Urban Digital Twins in Practice: A Survey-Based Analysis of Data Integration Levels and Approaches

Imane Jeddoub [1], Binyu Lei [2], Filip Biljecki [3,5], Rafika Hajji [4], Roland Billen [1]

[1] GeoScITY, Spheres Research Unit, University of Liège, 4000 Liège, Belgium – (i.jeddoub, rbillen)@uliege.be
[2] Urban Planning and Data Analytics, School of Geography, Earth and Environmental Sciences, University of Birmingham, Dubai, United Arab Emirates – b.lei@bham.ac.uk
[3] Department of Architecture, National University of Singapore, Singapore – filip@nus.edu.sg
[4] College of Geomatic Sciences and Surveying Engineering, Hassan II Institute of Agronomy and Veterinary Medicine, Rabat 10101, Morocco – r.hajji@iav.ac.ma
[5] Department of Real Estate, National University of Singapore, Singapore

Keywords: Urban Digital Twin, Data Integration, Lifecycle, CityGML, Standardization, Interoperability.

Abstract

Data integration is a significant technical challenge that hinders the full implementation of Urban Digital Twins (UDTs). Given their data-driven nature, UDTs rely on the integration of heterogeneous data sources, schemas, and structures. This process is inherently complex due to the diversity of techniques and systems involved. Furthermore, with the emergence of UDTs, the focus has shifted from acquiring new data through sensors to exploring how existing data can be integrated, restructured, and connected to meet the evolving requirements of digital twins. In this context, we raise fundamental questions about data integration within UDTs. Building on a previously proposed conceptual framework defining levels of data integration, this study gathers insights from practitioners actively engaged in UDT implementations. The aim is to understand how these conceptual integration levels are applied in practice, how they are selected across the UDT lifecycle, what tools support their implementation, and how they enable effective integration depending on the availability and nature of the data. By aligning theoretical frameworks with practical implementations, this research aims to enhance the effectiveness and adoption of scalable and flexible data integration strategies in the development, use, and maintenance of UDTs. The findings highlight the relevance of data integration levels for UDT practitioners, demonstrating that they were able to identify which levels are involved in their projects. Additionally, the results provide insights into how each level of integration is commonly applied across different stages of the lifecycle, as well as the tools and methods associated with each level.

1. Introduction

Multisource data integration has long been a field of significant interest within the geospatial and urban communities (Noardo, 2022). Many projects and standardization initiatives have been launched to address data integration and interoperability challenges, especially at a time when use case applications are diverse and rapidly evolving (AlBalkhy et al., 2024; Santhanavanich et al., 2022).

While integration challenges have traditionally been explored in the context of 3D city models, BIM-GIS integration, the convergence of GIS-IoT or BIM-IoT systems, spatial data infrastructure and City Information Models for smart cities (Xia et al., 2022), these issues are increasingly reframed within the paradigm of Urban Digital Twins (Xu et al., 2024). Urban digital twins (UDTs) are conceptualized as complex systems of systems, in which the bidirectional synchronization between the physical and digital worlds must be maintained through standardized, interoperable, and real-time data exchange mechanisms supporting various analyses through simulation, prediction, and visualization tools. Despite their transformative potential, UDTs remain hindered by both technical and non-technical challenges (Lei et al., 2023). Recent studies have highlighted key issues related to digital twins, including the ambiguity around the definition in the built environment, as well as challenges concerning data integration, completeness, accessibility, and overall data quality, among others (Lei et al., 2023).

