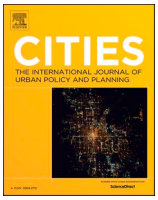
Cities 148 (2024) 104859

Contents lists available at ScienceDirect

Cities

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



Smart city development: Data sharing vs. data protection legislations

Aaron Joyce, Vahid Javidroozi *

Faculty of Computing, Engineering, and Built Environment, Birmingham City University, Birmingham B4 7XG, UK

ARTICLE INFO

Keywords:
Smart cities
Smart city development
Data sharing
Data protection
GDPR
Systems integration

ABSTRACT

Fuelled by cutting-edge technology, smart cities offer vast potential to enhance urban living, elevate service delivery, and foster sustainability. Central to this digital transformation is seamless data exchange across city sectors like transportation, healthcare, energy, education, and so on. However, stringent data protection laws pose a significant hurdle to cross-sectoral data sharing in smart cities. This study explores the multifaceted aspects of cross-sectoral data sharing within smart cities. It investigates its significance, identifies barriers, and outlines prerequisites while navigating data protection regulations. The findings highlight a critical gap in the existing literature, emphasising the need for a comprehensive framework. Guided by a thematic analysis of insights from 14 interviews with expertise in smart cities, emphasising data-sharing and data protection legislation, this framework provides a systematic approach to enable cross-sectoral data-sharing in smart cities. It serves as a guide for urban planners and stakeholders. Following its initial development, the framework undergoes validation through interviews with six smart city experts. Their insights confirm the framework's robustness and efficacy in addressing the research question. Hence, this research advances the discourse on smart city development and offers a blueprint for balancing cross-sectoral data-sharing with data protection regulations, bridging the gap between innovation and compliance.

1. Introduction

Smart cities are comprised of data, processes, connected systems, and advanced technologies which together have the purpose of solving complex city problems, the key to improving the quality of life and providing improved services within the city (Bibri, 2018). Data plays a central role in the functioning of smart cities and is generated by various sources such as networks of sensors, traffic systems and citizens' everyday devices, enabling improved operations within the city (Souza et al., 2017). Collecting and analysing the massive amount of data generated by smart cities can provide valuable insights used in the development of services, such as reducing traffic congestion or tackling the inefficiency of energy usage (Vanli & Akan, 2023). However, data sharing is essential to create significant value from the data generated and achieve the objectives of a smart city. This is because it facilitates collaboration between connected systems such as Smart Environment, Smart Energy, Smart Building, Smart Transportation, Smart Health, and Smart Education (Lim & Maglio, 2018).

Cross-sectoral Data Sharing (CDS) has been considered vital for the successful development of smart cities and necessary to provide efficient solutions that can be created within cities (Correia et al., 2021; Guenduez & Mergel, 2022; Kieu et al., 2022). Valuable insights derived from combined data sets can promote collaboration between stakeholders, as often this leads to more informed decision-making (Dorri et al., 2017). CDS contributes to breaking down data silos, addressing some of the key challenges faced by smart cities and the development of holistic data approaches including data integration, data governance, data privacy and data security. However, achieving the successful sharing of data between sectors is faced with challenges, such as interoperability between the systems of different sectors and ensuring data quality (Asswad & Gómez, 2021).

Data Protection Legislation (DPL) adherence is one of the most significant challenges to CDS, especially considering the amount of sensitive data collected within a smart city (Abdalla et al., 2019; Rao & Deebak, 2023). DPLs are essential for safeguarding the privacy and ensuring the security of sensitive data within smart city infrastructures (Cui et al., 2018; Gruschka et al., 2019). However, it is crucial to note that while these regulations aim to protect personal data, they can also introduce complexities that impact the efficacy of CDS (Paskaleva et al., 2017).

Under GDPR, sensitive data refers to personal information of a particularly sensitive nature that requires higher protection. It includes

* Corresponding author.
E-mail addresses: Aaron.Joyce@mail.bcu.ac.uk (A. Joyce), Vahid.Javidroozi@bcu.ac.uk (V. Javidroozi).

https://doi.org/10.1016/j.cities.2024.104859
Received 28 October 2023; Received in revised form 30 January 2024; Accepted 1 February 2024
Available online 23 February 2024
0264-2751/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data for uniquely identifying a person, health-related information, or data concerning a person's sex life or sexual orientation. Non-sensitive data, on the other hand, encompasses all other personal information that does not fall within the sensitive data categories. It covers general personal data such as names, contact details, identification numbers, online identifiers, and any other information that does not fit into the sensitive data classification (A&L Goodbody, 2016).

1.1. Problem statement

Data protection and privacy regulations play a critical role in safeguarding personal data throughout the CDS process. They establish necessary boundaries and frameworks to ensure data security and protect individuals' privacy rights. However, the strictures imposed by these regulations might inadvertently pose challenges by restricting certain aspects of data access, processing, and sharing, potentially impeding the smooth execution of CDS initiatives within smart cities. On one hand, CDS is necessary for the functioning of smart cities as it enables data-backed decision-making and innovative solutions across various sectors. On the other hand, DPL has the purpose of protecting the privacy of data subjects and ensuring that personal data is used ethically (Vandercruysse et al., 2020). The tension between regulatory compliance and the seamless execution of CDS is an intricate balance. While regulations are fundamentally designed to protect sensitive data, their stringent enforcement can at times present operational hurdles in achieving optimal CDS effectiveness. Striking the right balance between compliance and facilitating effective data sharing processes remains a notable challenge in the landscape of smart city development. Therefore, city sectors sharing data within smart cities must apply measures to cover the key principles of protecting data, such as anonymisation, encryption and data access controls to name a few (Yang et al., 2019).

Accordingly, several established standards and recommendations have been developed and they play an essential role in defining the guidelines for CDS while adhering to DPLs. For example, ISO/IEC 30182:2017 focuses on conceptualising a model for data interoperability within the smart city context (ISO, 2017). This standard underscores the significance of data sharing in driving city initiatives and highlights the role of private sector engagement, emphasising the implications of such participation concerning data protection legislation. However, while this standard outlines the advantages of a shared framework and common data understanding, it lacks explicit guidelines or detailed measures for ensuring data protection compliance in CDS practices. Another example is the ITU-T Recommendation X.1252 (International Telecommunication Union (ITU), 2021), which aims to define key terms in identity management, crucial for comprehending and standardising identity-related processes. While the terms outlined in this recommendation form a baseline understanding of identity management work, their direct applicability to the specific measures required for data protection within smart city data sharing may necessitate a more tailored approach. The recommendation's focus on defining terms central to identity management might offer limited direct guidance concerning broader data protection and sharing within smart cities.

Moreover, there are other standards which offer standard definitions for various terminologies within the area of data sharing for different purposes including smart city development. For example, the ITU-T Recommendation Y.101, designated by the International Telecommunication Union (ITU) (2000), provides fundamental terms and definitions crucial for understanding the Global Information Infrastructure.

Incorporating these standards and recommendations into the discourse on data protection and city data sharing within smart cities provides a foundational understanding. However, a critical analysis reveals a potential gap between these standards' broader definitions and the specific measures essential for safeguarding data in the context of CDS within smart city environments. Hence, while these standards offer valuable insights, a more nuanced exploration and integration of domain-specific guidelines may be required to comprehensively address data protection concerns in smart city data sharing initiatives.

Despite the importance of successfully achieving CDS whilst adhering to DPL for the development of smart cities, there is a clear lack of understanding within the existing literature on overcoming this challenge. To elaborate on this point, the importance, challenges, and strategies in achieving CDS within smart city developments have been partially covered in the literature. However, the impacts of adhering to DPL and the relationship between the two have not been thoroughly researched. Most existing literature covering this research topic has explored various aspects of smart city development such as broad data sharing frameworks or research into smart city process integration.

For example, Chiara (2021) underscores the intertwined yet distinct nature of privacy and data protection, highlighting the complexity of their practical differentiation, which echoes the sentiments echoed by Loideain (2019). Bargh et al. (2014) and Loideain (2019) both emphasize the necessity of dynamic, feedback-driven mechanisms for privacy policy adaptation and the application of GDPR provisions to empower individuals and regulate data sharing. However, despite these significant strides, a considerable gap persists. While the GDPR marks a critical advancement in enhancing user control and defining responsibilities for data controllers and processors (Loideain, 2019), it falls short of fostering a robust culture of compliance within the IoT ecosystem. Loideain (2019) emphasises the importance of effective GDPR application, oversight, and enforcement, pointing out the challenges of ensuring accountability, enforcement across the IoT supply chain, and providing relevant guidance. Furthermore, a smart city data framework by Liu et al. (2017) focuses on data quality and privacy for smart city data. The framework provides insights into managing smart city data; however, it does not address CDS or the challenges involved with adhering to DPL, supporting the need for research into the relationship between CDS and DPL. Furthermore, a research paper by Javidroozi et al. (2019a) explores Business Process Change in smart city developments. Although the paper presents a strong discussion, due to the focus of the paper it does not address CDS or DPL, further stressing the need for research in this area. Previous smart city research has also revealed several technical solutions which have attempted to solve CDS and DPL. However, it must be noted that these technical solutions do not provide an inclusive solution to how CDS can be achieved, such as by providing a practical guideline covering the challenges and requirements. Technical solutions include a solution by Anisetti et al. (2021), where they proposed an access control system that is supposed to solve data sharing and data protection within a big data environment. Although the solution is a step forward in the technical data sharing aspect of CDS, it does not cover achieving CDS comprehensively. This collective body of work illustrates the evolving landscape but also underscores the need for a more comprehensive approach to bridge the gap between legal frameworks, industry practices, and effective regulatory oversight within the complex smart city development. Further research and practical application are imperative to foster a culture of compliance and to address the multifaceted challenges associated with data sharing, privacy, and security in the smart city realm.

1.2. The research purpose

To mitigate these issues, the study intends to explore the multifaceted concepts and requisites integral to enabling CDS, while duly acknowledging the significance of adhering to DPLs. Through an extensive examination of these facets, the outcomes will culminate in the construction of a framework, which will function as a comprehensive guideline, facilitating the efficient execution of CDS initiatives within the diverse sectors of smart cities. As a consequential resource, this framework is poised to offer significant advancements for urban planners and stakeholders engaged in the development and execution of

CDS-oriented smart city endeavours.

The rest of the paper is structured as follows: Section 2 presents a comprehensive literature review, to find, justify, address, and support the research gap and involves investigating the current state of knowledge on CDS and DPL in the context of smart cities. Section 3 presents the research methodology used to address the research gap. Sections 4 and 5 present the key findings from thematic data analysis and the developed conceptual framework guideline for achieving CDS. Section 6 concludes the study, with validation of the conceptual framework, conclusions of the research, and recommendations for future work.

2. Literature review

The literature review was conducted by targeting two key academic databases, employing specific filters to ensure the retrieval of high-quality and relevant research. Emphasising recent studies aligned with the rapidly evolving nature of smart city trends, the review mainly encompassed papers published between 2018 and 2022. The screening process followed the PRISMA approach, systematically assessing and applying criteria. This methodological approach aimed to ensure the inclusion of contemporary, credible, and impactful studies relevant to the study's objectives. Hence, the following section provides the results of a thematic literature review conducted in this study.

2.1. The role of data in smart city

A smart city is defined as an urban area that utilises a variety of technologies to collect data from different domains of the city, aiming to enhance the quality of life for its residents and optimize city operations (Khan & Zia, 2021). With over half of the global population residing in urban areas, a projection suggests that by 2050, 68 % will be urbanised (United Nations (UN), 2018). Smart cities encompass various components, including Smart Environment, Smart Energy, Smart Building, Smart Transportation, Smart Health, and Smart Education, serving as indicators of a city's level of “smartness” (Cui et al., 2018).

The field of smart cities is continuously evolving, leading to the development of numerous frameworks, ranging from initial frameworks such as a general smart city framework developed by CISCO (Falconer & Mitchell, 2012) to more specific frameworks such as Technology-Organisation-Environment (TOE) based framework developed by Ullah et al. (2021) for risk management in smart sustainable cities, and a process-centric framework developed by Javidroozi et al. (2023). However, while several frameworks for smart city data sharing have emerged recently, few focus on achieving Cross-Sectoral Data Sharing (CDS) while adhering to DPLs.

Smart city data is pivotal in transforming sectors like transportation, health, environment, and energy. For instance, in transportation, it optimises traffic flow and commuting experiences. Health benefits from predictive analytics and personalised treatments. Energy sees real-time monitoring and sustainable strategies. In the environmental sector, analysing extensive weather data has proven valuable for agricultural development (Bibri & Krogstie, 2017; Hashem et al., 2016). This underscores data's role as a critical asset in smart cities. Additionally, smart city data allows real-time data collection, aggregation, and analysis from sensors across different sectors, facilitating informed decision-making (Khan & Zia, 2021). Traffic sensor data, as one example, can enhance city operations and services (Lim & Maglio, 2018).

As cities rapidly evolve into smart cities, leveraging data across various sectors becomes imperative for fostering collaboration, enhancing decision-making, and improving urban development and services, including transportation and environmental sustainability (Arroub et al., 2016; Khawaja & Javidroozi, 2023). This highlights the potential of utilising data generated within smart cities to address challenges like traffic congestion and environmental sustainability.

2.2. Open data and data sharing in smart cities

In the dynamic landscape of smart city development, open data emerges as a linchpin promoting collaboration and innovation across diverse sectors. It's not just about data accessibility; it's about empowering citizens and stakeholders alike. With data at their fingertips, citizens engage more actively, influencing decisions and shaping the city's future. Beyond transparency, open data fuels innovation, inspiring new solutions to age-old urban challenges. By intertwining datasets from various sectors, it's not merely about optimising traffic or improving services; it is about creating a city that thrives a city that is smarter, more efficient, and fosters economic growth. This open data is not just numbers and statistics; it is the bedrock for job creation, sparking new industries and nurturing an ecosystem of innovation (Nikiforova & Zuiderwijk, 2022). For developing smart cities, open data serves as a key component, facilitating multi-scale urban management and enhancing several key qualities like adaptability, efficiency, and transparency (Lnenicka et al., 2022). This accessible data stands as a cornerstone, fostering community participation and trust, prerequisites for the establishment of smart, participative communities (Berardi et al., 2019). The significance of open data in smart cities extends to its potential to address previously underexplored ecosystem processes and dynamics, offering more sophisticated modelling capabilities (Nitoslawski et al., 2019).

Transparency, a cornerstone of open government and open data initiatives, synergises with participation and collaboration processes in the context of smart cities (Lnenicka et al., 2022). Open data repositories, often incorporated into urban data platforms, embody a wealth of diverse data types, essential for addressing civil society issues, enhancing transparency, and bridging gaps between local governments and citizens. The deployment of these data portals and platforms aligns with the broader goal of encouraging transparent, responsive governance and furthering community involvement (Barns, 2018).

However, despite the potential and growing recognition of the value of open data in smart cities, progress has been slower than anticipated, impeding the advancement of these cities and diminishing their socio-economic values. The challenges in harnessing the full potential of open data are multi-faceted. They encompass not only the technical aspects, such as data quality and accessibility (Nikiforova & Zuiderwijk, 2022) but also social and cultural elements, including fostering public engagement and addressing communication gaps between governments and citizens (Berardi et al., 2019).

Various models and platforms, from university-led initiatives to government-run services, have emerged to leverage open data effectively in smart cities. These platforms range from simple open data listings to more sophisticated urban data platforms and data marketplaces (Barns, 2018). While some focus on benchmarking city performance, others prioritise making data more accessible and relevant, offering a diverse range of real-time and static datasets. The integration of open data into these platforms not only fosters transparency but also enables better decision-making, community engagement, and innovation.

2.3. Cross-sectoral data sharing

Collaboration among sectors within smart cities is integral to their development, performance, and overall success (Clement, Manjon, & Crutzen, 2022). The establishment of collaborative data ecosystems is central to smart cities, offering a potential solution for Cross-Sectoral Data Sharing (CDS). Although an official definition is lacking, CDS can be understood as the sharing and amalgamation of data across various sectors (e.g., transport, energy, healthcare) to foster collaboration, enhance decision-making, and address smart city challenges (Lim & Maglio, 2018; Mukherjee, Gupta, et al., 2022; Van Der Hoogen et al., 2019). However, limited insight exists regarding the factors that promote collaboration among stakeholders in achieving CDS (Clement,

Manjon, & Crutzen, 2022). Stakeholder involvement is pivotal in CDS initiatives, necessitating a stakeholder classification model to delineate their functions within data sharing (Van Der Hoogen et al., 2019). The model underscores the importance of involving all stakeholders throughout the data sharing process to encourage collaboration. Nonetheless, a substantial knowledge gap persists in developing cross-sectoral smart city systems, stemming from the complex and often non-interoperable nature of solutions and systems within sectors like energy, transport, and governance (Tsampoulatidis et al., 2022).

The imperative for CDS becomes apparent when considering the shared goals of sustainability and efficiency in smart cities. Achieving CDS remains a challenge, but recent advances such as IoT, 5G, and big data analytics have begun to enable it (Bibri, 2018). These technologies facilitate data collection, analysis, and sharing, paving the way for improved decision-making and services (Mukherjee, Gupta, et al., 2022). Future developments may incorporate AI and machine learning for enhanced data analysis and decentralised data storage solutions like blockchain to bolster security and privacy in CDS (Dorri et al., 2017).

Smart cities can be regarded as complex systems encompassing various subsystems, including healthcare, transport, education, and energy (Javidroozi et al., 2019b). Interoperability among these sectors is vital for smart city success, necessitating data sharing. For example, transport data can be valuable for pollution control and traffic management, and energy data may complement it. However, sensitive data, such as personal health records, adds complexity to data sharing, requiring robust data governance frameworks (Asswad & Gómez, 2021).

