**Author(s):** Lnenicka, Martin; Nikiforova, Anastasija; Luterek, Mariusz; Azeroual, Otmane; Ukpabi, Dandison; Valtenbergs, Visvaldis; Machova, Renata

**Title:** Transparency of open data ecosystems in smart cities : Definition and assessment of the maturity of transparency in 22 smart cities

**Year:** 2022

**Version:** Accepted version (Final draft)

# Transparency of open data ecosystems in smart cities: Definition and assessment of the maturity of transparency in 22 smart cities

Martin Lnenicka, Anastasija Nikiforova, Mariusz Luterek, Otmane Azeroual, Dandison Ukpabi, Visvaldis Valtenbergs, and Renata Machova

**Abstract**: *This paper focuses on the issue of the transparency maturity of open data ecosystems seen as the key for the development and maintenance of sustainable, citizen-centered, and socially resilient smart cities. This study inspects smart cities' data portals and assesses their compliance with transparency requirements for open (government) data. The expert assessment of 34 portals representing 22 smart cities, with 36 features, allowed us to rank them and determine their level of transparency maturity according to four predefined levels of maturity - developing, defined, managed, and integrated. In addition, recommendations for identifying and improving the current maturity level and specific features have been provided. An open data ecosystem in the smart city context has been conceptualized, and its key components were determined. Our definition considers the components of the data-centric and data-driven infrastructure using the systems theory approach. We have defined five predominant types of current open data ecosystems based on prevailing data infrastructure components. The results of this study should contribute to the improvement of current data ecosystems and build sustainable, transparent, citizen-centered, and socially resilient open data-driven smart cities.*

**Keywords**: *open data; smart city; transparency; maturity; ecosystem; expert assessment*

## 1 Introduction

Today, many cities worldwide join the smart city concept that relies on using Information and Communication Technologies (ICT) and Internet of Things (IoT) solutions in citizens' everyday lives to improve their quality of life and assist local governments in overcoming challenges in the urban resource' usage, reallocation, and delivery of services. These improvements are achieved through ICT, which transform cities into more sustainable and smart entities (Abbas et al., 2018; Gao et al., 2021; Patrao et al., 2020), using mostly human-oriented citizen-centered development (Nitoslawski et al., 2019). As a result, for a city to be sustainable, it must be smart, as many researchers claim (Bibria and Krogstie, 2017), which is only possible through the sociotechnical transition by adopting technological innovations in a complex social ecosystem (Kroh, 2021).

With time, the definition of a smart city, which at the beginning referred mostly to its technological context, shifted to focus mostly on people, with ICT acting as tools to drive citizen engagement and participatory governance schemes, as the city cannot be truly smart without effectively harnessing the power of its social capital, entrepreneurship, and innovation (Kourtit et al., 2012; Møller et al., 2019; Nitoslawski et al., 2019). ICT can also support transparency, ensure the accountability of decision-makers, and promote the participation of citizens in governance (Bonina and Eaton, 2020; David et al., 2018; Nitoslawski et al., 2019). Tapping to that potential requires citizens and local innovators to be informed, which means that open data and Open Government Data (OGD) are indispensable components in open innovation, community engagement, and smart city development (Mak and Lam, 2021). They are used as reference points in local hackathons, citizen labs, and online platforms allowing citizens to share their opinions on matters relevant to the city's life.

Real-time open data is a major component of smart cities that is critical for facilitating multi-scale urban management and improving other qualities such as adaptability, efficiency, interoperability, flexibility, transparency, and real-time response capacity (Sharifi, 2020). In addition, they are seen as a major enabler for transparency and trust, which is a pre-condition for the development of participative and smart communities. In both academia and practice, it is increasingly recognized that data, particularly open data, can provide clarity on previously underexplored ecosystem processes and dynamics and create more sophisticated modeling capabilities (Nitoslawski et al., 2019). At the same time, recent studies stress that the slow progress in open data initiatives has hindered the movement of smart cities' development lowering their social and economic values (Mak and Lam, 2021). According to Orejon-Sanchez et al. (2022), there is a list of smart city-related actions to be conducted under the commitment to open data in the upcoming years with a strong focus on the availability of these data as open data and a detailed plan to attract technical talent and support the digital formation of its population.

Transparency-by-design, together with participation and collaboration processes, are key concepts in the current debate on open government and open data initiatives in smart cities (Lněnička and Nikiforova, 2021; Lněnička and Saxena, 2021; Neves et al., 2020). Several studies have shown how data reuse can generate promising benefits in the smart city environment (Gupta et al., 2020). Open data are one of the most valuable resources in this context. They can be used as the only source or in combination with other data types to solve civil society problems, improve transparency and close the gap between local government and its citizens (Carrara et al., 2020; Johannessen and Berntzen, 2018). The report by Bertends et al. (2020) showed that more and more Member States of the European Union recognized the potential value of open data, and various open data portals are becoming increasingly supported by robust open data policies.

Therefore, together with the pressures for more transparent and responsive government, open data policies have become increasingly widespread (Berrone et al., 2016; Gao et al., 2021), and open data portals are commonly developed to enable access to relevant data (Bonina and Eaton, 2020; Lněnička and Nikiforova, 2021). Such platforms include closed and open data, which value potential may vary (Gupta et al., 2020; Pereira et al., 2017). Data portals and other platforms usually provide features to search, filter, download, analyze, link, visualize, evaluate, discuss, request, and share datasets. Each city usually has more of these platforms, but their levels of integration and openness can be different (Buchinger et al., 2021; Davies, 2020; Nikiforova and Lněnička, 2021). The dynamics of the diffusion of open data are given by the characteristics of social systems in which they occur. These are technology-enabled and rely on the intensities of information flows (Harrison et al., 2012). The concept of delivering transparency relies on the technical standards of OGD and how they are implemented by corresponding features of data infrastructures (Davies, 2020). For this paper OGD can be defined as data produced by public sector agencies and institutions, freely available on data portals in open formats and under open licenses for everyone to be reused (Lněnička et al., 2021).

Abella et al. (2017) defined a smart city as a public-private ecosystem that provides services to citizens and organizations with strong technological support and considers the economic and social impact on society. According to Abbas et al. (2018), smart cities are characterized by complex systems in openness, heterogeneity, complexity, dynamic work environments, and large-scale nature. In this regard, using an ecosystem approach enables us to understand and clarify the relationships and interactions between components (Bagheri et al., 2021; Caputo et al., 2019; Dawes et al., 2016; Lněnička et al., 2017). Nurmi et al. (2019) reported that this approach inspired new models of public information and data flows, services, and products delivery, in which ecosystems-enabled co-creation is considered a key innovation to increase transparency. Therefore, together with increasing pressures on the adoption of sustainable development goals and practices, the concept of the smart city seems to interlink all efforts and form the ecosystem, in which the expected impacts and values can be achieved for stakeholders (Caputo et al., 2019; Gao et al., 2021; Lněnička et al., 2017; Neves et al., 2020; Pereira et al., 2017).

Researchers identify different ecosystems aiming to provide solutions to specific problems (Ooms et al., 2020); considering the data-centric nature of the smart city (Bagheri et al., 2021; Korachi and Bounabat, 2018; Lněnička and Saxena, 2021), data ecosystems are especially relevant. They are represented by data portals, platforms, and other repositories and data sources providing OGD and other open data, in which citizens and other stakeholders can find datasets and features that enable them to work with these data making them actionable rather than static. The concept of transparency requires openness, accountability, and public participation (David et al., 2018; Davies, 2020; Lněnička and Nikiforova, 2021; Nitoslawski et al., 2019). As reported by Korachi and Bounabat (2018), smart cities need an assessment system to manage and assess the progress of smart city projects and the degree of ecosystem maturity (Danneels et al., 2017). Benchmarks and models should provide a way to identify the development paths of smart cities and enable their comparisons with other cities (Luterek, 2020). Considering these challenges, we aim to explore the transparency maturity of open data ecosystems in smart cities.

Since smart cities are closer to their citizens and provide relevant services accordingly, they should consider the ecosystem, and transparency maturity approaches to prove that they as smart cities are efficient and sustainable. Regarding the current smart city approaches, we inspect different data portals in selected smart cities and assess their compliance with open (government) data transparency requirements. This allows us to define the concept of the open data ecosystem in the smart city context and determine its key components. The assessment results enable us to rank the ecosystems in selected smart cities, identify and discuss their transparency maturity, and introduce different types of open data ecosystems. For this study, we establish and attempt to answer the following Research Questions (RQs):

**RQ1:** What components and relationships form open data ecosystems in smart cities?

**RQ2:** Whether and how open data ecosystems in smart cities comply with transparency requirements for open (government) data?

**RQ3:** What is the maturity level of transparency of open data ecosystems in smart cities, and how can it be assessed?

Most of the large-scale research done on the OGD and data portals refers to the national level and solutions provided by the central government. The contribution of the paper is six-fold: (1) a benchmarking framework to assess the level of transparency of open data ecosystems in smart cities consisting of 36 features has been developed by adapting transparency-by-design framework for open data portals by Lněnička and Nikiforova (2021); (2) the developed framework has been applied to 34 portals representing 22 smart cities, allowing determination of the level of transparency maturity at general, individual, and group levels; (3) four-level transparency maturity model has been defined to allow the classification of the portal as developing, defined, managed, and integrated, thereby allowing to identify key issues to be transformed into corrective actions to be included into agenda and navigate to the set of more competitive portals; (4) the portals concerned have been ranked based on their transparency maturity, thereby allowing more successful portals to be identified in order to be used as an example for improving overall or feature-wised performance by providing recommendations for the identification and improvement of current maturity level and specific features; (5) an open data ecosystem in the context of a smart city has been conceptualized and its key components were determined considering the data-centric and data-driven infrastructure and other components and relationships, using the system theory approach; (6) on the basis of the dominant components of data infrastructure, five types of current open data ecosystems have been defined, thereby opening up a new horizon for research in the area of sustainable and socially resilient smart cities by means of open data and citizen-centered open smart city governance.

To meet the objectives of this study, the paper is organized as follows: Section 2 establishes the theoretical background, section 3 provides the methodology for the research, the initial results, and discussion, as well as limitations, are brought forth in section 4, 5 and 6, respectively, and the paper concludes in section 7 highlighting future directions for research and clarifying the primary contributions of this paper.

# 2 Theoretical background

## 2.1 Concepts – smart city, (open) data portal/platform, and ecosystem

Smart cities have developed mainly due to innovative ICT industries and markets and began using and taking advantage of the IoT, cyber-physical systems, Artificial Intelligence (AI), big data, high-performance computing, and cloud computing to establish a link between each component and layer of a city (Kirimtat et al., 2020; Ramu et al., 2022). They use ICT to improve urban services' quality, performance, and interactivity, reduce costs and resource consumption, and improve relationships between citizens and government (Abella et al., 2017). All these processes and efforts produce large volumes of data that are fully, partly, or not processed (Caputo et al., 2019). Since it is widely agreed among researchers as well as public officials that these data can provide value for the development of the smart city and improve the quality of life of citizens, the principles of open data are recognized to enable the reuse of data (Abella et al., 2015; Abella et al., 2017). This relies on the assumption that the center of these efforts are citizens and their needs, so they can provide relevant feedback on what data should be published and what services should be provided or improved.

Data sources in cities may be divided based on the basic actor and the nature of the process they were obtained with (Arribas-Bel, 2014). These are data from 1) individuals holding location-aware devices, 2) databases used to provide (usually free) services through the internet by web companies, and 3) public and government organizations that release data in an open format. Data platforms are essential infrastructure for data management, innovative data-based services, and smart city initiatives (Buchinger et al., 2021). Most data in smart cities are published under open data principles. Danneels et al. (2017) defined an open data platform as a data services architecture, along with the management of access and (re)use, created to allow third parties to create new value. According to Corrêa et al. (2017), the term data portal/platform can refer to local government endpoints used to disclose data or information. Data portals may vary from a typical implementation aimed at collecting and publishing datasets to provide a one-stop-shop for data consumers, maximizing their reuse, to a simple web page that does not provide any reuse. Citizens and other stakeholders can supply and sell data using data platforms or create new business models on top of them (Bagheri et al., 2021).

The concept of open data, especially OGD and transparency, are closely related terms in which ICT play as much an important role as in the case of smart cities. The current path of OGD evolution is towards open data sustainability and smartness (Gao et al., 2021). Since all processes should be transparent, the design of data infrastructures is a crucial task (Lněnička and Nikiforova, 2021). Johannessen and Berntzen (2018) provided an overview of the available technologies and tools for achieving different forms of transparency in smart cities. The relation between transparency and smart cities is linked to both technology and objectives that should follow the release of government data. Transparency data standards operate at the border between the internal data framework of the state and the public realm (Davies, 2020). Different infrastructures, especially social infrastructure, in the smart city need to be integrated to make it easy to share information between different city services and stakeholders (Dinah et al., 2019). By building data infrastructures, two-way communication channels between involved stakeholders can be established to discuss data collection, management, and use within governance. According to Davies (2020), this can allow citizens to build the data infrastructures that can shape the operation of smart cities and a modern data-driven policy environment. Similarly, ICT-enabled networks of interacting stakeholders can be considered the ecosystem's social infrastructure, the socio-technical pipelines sharing and transmitting data, information, and ideas (Harrison et al., 2012). According to Abella et al. (2015), smart cities can be modeled as data sources and service providers' ecosystems: populated by the main agent, the city, and shared with other stakeholders who will reuse open datasets.

In general, open data ecosystems are built around OGD and other types of open data such as Open Citizen Data (OCD), Open Business Data (OBD), and Open Science Data (OSD), and information and data flows resulting from relations and interactions among involved stakeholders. The term open data ecosystem also encompasses many attributes, which further define its purpose and goals, such as smart, value-creating, sustainable, etc. According to Van Loenen et al. (2021), sustainable and value-creating open data ecosystems need to be user-driven, inclusive, circular, and skill-based. Ecosystems in smart cities are more service-oriented since their main objectives are to transform individuals' lives and improve society's well-being (Ooms et al., 2020).

