TOWARDS A DATA-DRIVEN CITY: AN ANALYSIS OF CITIES' DATA STRATEGIES

Yusuf Bozkurt[1,2], Simon Flaisch[2], Dieter Hertweck[2] and Alexander Rossmann[2]
[1]*School of Computing, Engineering & Physical Sciences, University of the West of Scotland*
Paisley, PA1 2BE, Scotland, United Kingdom
[2]*Department of Computer Science, Reutlingen University*
72762 Reutlingen, Germany

ABSTRACT

Smart cities are considered data factories that generate an enormous amount of data from various sources. In fact data is the backbone of any smart services. Therefore, the strategic beneficial handling of this digital capital is crucial for cities. Some smart city pioneers have already written down their approach to data in the form of data strategies, but what should a city's data strategy include, and how can the goals and measures defined in the strategies be operationalized? This paper addresses these questions by looking closely at the data strategies of cities in Germany and the top three countries in the EU Digital Economy and Society Index. The in-depth analysis of 8 city data strategies has yielded 11 dimensions that cities should consider in their data strategy. These are (1) relevance of data, (2) principles, (3) methods, (4) data sharing, (5) technology, (6) data culture, (7) data ethics, (8) organizational structure, (9) data security and privacy, (10) collaborations, (11) data literacy. In addition, data governance is a concept to put these 11 strategic dimensions into practice through standardization measures, training programs, and defining roles and responsibilities by developing a data catalog.

KEYWORDS

Smart Cities, Data Governance, Urban Data Strategy, Urban Data Governance

1. INTRODUCTION

Digital transformation is not feasible without data. Therefore, proper governance of data assets is essential for any organization to create value (Anthony Jnr, 2021; Dremel et al., 2017; Vial, 2019). Nowadays, there is no denying the importance of data, with the EU considering it a vital source of "growth, competitiveness, innovation, job creation, and societal progress" (European Commission, 2022b), benefiting citizens and businesses in many ways through data-driven applications. Today, cities are considered data factories, representing an enormous, rapidly growing, but also heterogeneous data landscape (Barns, 2018; Moustaka et al., 2018). A great deal of research addresses smart services used in different city dimensions, as shown in a literature review by (Bozkurt et al., 2020). Data is the backbone of many services, but what is the right way to deal with it? Cities address this question by publishing data strategies. In this work, we aim to analyze the openly published data strategies of EU cities in the top three countries of the EU Digital Economy and Society Index (DESI) (European Commission, 2022a) (Finland, Denmark, and the Netherlands) and Germany to get an integrated overview of the dimensions of urban data strategies, providing a holistic view of cities' data domains. The results aim to contribute to the body of knowledge of smart cities by providing an overview of the strategic dimensions of data to help practitioners approach the topic of data strategy and provide guidance on the areas in which they should examine their data. In doing so, we explored the following two research questions (RQ) in this paper:

RQ 1: What are the dimensions of cities' data strategies?

RQ 2: How can data governance in cities contribute to achieving the goals of data strategies?

RQ 1 is answered using content analysis of data strategies. The results provide relevant dimensions for a data-driven city. These dimensions allow data governance to be further defined and aligned with the data strategy. RQ 2 views the whole from the perspective of data governance, as data governance can be a promising concept for developing a data-driven smart city.

An analysis of previous scientific publications on urban data strategy yielded no results. We searched the IEEE Xplore, Web of Science, Emerald, and ACM scientific databases for peer-reviewed articles with the search terms "data strategy" AND ("smart city" OR "city"). In total, we found 139 hits, but after analyzing the titles and abstracts, none of the articles met the criteria of an analysis of city data strategies or guidance for developing city data strategies.

The rest of the paper is organized as follows. The next section explains the background concepts of smart cities, data governance, and data strategy. Section 3 describes the research approach to illustrate the data collection and analysis. Section 4 presents the results of the data strategy analysis by describing the dimensions of data strategies. Next, section 5 discusses how data governance contributes to achieving urban data strategies' goals, followed by the implications and limitations of this work. The final section provides a conclusion.