Facilitating CDS in smart city development yields numerous benefits, including enhanced decision-making, urban planning, energy efficiency, and service improvement (Bakıcı et al., 2013; Paiho et al., 2022). It broadens access to diverse data types, from traffic to environmental data, fostering the development of innovative services and problem-solving. Effective data governance is crucial to address barriers like privacy and data protection (Abdalla et al., 2019). While some view technical challenges as primary obstacles, others emphasize organisational and cultural factors (Cui et al., 2018). Strategies like data encryption and secure data sharing protocols are proposed to fortify data security and privacy in smart cities.

According to Kieu et al. (2022), "Multi-sector collaboration is key to moving smart cities toward their goals of being smart and liveable." Therefore, three CDS technical solutions will be reviewed here to gain an understanding of the current state of knowledge towards achieving the interoperability of city sectors. The criteria used to review these three CDS technical solutions were based on their capacity to facilitate data sharing across multiple sectors within a smart city. These solutions were selected based on their relevance to the enhancement of cross-sectoral data sharing capabilities. The analysis focused on evaluating the proposed solutions' strengths and limitations concerning their applicability, scalability, security, adherence to data protection regulations, and their potential to enable effective CDS.

The first solution identified is a hybrid blockchain-based platform integrating different crucial sectors involved with the functioning of a smart city. They state that "existing blockchain-based smart city solutions are restricted to individual sectors only, and there is a lack of connectivity and composability among them" (Mukherjee, Sahoo, & Halder, 2022). Therefore, this solution provides the ability for CDS to be facilitated, with sectors being able to share data securely. The blockchain solution also considers data integrity, which is very important, making it secure and transparent for data sharing. However, there are concerns about data privacy and adherence to data protection regulations for blockchain technology. Yet, Kosba et al. (2016) have proposed privacy-maintaining techniques for blockchain solutions such as zero-knowledge proofs (Kosba et al., 2016). Therefore, if these privacy technologies were applied into the blockchain solution presented by Mukherjee, Sahoo, and Halder (2022), this could overcome the limitations of not adhering to the current DPL.

The second solution identified is by Anisetti et al. (2021), where they proposed an access control system that is supposed to solve data sharing and data protection within a big data environment. This solution provides a partial solution for CDS, as although it is mentioned that the system 'should' protect the privacy of sensitive data, it is not guaranteed. The Smog Analysis Service and Traffic Monitoring Service are based on the same data sources and ingestion pipeline, and the access control system enables both services to access the dataset with restrictions set on the sensitive data (Fig. 1). In this example, the traffic monitoring service can access sensitive number plates and CCTV, and smog analysis services are not able to, due to restrictions. However, unlike the blockchain solution presented, it is not clear if the system would be scalable or secure, which suggests it may not be applicable for CDS within smart city developments.

The third solution identified is by Pomp et al. (2021), proposing a semantic data marketplace for data sharing within a smart city (Fig. 2). The data marketplace makes it easier to find and access data which improves the effectiveness of CDS. However, it is unclear if the solution meets the requirements of DPL or if the marketplace is scalable to meet the needs of multiple sectors within smart city developments.

After investigation of three solutions, each solution has its strengths and limitations when it comes to facilitating CDS. The access control system proposed by (Anisetti et al., 2021) could be the most suitable method for CDS, but there are concerns about its scalability. Additionally, the blockchain solution by Mukherjee, Sahoo, and Halder (2022) could be a suitable method for facilitating CDS while adhering to DPL if it includes privacy techniques. Lastly, the data marketplace solution by Pomp et al. (2021) could also be a suitable method for CDS, although it is unclear if DPL requirements have been met. Therefore, further research into these solutions and other emerging technologies is needed to provide a comprehensive understanding of the most effective approaches for facilitating CDS whilst adhering to DPL.

2.4. Data sharing and data protection legislations

According to the contents explained above a multifaceted view of the challenges and intricacies surrounding data sharing, privacy, and regulatory frameworks, particularly within the context of developing smart cities and the use of emerging technologies is evident. It means, that navigating data sharing within smart cities necessitates a nuanced understanding of privacy, adaptive mechanisms for policy formulation, and a collaborative, flexible regulatory framework that fosters innovation while ensuring data protection. Chiara's (2021) observations regarding the conflation of privacy and data protection within the technical community underscore a prevalent challenge. While these concepts share common ground in safeguarding individual rights and autonomy, they possess distinct scopes and rationales. The European Union's Charter of Fundamental Rights distinctly separates the two, yet practical application and legal interpretations often intertwine these rights, complicating their enforcement.

In addition, Bargh et al. (2014) highlight the necessity of feedback mechanisms to dynamically adapt privacy policies, especially in situations where predefined policies might fall short due to contextual complexities. Their approach aligns with the need for adaptable frameworks that can respond to evolving privacy concerns in data sharing environments.

Moreover, Graef, Tombal, & de Streel, 2019 underscore the dichotomy of data sharing, emphasising the myriad benefits while acknowledging associated risks. The regulatory landscape needs to strike a delicate balance between promoting data sharing for innovation while safeguarding against privacy infringements and market distortions. Achieving this balance necessitates a nuanced approach, integrating horizontal legal instruments with sectoral rules while ensuring consistent interpretation and cooperation among regulatory bodies.

Hence, the overarching need emerges for clear delineations between privacy and data protection, dynamic mechanisms for addressing privacy concerns in evolving data sharing scenarios, and a robust

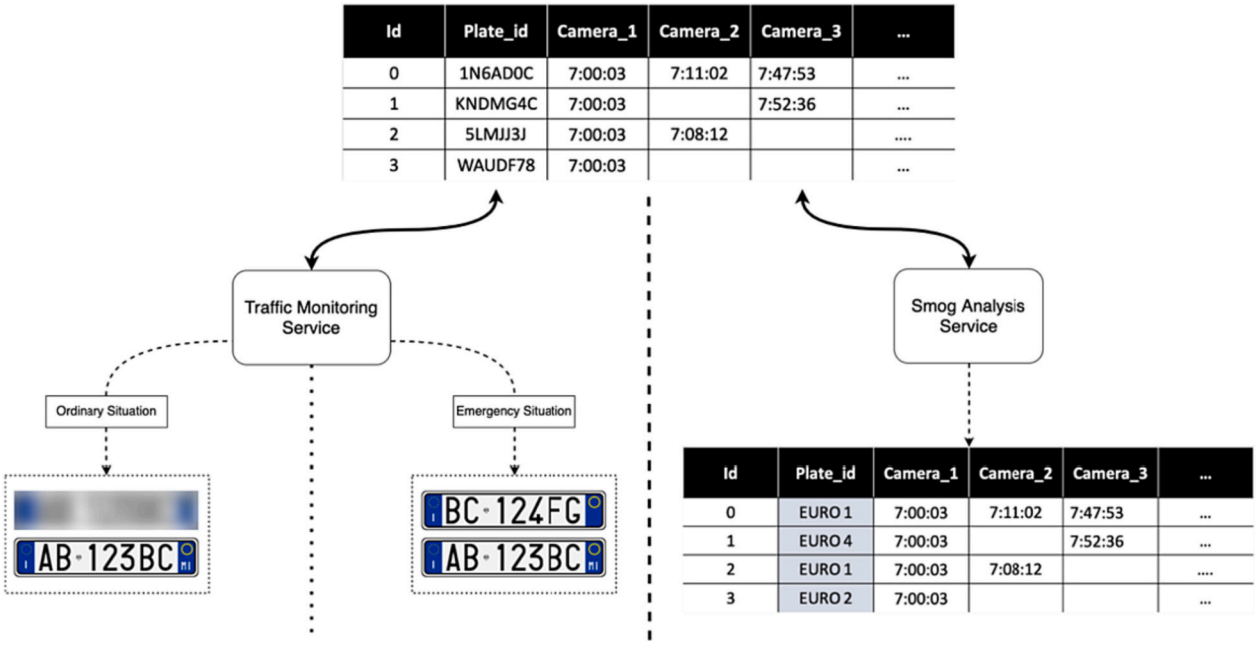


Fig. 1. Dynamic and scalable enforcement of access control policies for big data (Vandercruysse et al., 2020).

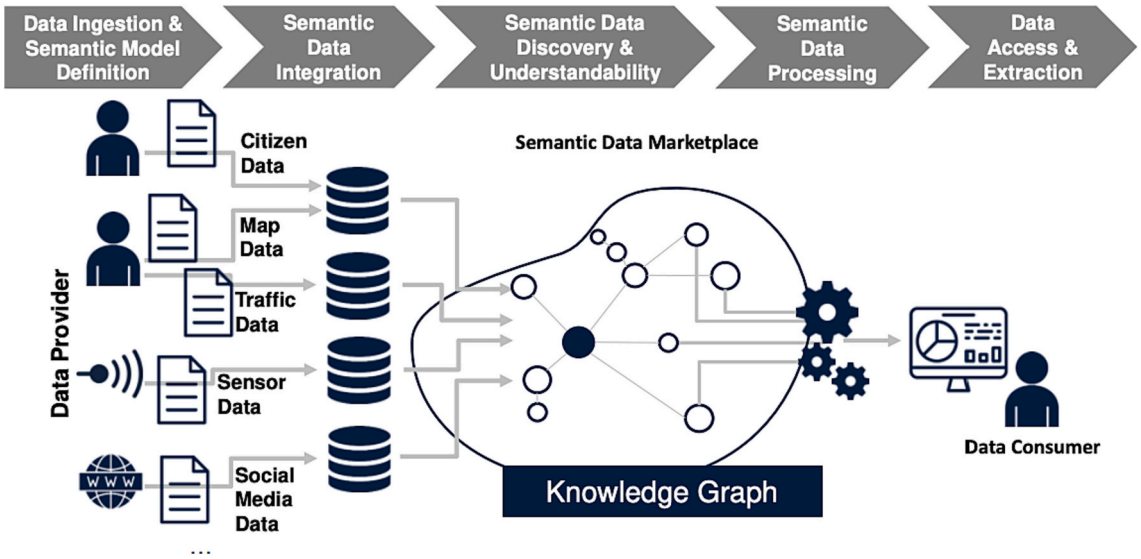


Fig. 2. A semantic data marketplace for easy data sharing within a smart city (Pomp et al., 2021).

regulatory framework that balances innovation with privacy safeguards. In the context of smart cities, where data sharing across sectors is pivotal for efficiency and innovation, the challenge magnifies. A comprehensive step-by-step guide and framework become imperative.

2.5. The impact of data protection legislations

The UK Data Protection Act and General Data Protection Regulation (GDPR) are two examples of legislation that have the purpose of protecting personal data in the context of smart cities. DPLs are necessary for the sharing of data to ensure the privacy of data subjects. These legislations require the protection of data from unauthorised access and processing to ensure that data is secure and accurate. DPL can enforce additional requirements on data controllers and storage systems, making it necessary for the use of techniques such as encryption, tight access control, and data protection impact assessments (Christofi et al., 2019). However, due to the sharing and use of personal data being an issue for smart city stakeholders, a central agent could be necessary to control and manage access to personal data required by other stakeholders. Furthermore, data governance policies would be required within CDS initiatives to prevent unwanted data sharing and accountability for the stakeholders involved (Denker, 2021).

Regarding the open data platforms (such as Greater London Authority Datastore), while open data principles promote broader access to city data, privacy regulations pose a significant barrier to making all potentially useful data fully open. As Lnenicka et al. (2022) highlight, transparency-by-design approaches to open data portals must be balanced with legal data protection frameworks that limit the sharing of potentially sensitive personal data. Anonymisation techniques can enable the sharing of some kinds of granular data related to mobility or energy consumption patterns while protecting individual identities. However, truly anonymous datasets are difficult to produce from many

real-world sources. Ultimately, open data platforms require carefully designed privacy preservation policies and governance processes to assess and manage the risks of making different datasets public. Technical controls like differential privacy, aggregation and data scrubbing can reduce re-identification risks to some degree. However substantial investments in both technological infrastructure and human-driven data review processes are necessary to facilitate cities in the disclosure of data while upholding the legislated safeguards concerning personal data. Achieving the right balance remains an ongoing challenge.

The application of data protection and privacy measures within smart cities confronts numerous challenges and inherent limitations. A primary hurdle involves the intricate management of data originating from diverse sources, compounded by potential conflicts arising from varying data protection legislations. Additionally, the process of anonymising data, while pivotal for preserving individual privacy, presents considerable implications for the effective functioning of smart city infrastructures. This complexity is particularly pronounced in sectors like healthcare, where the utilisation of personal data is integral, thereby posing significant hurdles to the seamless application of anonymisation practices (Gruschka et al., 2019). However, using data aggregation techniques can decrease the amount of personal data travelling throughout the smart network (Braun et al., 2018). Although considering the challenges of ensuring adherence to DPL, it is clear a balanced approach that ensures data protection and privacy while enabling effective cross-sectoral collaboration is necessary for the successful development and functioning of smart cities.

Subsequently, the impact of DPLs on CDS can be summarised as follows:

- Regulatory demands: DPLs necessitate strict security measures for shared personal data.
- Compliance requirements: Regulations impose added responsibilities on data systems and controllers.
- Centralised oversight: Potential need for a central authority to manage data access.
- Open data challenges: Balancing transparency and privacy within open data platforms.
- Technical safeguards: Measures like differential privacy to mitigate re-identification risks.
- Complexity and balance: Challenges in complying with varied data protection laws.

3. Research methodology

In response to the literature gap, this study investigates the following research question: “How can cross-sectoral data sharing be effectively accomplished in the context of smart city development while maintaining compliance with Data Protection Legislation (DPL)?” To address this research question, the study aims to construct a comprehensive conceptual framework, which functions as a practical guide for smart city initiatives, delineating the strategies and practices necessary to facilitate cross-sectoral data sharing while adhering to prevailing DPL. To address this aim, the following objectives are addressed in this research:

1. To explore the importance, requirements, challenges, and existing solutions for cross-sectoral data sharing in smart city developments;
2. To explore current data protection legislation and its impact on cross-sectoral data sharing within smart city developments;
3. To gain key insights into the concept of data sharing for smart city development, the challenges, steps, and success factors as the key components to guide developing a framework;
4. To develop a conceptual framework, outlining the steps required in achieving cross-sectoral data sharing, considering the challenges that must be addressed and the impacts of data protection legislation;
5. To perform validation of the developed framework, gaining insights into the framework's effectiveness, usefulness, and necessary improvements.

An inductive qualitative study allows for the finding of new insights and is suitable for complex problems with limited existing literature. Hence, this study employed this as an overarching approach to the study methodology.

3.1. Literature review methodology

The first two objectives are addressed through a literature review. Two academic databases were used; Scopus and ACM Digital Library, due to their unique strengths in providing comprehensive coverage within specific domains, offering a specialised focus on computing and information technology, and aligning closely with the subject areas pertinent to smart city research, such as computer science, urban planning, and technology integration within urban environments. This selection was based on the recognition that ACM Digital Library caters specifically to these niche areas, housing a wealth of scholarly resources and technical reports that may not be as extensively covered by broader databases like Web of Science. Moreover, ACM's repository emphasises publications related to smart city domains, ensuring a more targeted and in-depth exploration of literature within this specific research realm.

In the preliminary phase of this research, a detailed search strategy was employed to gather relevant literature essential for investigating the complexities surrounding smart cities, data sharing, data protection, and privacy measures. The search commenced with a carefully constructed query designed to capture a comprehensive spectrum of scholarly articles and technical reports pertinent to the subject matter, as follows:

(“data sharing” OR “data exchange” OR “data collaboration”) AND (“data protection” OR “privacy regulations” OR “legislation” OR “GDPR”) AND (“smart city” OR “smart cities” OR “urban development” OR “future cities”)

Utilising this search term, a robust initial collection of 195 papers was identified within the Scopus database. Concurrently, the exploration within the ACM Digital Library yielded an additional 129 papers, further enriching the pool of prospective literature. These figures represented the starting point of the research journey, indicating the expansive breadth of available literature within these databases. Subsequently, the duplications were reduced, resulting in the identification of 296 distinctive research papers. The identified 296 papers underwent a screening stage, when the titles, abstracts, and conclusions of the papers were examined to assess their relevance to the research. If needed for further clarification, we skimmed through the full texts of the papers. To narrow down further and ensure the relevance of searches, filters of source type were applied to ensure only relevant and useful research was found:

- Only journal articles, conference papers, and technical reports are included, instead of lecture notes; due to their robust review processes, which enhance quality and credibility. These sources offer completed analyses, undergo stringent peer or conference review, and generally follow standardised formats, ensuring authoritative and validated information for a rigorous academic literature review, while lecture notes, often lacking such review, may provide background but typically aren't core sources for substantiating a literature review (Haddaway et al., 2015);
- The article's publication date is no older than 2015, as it was important to maintain the relevance and accuracy of the research, especially considering smart city trends are rapidly developing. However, during the preliminary search, it was proved that most of the papers retrieved ranged from 2018 to 2022, which can be

justified by high research movements within the last 2–5 years within the fields of study;
- Subject areas of papers, comprising computer science, information technology and smart city subjects such as urban planning, transportation, and security;
- The language of papers is English to narrow down relevant results;
- The papers that could be accessible for free or via our institution contracts were included.

Results were also sorted to ensure that articles with high numbers of citations were prioritised, as this can often mean they are of high quality and relevance within research as they have been used the most. However, we also assessed the quality of the identified papers through an evaluation encompassing various facets, including the rigour of methodology employed, the calibre of writing, the publication venue, and the impact factor of the journals where the papers were published.

The PRISMA diagram as shown below (Fig. 3) was also utilised as an effective method to conduct a step-by-step assessment of screening papers based on the criteria applied (Moher et al., 2009).

