DATA PROTECTION AUTHORITY

Authorisation and Opinion Service

Report on the study day

“Data protection in smart cities – Focus on mobility – How to collect data, anonymize it, extract insights while preserving privacy, and make it available to relevant stakeholders?”

1. Introduction

1.1 Objective

In line with the priorities set in the Belgian Data Protection Authority's 2023-2024 action plan, a study day titled "Data Protection in Smart Cities with a Focus on Mobility - How to Collect Data, Anonymize It, Extract Insights While Preserving Privacy, and Make It Available to Relevant Stakeholders?" was organized on March 1, 2024.

The aim* of this study day was to provide a platform for discussion, where diverse actors could share their experiences, best practices, challenges, solutions and visions for the future of "smart cities". While the audience of the conference consisted mainly of specialized legal and technical experts active in the field, this report[1] aims to be accessible for both smart city experts, as well as citizens. It provides an overview of the discussed topics, key findings and refers to additional recordings made by several speakers.

*The aim was not to assess, nor endorse or criticize any particular project, but to draw general conclusions on the smart city actor's challenges.

1.2 Summary of the key takeaways

Hereunder is an overview of some of the noteworthy aspects highlighted by the participants during the study day concerning the smart city stakeholders.

"Smart city stakeholders" refers to diverse actors, such as developers, researchers, citizens, local administrators, investors, organisations, policy makers, scientists and others, that are required for the planning, implementation and operation of smart city projects.

Study day Key Takeaways - Focus on controllers and processors' duties	
Early involvement of Data Protection Officers (DPOs)	**Ensure the Rights of Citizens**
■ Commitment to data protection can be showcased by involving DPOs early in processes. ■ Information to DPOs related to smart cities projects should encompass all the relevant GDPR and privacy aspects of the projects ■ Enhancing collaboration between DPOs from both private and public sectors.	■ Citizens should be made aware of the sensitivity of mobility data. ■ Right to transparency, informed consent, and access regarding smart city data, including their metadata.
Data Protection & Public Procurement	**Obtain Citizen Engagement**
■ Data protection to be taken seriously into account	■ Enable citizens to play a role in accountability.

[1] This report was drafted by the Secretariat of the Authorisation and Opinion Service, and reviewed by Mrs. Cédrine Morlière (Director), Mr. Bart Preneel and Mr. Yves-Alexandre de Montjoye (external members of the Authorisation and Opinion Service).

■ Early DPO involvement and substantive broadening of evaluation criteria (regarding data protection) in public procurement.	■ Collect meaningful consent where necessary. ■ Ensure adequate transparency including where processing is framed by law.
Data processing (collection, analyses, decisions, etc.)	**Data Architecture and Security**
■ Minimize data to necessary levels, especially location data. ■ Deploy pseudonymisation, anonymisation and other PETs (Privacy Enhancing Technologies) whenever possible ■ Clarify purpose of processing and repurposing data in public-private partnerships. ■ Submit DPIA to Belgian DPA under art. 35-36 GDPR (processing resulting in a high risk).	■ Avoid single point of trust that could become a single point of security failure ■ Take into account the interplay between data architecture, security, privacy requirements and available/adequate privacy enhancing technologies.

1.2.1 Background "smart cities" and focus on smart mobility data

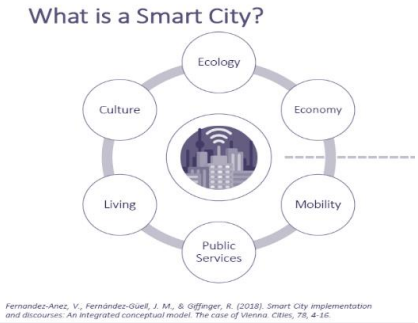
Cities are increasingly collecting data for many aspects of citizens' life. The interaction between people and urban environments is being transformed by innovative technological solutions and is driving the transition from traditional cities to smart, data-driven cities.

"A 'Smart City' *is a multi-stakeholders' ecosystem composed with local governments, citizens' associations, multinational and local businesses, universities, international institutions, …".*

- Audrey Lebas -

"Smart city" projects are based on data processing for many aspects of citizens' lives: economy and environment (e.g. resource management), people, governance (digitisation of administration), living (e.g. management of housing policies) and mobility.

It is, therefore, safe to say that smart cities are inherently interlinked with personal data and the data subjects.



(Extract from the presentation of Prof. Simonofski (University of Namur)

Article 4(1) of the GDPR defines "personal data" as being *any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.*

Due to the integration of advanced technologies in urban areas, some speakers outlined that a lot of smart cities projects could contain location information. For instance to optimize transportation systems, sensitive personal data such as travel patterns and location information are inevitably collected. Consequently, this generates concerns related to data security and privacy which are essential to ensure citizens' trust and compliance with smart mobility solutions.

The question arises whether **mobility data**, such as journeys, do not deserve a specific protection, as they enable to infer very sensitive information about the citizen (e.g. health status or religious preference through the places visited).

As illustrated by several speakers, mobility data are very hard if not impossible to anonymise, and it is likely to reveal very sensitive and granular information about the citizen.

Should mobility data be treated differently and with more caution than other smart city data? And if so how? To what extent should a city be smart and process mobility data while preserving the citizen's privacy?

1.2.2 Overview of both the study day and the additional recordings

It is essential to note that participants, both from private and public sectors, engaged on a voluntary basis, sharing aspects of their projects they deemed pertinent for discussion, and did so with permission to disclose general information to the public. Each speaker volunteered their insights and initiatives independently. The BE DPA (Belgian Data Protection Authority) does not endorse, nor criticize any specific project or initiative discussed during these sessions. The objective of the BE DPA is to draw general conclusions and insights from the collective discussions, not to provide assessments or endorsements of individual projects. Therefore, recordings of these discussions should not be construed as endorsements or assessments of the content presented, as the aim is to derive general insights rather than evaluate specific projects.

Several participants also provided recordings which cover a broad spectrum of discussions on smart cities data and privacy challenges, which are available on our website.

Keynote

Cédrine Morlière, President of the Belgian Data Protection Authority and Director of the Authorisation and Opinion Service, introduced the debates with a keynote speech tackling the rich and evolving regulatory landscape and the specific protection applying to traffic and location data.

"Be it generated by vehicles, by the city's cameras and sensors, or by the citizen themselves using geolocation applications, mobility data such as journeys enable to infer the place of work and of residence, as well as a citizen's centres of interest. It may possibly reveal sensitive information such as religious preference or health status through the places visited. Any disproportionate collection of mobility data could also create a sentiment of general surveillance from the viewpoint of the right to privacy."

Mobility data and privacy challenges

- **Yves-Alexandre de Montjoye ,** Prof. Imperial College of London, *Privacy risks of mobility traces*
- **Markus Sperl**, Project manager, Technologiestiftung Berlin, *Smart city and mobility data in Berlin*
- **Peter Lewyllie**, Project engineer, Agentschap Wegen&Verkeer, *Mobilidata: sharing road user data in Flanders - Experiences and challenges in data protection and privacy*

This technical panel aimed to explore how to reduce privacy risks from a technical point of view (pseudonymize, minimize data, go towards anonymization), with a focus on specific challenges regarding mobility data. Speakers shed light on the privacy risks associated with mobility traces, emphasizing the importance of safeguarding individuals' privacy amidst the vast troves of data generated by mobility patterns. In addition, insights were provided into the utilization of smart city and mobility data in Berlin, offering a comprehensive overview of its applications and implications. Participants shared valuable experiences and challenges encountered in sharing road user data in Flanders, underscoring the critical importance of data protection and privacy measures.



Smart cities data and privacy challenges for citizens

- **Anthony Simonofski**, Prof. Université de Namur - Trustworthy AI in Smart cities: discussing the role of citizens
- **Paul-Olivier Dehaye**, CEO Hestia.ai, member of PersonalData.IO – Smart cities and data politics: a citizens' perspective from Geneva
- **Marie-Charlotte Roques-Bonnet**, Senior Data Protection Legal Advisor & Researcher in Data Protection engineering (ENISA, IAPP, EDPB Individual Support Pool Expert, ID side project) – No Smart cities without smart Privacy compliance: from targeted risks' assessments to efficient Privacy Management Programs (by video)
- **Paul de Hert,** Prof. VUB, Should cities be smart?