The impact of open data on smart cities ecosystems was explored by Neves et al. (2020). They highlighted their role in generating and analyzing actionable data and open data management to understand, manage, and plan the city. Gupta et al. (2020) reported that the open data ecosystem should allow the development of data innovation capabilities while facilitating data literacy and considering cross-boundary data collaboration and data access. Open and user-driven innovation ecosystems, consistent with smart city stakeholders' interests and needs, support a data-driven economy where data can help improve policy and business decisions (Berrone et al., 2016). The data-driven smart city ecosystem developed by Abella et al. (2017) comprises (1) the city as the source of data, (2) the citizens as end-users of data and innovative services, and (3) the agents as reusers of data. These ecosystems are framed by existing policies and practices that need to be managed and reconfigured over time to support innovation cultures and citizen interactions (Harrison et al., 2012). To reduce the complexity of data ecosystems, a data management framework that considers all relevant data lifecycle phases is needed to integrate processes, stakeholders, and systems (Sinaeepourfard et al., 2016). It should also help address this ecosystem's challenges, which need to integrate different components and help with stakeholders' communication (Lněnička et al., 2017).

## 2.2 Smart cities benchmarks, indices, and rankings

Benchmarking is a performance measurement process with the top performers in the field as a reference point. In this case, it is conducted by obtaining a benchmark used to assess the city's success in being smart(er) against other cities. Benchmarks are usually presented as indices, rankings, and reports that include analysis of the city's smart characteristics, which level of detail may vary. Various indices and rankings have been presented in recent years (Berrone and Ricart, 2018; Patrao et al., 2020). Most of the benchmarks include indicators aligned with the dimensions defined by Giffinger et al. (2007). However, the structure and components of these benchmarks are still the subjects of debates among city leaders and other stakeholders, as it is difficult to identify sets of parameters acceptable among all cities assessed (Luterek, 2020).

One of the primary challenges for smart city benchmarking is the diversity of concepts that are difficult to translate effectively into a single measurement method. As a result, many of those tools do not progress beyond the infancy stage – their publication stops after one-two editions, and the discussion on smart city assessment tools and their main gaps to improve future methodologies and tools is still ongoing (e.g., Patrao et al., 2020). The most frequently indicated shortcomings of these tools are the lack of recognition of the comprehensive determinants of specific phenomena, the lack of transparency in the process of data collection, their aggregation, and the assignment of weights to individual indicators (Sáez et al., 2020)

Still, there is an obvious need for those benchmarking efforts – they can provide general guidance for planning the city's further development into "smartness" and allow comparison with others, which can be an effective tool to keep various stakeholders involved. Sharifi (2019) critically analyzed selected smart city assessment tools, highlighting their strengths and weaknesses and examining their potential contribution to the development and evolution of a smart city concept. They argue that to develop better strategies, assessment tools should be based on the advancements of smart solutions and big data analytics.

Others emphasize the importance of the human dimension. Cortés-Cediel et al. (2020) analyzed case studies reported by the EUROCITIES network that represents smart initiatives implemented in major European cities. They found that the top smart cities implement initiatives to improve people's well-being and increase the citizens' opportunities and participation and are complemented with initiatives that build and improve cities' technological and physical infrastructures. The report by Carrara et al. (2020) presented 17 different topics, including around 100 indicators for the standardization of city-data.

## 2.3 Smart cities and open data – performance measurement and maturity models

Various performance measurement and maturity models are used to overview the smart city's current state, highlight strengths and weaknesses, and provide city leaders recommendations and guidelines towards its development. Smart cities are also compared and ranked against each other based on these models. Each model consists of different domains (dimensions), phases, and corresponding indicator(s) used to assess the city's performance and maturity. The main difference between these models and benchmarks, indices, and rankings discussed in the previous section is their theoretical orientation and focus on the selected domain(s) and a sample of cities. Global indices and rankings can be considered more established and widely accepted than those models. Performance measurement and maturity models usually aim to introduce, assess, and validate new approaches domains or delimit themselves from existing models. They also provide only limited findings, recommendations, and guidelines for developing smart cities. They usually focus on a small sample of cities and cannot be generalized properly to a larger sample of cities.

Existing research presents various models and assesses different domains of smart cities. Still, the central pillar of these models is data-centric. Many researchers recognize the importance of information and data flows, and the efficient management of this domain is crucial to successfully implementing other domains and smart cities' strategies. Thus, this is one of the main reasons this paper focuses on the domain of the open data ecosystem and its maturity.

Danneels et al. (2017) applied a set of OGD ecosystem dimensions to assess the degree of ecosystem maturity. They considered three OGD platform types – cognitivist, connectionist, and autopoietic, including involved actors and their interrelationships. Bonina and Eaton (2020) explored the management of open data ecosystem processes. They focused on the governance of the demand and the supply side of open data portals by their owners to foster ecosystem development. They evaluated the maturity of the ecosystem in four dimensions: (1) enabling actors, (2) governance intervention, (3) interactions, and (4) dynamics over time. Lee and Kwak (2012) proposed the Open Government Maturity Model for assessing and guiding government agencies that aim to plan and implement open government enabled by social media and other relevant technologies. Five maturity levels represent the model: initial conditions (Level 1), data transparency (Level 2), open participation (Level 3), open collaboration (Level 4), and ubiquitous engagement (Level 5). It informs the government agencies of each maturity level's focuses, capabilities, processes, outcomes, problems, best practices, and metrics.

Torrinha and Machado (2017) identified smart city maturity models and assessed them, considering an approach based on the design principles framework to develop maturity models. They compared each maturity model based on 1) description (purpose, scope, focus), 2) domains, and 3) maturity levels. Their analysis also considered design principles when developing a new model divided into basic principles, the descriptive purpose of use, and the prescriptive purpose of use. Ensuring data quality is an important step to transform a city into a smart city. It is also one of the key categories of data transparency (Lněnička and Nikiforova, 2021). Korachi and Bounabat (2018) proposed a model to evaluate the maturity of a smart city based on the quality of produced and consumed data. The model consists of these domains: connectivity, data center, data analytics, applications, and end-users. Corrêa et al. (2017) assessed selected local data portals in Brazil and how they comply with the OGD principles. They mapped items necessary for the disclosure for active transparency and technical requirements for Brazil's access to information legislation based on the principles of the OGD. Warnecke et al. (2019) developed a web application that allows self-assessment of the maturity level of city authorities in terms of smart city performance and determination of their competitiveness by benchmarking. It allows the exclusion of irrelevant indicators to a city's development priorities without compromising the benchmark function.

# 3 Research methodology

This study is based on (1) the review of key concepts such as smart city, (open) data portal/platform, and ecosystem, smart cities benchmarks and performance measurement and maturity models of smart cities and open data portals, based on a descriptive literature review, (2) developing an experiment design by adapting the benchmarking framework for assessing the compliance of open (government) data portals with the principles of transparency-by-design proposed by Lněnička and Nikiforova

(2021), (3) applying the developed framework to 34 portals that can be considered to be part of open data ecosystems in smart cities, thereby carrying out their assessment by experts in 36 features context, which allows us to rank them and discuss their maturity levels and (4) based on the results of the assessment, defining the components and unique models that form the open data ecosystem in the smart city context.

This section refers to the research methodology given in Figure 1, including methods used in each phase. The further sections explain these steps in more detail.

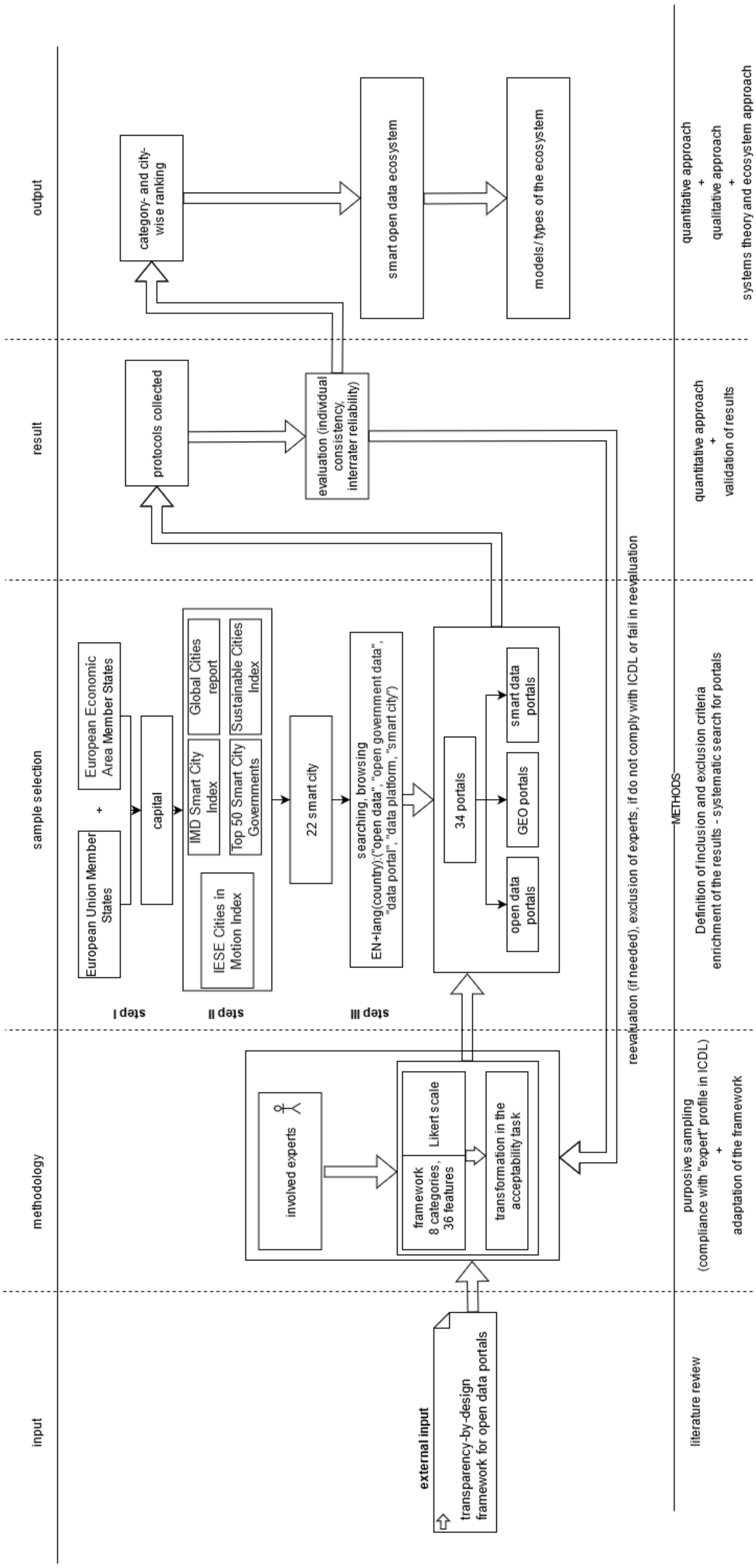


*Figure 1.* Steps of research methodology

## 3.1 Sample selection

There are many smart cities models. The most differences come from the weights assigned to dimensions defined by Giffinger et al. (2007), i.e., smart economy, smart mobility, smart environment, smart people, smart living, and smart governance. However, open data ecosystems rely heavily on transparency and citizen centricity, which the political context can limit. Thus, although some cities can be described as smart when the technological factors are used, the way they can be seen as open differs substantially. As our research is exploratory, in the first step of our sample selection process, we have limited our sample to the capitals of the Member States of the European Union and countries of the European Economic Area, which guarantees a more coherent political and legal framework. That includes citizen centered-approach, the law on the re-use of public data, transparency, and openness, which are often missing in the case of cities from countries located in Asia or South Africa.

The second step was to determine which cities are defined as smart. Due to the wide popularity of the smart city concept, various indices and rankings have been presented in recent years. Some of which are with a global reach (Berrone and Ricart, 2018), while others are limited to specific regions or countries (Patrao et al., 2020). Overall, they are expected to include indicators covering different aspects of the city's activities. Therefore, they can ensure that significant performance indicators are monitored to assess different benefits for different stakeholders and provide an important tool to identify a city's development paths (Luterek, 2020; Patrao et al., 2020).

Each of the cities defined in step 1 was mapped/cross-referenced with their rank in five smart city rankings (Table 1): IESE Cities in Motion Index (Berrone and Ricart, 2020), Top 50 Smart City Governments (SCG) (Top 50 Smart City Governments, 2021), IMD Smart City Index (SCI) (Bris et al., 2020), Global Cities Index (GCI) (Hales et al., 2019), and Sustainable Cities Index (SCI) (Batten, 2018). As a result, six cities have been excluded from our sample: Luxembourg, Valletta, and Nicosia, which have not achieved sufficiently high scores to be included in any of the above rankings, as well as Ljubljana, Sofia, and Zagreb, which were only included in IESE 2020, with very low ranks (99, 116 and 98, respectively).