A first step toward building consensus on the definition of digital twins in the built environment has been undertaken through robust and novel systematic methods, including Natural Language Processing (NLP) techniques and systematic reviews, to analyze definitions and components across disciplines and over time (Abdelrahman et al., 2025). The study defines the city/urban digital twin as a spatial-temporal digital representation of a real-world urban area or city across its lifecycle. In practice, digital representation is usually considered to be a 3D city model. 3D city models are a key component in the creation of a UDT (Stoter et al., 2020). They are considered part of a UDT rather than a UDT per se. In fact, the shift from a static to an enriched and dynamic city model under the umbrella of DT raises issues in terms of data integration and interoperability. Furthermore, the new paradigm of using federated data spaces and modular IT architectures facilitates the shift from traditional urban management tools to a distributed data platform allowing integration at the semantic level within a specific application domain (Gil et al., 2024). Although the state of the art identifies data integration as a major challenge for the full implementation of UDT, it is not explored in depth, and the integration is only tackled from a case study perspective, according to the UDT requirements and applications. This makes generalization and comparison across UDT projects difficult. Moreover, it remains unclear how practitioners implement update mechanisms and integration strategies in practice, and how these relate to the different levels of data integration across the UDT lifecycle. From this perspective, Yan et al. (2025) propose a common data environment (CDE) structured around three enablers: data sources, data management (comprising functional components), and data consumers, to address the scalability of DT applications from buildings to city-level implementations. Moreover, data integration is highlighted as a key functional component of the data management enabler. Furthermore, the authors identify commonly used data integration methods, namely schema- and ontology-based methods, service- and system-based methods, and process- and tool-based methods, by summarizing the characteristics of these approaches and illustrating them through examples involving heterogeneous data. Following foundational studies on data integration and interoperability challenges, this work is a complementary study that investigates current practices and perceptions of data integration among Urban Digital Twin practitioners. Building on prior theoretical work on levels of data integration in UDTs, we design and deploy a structured questionnaire targeting professionals engaged in real-world UDT initiatives. Through this, we aim to examine:

- How is data integration performed in current UDT implementations?

- To what extent are the defined levels of integration reflected in practice?

- How do practitioners approach update mechanisms and integration strategies across current UDT implementations?

By surveying responses across diverse practitioners, this study uses the questionnaire as its main instrument to explore integration and interoperability challenges within UDTs. The survey approach enables us to systematically capture expert insights on current integration practices, technical barriers, and strategic needs. These insights help validate and extend existing empirical studies and support a more structured understanding of integration levels and methods. Ultimately, this allows us to identify how integration and interoperability issues affect the scalability and long-term sustainability of UDTs, and what practical steps could address them. By synthesizing responses across diverse practitioners, this study complements existing empirical studies on data integration practices within UDTs, supporting a more systematic understanding of integration levels and approaches. Our primary objective is to address the challenges inherent in designing efficient, flexible, and standardized mechanisms for data integration in UDTs. These integration challenges are intrinsically linked to issues of data interoperability and scalability, which are critical for the effective implementation and sustainability of complex urban environments such as cities. This study builds on research conducted as part of a PhD thesis focusing on data integration in Urban Digital Twins (Jeddoub, 2026).

The structure of the paper is as follows: Section 2 reviews related work on Urban Digital Twins (UDTs), with a specific focus on data integration. Section 3 outlines the research methodology and the framework design of the questionnaire. Section 4 analyses the survey responses and offers detailed insights into current UDT projects, including their scale, lifecycle stages, and levels of integration. Section 5 discusses the findings and their implications in the context of standardization and interoperability. Finally, Section 6 concludes the paper and outlines future research directions.

2. Related work and Background

2.1 Research background

From 2016 until now, the corpus of papers on Urban Digital Twins has expanded significantly. While some publications delve into the technological frameworks and potential of UDTs (AlBalkhy et al., 2024; Weil et al., 2023), others address the conceptual, organizational, and technical challenges related to their integration within geospatial and urban fields. Motivated by the growing implication and demonstrated value of digital twins across various spatial contexts, an increasing number of cities and municipalities have started investing in UDT projects. These initiatives have evolved into national and international collaborations aimed at establishing interoperable, scalable, and robust spatial and urban data infrastructures to support the creation of "true" urban digital twins. This shift highlights a strategic move toward data-driven urban governance, sustainability, and real-time city management.
Inspired by existing studies on the adaptation and adoption of UDTs in the built environment, this research aims to explore practitioners' perspectives on the challenges of data integration and how they navigate issues of interoperability. To derive meaningful insights, this study extends earlier studies conducted by Jeddoub et al. (2023, 2024) and aligns with prominent works in the field of city and urban digital twins.