On the subject of smart cities data and privacy challenges for citizens: participants explored the vital role of trustworthy AI in smart cities. They discussed the role of citizens in smart cities, more specifically how citizens can participate in the smart city architecture (open data consumers; sensors or democratic participants), from a socio-technical perspective. Through case examples insights were provided from Geneva on citizen perspectives and smart city or smart mobility data control (how to facilitate citizen's access to data). Privacy-enhancing measures were highlighted for the evolving smart city data landscape. The notion of "smart cities" was analysed to answer one core question "should cities be smart?


Smart cities and European data spaces

- **Malte Beyer-Katzenberger**, Team leader, European Commission, DG Connect – *Privacy in the age of data spaces*

Through the analysis of four core questions relating to the smart city architecture, the individuals, key criteria and benefits, the speaker discussed the concept of data spaces as governance frameworks for voluntary data exchange, promoting data reuse and collaboration. Decentralization, the role of third-party intermediaries is put forward for trust, and the potential of data spaces to balance power dynamics and support smaller players.

Smart cities and mobility data – challenges for public and private players

- **Raf Buyle,** Innovation lead, Athumi, Building new data-ecosystems on top of data-collaboration
- **Caroline Vandenplas**, Managing Partner, B12 Consulting , Automatic anonymization of unstructured data: dream or reality?
- **Davor Meersman,** CEO, Future Craft Habitats**,** Whose twin is it anyway? Towards a digital commons for post-competitive data ecosystems

Discussions led by smart city actors explored the challenges faced by both public and private players in navigating the complexities of smart cities and mobility data. Topics ranged from building new data ecosystems to attempts at the automatic anonymization of unstructured data and the quest for a digital commons used for legitimate public purposes and the use of digital twins.

Smart cities and mobility data - data governance & procurement

- **Audrey Lebas**, Researcher, Smart city Institute HEC Liège, University of Liège, The role of data in Smart Mobility - sharing good practices from a governance point of view
- **Karl-Filip Coenegrachts**, Head of Unit, Data, Governance and Communities, VUB, Citcom.ai – a Brussels mobility data project
- **Laurens Vandercruysse**, Senior researcher, VUB, Data protection in smart cities, incentive structuring in procurement processes

On the subject of data governance and procurement, participants shared insights on project cases to illustrate governance perspectives on smart mobility data. They addressed data protection challenges in public procurement processes and good practices for a durable and efficient privacy protection procedure.


Closing remarks

- **Bart Preneel,** Prof. University of Leuven, COSIC Research group

Finally, Professor Bart Preneel, from the University of KU Leuven, concluded the day with general conclusions and noteworthy takeaways.

"It is all about power in the society. Technology is changing power relations and if we make our community smart, we may change power relationships."

1.2.3 Which actions has the Belgian DPA taken so far on this topic ?

The Authorisation and Opinion Service[2] has for mission, on its own initiative or at the request of governments or parliaments, to:

- provide opinions on all issues relating to the processing of personal data (including in the context of preparing draft normative texts).
- make recommendations on social, economic and technological developments that may affect the processing of personal data.

[2] Previously named the 'Knowledge Centre'.

Opinions	Guidelines
Over the past few years, several opinions on **data-driven projects related to smart cities, smart regions and smart mobility** were delivered. The objective was to review whether the general principles of personal data protection, such as lawfulness, fairness, data minimization and transparency, were taken into account.[3] Some examples are: ■ **Opinion 11/2024** on the use of ANPR cameras for the purpose of pollution control in Flanders [FR and NL] ■ **Opinion 273/2022** on the roll out of smart meters or on the monitoring of road traffic for trucks in the Walloon region [FR[4]] ■ **Opinion 186/2021** on SMART MOVE in Brussels, related to an intended kilometre tax [FR and NL]	The Authorisation and Opinion Service is contributing to the drafting of the upcoming revised **guidelines at the EDPB level on how to effectively anonymise or pseudonymise personal data [Opinion 05/2014 on Anonymisation Techniques].** Anonymization and pseudonymization are important technical measures to consider in the context of smart city and smart mobility projects.

The Litigation Chamber is the administrative dispute body of the BE DPA (Article 32 WOG) and is tasked with taking enforcement decision in cases referred to it, based on a complaint from a citizen or following an inspection on the BE DPA's own initiative.

One of the investigations related to smart city's projects, is a decision dating back to 2021, and which **concerned smart cameras** deployed at the Belgian Coast in order to count passers-by during the Covid crisis [24/2021 FR and NL]. The smart cameras were found to be compliant with the GDPR rules on privacy by design and minimization.

1.2.4 Methodology and further literature

Methodology of the study day

This study day aimed to address some key questions inspired by the Working document on smart cities adopted in 2023 by the Berlin Group (International Working group on Data Protection in Technology)[5], such as how to:

- minimize data collection to necessary levels in smart city projects;
- reidentification risk of mobility data;

[3] The data minimization and proportionality of the processing will be further explained in Section 2.3

[4] Translation ongoing.

[5] The Berlin Group is an international working group dedicated to data protection in the field of technology, created in 1983 at the initiative of various national data protection authorities. The group's secretariat is provided by the Berlin Data Protection Authority (Berliner Datenschutzbeauftragten). Participation in the group is not limited to national data protection authorities; it is also open to representatives from the private sector and NGOs. See: https://www.bfdi.bund.de/SharedDocs/Downloads/EN/Berlin-Group/20230608_WP-Smart-Cities.html?nn=355094.

- enable citizens to play a role in keeping the smart city projects accountable for the data they process;
- collect meaningful consent.

In this context, the following sub-questions were put forward to the various speakers, be it smart city or smart mobility actors:

Smart city actors

- How to collect data, pseudonymize/ anonymize it, extract insights and make it available to relevant stakeholders?
- How are smart city data/smart mobility data structured and how can access be safely gained to them?

Smart mobility actors

- Should mobility data be treated differently and with more caution than other smart city data? And if so how?

The purpose of this study day was not to detail all the legal rules and technicalities relating to various projects, but rather to provide participants with the opportunity to address one or more of these issues and/or set out their own challenges and questions regarding the protection of personal data in the context of smart city projects.

It was clearly announced that the intent of the conference was not to assess nor endorse or criticize any particular project or presentation, but to increase the BE DPA and the audience's general understanding of challenges face by smart city actors and by the citizen.

Further literature

For a more thorough exploration of the subject matter, readers are encouraged to refer themselves to the following research results for further explanation and insight.

Several DPA's have already provided insightful general guidance on the data protection aspects of smart cities, including our French, Dutch and Spanish counterparts. EDPB also issued ambitious guidelines on connected vehicles and mobility related applications, dating back from 2021.

More recently, the International Working Group on data protection and technology, referred to as the 'Berlin Group', adopted a Working Paper on Smart cities in September 2023, which formed the framework structure of the study day.

The Authorisation and Opinion Service draws inspiration and builds upon the methodology of the Berlin Group's working Paper[6] to deliver further clarifications based on the discussions held during the study day.

The importance of anonymization as a technique to share personal data safely and the difficulty to anonymize location data were key considerations underlying the study

[6] International Working Group on Data Protection in Technology, Working Paper on "Smart Cities", adopted 29th-30th November 2022, last accessed on 25th of July 2024, https://www.datenschutz-berlin.de/fileadmin/user_upload/pdf/publikationen/berlin-group/2023/20230608_WP-Smart-Cities.pdf.

day, as summarized by panellist **Yves-Alexandre de Montjoye** in the following co-authored papers :

- **"Anonymization: the imperfect science of using data while preserving privacy"**[7]

Information about us, our actions, our location and our preferences is created at scale through surveys or scientific studies or as a result of our interaction with digital devices such as smartphones and fitness trackers. The ability to safely share and analyse such data is key for scientific and societal progress. Anonymization is considered by scientists and policymakers as one of the main ways to share data while minimizing privacy risks. This review offers a pragmatic perspective on the modern literature on privacy attacks and anonymization techniques (traditional de-identification techniques and their strong limitations in the age of big data). The authors turned their attention to modern approaches to share anonymous aggregate data. They found that, although no perfect solution exists, applying modern techniques while auditing their guarantees against attacks is the best approach to safely use and share data today.