2. BACKGROUND

Increasing urbanization poses new challenges for cities. The UN estimates that ¾ of the world's population will live in cities (United Nations, 2018). To meet the challenges of urbanization, e.g., public transport, waste management, and urban utilities, the concept of smart cities is of great relevance (Ahvenniemi et al., 2017; Jin et al., 2014; Zanella et al., 2014). Smart city research has yielded many definitions, yet providing a unified definition is difficult as perspectives range from information and communication technology (ICT) focus to socioeconomic focus (Albino et al., 2015; Manville et al., 2014). Although there are many definitions, a common denominator seems to be the use of ICT for the operation of smart cities, as shown in several papers (Bozkurt et al., 2020; Caragliu et al., 2011; Lombardi et al., 2012; Washburn & Sinduh, 2010). Nevertheless, just using ICT is not enough; instead, the focus must be on how ICT uses data to drive growth and guide urban development processes (Caragliu et al., 2011). Thus, the success of smart cities is highly dependent on how sound data can be used to decide and shape urban development and growth (Albino et al., 2015). That leads to the need for appropriate handling of data.

Data governance is one concept that addresses the data asset in a way consistent with an organization's overall strategy by defining principles, standards, responsibilities, and processes for the entire data lifecycle. However, it should not be confused with the technical management of data. Instead, it is the decision-making in all data-related issues, which is then implemented in data management (Mahanti, 2021b, 2021a; Seiner, 2014; Weber & Klingenberg, 2021). The literature on data governance in the urban context is sparse (Bozkurt et al., 2022; Kvalvik et al., 2022), although there is already a lot of practical and scholarly work on data governance in the enterprise context (Khatri & Brown, 2010; Mahanti, 2021b, 2021a; Otto, 2011; Seiner, 2014; Weber et al., 2009; Weber & Klingenberg, 2021). In the smart city context, however, only two literature reviews on urban data governance analyze only a handful of relevant literature (Bozkurt et al., 2022; Kvalvik et al., 2022). There is also little work with primary data, e.g., Paskaleva et al. (2017) developed a framework for data governance in smart city initiatives based on interviews with Manchester, Eindhoven, and Stavanger, finding that stakeholder collaboration is a key aspect in urban data governance. In addition, data governance in the city can define the direction for the city's generation, collection, and utilization of data consistently aligned with the city strategy and across all stakeholders (Paskaleva et al., 2017).

Maximizing the business value of data requires a cohesive, enterprise-wide data strategy that addresses how data is collected, stored, managed, and used (Adelman et al., 2005). Data is a cross-cutting issue, especially in smart cities, as many stakeholders and actors are involved in the data ecosystem. Thus, a city must know its data infrastructure and act accordingly to set a strategic direction. The lack of a data strategy can lead to reactive rather than proactive initiatives or to a poorly planned and designed system, which later leads to interoperability issues or vendor lock-ins, and to repeatedly solving the same organizational and technical problems. Therefore, the need for an overarching data strategy that involves all sectors is a collaborative effort to ensure that goals and technology capabilities are aligned to reap the benefits of data (Fleckenstein & Fellows, 2018).

3. RESEARCH APPROACH

In data collection, we followed the basic outline of the gray literature analysis of Garousi et al. (2017), as cities themselves publish data strategies and do not fall under the category of peer-reviewed scientific literature.

For the data analysis, used an inductive coding approach (Chandra & Shang, 2019; Corbin & Strauss, 1990; Gioia et al., 2013) and captured concepts, ideas, and action items that emerged in individual articles through full-text reading.

Figure 1. Research Process

Figure 1 shows the research process. The first phase is planning the review. That includes the step of defining the search words. We defined our search words in a brainstorming session within the author team. First, we created a list of terms that could be relevant to our search. Then, we iteratively adjusted these search terms to obtain relevant search results and cover our research questions. As mentioned in the introduction, our search was limited to the top three countries according to DESI (Finland, Denmark, and the Netherlands) and Germany. In the end, the following search strings emerged:

- *<Denmark/Finland/Germany/The Netherlands>: "**data strategy** <country> **city**" OR "**data governance** <country> **city**" OR "**urban data governance** <country>" OR "**smart city data strategy** <country>"*

The next step in the planning phase of the review is the definition of inclusion and exclusion criteria. We applied the inclusion and exclusion criteria to distinguish the relevant search results from the irrelevant ones, as shown in Table 1.