2.2 Survey-Based Investigations of Urban Digital Twins: A Review of Recent Studies

Many scholars have recognized the value of combining literature reviews with survey-based analyses to gain comprehensive insights into UDTs. For instance, Ellul et al. (2022), under the auspices of EuroSDR, investigated the challenges and opportunities associated with digital twins for National Mapping and Cadastral Agencies (NMCAs) and other stakeholders in the location sector. Their findings, derived from an international survey and workshop focused on digital twin modelling, shed light on the geospatial implications in such DT processes. Similarly, Lei et al. (2023) conducted an exhaustive review of UDT-related challenges, emphasizing data integration as one of the most critical issues practitioners encounter when creating digital twins. Moreover, a study leveraging Natural Language Processing (NLP) techniques analysed over 15,000 scientific publications, combining the results with expert surveys to reach a consensus on the definition of digital twins in the built environment (Abdelrahman et al., 2025). This approach helped clarify the evolving nature of this concept. Another related study develops a governance assessment framework for UDTs using a mixed-methods approach: an extensive literature review, a four-round Delphi expert survey among 9 international panelists, and the Best-Worst Multi-Criteria Decision-Making (MCDM) method (Diaz-Sarachaga, 2025). Furthermore, Yan et al. (2025) emphasize the importance of implementing structured stakeholder engagement methods, such as surveys, to capture stakeholder use case capabilities in the development of Common Data Environments (CDEs) for UDTs. Continuing within the framework of EuroSDR's initiatives, a recent publication explores the growing integration of Geographic Information Science (GIS) with National Digital Twins (NDTs), emphasizing key fields of interest, namely geospatial data management, standardization, governance and public values (Ellul et al., 2024). It presents six expert perspectives on the challenges facing National DTs, offering a research agenda for advancing their implementation at the national scale.

2.3 Levels of integration

Building on previous work on data integration levels (Jeddoub et al., 2024), we have defined three levels of data integration for UDT in practice.

Level 1 (Conceptual Data Model) focuses on extending, profiling, or combining the existing Conceptual Data Model to incorporate new information and features required by specific applications, such as those needed in Urban Digital Twins (UDTs). For instance, Level 1 has been applied in the early development of the City of Vienna's Digital Twin by extending the CityGML 2.0 data model through Application Domain Extensions (ADEs) (Lehner et al., 2024).

https://doi.org/10.5194/isprs-annals-XII-4-W2-2026-65-2026 | © Author(s) 2026. CC BY 4.0 License.

Level 2 (Database Level) involves aligning data with the existing Conceptual Data Model and standards without altering the underlying data model. This process ensures that new data conforms to predefined structures and relationships, while maintaining the semantic consistency of the model. Diakite et al. (2022) performed data integration at this level by mapping heterogeneous spatial datasets (e.g., buildings, roads, waterbodies) to CityGML 2.0 classes and storing them in 3DCityDB using PostgreSQL/PostGIS. Although a front end (CesiumJS) is used for visualization, it primarily queries the backend without performing real-time data integration.

Level 3 (Front-End Integration) emphasizes the integration of data through various system architectures, including web interfaces and tools such as GIS applications, web applications, geoportals, dashboards, linked data, federated platforms, data spaces, and game engine applications. This level facilitates the visualization and management of complex 3D data directly within modern web browsers, addressing challenges related to data standardization and syntactic interoperability. Beyond visualization, this level facilitates the development of customized applications that support urban planning, infrastructure management, and environmental monitoring by integrating diverse data sources into a unified or connected platform. In this perspective, Alva et al. (2024) present a bottom-up approach for developing a UDT platform using openly available 3D datasets, primarily focused on energy systems and decarbonisation scenarios. The integration process involves harmonising diverse spatial and temporal datasets into a unified 3D model, enabling real-time querying, simulation, and visualisation through a user-friendly dashboard.

These data integration levels were illustrated earlier by Jeddoub et al. (2024). The authors provide an analytical framework for understanding these integration levels within the UDT lifecycle phases: creation, use, and update, and highlight the challenges and potential benefits associated with each level. While a literature review was conducted based on scientific publications, project reports, and websites, engaging directly with UDT practitioners will provide a more detailed perspective on their actual implementation.

3. Methodology: Context and survey protocol

Online questionnaires have become a widely used and efficient tool for data collection in both academic and applied research contexts, including social and urban studies (Krause et al., 2024). They offer flexible and cost-effective means of gathering diverse and large volumes of data within a short period, with reduced error rates and minimal resource requirements. Online surveys support a range of question types, both quantitative and qualitative (e.g., multiple-choice and open-ended), and can incorporate logic-based pathways to maintain thematic focus and ensure relevance across different respondent groups. Moreover, they are increasingly recognized as a valuable instrument for community engagement. For instance, urban planners use online surveys to collect demographic data, user feedback, community needs, and citizen perceptions regarding services, infrastructure, and public facilities, thereby informing and enhancing the decision-making process.