- **"A Zero Auxiliary Knowledge Membership Inference Attack on Aggregate Location Data »**[8]

This paper shows that individuals can be singled out from aggregate location data without even any need to crossmatch the data with other auxiliary datasets. An algorithm can create a synthetic datasets based on aggregate location data in order to single out individuals among those data.

2. Key takeaways on the challenges put forward by participants

Although the discussions covered various subjects, some topics were present in all presentations, such as data minimization, transparency, consent, the governance and architecture of smart city initiatives.

2.1 The challenge of data minimization – Privacy Enhancing Technologies (PETs) and current anonymisation/pseudonymisation framework

Data minimization should be understood as an obligation to ensure that personal data is adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed. This principle of data minimization gives expression to the *principle of proportionality* and is operationalized through the implementation of *data protection by design and by default* (Article 25 of the GDPR).[9]

These principles are to be applied from the outset, *i.e.*, before the processing.

---

[7] *A. Gadotti, L. Rocher, F. Houssiau, A-M. Creţu, and Y-A de Montjoye. 'Anonymization: The Imperfect Science of Using Data While Preserving Privacy." Science Advances 10, no. 29 (2024): eadn7053-. doi:10.1126/sciadv.adn7053.*

[8] V. Guan, F. Guépin, A-M. Cretu and Y-A. de Montjoye, *Zero Auxiliary Knowledge MIA against Aggregate Location Data,* 2024, *arXiv.Org*. https://doi.org/10.48550/arxiv.2406.18671.

[9] CJEU June 2021, nr. C-439/19, *Latvijas Republikas Saeima (Penalty points)*, EU:C:2021:504, para. 98; M. Krzysztofek, *GDPR,* Wolters Kluwer Law International, 2021, 65.

Participants overall acknowledged the importance of implementing these principles from an early stage. A particular attention was given to mobility data, anonymisation, pseudonymisation and other PETs[10].

2.1.1 The specific risks of mobility data

In the context of smart cities and in particular smart mobility, information [will be] collected from various sources, such as smartphones, GPS devices, and transportation systems, [providing] insights into people's movement patterns and transportation behaviours.

This information is often referred to as '*smart mobility data*' and is widely used in urban planning, traffic management, transportation optimization. The term, *mobility data*, thus, contains a broad category of data, such as traffic data, location data, bicycle data, images.

In some cases, it can even pertain to sensitive categories of personal data. For instance, *location data*, such as travel itineraries, could reveal information about someone's health condition, sexual orientation, racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership and sometimes even lead to (re-)identification.[11]

Location data is defined by the e-privacy directive[12] as being any data processed in an electronic communications network or by an electronic communications service, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service.

In addition, some mobility data could also overlap with **Big Data** and/or **Metadata**, leading to potentially broad insights of individuals their lives.

Metadata provides information about other data or in other words, it is data about data[13].

On the one hand, it describes the characteristics or attributes of a dataset, such as its structure, format, source, or content and on the other hand it provides context and facilitates the organization, discovery, and management of data. Examples of metadata include file names, timestamps, authorship, data types, and data relationships. It is defined by the Data Act[14] in article 2(2) as being "*a structured description of the contents or the use of data facilitating the discovery or use of that data*."

---

[10] ENISA, *Data Protection Engineering*, January 2023 : PETs are defined as software and hardware solutions, i.e., systems encompassing technical processes, methods or knowledge to achieve specific privacy or data protection functionality or to protect against risks of privacy of an individual or a group of natural persons.

[11] CJEU, 8 October 2020, nr. C-623/17, *Privacy International*, ECLI:EU:C:2020:790, para. 71; CJEU 8 April 2014, *Digital Rights Ireland and Others*, C 293/12 and C 594/12, EU:C:2014:238, para 27; CJEU 21 December 2016, *Tele2*, C 203/15 and C 698/15, EU:C:2016:970, paras 99 and 100.

[12] Article 2(c) E-Privacy Directive.

[13] R. L. Lubas, A. S. Jackson and I. Schneider, *The Metadata Manual : A Practical Workbook*, 1st edition, Chandos Publishing, 2013, 2.

[14] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).

It has been acknowledged that it could potentially reveal more information – even sensitive personal data – than the content itself. Indeed, the Court of Justice, confirmed that "*those data, taken as a whole, may allow very precise conclusions to be drawn concerning the private lives of the persons whose data has been retained, such as the habits of everyday life, permanent or temporary places of residence, daily or other movements, the activities carried out, the social relationships of those persons and the social environments frequented by them.*"[15] This shines a different light on it, considering the jurisprudence on "personal data" and "identification". Moreover, as a participant explained, users often lack awareness on what information they are providing based on the metadata by using services. Consequently, it is essential for smart city actors to act cautiously when processing metadata, and apply adequate privacy enhancing techniques.

Big Data is inherently linked with smart cities. As mentioned by one of the participants, it is an integral part of the architecture for intelligent communication technologies. And yet, there is to date, no universally agreed definition on Big Data. It is often described by three v-words: the volume of data stored, the velocity at which new data is acquired, and the variety of data that can be acquired.[16] In simplified terms, Big Data relates to large and complex datasets both structured and unstructured, which require storage, analysis, use and disposition.[17] Unstructured data, typically compromises, images, videos, sensor data, and so on, while structured data is typically represented by a table, where each row corresponds to a user, a building, a traffic intersection, a digital camera, etc. It follows that in a smart city context, a massive amount of data will be collected, processed, stored and analysed. This gives, as participants acknowledged, rise to many issues and concerns especially, since personal sensitive data is likely to be involved.[18]

Table 1. Definitions and examples for the main categories of record-level datasets.

Concept	Description	Examples
Structured data	Datasets that are often represented as a table, where each row is a user's record. A column contains each user's value for the same attribute.	Survey, census, and health care cross-sectional data with demographic attributes (such as sex and age), observations (such as medical condition).
Set-valued (semistructured) data	Datasets where each user is associated with a collection of data points. Each data point typically contains information about "actions" or "events" relating to the user. The most prominent types of set-valued data are time series, where each point includes both the timestamp and the description of an action or event.	Most types of behavioral data, such as location data (containing individual trajectories, with time and location from GPS coordinates or cell towers, for each place visited), movies and videos watched online, and supermarket shopping data.
Unstructured data	Datasets that do not have a natural structured representation.	Text data (messages, tweets, letters), graphs (social networks), images and videos (face, body posture, fingerprint, iris).

Source: A. Gadotti, L. Rocher, F. Houssiau, A-M. Creţu, and Y-A de Montjoye. "Anonymization: The Imperfect Science of Using Data While Preserving Privacy.", *Science Advances 10*, no. 29 (2024): eadn7053-. doi:10.1126/sciadv.adn7053. 3.

[15] CJEU 8 April 2014, nrs. C-293/12 en C-594/12, *Digital Rights Ireland and others*, ECLI:EU:C:2014:238.

[16] D. Grimaldi, K. Shalla, I. Fontanals, C. Carrasco-Farré, "From Smart City to Data-Driven City", In D. Grimaldi, C. Carrasco-Farré (Eds.) *Implementing Data-Driven Strategies in Smart Cities*, United States: Elsevier Science & Technology, 2021. doi:10.1016/B978-0-12-821122-9.00005-1, 25.

[17] E. M. Mimo and T. McDaniel, "Smart Cities: A Survey of Tech-Induced Privacy Concerns", in Richard Jiang, Ahmed Bouridane, Danny Crookes, Feng Hao, Said Boussakta, Chang-Tsun Li (Eds.) *Big Data Privacy and Security in Smart Cities*, Springer Nature Switzerland AG 2022 5.

[18] E. M. Mimo and T. McDaniel, "Smart Cities: A Survey of Tech-Induced Privacy Concerns", in R. Jiang, A. Bouridane, D. Crookes, F. Hao, S. Boussakta and C-T. Li (Eds.) *Big Data Privacy and Security in Smart Cities*, Springer Nature Switzerland AG 2022 5.