At the beginning of the screening process, 296 papers were retrieved from ACM and Scopus. However, results of the screening and eligibility check process concluded, after all the criteria were applied with 37 papers to be included. This multi-stage screening process ensured a thorough evaluation of the available literature and contributed to the selection of studies that align with the objectives of our review.

Furthermore, to ensure the timeliness and comprehensiveness of our literature review, we conducted a re-verification process prior to the publication in 2024. This involved reassessing the literature landscape for any new publications beyond our initial review period. Despite identifying 25 relevant papers published in 2023 and 2024 through Scopus, a thorough examination revealed that none introduced novel perspectives that significantly impact the core findings or contributions of our research. Notably, only very few comprehensive papers were identified, mainly focusing on blockchain technology, and none of them were in the context of Smart city development. Those papers did not offer a comprehensive framework or guideline pertaining to data sharing within the context of DPL. This reaffirms the validity of our chosen timeframe and ensures that our study captures a relevant snapshot of the literature up to the point when our research commenced.

3.2. Interviews

Objective 3 was addressed through online semi-structured interviews with experts in the field of smart cities (e.g. smart city board members, strategy management for smart city initiatives, consulting roles, urban planning roles, geospatial data sharing roles and digital development roles) that provided the ability to reach a larger pool of experts internationally. We intended to capture a broad spectrum of expertise and perspectives through participants' diverse professional backgrounds across various cities, relevant industries, and roles. This range encompassed experiences in smart city development, including roles within city councils, as well as contributions within solution-provider companies dedicated to smart city initiatives. The aim was to solicit insights and perspectives from a diverse range of individuals possessing extensive experiences across varied geographical landscapes, relevant professional sectors, and capacities. To achieve this, an extensive outreach effort was undertaken, leveraging both personal networks and the professional platform LinkedIn, resulting in the identification of over 200 potential participants. Subsequently, outreach was conducted with all identified individuals, culminating in 20 respondents expressing interest in contributing to our research initiative.

Careful consideration was exercised in the selection process, meticulously evaluating the alignment of participants' profiles with the core focus of our study. This involved a stringent vetting process where 14

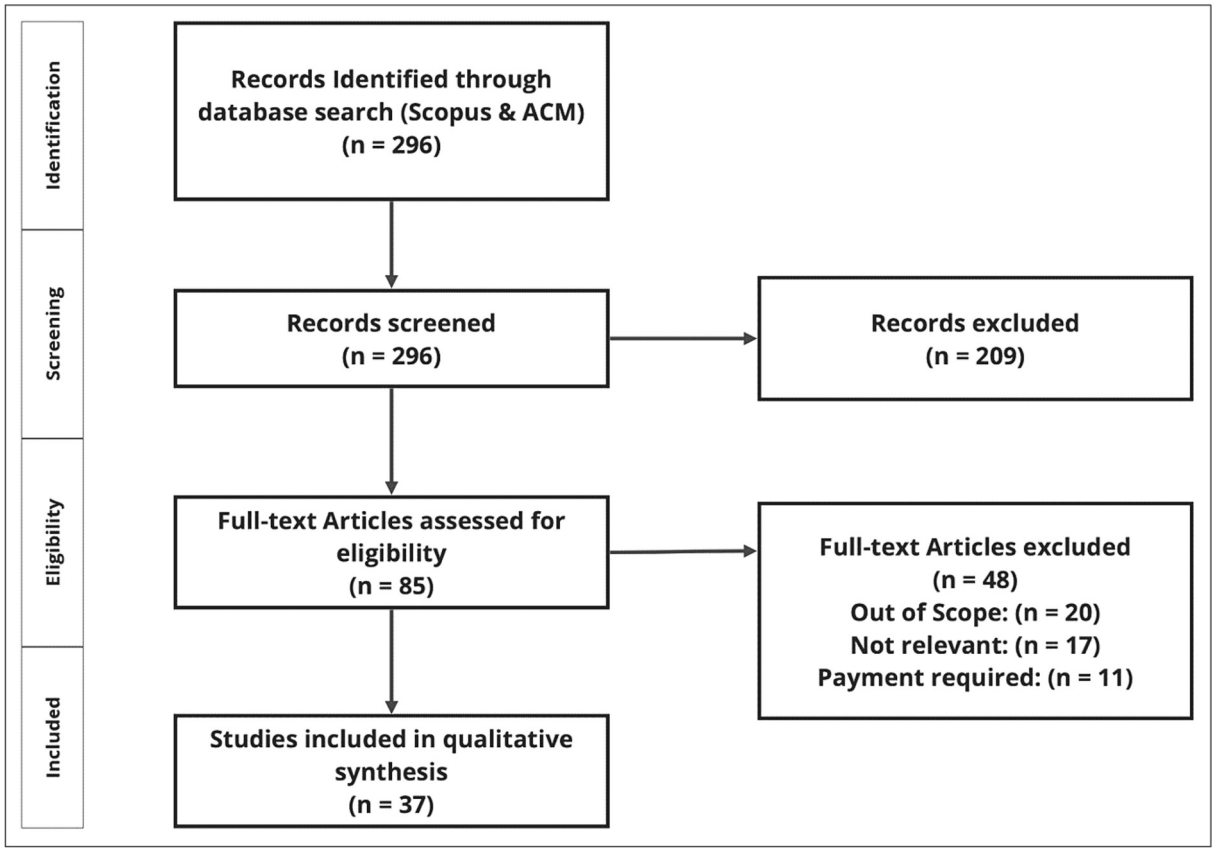


Fig. 3. PRISMA approach for the literature analysis of this study.

participants were ultimately chosen based on the relevance of their experiences and expertise within the domain of smart cities, particularly emphasising proficiency in data sharing paradigms and comprehension of data protection legislation. Tables 1 and 2 demonstrate the profile of the interviewees:

Spanning across a global spectrum, the interviewees have been actively engaged in smart city projects in various cosmopolitan centres such as Helsinki, Vienna, London, Stockholm, Bern, Riyadh, Dubai, Barcelona, Birmingham, and three Indian cities, including Delhi, Bangalore, and Chennai. Notably, a subset of the interviewed experts could be classified as stalwarts within the realm of smart cities, boasting significant expertise in spearheading data-driven solutions, serving on advisory panels, or orchestrating the management and execution of substantial smart city projects. This amalgamation of seasoned experiences, either derived from active involvement in smart city development or as contributors to solutions catering to smart city ecosystems, contributed multifaceted perspectives on critical facets pertinent to CDS in smart cities, particularly accentuating the intrinsic significance of robust data protection measures.

Conducting online interviews using Microsoft Teams or Zoom, as the interview platforms, depending on the interviewees' preference provided a flexible medium conducive to engaging professionals across diverse geographical locations and time zones. This accommodative nature facilitated seamless interactions and discussions with participants, irrespective of their disparate locations and temporal constraints. Furthermore, utilising qualitative methods enabled a thorough exploration of key concepts and variables necessary for the framework's development. This approach facilitated the acquisition of in-depth and nuanced data, enriching the investigation.

Participants were furnished with a detailed participant information sheet elucidating the project's objectives, the necessity for interviews, and related details. Additionally, a consent form was dispensed to participants well in advance of the interview date, mandating their explicit agreement before the scheduled interview. The protocol rigorously adhered to ethical standards and regulatory requisites concerning data acquisition and participant consent. The process for recording interviews was meticulously executed using the Microsoft Teams platform's recording and transcription feature. This facilitated the live transcription of interview proceedings, ensuring the accurate preservation of participants' responses. Furthermore, the recorded interviews were intended for subsequent analysis, allowing for comprehensive scrutiny and thematic exploration of the data captured during the sessions. The acquisition of consent constituted a fundamental step preceding the interviews, ensuring that participants were explicitly granted permission for the recording and potential analysis of their contributions. This stringent adherence to ethical procedures aimed to uphold the confidentiality, privacy, and informed participation of each interviewee throughout the research process.

The interview sessions were initially planned for a duration of 60 min each; however, the actual duration varied due to differing response rates among respondents, extending the sessions to a range between 60 and 90 min based on the pace at which participants addressed the questions. This duration was sufficient to explore into each participant's experiences and viewpoints. All questions posed during the interviews are listed in Appendix A.

Fourteen interviews were undertaken, and the results indicated that the saturation point, where no further novel insights were obtained, had been reached (Guest et al., 2006). The saturation point for this study was found after 13 interviews which covered all necessary aspects for the framework. A 14th interview then reinforced that no new perspectives were being found in the research of achieving CDS and adhering to DPL.

Table 1
Interviewees' profile.

Interviewee type	Number of interviewees
Director of smart city projects	5
Government advisors regarding smart sustainable cities	2
Smart city consultants	3
Solution providers for smart cities from the industry	4

Table 2
Interviewees' level of experience.

Experience level	Number of interviewees
More than 8 years	5
3–5 years	6
2–3 years	3

Data collected from the interviews were analysed thematically to identify patterns, trends, and themes, which will be then grouped into categories for the development of the framework and answering the research question (Braun & Clarke, 2006).

3.3. Data analysis methods

After the data from interviews had been collected and transcribed, expert participants' responses were analysed thematically. Braun and Clarke's (2006) six-phase approach was used to identify coding patterns, themes, and categories within the data collected from interviews. Hence, different perspectives and insights from 14 experts were categorised and compared to gain a full understanding of what needed to be applied into the framework. Analysis has allowed for data collected from the interviews to be put together, providing an in-depth understanding of the research question, the concepts, and variables to apply within the conceptual framework and their relationships. Key findings from the literature review were also analysed and compared with the data collected from experts to gain an even more rounded understanding of the framework, by comparing the different perspectives of researchers against the expert participants. For example, this included comparing experts' insights on applying blockchain as a solution to facilitate CDS with key findings for the suitability of blockchain as a solution highlighted within the literature review.

To ensure the reliability of the research, an identical copy of the data collected through transcripts was sent to a second coder to review, ensuring consistency and that there were no biases in the data analysis performed (Maxwell, 2012). Additionally, it was essential that detailed records of the coding process were kept and known guidelines for completing thematic analysis were followed.

Conducting a thematic analysis involved the following steps:

Familiarisation with the data by reading transcripts previously prepared and gaining an understanding of the insights, key information, and perspectives from expert participants;

- The coding process encompassed the meticulous categorisation of excerpts from the transcripts into distinct codes, delineating various aspects such as data standardisation, data quality, stakeholder alignment, and other pertinent themes. This process was facilitated by utilising NVivo, a sophisticated qualitative data analysis software. NVivo provided a structured platform for organising, sorting, and labelling the data excerpts, ensuring systematic categorisation and enabling comprehensive analysis of the multifaceted aspects inherent in the transcripts;
- Themes within the data collected were identified by looking for recurring topics or patterns from multiple participants, such as the requirement of data standardisation being mentioned consistently;
- Themes were then defined and reviewed to check for any inconsistencies or repeated themes; the process of defining themes involved two rounds of careful examination by two authors to extract and categorise recurring patterns from the data. Following this initial phase, intercoder reliability was conducted to ensure robustness and consistency in theme identification. To achieve this, an external researcher, a colleague not directly involved in the

initial coding, was engaged. This step aimed to mitigate the possibility of overlooking or inadvertently merging themes during the coding process. After defining themes, a rigorous review ensued to identify any inconsistencies or redundancies within the themes, ensuring the accuracy and reliability of the final thematic framework;
- The analysis delved into establishing relationships and connections among the identified themes, forming a comprehensive understanding of their interplay within the dataset. As this elaborate web of connections unfolded, the analysed data served as the foundational material for crafting a robust conceptual framework. This framework was designed as a guideline, delineating the strategic steps and pivotal components necessary to effectively realise Cross-Sectoral Data Sharing (CDS) within the context of smart city development.

3.4. Validation strategy

After the development of the framework and addressing objective 4, the study validates the framework through expert review interviews (Newcomer et al., 2015), to assess the effectiveness of the framework and identify areas for improvement. The data collected from the validation strategy were analysed and final adjustments were made to ensure the framework was complete, addressing the research problem and gap within the literature.

Interviews were conducted with six smart city experts who were previously interviewed during data collection, by conducting a second MS Teams interview. During the online validation interview, participants were presented with the developed framework and were asked four key questions about the framework's usefulness, applicability, implement-ability, and areas for improvement within the smart city field based on the following criteria:

- The developed framework was assessed against its usefulness and ability to be applied in a real-world smart city context. Smart city experts were asked what changes may need to be taken into consideration for the framework to be applied in different contexts.
- The developed framework was assessed against its ability to adhere to the current DPL. Smart city experts were asked if they believe the framework addresses adherence to current data protection laws and if not, what changes needed to be made.
- The developed framework was assessed on its ability to address and outline how cross-sector data sharing can be achieved, including the challenges involved and potential strategies. Smart city experts were asked if they believe the framework successfully outlines how CDS can be achieved and if not, what changes needed to be made.
- The developed framework's design was assessed with any suggestions for improvements. Smart city experts were asked to provide overall feedback on the developed diagram and if any changes needed to be made.

Supplementary inquiries were also posed to obtain detailed elaborations and precise elucidations of their perspectives. A live transcript of the interviews was recorded to ensure all responses could be later analysed. Validation interviews followed very similar methods to the data collection interviews to ensure consistency across all interviews within the research project. After validation of the framework and thematic analysis have been conducted, any insights and feedback received from participants were incorporated into the previously developed framework, producing the finished version of the framework. Any improvements or adjustments that could not be made due to the limited timeframe were also suggested for further work. The participant sampling criteria for interviews was purposive, specifically expert sampling. The sampling method of participants was conducted by re-contacting previous experts which provided highly valuable insights into the development of the framework. This was to ensure that experts validating the framework were the most knowledgeable in key areas which may need improvements, leading to high-quality feedback.

4. Research findings

After conducting a thematic analysis of interviews, eight themes were identified. These themes are directly related to the research question. Additionally, these findings have reinforced objectives 1–2 of this research, which were already addressed partially by conducting a literature review. The following part of this results section provides an exploration of the eight themes developed and the sub-themes found.

4.1. Key themes – Interview findings

This section includes the eight key themes identified, covering the findings and interviewee quotes from each theme. The themes collectively provide a clear understanding of the key aspects involved with CDS within a smart city:

1. Opportunities: Identifying the opportunities of CDS provides an understanding of the potential benefits and value creation for all stakeholders (sectors). Highlighting these opportunities ensures collaboration and can inform decisions.
   - Findings: Thematic analysis has revealed opportunities for informed decision making, enhanced collaboration, operational efficiency, and value creation. Key outcomes include improved data access, better insights, aligned goals, streamlined processes, and advanced solutions.
   - Interviewee key quotes:
     - “Cross-sectoral data sharing can help decision makers, make more informed decisions” (Interviewee 6).
     - “Whilst monetising data is a benefit, the insights pulled from the data are what actually creates value” (Interviewee 2).
     - “One of the benefits of sharing data across sectors is the wealth of knowledge and information that can be gained from merging different data sets” (Interviewee 8).
2. Challenges & Addressing Challenges: The understanding of challenges within CDS is crucial for developing solutions and addressing potential barriers to CDS. Therefore, providing strategies for overcoming these challenges within the framework ensures that it is comprehensive.
   - Findings: Thematic analysis has revealed challenges of data quality, data standardisation, interoperability, privacy concerns, data protection compliance barriers, and stakeholder collaboration factors, demonstrating value and data security. Addressing these challenges can require investment in technology, strong infrastructure, constant communication, data protection measures, and displaying successful examples and value propositions.
   - Interviewee key quotes:
     - “Data quality, standardisation of data, and the lack of a clear data owner are key challenges, especially for instant data sharing” (interviewee 3)
     - Interviewee 7 had a different viewpoint: “The main challenges are the storage of large amounts of data, data security and the prediction of potential issues” and suggested “consistent classification of data across sectors”.
3. Technologies/Solutions: Investigating the different technologies and solutions that provide secure and efficient data sharing can help inform the framework's recommendations and steps, supporting the successful execution of CDS initiatives.
   - Findings: Thematic analysis has revealed the following technologies for CDS: AI-enabled data cleaning, data cooperatives and trusts, dedicated sharing platforms, blockchain technology, data conversion software, open-source solutions, and integration layers. These technologies overall ensure data compatibility and interoperability of different data sources.