Table 1. Inclusion and exclusion criteria

Inclusion	The article contains a data strategy of a city that belongs to Germany or to the top three countries, according to DESI.
	The article is published in English or German.
	The article was published by or on behalf of the cities.
	The URL is working and freely available.
Exclusion	The source is non-text based.
	The article contains the duplicated content of a previously examined article.

After the planning phase of the review, applying the search strategy was the next phase of the research process. We applied the search terms in the Google search engine during the period March-April 2023. To conduct a proper analysis, we included the yielded results relevant to our research questions in an Excel file according to our inclusion and exclusion criteria. In this way, we identified a total of 15 relevant articles. We then reviewed them for duplicates and analyzed the articles in detail for their contribution to our research questions. Thus, at the end of the study selection process, eight relevant articles (Table 2) remained for in-depth analysis, which were imported into MAXQDA qualitative data analysis and mixed methods software (MAXQDA, 2022).

Table 2. Identified cities

Country	City	Strategy document
Germany	Cologne	(Blauhut & Stadt Köln, 2021)
Germany	Soest	(Stadt Soest, 2021)
Germany	Monchengladbach	(Tillmanns et al., 2022)
Germany	Berlin	(Land Berlin, 2018)
Germany	Hamburg	(Stadt Hamburg, 2020)
Finland	Helsinki	(The city of Helsinki, 2023)
The Netherlands	Amsterdam	(The city of Amsterdam, 2021)
The Netherlands	The Hague	(The city of The Hague, 2020)

Data analysis was performed using MAXQDA software, in which each document was coded, and the codes created were discussed within the team. In this way, data analysis iteratively produced a series of codes that were eventually combined into broader categories of similar themes in a workshop with the author team, resulting in the 11 dimensions presented in section 4.

4. RESULTS

The following section presents the analysis results of the data strategies, which are synthesized into 11 dimensions of urban data strategies.

(1) Relevance of data:
The relevance of data finds an introductory place in the data strategies. A common understanding is established, and the goals, motivation, and vision are described. That includes communicating the goals of optimizing urban services using data, creating new innovative solutions, making urban planning data-driven, and being an attractive location for businesses. The relevance of data protection, data sovereignty, data sharing, and data quality are addressed in that context (Blauhut & Stadt Köln, 2021; Land Berlin, 2018; Stadt Hamburg, 2020; Stadt Soest, 2021; The city of Amsterdam, 2021; The city of Helsinki, 2023; The city of The Hague, 2020; Tillmanns et al., 2022).

(2) Principles:
The Principles address the fundamental aspects of handling data. (1) Citywide standards for high data quality along the entire data lifecycle must be established with a focus on harmonizing specific data models and visible interface alignment and guidelines. Master data management, including reference data, must be revised to develop administrative and citywide standards for elementary attributes of data sets (Blauhut & Stadt Köln, 2021; The city of Helsinki, 2023; Tillmanns et al., 2022).

Next, (2) Data quality must be ensured over the long term, including syntactic and semantic data quality, and redundancies and non-meaningful data must be minimized. KPIs are a tool to measure data quality based on criteria such as timeliness, control intensity, etc. (Stadt Hamburg, 2020; Stadt Soest, 2021; The city of Helsinki, 2023; The city of The Hague, 2020; Tillmanns et al., 2022).

Ensuring (3) sovereignty is another principle: The data about the city must be available to the city and remain so. Citizens and visitors have a right to know how their data is being used. They must always have easy access to their data and be able to use the opportunities offered by laws and regulations. Control over data must be well-organized and transparent. In doing so, partners working with the city must be transparent about using technology and data to solve urban challenges. Dependence on specific technologies and data must be reduced, and the public sector must be able to decide on access rights to the holistic urban data space (Blauhut & Stadt Köln, 2021; The city of Helsinki, 2023; The city of The Hague, 2020).

Finally, the 4) data in and about the city must be classified. The goal of data classification is to identify data sets characterized by standard rules and working methods in order to finally define data usage guidelines adapted to the data classes, improve the usability of the data and design data architectures tailored to the application (The city of Helsinki, 2023).

(3) Methods:
The methods address approaches and tools in handling data. Thus, a roadmap defines the next steps for realizing the strategic goals and provides a transparent common understanding. This ensures that the strategy is put into action and avoids unclarity due to the overwhelming volume of goals (Stadt Soest, 2021; The city of Helsinki, 2023). The goals are to be addressed step by step in an agile working method with all stakeholders in the city. Data governance is seen as a dynamic and continuously evolving framework for implementing the data strategy. This approach improves operational procedures, and the data strategy adapts and scales (Stadt Hamburg, 2020; Tillmanns et al., 2022). Furthermore, success must be measurable by defining KPIs and regular benchmarking (The city of Helsinki, 2023).