Participants identified several risks during the study day:

- **Re-utilisation and diverse processors using mobility data**: when diverse actors (*e.g.* traffic police, public transport planners, city planners, health care officials) process mobility data subsequently, it creates a risk for unfair use of data due to lack of transparency.
- **Complexity and Ambiguity of Big Data including mobility data**: Participants emphasized the potential risks associated with the fast-evolving technological environment, where previously unknown datasets and new combinations of data & metadata could potentially lead to the (re)-identification of individuals. As further illustrated under Section 2.1 III Example from participants.

As a result, various participants stressed that open mobility data by default leads to concerns because it is a sensitive type of data which needs to be used with extra care.

2.1.2 Mitigating measures (PETS)

Various participants discussed the implementation of several technological measures aimed at enhancing data minimization and limiting re-identification, commonly referred to as *privacy-enhancing technologies* (PETs). These techniques encompass a range of strategies and tools designed to minimize the collection, use, and retention of personal data, thereby reducing privacy risks and enhancing individuals' control over their information.

Although participants touched upon diverse approaches such as aggregation, computing on encrypted data[19] (fully homomorphic encryption, multi-party computation), differential privacy, adding noise on statistics, synthetic data, and so on this report will be limited to a general overview of the pseudonymisation and anonymisation techniques. For a more in depth reading readers are encouraged to refer the opinions of the Authorisation and Opinion Service, the Working paper of the Berlin Group and the guidelines on anonymisation and pseudonymisation of the EDPB dating 2014. Discussions on the technical update of these guidelines have been ongoing since several years at the EDPB level. Awaiting the outcome of the ongoing review of these guidelines, the BE DPA refers by way of technical update to the study published in July 2024 by panellist and external member of the BE DPA Authorisation and Opinion Service Y-A. de Montjoye, as a co-author of "*Anonymization: the imperfect science of using data while preserving privacy*".[20]

The principle of minimum data processing requires the use of anonymized or pseudonymized data if the purpose of the processing can be achieved on the basis of that data.[21] The difference between these concepts is, therefore, crucial since anonymized data is not considered to be personal data, and is thus exempt from the GDPR scope, contrary to pseudonymized data.

---

[19] For more information on computing on encrypted data, see : N. P. Smart, J.W. Baron, S. Saravanan, J. Brandt, A. Mashatan: *Multiparty Computation: To Secure Privacy, Do the Math: A discussion with Nigel Smart, Joshua W. Baron, Sanjay Saravanan, Jordan Brandt, and Atefeh Mashatan*, ACM Queue 21(6): 78-100 (2024); N. P. Smart, "Computing on Encrypted Data", *IEEE Secur. Priv*. 21(4): 94-98 (2023)

[20] A. Gadotti, L. Rocher, F. Houssiau, A-M. Creţu, and Y-A de Montjoye, "Anonymization: The Imperfect Science of Using Data While Preserving Privacy", *Science Advances*, 10, no. 29 (2024), eadn7053-, doi:10.1126/sciadv.adn7053.

[21] Opinion 247/2022, para. 73.

Pseudonymization is defined in article 4(5) GDPR as being "*the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person*." It is a practice of replacing personal identifiers with substitutes, such as pseudonyms or codes. [22] It follows, that pseudonymized data should still be considered information on an identifiable natural person.[23]

Anonymisation is the process of rendering personal data anonymous. According to 26 of the GDPR, anonymous data is "*information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable*".

At the core of smart cities and the use of data, lies the question of how to unlock the potential of mobility data while guaranteeing protection: is there a way to anonymise this data in a way that it can be safely shared with 3rd parties?

Participants identified several challenges:

- **Identification of Risks and Privacy Concerns**: Despite privacy-enhancing technologies (PETs) being employed, there is no guarantee of 100% solid proof against identification risks, especially where mobility data are involved. Anonymization and pseudonymization, while valuable, pose challenges in determining the threshold of identifiability, leading to potential risks in data processing. In particular, it is very difficult to identify which future services or applications will provide additional data that can help in re-identification.

- **Complexity and Ambiguity of Big Data**: The inherent complexity and constantly evolving nature of big data present challenges in finding and recruiting experts with the knowhow to implement pseudonymization/anonymization/PETs techniques efficiently (and ensure they achieve their goals and meet thresholds).

- **Relevance in Anonymization:** The excessive pursuit of anonymization in data processing may result in diminished relevance and fewer useful or correct insights.

2.1.3 Example from participants

It is evident from the following example that data minimization and the implementation of data protection by design and default are closely intertwined and must be carefully considered before implementing smart city initiatives and during its lifecycle.

[22] Article 4(5) GDPR : the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

[23] Rec. 16 GDPR.

One of the projects of a participant clearly illustrates the importance of implementing data protection by design and by default and data minimization through suitable planning and development stages.

In that particular project a solution was sought to overview the maintenance of overhead line, in an efficient manner. The participant explained that they came up with the idea of a camera with a specific angle and a limited perimeter view. Since the aim of this tool was to overview the overhead line, the necessary data could not go as far as filming people in public spaces.

The participant illustrated that though the scope of the camera view was limited, there could be a chance that someone might still appear in the angle. To tackle this eventuality, they came up with the design of an algorithm which blurred faces.

Yet as pointed out by the participant there is always a possibility that some elements are overlooked. Take for instance the challenge of anonymisation of unstructured data: an algorithm trained to blur faces, might miss some faces (the blurring might be incorrect) leading to possible identification.

2.1.4 Recommendations put forward by participants

- **Good practices:** adhere to the notion that all data and metadata should be treated as personal data unless properly anonymized and consider asking consent or refraining from starting the processing where risks cannot be mitigated.

- **Adoption of Privacy-Enhancing Technologies**: Continued investment in and adoption of privacy-enhancing techniques, such as anonymization, pseudonymization, computing on encrypted data and localized processing are essential to mitigate identification risks and uphold data minimization principles.

- **Case-by-Case Evaluation**: The selection of technical anonymization methods should involve a thorough evaluation of privacy utility tradeoffs on a case-by-case basis, considering factors such as noise tracking mechanisms, risks of unanticipated queries, new services or applications that provide additional data and the balance between privacy protection and data usability.

- **Early Integration of GDPR Principles**: Stakeholders should prioritize the early integration of data minimization principles and data protection by design and default in smart city initiatives, ensuring that these principles are considered from the initial planning and development stages.

> *"There is no silver bullet. If you think you have a perfect solution that no one can ever attack and that there is absolutely no risk; or if someone wants to sell you a solution that is absolutely perfect and has absolutely no risk, either you won a Nobel prize in computer science, either and more likely, you don't know enough. Correctly understanding and mitigating the remaining risks, and communicating about them is an absolutely essential element. If the risks are mitigated enough, maybe it can be anonymous data or otherwise it is pseudonymous data. Anonymous data is not the only way to use data.*
>
> *Some techniques can help depending on the conditions. In order to share publicly aggregate data, for large datasets, differential privacy might be able to help to reach an acceptable level of anonymization."*
>
> - Yves-Alexandre de Montjoye -

2.2 The transparency, consent and citizens involvement challenges

The *transparency* concept, has been touched on by several speakers, illustrating its significance in relation to smart cities. The concept is integral to the GDPR. Transparency is associated with various articles within the regulation, particularly in relation to the exercise of data subjects' rights and communication practices towards the data subject.[24] Aside from being an obligation, it is also a necessary condition for trust.[25] Hence, it is intrinsically linked to :

- **Fairness and consent**, because transparent communication or information to subjects, is considered to be a minimum act of fairness. Consequently, it should also include, transparency on **risks**, to enable informed consent, as indicated by one participant.

To process data lawfully you need a legal basis as provided for in the GDPR. Without one, the processing of data would be unlawful, unless you can establish an exemption. One of the 6 legal basis is *consent (see infra 2.3).*

To be valid, consent must be specific, unambiguous, informed, hence the transparency, and it must be freely given. It follows that the transparency is key for a valid consent.

Article 4(11) of the GDPR defines *consent* as: "*any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her*."

Some participants considered consent as being a feasible option with the advancement of technology. The technologic app improvement and lay out should enable users to use it, even in a multi-user environment such as car sharing services.