*Table 1*. Overview of smart cities' positions in selected rankings

| City | IESE 2020 | TOP 50 SCG 2021 | IMD SCI 2020 | GCI 2019 | SCI 2018 |
|---|---|---|---|---|---|
| **Amsterdam** | 8 | 10 | 9 | 20 | 12 |
| **Athens** | 96 | n/a | 99 | n/a | 75 |
| **Berlin** | 7 | 23 | 38 | 14 | 18 |
| **Bratislava** | 62 | n/a | 76 | n/a | n/a |
| **Brussels** | 41 | n/a | 60 | 12 | 47 |
| **Bucharest** | 103 | n/a | 87 | n/a | n/a |
| **Budapest** | 74 | n/a | 77 | 62 | 57 |
| **Copenhagen** | 6 | 35 | 6 | 45 | 11 |
| **Dublin** | 33 | 26 | 34 | 46 | 20 |
| **Helsinki** | 22 | 5 | 2 | n/a | n/a |
| **Lisbon** | 52 | 48 | 75 | n/a | 62 |
| **Ljubljana** | 99 | n/a | n/a | n/a | n/a |

| | | | | | |
|---|---|---|---|---|---|
| **Luxembourg** | n/a | n/a | n/a | n/a | n/a |
| **Madrid** | 25 | n/a | 45 | 15 | 21 |
| **Nicosia** | n/a | n/a | n/a | n/a | n/a |
| **Paris** | 3 | n/a | 61 | 3 | 15 |
| **Prague** | 39 | n/a | 44 | 48 | 23 |
| **Riga** | 85 | n/a | n/a | n/a | n/a |
| **Rome** | 67 | n/a | 101 | 36 | 40 |
| **Sofia** | 116 | n/a | 89 | n/a | n/a |
| **Stockholm** | 14 | 50 | 16 | 39 | 2 |
| **Tallinn** | 55 | 12 | 59 | n/a | n/a |
| **Valletta** | n/a | n/a | n/a | n/a | n/a |
| **Vienna** | 18 | 9 | 25 | 25 | 5 |
| **Vilnius** | 65 | n/a | n/a | n/a | n/a |
| **Warsaw** | 54 | n/a | 55 | 55 | 54 |
| **Zagreb** | 98 | n/a | n/a | n/a | n/a |
| **London** | 1 | 3 | 15 | 2 | 1 |
| **Oslo** | 12 | 27 | 5 | n/a | 8 |
| **Reykjavik** | 5 | n/a | n/a | n/a | n/a |
| **Zurich** | 11 | 45 | 3 | 30 | 6 |

Finally, to identify relevant websites for each city, we have used a purposive sampling method and systematic search for portals using two complementary techniques: browsing and searching. Browsing was initiated on the main website of each city to identify the relevant services in the city's infosphere (open data portals, smart portals, geodata portals, dashboards, etc.). Additionally, we have used a Google search engine and queries for each city consisting of the city name and search key phrases "open data,” "open government data,” "data portal/platform/repository,” and "smart city" in both English and local language, where appropriate. The sampling was conducted in August 2021. It resulted in 34 portals for 22 cities, which depending on their nature, were divided into three categories - open data portals, geodata portals, and smart data portals, to be further examined in this study.

## 3.2 Benchmarking framework

To evaluate the transparency maturity of data ecosystems in smart cities, we have used the transparency-by-design framework (Lněnička and Nikiforova 2021). The framework was developed based on the analysis of related works that deal with open data ecosystems and transparency requirements on open data and deriving the key concepts they found important. The framework is based on a descriptive literature review of studies that deal with the transparency of open data and consists of 8 categories (from now on, dimension) and 47 corresponding features, 36 of which are unique (from now on, sub-dimension) promoting transparency-by-design, where both actors – data portal, data user and data publisher, were considered. This list is intended to constitute the checklist of the open data portal aiming to achieve the highest level of transparency.

We found the respective framework appropriate for this study because it is up to date, based on the extensive literature review and covering major-related features. As its authors state, it extends the scope to be broader than the classical understanding of transparency. In other words, some dimensions do not relate to transparency in a very direct way, focusing more on participation, cooperation, and collaboration, which are sometimes separated from it but are still included as their implementation promotes transparency. All dimensions and sub-dimensions of the framework are equally weighted to make the proposed benchmarking as simple as possible, thereby increasing its chances to be used and maintained despite the person or organization and respective skills. Otherwise, there would be a risk for manipulations to achieve the best or better level in addition to the above-mentioned. Those (sub-)dimensions with the higher weighting are improved, leaving those with the lower "value" for the benchmarking untouched. This, however, would have a negative impact on the sustainability and overall quality of smart city's data portals (reducing their actual value) and stakeholders' satisfaction and intention to use them. The categories with relevant features are presented in Figure 2.

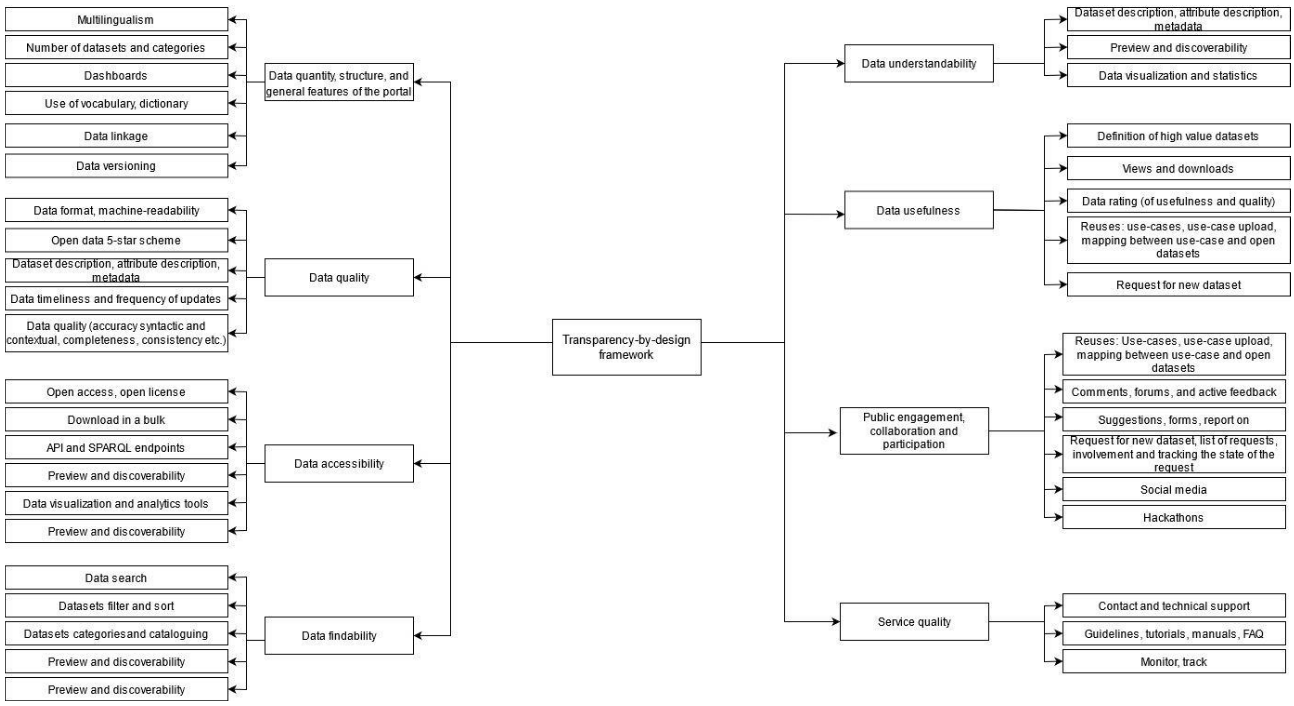


*Figure 2.* Constructs of the transparency-by-design framework

Given our goal to carry out benchmarking, the framework required adaptation, i.e., transformation, to be suited for ranking portals. The benchmarking supposes the collection of quantitative data, which makes this task an acceptability task. Thus, a six-point Likert scale was applied for evaluating the portals. Each sub-dimension was supplied with its description to ensure the common understanding, a drop-down list to select the level at which the respondent (dis)agree, and a comment to be provided, which has not been mandatory. This formed a protocol to be fulfilled on every portal. Each sub-dimension/feature was assessed using a six-point Likert scale, where strong agreement is assessed with 6 points, while strong disagreement is represented by 1 point.

The use of the 6-point Likert scale allowed us to include both levels of (dis)agreement – strong, usual, and neutral opinion. This should potentially allow us to ensure better differentiation of the results. Using more typical 3- or 5- point Likert scales is challenging, as our previous experience shows. The six-point scale encourages participants to consider the question more carefully and choose that deviates either positively or negatively. As Hills and Argyle (2002) found, the equal number of positive and negative items are less sensitive to questionnaire and respondent bias. This choice has been a very popular choice for acceptability tasks. It is recognized to be as such now, in most recent studies in the related research area, i.e., development of sustainable cities and societies (Uchehara et al.,

2022) and identification of factors affecting the seeking and sharing of information on the smart city platforms (Kusumastuti et al., 2022). To avoid the "learning effect," "carry-on," biasing, order effects by which repeated-measures designs are characterized, all stimuli that are portals to be assessed were randomized.

By combining the scores across dimensions and sub-dimensions, it is possible to (1) rank the transparency of portals, (2) carry out the relevant statistical analysis, and (3) direct and explore analysis and discussion at the definite and criteria levels of abstraction as well as derive from the experience gained the components ensuring a competitive and transparency-by-design compliant portal(s) for the smart city. In addition, an opportunity of providing the comment on each feature was provided for cases the evaluator would like to motivate the provided assessment or would like to emphasize something, i.e., good or bad practice, and get both quantitative and qualitative input.

### 3.2.1 Data collection and validation

Each website (portal) was evaluated by experts, where we consider a person to be an expert if a person works with open (government) data and data portals daily, i.e., it is the key part of their job, which can be public officials, researchers, and independent organizations. In other words, compliance with the expert profile according to the International Certification of Digital Literacy (ICDL) and its derivation proposed in (Lněnička et al., 2021) is expected to be met. Considering these requirements, purposive sampling was used. At the beginning of the experiment, 15 evaluators were involved and carried out the assessment. However, examining those protocols revealed that five evaluators failed to perform the reassessment round. Therefore, they cannot be considered compliant with the expert profile we defined above. Thus, their results were excluded from further processing. As a result of the assessment process, we have collected 340 protocols (ten of each expert on each of 34 portals).

Otherwise, 7 of 10 experts hold PhD and are researchers (some are public sector workers at the same time), 3 of them were master students who work in the private sector. Regarding their expertise, 4 represent computer science and information systems, 4 represent business management and economics, 2 represent social and political sciences, where most of them have experience in interdisciplinary studies. Therefore, the sample of experts was heterogeneous and allowed for more objective analysis.

When all individual protocols were collected and processed, mean values and standard deviations (SD) were calculated, and if statistical contradictions/inconsistencies were found, reassessment took place. This allowed us to ensure individual consistency and interrater reliability among experts' answers. For this purpose, the SD between answers gained from two evaluators had to be lower than 1.0 (0.75 for the portal – calculated as the mean value from all answers). If reassessment did not work, i.e., the expert was sure in their results, the evaluator was asked to comment on the assessment. Thereby, we managed to obtain both quantitative and qualitative results to be further explored.

Supporting the principles of open data science, the set of protocols is made available in the ZENODO open repository under the CC-BY license and supplied with the metadata. The assessment took place between June 2021 and October 2021.

### 3.2.2 Interpretation of results

The procedure for assessing the maturity level of the transparency is as follows: (1) perform an assessment of each dimension using sub-dimensions, mapping out the achievement of each indicator; (2) all sub-dimensions in one dimension are aggregated, and then the average value is calculated based on the number of sub-dimensions – the resulting average stands for a dimension value, i.e., eight values per portal; (3) the average value from all dimensions are calculated and then mapped to the maturity level – this value of each portal is also used to rank the portals.

Based on the assessment of maturity models for smart cities performed by Torrinha and Machado (2017) and the data-driven maturity model for assessing smart cities introduced by Korachi and Bounabat (2018), we developed a four-level maturity model to determine the transparency maturity. The model aims to assess the current state of the open data ecosystem in the smart city context and provide improvement guidelines and suggested actions to make progress at all levels. The model's focus is the data-centric and data-driven infrastructure of the smart city. The transparency-by-design framework represents indicators enabling quantification of the current state and corresponding level by allowing scoring from 1 to 6, i.e.,1.5 points for each maturity level. We do not consider the "initial" level of maturity since we already provided initial conditions that must be met to select the smart city for the assessment process. The levels are described in Table 2.

*Table 2.* Transparency maturity levels of the open data ecosystem

| Level | Description | Points |
|---|---|---|
| **Level 1** (Developing) | There are no formal procedures for publishing open (government) data, and the transparency efforts fall to each data provider (publisher). This results in missing relationships between the components of the ecosystem and no or low engagement of stakeholders. | 0–1.5 |
| **Level 2** (Defined) | There are formal procedures for publishing open (government) data. These procedures are defined, documented, and communicated. Although the data infrastructure is implemented, the processes of involving stakeholders to reuse open data are lacking. | [1.5]–3 |
| **Level 3** (Managed) | There are standardized processes to be followed in the open government and transparency vision achievement. The open data ecosystem and its components are mainly automated. Stakeholders are active in the ecosystem and provide feedback to improve it. | [3]–4.5 |
| **Level 4** (Integrated) | Procedures are based on best practices. Components and relationships between them are optimized for the city's environment and the requirements and needs of involved stakeholders, which are constantly being identified and monitored. | [4.5]–6 |

Data and comments collected from the application of the benchmarking framework to portals allow us (1) to assess the current maturity of the selected portals and their ranking accordingly, indicating the most successful and competitive examples, both in general and in the case of a particular (sub-)dimensions, (2) to provide a definition of a general open data ecosystem and its key components in the smart city context, (3) to determine existing types of open data ecosystem for smart cities, and 4) provide recommendations to improve the ecosystem and its potential to reuse data and services. We used qualitative and quantitative approaches to achieve these goals and the systems theory and ecosystem approaches.

# 4 Results and recommendations

The experiment resulted in a set of 340 protocols from 10 experts. They were analyzed from several perspectives. First, a category-wise analysis of the results for (1) open data portals, (2) geodata portals, and (3) all portals was carried out. A city-wised analysis has then been carried out to provide a detailed insight into the results of each examined city its strengths and weaknesses in the context of the applied framework. Finally, a ranking of assessed cities was conducted, thus identifying portals that comply with the transparency-by-design principle the most and should include some corrective actions.

