities.2020.102664

Naafs, S. (2018). "Living laboratories": the Dutch cities amassing data on oblivious residents Using a smartphone in Utrecht, where €80m has been invested in data-driven management. Retrieved May 2, 2021, from The Guardian website: https://www.theguardian.com/cities/2018/mar/01/smart-cities-data-privacy-eindhoven-utrecht

Naughton, J. (2021, May 29). Data isn't oil, whatever tech commentators tell you: it's people's lives. Retrieved June 25, 2021, from The Guardian website: https://www.theguardian.com/commentisfree/2021/may/29/data-oil-metaphor-tech-companies-surveillance-capitalism

Neisse, R., Baldini, G., Steri, G., & Mahieu, V. (2016). Informed consent in Internet of Things: The case study of cooperative intelligent transport systems. *23rd International Conference on Telecommunications (ICT)*, 1–5. https://doi.org/10.1109/ICT.2016.7500480

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *Nternational Journal of Qualitative Methods*, *16*(1), 1–13. https://doi.org/10.1177/1609406917733847

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, MA: Harvard University Press.

Perry, C., & Jensen, O. (2001). Approaches to Combining Induction and Deduction In One Research Study. *Social Research*, 1–8. Retrieved from https://www.researchgate.net/publication/255654388_Approaches_to_Combining_Induction_and_Deduction_In_One_Research_Study

Pertzel, F. (2020). *The Signs on The Walls: Gender and Sexuality in Public Space*.

Redden, J., Brand, J., & Terzieva, V. (2020). Exploring social justice in an age of datafication. Retrieved May 5, 2021, from Data Justice Lab website: https://datajusticelab.org/data-harm-record/

Roshchupkin, G. (2020). LGBT Migration in the EECA Region. Retrieved April 14, 2021, from Aids Action Europe website: https://www.aidsactioneurope.org/en/publication/lgbt-migration-eeca-region

Savinykh, V. P., & Tsvetkov, V. Y. (2014). Geodata as a systemic information resource. *Herald of the Russian Academy of Sciences*, *84*(5), 365–368. https://doi.org/10.1134/S1019331614050049

Selm, J. van. (2019). Migration in the Netherlands: Rhetoric and Perceived Reality Challenge Dutch Tolerance. Retrieved July 2, 2021, from Migration Information Source website: https://www.migrationpolicy.org/article/migration-netherlands-rhetoric-and-perceived-reality-challenge-dutch-tolerance

Shaw, J. (2017). The Watchers: Assaults on privacy in America. Retrieved June 11, 2021, from Harvard Magazine website: https://www.harvardmagazine.com/2017/01/the-watchers

Shield, A. D. (2019). Immigrants on Grindr. In *Immigrants on Grindr* (1st ed.). https://doi.org/10.1007/978-3-030-30394-5

Solove, D. J. (2011). *Nothing to Hide: The False Tradeoff between Privacy and Security Nothing to Hide: The False Tradeoff between Privacy and Security (Introduction) (Introduction)*. Retrieved from http://ssrn.com/abstract=1827982

Szulc, L. (2020). Queer Migrants and Digital Culture. In K. Smets, K. Leurs, M. Georgiou, S. Witteborn, & R. Gajjala (Eds.), *The SAGE Handbook of Media and Migration* (1st ed., pp. 220–232). SAGE Publications Inc.

Taylor, L. (2016a). 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. (2016b). The ethics of big data as a public good: Which public? Whose good? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 1–13. https://doi.org/10.1098/rsta.2016.0126

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data and Society*, *4*(2), 1–14. https://doi.org/10.1177/2053951717736335

Taylor, L. (2018). On the presumption of innocence in data-driven government. Are we asking the right question? In E. Bayamlioglu, I. Baraliuc, L. Janssens, & M. Hilderbrandt (Eds.), *Being Profiled: Cogitas Ergo Sum, 10 years of "profiling the European citizen"* (pp. 72–78). Amsterdam: Amsterdam University Press.

Taylor, L., Floridi, L., & Van Der Sloot, B. (2017). Introduction: a new perspective on privacy. In L. Taylor, L. Floridi, & B. Van Der Sloot (Eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer.