3.1 Survey questions: contextualization and definition

With the emergence of Digital Twins, data management practices have undergone significant transformation (federated data spaces ecosystem, Open Urban Platform, Spatial Data infrastructure). Unlike traditional approaches where data was managed in isolated silos, Digital Twins promote integrated and interoperable data federated system. This paradigm shift encourages the identification of relevant datasets and the adaptation of tools to enable their processing, harmonization, and integration within a unified system. However, given the diverse purposes of various applications, establishing a single, comprehensive model that accommodates all downstream use cases remains a significant challenge. As a result, current efforts tend to rely on adaptable mechanisms that leverage existing standards and data structures to progressively enrich Digital Twin implementations. Given their multi-disciplinary nature, UDTs are data-driven tools that integrate multi-source data, various spatial and temporal scales and domain-specific information. The heterogeneity and the complexity of data in practice raise the data integration challenges, which have led us to fundamental questions regarding how data are or will be effectively integrated into an urban digital twin, and what challenges are associated with these integrations. Our primary goal was to assist practitioners in documenting their initiatives and identifying their lifecycle stages, familiarizing them with various levels of integration, and understanding the associated advantages and challenges. Additionally, we aimed to explore the update mechanisms adopted by practitioners and how these relate to different data integration levels and implementation practices. In this context, we formulated the following questions (refer to Table 1). The survey is designed to collect responses anonymously; nevertheless, we included a question requesting participants' email addresses and serves the purpose of enabling follow-up communication for clarifying survey responses, if necessary.
To maximize survey responses and increase the response rate, we compiled a list of participants actively involved in the digital twin field and contacted them directly by email to explain the objectives of the questionnaire. The survey remained accessible to the broader community, resulting in additional responses from individuals not directly contacted. A total of 166 participants were approached, but only 35 responses were collected. While we initially anticipated a higher response rate, the 35 completed surveys provide valuable insights. In particular, these respondents are experts in the field of digital twins, which increases the relevance and quality of the data collected. Moreover, when comparing our study with similar studies in terms of approach and analysis (see Section 2), the number of responses and the resulting findings can be considered reasonable and sufficient for the purposes of this analysis.

Questions	Purpose
Participants profile	
Q1: What is your background?	Identify participants' background information (i.e., Academia, Industry, or Government/Policymaking).
Q2: What are your fields of expertise?	Identify participants' fields of expertise (Urban modelling, Geoinformation, Engineering technology, Building/City analysis, ICT).
UDT project: Basic information	
Q3: What is the project's name? Could you please provide a short description of the	Obtain a concise overview of the project's purpose, scope, and goals.

project or a link to the project?	
Q4: What is the project's spatial scale?	Understand the project's spatial extent.
Q5: What is your current stage in the DT lifecycle?	Determine the current stage in the DT lifecycle.
UDT: data integration levels	
Q6: What is your level of data integration?	Introduce the data integration levels and evaluate their applicability for practitioners to determine whether their initiatives operate through one or more integration levels.
Q7: How do you perform such data integration?	Identify common methods and tools used for data integration.
Q8: Do you find these data integration levels relevant to your UDT implementation?	Identify positive, neutral, or negative feedback about the relevance of data integration levels.
Q9: Do you have any proper strategies and mechanisms to support the update of your UDT?	Identify and understand current update strategies and practices if applicable.

Table 1. Framework design of the survey questions.

4. Results

4.1 Participants profile

We initially collected participants' background information. 24 respondents are from Academia, 7 from Government/Policymaking and 4 from Industry (Figure 1.a). Participants' field of expertise is mostly from Geoinformation with 30 experts (Figure 1.b). The participant's geographical locations are deduced from the project initiatives with most UDT projects conducted in Europe (22) (Figure 1.c). Many European countries have reported their initiatives, including the Netherlands (6), Germany (6), Switzerland (1), Poland (1), Austria (1), Sweden (1), Croatia (1), Slovenia (1), Bulgaria (1), and the United Kingdom (1). Additionally, two projects were reported at a European scale by participants from Italy.
Beyond Europe, we had 7 participants from the Asia-Pacific region, including Singapore (2), Australia (3), and South Korea (2), as well as 4 participants from North America, comprising (3) from the United States and (1) from Canada.