In presenting the results, the scale mentioned above was used, i.e., the level of agreement with the following statements: 1 (Strongly disagree), 2 (Disagree), 3 (Slightly disagree), 4 (Slightly agree), 5 (Agree), and 6 (Strongly agree), where agreement (4.5...6 of 6 points) is visualized in blue and corresponds to maturity level #4, disagreement (0…1.5 points) – red, corresponding to maturity level #1, while slight disagreement (1.5...3) – gray, corresponding to maturity level #2, and slight agreement (3...4.5) – light green, corresponding to maturity level #3. The best and worst results are depicted by dashed borders and result in either blue or red color for the best and the worst results, respectively.

## 4.1 Results by categories

Regarding the open data portals, we have identified that the worst result is demonstrated by features representing the public engagement, collaboration, and participation dimensions. In contrast, the best result belongs to the data findability dimension, followed by data usefulness (see Figure 3). However, the difference between these two results is less than 1,5 points. The best-demonstrated result is 4.6 points out of 6, i.e., tends to be assessed by experts as between agree and slightly agree, i.e., there has been no strong agreement on the fulfillment of these assessed features. As shown in Figure 3, all open data portals assessed have significant room for improvement since none of the categories has achieved a strong positive result. This, however, can also be said about the negative trend, i.e., none of the categories have been assessed by less than 3 points.

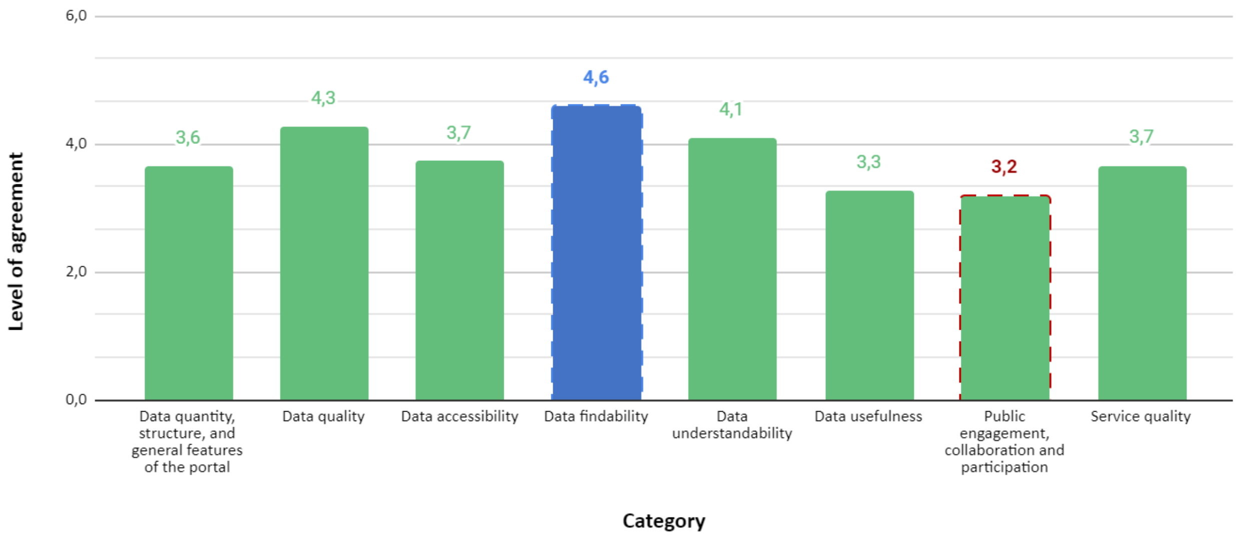


*Figure 3.* Mean values for open data portals (by category)

As regards geodata portals, the results are even worse for all categories assessed, while the trends for best and worst results remain the same. However, if we refer to points that these dimensions have gotten from experts, they are 0.5 points and 0.3 points worse for geodata portals than open data portals. As a result, 2 of 8 dimensions were assessed as poorly implemented, i.e., below 3 points and gray bars in Figure 4. The most significant difference between the results is the "data quantity, structure and general features of the portal," which refers to features such as data organization (the number of datasets and categories, i.e., an overview of the number of datasets and categories, thus demonstrating that the data are organized and available, proving that the portal is active), multilingualism, dashboards, use of vocabulary, data linkage and data versioning, followed by data findability, which refers to data search, filter, and sort, as well as datasets categories and cataloging. The third major difference between the results is the data quality dimension, which focuses not only on the quality of data, such as data accuracy, but also on assessing the dataset description, attribute description, metadata presence, quality, and data timeliness frequency updates. This means that geodata portals are less organized than open data portals. This, however, can be easily explained by the current paradigm, when most of the efforts and resources are spent on the national OGD portal and local portals. In contrast, geodata portals are not meant to be primarily the source of open data. This is also seen in the literature and current benchmarks and rankings, where geodata portals are rarely an object of interest.


*Figure 4.* Mean values for geodata portals (by category)

For the last category represented by smart portals, the results are slightly different from what we have seen before (Figure 5) when referring to the best result now belonging to data quality. The worst result remains valid, although it should be noted that it is even lower compared to the two categories mentioned above. Surprisingly, the data quantity, structure, and general features of the portal, which were very challenging for the geodata portal, have been assessed relatively high (3.8 compared to 3.1), exceeding the result of the open data portals. In addition, better results were achieved by service quality dimension (referring to contact and technical support, guidelines, tutorials, etc., as well as monitoring and tracking features), and data usefulness dimension (which mainly refers to features that allow gaining insight into the result of a service or product creation (reuse/showcase or co-creation), mapping it to the data used, and whether the identification of the most up-to-date trends such as high-value datasets takes place and requests of dataset or service of interest for public).

At the same time, data accessibility, which refers to open access, open license, download in bulk, API and SPARQL endpoints, preview and discoverability, data visualization, and data analytic tools, is assessed worse than the above dimensions. This may be partly because most smart data portals focus on the services and provide a reference from the service to the open data portal. Unfortunately, this is not always the case and sometimes not a reference to the open data portal or dataset, nor a license and preview and discoverability or data analytic tool take place, which significantly affects actionability and, consequently, users' interest in both the service presented and the portal as a whole. Also, the service's transparency and its creation are limited since the users cannot access the data on which the service was built.

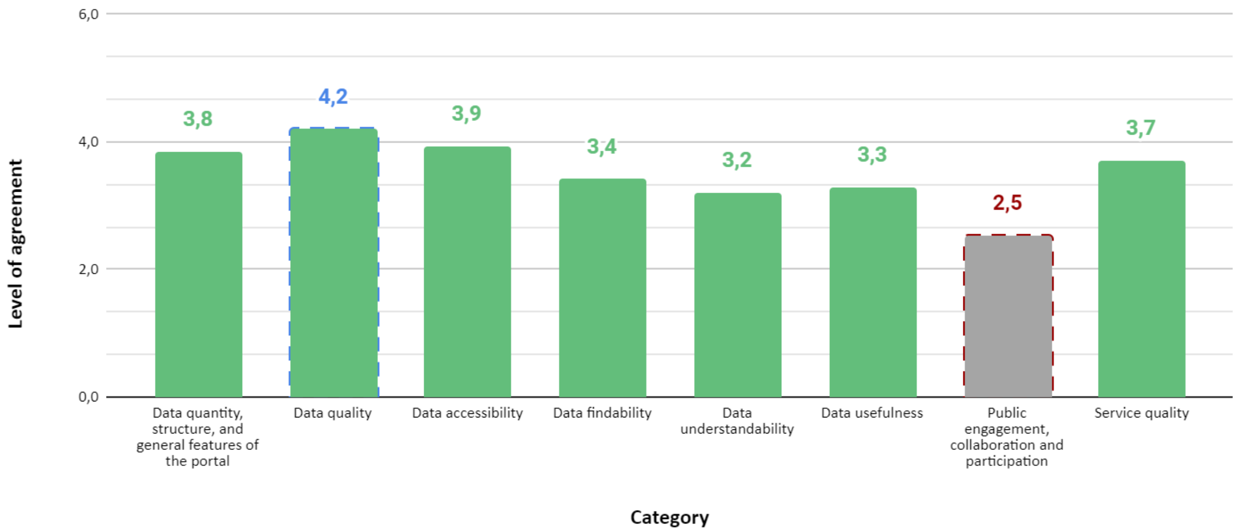


*Figure 5.* Mean values for smart data portals (by category)

When combining the results on all the portals we have analyzed, the overall trend observed for the open and geodata portals is no longer valid, as the best result is demonstrated by data quality dimension with 4.1 points, followed by data findability (4.0 points), as can be seen in Figure 6. However, the most negative result remains public engagement, collaboration, and participation with only 2.9 points, i.e., tending to disagree slightly and can be seen as the most critical. This is followed by data usefulness (3.2), data quantity, structure, and general features of the portal (3.5), service quality (3.6), data understandability (3.7), and data accessibility (3.7), which although have been assessed as partly fulfilled, still have less than 4 points of 6.

This means that changes and improvements should be subject to all dimensions and corresponding features, but public engagement, collaboration, and participation should become central. This is due to such low results and the importance of this category, considering data portals of all types, i.e., open, geodata, and especially smart data portals. Otherwise, if no features support public engagement, collaboration, and participation or the respective features are not well implemented, there are minor changes for any changes, value creation, and meeting the objectives of the initiatives concerned.

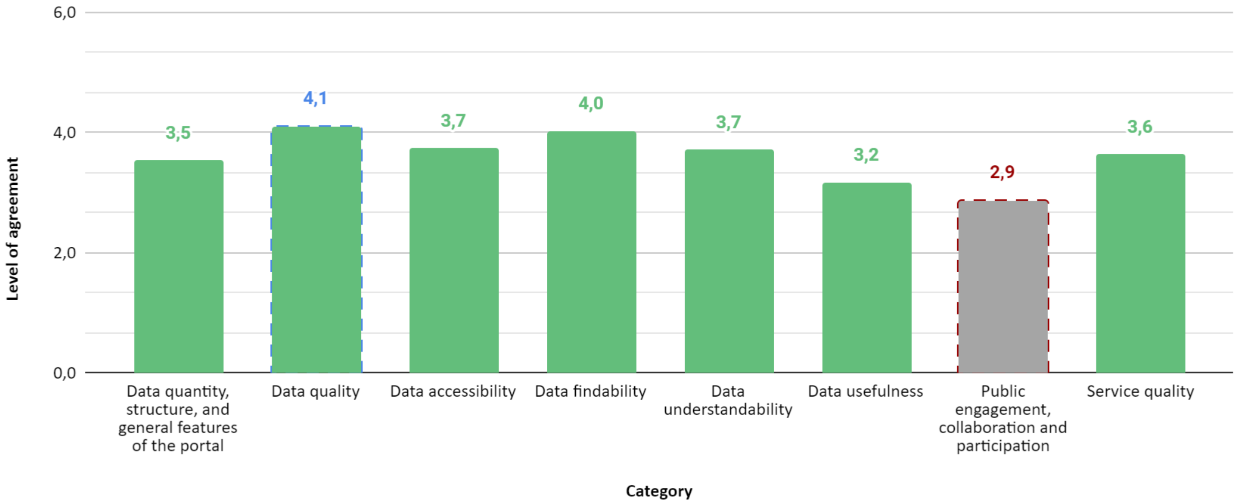


*Figure 6.* Mean values for all portals (by category)

While in this section we have covered the dimension-wise results, let us refer to the results by portal and city.

## 4.2 Results by portals

Although most dimensions with corresponding features were assessed as weakly implemented, i.e., below 4 points, this trend is not valid for all smart cities. More precisely, Figures 7, 8, and 9 provide a city-wise insight on the results, from which we see that the three smart cities, namely Helsinki, Madrid, and Paris, have demonstrated relatively high results with an average result of 4.52 for Helsinki and 4.49 for Madrid and Paris. All portals belong to open data portals. Although the results for the open data portals of Madrid and Paris are expected because the corresponding countries are constant leaders in the context of the maturity of open data (portal), as demonstrated by the recent Open Data Maturity Report (Hesteren and van Knippenberg, 2021), the city of Helsinki is something that is not so self-evident, particularly given that the report mentioned above states that Finland is among six countries, which were moved down from fast-trackers to followers this year. However, this shows that cities and smart cities can develop more rapidly than the whole country when human- and financial resources are allocated wisely.

When different categories of portals are compared, it can be noticed that open data portals tend to demonstrate better results, mostly representing maturity level #3 (Figure 7), while geodata portals (Figure 8) – the worst, mostly falling in maturity level #2, with the best result of 3.9 points for Brussels. In comparison, Brussels' open data portal gained 4.4 points, and the worst assessment was gained by the geodata portal of Sofia and Stockholm with 2.9 points of 6. Open data and geodata portals generally have three and two "outsiders," respectively, tending to be poorly implemented, i.e., gaining below 3 points. In contrast, open data portals have one expressed leader assessed at a little more than 4.5 points and a further three portals with almost similar points. Smart data portals (Figure 9), however, although not showing very strong positive results, were assessed at 3.6 to 3.8 points, with Dublin taking a leading position among the inspected smart data portals. This makes their average results slightly better than the average geodata portals' results, having 3.7 out of 6 points for smart data portals, 3.7 for open data portals, and 3.4 for geodata portals.

