



Redefining urban digital twins for the federated data spaces ecosystem: A perspective

Downloaded from: <https://research.chalmers.se>, 2024-12-20 11:26 UTC

Citation for the original published paper (version of record):

Gil, J., Petrova-Antonova, D., Kemp, G. (2024). Redefining urban digital twins for the federated data spaces ecosystem: A perspective. *Environment and Planning B: Urban Analytics and City Science*, In Press. <http://dx.doi.org/10.1177/23998083241302578>

N.B. When citing this work, cite the original published paper.

Redefining urban digital twins for the federated data spaces ecosystem: A perspective

EPB: Urban Analytics and City Science
2024, Vol. 0(0) 1–16

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/23998083241302578

journals.sagepub.com/home/epb

S Sage

Jorge Gil 

Chalmers University of Technology, Sweden

Dessislava Petrova-Antonova

Sofia University, Bulgaria

Graham JL Kemp

Chalmers University of Technology and University of Gothenburg, Sweden

Abstract

Urban Digital Twins (UDTs) promise to facilitate the transition to a smarter planning and decision-making process, but they face many challenges to meet stakeholders' expectations and to support delivering a better living environment for citizens. Their full potential can only be reached through more flexible sharing and integration of data and data models. From this perspective, we argue that UDTs can be realised in an ecosystem of data spaces that support a federated data architecture. We present the conceptual architecture for UDTs in a federated data spaces ecosystem, describe the different components and layers, and reflect on the UDT mediator role in relation to other actors in the ecosystem. We aim to contribute to the field of UDTs by showing the way forward for their development, connected to the concepts of data spaces and federated database systems, and to the field of data spaces by introducing UDTs as a new component playing a mediator role.

Keywords

Digital twins, federated database systems, data spaces, smart cities

Introduction

The enormous environmental and societal challenges we face today, particularly in cities, require the support of digital technologies to open the existing siloes of data, and multidisciplinary methods and knowledge to tackle the complexity of urban dynamics. Urban Digital Twins (UDTs), also called

Corresponding author:

Jorge Gil, Department of Architecture and Civil Engineering, Chalmers University of Technology, Sven Hultins gata 6, Göteborg SE-412 96, Sweden.

Email: jorge.gil@chalmers.se

Local Digital Twins or Digital Twins of Cities, are a technology that promises to facilitate this transition to a smarter planning and decision-making process (Bolton et al., 2018). To support data-driven planning and decisions, a bi-directional connectivity providing a seamless flow of data (both real-time and historical) between the physical city and its digital counterpart is required (Figure 1). The cyclical process can consist of six phases: (1) Aggregate a collection of heterogeneous multi-scale and multi-temporal data; (2) Analyse present, past and future urban scenarios; (3) Deliver insights to the stakeholders; (4) Support decisions by the stakeholders; (5) Create dynamic changes in the city; (6) Interact with physical assets and systems through everyday operation and provide information through sensors. A critical aspect of urban digital twins is their need to be responsive to near-real-time changes in the city. This process requires extensive input datasets, feedback and a high-frequency information flow throughout its lifecycle.

UDTs have been inspired by, and naturally evolved from, the established field of 3D city models (Ketzler et al., 2020), which still plays an integral part in their development. The focus of this field has been on the three-dimensional geometric representation, visualisation and simulation of the physical objects and structures of the city, with a particular focus on the representation of buildings, developing theories, methods and tools around the acquisition, processing and management of geographic data for the production of semantic 3D city models, in a convergence of the fields of Geographic Information Systems (GIS) and Building Information Modelling (BIM), into what one might call geo-BIM. These models have been applied in numerous use cases of different application domains, for example, infrastructure planning, disaster management, energy, and real estate (Biljecki et al., 2015) driven by the aspirations of the smart cities concept. 3D city models consist of a centralised repository (often file based) using a single data model (often CityGML or IFC based), in a tight coupling between their application use case and the data on which they operate, that is ingested, processed and stored in large quantities. Notable examples are the 3D buildings model of the Netherlands,¹ the 3D models of Helsinki,² Berlin,³ and Calgary,⁴ or historical 3D models created of the city of Zurich.⁵ The current challenges of the 3D city model paradigm, that is, consistency between models, standardisation, data quality, data interoperability, data maintenance/governance,

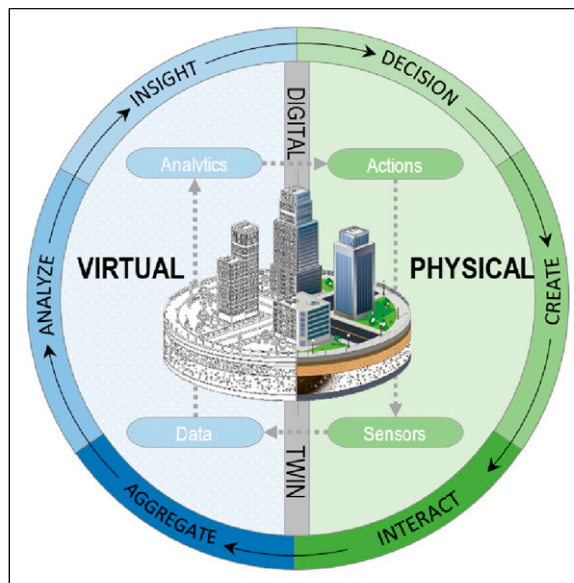


Figure 1. The role of the digital twin in the urban planning and decision-making process of the physical city.

and real-world use cases (Stoter et al., 2020), must also be addressed for UDTs to reach their potential.

While UDTs rely on a 3D city model for interactive 3D visualisation and physical simulation, they are expected to address a wide range of application domains, such as transportation, utilities, and public services, relevant to urban planning and decision-making that requires diverse representations (Stoter et al., 2021). In many of these use cases, 3D data (Herbert and Chen, 2015) and realistic-looking visuals (Stoter et al., 2021) are not a requirement. Equally important are 1D data, for example, traffic counts, weather measurements or energy consumption, and 2D data, for example, population statistics or land use information, and in particular temporal and real-time data (Stoter et al., 2021). In this context, interactive dashboards are suitable interfaces for UDT implementations (Calzati, 2023). However, for UDTs to reach maturity, they must address both technical and societal challenges around diverse application domains, a multitude of use cases, heterogeneous data sources and the stakeholders involved, which, beyond municipalities, include public and private sector organisations and citizens (Calzati, 2023; Ketzler et al., 2020). As UDTs evolve, one might need to increase their scope and utility by connecting the data in a municipal city model with data provided by other organisations, for example, data related to transport or energy consumption in the city. While one could consider extending a city model and importing additional static information into a city database, it would be hard to develop and to work with a single unified data model. Further, it can be advantageous for these other data sets to be managed and updated autonomously by provider organisations and then to connect to these external data sets so that up-to-date data can be accessed when needed. Addressing this requires flexible sharing and integration of data, metadata, data models, and solutions that support the growth of rich and complex UDTs, enabling the development of new services, applications, technologies, and business models to meet stakeholders' diverse needs.