A data catalog is a central tool in the data strategy, which contains the data with descriptive metadata and evolves constantly. The goal is to create a citywide overview of the city's data assets, what quality parameters it has, and who the data owners are. Cities can improve data quality, clean up redundant data sets, know about the current data sets in various departments, and provide all interested stakeholders with an overview of the data that is relevant to them and its sources using the data catalog (Stadt Soest, 2021; Tillmanns et al., 2022).

(4) Data exchange:
The data-exchange aspect is divided into open data and non-open data. Cities want to make more data available to the public, e.g., by making frequently requested data available online by default. Open data is the basis for

information and participation and for civil society to develop its own applications. In addition, open data lays the foundation for transparency and trust in city government. In addition to raw, machine-readable data, services based on open data must also be made openly accessible, i.e., applications created with public funds must also be accessible to the public. Cities have many creative businesses and residents that could use open urban data to develop practical data-based applications for the urban ecosystem. This is how to sustain the vitality of the urban ecosystem (Blauhut & Stadt Köln, 2021; Land Berlin, 2018; Stadt Soest, 2021; The city of Amsterdam, 2021; The city of Helsinki, 2023; Tillmanns et al., 2022).

Non-open data, on the other hand, refers to data that other units can use within the city administration, following legal considerations. By combining different data sources, new insights are possible, and at the same time, the effort for further data collection can be reduced and redundancies minimized. Data should not be considered in their departmental silos but rather across departmental boundaries. The classification of data sets as open data or internal use only, as well as access control, must be decided through data governance processes (Stadt Hamburg, 2020; Stadt Soest, 2021; The city of Helsinki, 2023; Tillmanns et al., 2022).

(5) Technology:

The technological dimension is primarily about modernizing the data infrastructure regarding homogeneity, security, reliability, and integrability. A city's data infrastructure must be designed so that there is no lock-in to individual vendor solutions by relying on open source. In the long term, scalability must be ensured to deal with ever more and larger volumes of data. In addition, modern analytics environments and tools are needed to process data faster and more securely (Blauhut & Stadt Köln, 2021; The city of Amsterdam, 2021; Tillmanns et al., 2022).

In terms of analytics environments and data-driven decision-making, machine learning methods are applied to develop forecasting models and urban planning simulations visualized in user-centric dashboards (The city of Helsinki, 2023). Generally, an urban data platform is a key concept in urban data infrastructure. It is considered the backbone of the urban data infrastructure and facilitates data-driven decision-making through an integrated view and standardized interfaces. As a system of systems, an urban data platform connects various data sources through appropriate interfaces and enables regulated access to each data set. Based on this infrastructure, for example, city subsidiaries and the administration can share data, and an open data portal can be connected to the urban data platform to publish specific data sets (Stadt Hamburg, 2020; Stadt Soest, 2021; Tillmanns et al., 2022).

(6) Data culture:

Data is not merely about processes or technology but also people and culture. This is reflected in cities' data strategies. The focus is on communicating the benefits of data use and sharing in the course of data-driven applications and decisions to citizens, employees, and all other stakeholders in the city. It is vital to demonstrate real-world example projects to make them tangible for everyone, as data projects are abstract for many people, therefore it is often not directly apparent why one should engage with data. Communication of the benefits of data must be incorporated into city activities (The city of The Hague, 2020), so prototypes and showcases are essential means of communication (Tillmanns et al., 2022).

Further, thinking in domain silos must be broken down. The reluctance to share one's data with others must be overcome, and the understanding must increase that data and information often improve when shared with others. Combining different data sources creates new insights and reduces the need to recreate data. Clarifying the benefits of data at all levels of the hierarchy fosters motivation and a shared sense of long-term direction for data-driven work (The city of Amsterdam, 2021; Tillmanns et al., 2022).