- Interviewee Key Quotes:
  - The majority of interviewees (2, 4, 7, 8, 9, 14) mentioned blockchain as a technology solution to facilitate CDS, such as Interviewee 9 said: "Technology solutions such as encryption and anonymisation can be used on a small scale for sharing data, but for industrial-level data sharing, blockchain technology is the key solution for encryption and maintaining privacy."
  - "The use of blockchain for ensuring data integrity in cross-sector data sharing is still in its exploratory stages." (Interviewee-2), suggesting that there are mixed opinions on the use of blockchain technology.
  - "AI-powered data cleaning helps smart cities share data between different sectors securely and according to regulations by improving data quality, protecting privacy, following rules, making data handling easier, and reducing mistakes." (Interviewee-13), suggesting that AI-enabled data cleaning holds a crucial role in enabling secure and compliant cross-sector data sharing within smart cities,

4. Stakeholder Collaboration/Alignment: The collaboration and alignment of stakeholder interests have been found to be essential to the success of CDS projects. Therefore, including strategies for collaboration and alignment into the framework helps ensure that stakeholders involved with CDS have their needs met and goals are achieved.
   - Findings: Thematic analysis has revealed the importance of recognising stakeholder interests, incentivising data sharing, constant communication, trust building, data sharing agreements (such as an MOU), engagement of all stakeholders, standardisation between sectors, data governance of the data shared between sectors, plans for long-term data maintenance, understanding all stakeholder needs, creating central governing bodies, and establishing a shared purpose. These factors create effective collaboration, alignment, and trust among stakeholders.
   - Interviewee Key Quotes:
     - "To ensure the data sharing is aligned with the needs of all stakeholders, it is important to show the positive value of sharing the data and the results of other projects as a positive example." (Interviewee-1);
     - Interviewee 2 had a similar view on stakeholder alignment stating the need to "Clearly define the purpose and scope of data sharing by outlining the specific data being shared, the intended use of the data and the expected outcomes from sharing."
     - Interviewee 5 provided another perspective on stakeholder alignment: "To align data-sharing efforts with the needs of all stakeholders, it is essential to understand their individual requirements, expectations and concerns."
5. Risks/Threats & Mitigating Them: Knowing the potential risks and threats associated with CDS is vital for being able to put measures in place to mitigate them. Therefore, including risk mitigation strategies within the framework supports the success of data-sharing initiatives while ensuring data protection, privacy, and compliance.
   - Findings: Thematic analysis has revealed the following risks in CDS: privacy concerns, security risks, data misuse, and data sharing costs. Mitigating these risks can involve measures such as anonymisation, encryption, secure storage, access controls, cybersecurity measures, ethical decision-making boards, clear data sharing agreements, data quality checks and measures and validation processes.
   - Interviewee Key Quotes:
     - Several interviewees (1, 2, 4, 5, 7, 8, 11, 12) mentioned data privacy as a key risk to CDS.
     - Interviewees 1 and 5 focused on the risk of "Aggregating data sets from different sectors can increase security risks" suggesting merging data sets should be carefully considered, especially if sensitive data is involved.
     - However, interviewee 7 states multiple risks "The potential risks of sharing data across different sectors in a smart city include security breaches, unauthorised access to sensitive information and data privacy violations" They also suggest that "These risks can be mitigated by developing a predictive framework that ensures data security".
6. Requirements: Understanding the requirements for CDS can be used for building the framework's guidelines and recommendations. By addressing these requirements this ensures that the framework is comprehensive and meets the needs of stakeholders sharing data across sectors.
   - Findings: Thematic analysis has revealed requirements including data standardisation, data minimisation, data accountability, data access control, data privacy protection, data protection compliance, the value being created from sharing, stakeholders and understanding of value and benefits for sharing. These factors help to ensure continuous data sharing, security, privacy, and effective collaboration for improved decision-making.
   - Interviewee Key Quotes:
     - "The key requirements for cross-sector data sharing in a smart city are standardisation of data formats, data minimisation, purpose limitation, data accountability and access policies" (interviewee 2), whilst also mentioning adhering to DPL is vital.
     - "To achieve successful cross-sector data sharing, it is essential to have a well-defined problem that the data sharing aims to solve and experienced team members" and that "Applying strict data standards is crucial for ensuring compliance with current data protection legislation" (interviewee 6).
7. Factors of the data being shared: Analysing the factors related to the data being shared is crucial for developing strategies that ensure effective data sharing for different types of data, such as applying measures for sharing highly sensitive data or data that could be triangulated.
   - Findings: Thematic analysis revealed factors such as understanding the data's sensitivity, assessing the benefits of sharing the data and ensuring technical compatibility of the data. Considering these factors and addressing the potential risks with sensitive data ensures effective CDS that meets the needs of each sector.
   - Interviewee Key Quotes:
     - The majority of interviewees (1, 3, 4, 5, 6, 7, 9, 10, 13) mentioned the importance of considering the sensitivity and privacy concerns of the data being shared. For instance, interviewee 9 stated, "For example, healthcare data contains sensitive information on patients personal data, demographics, and location, so there need to be strong justifications for using the data and effective governance of the data to ensure data security".
     - Interviewee 7 contributes to this point by suggesting "When determining what data can be shared across different sectors, it is important to consider the classification of the data, the different layers of data, and how the sectors are classified" and that "Not all data should be shared or exchanged between sectors and it is important to classify data into multiple layers".
8. Success Factors: identifying the key success factors for CDS helps inform the framework's recommendations and provides elements or aspects that need to be covered when applying CDS for successful initiatives/projects.
   - Findings: Thematic analysis has revealed success factors including a clear purpose, strong governance framework, appropriate technology, stakeholder involvement, transparent communication, effective technical solutions and having a clear problem definition before sharing data.
   - Interviewee key quotes:
     - "Having a clear understanding of the data involved, its quality, and its limitations is essential" and "Understanding the sources of the data and how it is collected and stored" (Interviewee 6).

- On the other hand, interviewee 9 stated some of the general success factors are (summarised) "having a clear purpose and desired outcome, a governance framework that ensures data protection and privacy, the use of data standards, agreement of all stakeholders, and transparency with the public."

4.2. Key themes & connections diagram

The diagram (Fig. 4) visually showcases the key themes identified in the thematic analysis and their connections. Rectangles represent themes, while coloured circles signify factors associated with each theme. For instance, within the challenge's theme, a green factor represents data privacy and security concerns, which are interconnected with stakeholder collaboration, requirements, risks/threats, and factors related to shared data. This underlines its significance across all aspects of CDS. Each factor in the diagram is depicted in green, red, purple, or blue if it is linked to other themes; white signifies no connection, illustrating the essential relationships between each theme. It should also be noted that the colours attributed to factors in the diagram denote distinct factors associated with multiple themes, emphasising their interconnection rather than indicating varying severity or frequency of relationships. Further discussions on the eight themes in the diagram and their connections follow in subsequent sections.

The relationships observed in the diagram between requirements, challenges, success factors, and factors related to shared data demonstrate the interconnected nature of different facets within CDS. For instance, requirements often stem from challenges encountered during data sharing, while meeting these requirements and applying strategies to address challenges can lead to the identification of success factors. These interconnections highlight the complexity and interdependence within CDS (Fig. 5).

Below are some of the significant findings in detail:

1. Data Quality is a recurring topic across several themes of requirements, challenges, and the data being shared factors. This could be because ensuring and maintaining high data quality is essential for effective CDS, as it establishes a foundation of reliability and accuracy in shared datasets, enhancing trust among sectors and streamlining collaboration. High data quality minimises errors, inconsistencies, and ambiguities, fostering seamless interoperability and enabling more meaningful and effective data exchanges among various sectors involved. In addition, it is essential to apply standardised practices such as data governance frameworks, quality assurance protocols, and validation procedures across all involved sectors. Data governance defines the policies and procedures guiding data management, ensuring compliance, security, and integrity. Quality assurance encompasses systematic evaluations and monitoring to confirm data accuracy, consistency, and completeness. Validation processes involve verifying data accuracy, relevance, and compliance with predefined standards, assuring data reliability and suitability for intended use within CDS initiatives. These practices collectively establish a structured approach to maintain high-quality data across diverse sectors engaged in data sharing.
2. Blockchain Technology has been mentioned by eight interviewees, especially in the technologies and solutions theme as the most suggested technology to facilitate CDS. Blockchain is very effective in ensuring data protection, suggesting its importance as a solution for addressing privacy, security, and data protection compliance concerns. Additionally, as blockchain is decentralised and secure this makes it a viable option for CDS.
3. Data Standardisation is a frequent sub-theme, found in requirements, challenges, and success factors tables. It is clear from the thematic analysis that having a consistent data format and data structures is essential for interoperability across sectors.
4. Stakeholder Communication was identified as a vital aspect in the requirements, data being shared factors, and success factors tables. Effective communication between different sectors can help the stakeholders involved. For example, constant communication can allow for them to better understand the different complex data sets between them, and ensure that the data is usable, and data is valuable for all stakeholders involved. Additionally, by prioritising communication, they can more effectively address potential challenges and simplify CDS processes.
5. Data Protection Legislation Adherence and Data Governance were recurring sub-themes in the requirements, challenges, and success factors tables. This could be because ensuring that CDS adheres to current DPL has been found to be essential to avoid legal penalties and to ensure data privacy and data accountability.

4.3. Thematic pillars for the CDS framework

The identified themes, derived from interview insights, underpin a comprehensive framework. Eight tables were formulated to articulate these essential themes, enriching our understanding and aiding comprehension of CDS and data protection compliance in smart city development. The tables were curated following a thematic analysis, guiding the selection of prominent factors for inclusion. Only the most frequently cited and significant factors were chosen, emphasising critical aspects like enhanced decision-making, and bolstering the opportunities of CDS. To ensure precision, factors mentioned by at least two interviewees were incorporated. For example, the technologies/solutions table encompasses three factors crucial to the CDS landscape. While these tables are not part of the primary framework, they serve as valuable supplementary resources, offering sector-specific insights. For instance, the technologies/solutions table aids in selecting suitable technology for data-sharing initiatives while aligning with the overarching framework's objectives (Tables 3–10).

Although many more were discussed by interviewees such as AI-enabled data cleaning or data conversion software, this table represents only the most significant solutions towards achieving CDS mentioned by multiple interviewees.

5. Research discussion: developing a framework for data sharing in developing smart cities

In addressing the challenges presented within achieving CDS, the final output of this research study, being the conceptual framework, has been developed from the key concepts, variables and factors found within the results tables. The CDS guideline diagram presents the steps required to achieve successful CDS initiatives whilst considering the important impacts and adherence to current DPL. The framework developed from these key findings addresses objective 4 of this research by serving as a complete guideline for sectors, helping them overcome barriers, manage risks, and achieve CDS successfully. This in turn addresses the identified research gap and question.

5.1. Findings influencing framework development

The results tables created from interviews have had a significant influence on the development of the conceptual framework, playing a key role in ensuring the guideline diagram is comprehensive and achieves the research aim. Each step to be included in the development of the framework was carefully considered and primarily derived from the results tables. Only the most significant factors from the results tables have been applied to ensure the guideline represents the key steps and actions for CDS. Each step developed within the framework can be justified by the insights of interviewees as discussed below:

- **Step 1: Strategic Planning and Stakeholder Alignment**: This step can be backed up by findings, as within the success factors table, having a desired outcome, clear problem definition, and effective stakeholder alignment was suggested by interviewees 6 and 9. In

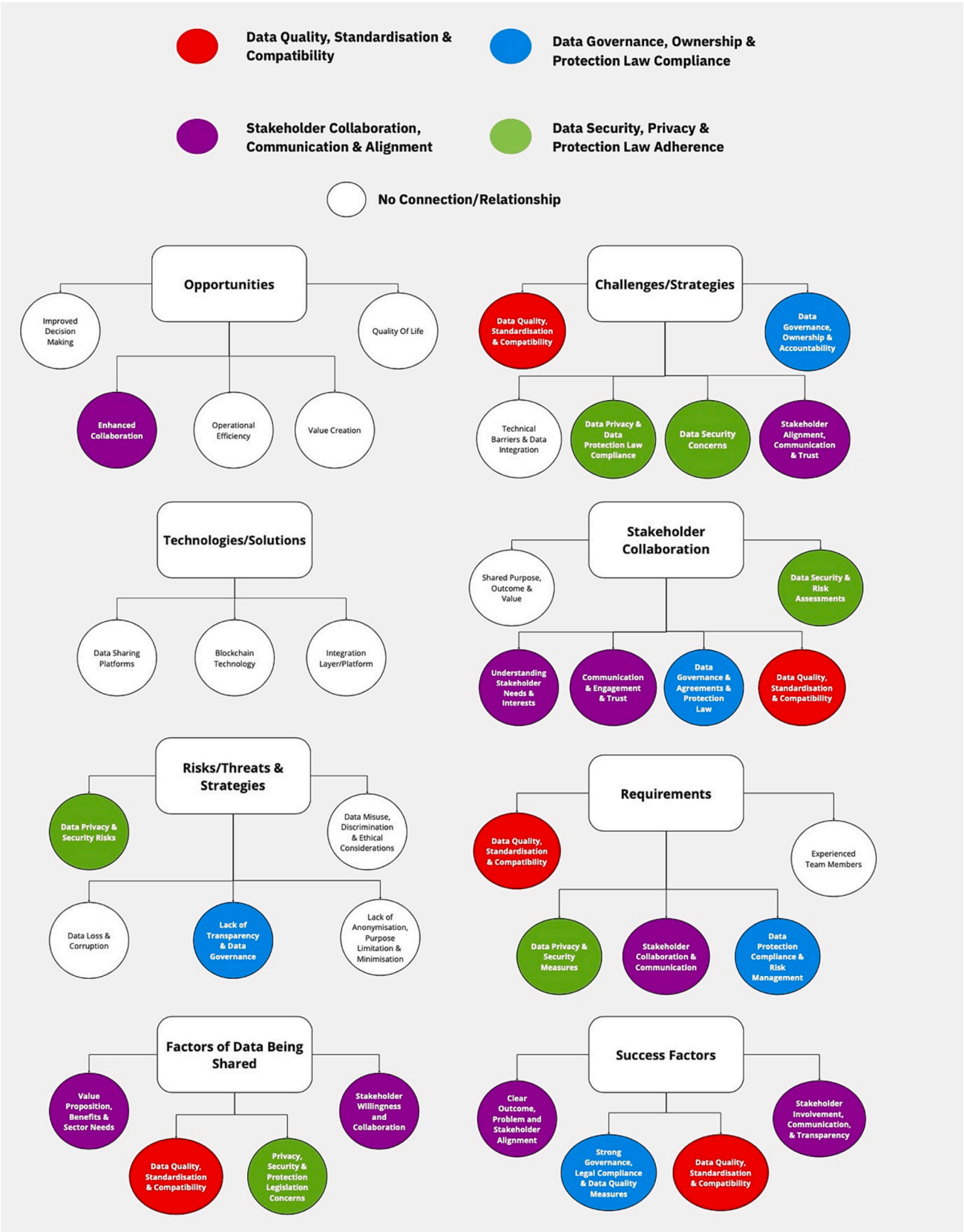


Fig. 4. Cross-sectoral data sharing: Key themes and connections.

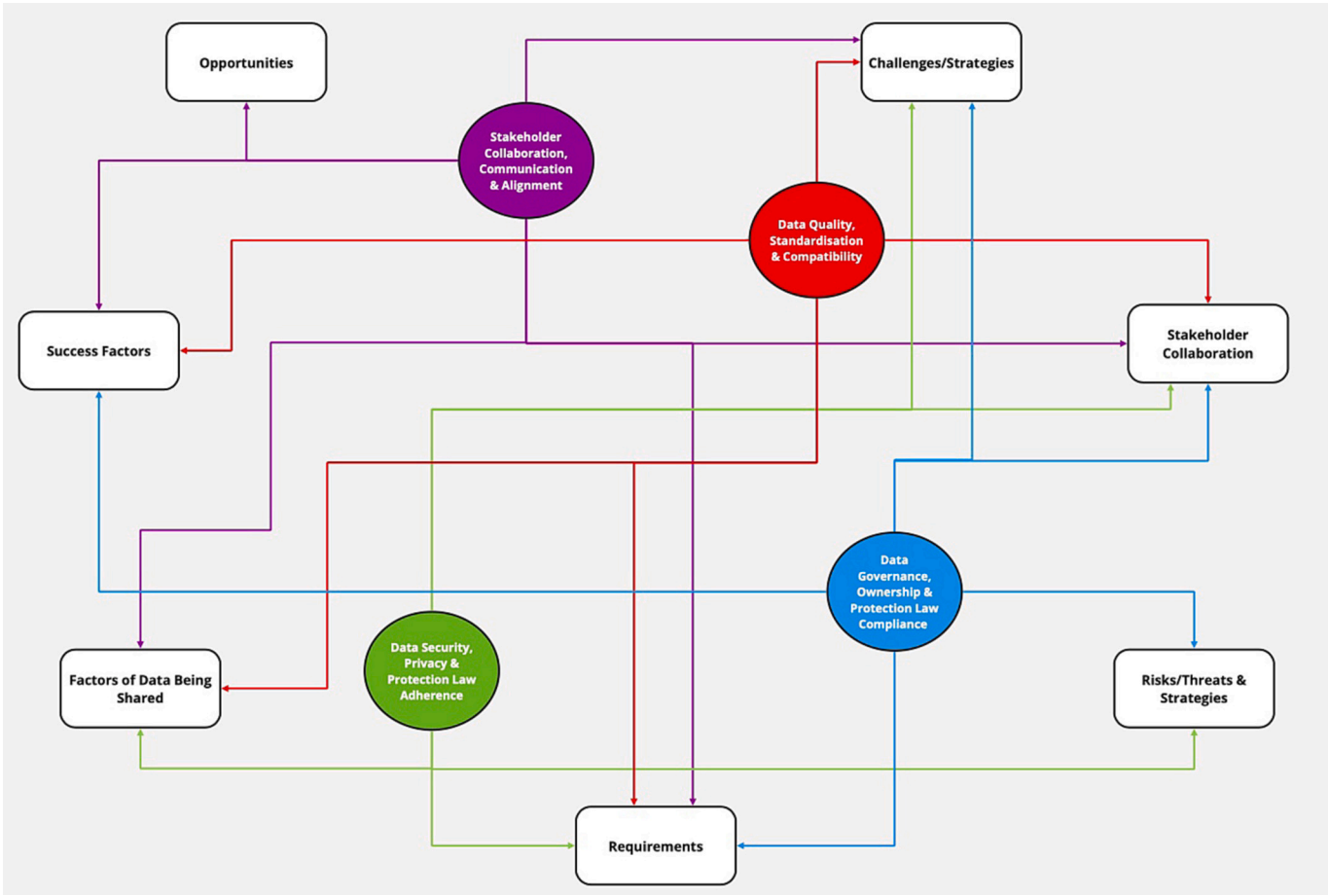


Fig. 5. Cross-sectoral data sharing: Key themes and the relationships between them.

addition to this, stakeholder involvement, clear communication, and transparency were also mentioned as key success factors by 3 of the interviewees. Further to this, stakeholder collaboration and communication is a key requirement backed up by interviewees 1, 5, 6 and 9. Lastly, a results table was developed for stakeholder collaboration factors due to the increasing importance of having effective stakeholder alignment, suggesting its importance.