The proposition that we explore in this article is that UDTs can be realised in an ecosystem offered by data spaces, a concept and set of technologies offering access to diverse data from multiple data providers and domains in a de-centralised system, and that a federated data architecture is necessary for a UDT to thrive in such an ecosystem. These insights have developed over the course of a collaboration with partners from academia, industry and the public sector, including technology developers, data owners, city planners, architects, construction companies, consultancies, and property managers, sharing an interest in UDTs. The importance of a federated approach is also recognised by the Centre for Digital Built Britain, who advocate connected digital twins that span across organisational and sectoral boundaries as 'tools to understand the complexities of interconnected systems and provide better insights to enable better decisions and interventions' (Gemini Council and Lamb, 2022).

UDTs rely on diverse datasets from various domains owned by diverse stakeholders. The federated data space architecture enables seamless integration and interoperability among datasets within one data space and across data spaces of a city, ensuring that data silo walls are broken down and information flows freely across different domains, to the benefit of all participants in the data space, and consequently all stakeholders engaged with a UDT. Moreover, it facilitates the trust between data space participants and data sovereignty. Data providers have confidence that their data will be used responsibly and securely, maintaining their ownership and control over their information. This trust is built through transparent data governance policies, robust security measures, and clear usage agreements that define how data can be accessed, shared, and utilised. Data sovereignty, which refers to the principle that data is subject to the laws and governance structures of the nation where it is collected, reinforces this trust by ensuring that data providers can impose and enforce local regulations on their data. By respecting data sovereignty, a federated data architecture allows for the protection of sensitive information, compliance with regional data protection laws, and the empowerment of data providers to manage their data in alignment with their own policies

and values. This alignment fosters a collaborative environment where data providers are more willing to participate.

In the next sections, we introduce the concepts of UDTs, data spaces and federated database systems. This is followed by an illustrated exposition of a conceptual data architecture for UDTs in a federated data spaces ecosystem. The article concludes with a reflection on the implications and ways forward.

Background

In this section, we start by reviewing the concept of UDTs towards a broader understanding of the requirements from different stakeholders and for diverse data. We then introduce the concept of data spaces as it is being developed internationally, as a data ecosystem with technological and organisational components that can sustain the needs of UDTs. We conclude with a brief description of federated database systems, as the distributed architecture that should be adopted by UDTs to engage in data spaces, and which gives a critical mediator role to UDTs in the context of data spaces.

What is (expected of) an urban digital twin?

UDTs, introduced in the previous section, can be considered as digital decision-making tools, representing the physical aspects, systems and processes of a city, and providing insights to various stakeholders through a variety of simulations and analyses based on recent technologies (Jeddoub et al., 2023). Their evolution and adoption have led to an increasing number of definitions that are continuously adjusted to meet changing stakeholder requirements. The lack of a clear and commonly accepted definition of UDTs is due to a lack of consensus among stakeholders with different points of view (Shahzad et al., 2022), variability in the characteristics (Sepasgozar, 2021), terminological ambiguity across domains (Ketzer et al., 2020), a variety of technical approaches focused on specific domains (Lehtola et al., 2022), and the availability of different forms and outputs (Ferré-Bigorra et al., 2022). In addition, there is a lack of data sharing and interaction frameworks, facilitating the full deployment of UDTs at a large scale (Shahat et al., 2021). Nevertheless, there are key components of UDTs that make them a backbone of the digital transition of cities (Bolton et al., 2018, Lei et al., 2023, Bauer et al., 2021, VanDerHorn and Mahadevan, 2021, White et al., 2021, Ye, 2023):

- Visualisation Models: provide digital representations or virtual replica at different levels of detail of the city's physical infrastructure, including buildings, roads, bridges, etc. as well as other city systems and processes.
- Bi-directional Data Exchange Models: collect and integrate historic and real-time data from various sources, providing information about the city's dynamics, including traffic flow, energy consumption, air quality, etc.
- Analytical and Simulation Models: utilise artificial intelligence (AI), machine learning algorithms, and other mathematical and analytic models, to run simulations, analyse data, identify patterns, and predict future trends.

While the 3D visual representation is an important component and the focus of current UDT implementations such as Gothenburg and Helsinki, it is merely the tip of the iceberg of what constitutes UDTs. Under the surface, other components must exist to provide data in machine usable standards, automated workflows, analytic models, and protocols for data interoperability, security and trust that enable the production, operation and maintenance of those UDTs. These set of

components, in turn, enable the functional expectations and requirements of UDTs to support the decision-making and planning process, illustrated in [Figure 2](#).

Data functionality provides the interaction between the physical and digital twins and supports all the other functional aspects. A physical twin is equipped with sensors for capturing real-time data of its state, and the data functionality stores the captured data for later use. Reactive functionality uses this data stream to monitor present city operations and the urban environment or to run historical analysis. The derived insights can be fed back to the data functionality. Predictive functionality looks at a future state of a city's operations and the urban environment and can be used for planning purposes. For instance, when a new city quarter is planned, it allows planners to assess its impact before implementation. Finally, 'What if' functionality can be initialised using the current information from the digital twin and combined with simulation models, to subsequently evaluate the potential outcome of alternative planning interventions and policy changes. All these functional requirements (and expectations) of the UDT revolve around a multitude of data from different domains, diverse models and simulations, and should rely on data and model federation at building, neighbourhood, district and city levels.

Data spaces

Data spaces are an approach to data management for large scale scenarios involving numerous data sources ([Curry, 2020b](#)), aiming to facilitate data exchange and collaboration between multiple independent organisations in a given knowledge domain, for example, urban planning, mobility, or utilities. The data spaces are based on a data federation (de-centralised) model because, in these cases, it is difficult and expensive to have a unifying data schema across all data sources to represent the given domain ([Curry, 2020b](#)). Particularly relevant to UDTs is the concept of real-time linked data spaces, designed to enable data management for intelligent systems within smart environments (e.g. smart cities), combining the paradigm of data spaces with real-time data querying capabilities ([Curry, 2020a](#)).

The basic idea of a data space is to provide a software infrastructure that enables data users to connect with data providers ([Figure 3](#)). Data users do not access a data provider's data files or databases directly; rather, they communicate with the data provider via data space connectors based on open standards, and data space software that handles authorisation and provides services that ensure that data are accessed under conditions agreed between data providers and data users.