4.2 UDT projects

4.2.1 UDT scale: We asked participants to specify the spatial scale of their UDT projects by selecting from various predefined options: National, Regional, City, District, Quarter/Block. This approach allowed us to categorize projects based on their spatial scale, facilitating a better understanding of the distribution of current UDT implementations. The findings reveal that many initiatives (13) are taking place on a city-wide scale, (9) at national levels and (6) at district level. However, 7 initiatives have identified other scales, namely: Global (1), European (3), Campus (1), and multiscale projects, such as those encompassing various municipalities (1), testbeds (1), and city blocks (1) (refer to Figure 2.a).

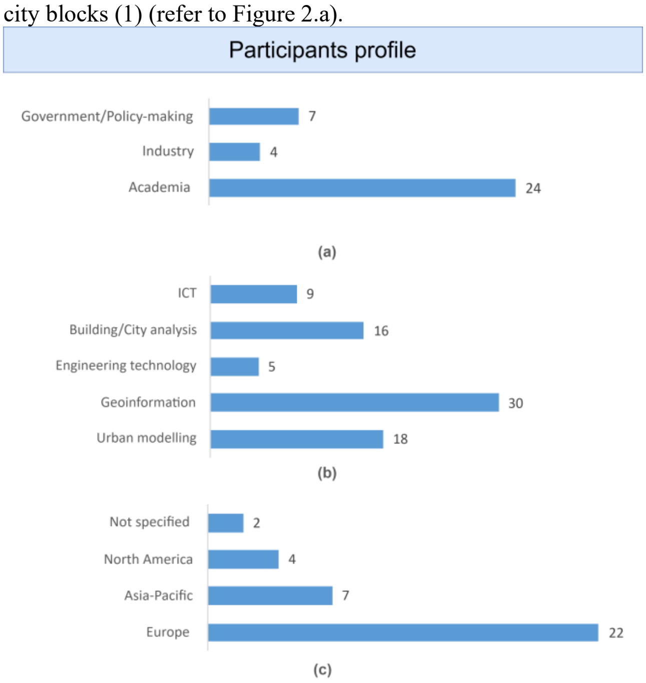

Figure 1. Participants profiles (a) Background, (b) Fields of Expertise and (c) Geographical Locations.

4.2.2 UDT lifecycle: UDTs are not static representations of reality at a single snapshot; rather, they are dynamic systems that evolve through continuous integration and real-time data enrichment. Understanding the current position of various initiatives within the digital twin lifecycle is crucial for assessing their maturity and operational status. Our analysis reveals that out of the surveyed initiatives (see Figure 2.b): (18) projects are in the creation phase, indicating they are currently under implementation. (13) projects have reached the use phase, demonstrating operational functionality and (4) initiatives are in the update phase. This distribution underscores the progressive development of UDT projects, with a significant number transitioning from conceptualization to active use and continuous maintenance.

4.2.3 Cross-analysis of the scale and lifecycle: We have presented the lifecycle stages per spatial scale and concluded that several initiatives are city-wide and span the entire lifecycle of the digital twin (Figure 3). National-scale DTs are at the creation stage, although a few initiatives are at the usage or update stage. While for other urban scales, notably districts, blocks, and municipalities, they are more at the creation and use phases. From this analysis, we can conclude that efforts are highly concentrated at the city scale, with 13 initiatives, including 5 in the creation phase, 5 in the use phase, and 3 in the update phase.

4.2.4 UDT levels of integration: Participants were asked to specify the integration level(s) adopted in their UDT projects. Overall, (16) projects reported using Level 1, (20) selected Level 2, and (16) adopted Level 3. Figure 2.c shows individual counts per integration level, as several projects reported using multiple levels simultaneously. Consequently, the total number of reported integration levels exceeds the total number of surveyed projects (35). This indicates that UDT projects often combine several integration levels rather than relying on a single one.