- **Step 2: Data Inventory, Standards, Compatibility and Quality Management**: This step can be backed up by findings, as 70 % of interviewees suggested data quality, data standards and interoperability are challenges that must be addressed in achieving CDS. Additionally, data integration was also suggested as a challenge by 70 % of interviewees. Further to this, data management and standardisation were found to be a key requirement by interviewees 2, 3 and 4. Lastly, having appropriate data standards and data quality measures was found to be a key success factor by interviewees 2, 5, 6 and 9.
- **Step 3: Data Governance, Privacy and Security Measures**: This step can be backed up by findings, as data governance, data ownership and accountability were identified as a challenge by 50 % of interviewees. Additionally, data security and privacy concerns were also highlighted as key challenges in achieving CDS. Linking back to step 1, data sharing governance is a significant stakeholder collaboration factor and was mentioned by 50 % of interviewees. Additionally, data security and privacy measures were mentioned by interviewees 2, 4 and 9 as key requirements for CDS. Lastly, not having transparency and effective governance has been identified as significant risk a by interviewees 2, 3 and 8.
- **Step 4: Compliance with Data Protection Legislations**: This step can be backed up by findings, as data protection compliance was identified as a key requirement by 4 interviewees. In addition to this adherence to DPL was also mentioned in almost every table, such as being a factor for data being shared, a success factor, a challenge, a risk, and a factor in the stakeholder collaboration table. This suggests it is an essential consideration for CDS success.
- **Step 5: Data Sharing Agreements and Ethical Considerations**: This can be backed up by findings, as data misuse, discrimination and ethical concerns have been identified as a risk within CDS. In addition to this, data-sharing agreements are one of the requirements identified by interviewees 2, 4, 6 and 9 for effective stakeholder collaboration. This suggests that having effective data-sharing agreements and considering ethical considerations are a must for CDS.
- **Step 6: Sharing Infrastructure and Technology Solutions**: This can be backed up by findings, as the technologies and solutions table has highlighted the importance of using appropriate and secure solutions for CDS. In addition to this, risks of data loss, corruption and storage have been stressed by interviewees which links back to data infrastructure. Further to this, within the factor of data privacy and security measures, the importance of secure shared data infrastructures was highlighted by interviewees 2, 4 and 9.
- **Step 7: Monitoring, Evaluation, and Communication**: This step has not been backed up by the results tables. However, during interviews 10 participants stressed the importance that CDS is a continuous operation, and that measures must be in place such as monitoring and evaluating the successes of the sharing initiative, alongside constant communication.

Table 3
The 5 key opportunities for achieving CDS within smart city developments.

Opportunity category	Key benefits and outcomes	Derived from interviewees
Informed decision making	1. Improved data access, providing more accurate and widespread data-driven decisions. 2. Data merging data sets of different sectors leads to higher value insights, extensive knowledge, and further understanding of complex issues. 3. Supports long-term planning within sectors within the city	Interviewees 1, 2, 3, 5, 6, 9, 10, 14
Enhanced collaboration	1. Facilitates the alignment of goals between different sectors. 2. Encourages data sharing and collaboration between public and private organisations and government. 3. Improved functioning of the city through collaboration between various city sector departments and open data linking processes. 4. Communicating between sectors and understanding their different workflows has been found to be crucial for successful smart city initiatives.	Interviewees 1, 5, 6, 9, 11
Operational efficiency	1. Significant time and cost savings by enabling the reusing of data, cutting out repetitive tasks and reducing the need for redundant data collection across different sectors. 2. Develops holistic approaches to city management, addressing complex issues like climate change and resource allocation through data sharing and mapping decisions. 3. Increases overall system (city) efficiency by connecting all sectors and optimising resources, streamlining processes, and providing efficient urban planning.	Interviewees 1, 3, 4, 5, 6, 12
Value creation	1. Insights gained from shared data can lead to better decision-making and higher-impact outcomes. 2. Unified goals and increased collaboration can often result in higher overall value creation. 3. Supports solutions to complex urban challenges and the development of sustainable cities.	Interviewees 1, 5, 6, 9, 13
Citizen benefits and quality of life	1. Improved quality of life through improved solutions for emergencies, and optimising various city issues e.g., high energy usage. 2. Responsive city solutions by leveraging cross-sectoral data sharing e.g., smart traffic management.	Interviewees 2, 3, 6, 9, 11, 13

5.2. A structured approach to developing a framework for CDS adhering to DPLs

To develop the framework, a structured approach was used to ensure that the final version of the design covers all the key components identified through data collection interviews. The development of the framework involved the following key considerations:

1. **The Design approach** chosen for the diagram was suggested by interviewee (1) stating that a "Ladder" diagram was a viable option to effectively provide a guideline to achieving CDS, adapted slightly to form the "Steps" of CDS. This approach effectively represents the various components/factors involved with CDS and allows for a clear visualisation of the steps sectors must take.
2. **The Design tools** used to develop the diagram were Miro and Canva. Miro was used to develop ideas and first attempts, an example of this is Version 1, discussed in the following part of this section. Once all the steps and the actions within the steps had been finalised and categorised, the initial Miro designs were moved to Canva which provided the ability to develop a more professional design of the diagram.
3. **Identifying** the steps to be included within the diagram was supported by results tables. Ensuring the most important components/actions were included within the diagram to successfully represent the steps in achieving CDS. Once identified, the key factors identified from the thematic analysis were transformed into steps, such as stakeholder alignment, with some of the key factors included as actions within steps, such as defining the outcome of the data sharing, part of the strategic planning step.
4. **Incorporating the key components** previously identified into the guideline diagram involved careful consideration to ensure a well-structured representation of the framework. Each component was organised based on its relevance and corresponding order in achieving CDS. Afterwards, categorisation and combining of key factors and steps were performed to provide comprehensive but clear steps. This was especially important for the development of Version 2 and the changes made to create the final version of the diagram, as steps were reduced from 14 steps to 7 steps through strategic categorisation, to provide a clearer guideline. Additionally, for the final version of the framework, connections derived from interview results were used to represent the different connections between each step, symbolised with coloured circles.
5. **Multiple iterations** of the diagram were to refine the diagram and ensure its design accurately represented the data found through interviews with experts and provided an effective guideline for sectors.

The next step in the development of the framework involves performing validation interviews with return expert participants to provide rounded feedback and improvements based on a set of criteria and questions. The following part of this section involves the changes made through different versions of the diagram and the final developed version of the framework before validation.

Version 1 of the diagram was created on Miro and served as the initial idea/design for the step diagram (Fig. 6).

5.2.1. Key takeaways

- The foundational design required further development to enhance comprehensiveness.
- The diagram lacked detailed explanations or step-by-step actions for execution, limiting its usefulness as a guideline for sectors to achieve CDS.
- The diagram lacked a structured approach concerning the order and categorisation of the steps, which is crucial for effective guidance.

Version 2 of the framework was created on Canva, and improved based on the design of Version 1 (Fig. 7).

5.2.1.1. Changes made.

- The diagram's visual design was enhanced by incorporating step-specific icons, a colour scheme, and distinct block shapes representing each step. These modifications aimed to create a more engaging and user-friendly representation, ensuring greater clarity and ease of understanding.

Table 4
Some of the key challenges and strategies to overcome challenges in achieving CDS within smart city developments.

Challenge category	Challenges	Strategies	Derived from interviewees
Data quality, standardisation, and interoperability	1. Incomplete and poor-quality data 2. Lack of standardisation with different data formats, structures, or systems preventing cross-sectoral data sharing 3. Data stored and locked in proprietary systems	1. Apply measures for improving data quality within sectors. 2. Creating standards for data formats and systems 3. Having dedicated data owners and ensuring data interoperability across sectors	Interviewees 1, 2, 3, 4, 5, 7, 8, 14
Technical barriers and data integration	1. Legacy systems and lack of APIs 2. Merging and preparing data for sharing. 3. Complex data integration processes due to different data systems and formats	1. Adopting flexible and open systems and Developing APIs 2. Investing in effective data sharing technologies and infrastructure to support data integration. 3. Streamlining data integration processes for more efficient data sharing between sectors	Interviewees 1, 2, 4, 5, 6, 7, 8, 12
Trust, data privacy concerns and compliance with data protection laws	1. Trust issues between stakeholders 2. Data privacy and security concerns 3. Compliance with data protection and privacy laws across sectors 4. Complex, outdated and restrictive data protection legislation.	1. Transparent communication between stakeholders creates a culture of collaboration. 2. Comprehensive data policies 3. Establishing clear protocols for data minimisation, purpose limitation and data anonymisation. 4. Establishing clear data sharing agreements.	Interviewees 1, 2, 4, 6, 9, 11
Organisational and cultural barriers	1. Convincing stakeholders of the benefits of data sharing 2. Traditional data silo approach 3. Decentralised decision-making.	1. Communicating the value proposition, showcasing successful examples 2. Breaking down silos and using decentralised storage 3. Promoting collaboration and centralising decision-making	Interviewees 1, 4, 5,13
Time and resource constraints	1. Time-consuming and complex data sharing processes	1. Weighing the potential benefits against the time and resources required 2. Investing in necessary infrastructure and resources	Interviewees 6, 9,10
Data security concerns	1. Ensuring consistent security across all data sources 2. Security risks during data storage and data sharing	1. Applying strong security protocols and maintaining security across data. 2. Applying strong cybersecurity measures, ensuring secure storage, and sharing of data	Interviewees 2, 9, 10
Stakeholder coordination and collaboration	1. Difficulty in ensuring coordination and collaboration between different stakeholders from various sectors	1. Creating a culture of collaboration and clear communication channels 2. Setting up dedicated teams for cross-sectoral data sharing	Interviewees 1, 2, 3, 5, 6, 9, 11
Incentives and motivation	1. Lack of incentives or motivation for sectors to participate in cross-sectoral data sharing	1. Establishing clear benefits and rewards for participating sectors, promoting the value of data sharing, and providing support for data sharing	Interviewees 2, 4, 5, 7, 14
Data governance, data ownership and accountability	1. Identifying and assigning data ownership and accountability, especially when multiple stakeholders are involved. 2. Managing and maintaining data collected from different sectors.	1. Clearly defining data ownership roles 2. Develop data governance frameworks. 3. Create guidelines for data usage and management e.g. clear policies for data access, storage, and usage. 4. Ensuring data integrity, standardisation, compatibility and security.	Interviewees 3, 4, 5, 7, 9

Table 5
The key technologies and solutions in facilitating CDS within smart city development.

Technology/ solution	Benefits	Drawbacks	Interoperability considerations	Derived from interviewee(s)
Data sharing platforms	1. Facilitates consistent data exchange and integration across different sectors. 2. Improves overall system efficiency.	This may require standardisation of data formats across different sectors.	Platforms must be adaptable to the needs of different sectors.	Interviewees 3, 4, 10
Blockchain technology	1. Provides secure and transparent data sharing across sectors. 2. Ensures data integrity and traceability. 3. Enables monetisation of data sharing. 4. Supports execution of smart contracts.	1. May be slow and not scalable for some data sharing 2. Execution is still in its early stages. 3. Integration with existing data platforms may be challenging.	Requires a clear data governance structure to manage the data exchange.	Interviewees 1, 2, 4, 8, 9
Integration layer/ platform	1. Simplifies the process of connecting diverse data sources. 2. Enables merging of different datasets while maintaining their original systems. 3. Avoids over-integration.	1. Requires significant investment in infrastructure and development. 2. Integration with existing data platforms may be challenging.	Requires a clear governance structure to manage the data exchange.	Interviewees 4, 6, 13, 14

- More comprehensive with the addition of actions for each step. Adding more depth into the actions sectors must take. This was a key improvement to ensure the diagram met the requirements of serving as a guideline. These actions were derived from the results tables for what needs to be achieved in CDS based on the findings.

5.2.2. Key takeaways

- Little consideration was taken for the categorisation of steps.
- The diagram initially featured 14 steps, which, seemed to overcrowd the visual representation, potentially hindering its ease of reading, use, and execution.

The final version of the framework before validation was created on Canva, with final improvements made from Version 2 of the design (Fig. 8).

5.2.2.1. Changes made.

Table 6
The key factors towards achieving collaboration between sectors during CDS initiatives.

Collaboration category	Explanation	Derived from interviewee(s)
Understanding stakeholder interests, needs, and incentives	1. Addressing each stakeholder's unique needs, benefits, and values is vital for effective data-sharing collaboration. Understanding their requirements and expectations can be achieved through in-depth discussions and an inclusive approach. 2. Stakeholders must evaluate their own benefits when deciding to share data. In cases where a sector is unwilling to share data, presenting positive outcomes and similar success examples can encourage sharing. 3. Incentivising data sharing by demonstrating the positive value, providing payments, or cost-sharing agreements can encourage stakeholders to share their data willingly.	Interviewee 1, 5, 11
Communication, trust building, engagement	1. Open and transparent communication about the benefits, process, and importance of data sharing creates trust and collaboration. This involves clearly outlining the purpose and value created by the data sharing initiative. Regular meetings and open communication can help keep stakeholders informed and engaged throughout the process. 2. Actively including all stakeholders in the decision-making process helps align data sharing initiatives with their needs, building trust and encouraging collaboration. This may require collaboration and coordination between different sectors to ensure that the data sharing aligns with their needs. 3. Regular communication and collaboration can create a sense of trust and mutual understanding, enabling stakeholders to voice their concerns, identify potential challenges, and collaboratively develop solutions.	Interviewee 1, 4, 5, 6, 9
Data sharing governance: data agreements, protocols, and central governing body	1. Establishing clear agreements, protocols, and procedures for data sharing, including data ownership, roles, responsibilities, and data quality standards, is essential for effective stakeholder alignment. This should be outlined in a formal agreement or memorandum of understanding that is agreed upon by all sectors/ stakeholders involved. 2. A central governing body can be used to oversee data sharing agreements and ensure compliance with regulations, helping align the interests of different sectors and streamline the data-sharing process. The governing body would act on behalf of the different sectors involved in sharing the data, helping them with the push and pull-on making agreements about what data can be used and shared.	Interviewee 2, 4, 6, 9, 10, 14
Data management: standardisation, quality, and long-term maintenance	1. Applying standardised data formats, structures, and quality management processes ensures consistency and accessibility, meeting the needs of all stakeholders. This involves applying data validation, cleaning, and documentation requirements to ensure data quality and accuracy standards are met. 2. Ensuring the long-term maintenance of shared data, including data quality and accessibility, meets the needs of all stakeholders. This involves not only maintaining data quality but also ensuring that the data remains relevant, updated, and accessible over time. 3. Regular data reviews can help identify areas for improvement and ensure that shared data continues to meet the needs of all stakeholders.	Interviewee 2, 4, 11
Risk assessment, data security, and compliance	1. Identifying the potential risks of combining and aggregating datasets, applying appropriate measures, and adhering to relevant data protection legislation is crucial when sharing data across sectors. 2. Strict data security measures protect the data from unauthorised access or misuse. This can ensure that cross-sectoral data sharing meets the needs of all stakeholders while maintaining trust and compliance. 3. Organisations must take the time to address the different legal aspects and implications of data sharing, ensuring that all stakeholders' interests are taken into consideration.	Interviewee 5, 6, 13
Shared purpose and value	Agreeing on a central outcome of value and purpose for data sharing can help ensure that it meets the needs of all stakeholders and aligns with their interests. It is important to establish a central outcome of value and purpose before the data is shared, with all sectors	Interviewee 1, 9, 14

(*continued on next page*)

Table 6 (*continued*)

Collaboration category	Explanation	Derived from interviewee(s)
	involved agreeing upon the intended value and purpose of the data sharing.	

- The guideline underwent strategic categorisation, where similar steps were thoughtfully combined to ensure clarity and improve the diagram's readability. This process effectively reduced the total steps from 14 to 7, enhancing its overall coherence and ease of reading and execution. This has been carried out as many different or similar actions could be included within one step. For example, data inventory, data standardisation, data compatibility/interoperability and data quality have been combined into one step as they relate to each other,

- Connections and relationships between the 7 steps have been added. This has been displayed by the coloured circles representing each different step and the connections key, such as Strategic Planning & Stakeholder Alignment is related to steps 2–7. This improvement was essential to ensure that a comprehensive and interconnected view of the data sharing process is met. Providing sectors with the understanding that many processes within achieving CDS are connected, also shows the importance that all the steps are addressed to achieve successful CDS.

The final version of the framework comprises seven distinct steps, offering a structured approach to facilitate CDS in smart city development. These steps, ranging from strategic planning to monitoring and evaluation, outline clear actions for achieving successful CDS. The framework underscores the interconnectedness of these steps, emphasising their collective importance. This framework equips sectors to navigate the complexities of CDS, ensuring compliance with data protection legislation and fostering effective data sharing within smart cities.

Given that smart cities rely heavily on data, effective CDS is pivotal in breaking down data silos across different domains or sectors. This framework, by addressing key CDS components, contributes significantly to the advancement of smart city development.

5.3. Validation of the framework for CDS adhering to DPLs

This section details the findings from validation interviews with six expert participants, based on the framework's applicability, usefulness, execution, and ability to effectively outline how CDS can be achieved whilst considering adherence to DPL. This was then followed by improvements and changes made to the guidelines based on expert feedback.