Taylor, L., Richter, C., Jameson, S., & Perez del Pulgar, C. (2016). Customers, Users or Citizens? Inclusion, Spatial Data and Governance in the Smart City. In *Map4Society Final Project Report*. https://doi.org/10.2139/ssrn.2792565

The Economist. (2017). The world's most valuable resource is no longer oil, but data. Retrieved April 5, 2021, from https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

Traunmueller, M. W., Johnson, N., Malik, A., & Kontokosta, C. E. (2018). Digital footprints: Using WiFi probe and locational data to analyze human mobility trajectories in cities. *Computers, Environment and Urban Systems*, *72*(July), 4–12. https://doi.org/10.1016/j.compenvurbsys.2018.07.006

United Nations. (2014). A World that Counts: Mobilising the Data Revolution for Sustainable Development. Retrieved July 4, 2021, from https://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf

Wadhwa, T. (2015). Smart Cities: Toward the Surveillance Society? In D. Araya (Ed.), *Smart Cities as Democratic Ecologies* (pp. 125–141). https://doi.org/10.1057/9781137377203_9

Wahba, S. (2019). Smarter cities for an inclusive, resilient future. Retrieved July 4, 2021, from World Bank Blogs website: https://blogs.worldbank.org/sustainablecities/smarter-cities-inclusive-resilient-future

Witanto, J. N., Lim, H., & Atiquzzaman, M. (2018). Smart government framework with geo-crowdsourcing and social media analysis. *Future Generation Computer Systems*, *89*, 1–9. https://doi.org/10.1016/j.future.2018.06.019

Wu, D., & Cui, Y. (2018). Disaster early warning and damage assessment analysis using social media data and geo-location information. *Decision Support Systems*, *111*, 48–59. https://doi.org/10.1016/j.dss.2018.04.005

Zijlstra, J., & Liempt, I. Van. (2017). Smart(phone) travelling: understanding the use and impact of mobile technology on irregular migration journeys. *International Journal of Migration and Border Studies*, *3*(2/3), 174. https://doi.org/10.1504/ijmbs.2017.083245