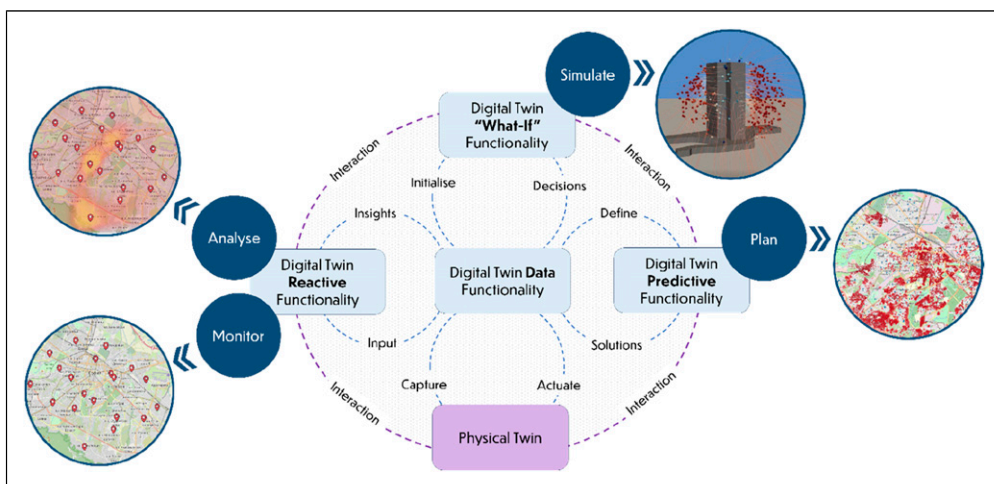


Figure 2. Functionality of a UDT required to deliver the various expectations of stakeholders.

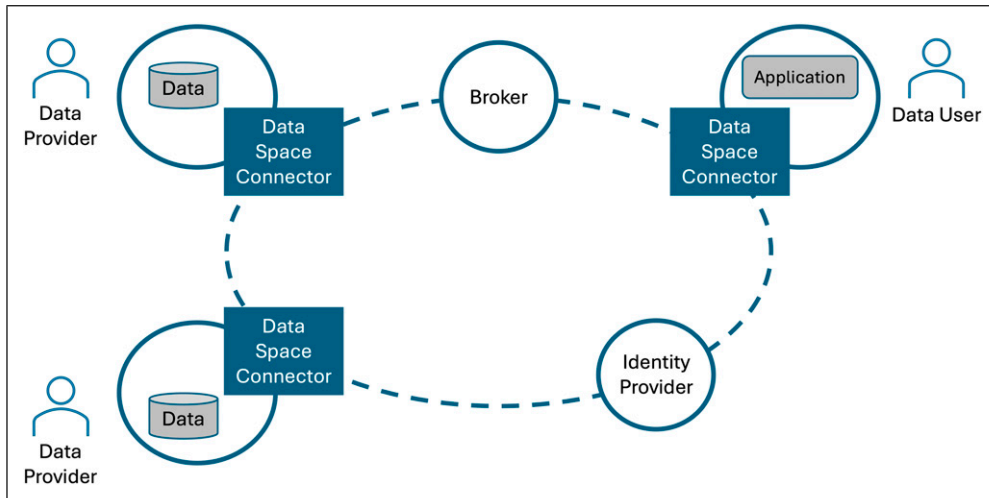


Figure 3. Simple data space representation of a data user connecting to two different data providers.

Data spaces involve various actors that can play one or more roles (Data Spaces Support Centre DSSC, 2023b). The main roles include (1) data space participants, that is, data rights holders, data providers, and data users; (2) data space intermediaries, that is, connection-providing, marketplace, clearing house, and personal data handling; (3) data space services for enabling transactions and for data processing; and (4) the governance authority. For example, a municipality participating in a UDT data space can be simultaneously: a data rights holder of data on municipal assets managed by another data provider organisation; a data provider of open data sets covering other municipal assets; and a data user of traffic congestion and air quality data collected by another data space participant. The governance authority can be a national institution, such as a cadastre or a national administration, since data space participants can operate at the national level.

The data space building blocks, introduced by OPEN DEI (Nagel and Lycklama, 2021) and further developed by the Data Space Support Centre (Data Spaces Support Centre DSSC, 2023a) into a Building Block Taxonomy, are divided into two primary groups, namely, organisational and business building blocks and technical building blocks. The organisational and business aspects must be developed for these types of complex sociotechnical solutions such as data spaces to thrive, and one should not mistake them for simple technical solutions. At the organisational and business level, the data space building blocks support the development of business models, use cases and data products for participants in the data space; define governance rules between data providers, recipients, and intermediaries; deliver organisational governance through the data space governance authority; and support regulatory compliance and contractual frameworks to manage the rights, obligations and contractual resources for data space participants (Data Spaces Support Centre DSSC, 2023a). At the technical level, the building blocks relate to *data interoperability*, providing functionality for data exchange, including data models, formats and interfaces (APIs), and for provenance and traceability; to *data sovereignty and trust* including the identification of participants and assets, the establishment of trust and policies for data access and usage control; and to value-creation by facilitating the registration and discovery of data product offerings or services and enabling the monetisation of data sharing within a marketplace (Data Spaces Support Centre DSSC, 2023a).

Data spaces are being proposed in the European Strategy for data (European Commission, 2020) and the EU Data Act, that are supported through various initiatives, such as the International Data Spaces (IDS) Association,⁶ Gaia-X⁷ and the Data Spaces Support Centre.⁸ Implementations of data

spaces exist at different stages of development at the European scale for different knowledge domains directly relevant to UDTs, for example, smart community⁹, remote sensing,¹⁰ green deal,¹¹ or energy (Dognini et al., 2024). The mobility data space¹² is a leading example, providing an extensive data catalogue covering categories such as traffic information, road works, parking, road signage, weather, public transport, car and bike sharing, and infrastructure, among others. The various data products within each data category have regional, national, European and even global coverage, and are provided by a wide range of partners in the data space, such as public authorities, car manufacturers, map and mobility service providers, sensor developers, consultancy, telecommunication, and insurance companies, among others. Each partner, by joining the data space in an equal, secure and de-centralised form, benefits from the access to other data sets that can enable new business models and products. In the case of municipalities, joining the data space gives access to the data without having to own it or negotiate individually with the various data owners, which could prove costly or impossible.