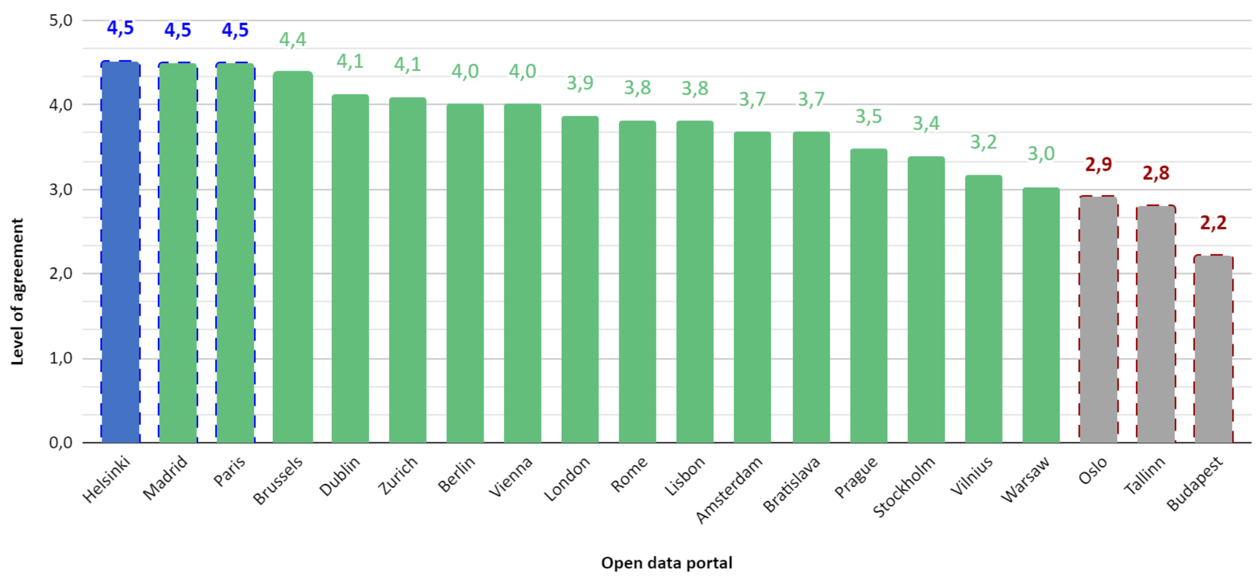


*Figure 7.* Open data portals' ranking

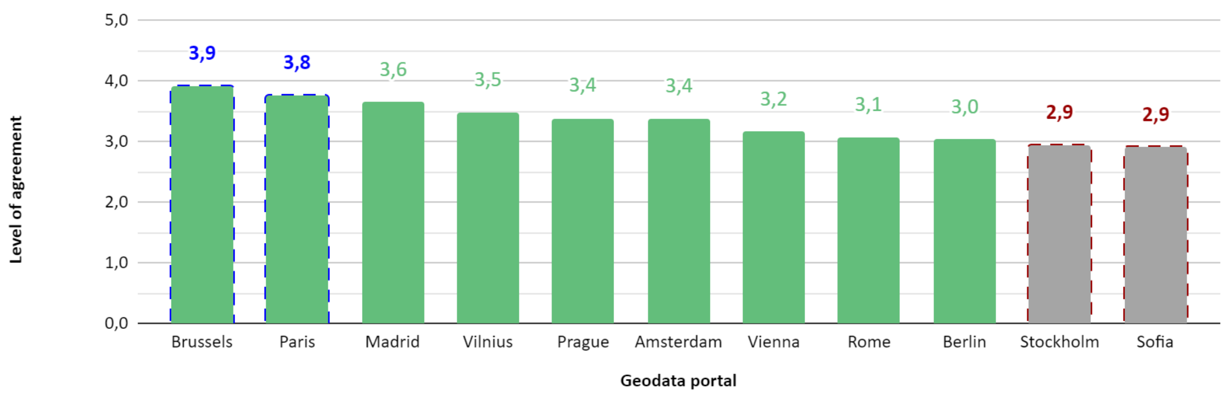


*Figure 8.* Geodata portals' ranking

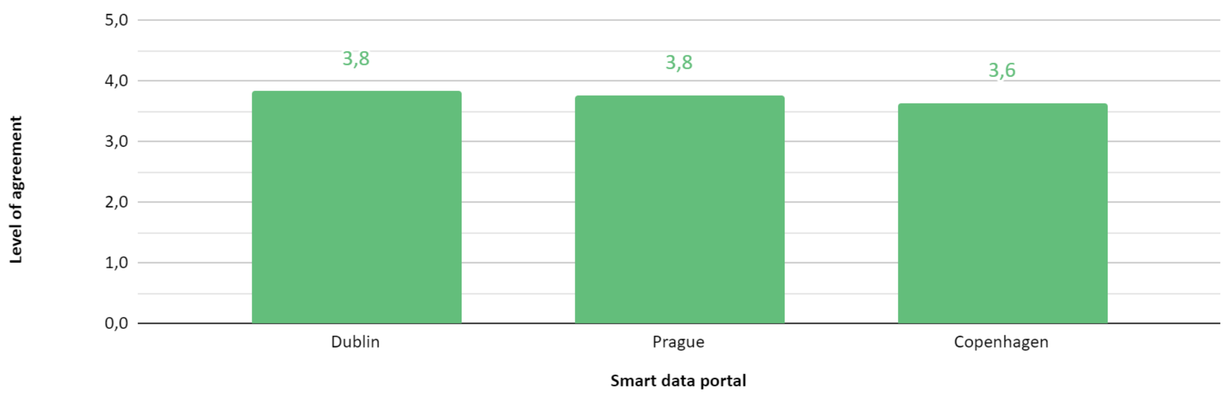


*Figure 9.* Smart data portals' ranking

Figure 10, however, provides an overall ranking of all inspected portals, where open data portals are shown in blue, geodata portals – in green, and smart data portals – in orange. This ranking proves that open data portals are assessed best, with average results demonstrated by smart data portals and mostly weakest results compared to the general results of the above categories shown by geodata portals. None of the portals have gained more than 4.5 points; all have some room for improvements, while five portals have been assessed below 3 points with two geodata portals and three open data portals. Figure 11 then provides a more detailed look at assessed categories for the three best-performing portals and respective smart cities.

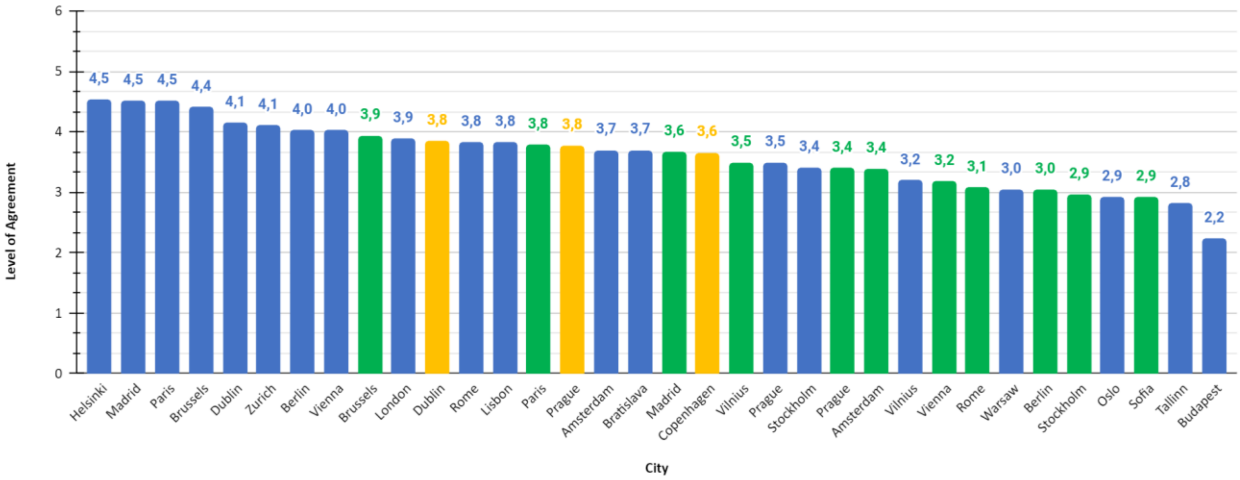


*Figure 10.* All assessed portals' ranking

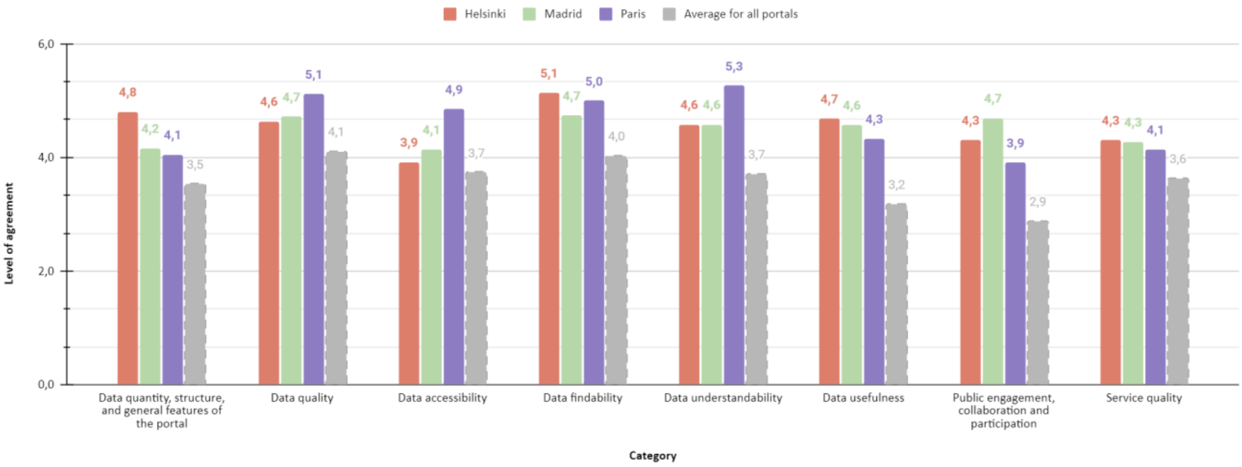


*Figure 11.* Leading portals' results by category

However, although we observed that open data portals are generally assessed higher compared to other data portals, it should be noted that the framework we have applied, i.e., dimensions and sub-dimensions/features assessed, were originally defined for open data portals specifically, which could lead to the above-discussed trend.

## 4.3 Definition of the open data ecosystem in the smart city context

We can derive the open data ecosystem in the smart city context from the above. However, we should first set our findings into the existing body of knowledge on open data ecosystems. Van Loenen et al. (2021) argued that the scale of these ecosystems might vary: within institutions, countries, regions, worldwide, and within different disciplines and domains. Most of the existing definitions are based on the global or country-level views of components and relationships between them, which prevent them from considering other regional and local components and characteristics that can be useful in exploring the dynamics and maturity of the ecosystem. Moreover, they often include a data domain as one of the components at the same level of importance, despite the components of data infrastructures and their full delimitation being the key ones around which other processes should be identified and formalized to achieve ecosystems' goals. Dawes et al. (2016) reported that the components of the city- and municipal-level ecosystems are more evident and easier to analyze compared to more diffuse national systems.

In general, the open data ecosystem, without considering specific levels, disciplines, or domains, can be defined as a set of components that constitute this ecosystem, a set of stakeholders involved and interacting with the system, which should be taken into account and affect those components, and existing environments (economic, social, technical, environmental, cultural, and political) that shape the ecosystem's purpose, goals, processes occurring in it (transparency, participation, collaboration, and cooperation), and services delivering in it.

Since the key asset of the open data ecosystem is the data, the first step in the definition's formulation is to introduce the concept of the data-centric and data-driven infrastructure in the smart city context. This infrastructure can be defined as "*a collection of online data sources providing city-level data for free in open formats and under open licenses for everyone to be reused*." Its main components are data portals, platforms, and other data repositories in the smart city and its administration and other public authorities, which vide OGD, OBD, OCD, and OSD, provided by other stakeholders and freely available for reuse. The data sources cover every data provider who publishes data under open data principles. They do not include institutions with separate tools/policies/approaches to providing data. The key data sources at the smart city level we found are (1) open data portal – publishes OGD and reuses, provides features to work with them, etc., (2) smart data portal – publishes data relevant to smart services and smart projects, (3) geodata portal – publishes spatial data in open formats, provides features and services to work with them, (4) IoT and big data portal – provides raw data and data streams, and (5) domain-specific portals such as smart education, smart transportation, smart energy, etc., usually correspond to domains listed by Giffinger et al. (2007). The last data source type can also be implemented as one platform, such as https://dublindashboard.ie/. The infrastructure dynamics are driven by information and data flow between these components represented by datasets' requests, downloads, processing, sharing, etc. The intensity of these actions limits or enhances the flows in the ecosystem.

The components and relationships can be described using a systems theory approach. This conceptual view enables us to take a second step, in which stakeholders are involved in the ecosystems' dynamics to perform the actions defined above. The data portals vary considerably regarding their orientation towards their primary target groups. Since different stakeholders have different needs, goals, skills, and other characteristics, their interactions and activities result in prioritizing the importance of different actions that lead to the use of different online data sources. Our findings showed that we need to separate users from stakeholders as consumers of this ecosystem. A user can be defined as an individual or group of individuals who may not represent any organization, those who represent it, governments of municipalities, or governments of the country. However, their number and nature should not be limited. According to the concept of Society 5.0 or Super Smart Society, each individual or group of individuals can create added value by utilizing data and digital transformations, thereby applying creative thinking. To effectively and efficiently manage and gain value from the ecosystem, the ecosystem orchestrator (city leaders) must consider these dynamics in their decisions on developing the open data ecosystem.

Further, except for the core components that form the data-centric and data-driven infrastructure of the open data ecosystem, we can identify other components that form the surrounding environment outside the ecosystem boundary but still affect the dynamics of the ecosystem. The first category of these components consists of websites and platforms that do not directly provide open data and features to work with them. Still, support transparency, participation, collaboration, and cooperation processes to enable stakeholders' communication, discussion, and information sharing. For example, https://transparencia.madrid.es/, https://www.partizipation.wien.at/ or https://projektzukunft.berlin.de/. We can also include other digital platforms and channels in this category, such as wikis, forums, microblogging, video sharing, community groups, and other tools to enable citizens and other stakeholders to share their views and opinions. The second category includes online data sources that provide regional and national data, such as the national OGD portal. These sources can also provide city-level data and other services. Still, from the city's point of view, these sources are not usually relevant for the consideration regarding the development of the open data ecosystem in the smart city context. We found a specific type of portal is, a smart cities' platform, that centralizes datasets from more cities in one place and categorizes them according to projects or targeted areas. These components and their relationships are shown in Figure 12.