(7) Data ethics:

Dealing ethically with data is equivalent to being ethical in the physical world, so cities want to ensure peace, justice, and security in the digital world (The city of The Hague, 2020). Cities must respect societal values when implementing data-driven activities, as not everything legal is legitimate (Tillmanns et al., 2022). The intention is to prevent unintended negative consequences of data use (Stadt Soest, 2021). Ethics committees ensure the ethical handling of data by reviewing the entire data lifecycle from openness, transparency, and fairness aspects (Blauhut & Stadt Köln, 2021; Stadt Hamburg, 2020; Stadt Soest, 2021; The city of Amsterdam, 2021; The city of Helsinki, 2023; The city of The Hague, 2020; Tillmanns et al., 2022). In this way, cities aim to ensure the traceability of results with high transparency and away from data-based applications being a black box (Tillmanns et al., 2022). The results and workflows must be transparently

traceable and reproducible for third parties. This will minimize potential disadvantages or discrimination against groups or individuals (Stadt Soest, 2021).

The data ethics dimension is mainly guided by the FAIR principles, which consist of the principles of (1) findability, meaning that users can easily find the data they need. The (2) accessibility, by being retrievable by their identifier through a standardized communication protocol. The (3) interoperability, when the data use commonly accepted identifiers, data models, and ontologies and finally the (4) reusable use of the data, in which the attributes are precisely described, use licenses are clearly defined, and the source is given (The city of Helsinki, 2023).

(8) Organization structure:

Dealing with data requires structural and organizational changes. Data governance and management teams within existing structures are required, and responsibilities and roles must be defined (Blauhut & Stadt Köln, 2021; Stadt Hamburg, 2020; Stadt Soest, 2021; The city of Helsinki, 2023; The city of The Hague, 2020; Tillmanns et al., 2022). For example, roles such as data owner and data manager within departments must be defined, allowing for decentralized implementation, but data matters must be governed centrally (Stadt Hamburg, 2020). For this reason, a steering group made up of cross-departmental representatives from each department derives specific measures from the data strategy centrally (Soest, 2021). Thus, data governance defines how strategic use can be ensured regarding organization and processes (Tillmanns et al., 2022).

Data governance controls data management, which executes data processes operationally. Data governance forms the basis for coordinated data management through rules, organization, and processes on a functional and technical level (Stadt Soest, 2021; The city of Helsinki, 2023; Tillmanns et al., 2022). In contrast, the city's data management primarily ensures that the data processed in the city is available and provided in the desired form through predefined rules and quality controls (Stadt Soest, 2021; The city of Amsterdam, 2021; The city of Helsinki, 2023; Tillmanns et al., 2022).

(9) Data security and privacy:

Data security and privacy are basic requirements for gaining citizens' trust. Cities must ensure that data collected as private is also processed as confidential (Blauhut & Stadt Köln, 2021; Stadt Hamburg, 2020; Stadt Soest, 2021; The city of Amsterdam, 2021; The city of The Hague, 2020). The principle of "privacy by design" ensures that all hardware and software used in collecting and processing personal data is designed to keep the data secure during its entire lifecycle (Stadt Soest, 2021). However, data security and data protection should not be seen as an obstacle to innovation but rather as a driver of it. There is a high degree of uncertainty about who may use which data and how, and what formal requirements must be met. This uncertainty leads to data being shared and used at a lower rate than it could be (Tillmanns et al., 2022).

Furthermore, not all data need the same level of protection and must be treated differently. Data protection classes can help define a coherent set of rules that determine which data needs which level of security, considering the criteria of confidentiality, integrity, and availability (Tillmanns et al., 2022).

(10) Cooperations:

A stable and sustainable data ecosystem relies on coopeations with all internal and external stakeholders (The city of The Hague, 2020). Stakeholders must be identified through iterative dialog processes, and the city's data ecosystem must be strengthened through collaborative processes, ultimately leading to the basis for better services (Blauhut & Köln, 2021). This is not only to create a good foundation for successful project implementation but also to ensure a high level of acceptance and participation of all interested parties so that projects can be deeply embedded in the city's overall development (Blauhut & Stadt Köln, 2021).

Inter-municipal cooperations are a great way to benefit from the experiences of others, as cities often face the same challenges, saving time and costs. Moreover, data is not bound by city boundaries, so the urban data ecosystem must also provide the ability to implement data-driven solutions across cities. In the end, collaboration with neighboring cities helps tackle complex problems that cannot be solved alone due to limited resources (Stadt Soest, 2021; The city of Amsterdam, 2021; Tillmanns et al., 2022).