But some issues were also noted by participants:

- in current apps, consent is mostly used for advertising purposes, so that smart city actors (*e.g.* road operators, car manufacturers, etc.) find it hard to get a separate consent for the processing and sharing of data – it is a challenge to convince apps developers to build consent solutions (opt-in);
- collecting the consent of users (bike users, car users, etc.) is also a challenge. The solution is to be clear about the purposes, about the advantages and with whom the data are shared;
- identity management for consent via apps (in case of several vehicle's user) is also a special challenge;
- getting parental consent for bike apps used by children is a special challenge;

In addition, participants highlighted the challenge of proactive involvement of citizens. This phenomenon stems partly from the transparency issues towards the citizens and the lack of trust among citizens towards smart city initiatives. Additionally many individuals do not fully understand how their data is processed, leading to a sense of

[24] Article 29 WP, WP260 rev.01, adopted 29 November 2017, Guidelines on transparency under Regulation 2016/679, last accessed on 25th of July 2024, https://ec.europa.eu/newsroom/article29/items/622227.

[25] I. Varošanec, On the path to the future: mapping the notion of transparency in the EU regulatory framework for AI, *International Review of Law, Computers & Technology*, 36:2, 95-117, (2022), DOI: 10.1080/13600869.2022.2060471, 97.

diminished control. As a result citizens often do not exercise their rights regarding their smart city related data.

Concretely, participants pointed out diverse challenges:

- **Difficulty to communicate about risks**: tackling e.g. the opaqueness or black box nature of algorithms, highlighting the need for transparent algorithmic decision-making processes where smart city projects involve such type of decisions.
- **Citizen reluctance towards opt-in opportunities**: Concern that citizens would be unlikely to consent, when offering this opt-in opportunity.
- **Literacy hurdle**: Challenges related to citizens' understanding of their data and metadata being processed (which input, which output, which systems are processing their data), despite information platforms being available in some smart city projects.
- **Engagement struggle**: Encouraging citizen interaction and participation in data access and testing, especially considering the lack of engagement observed in some cases.

2.2.1 Examples from participants

As previously mentioned, participants indicated that transparency also implicated challenges, in particular regarding the ability of citizens to understand the input and output from their data.

The use case with **Uber drivers** highlights the impact of **transparency** and giving valuable meaning to data. The participant referred to the viewpoint of the Swiss Federal Court in the Uber case of 2022[26] where the Court upheld a decision that classified the drivers as being in an employment relationship. Unpaid overtime and others uncovered benefits (such as holiday, sick pay, car expenses) could be quantified by the drivers by accessing their data collected by Uber.

In this particular case, the participant pointed out the initial challenges in data retrieval due to the complex format and sheer volume, efforts were made to empower drivers in understanding and using their data. By leveraging open-source software and introducing AI-driven consent mechanisms, drivers were able to scrutinize their data comprehensively. This process not only assisted in legal negotiations but also shed light on potential discrepancies in compensation offered by Uber. According to the participant, this experience, which is an employment case, could be replicated elsewhere in other contexts.

This example shows the importance of transparency and the need to receive the requested data in an understandable and accessible format, in order for the data subjects to be able to interpret it and take benefit from it.

In this context, it was also reminded that the Dutch Data Protection Authority (AP), following several drivers' complaints, imposed a fine of 10 million EUR on Uber Technologies, Inc. and Uber B.V.[27] mainly for lack of transparency, obstacles for

[26] Swiss Federal Court, 30th Mai 2022, nrs. 2C_575/2020, 2C_34/2021, https://www.bger.ch/files/live/sites/bger/files/pdf/fr/2c_0575_2020_yyyy_mm_dd_T_f_14_32_25.pdf.

[27] Dutch Data Protection Authority, Decision of 11 December 2023 imposing a fine of 10 million EUR on Uber Technologies, Inc. and Uber B.V. ('Uber'), https://www.autoriteitpersoonsgegevens.nl/en/current/uber-fined-eu10-million-for-infringement-of-privacy-regulations.

exercising the **right to access of drivers**, failure to disclose the full details of its retention periods for data concerning European drivers, or to name the non-European countries in which it shares this data.

On the one hand speakers mentioned that transparency could be improved if citizens could be involved in the access, testing of the data and their results. On the other hand, however, it turned out that citizens were not interacting despite having platforms available to them. Previously mentioned Swiss court case is also a good example, showcasing that drivers appeared to gain interest once their understanding improved on the usage of their data.

It follows that understanding and access to data, go hand in hand. This shows the necessity of another actor to enhance transparency. As put forward by one of the speakers, stakeholders need to be incentivized, to increase transparency. Meaning that not only should stakeholders have a deep understanding of the data protection and privacy principles, but they must also encourage this further down the line, enhancing an environment based on transparency.

Consequently, it is essential to increase knowledge and to create a transparency environment at the level of stakeholders to incentivize and enhance participation of citizens. Participants highlighted the possible role for third parties such as NGOs and privacy authorities to inform and support users in their access rights.

In the discussions on the feasibility of introducing a **consent option for connected cars** or utilizing default settings, a panellist expressed significant skepticism regarding the possibility of a default "on" option, emphasizing that consent should be the preferred approach in such scenarios.

While acknowledging the complexity arising from multiple users sharing a single vehicle, the panellist suggested that this issue could be addressed by car manufacturers. Yet it was highlighted that the challenge of obtaining informed consent, noting that the percentage of successfully collected consents tends to be low in such contexts.

This underscores the difficulty of obtaining consent, particularly in situations where in-car applications may be designed by an app-developer to automatically gather consent in order to collect personal data for their own purposes (e.g. targeted advertising). It is crucial for these applications to transparently disclose information about the collected data, the purposes, the parties with whom they share data, etc.

2.2.2 Recommendations put forward by participants

- **Transparency mechanisms**: Implement mechanisms to ensure transparent communication and information dissemination to subjects, including clear explanations of risks to enable informed consent.
  Example: A panellist advised to experiment risk communication formats that stem from the medicine area to communicate risks on medication. In a use case they experimented with easy-to-understand formulations and icons stemming from these medical procedures to communicate about risks of reidentification.

- **Incentivization**: Incentivize stakeholders to prioritize transparency, accountability, and data protection principles, encouraging their dissemination and implementation throughout the data ecosystem.
- **Transparent algorithms**: Advocate for the adoption of transparent algorithms and decision-making processes, promoting accountability and trust in algorithmic systems.
- **User Education and Awareness**: Efforts should be made to raise awareness among users about the data and metadata they generate while using services, enabling them to make informed decisions regarding their data privacy.
- **Learning tools**: Develop educational programs and initiatives to improve citizens' understanding of their data and its implications, fostering engagement and participation in data-related activities.

> *"If risks cannot be eradicated, a solution could be to communicate these risks to people sharing the data. If smart city stakeholders want to use the data, where there are risks involved, a solution can be to communicate these risks to data subjects and let the data subject take the decision. In order to give informed consent, people should know about risks."*
>
> - Yves-Alexandre de Montjoye -

2.3 The challenge of proportionality and lawfulness of smart city projects framed in laws

Some participants indicated that consent is not always an appropriate legal basis for data processing in smart city projects, particularly because mandatory consent should be regulated by a (formal) law and aligned with the authorities' public mission. When smart city data is processed by a public authority, in order for such project to be lawful, the features and goals of the processing must be stated clearly in the law ruling the public body's mission or in specific legislation framing the processing in line with the public authority's mission. Some participants emphasized that governments and public authorities must set the right balance between the risks, the volume of collected data and the real improvement in citizens' lives. Legislators bear the obligation and responsibility of assessing both necessity and proportionality, *ab initio*. Initiatives aimed at enhancing citizens' lives must adhere to this principle, ensuring that advancements are not implemented at the expense of their rights and freedoms. Additionally, this deliberation, leads to one of the core questions, according to a participant, namely "*how smart cities really need to be*?" and "*can cities be smart while protecting privacy*"? While the objective is to improve citizens' lives, this should not serve as carte blanche for unrestricted implementation. Rather, the pursuit of improvement must always be accompanied by a careful balance that safeguards citizens' fundamental rights and freedoms.