APPENDICES

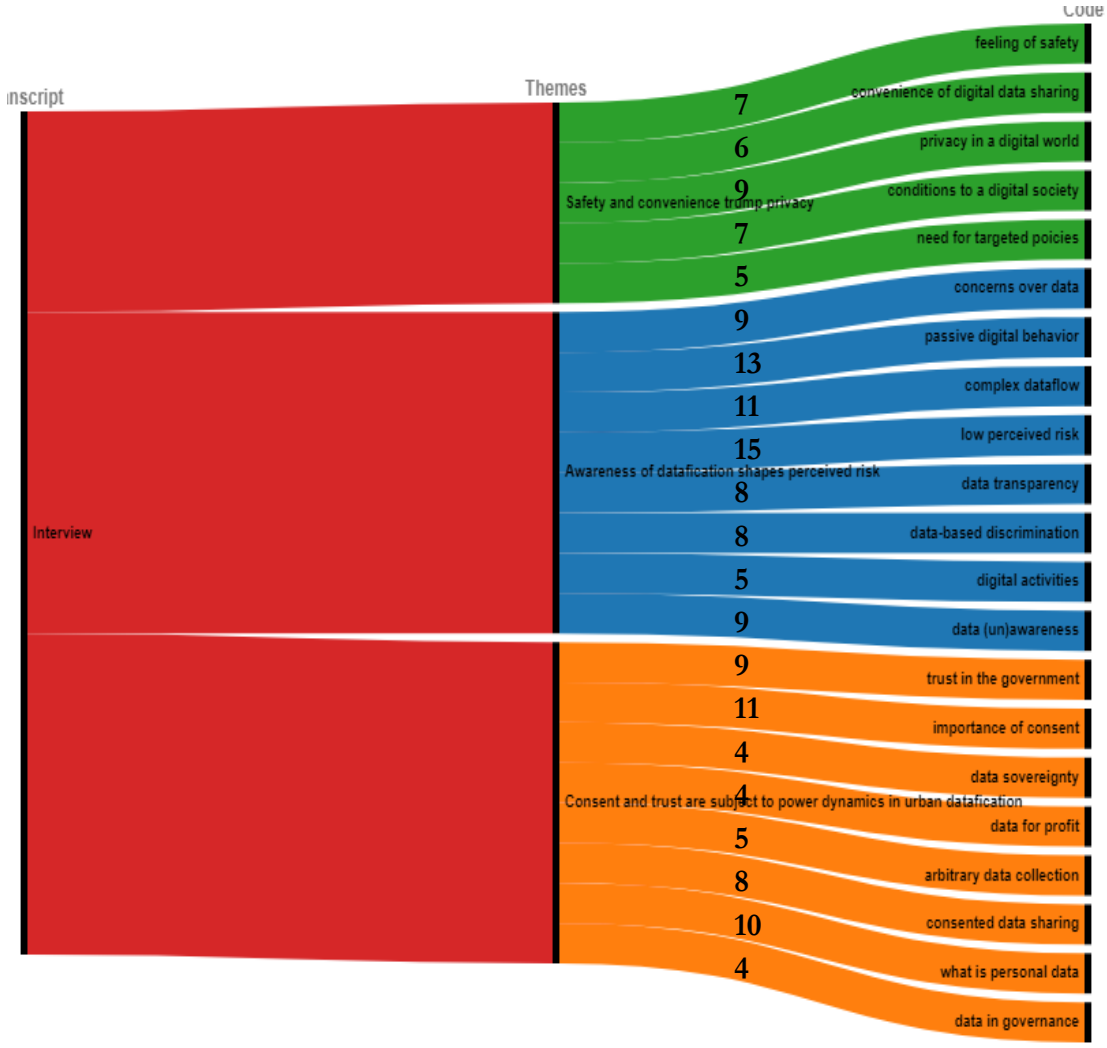
Appendix 1: Organizations contacted in the first phase of participant recruitment.

Name of the organization	Who are they?
LGBT Asylum Support	A Dutch NGO that links LGBT refugees to other Dutch organizations with a view to helping asylum seekers.
COC Amsterdam	The World's oldest LGBTQIA+ interest association and work for the "decriminalization of sexual orientation and gender identity and for equal rights, emancipation and social acceptance of LGBT's in Amsterdam, the Netherlands and all over the world".
Vluchtelingenwerk	An organization that guide asylum seekers through the asylum procedure, help them with housing, social integration, etc.
SEHAQ Queer Refugees Group	"…a refugee-led collective that raises awareness, hosts community events, and creates safer spaces for lesbian, gay, bisexual, trans* and queer (LGBTQ) asylum seekers, refugees and undocumented people in the Netherlands."
Connecting Hands (University of Twente)	An organization at the University of Twente that helps refugees and asylum seekers "fit into their new life, integrate into their new social environment and to accomplish their goals".
IND (Immigratie- en Naturalisatiedienst)	The official immigration and naturalization service of the Netherlands.
COA (Centraal Orgaan opvang Asielzoekers)	An organization guiding the asylum seekers in their asylum process and provides them with safe housing and other resources.
Refugee Academy, Vrij University	An institute at the Vrij Universiteit Amsterdam that facilitates and promotes research to address the refuse crisis.