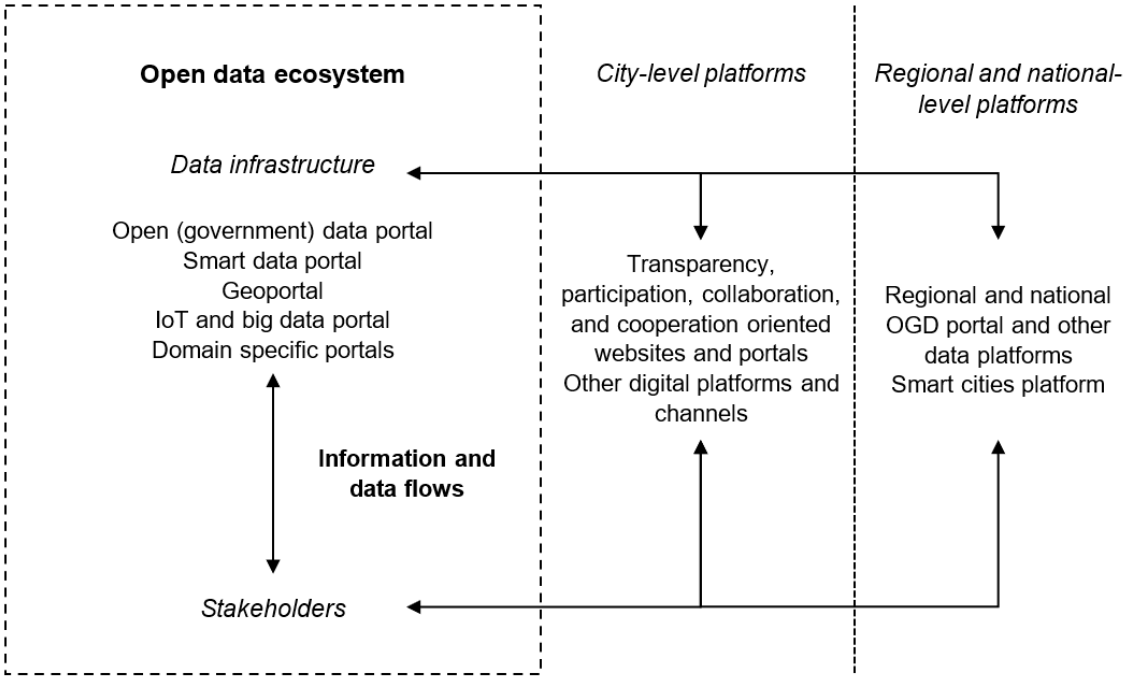


*Figure 12.* Components and relationships of the open data ecosystem in the smart city context

Finally, we found that concepts that should be considered since they affect/shape the ecosystem are: 1) stakeholders and their roles, 2) phases of the data lifecycle, in which a stakeholder participates in the ecosystem, 3) technical and technological infrastructure, 4) generic services and platforms, 5) human capacities and skills of both providers and consumers, 6) smart city domains (thematic categories) as the targeted areas for data reuse, 7) externalities affecting goals, policy, and resources, 8) level of (de)centralization of data sources – development, restrictions, 9) perception of importance and support from public officials, and 10) user interface, user experience, and usability.

Our definition is established based on the knowledge and experience of the experts involved and observations made during the above-described study. The open data ecosystem in the smart city context can be defined as "*systematic efforts to integrate ICT and technologies into city life to deliver citizen-centric, better-quality services, solutions to city problems with open data published through the data-centric and data-driven infrastructure*." It can also be viewed as a part of the transition to the knowledge economy. It is also a part of a local e-government system, and it is usually considered one of the e-government services. Generally, all these approaches to smartness and smart open data services evolved from the concept of e-government and respective websites that have been upgraded to meet the needs of smart cities.

The definition of the open data ecosystem and its description aims to be general and include all components we found. However, some variations of this ecosystem can be identified based on the predominant components of the data-centric and data-driven infrastructure. Our study has identified the following types of ecosystems:

- **Type 1**: The city's OGD portal is the center of the data infrastructure, and all OGD, including those labeled as smart, are published and centralized through it. For this type of open data ecosystem, other websites that had previously provided open data or other services to access public sector information have been replaced by the OGD portal. The focus is on datasets, providing features to work with them, reuse them, and make all data requests transparent in one place.
- **Type 2**: This ecosystem also usually has the OGD portal as the central point, but other portals and platforms publish open data. The smart data portal and online city dashboards focusing on different dimensions such as transport, health, air quality, etc., are important components of this ecosystem.
- **Type 3**: A decentralized type of ecosystem that includes many components such as OGD portal, smart data portal, geodata portal, etc. However, it increases the ecosystem's complexity, which is more difficult to manage and less usable for stakeholders.
- **Type 4**: The smart city portal focused on projects and services is usually the center of this ecosystem, but it is not the priority to provide data and appropriate features to reuse them. Most services are developed by public sector organizations, research institutions, or businesses and provided to citizens.
- **Type 5**: Apart from the city's OGD portal, there are additional transparency-, participation-, collaboration-, and cooperation-oriented websites and portals to support the formation and improvement of relations between stakeholders. This type of ecosystem is focused on processes to improve open data reuse.

## 4.4 Recommendations and best practices

Based on the portals' assessments, we can assign each city a maturity level (following the score scale presented above) and suggest recommendations and best practices to move from the current maturity level to the aimed (higher) one. Our study makes it possible to draw general conclusions and define specific, case-by-case recommendations. The self-determination of recommendations based on the results obtained may be done in at least two ways: (1) the sub-dimensional or dimensional, or (2) portal-wised. The first approach supposes the identification of specific sub-dimensions that have been assessed as poorly implemented and which should be a subject for improvements, their further inspection, and improvement. The latter refers to the identification of better portals, generally or in terms of a specific (sub-)dimension, which can be used by the holders of the relevant smart city portals as an example to be examined, and which best practices to be adopted on the portal concerned. This may include passive and isolated investigation of the portal concerned and the leading one(s) and establishing cooperation with them, where this seems possible, e.g., neighboring cities or countries. However, it is possible to provide more general recommendations, rather than the portal-wised, considering the pre-defined transparency maturity levels and open data ecosystems and the current body of knowledge.

In addition to the (sub-)dimension-wised assessment of portals, the appropriate maturity level can be identified. Depending on the maturity level of the portal, the general recommendations can be defined. They are provided in Table 3. In other words, very poorly performing portals falling at Level #1 of maturity should establish formal procedures for publishing data. In this case, the transparency efforts are likely falling to each data provider, which should be changed by establishing relationships between the components of the ecosystem and by establishing or improving the engagement of stakeholders. In this way, the current gaps and the way to make improvements can be determined, which can and should then be accompanied by feature-wised improvements based on the results of the experts' assessment, as well as identifying actions to be made, e.g., what features affect the engagement of stakeholders.

*Table 3.* Recommendations for improving the maturity level

| **Current and targeted levels** | **Recommendations** |
|---|---|
| Level #1 to Level #2 | ● define formal procedures for publishing open (government) data,<br>● document and communicate these procedures with stakeholders,<br>● establish relationships between the components of the ecosystem,<br>● establish or improve engagement of stakeholders. |
| Level #2 to Level #3 | ● identify and implement actions and activities to involve stakeholders and encourage them to reuse data,<br>● ensure the possibility to provide feedback, collect it and use for defining agenda,<br>● determine the current and improve the level of automation of the open data ecosystem and its components. |
| Level #3 to Level #4 | ● ensure that procedures are based on the best practices,<br>● constantly identify and monitor stakeholders and their needs,<br>● optimize components and relationships between them for the city's environment and the requirements and needs of involved stakeholders. |

Berends et al. (2020) reported that smart city strategies are important drivers for open data, as a more crosslinked city and the use of intelligent devices lead to many useful data that can be used to improve the quality of life in the city. This requires solid data management systems and emphasis on promoting the re-use of these data to release the value they contain. This is important because it facilitates the interoperability between different systems, and data portals can more easily overcome certain barriers by sharing and exchanging best practices and experiences. Thus, the most important step in developing the data-centric and data-driven infrastructure in which all components are provided for the designated purpose (data service), and all are well interfaced with each other.

Our study evinces critical compliance issues with transparency requirements for open government data. Of the eight dimensions, public engagement, collaboration, and participation received the lowest rank, with Bratislava, Budapest, Tallinn, and Oslo among the worst-performing cities. A critical requirement for OGD is to embed the necessary features to encourage public engagement in government activities (Lněnička and Nikiforova, 2021). To achieve open data reuse, data holders and publishers should provide a mapping between use-cases and (high-quality) open datasets and allow and embed features that enhance insights on the data reuse and provide relevant feedback. Two possible ways to enhance co-creation on data portals are through social media (Nikiforova and McBride, 2021) and hackathons (Kamariotou and Kitsios, 2017; Purwanto et al., 2019; Lněnička and Nikiforova, 2021). Social media fosters feedback mechanisms such that it elevates the citizens from passive recipients of government information to active contributors, thus helping to shape government programs and activities, while hackathons have proven to engender citizens' participation in contributing socially and economically to the cities (Kamariotou and Kitsios, 2017; Purwanto et al., 2019; Sigala and Ukpabi, 2019).

The next in the low-ranked dimensions are data usefulness, and Vilnius received the poorest ranking in this category. This points to the fact that the data availability alone is not sufficient to meet the needs of different stakeholders. Available data must specifically meet the critical needs and requirements of the data users. According to our definition of data usefulness, portals should establish mechanisms to determine high-value data (the definition of this term should be made available) and the most demanding data. This should be ensured by collecting requests from stakeholders and managing them transparently, i.e., other users can see the list of requested datasets, comment on and vote for them, and track the request's state. In addition, data insights to the data by incorporating ratings in the metadata so that visitors can easily identify and gain insight into useful data without going through a huge pile of information should be provided. As the number of downloads and views can represent the usefulness of a dataset, appropriate numbers should be provided in addition to opportunities for users to comment on their (dis)satisfaction with the data.

Lastly, improving the portal's quantity, structure, and general features is also important. In particular, Budapest, Oslo, Vilnius, and Warsaw fall behind in this category. Portals should provide opportunities for multilingualism with dynamic features to help data users reduce usage barriers, particularly for non-native speakers of the city data portals. It is also important to provide overviews on the landing pages of the portals. These overviews should highlight the categories of the datasets and their numbers. In addition, portals with dashboards and links to related datasets improve user experience. Similarly, users should be able to track updates where timestamps are especially valuable by ensuring the users' awareness about the data timeliness and currency.

While the above recommendations were tailored towards improving poorly ranked dimensions, it is equally important to indicate that dimensions that performed well need to be maintained and sustained to improve the maturity level. For instance, data findability stands out as the overall best performing dimension, with the cities of Brussels, Dublin, and Helsinki topping the chart. This result demonstrates the ease with which users can locate relevant and related datasets in the mentioned data portals. We thus recommend that other cities, for instance, Budapest, which ranked low in this category, can take a cue from this result and improve that category.

# 5 Discussion

On the level of smart cities and regions, open data publishing and reuse through data portals are relatively new services provided to citizens and other stakeholders (Buchinger et al., 2021; Neves et al., 2020). Thus, without both proactive and reactive open datasets publication and an active citizenry requesting, exploring, and engaging with data, opportunities to develop open data ecosystems in smart cities may be limited (Davies, 2020; Gupta et al., 2020). Similarly, Berrone et al. (2016) highlighted the importance of achieving a culture of openness, referring to both stakeholders – data users and providers. Prieto et al. (2021) developed a framework for prioritizing open data publication and introduced indicators to help choose the most relevant data to publish within the smart city context. Neves et al. (2020) recommended having a performance measurement tool or framework to reduce mismatches between the public needs and expectations and the government transparency on data publication. This recommendation is supported by Ooms et al. (2020), who reported that performance measurement affects the effectiveness of the ecosystem and its maturity.

The existing body of knowledge and our own experience show that the key actions in transparency efforts are participation, cooperation, and collaboration. Following the model introduced by Abella et al. (2017), building open data ecosystems requires implementing two-way, participatory channels dealing with the new data requirements and ideas, providing feedback on inadequate information, and achieving sustainability of these processes. Hivon and Titah (2017) explored the participation of citizens, defined as the activities performed by individuals to use open datasets. They concluded that open data reuse is enabled through four distinct dimensions: 1) hands-on activities, 2) greater responsibility, 3) better communication, and 4) improved relations between citizens and the open data portal development team. The recent studies in the smart city context stress the usefulness of Digital Twins allowing users to interact and report feedback on needed or planned changes (Ramu et al., 2022; White et al., 2021).

All these start with users' interactions with open data portals, where the users should be able to understand and process available data and information. In this regard, transparent smart cities should consider the user-centered design and usability evaluation (Johannessen and Berntzen, 2018; Nikiforova and McBride, 2021). This, however, should be done considering different types of users that might be interested in its use (Lněnička et al., 2021). In this regard, it means that the portal should be adapted to both beginners with a very limited set of skills and knowledge, i.e., low digital literacy, to advanced users/experts not only allowing but also facilitating data use and their transformation into the service or taking part in already launched service, thereby getting involved in co-creation. A multi-perspective knowledge-driven analysis of the OGD portal can be applied (Nikiforova and Lněnička, 2021).