5.3.1. Results & key findings

Interviewees collectively mentioned that the framework is useful because it provides a checklist and structured approach for CDS. Interviewee 2 believes that "the framework would be useful to help cities and other stakeholders understand what needs to be considered when it comes to data collection and data sharing". Leading on to a point where they stated, "This is important as more stakeholders express interest in data sharing". This suggests that the developed framework has achieved its intended purpose in providing an outline on how to achieve CDS, whilst considering the various challenges and the impacts of DPL. Interviewee 2 also said "It's important for all stakeholders to understand the steps needed for effective data sharing" reinforcing the framework's importance in facilitating CDS within smart city developments. Finally, interviewee 1 mentions that a barrier in applying the framework could be the number of stakeholders within the process and the fact that it

Table 7
The main risks and threats, alongside the mitigation strategies in addressing them during CDS initiatives.

Risk/threat	Description	Mitigation strategy	Derived from interviewee (s)
Data privacy, security risks, and classification	1. Unforeseen accumulation and triangulation of data can mean that individuals can be identified from two different datasets, posing a privacy risk. 2. Aggregating data sets from different sectors can increase security risks, potentially exposing sensitive data. 3. Data breaches and cyberattacks can compromise the confidentiality, integrity, and availability of shared data. 4. Inconsistent classification of data across sectors may result in security issues and complicate data-sharing initiatives.	1. Limit data granularity and ensure measures are in place to protect privacy. 2. Determine the appropriate level of data sharing, and limit access to sensitive data. 3. Apply strong data security measures like encryption and access controls to prevent unauthorised access. 4. Apply a consistent data classification system and ensure effective training for stakeholders.	Interviewee 1, 2, 5, 6, 7, 14
Data misuse, discrimination, and ethical concerns	1. The use of data can create a divide or discrimination. 2. Misuse of data or sensitive data leaks	Establish an ethical decision board and clear data sharing agreements outlining acceptable uses of data and the responsibilities of each sector involved.	Interviewee 3, 4, 12
Data loss, corruption, and storage	1. Data loss or corruption can occur due to accidental deletion, technical issues, or insider threats. 2. Decentralised storage can raise questions about data access and security.	1. Apply appropriate backup and recovery procedures to help mitigate the risk of data loss. 2. Design a secure data storage architecture and carefully manage and monitor data access to help ensure security.	Interviewee 3, 6
Data protection compliance and predictive security framework	1. Ensuring compliance with data protection legislation is a significant concern. 2. Rapid technological developments may not be addressed by current data protection laws. 3. The lack of a predictive framework for data security can	1. Comply with existing regulations but make further appropriate measurements based on the data being shared. 2. Develop a predictive framework that ensures data security, using vectorisation and AI approaches, and ensures consistent	Interviewee 4, 5, 7, 11

(*continued on next page*)

Table 7 (*continued*)

Risk/threat	Description	Mitigation strategy	Derived from interviewee (s)
	lead to unauthorised access and data breaches.	classification of data across sectors.	
Transparency and effective data governance	1. Citizens might not be informed about data collection, its usage, and potential privacy issues, leading to a lack of trust. 2. Poor data governance can compromise data privacy and security.	1. Inform citizens about data collection, enhance transparency in data usage, and follow legislations. 2. Establish trust by using technologies such as blockchain for secure data sharing. 3. Establish and enforce data governance and sharing agreements.	Interviewee 2, 3, 8
Data anonymisation, purpose limitation, and minimalization	1. Sharing sensitive data across sectors can lead to privacy risks. 2. Sharing data without clear purposes can increase the risk of misuse or unauthorised access. 3. Sharing large amounts of data can lead to increased privacy risks.	1. Apply data anonymisation. 2. Define and enforce purpose limitation for data sharing 3. Apply data minimisation techniques to share only necessary data.	Interviewee 2, 4, 7, 8, 14

requires everyone to be on board.

Further to this, interviewee 3 found the diagram useful as it provides a complete overview of the challenges and elements of creating a CDS process and mentioned the importance of not overlooking ethical considerations. They also mentioned that it could provide a useful guideline for initiating a data sharing project and provide a clear outline of what would be involved. The interviewee stated: "The framework addresses several current challenges for cross-sectoral sharing because it provides guidelines about what is involved in starting a project". This suggests the framework can be applied in applying CDS initiatives and addressing the challenges involved. However, interviewee 3 criticised the diagram suggesting it needs "a more structured approach on what should be done first, second and so on" highlighting a key flaw within the usefulness of the conceptual framework, which must be addressed by organising the steps in a more logical or structured flow. However, when interviewee 3 was asked if they believed all the key elements in achieving CDS had been included, they responded believing that the framework covers all the relevant steps and aspects that come to his mind.

Finally, interviewee 4 stated the framework was impressive and that it included the critical components for CDS when the steps and actions within the diagram were explained to them. Additionally, when asked about the applicability of the framework, they said, "Smart cities are ecosystems of different components and data is key in making any system smart, so a cross-sectoral data sharing framework like this one could be useful when planning a smart city" reinforcing the importance of the guidelines for CDS initiatives within smart city developments. Further to this, interviewee 4 stated that the framework is applicable and pointed

Table 8
The most vital requirements during CDS initiatives within smart city developments, highlighted by multiple expert interviewees.

Requirements	Explanation	Derived from interviewee(s)
Data management and standardisation	1. Ensuring that data formats are standardised across sectors to facilitate effective integration, sharing, and analysis of the data. 2. Applying data tagging and cataloguing systems to provide an overview of available data and the ability to discover data. 3. Maintain high quality data through validation and verification processes and ensure its accuracy and relevance by regularly updating shared data.	Interviewee 2, 3, 4, 11
Data privacy and security measures	1. Privacy protection through masking personal identifiers, anonymisation techniques and adhering to privacy-by-design principles. 2. Applying multiple layers of protection based on data sensitivity e.g., Health data should be highly protected with measures including encryption, blockchain, access controls, and other security measures. 3. Developing a secure shared data infrastructure, moving from a decentralised data storage system to a centralised data exchange platform to improve data security, privacy, and efficient sharing across sectors.	Interviewee 2, 4, 9, 10
Data protection compliance and risk management	1. Ensuring legal compliance with data protection legislation by applying processes to comply with them. 2. Collecting and sharing only the necessary data for the intended purpose (data minimisation) to reduce the risk of sensitive data exposure and ensure compliance with data protection legislation. 3. Establishing clear accountability for data handling and applying data access control to ensure that only authorised people can access the data. 4. Conducting privacy impact assessments to identify and mitigate potential privacy risks in data sharing.	Interviewee 2, 4, 5, 9, 10, 12, 13
Stakeholder collaboration and communication	1. Combining top-down and bottom-up approaches in applying cross-sectoral data sharing to ensure effective collaboration and involvement from all stakeholders. 2. Identifying the goals, scope, and objectives of data sharing initiatives, demonstrating the value created for different sectors, and aligning with their goals. 3. Promoting effective communication and collaboration between different sectors to better understand complex data sets and ensure that data is usable and valuable for all sectors involved. 4. Recognising the potential benefits of cross-sectoral data	Interviewee 1, 5, 6, 9, 14

(*continued on next page*)

Table 8 (*continued*)

Requirements	Explanation	Derived from interviewee(s)
	sharing and how it can lead to better decision-making, cost savings, and improved outcomes for the smart city.	
Experienced team members	1. Having team members and technology teams with experience in managing cross-sectoral data-sharing projects to ensure that the project stays on track, achieves its goals, and adheres to legal data protection requirements. 2. Team members should be provided with basic data analysis tools to help understand the value of the data available and facilitate data-driven decision-making.	Interviewee 3, 6, 14

out that the framework could be applied to specific sectors within a smart city, such as healthcare or transportation as it covers aspects like data governance and legislation compliance which vary across sectors.

5.3.2. Framework improvements (sensitive/non-sensitive data concept)

Interviewee 2 stated that the order of steps and actions might vary depending on the data sharing project and the framework might need to adapt to different scenarios. Contributing to this point, Interviewee 1 suggested that sharing non-sensitive data might not require all the steps, and some may not apply to different sectors, such as, geographical, or environmental data being shared would not require data protection compliance legislation. Lastly, Interviewee 3 suggested that the final guideline could allow for customisation based on the needs of a sector, allowing them to see what is relevant to them.

To address these suggestions, the conceptual framework has been categorised and split into two different layers (Fig. 9):

1. Providing sectors sharing sensitive data the original guideline with comprehensive steps covering aspects of data privacy and data protection compliance
2. Providing sectors sharing non-sensitive data the ability to follow the relevant steps to their specific data-sharing context. This included removing the data protection compliance step and actions involving data privacy within each step.

5.3.3. Framework improvements (supporting framework diagram)

When discussing the developed conceptual framework with interviewee 6, it was suggested that a separate diagram could be created to provide sectors with an overview of the factors involved with the relationship between CDS and adherence to DPL. The interviewee identified this improvement after the research gap was explained, as it was later highlighted that the framework meets the purpose of how to achieve CDS. However, it was said that the guideline is limited in considerations towards how DPL impact the CDS process. Therefore, an additional diagram has been developed to support the framework, providing stakeholders with a further understanding of the impact of DPL when CDS.

The diagram (Fig. 10) presents an understanding of the complex relationship between some of the relevant factors in achieving CDS whilst adhering to DPL. Finally, this additional conceptual framework considers a specific focus of this research, this being how CDS is impacted by DPL within smart city developments.

5.3.4. Final improved framework for CDS adhering DPLs

Based on the following improvements the final version of the framework for CDS while adhering DPLs has been developed (Fig. 11):

Table 9
The main factors towards determining what data can be shared within CDS initiatives in the context of a smart city.

Data being shared factors	Explanation	Derived from interviewee(s)
Value proposition, benefits, and sector needs	1. The potential benefits of sharing the data should be assessed, such as improved decision-making and collaboration. 2. Sectors should identify their specific data needs for effective data sharing, consider the challenges and consider the legal requirements of sectors like healthcare (sensitive data). 3. The value and need of sharing the data should be weighed against the costs and risks associated with it.	Interviewee 1, 4, 6, 11
Privacy and security concerns, data protection compliance and data agreements/monitoring	1. Compliance with data protection and privacy legislations such as GDPR should be ensured. 2. Effective governance, access controls, and security measures should be applied to protect sensitive data. 3. The potential risks associated with sharing sensitive data should be evaluated, and strategies should be developed to minimise the risks. 4. Privacy concerns should be addressed through anonymisation, feature granularity, and strict adherence to data usage agreements. 5. Agreements should be established to ensure that data is not used for unintended purposes. 6. Monitoring systems should be applied to verify adherence to data sharing agreements and compliance with relevant regulations.	Interviewee 3, 5, 9, 14
Data sensitivity, classification, and compatibility	1. The sensitivity of data should be determined and classified. 2. Data should also be classified into multiple layers to determine what data can be shared or exchanged between sectors. 3. Smart classification systems should be developed to streamline the sharing process while maintaining data security. 4. The technical compatibility of data should be evaluated, and it should be ensured that the data can be effectively used across sectors. 5. The potential risks of aggregating data from different sources should be considered, and appropriate measures should be in place to minimise these risks.	Interviewee 4, 7, 9
Stakeholder willingness, collaboration, and transparency	1. The willingness of sectors to share data should be assessed and potential barriers, such as privacy concerns.	Interviewee 1, 3, 9, 13, 14

(*continued on next page*)

Table 9 (*continued*)

Data being shared factors	Explanation	Derived from interviewee(s)
	2. Communication between sectors should be encouraged to determine what data can be shared and under what conditions. 3. Incentives of payment or demonstrating the value of data sharing, to encourage the different sectors to share data. 4. Transparency should be promoted by using an open data approach, making the data public unless there is an important reason not to. 5. Data subjects should be aware of how their data is being used and the value it adds.	

Table 10
The most significant success factors, based on the experience of smart city experts, for CDS within smart cities.

Success factors	Explanation	Derived from interviewee(s)
Desired outcome, clear problem definition, and stakeholder alignment	1. Align the project's focus with stakeholder goals to create value for all the sectors involved. 2. Identify the specific goals, objectives, and benefits. 3. Ensure that all stakeholders agree on the problem to be solved. 4. Collaboratively address concerns, and potential challenges, and develop solutions.	Interviewees 6, 9, 11
Strong governance, legal framework, and data quality measures	1. Apply measures of encryption and access control. 2. Ensure data protection, data privacy, and compliance with legislation. 3. Ensure data interoperability, integration into existing processes, and address inaccuracies to maintain data quality.	Interviewee 2, 5, 6, 9, 13
Appropriate data standards and technical solutions	1. Facilitate compatible and compliant data sharing between sectors. 2. Apply data management systems, data conversion tools, and data sharing technologies for secure data sharing.	Interviewee 2, 5, 6, 9, 12
Stakeholder involvement, clear communication, and transparency	1. Establish trust and understanding between stakeholders through clear communication and transparency. 2. Keep the public informed about the data sharing process and its implications.	Interviewee 5, 6, 9, 11

- Removing Connections Key: Interviewees 1, 2, 3, 4 and 5 suggested the connections key should be removed as it complicated/overwhelmed the diagram, making it unclear and difficult to understand.
- Strategic Categorisation: Based on the expert interviewee's feedback, it was understood that data sharing agreements and data protection compliance had overlapping actions, were unclear and placed incorrectly. Therefore, they were combined, resulting in 6 steps instead of the original 7 steps (mainly interviewee 1).
- Data Catalogue Action: The importance of having a data catalogue was highlighted by the interviewee, stating that gaining a full overview of the data even within just one organisation can be challenging, even with data inventory in place, so applying a data catalogue can help overcome this issue. Therefore, it has been added as part of an action in the 2nd step (interviewee 1–5).
- Data Sharing Architecture: The data sharing architecture was overlooked in the initial diagram, as highlighted by the interviewee. Therefore, it was included in the 5th step (interviewee 1).
- Planning Data Protection: The interviewee mentioned that an area of concern for the diagram could be DPL being late in the process. This suggests that the following step should be included in Step 1 "Consider data protection legislations at this stage to ensure compliance later on and identify any potential limitations or conflicts with data protection legislations". The interviewee also mentioned that by doing so sectors can navigate legislation better, identify limitations early, and understand if the data sharing will be possible which is crucial in deciding the path to take or data to share (interviewee 3).
- Data Validation: Interviewee 3 stated the importance of data validation when ensuring data quality and potential inaccuracies. Therefore, "Execute rigorous data validation procedures to verify that the data can be used for the described purpose" has been added to the 2nd step of the diagram as this is significant in CDS.
- Data Sharing Incentives: The initial diagram did not fully consider the need for incentives, to ensure stakeholder participation in data sharing. Therefore, the addition of considering incentives has been added into the 1st step (interviewee 3).
- Data Quality Management: The interviewee highlighted the importance of data quality measures; therefore, this action was expanded in greater detail (interviewee 3).
- Continuous Process: Interviewee 6 suggested the need to highlight that the CDS process should be continuous and is key to any project he has been involved in. Mentioning the process being "Living" and that data should be open where possible. Therefore, an action has been added to step 1 "Ensure the data sharing initiative continues to be adaptable and responsive to developing data sharing needs and contexts".

As shown in Fig. 11, conducting validation interviews has established and reinforced the framework meets its intended purpose. Although the framework developed previously within the discussion section was comprehensive, there were many improvements to be made, highlighted by return expert interviewees. Validation of the developed framework has provided various areas to improve the diagram and make changes to ensure each step contributes towards achieving CDS. Some of the most significant improvements made to the framework to ensure it meets its purpose are; further strategic categorisation reducing the guideline to 6 key steps, removing the connections key for clarity, developing a second version of the diagram presenting the steps for non-sensitive CDS, developing a supporting diagram consisting of the factors influencing the relationship between CDS and DPL and various additions to actions within steps such as the importance of data catalogues.