The characteristics of data spaces, as these platforms become implemented for different urban knowledge domains and in different territories, make them the natural ecosystems for UDTs to develop. Emerging UDT implementations, such as those of Helsinki, Rotterdam or Flanders,¹³ are exploring the development of data ecosystems with multiple actors (D'Hauwers et al., 2022), and in this context, data spaces provide a suitable framework. However, this requires that the technical implementation of UDTs considers a federated database systems approach.

Federated database systems

In the development of UDTs, it is understood that using a central data repository poses multiple problems: storage space; resources and effort required for cleaning and restructuring; maintenance and access to the latest data updates; and reduced flexibility of interface and querying options. In contrast, federated database systems leave the various data sets at their origin, and the system manages the user queries to retrieve results from the relevant component databases in the federation. This represents a paradigm shift from the heavy data processing and transfer operations that are needed to build and maintain a centralised data repository, to complex querying operations through the network of federated databases.

The federated database system is a more robust and data independent system, but it requires us to know how the data backend is developed, and how UDTs can mediate between users and the rich ecosystem of data sources.

Federated database systems adopt a multi-level schema architecture between the component databases where the source data reside, and the user interface application(s) (Sheth and Larson, 1990). Schemas are descriptions of the data, consisting of classes or entities and their relationships. Each schema provides a specific description of the data suitable for a purpose: *local schemas* are optimised for storage in a specific data source; *component schemas* harmonise and extend local schemas in a common data model; *export schemas* control access to the data source and exchange the data based on standards; *federated schemas* link the various data sources. The federated schema at the top level provides a linked catalogue mapping the federated data space. This schema enables querying the data required to answer complex questions or to run complex models, which can be stored in different database management systems (e.g. relational, non-relational or graph databases) suitable for each specific data source, while at the same time hiding this complexity from the user.

The multi-level schema architecture of federated database systems ensures greater independence between the specific applications, for example, a UDT, and the data sources, where changes on one level do not imply changes on all levels. Thus, each component can be developed more or less independently, with new data sources and applications being added to the federation, for example, data space, as users' needs evolve.

In federated database systems, mediators are software components that perform a variety of functions in the layers between user applications and data resources. [Wiederhold \(1992\)](#) describes a modular information architecture in which mediators are structured into hierarchies, performing tasks between layers such as transforming databases using view definitions and supporting abstraction and generalisation over underlying databases. This architecture became an important part of the Knowledge Sharing Effort ([Neches et al., 1991](#)). [Figure 4](#) illustrates how a mediator can fit into the data space picture. Data users’ applications can send queries expressed against a federated schema to a mediator which, in turn, sends queries expressed against the component schemas of the data providers’ databases (blue dashed arrows). The mediator combines the results from the component databases and returns these to the users’ applications (green dashed arrows).

With data spaces and federated database systems in mind, new architectures for UDTs have been discussed ([Lefever and Michiels, 2019](#)) and various technical implementations are currently being developed (e.g. [Coenen et al., 2021](#); [Raes et al., 2022](#)). One notable effort is by the DataBri-X consortium,¹⁴ aiming to provide a toolbox of practical, robust and scalable ‘bricks’ (i.e. processes, technologies and tools), building on existing and emerging initiatives, to make data available in the context of European Data Spaces. These efforts will offer technical solutions for the various components required by UDT applications to have a federated data structure and to connect to data spaces playing a critical mediator role.

Urban Digital Twins in the federated data space ecosystem

At this stage, we illustrate how UDTs co-exist with other actors in the federated data space ecosystem, presented in the previous section, to grasp the architecture, functioning and implications for UDTs of this new paradigm. A UDT can be seen as an interface to the autonomous, distributed, and heterogeneous data sources and data spaces from different knowledge domains that are required for digital representations of city assets, systems and processes.

We start with an overview of urban data spaces, for example, mobility and environment, with nodes representing actors playing different participant and service roles (i.e. data rights holders, data

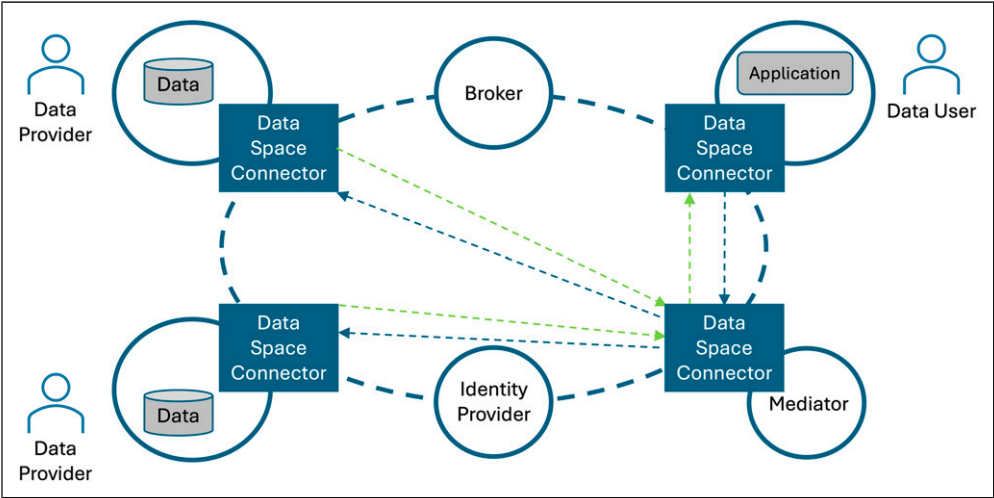


Figure 4. Data space representation with the inclusion of a mediator, to which data users and data providers connect, performing queries from user applications across multiple databases.

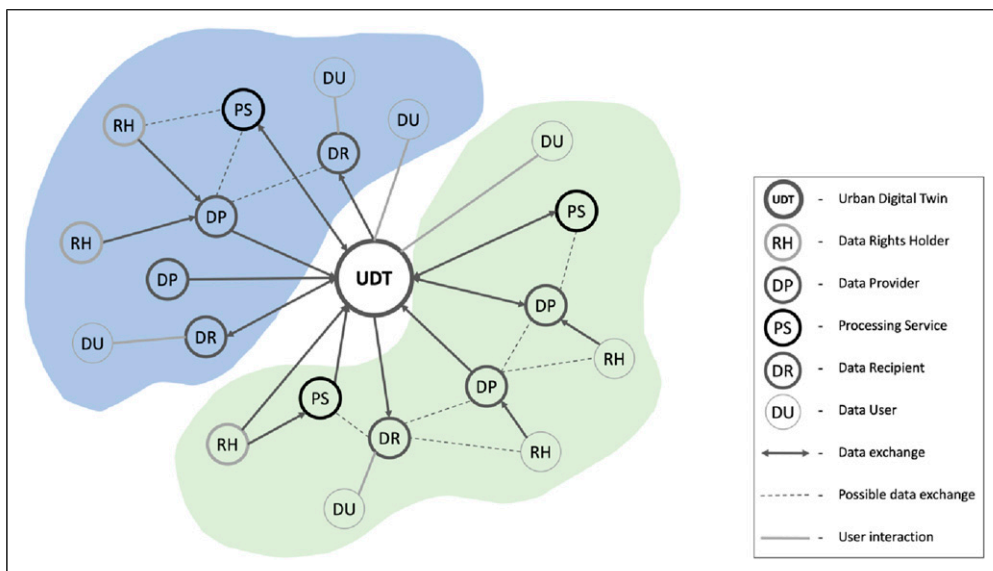


Figure 5. Conceptual diagram of urban data spaces with multiple actors, and a UDT as a new node with a mediator role connecting different data spaces.