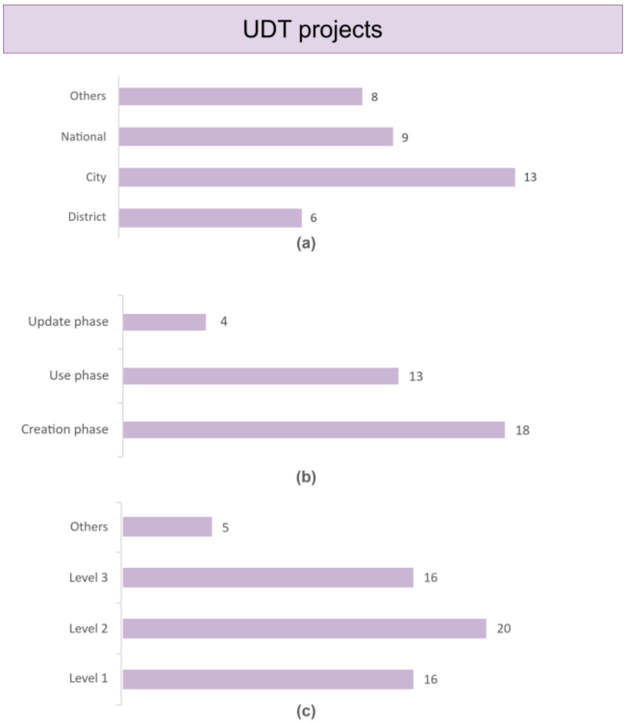

Figure 2. UDT projects: (a) scale, (b) lifecycle stages and (c) level of integration.

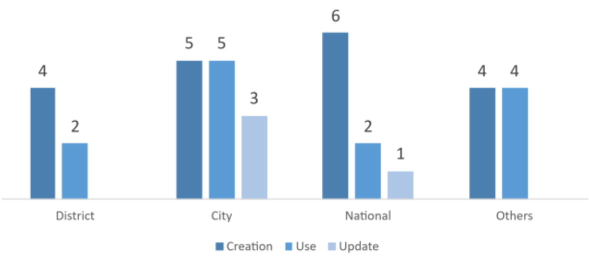

Figure 3. UDT projects: comparative analysis of the spatial scale and the lifecycle.

Overall, 32 of the 35 respondents were able to associate their projects with one or more of the proposed integration levels. The remaining three respondents selected the "Others" option and provided explanations that did not fit the current classification. These responses reveal three distinct situations. One respondent explained that different components of the same UDT belong to different integration levels, indicating that the proposed levels are complementary rather than mutually exclusive. Another respondent described their work as linking geospatial and management databases at the object level rather than performing data integration as defined in this study, suggesting an alternative interpretation of integration. The third respondent indicated that the question itself was not sufficiently clear, highlighting difficulties in interpreting the proposed classification. Beyond these three responses, additional comments from respondents who selected one or more integration levels also revealed uncertainty regarding the boundaries between the proposed levels. For example, one practitioner questioned whether a front-end that only queries an existing database without performing data integration should be classified as Level 3. These comments suggest that the distinction between database-level and front-end integration may require further clarification.

These responses provide insight into the current boundaries of the proposed framework. The coexistence of multiple integration levels within the same project indicates that integration is often layered, while the object-level linking response suggests that some integration practices are not explicitly represented by the current classification. Together, these findings show that the survey not only supports the proposed framework but also identifies areas where it can be refined. Clarifying the boundaries between integration levels and considering additional integration practices could improve the framework's clarity and applicability. They also reflect the diversity of data integration approaches among current UDT initiatives, suggest opportunities to improve future survey design, and contribute to refining the proposed framework and extending the current understanding of data integration levels within UDTs.

4.2.5 Cross-analysis of the levels of integration and the lifecycle: To gain a better understanding of these integration levels and their distribution throughout the digital twin's lifecycle, we conducted a cross-analysis to identify how many projects utilize each integration level during each phase (refer to Figure 4). The analysis indicates that projects in the creation phase predominantly rely on Level 1 (extension of conceptual models) and Level 2 (adaptation of data to meet model requirements) integration methods. In contrast, projects in the use phase extensively employ Levels 2 and 3. Notably, in the update phase, Level 1 integration is absent. This suggests that if Level 1 was utilized during creation, subsequent updates are likely performed using Level 2, as the conceptual model can now accommodate new data, or Level 3, when external APIs are employed. Identifying these patterns highlights the importance of selecting appropriate integration levels at each phase of the lifecycle to ensure seamless data flow and system interoperability.