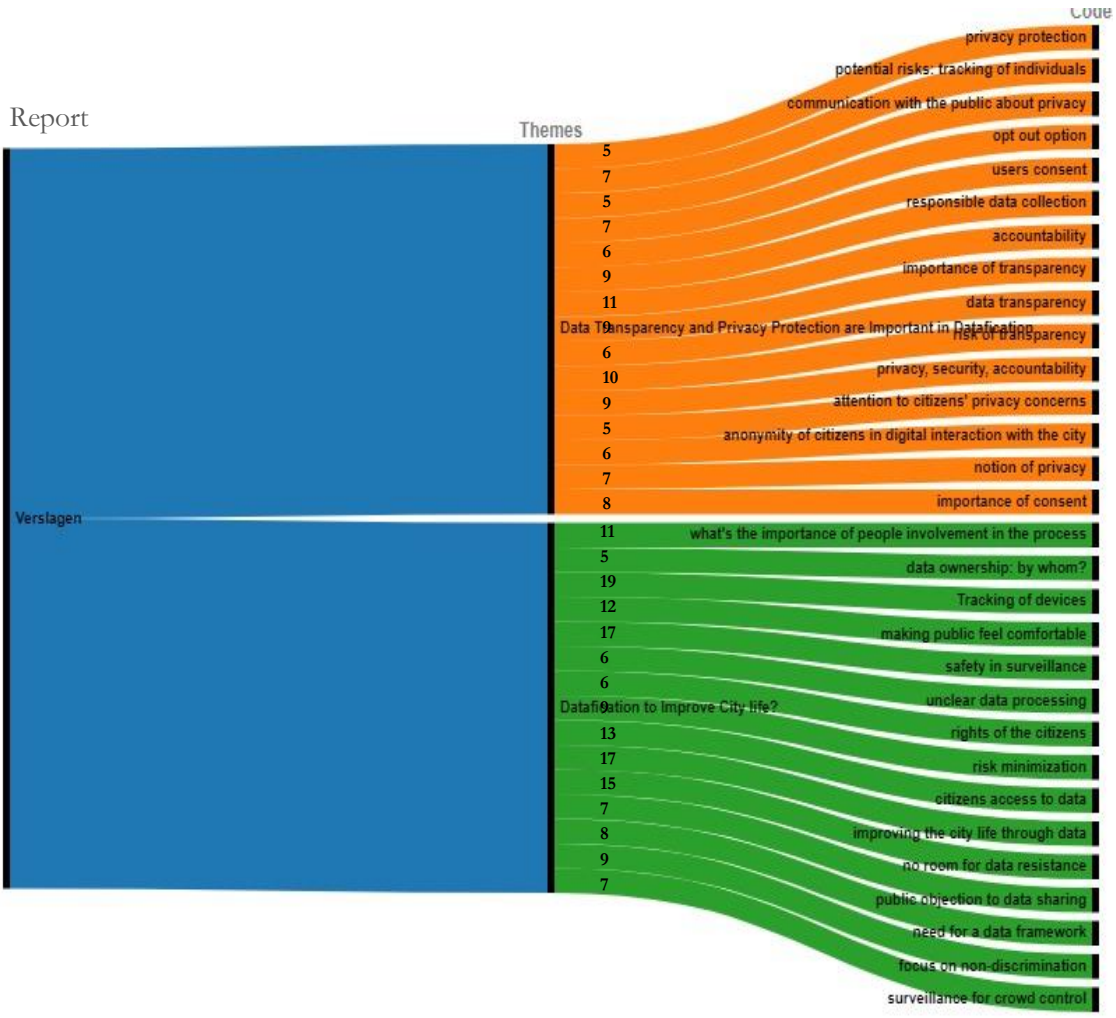
Appendix 2: Organizations contacted to reach out to LGBTQIA+ migrants living in Amsterdam

Name of the organizations
IHLIA
ASV Gay
Gay Expats in Amsterdam
The Amsterdam Expats Meetup Group
De Regenboog Groep
Wij Zijn Hier
Amsterdam City Rights
WhatsApp groups