(11) Data literacy:

Building data literacy is critical for cities as data becomes prevalent in traditional areas that were not previously data-driven. Building data literacy aims to ensure that the knowledge and skills needed to gain insights from data and use data consciously and responsibly are in place (The City of The Hague, 2020; Tillmanns et al.,

2022). Data literacy applies to all disciplines but also aims to provide citizens with a basic understanding of using technology to make better decisions and know the risks and consequences of sharing their data (The city of The Hague, 2020).

Cities must invest in staff training and empowerment, both in individual training and in the development of a learning organization (Tillmanns et al., 2022). Training programs must be established divided into two categories. The (1) basic training program for all employees builds a common understanding and knowledge about the use of data. Only on the basis of this common ground will it be possible to develop the city government into a data-driven organization (Stadt Soest, 2021; Tillmanns et al., 2022). In addition to the basic training program, employees require specific data skills, depending on their position and task. Employees responsible for data governance and management processes must complete training in the specific areas of responsibility by (2) professional advanced training programs. Investing in their employees will increase the cities' attractiveness as an employer in the long term (Stadt Soest, 2021; The city of Helsinki, 2023; Tillmanns et al., 2022).

5. DISCUSSION

Urban data strategies divide into the 11 dimensions detailed previously. The relevance of data is at the forefront of data strategies. It establishes a shared commitment to data utilization, lists possible scenarios for data utilization, and identifies how the city, citizens, businesses, and other stakeholders can benefit from data. Data utilization must be deeply rooted in the culture and mindset to be successful in a sustainable manner. All stakeholders must be made aware of the benefits that proper use of data brings to the public good. The data strategy can only be accepted if all have a common understanding. To this end, it is necessary to communicate fundamental knowledge and present comprehensible application scenarios from everyday practice demonstrating the benefits. This also includes a transparent portrayal and educational campaign concerning data security and protection, as a large amount of sensitive data is in municipal hands – incorrect handling of data damages the trust of stakeholders. The city needs to ensure that data is handled properly throughout its entire lifecycle but also that the right measures are taken for the further innovative development of data ecosystems, e.g., through open data, collaboration with municipal stakeholders, and the development of innovative technologies, including data portals and digital twins. Therefore, data governance is the executive hand of the data strategy, assuming a data governance organizational unit exists in the city (Stadt Hamburg, 2020; Stadt Soest, 2021; The city of Amsterdam, 2021; Tillmanns et al., 2022).

Bozkurt et al. (2022) define urban data governance as a unit that deals with all data-related issues from a holistic, overarching perspective to pursue sustainable urban development by governing data assets in the interest of citizens and promoting business and services. The focus is on sustainable urban development, which is consistent with the overall strategy of cities. The city's data governance unit can operationalize the dimensions listed in the data strategy in collaboration with relevant stakeholders (internal and external) through defined actions. For example, the data governance unit can define city-wide standards and data quality principles according to the city's needs to ensure the highest level of interoperability and integrability. In addition, given its data access scope (Bozkurt et al., 2022), the data governance unit aims to establish policies and mechanisms for data exchange in an internal sense, external sense, and open data. To this end, it plays a crucial role in data classification and the creation of a citywide data catalog to promote the sharing of data of different uses and, through the data catalog, provide a standardized view of the city's data resources, promoting not only technological and semantic compatibility, but also data ethics issues such as trust, ownership, and transparency. This ultimately benefits the implementation of data platforms and other digital services. Bozkurt et al. (2022) eight dimensions of urban data governance have many connecting elements and parallels to the, in this paper, identified dimensions of data strategy. Thus, data governance drives the city's data strategy and hinders the strategy from becoming a dusty document on the shelf. Data governance is a concept that needs to be practiced continuously by all stakeholders and requires a cultural change and transformation process. This can be achieved by regularly educating the staff and building data literacy in the long term.