providers, data recipients, and data processing), and data users being final recipients engaged in specific urban planning and decision-making use cases. A UDT is a new type of node in the data space playing a mediator role between different actors and data spaces (Figure 5).

Data sharing in a typical data space happens through direct data exchange and processing between two actors (Figure 3), for example, a *data provider* and a *data recipient*, which implies that potentially every actor can be connected to every other actor in the data space. Interoperability is ensured through secure and trusted connections based on agreed open standards. However, richer applications, beyond a singular data transaction, require the integration and processing of multiple data sources from different *data providers* in a complex value chain (Figure 4). Such is the case of UDTs, that bring together the data from a multitude of physical, technical, ecological and socio-economic urban systems.

In this case, a UDT node will connect to various *data provider* nodes of historic and real-time data from different data spaces (Figure 5), and in turn *data recipients* that have complex queries can connect directly to the UDT node. This way *data recipients* can obtain information on the road network, buildings, current traffic, and weather conditions to estimate air quality levels. This estimation is fed back to the UDT to support decision-making by *data users* regarding public health. Despite a UDT playing this mediator role, it does not mean that the data space becomes a centralised system. The data and analytic models remain with their owners, and direct connections between actors are still possible. Furthermore, each actor in the data space, including UDTs, can play different roles, which we shall examine next.

The federated architecture of urban digital twins

By definition, a UDT provides an overview of, and access to, the multifaceted dimensions of the city, offering a visual and interactive querying interface to the urban data space. A municipality will require a UDT to address a wide range of planning and decision support use cases, for example, traffic management and air quality, which requires access to multiple data sources in different data

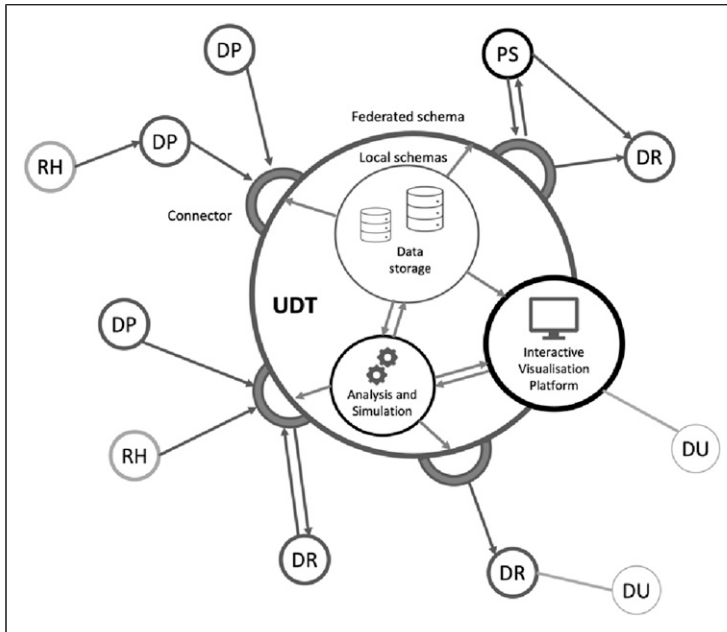


Figure 6. Conceptual diagram of the architecture of a UDT, including the three main components of a UDT, and connections to other actors in data spaces.

spaces as described in the previous section. The federated architecture of a UDT fulfilling those roles is presented here in a conceptual diagram (Figure 6).

At its core, a UDT developed by a municipality can have one or more databases of different types storing internally the data required for analysis, simulation and visualisation in the UDT platform, using local schemas that are optimised for the specific use cases. For example, this data can include terrain, buildings, roads and vegetation data for 3D visualisation coming from municipal data sets; traffic and road quality data provided by a fleet of vehicles from mobility service providers; and air quality and weather data from local sensors. The databases are not accessible in this form to external actors. Instead, the UDT platform uses export schemas of the federation's component databases to define which data and in what form they are provided to the linked data layer.

The municipality might want to make its city model, or parts of its city model, available for other organisations (e.g. private companies, cadastre, and academia) in a data space ecosystem. The outer layer of a UDT offers linked data of the urban data space in a mediator role, that is, the federated schema that links the various data assets and services available in the data space that are relevant for the municipality and the UDT use cases. This layer has multiple connectors to allow data exchange and data processing between the various actors in the data space, which use a vocabulary of ontologies, reference data models and metadata based on open standards. In many cases, a connector provides access to linked data to a *data recipient* (DR), meaning that the data are composed of elements from different data providers via a single query and interface.

As for the UDT interactive 3D user interface, it accesses the local schemas for optimised performance, and it can also connect to the linked data layer via external connectors to obtain additional data. *Data users* (DU) can access this interface directly, and it functions like an application in the urban data space.

The roles of different actors in an urban data space

To make the UDT concepts presented here more concrete, we can consider the use case of planning an urban district where private and public sector actors need to analyse aspects of urban comfort, such as air quality and temperature, wind and noise, combined with pedestrian and vehicular flows. Based on the different roles that all actors can play in this use case, we focus on two main groups: *data rights holders* and *data providers* (Figure 7), and *data recipients*, *processing services* and *data users* (Figure 8).

The *data rights holder* (RH) is a simple role, where, for example, the cadastre owns data about the buildings and properties, a transport authority owns data about road infrastructure and traffic counts, and a mobility services provider collects air quality real-time data from a network of sensors installed on vehicles. The data are assets of these actors' operations, and they store large quantities of data in databases of different types using their local schemas, retaining ownership and control over how the data are used.