This requires even more actions from portal holders and data publishers than ever before. However, these actions are not very resource-consuming, where the portal holders are required to adapt the current system making it sustainable and transparent, and maintain it. At the same time, data publishers are invited to follow the requirements imposed by the system and provide the support expected. It mainly means that it is not sufficient to make a data artifact available and findable; it should be made accessible, interoperable, and reusable – both FAIR principles should be fulfilled. It should be manageable and actionable, allowing the user to understand it and get involved if interested. Therefore, along with the comprehensive description of the artifact, active support and call for the engagement should be made available. In addition, the city's agencies, departments, and employees should be understood with common tools and information resources that can be used in collaboration activities and to share their experiences and lessons learned (Lee and Kwak, 2012).

However, smart cities may face various challenges in reaching these goals. The comparative case study conducted by Van Loenen et al. (2021) revealed that the open data supply and demand in open data ecosystems are often unbalanced, excluding certain user groups and domains. In addition, they might be linear and lack skill training. Also, despite the importance of features supporting the users and facilitating their engagement, the portal's content is still crucial in a data content sense. The physical world we live in continuously produces huge amounts of data of different types and structures (if any), i.e., big data. These data are used to make data-based decisions and services, where cloud and fog computing are typical solutions (Kirimtat et al., 2020). This trend has now come to OGD since this data becoming available for everyone without restrictions can significantly facilitate the movement towards smart cities, allowing a wider audience to get involved and produce OGD-based solutions and services. Therefore, stakeholders and their engagement in activities supporting OGD reuse are the key challenges those public officials face.

However, although some countries are actively opening their data, others are not as competitive in this sense. This, in turn, refers to the importance of identifying and opening high-value datasets, which, although in many cases may differ from one country to another, in the context of such emergent topics as the smart city will likely be similar for most countries and cities. Data quality is also an important factor that can limit the development and performance of smart cities. Data consumed by cities should be cost-effective and positively impact smart cities outcomes (Korachi and Bounabat, 2018). Effective data preparation was demonstrated by Azeroual (2021) to deal with challenges in providing open data in various data portals and databases. The author introduced two strategies for improving the data quality of OGD: data-driven or process-driven. The first is to directly change data values, such as correcting invalid data or normalizing data. The second pursues redesigning the process of creating and updating to identify and prevent poor quality causes. Nikiforova (2020), in turn, stressed the importance of user-centered data quality analysis, which both individual users can carry out before data are used, checking the compliance of the data and their quality with the task to be conducted, data publishers when preparing data for publishing or data portals.

Among other challenges, according to Sołtysik-Piorunkiewicz and Zdonek (2021) and the literature review they have conducted, Society 5.0 and Industry 4.0 showed open data and open innovation as key factors in the sustainable environment. Gao et al. (2021) emphasized the emerging fields of cloud computing and AI to provide opportunities for governments to employ data analytics based on these approaches. Following these trends facilitates countries and their citizens in creating innovative solutions and, in many cases driving cities to smart cities (Pereira et al., 2017; Slobodova and Becker, 2020). Supporting this idea, Nikiforova (2021) inspected selected OGD portals to understand the relevance of their content to the Society 5.0 expectations.

Another point to be mentioned supported by the results we have presented is that the city's maturity does not necessarily correspond to the open data maturity on the country level, as is the case for Helsinki and Brussels, which demonstrated the best results. In contrast, corresponding countries are not as competitive as the recent Open Data Maturity Report 2021 (Hesteren and van Knippenberg, 2021). This is also the case for Tallinn, i.e., while in the past two years Estonia has managed to improve its results sufficiently and took leading positions, Tallinn has demonstrated very weak results. This, however, can be explained by a not very mature ecosystem established at the country and cities levels. Thus, it can be expected that Tallinn will be able to improve its results in the coming years when the understanding of how the data ecosystem should look like will be established, i.e., this area is currently very new for Estonia, and all forces are now focused on its national OGD portal. These differences are not unexpected and unexplainable because the cities are usually more independent and may have more financial resources, which can be allocated wisely.

These findings also contribute to the discussion between those who support centralization and interconnection of OGD and data infrastructures on all administrative levels with a national data portal as the central point and those who tend to decentralize data sources on local and regional levels. Carrara et al. (2020) reported that not all city portals have their data harvested by the national open data portal. One of the reasons is the city data infrastructure that is not fully developed or suffers data quality issues. In addition, according to Azeroual (2021), the creation of portals and infrastructures is expensive and requires resources, governmental and institutional obligations, and guidelines.

# 6 Limitations

The limitations of this research can be categorized into three groups: 1) conceptual referring to the evolution of the ecosystems and diverse dimensions of transparency and openness; 2) technical referring to the development of data infrastructures and availability of relevant data, and 3) related to data collection (multilingualism). While some of these limitations are study-specific, others are more general and point to the complexity of conducting studies of that nature. Some open issues have emerged from our study and could limit the applicability of the findings.

*Dimensions of transparency and openness*. In this study, we have referred to the framework proposed by Lněnička and Nikiforova (2021). Still, the divergent views on transparency and how it can be measured can affect the results dimensions of transparency, and their definition may vary from one framework to another. For example, Johannessen and Berntzen (2018) identified and categorized six different dimensions of transparency: document, process, benchmarking, decision-maker, meeting, and disclosure transparency. The sample selection process also attempted to mitigate this limitation, selecting cities from the European Union to guarantee more coherent transparency and open framework to which they refer to and aspire to. In addition, the selected framework is based on an extensive study of literature and reports, extracting features that may affect the transparency in both direct and indirect ways, which were further mapped onto the transparency cycle proposed by Janssen et al. (2017). This should potentially reduce the differences between these dimensions significantly.

*Evolution of the ecosystems*. Informational ecosystems change over time, making it more challenging to compare cities at different stages of their development. Ooms et al. (2020) explored the development of elements of smart city ecosystem governance structures and how these elements contribute or inhibit the success of such ecosystems at different stages of development. At the initiation phase, governance structures aim at strengthening internal relations. The ecosystem builds external relations with other parties during the growth phase, such as competitors and suppliers. At this phase, governance elements such as the co-creation strategy and the dedicated promotion organization play a major role in facilitating communication with external parties. The mitigation of this limitation was attempted through the sample selection process.

*Development of data infrastructures.* Similarly, data infrastructures are evolving quickly, which means that the benchmarking results provided can become outdated. To deal with this limitation, like for other reports, a reassessment should be carried out to get the most up-to-date results. To make it possible not only by ourselves but the methodology of the study was also presented in sufficient detail, thereby providing the possibility of both reproduction and replication of this study.

*Availability of relevant data.* Due to privacy, security, and other reasons, not all data are published in open formats. It limits the ability of experts to assess transparency. Ghahremanlou et al. (2019) indicate that cities do not provide open datasets that could be reused, and services could be created in the domains of projects launched by the cities. Also, even if open datasets are collected and published on open data portals, no particular action or project is planned towards this goal. According to Sinaeepourfard et al. (2016), there are various resources and technologies in the modern city for generating data from which these data should be available to facilitate and optimize users' interaction with the smart city. This issue is also closely related to data quality and high-quality datasets, making it difficult to provide datasets that different stakeholders request with different skills and needs (Lněnička et al., 2021).

*Multilingualism*. One of the major limitations in collecting data was the multilingualism of web portals included in the sample. Although the evaluators included experts with different linguistic backgrounds, most of the portals were not fully available in English, making it harder to understand and contextualize information during the collection process. Furthermore, the technology used on some portals made it sometimes impossible to use automated translating tools effectively. This points to the need to ensure that the portal is available in more than the local language (if it is not English), increasing the potential audience at both global and local levels.

Other open issues include especially the dynamic environment in which smart cities must be prepared to provide services to meet the current needs of their citizens and adapt and become resilient. This requires data infrastructures to provide relevant data in real or near-real-time. We recommend considering architectures that can deal with big data analytics and could resolve these issues (Caputo et al., 2019). The platform perspective and the choice of the most suitable types to be deployed in the open data ecosystem affect the portals' target audience and other components of data infrastructures. The actual value for citizens and their skills to reuse data should be considered (Bagheri et al., 2021; Lněnička et al., 2021).

# 7 Conclusions

Smart cities enable and empower citizens and support individual and communal demand for well-being by integrating intelligent technologies with a natural and built environment. However, although cities can develop more rapidly than the whole country when resources are allocated wisely, smart cities' sustainable development cannot be fully achieved and further developed if the data ecosystems on which they rely are not fully open and transparent. In other words, as Kroh (2021) stressed, the development of ecosystems through the sociotechnical transition by adopting technological innovations is crucial to achieving sustainability.

The open data ecosystem in the smart city is a valuable opportunity for providers with available solutions and standardized technology to provide customers with solutions to enable them to link devices from different providers to a large system. These ecosystems open new opportunities for organizations to facilitate access to information for stakeholders and to credibly and reliably justify sustainability reports with raw data, information on the city's development projects, or projects developed by third parties implemented using open data. This helps stakeholders develop new digital solutions, as they can find a lot of open data and already existing digital tools that can be used free of charge to facilitate economic, social, technological, and environmentally sustainable development. However, to meet these objectives, the data ecosystem should be open and transparent, which is the key prerequisite for developing and maintaining sustainable, citizen-centered, and socially resilient smart cities.

This paper, therefore, inspected smart cities' data portals and assessed their compliance with the transparency-by-design framework imposed for open data portals. We have raised and answered three research questions. First, we have identified components and relationships that form open data ecosystems in smart cities. This was done by applying the transparency-by-design framework for open data portals to 34 smart city portals of 22 smart cities and by performing a qualitative analysis of the results obtained while originally carrying out quantitative analysis and benchmarking of portals. Quantitative analysis has shown that all portals assessed have significant room for improvements since none of the 36 sub-categories has achieved a strong positive result. This, however, can also be said about the negative trend, i.e., none of the categories have been assessed by less than 3 of 6 points. The worst result is demonstrated by subcategories representing the public engagement, collaboration, and participation category, while the best result is the data findability category, followed by data usefulness. Most analyzed portals belong to the maturity level #3, which according to our predefined classification of levels of maturity, stands for "managed," i.e., there are standardized processes to be followed in the open government and transparency vision achievement, where open data ecosystem and its components are mainly automated. At the same time, stakeholders are active in the ecosystem and provide feedback to improve it. This has also allowed us to answer the second RQ, within which we have set up a discussion on how these ecosystems meet the transparency requirements of the open (government) data.

And this led us to the third RQ on the maturity of the transparency of these ecosystems and how it can be assessed, which has been answered by proposing the open data ecosystem and its key components. It has been developed based on the contribution mentioned above and extensive literature analysis on the issue. Most reviewed data portals also have a normative orientation towards the idea of a sustainable city and a sustainable environment. The maturity levels also enabled us to compare different smart cities regarding their transparency in making open data freely available through data portals to citizens and focusing on concrete categories to improve. Our definition of the open data ecosystem in the smart city context is unique. Using the systems theory approach, it considers the components of the data-centric and data-driven infrastructure and other components and relationships. The proposed definition establishes a general understanding of this ecosystem. Based on the predominant components of the data-centric and data-driven infrastructure, our study has allowed us to identify five types of open data ecosystems.

To sum up, the contribution of the paper is as follows:

- a benchmarking framework to assess the level of transparency of open data ecosystems in smart cities consisting of 36 features;
- the developed benchmarking framework has been applied to 34 smart city data portals, which allowed determination of the level of the transparency maturity at three-level – general, individual, and group levels;
- a four-level transparency maturity model has been defined, classifying the portal as developing, defined, managed, and integrated. This allowed us to define and provide the reader with the recommendations of identifying key issues with their further transformation into corrective actions to be included into plan and navigating to the set of more competitive portals;
- concerning smart city data portals have been ranked based on their transparency maturity, thereby allowing more successful portals to be identified to be used as an example for improving overall or feature-wise performance by providing recommendations for the identification and improvement of current maturity level and specific features;

- an open data ecosystem in the context of a smart city has been conceptualized, and its key components were determined considering the data-centric and data-driven infrastructure and other components and relationships, using the system theory approach;
- five types of current open data ecosystems have been defined based on the dominant components of data infrastructure.

The latter opens a new horizon for research in sustainable and socially resilient smart cities through open data and citizen-centered transparent smart city governance. This should contribute to improving current data ecosystems and building sustainable, transparent, citizen-centered, and socially resilient open data-driven smart cities. It also makes a call for further studies in this area. Our other papers will introduce a more detailed classification and description of various types of open data ecosystems, their components, and recommendations for their improvements. Also, regarding the weighting of (sub-)dimensions, we plan to explore different weighting schemes and how they could affect the transparency maturity and development of open data ecosystems.

## References

Abbas, H., Shaheen, S., Elhoseny, M., Singh, A. K., & Alkhambashi, M. (2018). Systems thinking for developing sustainable complex smart cities based on self-regulated agent systems and fog computing. Sustainable Computing: Informatics and Systems, 19, 204–213.

Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2015). Information reuse in smart cities' ecosystems. El Profesional de la Información, 24(6), 838844.

Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2017). A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems. Cities, 64, 47–53.

Arribas-Bel, D. (2014). Accidental, open and everywhere: Emerging data sources for the understanding of cities. Applied Geography, 49, 45–53.

Azeroual, O. (2021). Datenqualität und -kuratierung als Voraussetzung für Open Research Data. Information - Wissenschaft & Praxis, 72(4), 204–211.

Bagheri, S., Brandt, T., Sheombar, H., & Van Oosterhout, M. (2021). Value creation through urban data platforms: a conceptual framework. In Proceedings of the 54th Hawaii International Conference on System Sciences (pp. 2464–2473).

Batten, J. (2018). Citizen Centric Cities. The Sustainable Cities Index 2018. Available at: https://www.arcadis.com/campaigns/citizencentriccities/images/%7B1d5ae7e2-a348-4b6e-b1d7-6d94fa7d7567%7Dsustainable_cities_index_2018_arcadis.pdf

Berends, J., Carrara, W., & Vollers, H. (2020). Analytical Report 6: Open Data in Cities 2. Luxembourg: Publications Office of the European Union.