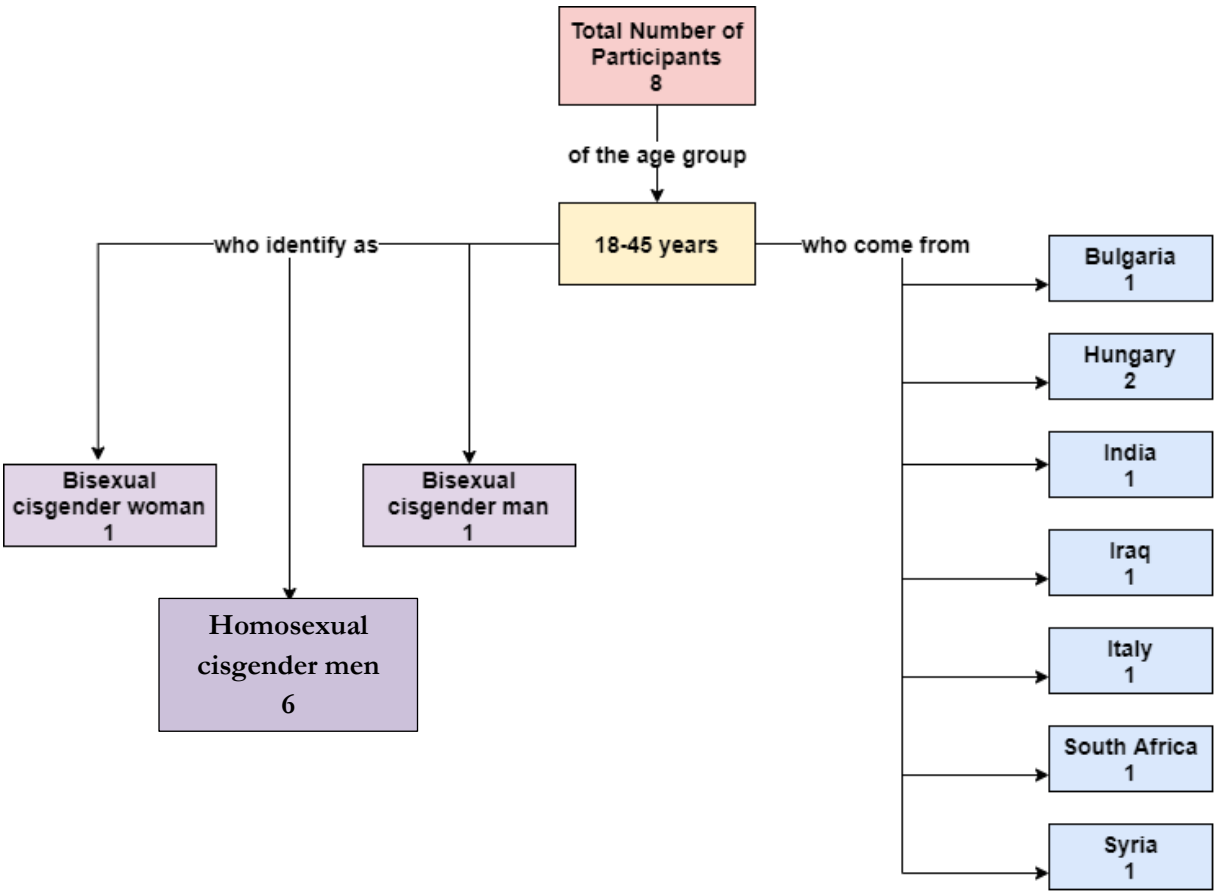
Appendix 3: Codes used in the interview data



Appendix 4: Codes used in Verslagen data



Appendix 5: Research participants details



Appendix 6: Interview questions

Interview Questions

Warming-up section

- Hi! Thank you so much for participating in this study. How are you?
- How is this coronavirus situation going for you?
- How would you say it has affected your daily routine?
- Would you say it has also affected the way you use technology? (like your phone, computer, travel cards like the OV-Chipkaart, digital payment, etc.). If so, how?

Data practices section

- So now, can you walk me through a typical day of your life (*particularly focusing on the use of these devices/technology*)?
- And do you get asked or do you need to share your personal data during such a day? Can you explain a little bit further maybe with examples?
- And how else do you think someone may get your personal data?
- While we are on it, can you maybe describe your idea about what you consider as personal data?
- Do you think the way you use such devices/technology is influenced by your sexual orientation/gender identity and your cultural background?
- Do you also feel like your ways could be different from someone who does not belong to your community? Why?

Topics to be covered in this section:

How are they using technology?
What data are they producing?
Has the "geo" element come up in the discussion?