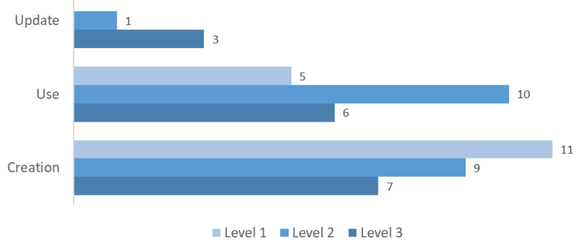

Figure 4. UDT projects: comparative analysis of the lifecycle and the levels of integration.

4.2.6 Data integration - Methods and tools: The aim of this section is to identify and understand the common methods and tools used for data integration (e.g., APIs, GIS, BIM, ETL processes, database or cloud-based solutions). The open-ended responses were analyzed to identify the software, standards, and methodologies reported by practitioners. These tools were then compared across the proposed integration levels to identify common implementation patterns (refer to Figure 5). Our observations indicate that participants emphasized different aspects:

Software and Tools: 3DCityDB, cloud storage, APIs, transactional Web Feature Service (WFS-T), semantic databases, CesiumJS, ETL tools, ArcGIS Pro, data catalogs, container applications with Kubernetes, and Open Urban Platform.

Standards: OGC standards, CityGML, CityJSON, IFC, SensorThings API, MUDDI, EnergyADE, LADM, 3D Tiles, and new data models.

Methodologies and Workflows: DDL with SQL scripts, self-written processing pipelines, Python and C++ pipelines, profiles, and extensions.

To better understand how these methods are applied in practice, the reported software, standards, and methodologies were compared across the proposed data integration levels. The comparison shows that most initiatives operating at Level 1 primarily focus on the common standards, aiming to meet the specific needs related to digital twin implementation. When it comes to workflows and methodologies, the emphasis is more specifically on the adaptation of existing standards or the definition of new models, particularly in relation to extensions and profiling. At Level 2, all three categories (software, standards, and methodologies) are represented. Regarding the methodologies and workflows category, the focus is mainly on the use of FME pipelines, which support the transformation and conversion of data to align with the defined standards. These pipelines enable efficient conversion and alignment of datasets with the existing standards at Level 2, promoting interoperability and consistency across systems. At Level 3, the emphasis shifts clearly towards software and tools. This reflects a higher level of technological maturity, where standards and workflows have already been integrated, and the primary focus is now on automation, optimization, and operational efficiency.

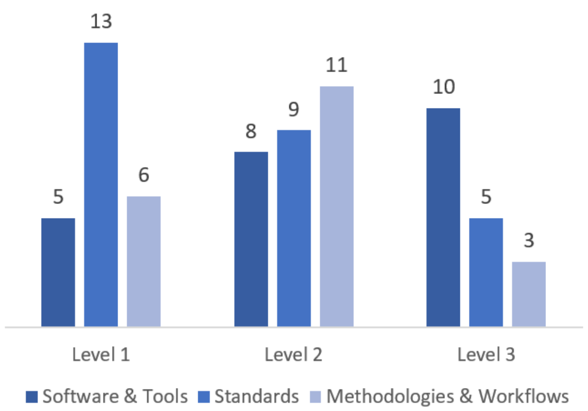

Figure 5. UDT projects: Analysis of the levels of integration based on the tools and methods.

4.2.7 Relevance of Data Integration Levels to UDT Implementation: We asked participants about the relevance of the data integration levels and whether they resonate with their own implementations (refer to Figure 6). 89% of respondents from diverse backgrounds confirmed the relevance of the data integration levels (refer to Figure 6.a). The high proportion of "Yes" and "Maybe" responses further emphasizes that, while some participants fully recognized the relevance of the levels (54% responses) (refer to Figure 6.b), others may have experienced partial alignment while reflecting their alignment with ongoing Urban Digital Twin (UDT) development efforts. Indeed, 34% of practitioners responded "Maybe," indicating some level of uncertainty (See Figure 6.b). However, these participants were still able to recognize the levels of integration deployed in their implementations, suggesting that the classification is both applicable and understandable, implying continued interest in the UDT development process. A minority (11%) responded "No," highlighting potential gaps in how the classification aligns with the tools, workflows, or data practices used in their specific contexts.