Berrone, P., Ricart, J. E., & Carrasco, C. (2016). The open kimono: Toward a general framework for open data initiatives in cities. California Management Review, 59(1), 39–70.

Berrone, P., & Ricart, J. E. (2018). IESE Cities in Motion Index 2018. Available at: https://media.iese.edu/research/pdfs/ST-0471-E.pdf

Berrone, P., & Ricart, J. E. (2020). IESE Cities in Motion Index 2020. Available at: https://media.iese.edu/research/pdfs/ST-0542-E.pdf

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

Bonina, C., & Eaton, B. (2020). Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo. Government Information Quarterly, 37(3), 101479.

Bris, A., Chee, C. H., & Lanvin, B. (2020). Smart City Index. Lausanne: IMD World Competitiveness Center.

Buchinger, M., Kuhn, P., Kalogeropoulos, A., & Balta, D. (2021). Towards interoperability of smart city data platforms. In Proceedings of the 54th Hawaii International Conference on System Sciences (pp. 2454–2463).

Caputo, F., Walletzky, L., & Štepánek, P. (2019). Towards a systems thinking based view for the governance of a smart city's ecosystem: A bridge to link Smart Technologies and Big Data. Kybernetes, 48(1), 108–123.

Carrara, W., Engbers, W., Nieuwenhuis, M., & van Steenbergen, E. (2020). Analytical Report 4: Open Data in Cities. Luxembourg: Publications Office of the European Union.

Corrêa, A. S., de Paula, E. C., Correa, P. L. P., & da Silva, F. S. C. (2017). Transparency and open government data: a wide national assessment of data openness in Brazilian local governments. Transforming Government: People, Process and Policy, 11(1), 58–78.

Cortés-Cediel, M. E., Cantador, I., & Rodríguez Bolívar, M. P. (2020). Technological and Human Development of Smart Cities: An Empirical Characterization of EUROCITIES Case Studies. In Proceedings of the 53rd Hawaii International Conference on System Sciences (pp. 2293–2302).

Danneels, L., Viaene, S., & Van den Bergh, J. (2017). Open data platforms: Discussing alternative knowledge epistemologies. Government Information Quarterly, 34(3), 365–378.

David, N., McNutt, J. G., & Justice, J. B. (2018). Smart cities, transparency, civic technology and reinventing government. In Rodríguez Bolívar M. (Ed.), Smart Technologies for Smart Governments. Public Administration and Information Technology (pp. 19–34). Springer, Cham.

Davies, T. (2020). Shaping participatory public data infrastructure in the smart city: Open data standards and the turn to transparency. In The Routledge Companion to Smart Cities (pp. 74–90). Routledge.

Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: an ecosystem approach. Government Information Quarterly, 33(1), 15–27.

Dinah, W., Lefika, P. T., & Joseph, B. K. (2019). The role of open data in smart cities: exploring status in resource-constrained countries. In Rodríguez Bolívar M., Bwalya K., Reddick C. (Eds.), Governance Models for Creating Public Value in Open Data Initiatives. Public Administration and Information Technology (pp. 105–121). Springer, Cham.

Gao, Y., Janssen, M., & Zhang, C. (2021). Understanding the evolution of open government data research: towards open data sustainability and smartness. International Review of Administrative Sciences, https://doi.org/10.1177/00208523211009955

Ghahremanlou, L., H Tawil, A.-R., Kearney, P., Nevisi, H., Zhao, X., & Abdallah, A. (2019). A survey of open data platforms in six UK smart city initiatives. The Computer Journal, 62(7), 961–976.

Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanović, N., & Meijers, E. (2007). Smart cities: Ranking of European medium-sized cities. Centre of Regional Science, Vienna UT. Available at: http://www.smart-cities.eu/download/smart_cities_final_report.pdf

Gupta, A., Panagiotopoulos, P., & Bowen, F. (2020). An orchestration approach to smart city data ecosystems. Technological Forecasting and Social Change, 153, 119929.

Hales, M., Pena, A. M., Peterson, E., & Dessibourg-Freer, N. (2019). A Question of Talent: How Human Capital will Determine the next Global Leaders. 2019 Global Cities Report. Available at:

https://www.kearney.com/documents/20152/2794549/A+Question+of+Talent%E2%80%942019+Global+Cities+Report.pdf/106f30b1-83db-25b3-2802-fa04343a36e4?t=1561389512018

Harrison, T. M., Pardo, T. A., & Cook, M. (2012). Creating open government ecosystems: A research and development agenda. Future Internet, 4(4), 900–928.

Hesteren, D. & van Knippenberg, L. (2021). Open Data Maturity report 2021. Available at: https://data.europa.eu/sites/default/files/landscaping_insight_report_n7_2021.pdf

Hills, P., & Argyle, M. (2002). The Oxford Happiness Questionnaire: a compact scale for the measurement of psychological well-being. Personality and individual differences, 33(7), 1073–1082.

Hivon, J., & Titah, R. (2017). Conceptualizing citizen participation in open data use at the city level. Transforming Government: People, Process and Policy, 11(1), 99–118.

Janssen, M., Matheus, R., Longo, J., & Weerakkody, V. (2017). Transparency-by-design as a foundation for open government. Transforming Government: People, Process and Policy, 11(1), 2–8.

Johannessen M. R., & Berntzen L. (2018). The transparent smart city: how city councils and city administrations can apply smart technology for increased transparency - with case evidence from Norway. In Rodríguez Bolívar, M. P. (Ed.), Smart Technologies for Smart Governments. Public Administration and Information Technology (pp. 67–94). Springer, Cham.

Kamariotou, M., & Kitsios, F. (2017). Open data hackathons: a strategy to increase innovation in the city. In Proceedings of International Conference for Entrepreneurship, Innovation and Regional Development (ICEIRD 2017) (pp. 231–238).

Kirimtat, A., Krejcar, O., Kertesz, A., & Tasgetiren, M. F. (2020). Future trends and current state of smart city concepts: A survey. IEEE Access, 8, 86448–86467.

Korachi, Z., & Bounabat, B. (2018). Data-driven maturity model for assessing smart cities. In Proceedings of the 2nd International Conference on Smart Digital Environment (pp. 140–147). ACM.

Kourtit, K., Nijkamp, P., & Arribas, D. (2012). Smart cities in perspective – a comparative European study by means of self-organizing maps. Innovation: The European journal of social science research, 25(2), 229–246.

Kroh, J. (2021). Sustain(able) urban (eco)systems: Stakeholder-related success factors in urban innovation projects. Technological Forecasting and Social Change, 168, 120767.

Kusumastuti, R. D., Nurmala, N., Rouli, J., & Herdiansyah, H. (2022). Analyzing the factors that influence the seeking and sharing of information on the smart city digital platform: Empirical evidence from Indonesia. Technology in Society, 68, 101876.

Lee, G., & Kwak, Y. H. (2012). An Open Government Maturity Model for social media-based public engagement. Government Information Quarterly, 29(4), 492–503.

Lněnička, M., Máchová, R., Komárková, J., & Pásler, M. (2017). Government enterprise architecture for big and open linked data analytics in a smart city ecosystem. In International Conference on Smart Education and Smart E-Learning - SEEL 2017: Smart Education and e-Learning 2017 (pp. 475–485). Springer, Cham.

Lněnička, M., Máchová, R., Volejníková, J., Linhartová, V., Knězáčková, R. & Hub, M. (2021). Enhancing transparency through open government data: The case of data portals and their features and capabilities. Online Information Review, 45(6), 1021–1038.

Lněnička, M., & Nikiforova, A. (2021). Transparency-by-design: What is the role of open data portals?. Telematics and Informatics, 61, 101605.

Lněnička, M., & Saxena, S. (2021). Re-defining open government data standards for smart cities' web-sites: a case study of selected cities. Digital Policy, Regulation and Governance, 23(4), 398–411.

Luterek, M. (2020). Why are rankings of 'smart cities' lacking? An analysis of two decades of e-government benchmarking. In Themistocleous, M. et al. (Eds.), EMCIS 2020 (pp. 238–255). Springer, Cham.

Mak, H. W. L., & Lam, Y. F. (2021). Comparative assessments and insights of data openness of 50 smart cities in air quality aspects. Sustainable Cities and Society, 69, 102868.

Møller, M. S., Olafsson, A. S., Vierikko, K., Sehested, K., Elands, B., Buijs, A., & van den Bosch, C. K. (2019). Participation through place-based e-tools: A valuable resource for urban green infrastructure governance?. Urban Forestry & Urban Greening, 40, 245–253.

Neves, F. T., de Castro Neto, M., & Aparicio, M. (2020). The impacts of open data initiatives on smart cities: A framework for evaluation and monitoring. Cities, 106, 102860.

Nurmi, J., Seppänen, V., & Valtonen, M. K. (2019). Ecosystem architecture management in the public sector–from problems to solutions. Complex Systems Informatics and Modeling Quarterly, (19), 1–18.

Nikiforova, A. (2020). Definition and evaluation of data quality: User-oriented data object-driven approach to data quality assessment. Baltic Journal of Modern Computing, 8(3), 391–432.

Nikiforova, A. (2021). Smarter open government data for society 5.0: Are your open data smart enough?. Sensors, 21(15), 5204.

Nikiforova, A., & Lněnička, M. (2021). A multi-perspective knowledge-driven approach for analysis of the demand side of the open government data portal. Government Information Quarterly, 38(4), 101622.

Nikiforova, A., & McBride, K. (2021). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. Telematics and Informatics, 58, 101539.

Nitoslawski, S. A., Galle, N. J., Van Den Bosch, C. K., & Steenberg, J. W. (2019). Smarter ecosystems for smarter cities? A review of trends, technologies, and turning points for smart urban forestry. Sustainable Cities and Society, 51, 101770.

Ooms, W., Caniëls, M. C., Roijakkers, N., & Cobben, D. (2020). Ecosystems for smart cities: tracing the evolution of governance structures in a dutch smart city initiative. International Entrepreneurship and Management Journal, 16, 1225–1258.

Orejon-Sanchez, R. D., Crespo-Garcia, D., Andres-Diaz, J. R., & Gago-Calderon, A. (2022). Smart cities' development in Spain: a comparison of technical and social indicators with reference to European cities. Sustainable Cities and Society, 81, 103828.

Patrao, C., Moura, P., & Almeida, A. T. D. (2020). Review of smart city assessment tools. Smart Cities, 3(4), 1117–1132.

Pereira, G. V., Macadar, M. A., Luciano, E. M., & Testa, M. G. (2017). Delivering public value through open government data initiatives in a Smart City context. Information Systems Frontiers, 19(2), 213–229.

Prieto, A. E., Mazón, J. N., & Lozano-Tello, A. (2021). Framework for prioritization of open data publication: an application to smart cities. IEEE Transactions on Emerging Topics in Computing, 9(1), 131–143.

Purwanto, A., Zuiderwijk, A., & Janssen, M. (2019). Citizens' motivations for engaging in open data hackathons. In International Conference on Electronic Participation (pp. 130–141). Springer, Cham.

Ramu, S. P., Boopalan, P., Pham, Q. V., Maddikunta, P. K. R., The, T. H., Alazab, M., ... & Gadekallu, T. R. (2022). Federated Learning enabled Digital Twins for smart cities: Concepts, recent advances, and future directions. Sustainable Cities and Society, 79, 103663.

Sáez, L., Heras-Saizarbitoria, I., & Rodríguez-Núñez, E. (2020). Sustainable city rankings, benchmarking and indexes: Looking into the black box. Sustainable Cities and Society, 53, 101938.

Sharifi, A. (2019). A critical review of selected smart city assessment tools and indicator sets. Journal of Cleaner Production, 233, 1269–1283.

Sharifi, A. (2020). A typology of smart city assessment tools and indicator sets. Sustainable Cities and Society, 53, 101936.

Sigala, M., & Ukpabi, D. (2019). Citizen Engagement and Entrepreneurship: implications for sustainable tourism development. In Information and Communication Technologies in Tourism 2019 (pp. 396–407). Springer, Cham.

Sinaeepourfard, A., Garcia, J., Masip-Bruin, X., Marin-Tordera, E., Yin, X., & Wang, C. (2016). A data lifecycle model for smart cities. In 2016 international conference on information and communication technology convergence (ICTC) (pp. 400–405). IEEE.

Slobodova, O., & Becker, S. (2020). Zooming into the ecosystem: agency and politics around open data platforms in Lyon and Berlin. Frontiers in Sustainable Cities, 2, 20.

Sołtysik-Piorunkiewicz, A., & Zdonek, I. (2021). How society 5.0 and industry 4.0 ideas shape the open data performance expectancy. Sustainability, 13(2), 917.

Top 50 Smart City Governments. (2021). Available at: https://www.smartcitygovt.com/202021-publicationdownload

Torrinha, P., & Machado, R. J. (2017). Assessment of maturity models for smart cities supported by maturity model design principles. In 2017 IEEE International Conference on Smart Grid and Smart Cities (ICSGSC) (pp. 252–256). IEEE.

Uchehara, I., Moore, D., Jafarifar, N., & Omatayo, T. (2022). Sustainability rating system for highway design:—A key focus for developing sustainable cities and societies in Nigeria. Sustainable Cities and Society, 78, 103620.

Van Loenen, B., Zuiderwijk, A., Vancauwenberghe, G., Lopez-Pellicer, F. J., Mulder, I., Alexopoulos, C., ... & Flores, C. C. (2021). Towards value-creating and sustainable open data ecosystems: A comparative case study and a research agenda. JeDEM-eJournal of eDemocracy and Open Government, 13(2), 1–27.

Warnecke, D., Wittstock, R. & Teuteberg, F. (2019). Benchmarking of European smart cities – a maturity model and web-based self-assessment tool. Sustainability Accounting, Management and Policy Journal, 10(4), 654–684.

White, G., Zink, A., Codecá, L., & Clarke, S. (2021). A digital twin smart city for citizen feedback. Cities, 110, 103064.