This is also underscored in article 22 of the Belgian Constitution which stipulates that "everyone has the right to respect for his private and family life, except in the cases and under the conditions determined by law. The law, decree or rule referred to in Article 134 guarantees the protection of this right." Indeed, consequently, any norm regulating the processing of personal data (which by its nature constitutes an interference with the right to the protection of personal data) must on the one hand

be necessary and proportionate. It follows that in order to be lawful, this interference must correspond to a pressing social need and be proportionate to the aim pursued.[28]

As put forward by the Belgian DPA, any disproportionate collection of mobility data could also create a sentiment of general surveillance from the viewpoint of the right to privacy.

The core question with regard to the legislative proposals, is whether or not the rights of data subject's will be impacted significantly. Depending on the answer, the legislative proposal will need to include the essential elements in the proposal.

Hereunder is an overview of the views expressed by some of the participants on what legislation should include:

- Legislation should explicitly specify the purposes of processing location data by public actors. This clarity in legislation ensures that debates about purposes occur at a legislative level, rather than being left ambiguous. Legislation might simplify the privacy assessment balance on the side of smart city actors.
- The legal basis to treat data in the context of public interest must be foreseen in the public body's mission through its legislated goals (public task described in a law or decree). It is then the task of the Belgian Data Protection Authority to see to it that the legislated goal is described in a sufficiently foreseeable manner.

2.3.1 Examples from the BE DPA

The following cases highlight that legislators bear the obligation and responsibility of assessing both necessity and proportionality, ab initio, in the context of lawfulness.

The Belgian Data Protection Authority issued on 20 February 2024 an opinion regarding a draft decree on data processing and exchange in the context of **emission monitoring of road vehicles (ANPR cameras)** where a proposed initiative failed to adequately prove the necessity and proportionality[29]. Specifically, the proposal aimed to establish a decree-based framework for emissions monitoring and subsequent data exchange, aligning with the Flemish Government's goal of enhancing air quality as outlined in the 2030 Air Policy Plan.

- The preliminary policy research suggests the potential for monitoring vehicle emissions, which could yield valuable insights into the actual emissions produced by vehicles.
- Additionally, the proposal outlines a framework for conducting measurement campaigns on public roads to assess vehicle emissions in Flanders. Through emission monitoring, vehicles exhibiting concerning emission levels can be identified for roadside inspections.
- Furthermore, the gathered measurement data would contribute to environmentally relevant research and policy evaluation efforts. One core issue was that neither the draft (including the Explanatory Memorandum) nor the 2030 Air Policy Plan unambiguously established that remote sensing would be effective

[28] M. Krzysztofek, *GDPR*, Alphen aan den Rijn: Wolters Kluwer Law International, 2021, 56 ; CJEU 21 december 2016, nrs. C-203/15 and C-698/15, *Tele2 Sverige AB and others*, ECLI:EU:C:2016:970, para. 112. Pro memorie, the CJEU precluded national legislation which, for the purpose of fighting crime, provides for the general and indiscriminate retention of all traffic and location data of all subscribers and registered users relating to all means of electronic communication.

[29] https://www.autoriteprotectiondonnees.be/publications/avis-n-11-2024.pdf

in combatting emission fraud by vehicle owners, especially considering the significant interference with the rights and freedoms of data subjects that the use of ANPR cameras (and the creation of a new, widely accessible database) represents.

Indeed, it is evident that the lawfulness is called into question when the effectiveness of the proposed systems remains uncertain.

Similarly in Opinion 186/2021 on SMART MOVE in Brussels, related to an **intended kilometre tax** [FR and NL], the BE DPA observed that the introduction of a kilometre charge to combat traffic congestion in the Brussels Capital Region, was a legitimate objective. The BE DPA however advised against a continuous collection of location data generated by vehicles, and recommended using an on board unit that would perform a limited and local processing of data relating to the trips of the vehicle.

In its Opinion, the Belgian DPA required that the data relating to the trips of the vehicle always stays within the on board unit that performs the tax calculations. In that way, the sending of the location data to the collecting authority only takes place in exceptional situations (e.g., audit of the system in the case of a challenge and litigation).

J. Balasch, A. Rial, C. Troncoso, B. Preneel, I. Verbauwhede, C. Geuens:
PrETP: Privacy-Preserving Electronic Toll Pricing. USENIX Security Symposium 2010: 63-78

So, local processing is a possible way of meeting the minimalization requirement, and consequently, defining a legitimate and proportionate objective.

As regard the legal basis of this project, it can be noted that the objective of introducing an intelligent and fair charge based on the actual use of a vehicle, monitored through an onboard application that citizens' are required to download on their smartphones, was found disproportionate, having regard to the interference which the achievement of that objective involves to the rights and freedoms of the persons concerned, in order to reduce the congestion of vehicular traffic in the Brussels Capital Region.

In the same line opinion 273/2022 [*FR* and NL[30]] on the **rollout of smart meters** or on the **monitoring of road traffic for trucks** in the Walloon region required from the legislators necessary and proportionate nature of the collection of real-time location data for the purposes of real-time traffic monitoring and, if necessary, to justify this in the explanatory memorandum of the decree.

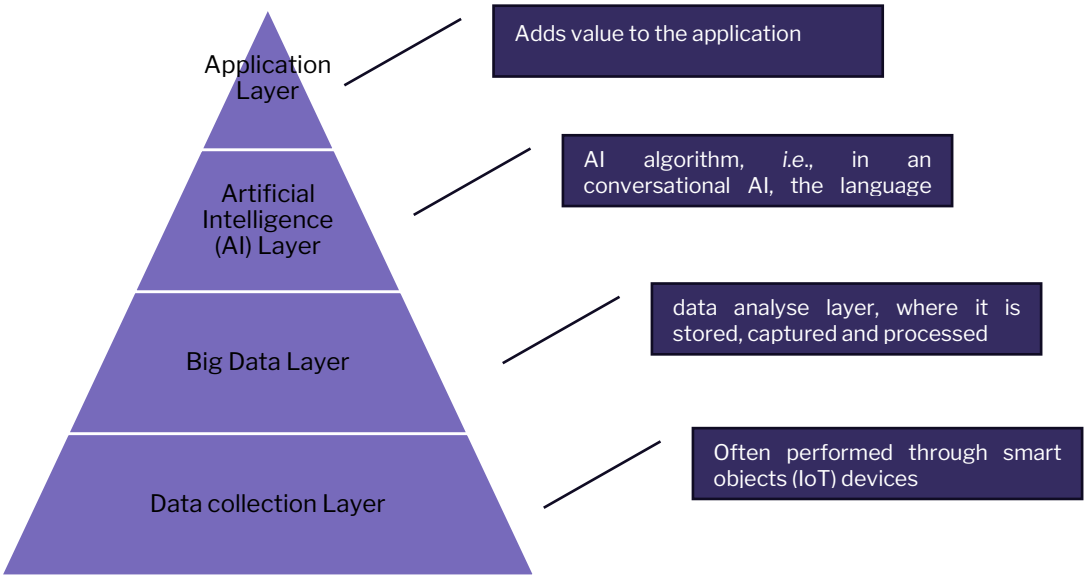
As these illustrate, it is crucial to consider technical measures which ensure that data collection is proportionate, such as anonymization and pseudonymization of data.

[30] Translation ongoing.

2.4 The Challenge of data infrastructure and data access

The architecture of data collection and access in smart cities is inherent to the functioning of smart cities. It is, therefore, unsurprising that diverse participants touched upon this subject.

A data architecture, should be understood as the underlying structure, rationale or framework of a system.[31] As has been emphasized by participants, there are diverse domains (mobility, homes, …) in smart cities. Hence, architectural designs and layouts will be tailored to the unique needs and contexts, and can therefore focus on various aspects, such as, technology, human-system interaction, logic, or other.[32] Despite these differences, a participant argued that smart city projects involving AI often share a common technical architecture at their core, which is structured in diverse layers that cannot be separated from each other.[33] (fig)



** For an example in which such an architecture is used, readers are referred to Jin-ho Park, Mikail Mohammed Salim, Jeong Hoon Jo, Jose Costa Sapalo Sicato, Shailendra Rathore, Jong Hyuk Park*, CIoT-Net: a scalable cognitive IoT based smart city network architecture, Hum. Cent. Comput. Inf. Sci. (2019) 9:29 https://doi.org/10.1186/s13673-019-0190-9 (see Fig. 2 on page 7).*

The challenge for public and private smart city actors are that the architecture of smart city projects should allow for effective data sharing from the data subject at the data collection layer, to the application layer. To enable this interaction, as highlighted by some of the participants, *interoperability* plays a key role. Interoperability ensures that different systems, devices, and applications can work together and exchange information effectively, despite being developed by different manufacturers or using different technologies.