Data concern section

- What are your thoughts about you having to share your personal data?
- Can you think of any examples when you feel at any point during your daily use of such technology that someone may have access to your personal data like your sexual orientation or gender identity?
- Can you walk me through your ideas about how the city government may collect your personal data?
- Do you think such data collection would affect your life in any way as an LGBTQIA+ migrant living in Amsterdam? Could you explain further?

Topics to be covered in this section:

Are they familiar with the city's data collection methods?
What particular concerns do they have?
Are they aware of any potential risks that may be related to their sexual orientation or gender identity and/or their migration background?

Thank you very much for your time. Do you have any questions? You can always contact me later if you have any questions, or to know about the outcome of my research. I wish you a good day. Bye!

Appendix 7: Consent form

Consent Form for participants taking part in the master thesis project by Udipta Boro from ITC, University of Twente.

YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes	Yes	No
Taking part in the study		
I have read and understood the study information dated _____ [DD/MM/YYYY format], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	○	○
I consent voluntarily to be a participant in this study, and I understand that A) I can refuse to answer questions. B) I can withdraw from the study at any time, without having to give a reason. C) I can ask for my part of or all the information I provided to be withdrawn from the study after the interview without needing to give a reason.	○	○
I understand that taking part in the study involves sharing personal information such as names, email address, my sexual orientation or gender identity with the researcher and that my responses may also be recorded (audio or in written form by taking notes) with my permission. If recorded, the recording will be transcribed as text and the original audio recording will be deleted and anonymized after the transcription.	○	○
Risks associated with participating in the study		
I understand that taking part in the study involves the following risks: learning about third parties' personal data collection methods which may cause emotional distress.	○	○
Use of the information in the study		
I understand that information I provide as part of the interview, will be analyzed and used in an anonymized form for the preparation of Udipta Boro's Master thesis and possible related publications.	○	○
I understand that personal information collected about me that can identify me, such as my name, email id, address will not be shared beyond the researcher.	○	○
I agree that the information I give during the interview can be quoted in research outputs without using my real name or other personal identifiers.	○	○
Consent to be recorded (please consent to one of the options)		
I agree to be audio recorded.	○	
I agree to my responses being noted down with pen and paper.	○	

UNIVERSITY OF TWENTE.

Signatures

Name of participant [optional]	Signature	Date

For participants unable to sign their name, mark the box instead of a signature.

I have witnessed the accurate reading of the consent form with the potential participant and the individual has had the opportunity to ask questions. I confirm that the individual has given consent freely.

Name of witness [optional]	Signature	Date

I have accurately informed the potential participant about my research and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

Researcher's name [printed]	Signature	Date

Study contact details for further information:
Udipta Boro
u.boro@student.utwente.nl
+31 6 84515817

UNIVERSITY OF TWENTE.

Appendix 8: Data Management Plan (DMP)

Sources of data (more than 1 source is possible)	Interviews Policy document review
If various data sources are combined, could you identify the individual datasets? If yes, mention them.	The data will be stored separately. So, the individual datasets can be identified when needed.
Which organization (s) owns the data you are going to use? (more than 1 organization is possible) Give the names and the website of the organizations	The primary data I am collecting through interviews do not belong to any organization. But the policy documents belong to the Personal Data Commission Amsterdam.
In what ways are you permitted to use the data (e.g., only for your own analysis, disseminate original data, disseminate derived data, etc..) ?	I asked for prior consent from the participants to use the primary data for my own analysis. These data were kept with me only during the length of the MSc.Thesis. The data collected from the policy documents review can also be used for my own analysis.
How will you organize your data during the project?	Some of data I collect are very sensitive, e.g., the personal data of the participants and their interview responses. I anonymized the data immediately after collection by assigning random numbers (e.g., PN3, PN11) to the participants. During the thesis writing period, these data were kept in an encrypted folder created using VeraCrypt. Any mention of name, places, etc. during the interviews were also anonymized and replaced with N1, N2, etc. for names and P1, P2, etc. for places. After the thesis writing phase, these data needed to be archived in the University data repository. With the persons associated with the repository, I made sure that these data are stored very securely and that no one can consult those data without my authorization first. For the policy documents, as they were open data, kept them in a separate folder on my personal laptop and archived them following the standard procedures.