Being active in the data space, the *data rights holders* implement export schemas to make (parts of) the data available to other actors in the data space, thus contributing to the UDT and municipal operations. For example, the transport authority shares the traffic counts via an application programming interface (API), taking, in this case, a *data provider* (DP) role. Or these actors can establish an agreement with a *data provider*, for example, the national cadastre agency that transfers part of the data, processes it to align with a given open standard, and gives controlled external access to it in the urban data space and consequently to the UDT.

The *data recipient* (DR) role is played by actors that not only access the data but process it, creating new data assets available in the data space. For example, this can be a scientific institution that runs wind simulations for the urban district planned by the municipality, based on built environment geometry provided by the cadastre agency or obtained from the UDT, or a consultancy

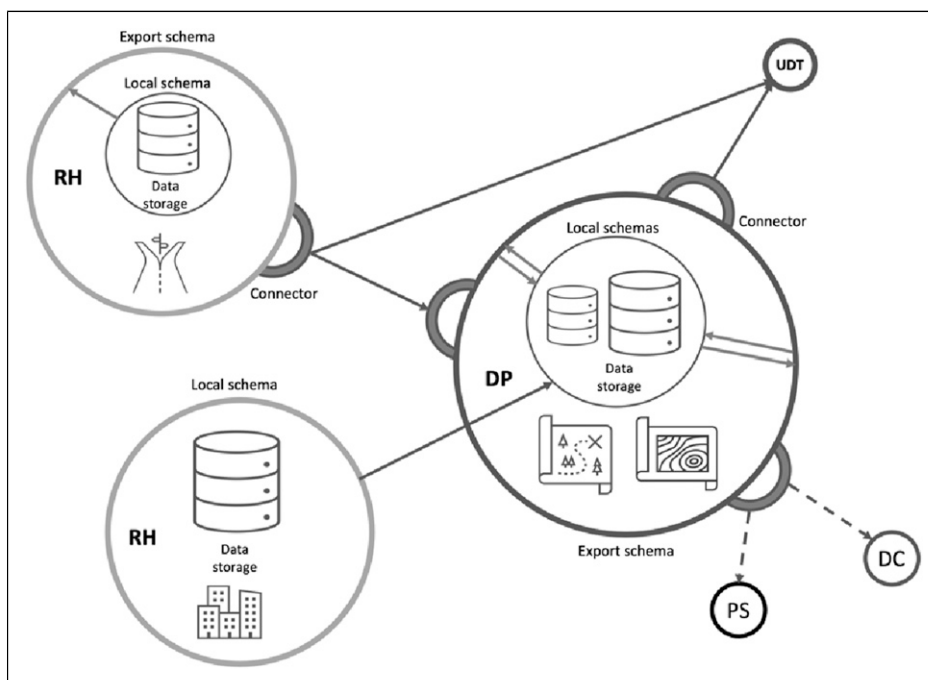


Figure 7. Conceptual diagram of the architecture of data providers and data rights holders.

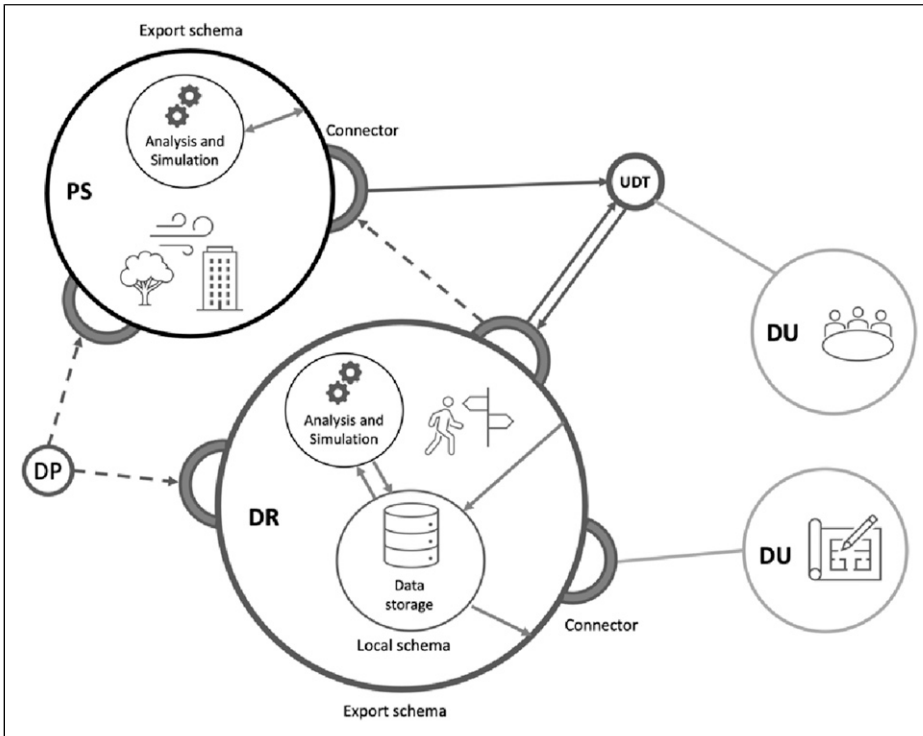


Figure 8. Conceptual diagram of the architecture of data recipients, processing services and data users.

company that runs pedestrian micro simulations of the same area. It can even be the UDT consuming cadastre data to generate the 3D geometry of the built environment for visualisation purposes. Some actors can be simply *processing services* (PS) that make analytical or computational models available in the dataspace, for example, wind or pedestrian simulations, that are applied by *data recipients* or the municipal UDT to obtain information relevant to the urban district planning use cases. In the PS role, there is no permanent data storage, only to the extent that some data must be cached for processing.

Elsewhere, we have the *data users* (DU) that access data and do not produce outputs for the data space. These actors can be municipal planners accessing information via the 3D interface of the UDT to support decision-making regarding the urban comfort of the new planned district, or they can be an architecture company that uses the information to support the design of their buildings and public spaces for the same urban district.

The stakeholders with different roles can, in one way or another, connect to the same UDT to fulfil their specific use case. A UDT is thus a mediator, providing access to the multifaceted data, analytical and simulation models and visualisations of the city required for the assessment, planning and design of complex problems like the urban comfort of a new urban district.

Discussion and conclusions

Given the emerging context of data spaces providing heterogeneous data federation for different knowledge domains relevant for UDTs, it is important to consider what they offer and the roles that a UDT can play in such an environment, enabling UDTs to meet the expectations of various stakeholders: a *data provider* of linked municipal data to architecture companies; a *data recipient* of

the latest municipal data for processing; a *data rights holder* of internal simulation results (e.g. wind, noise, temperature) for consultants and analysts; a *data user* of real-time traffic or weather data displayed in dashboard visualisations for planning agencies; and a *data processing service* processing cadastral data to produce a 3D city model for simulation and visualisation.