---

[31] A. Das, Sumanta Chandra Mishra Sharma, Bikram Kesari Ratha, Chapter 1 - The New Era of Smart Cities, From the Perspective of the Internet of Things, in Rawat, Danda B., and Kayhan Zrar Ghafoor, editors. *Smart Cities Cybersecurity and Privacy*. First edition., Elsevier, 2019, 5.

[32] R. Wenge, X. Zhang, C. Dave, L. Chao and S. Hao, "Smart city architecture: A technology guide for implementation and design challenges," in *China Communications*, vol. 11, no. 3, pp. 56-69, March 2014, doi: 10.1109/CC.2014.6825259. 58.

[33] P. James, R. Astoria, T. Castor, C. Hudspeth, D. Olstinske and J. Ward, Smart Cities: Fundamental Concepts, in J. Carlos (Ed.) Handbook of Smart Cities, Springer Nature Switzerland AG 2021, 14.

As stated on the website of the European Commission, the DGA aims at increasing trust in data sharing, strengthening mechanisms to increase data availability and overcome technical obstacles to the reuse of data.[34] It defines '**interoperability'** *as the ability of two or more data spaces or communication networks, systems, connected products, applications, data processing services or components to exchange and use data in order to perform their functions"*[35]

Several participants emphasized that the vast amounts of data held by big tech companies are often closely guarded, with little willingness to share it with competitors and citizens. The reluctance of big tech to share data underscores the importance of rethinking how data ownership, access, and control is approached. Technology plays a pivotal role in shaping power relationships, and the architecture of our technological systems can either reinforce or challenge existing power structures. By embracing diverse architectures, particularly within smart communities, there is the potential to reshape power relationships in significant ways, according to participants.

In that regards some participants fervently put forward the new EU legislative proposals that enhance data sharing, among others through data spaces and personal data stores.[36]

Data space is an umbrella term for a data management community that contains all the data sources for an organization regardless of its format, location or model. Each data source is considered a participant of the dataspace.[37]

This coincides with the main features, provided by a participant, to recognise data spaces: namely a stable (governance) framework, that supports the exchange/sharing of data, which is voluntary in both B2B (Business to Business) and B2C (Business to Consumer) situations and which in principle is open to any participant.

Datapods as defined by a participant refer to a "Pod" or "Personal Online Datastore" and coincides with a virtual storage space on the internet that citizens or consumers use to store and manage personal data. The participant highlighted that storage of data in one personal data pod instead of data stored by several service providers would give users more control over their own data, including personal or confidential information.[38]

Participants illustrated challenges such as:

---

[34] European Commission, last accessed on 27 September 2024, https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

[35] Article 2(40) Data Act.

[36] The European Commission is advancing the development of various data spaces, and the EDPB and the EDPS recently published an opinion on the European Health Data Space (EHDS). The BE DPA supports the EDPB's approach.

[37] E. Curry, S. Scerri and T. Tuikka, Data Spaces: Design, Deployment, and Future Directions, In: E. Curry, S. Scerri and T. Tuikka (eds.), *Data Spaces*, Springer, Cham, 2022, https://doi-org.kuleuven.e-bronnen.be/10.1007/978-3-030-98636-0_1, 3.

[38] Athumi, https://athumi.eu/en/data-collaboration-platforms/generic-information-platform-public-domain-2. Athumi is a public company that aims at encouraging and facilitating secure data exchange and data collaboration between consumers, businesses and public agencies, https://athumi.be/en/about-us [accessed on 12 July 2024].

- **Interoperability**: which persist across diverse stakeholders, including technical experts, public sector entities, and communities, hindering effective data sharing efforts.
- **Complexity**: the complexity of the data architecture design impedes efforts to empower data subjects and mitigate passive involvement in data processes.
- **Data centralization**: which can be a concern due to the risk of data breaches, potentially compromising the security and privacy of sensitive information.

2.4.1 Examples from participants

As illustrated by a participant, the core technology architecture of smart cities can lead to a passive involvement of data subjects in smart city projects, primarily consuming services or passively receiving them. This can lead to various issues, such as acceptance challenges, as citizens feel disconnected from the design process, leading to a lack of trust. However, by rethinking the role of citizens in these projects, they could be empowered and gain a proactive function.

> *"Avoid single point of trust that becomes single point of failure"*
>
> - Bart Preneel -

One illustration provided by a participant was the implementation of active users in the data architecture. Citizens can voluntarily provide their data to feed smart city projects. This idea which has been implemented in Flanders, consisted of distributing some kits to citizens to collect data on environmental and sound pollution. These kits not only facilitated data collection but also served as educational tools, allowing citizens to better understand the data and contribute to large-scale organizations. In such projects, awareness and transparency are key: the citizens need to be clearly informed about the nature of the data that is collected and the fact that they can access this data themselves and decide on access to it.

Another example given by panellists is the concept of open data consumers[39], which highlights the potential to involve citizens by providing transparency on how data is processed. The challenge, however, is that data shared on open data portals (federal, regional, local, and EU) – such as websites, applications –often goes underutilized.[40] These portals aim to serve two types of users: developers seeking to build innovative services and citizens interested in transparency. Unfortunately, the transparency goals are not being fully met. To address this, a collaboration with the University of Namur aims to engage citizens with open data portals through interactive quizzes.

---

[39] In the view of the panellist in particular this **concept of "open data consumers**" corresponds with individuals or groups who utilize publicly available data provided through open data portals. This concept highlights the potential role of these consumers in engaging with and making use of the data. The aim is to promote transparency by allowing citizens to see how data is processed and to use the information for various purposes, such as developing new services or gaining insights. In the given context, open data consumers include both developers, who might use the data to create innovative applications, and ordinary citizens, who are interested in understanding more about the data. The challenge mentioned is that these data sets often remain underutilized, particularly by the general public. To address this, initiatives like interactive quizzes are being introduced to help guide non-expert users to the relevant data, thereby encouraging greater engagement and understanding among open data consumers.

[40] For more information see: Bono Rossello, Nicolas; A. Simonofski, A. Clarinval and A. Castiaux, , "A Typology for AI-enhanced Online Ideation: Application to Digital Participation Platforms" (2024). *Hawaii International Conference on System Sciences 2024 (HICSS-57). 3.* https://aisel.aisnet.org/hicss-57/dg/ai/3 ; A. Simonofski, A. Zuiderwijk, A. Clarinval, & W. Hammedi (2022). *Tailoring open government data portals for lay citizens: A gamification theory approach.* Elsevier, 2022, International Journal Of Information Management; 2022, Vol. 65.

These quizzes direct citizens to the exact datasets with relevant information, re-engaging and educating non-expert users about the data.

Yet as a participant pointed out: the question is, what should be open? Sometimes decisions are made without showing how these have been made. One could say to open up everything, hence being transparent on everything. But there are limits to transparency, for instance, not everyone is willing to share what and how much benefits they receive or even their medical history. As put forward by a speaker, commonly people have to agree on what should be open and to whom. Because this entails the protection of privacy values and the control of it.

2.4.2 Recommendations put forward by participants

- **Investment creates trust and facilitates implementation:** some speakers noted that investment in data stores or platforms (be it by government bodies or public-private partnerships) are needed in order to establish citizens' trust in such frameworks for data sharing.
- **Empowering data control:** other speakers stressed the need to put citizens in control of their data, to enable access for innovative European companies, and to adopt common European data spaces to facilitate data flows. According to the same participants these measures will enhance data sovereignty[41], foster innovation, and ensure seamless data sharing across Europe while respecting fundamental rights to privacy and legal frameworks.
- **Data decentralization**: To ensure the security and resilience of smart city systems, some participants highlighted that it is essential to avoid a single point of trust that could become a single point of failure. This can be achieved by adopting a decentralized approach to data storage and management, emphasizing local data storage and processing wherever possible and encrypted storage and processing in centralized environments[42].