Validating the framework ensured that it is comprehensive in addressing the research question and filling the gap identified within the literature of a lack of understanding on how to achieve CDS successfully. Backed up by interviewees' expert opinions, the framework also achieves the research aim by clearly and comprehensively outlining almost every factor which must be considered if a sector is to begin CDS initiatives. However, it has also been discovered that the diagram may have execution barriers, as previously highlighted by some interviewees. Although, this is mainly due to the immaturity of the topic, and as smart city development continues, this guideline may become even more useful. Finally, although there are further improvements which must be

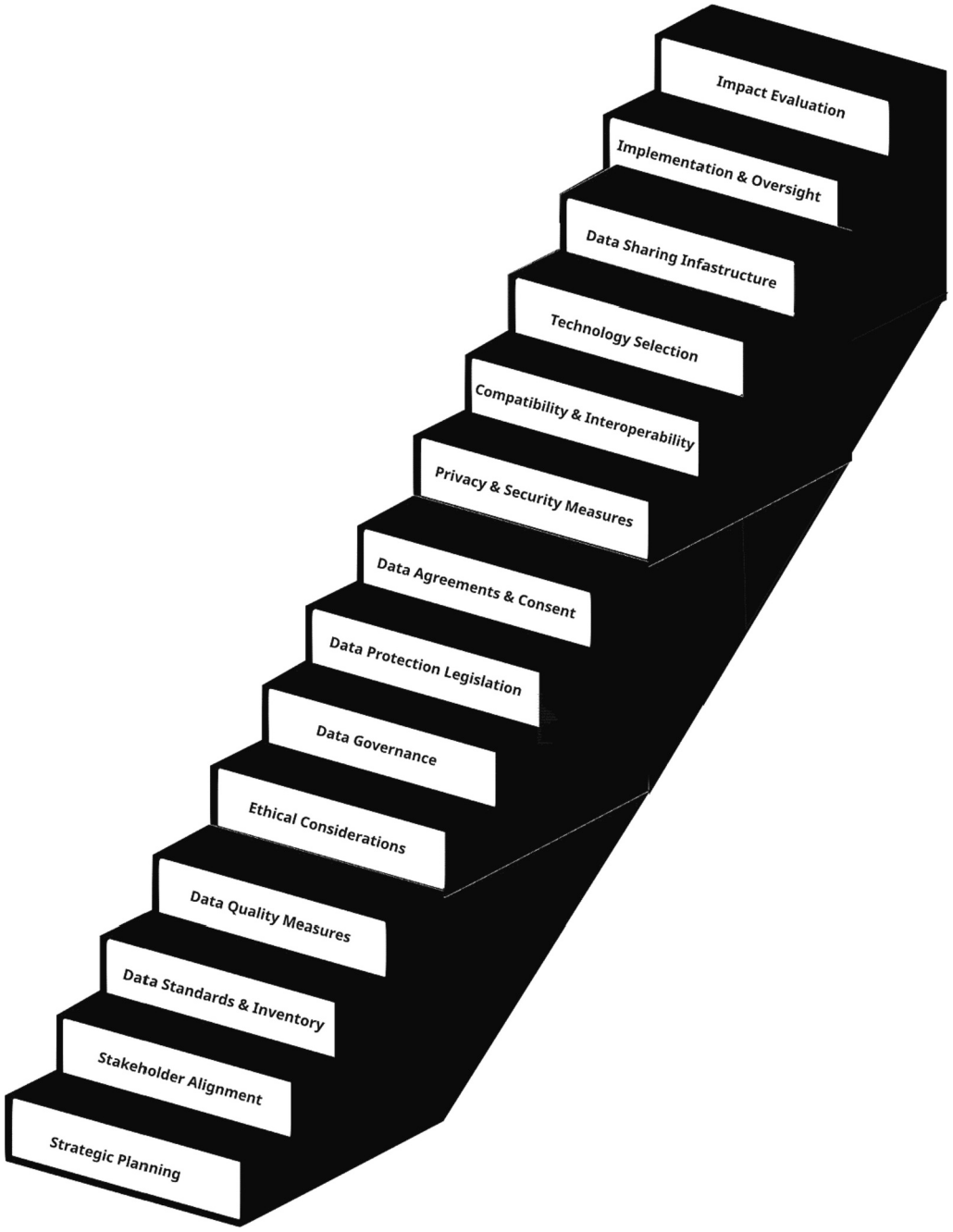


Fig. 6. Framework for CDS adhering to DPLs - Version 1.

addressed in the framework guideline development, the validation strategy of the framework has effectively addressed objective 5 of this research study.

The following are the key audiences for the framework:

- City planners and administrators: They can use this framework to understand the necessary steps and considerations for effective data sharing across various sectors in a smart city. It provides guidance on overcoming barriers, managing risks, and achieving successful CDS initiatives, aligning with the strategic planning and governance needed in city development.
- Data management professionals: Professionals responsible for handling data within different city sectors will find this framework beneficial. It emphasises the importance of data inventory, quality management, standards, security measures, and governance, aiding them in implementing these aspects effectively.
- Policymakers and regulatory bodies: Individuals involved in creating DPLs or those responsible for enforcing compliance can use this framework as a reference. It highlights the need for adherence to data protection laws and ethical considerations, contributing to the creation of effective policies that balance data sharing and privacy concerns.
- Technology developers and solution providers: Companies and innovators involved in developing technological solutions for data sharing in smart cities will find insights into key requirements and challenges, aiding them in creating solutions that align with the specific needs highlighted in the framework.
- Researchers and academics: Academics and researchers focusing on smart city development, data sharing, or data governance will

Fig. 7. Framework for CDS adhering DPLs - Version 2.

benefit from this framework. It provides valuable insights into the challenges and requirements, acting as a reference for further studies and research in the field.

- Smart city stakeholders and community: It also caters to various stakeholders involved in smart city initiatives, offering a clear understanding of the interconnected steps necessary for successful CDS. This enables informed decision-making and collaboration among stakeholders, fostering effective data sharing within smart cities.

5.4. *Linking our findings to the body of knowledge (some examples)*

The comprehensive investigation into CDS aligned with DPLs reveals critical challenges and requisites resonating with established research. The identified challenges, explained in Section 2, insights from interviews with experts from various global contexts, and the subsequent framework creation closely align with established literature's discussions on the need for comprehensive, adaptable frameworks. Our framework intends to provide a structured approach that navigates complexities while embracing sector-specific needs, aligning with the literature's emphasis on tailored strategies for successful CDS.

- For example, Lnenicka et al. (2022) underscore similar aspects, emphasising the significance of data quality, standardisation, interoperability, privacy concerns, and stakeholder engagement within smart city frameworks. Our research echoes these crucial elements, spotlighting their prominence in facilitating effective CDS initiatives. The complexities outlined, including the challenges in adhering to DPL, resonate with broader literature emphasising the nuanced nature of data sharing, especially in diverse sectors gathering sensitive data.
- Another example is the research conducted by Nikiforova and Zuiderwijk (2022) that highlights the evolution of research on Open Government Data (OGD), particularly concerning the drivers and inhibitors for its adoption. They emphasised that while previous studies primarily focus on the reuse of OGD by companies and citizens, their study notes a research gap in understanding the barriers to provision, especially the resistance of public organisations to share government data. They also discuss the Innovation Resistance Theory (IRT), which delineates functional and psychological barriers, encompassing challenges like data quality issues, complications in data preparation, value perceptions, risk concerns including misuse and privacy, traditional organisational culture, and image apprehensions, hence, presenting a nuanced understanding of barriers impeding data-sharing initiatives and reinforces the need for tailored strategies to overcome resistance among public organisations.
- Moreover, Zuo et al.'s (2023) study explores citizen discontent within smart cities, shedding light on active and passive dissatisfaction arising from technical, societal, and privacy-related issues. Their research resonates strongly with our findings from the CDS study, notably in the challenges concerning data quality, interoperability, and adherence to DPLs. Both Zuo et al.'s insights and our findings converge on the significance of addressing technical challenges, societal impacts, and privacy concerns in smart city initiatives. Our framework development, rooted in the insights garnered from interviews, aligns intricately with Zuo et al.'s identification of citizen discontent types and government responses. Their perspectives on discontent serve as a complementary lens to our framework, aiding in refining and enriching the steps delineated for effective CDS initiatives. The synthesis of Zuo et al.'s research with our findings contributes to a comprehensive understanding of citizen discontent in smart cities and bolsters the development of a robust framework for successful cross-sectoral data sharing initiatives, emphasising the critical components essential for navigating the complexities of smart city landscapes.
- The smart city data framework proposed by Liu et al. (2017), while provides valuable insights into managing data within smart cities, it

Fig. 8. Framework for CDS adhering DPLs – Final version before validation.

Fig. 9. Improving the framework for CDS adhering DPLs.

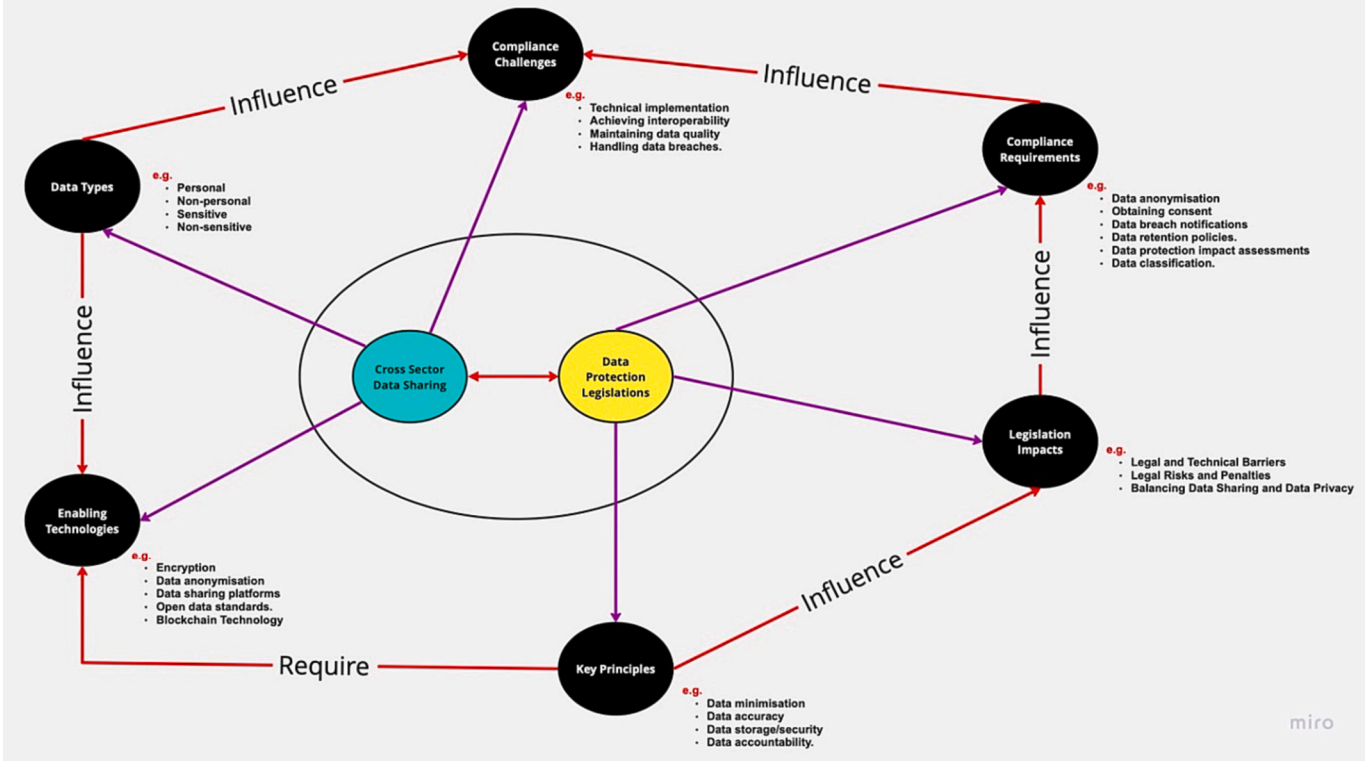


Fig. 10. Supporting framework diagram.

notably lacks coverage on CDS and the intricate challenges associated with adhering to DPL. Our developed framework addresses this gap by explicitly focusing on the relationship between CDS and DPL, offering a comprehensive guideline for smart city development that extends beyond the confines of data quality and privacy, providing a more holistic approach. In contrast to Liu et al.'s (2017) framework, which confines its scope to specific dimensions of data management, our framework takes a more holistic approach. It not only incorporates robust measures for ensuring data quality and privacy but also delineates a roadmap for effective and compliant data sharing across diverse sectors within a smart city. By explicitly addressing the challenges posed by CDS in conjunction with DPL, our framework becomes an essential tool for urban planners, policymakers, and practitioners navigating the complex landscape of smart city development. It not only builds upon the foundational aspects of data quality and privacy but extends its reach to foster a more inclusive, interconnected, and ethically sound approach to data sharing in the context of modern smart cities.

- Furthermore, the work of Javidroozi et al. (2019a) although contributes significantly to the discourse on smart city evolution, it primarily centres around business processes and lacks explicit consideration of CDS and DPL. Our framework, developed in this study, complements and extends the existing knowledge by providing specific guidance on CDS and DPL, offering a more comprehensive perspective that aligns with the evolving needs and challenges of contemporary smart city development.

Our research, thus, not only contributes to filling the existing void in literature but also aligns with established perspectives, creating a bridge between the identified challenges, frameworks, and the overarching discourse on CDS and DPL adherence within smart cities.

6. Conclusions

Establishing CDS stands as a pivotal element within the landscape of smart city advancements, bearing the potential for substantial value generation. Nevertheless, navigating the path to a successful realisation of CDS within smart city environments reveals profound complexity. Numerous hurdles demand attention to facilitate seamless data collaboration among city sectors, spanning issues of data quality, standardisation, interoperability, privacy preservation, and stakeholder involvement. Moreover, while adherence to DPL remains integral, it amplifies impediments to achieving CDS, notably considering the collection of sensitive data across diverse city sectors. Despite the critical significance of CDS and DPL in smart city contexts, scholarly literature reveals a distinct paucity of comprehending strategies to surmount these challenges.

This research has investigated the relationship between CDS and DPL, with the aim of understanding how CDS can be successfully facilitated in the context of a smart city. A literature review was conducted to identify, justify, and address this gap within the literature. This addressed objectives 1 and 2 of this project by gaining an understanding of the importance, requirements, and solutions towards CDS and further exploring the importance and impacts of DPL within smart cities. This was followed by semi-structured interviews with 14 smart city experts with experience and involvement in smart city projects, mainly in Helsinki, Viena, London, Stockholm, Bern, Riyadh, Dubai, Barcelona, Birmingham, and three cities in India, including Delhi, Bangalore, and Chennai. This resulted in heavily influential and valuable insights into the requirements, challenges, and strategies within achieving CDS, deriving the key concepts and requirements to be incorporated into the development of the conceptual framework.

The main outcome of this research is the development of a framework for CDS adhering to DPLs that consists of six steps, providing a structured approach to facilitating CDS initiatives and achieving the aim of this research project. By addressing the challenges and requirements of data sharing, the guideline contributes to the success of CDS initiatives whilst maintaining adherence to DPL and, ultimately addresses objective 4 of this research. Following the initial developments of the framework, a second round of interviews with previous participant

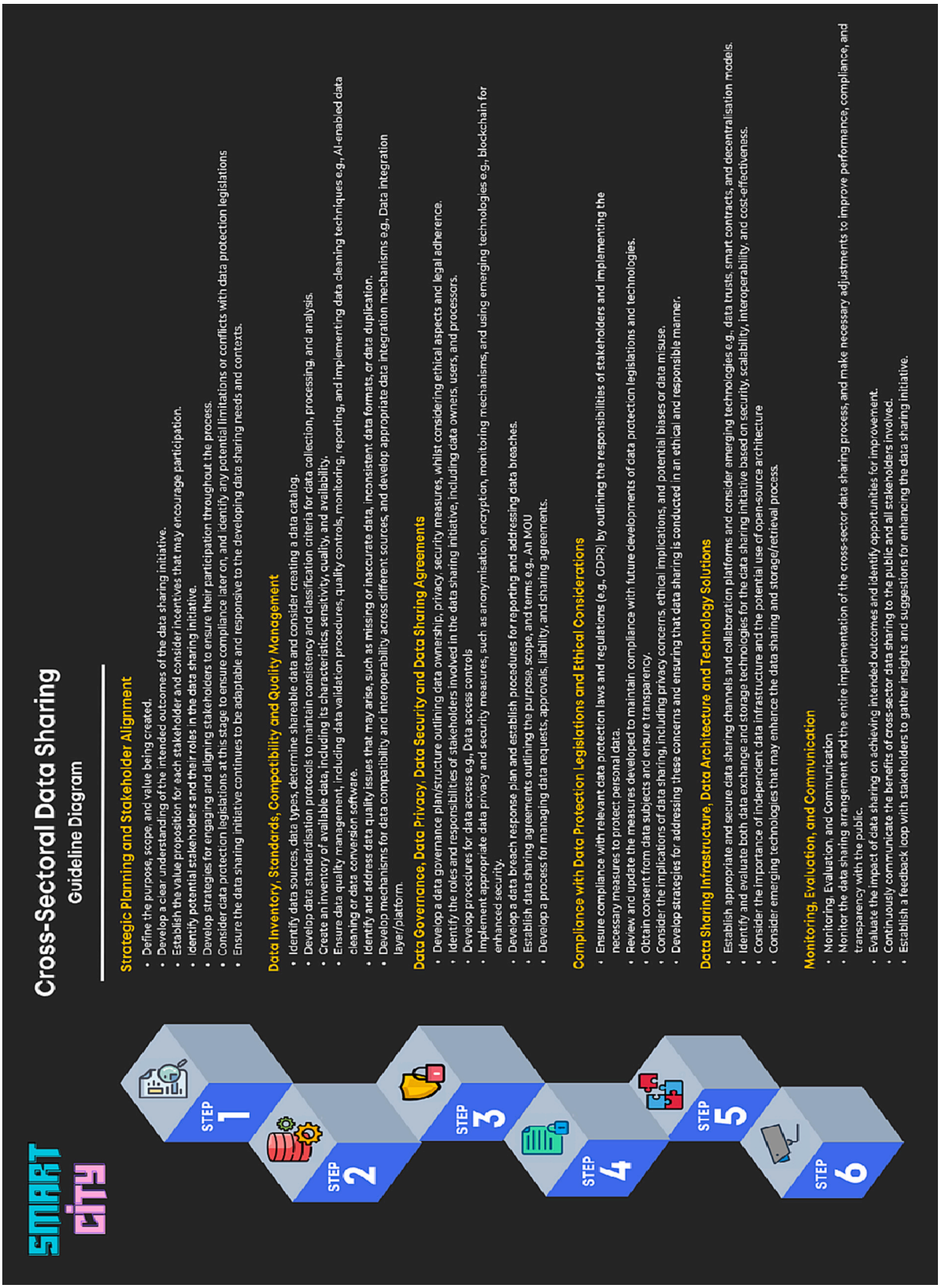


Fig. 11. Validated version of the framework for CDS adhering DPLs.

smart city experts helped further refine and validate the framework's effectiveness as a valuable resource within the development of smart cities, addressing objective 5 of this research study.

The significant findings of this study can assist in effective CDS, supported by many interviewees, including the importance of strong data governance, constant communication between stakeholders, addressing data quality issues, identifying the value being created by data sharing and applying data standardisation. These insights provide a general outlook on the crucial factors contributing to successful CDS and overcoming the barriers hindering data-sharing initiatives. However, more factors have been identified within the results tables of this research.