For the multitude of stakeholders in data spaces, accessing data and calling analytic applications via a UDT interface can prove advantageous. Actors only need to work with a UDT interface, and the diverse data and models required for complex applications will already be linked in the federated schema. Nevertheless, one can have different UDT interfaces for specific applications depending on each one's complexity and specificity. UDTs would be like 'The Canterbury Tales' by Chaucer, that is, collections of stories about the city with multiple and unique perspectives for different users.

At the same time, data sharing still faces several barriers and challenges, which requires a multifaceted approach if they are to be addressed. Organisations often hesitate to share data due to fears of data breaches, unauthorised access, and misuse of sensitive information. Ensuring robust security measures and compliance with privacy regulations is a significant challenge that the data space architecture aims to solve. Inconsistent data formats, protocols, and standards across different organisations and industries can hinder seamless data integration and interoperability. Establishing universal standards is crucial but difficult due to varying requirements and practices. In addition, the reliability and accuracy of shared data are critical for effectively using data for decision-making. Inconsistent data quality and the absence of trusted data sources can undermine the value of data spaces. Building trust among data providers and users is essential. Navigating diverse and sometimes conflicting data protection laws and regulations across different jurisdictions can complicate data sharing. Defining clear governance structures, data ownership rights, and responsibilities among multiple stakeholders is complex. Organisations must ensure compliance with all relevant legal frameworks, placing additional challenges to the wide adoption of the data spaces. The costs associated with developing, implementing, and maintaining data spaces, including investments in technology, training, and ongoing operational expenses, can be prohibitive for many organisations. Thus, there are organisational challenges to implementing federated systems which aim to provide generic solutions that address many use cases. For specific use cases, it is often faster and simpler to implement an ad hoc solution that reads in the data needed to answer that specific use case that is limited in its scope, rather than developing more general solutions with wider applicability, such is the case with 3D city models and BIM.

In this proposition, UDTs are seen as a mediator in a federated data space. This represents a change of paradigm from the implementation of carefully crafted and realistic 3D city models built on a centralised (file) system and a closed platform that one can develop or acquire. There is an obvious need for a computational 3D visualisation platform; however, that is only one component of a UDT. A UDT can be a node in a distributed and federated data system, an urban data space of many actors, where neither data nor analytic models leave their owners and providers, who retain ownership and control access. An urban data space is a living digital data ecosystem, providing relevant, trusted and up-to-date information to a UDT run by a municipality.

We also presented a federated architecture for UDTs and other actors in an urban data space. The separation between local data schemas, federated linked schemas, and export data schemas in different layers brings a more modular structure and offers greater independence in developing UDTs. Furthermore, the separation of external data and applications is also critical, freeing the applications to connect to data based on open standards, but also to use local data schemas for optimal performance of specific use cases.

With this cross-disciplinary perspective, we contribute to the field of UDTs, advancing the understanding of their role in supporting municipal planning through the concepts of data spaces and federated database systems; and to the field of data spaces, by introducing a UDT as a new functional component with a specific mediator role, and by associating data spaces with the

application domain of UDTs. And we hope to offer a road map that contributes to unlocking the potential and promises of UDTs in supporting smarter planning and decision-making processes towards sustainable cities, delivering a better living environment for its citizens.

Acknowledgements

The authors thank the insights from the partners in the Data Models Milestone Project, colleagues from DTCC, and members of the Local Digital Twins Forum.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Chalmers Digital Twin Cities Centre (DTCC) supported by Sweden's Innovation Agency Vinnova under Grant No. 2019-00041, the European Union's H2020 WIDESPREAD-2018-2020 TEAMING Phase programme under grant agreement no. 857155 and H2020 research and innovation programme under grant agreement no. 101135576.

ORCID iD

Jorge Gil  <https://orcid.org/0000-0001-6671-2578>

Notes

1. <https://3dbag.nl/en/viewer>.
2. <https://kartta.hel.fi/3d/>.
3. <https://www.businesslocationcenter.de/berlin3d-downloadportal/?lang=en#/export>.
4. <https://mapgallery.calgary.ca/apps/bcd22e7089a440e792628ac61f35f4c1/explore>.
5. <https://data.stadt-zuerich.ch/dataset?q=&tags=3d-stadtmodell>.
6. <https://internationaldataspaces.org/>.
7. <https://www.data-infrastructure.eu/GAIAx/>.
8. <https://dssc.eu/>.
9. <https://www.ds4sscc.eu/>.
10. <https://dataspace.copernicus.eu/>.
11. <https://green-deal-dataspace.eu/>.
12. <https://mobility-dataspace.eu/>.
13. <https://vloca-kennishub.vlaanderen.be/>.
14. <https://databri-x.eu/>.

References

- Bauer M, Cirillo F, Fürst J, et al. (2021) Urban digital twins – a FIWARE-based model. In: *At - Automatisierungstechnik*. De Gruyter O, Vol. 69, 1106–1115.
- Biljecki F, Stoter J, Ledoux H, et al. (2015) Applications of 3D city models: state of the art review. *ISPRS International Journal of Geo-Information* 4: 2842–2889. DOI: [10.3390/ijgi4042842](https://doi.org/10.3390/ijgi4042842).
- Bolton A, Butler L, Dabson I, et al. (2018) Gemini principles. In: *Gemini Papers, Report*. Centre for Digital Built Britain. Available at: <https://www.cdbb.cam.ac.uk/DFTG/GeminiPrinciples> (accessed 10 February 2024).