With this work, we intend to provide an overview of how urban data strategies are constructed (RQ1) and how they can be operationalized through data governance (RQ2). Data has immense importance for the development of cities, be it for data-driven urban planning or sensor-based services in the city. This paper aims to contribute to this critical smart city area, which has received little attention from the academic community

concerning data strategies and data governance in smart cities. The identified 11 dimensions of data strategies show that data must be strategically considered from a multifaceted perspective. This ranges from technology and data protection to cultural factors in dealing with data. Data governance as a tool for implementing data strategies can help cities determine gaps to ensure their data strategy is not just a strategic paper dusting away in a drawer. In this way, practitioners can use the results of this work as initial guidance on how to develop the data strategy and understand its implementation. In addition to this practical contribution, there is also a scientific contribution in further exploring the body of knowledge about smart cities from the data perspective and integrating data governance further into the smart city discussion. Scholars can use the results of this work, for example, to develop a maturity model of the data aspects of smart cities and transfer concrete data governance measures to cities. Still, we do not consider the identified 11 dimensions of data strategy for smart cities as final and take it as a starting point for further development in further research, such as broadening the sources by expanding the search criteria. In addition, it is quite possible that the data strategies are not publicly available and may be requested directly from cities; ideally, this can be expanded through primary data collection in the form of interviews with cities. Finally, smart city strategies, which are more accessible openly than data strategies, can be further analyzed to identify other strategic directions related to data, as the data aspect may also be present in some smart city strategies.

6. CONCLUSION

This paper has examined 8 data strategies of cities in Germany, Finland, Denmark, and the Netherlands to contribute to the knowledge of smart cities by providing an overview of the strategic dimensions of urban data and clarifying how data governance in cities can contribute to data strategy. The analysis revealed the data strategy dimensions of (1) relevance of data, (2) principles, (3) methods, (4) data exchange, (5) technology, (6) data culture, (7) data ethics, (8) organizational structure, (9) data security and privacy, (10) collaboration, (11) data literacy. In addition, it shows that data governance is a concept to put these 11 strategic dimensions into practice through standardization efforts, training programs, defining roles and responsibilities, and developing a data catalog. Data governance drives the city's data strategy and prevents the strategy from becoming a dusty document on the shelf.

Ultimately, the findings help practitioners approach the topic of data strategy and provide guidance on areas where they should examine their data. Further, the scientific community is expanding knowledge about urban data and the strategic dimensions of cities seeking to be data-driven and providing a basis for developing further work in areas such as the maturity model for data-driven cities and data governance concepts for smart cities.

REFERENCES

Adelman, Sid., Abai, Majid., & Moss, L. Terpeluk. (2005). *Data strategy*. Addison-Wesley.

Ahvenniemi, H., Huovila, A., Pinto-Seppä, I., & Airaksinen, M. (2017). What are the differences between sustainable and smart cities? *Cities*, *60*, 234–245.

Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology*, *22*(1), 3–21.

The city of Amsterdam. (2021). *City of Amsterdam Data Strategy*. https://assets.amsterdam.nl/publish/pages/1017819/data_strategy_city_of_amsterdam_2021-2022.pdf

Anthony Jnr, B. (2021). Managing digital transformation of smart cities through enterprise architecture–a review and research agenda. *Enterprise Information Systems*, *15*(3), 299–331.

Barns, S. (2018). Smart cities and urban data platforms: Designing interfaces for smart governance. *City, Culture and Society*, *12*, 5–12.

Land Berlin. (2018). *Gründbuch für die Digitalisierungsstrategie des Landes Berlin*.

Blauhut, D., & Stadt Köln. (2021). *Daten "machen" Strategie*. https://buergerinfo.stadt-koeln.de/getfile.asp?id=821964&type=do

Bozkurt, Y., Braun, R., Rossmann, A., & Hertweck, D. (2020, July 21). The Advent of Smart Cities: Status Quo and Future Research Directions. *6th IADIS International Conference on Connected Smart Cities (CSC2020)*.

Bozkurt, Y., Rossmann, A., & Pervez, Z. (2022). A Literature Review of Data Governance and Its Applicability to Smart Cities. *Proceedings of the 55th Hawaii International Conference on System Sciences 2022*, 2680–2689.

Caragliu, A., del Bo, C., & Nijkamp, P. (2011). Smart Cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82.

Chandra, Y., & Shang, L. (2019). Inductive Coding. In *Qualitative Research Using R: A Systematic Approach* (pp. 91–106). Springer Nature Singapore.

Corbin, J., & Strauss, A. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. In *Qualitative Sociology* (Vol. 13).