Overall, this study has made a substantial contribution to the literature by providing further insights into the challenges sectors must address, a comprehensive understanding of the various aspects of CDS and a framework towards achieving CDS in smart cities. By following the developed framework, sectors can navigate the complex issues involved with CDS, ensuring they apply all the requirements and strategies mentioned, paving the way for more effective collaboration within smart city development.

A summary of the study's contributions is as follows:

- Framework development from key findings: The framework and it's supporting content (including text and diagram), as the study's primary contribution, synthesises critical concepts, variables, and factors obtained from results tables:
- It guides successful CDS initiatives while addressing the implications of DPLs.
- The developed framework comprehensively serves as a guideline for sectors, aiding in barrier resolution, risk management, and successful CDS. This addresses the identified research gap in this study.
- The framework equips various stakeholders in smart city development with a structured approach to successful cross-sectoral data sharing initiatives while addressing complexities and ensuring compliance with DPLs.
- Enabling data-driven smart city advancements: The research outcome stands as a critical resource in advancing smart city development. Given the foundational importance of data sharing in smart city initiatives, the framework provides a structured approach that facilitates seamless collaboration among various city sectors. By emphasising the critical components necessary for effective data sharing, the framework enables the breaking down of data silos across diverse domains or sectors within a smart city. This contribution underlines how the developed framework acts as a catalyst for smart city advancements, empowering city planners, administrators, technology developers, policymakers, and other stakeholders involved in smart city initiatives. It ensures a strategic and cohesive approach to data sharing, ultimately fostering the realisation of data-driven solutions and innovations within the smart city landscape.
- Integration into smart city development roadmaps: The guide formulated within this study presents a significant asset that seamlessly integrates with smart city development roadmaps, especially concerning the data-centric aspects. As smart city initiatives evolve, the guide provides a structured pathway, ensuring the efficient utilisation and management of data across diverse sectors. This integration enhances the existing roadmaps by introducing a comprehensive approach to cross-sectoral data sharing, addressing challenges, managing risks, and promoting compliance with data protection legislation.
- Guiding data-focused smart city frameworks: Moreover, the guide serves as a foundational tool in shaping smart city frameworks, particularly those centred on data governance, data sharing, and interoperability. By presenting essential steps and considerations, it guides the development of frameworks that prioritise seamless data exchange while upholding privacy, security, and ethical standards. This integration supports the creation of robust data-centric frameworks essential for fostering innovation and sustainable growth within smart city environments.
- Literature integration: Enhanced integration of findings from empirical interviews with existing literature enriches the understanding of CDS complexities and their interaction with DPLs.
- Influence of findings on the framework: Results tables from interviews significantly shaped the framework development. Each framework step derives from pivotal factors identified, ensuring a comprehensive yet concise representation of key CDS actions.
- Structured approach in framework development: A systematic approach was employed to ensure the framework's inclusivity. It involved selecting essential steps from interview insights, categorising, refining, and iterating the framework design based on expert opinions.

6.1. Limitations of the research

Despite earnest attempts to mitigate constraints, it's imperative to recognise inherent limitations within the research process while elucidating the findings of this investigation. For example, a limitation inherent in this study stems from the nature of the data collection and analysis. The qualitative nature of the interviews, while instrumental in capturing nuanced insights, may not encompass the entire spectrum of perspectives prevalent within the domain of smart city initiatives.

In addition, while the sample size of 14 experts could be perceived as a limitation, it is important to note that efforts were made to reach a saturation point where the collected data became repetitive, and no new themes, ideas, or concepts emerged. This approach was aimed at ensuring a comprehensive understanding of the subject matter, thereby minimising the impact of this limitation.

Additionally, gathering data from cities primarily associated with smart city projects across various global locations might not fully encapsulate all the experiences and challenges encountered in diverse smart city contexts worldwide, hence considered as a limitation. However, this approach offered a unique advantage by providing insights from cities at different stages of smart city development, each with distinct challenges, progress trajectories, and specific requirements. This diversity significantly enhanced the generalisability and applicability of our findings, offering a more comprehensive understanding of the complex landscape of data-sharing initiatives within the smart city realm.

6.2. Recommendations for future research

While this study has covered how CDS can be achieved whilst adhering to DPL in great depth, there is much more to be addressed within this smart city research area. Specific to this study, further work could involve several improvements to the framework developed as follows:

- Adapting the framework to suit different sectors' needs by highlighting the actions to be taken for each specific sector,
- Including more dimensions within the framework such as the risks within each step or incorporating different perspectives as dimensions such as technical, economic, and political,
- As mentioned by a few interviewees, adding an ‘intelligence layer’ in a future version of the framework could be driven by AI. For example, this could involve the exploration of data standards with the help of AI.

Moreover, further work non-specific to this study, within this research area could include investigating the challenges and solutions for CDS in specific sectors such as healthcare, transportation, and energy to provide guidance for each sector's unique requirements. An additional framework could also be developed specifically to measure the value or

impacts of CDS initiatives within smart cities, which could include assessing the overall benefits.

CRediT authorship contribution statement

Aaron Joyce: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Vahid Javidroozi:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Interview questions

1. Can you briefly describe your background or expertise in the field of Smart Cities/Data Sharing?
2. How do you evaluate the importance of cross-sectoral data sharing and what are the main benefits from cross-sectoral data sharing for the purpose of smart city development?
3. How do you currently see cross-sectoral data sharing being implemented for the purpose of smart city development?
4. What do you think are the main challenges/barriers in facilitating cross-sectoral data sharing?
5. How do you suggest overcoming/addressing these challenges/barriers?
6. What are the potential risks of sharing data across different sectors in a smart city context? How do you suggest they can be mitigated?
7. What do you believe are the key requirements for achieving cross-sectoral data sharing, with consideration into adhering to DPL?
8. How do you think DPL impact the ability to achieve cross-sectoral data sharing in smart cities?
9. What DPL do you think are the most important to consider in this context?
10. Are there any technology solutions and any methods that you think can play a key role in facilitating cross-sectoral data sharing?
11. What factors do you think should be taken into consideration when determining what data can be shared across different sectors?
12. How do you suggest ensuring data quality when sharing data across sectors in a Smart City?
13. What are some of the challenges associated with storing and using data that has been collected from multiple sectors?
14. What are the main threats to data security when sharing data across different sectors?
15. What are the main concerns around data privacy when sharing data across different sectors in a Smart City?
16. How can we ensure that cross-sectoral data sharing is aligned with the needs of all stakeholders involved?
17. In your experience, what are some successful examples of cross-sectoral data sharing projects in Smart Cities? Additionally, what are some of the success factors?

References

A&L Goodbody. (2016). *The GDPR: A guide for businesses* (Belfast, London, New York, Palo Alto, San Francisco, Dublin).

Abdalla, W., Renukappa, S., Suresh, S., & Al-Janabi, R. (2019). Challenges for managing smart cities initiatives: An empirical study. In *Proceedings - 2019 3rd International Conference on Smart Grid and Smart Cities, ICSGSC 2019* (pp. 10–17).

Anisetti, M., Ardagna, C. A., Braghin, C., Damiani, E., Polimeno, A., & Balestrucci, A. (2021). Dynamic and scalable enforcement of access control policies for big data. In *ACM International Conference Proceeding Series* (pp. 71–78).

Arroub, Ayoub, Bassma Zahi, Essaid Sabir, and Mohamed Sadik. 2016. "A literature review on smart cities: Paradigms, opportunities and open problems." Proceedings - 2016 International Conference on Wireless Networks and Mobile Communications, WINCOM 2016: Green Communications and Networking 180–86.

Asswad, Jad and Jorge Marx Gómez. 2021. "Data ownership: A survey." Information 2021, Vol. 12, Page 465 12(11):465.

Bakıcı, T., Almirall, E., & Wareham, J. (2013). A smart city initiative: The case of Barcelona. *Journal of the Knowledge Economy, 4*(2), 135–148.

Bargh, M. S., Meijer, R., Choenni, S., & Conradie, P. (2014). Privacy protection in data sharing: Towards feedback based solutions. In *ACM International Conference Proceeding Series* (pp. 28–36) (2014-January).

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

Berardi, R., Belizario, M. G., Cristina, R., & Berardi, G. (2019). *Use of smart and open data in smart cities*.

Bibri, S. E. (2018). The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. *Sustainable Cities and Society, 38*, 230–253.

Bibri, S. E., & Krogstie, J. (2017). Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable Cities and Society, 31*, 183–212.

Braun, T., Fung, B. C. M., Iqbal, F., & Shah, B. (2018). Security and privacy challenges in smart cities. *Sustainable Cities and Society, 39*, 499–507.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Chiara, P. G. (2021). The balance between security, privacy and data protection in IoT data sharing: A critique to traditional 'Security&Privacy' surveys. *European Data Protection Law Review, 7*(1), 18–30.

Christofi, Athena, Rob Heyman, Laurens Vandercruysse, Valerie Verdoodt, Caroline Buts, Michael Dooms, Jo Pierson, and Peggy Valcke. 2019. "Smart city privacy: Enhancing collaborative transparency in the regulatory ecosystem." 2019 CTTE-FITCE: Smart cities and information and communication technology, CTTE-FITCE 2019.

Clement, J., Manjon, M., & Crutzen, N. (2022). Factors for collaboration amongst smart city stakeholders: A local government perspective. *Government Information Quarterly, 39*(4), Article 101746.

Correia, D. M., Teixeira, L., & Marques, J. L. (2021). Smart supply chain management: The 5W1H Open and Collaborative Framework. In *2021 IEEE 8th International Conference on Industrial Engineering and Applications, ICIEA 2021* (pp. 401–405).

Cui, L., Xie, G., Youyang, Q., Gao, L., & Yang, Y. (2018). Security and privacy in smart cities: Challenges and opportunities. *IEEE Access, 6*, 46134–46145.

Denker, A. (2021). Protection of privacy and personal data in the Big Data environment of smart cities. In *The international archives of the photogrammetry, remote sensing and spatial information sciences XLVI-4-W5-2021(4/W5-2021)* (pp. 181–186).

Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017). Blockchain for IoT security and privacy: The case study of a smart home. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2017* (pp. 618–623).

Falconer, G., & Mitchell, S. (2012). *Point of view smart city framework a systematic process for enabling smart+connected communities*.

Gruschka, Nils, Vasileios Mavroeidis, Kamer Vishi, and Meiko Jensen. 2019. "Privacy issues and data protection in Big Data: A case study analysis under GDPR." Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018 5027–33.

Graef, I., Tombal, T., & de Streel, A. (2019). Limits and enablers of data sharing. An analytical framework for EU competition, data protection and consumer law. *SSRN Electronic Journal*, 1–34.

Guenduez, A. A., & Mergel, I. (2022). The role of dynamic managerial capabilities and organizational readiness in smart city transformation. *Cities, 129*, Article 103791.

Guest, Greg, Arwen Bunce, and Laura Johnson. 2006. "How many interviews are enough?: An experiment with data saturation and variability." Field Methods 18(59).

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PLoS One, 10*(9), Article e0138237.

Hashem, I. A., Targio, V. C., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., … Chiroma, H. (2016). The role of big data in smart city. *International Journal of Information Management: The Journal for Information Professionals, 36*(5), 748–758.

International Telecommunication Union (ITU). (2000). *Y.101: Global information infrastructure terminology: Terms and definitions*.

International Telecommunication Union (ITU). (2021). *ITU-T recommendation database*.

ISO. (2017). *ISO/IEC 30182:2017 - Smart city concept model — Guidance for establishing a model for data interoperability*.

Javidroozi, V., Shah, H., & Feldman, G. (2023). FABS: A framework for addressing the business process change challenges for smart city development. *IEEE Access*, 1.

Javidroozi, V., Shah, H., & Feldman, G. (2019a). A framework for addressing the challenges of business process change during enterprise systems integration. *Business Process Management Journal, 26*(2), 463–488.

Javidroozi, V., Shah, H., & Feldman, G. (2019b). Urban computing and smart cities: Towards changing city processes by applying enterprise systems integration practices. *IEEE Access, 7*, 108023–108034.
Khan, U. T., & Zia, M. F. (2021). Smart city technologies, key components, and its aspects. In *4th International Conference on Innovative Computing, ICIC 2021*.
Khawaja, S., & Javidroozi, V. (2023). Blockchain technology as an enabler for cross-sectoral systems integration for developing smart sustainable cities. *IET Smart Cities, 5*(3), 151–172.
Kieu, T. K., Grattan, K., Goldman, B., Ha, T. T. T., Thi, T. T. T., Pomeroy-Stevens, A., & Bachani, D. (2022). Bringing sectors together in Da Nang, Vietnam: Participatory systems mapping. *Journal of Urban Health, 99*(4), 760–769.
Kosba, A., Miller, A., Shi, E., Wen, Z., & Papamanthou, C. (2016). Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In *Proceedings - 2016 IEEE Symposium on Security and Privacy, SP 2016* (pp. 839–858).
Lim, C., & Maglio, P. P. (2018). Data-driven understanding of smart service systems through text mining. *10*(2), 154–180. https://doi.org/10.1287/Serv.2018.0208
Liu, X., Heller, A., & Nielsen, P. S. (2017). CITIESData: A smart city data management framework. *Knowledge and Information Systems, 53*(3), 699–722.
Lnenicka, M., Nikiforova, A., Luterek, M., Azeroual, O., Ukpabi, D., Valtenbergs, V., & Machova, R. (2022). Transparency of open data ecosystems in smart cities: Definition and assessment of the maturity of transparency in 22 smart cities. *Sustainable Cities and Society, 82*, Article 103906.
Loideain, N. N. (2019). A port in the data-sharing storm: The GDPR and the internet of things. *Journal of Cyber Policy, 4*(2), 178–196.
Maxwell, J. A. (2012). *A realist approach for qualitative research*. SAGE Publications.
Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Journal of Clinical Epidemiology, 62*(10), 1006–1012.
Mukherjee, A., Sahoo, S., & Halder, R. (2022). A blockchain-based integrated and interconnected hybrid platform for smart city ecosystem. *Peer-to-Peer Networking and Applications, 15*(4), 2116–2141.
Mukherjee, S., Gupta, S., Rawlley, O., & Jain, S. (2022). Leveraging big data analytics in 5G-enabled IoT and industrial IoT for the development of sustainable smart cities. *Transactions on Emerging Telecommunications Technologies, 33*(12), Article e4618.
Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). *Handbook of practical program evaluation* (4th ed., pp. 1–864).
Nikiforova, A., & Zuiderwijk, A. (2022). Barriers to openly sharing government data: Towards an open data-adapted innovation resistance theory. *ACM International Conference Proceeding Series*, 215–220.
Nitoslawski, S. A., Galle, N. J., van den Bosc, & Steenberg, J. W. N. (2019). Smarter ecosystems for smarter cities? A review of trends, technologies, and turning points for smart urban forestry. *Sustainable Cities and Society, 51*, Article 101770. https://doi.org/10.1016/J.SCS.2019.101770
Paiho, S., Tuominen, P., Rökman, J., Ylikerälä, M., Pajula, J., & Siikavirta, H. (2022). Opportunities of collected city data for smart cities. *IET Smart Cities, 4*(4), 275–291.
Paskaleva, K., Evans, J., Martin, C., Linjordet, T., Yang, D., & Karvonen, A. (2017). Data governance in the sustainable smart city. *Informatics, 4*(4), 41.
Pomp, A., Paulus, A., Burgdorf, A., & Meisen, T. (2021). A semantic data marketplace for easy data sharing within a smart city. *International Conference on Information and Knowledge Management, Proceedings*, 4774–4778.
Rao, P. M., & Deebak, B. D. (2023). Security and privacy issues in smart cities/industries: Technologies, applications, and challenges. *Journal of Ambient Intelligence and Humanized Computing, 14*(8), 10517–10553.
Souza, Arthur, Jorge Pereira, Juliana Oliveira, Claudio Trindade, Everton Cavalcante, Nelio Cacho, Thais Batista, and Frederico Lopes. 2017. "A data integration approach for smart cities: The case of Natal." 2017 International Smart Cities Conference, ISC2 2017.
Tsampoulatidis, Ioannis, Nicos Komninos, Evangelos Syrmos, and Dimitrios Bechtsis. 2022. "Universality and interoperability across smart city ecosystems." Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) 13325 LNCS:218–30.
Ullah, F., Qayyum, S., Thaheem, M. J., Al-Turjman, F., & Sepasgozar, S. M. E. (2021). Risk management in sustainable smart cities governance: A TOE framework. *Technological Forecasting and Social Change, 167*, Article 120743.
United Nations (UN). (2018). *68% of the world population projected to live in urban areas by 2050, says UN*. UN DESA | United Nations Department of Economic and Social Affairs.
Van Der Hoogen, A., Scholtz, B., & Calitz, A. (2019). A smart city stakeholder classification model. In *2019 Conference on Information Communications Technology and Society, ICTAS 2019*.
Vandercruysse, L., Buts, C., & Dooms, M. (2020). A typology of smart city services: The case of data protection impact assessment. *Cities, 104*, Article 102731.
Vanli, T., & Akan, T. (2023). Mapping synergies and trade-offs between smart city dimensions: A network analysis. *Cities, 142*, Article 104527.
Yang, L., Elisa, N., & Eliot, N. (2019). Privacy and security aspects of E-government in smart cities. *Smart Cities Cybersecurity and Privacy*, 89–102.
Zuo, C., Ding, L., Liu, X., Zhang, H., & Meng, L. (2023). Map-based dashboard design with open government data for learning and analysis of industrial innovation environment. *International Journal of Cartography, 9*(1), 97–113.