- Calzati S (2023) No longer hype, not yet mainstream? Recalibrating city digital twins' expectations and reality: a case study perspective. *Frontiers in Big Data* 6: 2.
- Coenen T, Walravens N, Vannieuwenhuyze J, et al. (2021) Open Urban Digital Twins – insights in the current state of play. White paper, imec. Available at: <https://researchportal.vub.be/en/publications/open-urban-digital-twins-insights-in-the-current-state-of-play>.
- Curry E (2020a) Real-time linked dataspace: a data platform for intelligent systems within internet of things-based smart environments. In: Curry E (ed) *Real-Time Linked Dataspaces: Enabling Data Ecosystems for Intelligent Systems*. Cham: Springer International Publishing, pp. 3–14. DOI: [10.1007/978-3-030-29665-0_1](https://doi.org/10.1007/978-3-030-29665-0_1).
- Curry E (2020b) Dataspace: fundamentals, principles, and techniques. In: Curry E (ed) *Real-Time Linked Dataspaces: Enabling Data Ecosystems for Intelligent Systems*. Cham: Springer International Publishing, pp. 45–62. DOI: [10.1007/978-3-030-29665-0_3](https://doi.org/10.1007/978-3-030-29665-0_3).
- Data Spaces Support Centre (DSSC) (2023a) Building blocks, version 0.5. Available at: <https://dssc.eu/space/BBE/178421761/>.
- Data Spaces Support Centre (DSSC) (2023b) DSSC glossary, version 2.0. Available at: <https://dssc.eu/space/Glossary/176553985/>.
- Dognini A, Monti A, Kung A, et al. (2024) Blueprint of the common European energy data space. In: *Interoperability Network for the Energy Transition*. Munich, Germany. DOI: [10.5281/ZENODO.12609569](https://doi.org/10.5281/ZENODO.12609569).
- D'Hauwers R, Walravens N and Ballon P (2022) Data ecosystem business models: value and control in data ecosystems. *Journal of Business Models* 10(2): 1–30.
- European Commission (2020) *A European Strategy for Data*. Brussels, Belgium: COM, p. 66.
- Ferré-Bigorra J, Casals M and Gangolells M (2022) The adoption of urban digital twins. *Cities* 131: 103905.
- Gemini Council and Lamb K (2022) *Gemini Papers: What are connected digital twins? Gemini Papers*. Cambridge, UK: Centre for Digital Built Britain, 1, DOI: [10.17863/CAM.82194](https://doi.org/10.17863/CAM.82194).
- Herbert G and Chen X (2015) A comparison of usefulness of 2D and 3D representations of urban planning. *Cartography and Geographic Information Science* 42(1): 22–32. DOI: [10.1080/15230406.2014.987694](https://doi.org/10.1080/15230406.2014.987694).
- Jeddoub I, Nys G-A, Hajji R, et al. (2023) Digital Twins for cities: analyzing the gap between concepts and current implementations with a specific focus on data integration. *International Journal of Applied Earth Observation and Geoinformation* 122: 103440.
- Ketzler B, Naserentin V, Latino F, et al. (2020) Digital twins for cities: a state of the art review. *Built Environ, City Information Modelling: Digital Planning for Sustainable Cities* 46: 547–573.
- Lefever S and Michiels P (2019) City of things: an open smart city vision and architecture. In: *City of Things*. White Paper. Available at: https://www.imec.be/sites/default/files/inline-files/VLOCA_white_paper.pdf.
- Lehtola VV, Koeva M, Elberink SO, et al. (2022) Digital twin of a city: review of technology serving city needs. *International Journal of Applied Earth Observation and Geoinformation* 114: 102915.
- Lei B, Janssen P, Stoter J, et al. (2023) Challenges of urban digital twins: a systematic review and a Delphi expert survey. *Automation in Construction* 147: 104716.
- Nagel L and Lycklama D (2021) *Design Principles for Data Spaces*. Berlin: Position Paper. DOI: [10.5281/zenodo.5105744](https://doi.org/10.5281/zenodo.5105744).
- Neches R, Fikes RE, Finin T, et al. (1991) Enabling technology for knowledge sharing. *AI Magazine* 12(3): 36.
- Raes L, Michiels P, Adolphi T, et al. (2022) DUET: a framework for building interoperable and trusted digital twins of smart cities. *IEEE Internet Computing* 26(3): 43–50. DOI: [10.1109/MIC.2021.3060962](https://doi.org/10.1109/MIC.2021.3060962).
- Sepasgozar SME (2021) Differentiating digital twin from digital shadow: elucidating a paradigm shift to expedite a smart, sustainable built environment. *Buildings* 11(4): 4.
- Shahat E, Hyun CT and Yeom C (2021) City digital twin potentials: a review and research agenda. *Sustainability* 13(6): 4.
- Shahzad M, Shafiq MT, Douglas D, et al. (2022) Digital twins in built environments: an investigation of the characteristics, applications, and challenges. *Buildings* 12(2): 4.

- Sheth AP and Larson JA (1990) Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Surveys* 22(3): 183–236. DOI: [10.1145/96602.96604](https://doi.org/10.1145/96602.96604).
- Stoter J, Arroyo Ohori K, Dukai B, et al. (2020) State of the art in 3D city modelling - six challenges facing 3D data as a platform. In: *GIM International: The Worldwide Magazine for Geomatics*. Available at: <https://www.gim-international.com/content/article/state-of-the-art-in-3d-city-modelling-2>.
- Stoter J, Ohori KA and Noardo F (2021) Digital twins: a comprehensive solution or hopeful vision? Reflecting on the challenges in the geospatial domain. In: *GIM International: The Worldwide Magazine for Geomatics*. Available at: <https://www.gim-international.com/content/article/digital-twins-a-comprehensive-solution-or-hopeful-vision>.
- VanDerHorn E and Mahadevan S (2021) Digital twin: generalization, characterization and implementation. *Decision Support Systems* 145: 113524.
- White G, Zink A, Codecá L, et al. (2021) A digital twin smart city for citizen feedback. *Cities* 110: 103064.
- Wiederhold G (1992) Mediators in the architecture of future information systems. *Computer* 25(3): 38–49.
- Ye X, Du J, Han Y, et al. (2023) *Developing Human-Centered Urban Digital Twins For Community Infrastructure Resilience: A Research Agenda*. Sage Publications Inc, Vol. 38, pp. 187–199.

Jorge Gil is associate professor in Urban Analytics and Informatics at the Department of Architecture and Civil Engineering, Chalmers University of Technology, Sweden. His research focus is on the conceptualisation and development of integrated urban models for urban planning research and practice, in the context of Smart Cities, City Information Modelling (CIM) and Urban Digital Twins. He applies spatial analysis, network analysis, machine learning and visualisation methods to big data to advance research in the domains of sustainable mobility of people and goods, social inclusion, energy transition, and circular economy.

Dessislava Petrova-Antonova is a professor in Software engineering at the Faculty of Mathematics and Informatics and a Research leader of the Data Management Research Unit of GATE Institute, Sofia University. She is a coordinator of the GATE pilot project “City Digital Twin,” where a 3D City Model covering an area of Sofia is under development, along with simulations, analyses, and visualisations of the urban environment and processes. Her research interests include data management focusing on semantic interoperability, data enrichment, and domain-specific data models. She has over ten years of experience developing service-based software applications.

Graham Kemp is a professor in the Data Science and AI division at the Department of Computer Science and Engineering, Chalmers University of Technology and University of Gothenburg, Sweden. His main research interests are related to databases and bioinformatics, and include work with federated database architectures for life science data. He currently leads the data management and integration research area at the Chalmers Digital Twin Cities Centre.