Dremel, C., Herterich, M., Wulf, J., Waizmann, J.-C., & Brenner, W. (2017). How AUDI AG Established Big Data Analytics in Its Digital Transformation. *MIS Quarterly Executive*, *16*(2).

European Commission. (2022a). *Digital Economy and Society Index (DESI) 2022*. https://digital-strategy.ec.europa.eu/en/policies/desi

European Commission. (2022b). *A European Strategy for data*. https://digital-strategy.ec.europa.eu/en/policies/strategy-data

Fleckenstein, M., & Fellows, L. (2018). Modern Data Strategy. In *Modern Data Strategy*. Springer International Publishing.

Garousi, V., Felderer, M., & Mäntylä, M. V. (2017). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology*, *106*, 101–121.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, *16*(1), 15–31.

The city of Helsinki. (2023). *The City of Helsinki Data Strategy - Digitaalinen Helsinki*. https://digi.hel.fi/english/helsinki-city-data-strategy/

Jin, J., Gubbi, J., Marusic, S., & Palaniswami, M. (2014). An information framework for creating a smart city through internet of things. *IEEE Internet of Things Journal*, *1*(2), 112–121.

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152.

Kvalvik, P., Sánchez-Gordón, M., & Colomo-Palacios, R. (2022). Beyond technology in smart cities: a multivocal literature review on data governance. *Aslib Journal of Information Management*.

Lombardi, P., Giordano, S., Farouh, H., & Yousef, W. (2012). Modelling the smart city performance. *Innovation: The European Journal of Social Science Research*, *25*(2), 137–149.

Mahanti, R. (2021a). Data Governance and Compliance. In *Data Governance and Compliance*. Springer Singapore.

Mahanti, R. (2021b). *Data Governance and Data Management Contextualizing Data Governance Drivers, Technologies, and Tools Data Governance*. Springer Nature.

Manville, C., Cochrane, M., Cave, J., Liebe, A., & Massink, R. (2014). *Mapping Smart Cities in the EU*. https://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOL-ITRE_ET(2014)507480_EN.pdf

MAXQDA. (2022). *MAXQDA | All-In-One Tool for Qualitative Data Analysis & Mixed Methods - MAXQDA* (2022.4). https://www.maxqda.com/

Moustaka, V., Vakali, A., & Anthopoulos, L. G. (2018). A Systematic Review for Smart City Data Analytics. *ACM Computing Surveys (CSUR)*, *51*(5).

Otto, B. (2011). Data governance. *Business and Information Systems Engineering*, *3*(4), 241–244.

Paskaleva, K., Evans, J., Martin, C., Linjordet, T., Yang, D., & Karvonen, A. (2017). Data governance in the sustainable smart city. *Informatics*, *4*(4).

Seiner, R. (2014). *NON-INVASIVE DATA GOVERNANCE: The Path of Least Resistance and Greatest Success* (1st ed.). Technics Publications.

Stadt Soest. (2021). *DATENSTRATEGIE SOEST*. https://digital-soest.de/images/AIDW/Datenstrategie_Final_2021-05-14.pdf

Stadt Hamburg. (2020). *DIGITALSTRATEGIE FÜR HAMBURG*. https://static.hamburg.de/fhh/epaper/digitalstrategie/#0

The city of The Hague. (2020). *Towards a data-driven and responsible city - Data strategy 2020-2022*.

Tillmanns, K., Platzer, M., Lenzen, I., & Mönchengladbach. (2022). *Leitlinien Datensouveränität*. https://www.moenchengladbach.de/fileadmin/moendico/Leitlinien_Datensouver%C3%A4nit%C3%A4t_MG.pdf

United Nations, D. of E. and S. A. (2018). *World Urbanization Prospects: The 2018 Revision*. https://population.un.org/wup/Download/Files/WUP2018-F21-Proportion_Urban_Annual.xls

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. In *Journal of Strategic Information Systems*, 28(2), pp. 118–144. Elsevier B.V.

Washburn, D., & Sinduh, U. (2010). Helping CIOs Understand "Smart City" Initiatives. *FORRESTER*.

Weber, K., & Klingenberg, C. (2021). *Data Governance - Der Leitfaden für die Praxis*. Hanser.

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All---A Contingency Approach to Data Governance. *Journal of Data and Information Quality*, *1*(1).

Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of things for smart cities. *IEEE Internet of Things Journal*, *1*(1), 22–32.