2.5 Accountability and Governance

Overall, participants acknowledged the importance of implementing a human-centric approach in the field of smart cities, which aligns with the idea that individuals should have control, as mentioned previously. Smart city initiatives cover various aspects such as mobility, data protection, sustainability, social equity, and others, it follows that their implementation is intertwined with accountability and governance. Participants emphasized the necessity of involving Data Protection Officers (DPOs) from diverse actors (private and public)) in all stages of projects (from inception to implementation and evaluation) to enhance accountability and promote inter-organizational learning in data protection practices. Moreover, collaboration between the public and private sectors requires the implementation and adherence to regulatory standards.

---

[41] For more information on data sovereignty, see: https://digital-strategy.ec.europa.eu/en/policies/strategy-data https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf

[42] For more information on "trusted execution environments", see :P. Jauernig, A.-R. Sadeghi, E. Stapf, Trusted Execution Environments: Properties, Applications, and Challenges, IEEE Secur. Priv. 18(2): 56-60 (2020); X. Li, B. Zhao, G. Yang, T. Xiang, J. Weng, R. H. Deng, A Survey of Secure Computation Using Trusted Execution Environments, CoRR abs/2302.12150 (2023)

Participants shared some challenges they encountered in this context:

- Thorough testing and experimentation are required to verify compliance and address potential gaps in new digital European regulatory framework on data exchange and processing such as the Data Act, Data Governance Act[43], and AI Act[44]. Additionally, navigating complex regulatory frameworks and identifying suitable regulatory sandboxes poses significant challenges.
- Balancing innovation with privacy and accountability concerns within smart cities and communities appears to be a significant governance dilemma. Moreover, the perception among some bidders that tender evaluations often prioritize functionality and price over sustainability and data protection complicates governance efforts.
- The need for evidence-based decision-making support systems underscores the importance of robust governance structures and transparent processes. Yet, there is a risk of data protection requirements being treated as checkbox exercises, without substantive consideration of GDPR compliance, presenting additional governance challenges.

> *"Trust and governance for public-private partnerships: need to clarify the purpose of processing and repurposing."*
>
> - Bart Preneel -

2.5.1 Examples from participants

Participants highlighted concerns about the current regulatory landscape, indicating challenges in implementation. Additionally, they felt that the emergence of the new digital European regulatory framework introduced further complexities with various new initiatives. However, according to a participant, AI regulatory sandboxes could offer a solution to address the difficulties posed by existing regulations, which prove impractical in practice, and to ensure compliance.

The participant illustrated it with an example on the effectiveness of autonomous detection of infractions by ANPR cameras. As thus, the technology would be in place to enable ANPR cameras to detect mobile phone usage while driving, but despite having privacy-by-design features, a legal obstacle arises: only the police are allowed to process the images. Consequently, making the cameras non-autonomous for the detection of potential infractions. And requiring human oversight by police. To detect this phone usage while driving, the participant believed that AI regulatory sandboxes[45] – a controlled environment where companies and organizations can test their artificial intelligence (AI) technologies under the supervision of regulators. This setup allows them to experiment with new AI applications, products, or services while adhering to

---

[43] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)

[44] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

[45] Article 3(55) of the AIA defines an '**AI regulatory sandbox**' as being "*a controlled framework set up by a competent authority which offers providers or prospective providers of AI systems the possibility to develop, train, validate and test, where appropriate in real-world conditions, an innovative AI system, pursuant to a sandbox plan for a limited time under regulatory supervision.*"

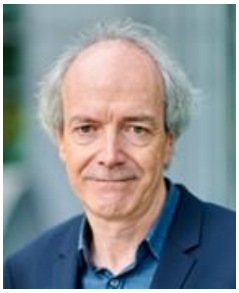
An example of such an AI regulatory sandbox is the pilot regulatory sandbox, presented by Spain, https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented

specific guidelines and regulations – could be beneficial to challenge such regulatory limitations, providing feedback on impractical regulations and facilitating adjustments to accommodate innovative solutions effectively.

2.5.2 Recommendations put forward by participants

- **Implementation of Data Protection Impact Assessment**: data protection is a collaborative effort which requires mapping all the risks, all the data flows, the diverse actors that process and this as early as possible, as well as on a continuous basis, when dealing with smart systems.
- **Encourage Early Involvement of Data Protection Officers**: Public sector buyers should signal their commitment to data protection by involving Data Protection Officers (DPOs) early in procurement processes. Substantive involvement of DPOs and broadening of evaluation criteria can mitigate tensions and ensure that data protection is prioritized throughout project lifecycles.
- **Facilitate Inter-organizational Learning**: Procurement procedures should be designed to promote inter-organizational data protection learning and knowledge exchange. Encouraging collaboration among suppliers and sharing best practices will enhance data protection capabilities and drive improvements in governance frameworks.
- **Enhance Stakeholder Collaboration:** Governance frameworks should prioritize stakeholder engagement and collaboration to ensure inclusive decision-making and effective implementation of data protection measures in smart city initiatives. Establishing multi-stakeholder forums and platforms for knowledge sharing can foster cooperation and alignment of interests.
- **Invest in Capacity Building**: Governments and organizations should invest in capacity building and training programs to enhance the skills and (technical) knowledge of stakeholders involved in data protection governance. Providing training on relevant laws, regulations, and best practices will empower stakeholders to address privacy challenges effectively.
- **Promote Transparency and Accountability including in the context of AI**: Governance frameworks should promote transparency and accountability in data handling practices, policies, and procedures. Establishing mechanisms for reporting privacy breaches, seeking redress for individuals, and fostering trust between stakeholders will enhance accountability in smart city initiatives. It should be noted that the GDPR requires organizations to ensure transparency, accountability, and the protection of personal data. In the case of AI-automated decision-making, organizations must provide clear explanations to data subjects about how decisions are made and establish mechanisms for individuals to contest and seek human intervention. It follows that the burden of proof relies on organizations to demonstrate compliance.

3. Final conclusions by Bart Preneel



It is evident that the discussion on data protection in smart cities has shed light on the intricate interplay between technology, governance, and privacy. Through insightful presentations and lively discussions, challenges and opportunities inherent in the architecture of smart cities have been explored.

One key takeaway is the inherent political nature of architecture. The imperative to avoid single points of trust, which could become single points of failure, underscores the need for robust and decentralized systems with local processing[46]. If data is collected centrally, it should be protected by encryption and processing should also perform in the encrypted domain using the latest cryptographic techniques. Additionally, the concept of differential privacy has been emphasized as a means to safeguard privacy while allowing for meaningful data analysis.

Moreover, the study day highlighted the importance of user trust and engagement in data handling processes. From consent to control over data and metadata, empowering users is paramount in building a foundation of trust in smart city initiatives. Furthermore, discussions surrounding data ownership, repurposing, and the role of Data Protection Officers have underscored the necessity for clear guidelines and governance structures.

In conclusion, navigating the evolving landscape of smart cities, all actors must strive to strike a balance between utility and privacy. By leveraging encryption and privacy-preserving models, smart city stakeholders can create urban environments that not only enhance efficiency but also prioritize the ethical treatment of data. Moving forward, it is imperative that continue to engage in dialogue and collaboration to ensure that smart cities are inclusive, transparent, and respectful of individual privacy rights.

[46] For more information, see also C. Troncoso, G. Danezis, E.Kosta, J. Balasch, Bart Preneel, PriPAYD: Privacy-Friendly Pay-As-You-Drive Insurance. IEEE Trans. Dependable Secur. Comput. 8(5): 742-755 (2011) ; C. Troncoso, D. Bogdanov, E. Bugnion, S. Chatel, C. Cremers, S. F. Gürses, J.-P. Hubaux, D. Jackson, J. R. Larus, W. Lueks, R. Oliveira, M. Payer, B.Preneel, A. Pyrgelis, M. Salathé, T. Stadler, M. Veale, Deploying decentralized, privacy-preserving proximity tracing, Commun. ACM 65(9): 48-57 (2022)

4. Nomenclature

AI Act	Artificial Intelligence Act
BE DPA	Belgian Data Protection Authority
DA	Data Act
DGA	Data Governance Act
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
EDPB	European Data Protection Board
ePrivacy Directive	Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)
GDPR	General Data Protection Regulation
PETs	Privacy Enhancing Technologies