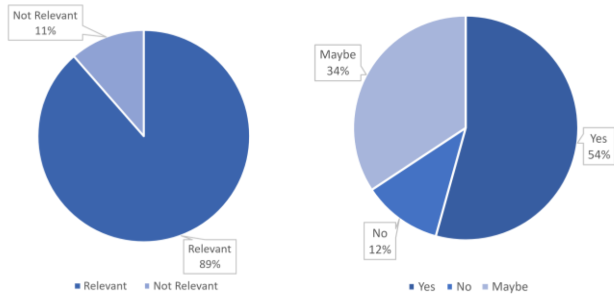

Figure 6. Relevance of data integration levels (a) among practitioners and (b) for UDT implementation.

4.3 Data integration levels assessment: A multi-dimensional approach

This section summarizes the multi-dimensional study that we followed to assess data integration levels within Urban Digital Twin (UDT) projects. The analysis is grounded in UDT projects and draws on multiple dimensions that correlate with integration, including **scale** (district, city, regional, national), lifecycle stage (creation, use, update), methods and tools (e.g., GIS, BIM, FME pipelines, standards), purpose and use case (e.g., visualization, simulation, prediction, monitoring). By analyzing integration through these dimensions, we aim to highlight the complexity and diversity of data integration approaches used in real-world projects, rather than relying solely on the systematic literature review. To support this analysis, various UDT projects were reviewed and analyzed. Each project illustrates how it aligns with the proposed integration levels based on the use case and highlights a specific scenario of integration. The list of projects is publicly available in a GitHub repository (refer to Appendix), providing a shared reference for ongoing UDT initiatives.

4.4 Exploring UDT update mechanisms across data integration levels

In the framework design of the survey, practitioners were asked whether they have set any proper strategies and mechanisms to support the update of their UDT. While the challenge of maintaining UDTs should ideally be addressed earlier, even before any actual updates occur (e.g., while designing and building UDTs), we wanted to investigate whether such update mechanisms are already being considered in practice. In addition, we aimed to identify the update mechanisms currently adopted by practitioners and explore how these mechanisms relate to the different data integration levels. This question was included based on the understanding that the level of integration adopted in the creation and use stages may influence the appropriate level of integration at the update phase. In some cases, the choice may even lead to a shift in the integration level later on, once the digital twin has already been established and/or utilized. The responses gathered from the survey reveal that strategies for updating UDTs are still evolving, with varying degrees of integration and automation. While some practitioners have clearly defined mechanisms, such as using Cesium Ion for cloud storage and APIs for real-time data updates, others have yet to formalize their update strategies, with some still relying on manual updates, simple front-end modifications, or retrieving data directly from their data owners.

In addition, some practitioners highlight the role of standards in bringing together the various DTs implemented in the Netherlands. Other projects rely on the implementation of APIs to connect to urban data platforms and external data sources, enabling automatic updates to the UDT whenever the underlying database is refreshed. These responses highlight the diversity in maturity levels and approaches across UDT projects. The reported update mechanisms vary across the different integration levels, suggesting a relationship between the adopted integration level and the way UDTs are maintained and updated. The reported update mechanisms also differ in their complexity and degree of automation across the three integration levels. For instance, at level 1, updates are primarily managed through versioning systems and ADEs, such as Scenario ADE. At level 2, ETL tools and predefined rules are employed to automate the data flow, ensuring that new data integrates smoothly with the established database. At level 3, updates are frequently automated through the integration of real-time data APIs, such as those from weather stations, IoT devices, or urban sensors. These updates are then visualized on the front end without requiring manual intervention. Additionally, practitioners may use external tools that allow for automated data ingestion, ensuring that new data is regularly pushed to the system. This level enables a dynamic and interactive digital twin, where updates happen in near real-time, and the data continuously reflects the evolving state of the urban environment. These findings characterize the update mechanisms currently adopted in UDT implementations. However, because the questionnaire did not explicitly distinguish between predefined strategies and those developed during project implementation, the results should be interpreted as reflecting current practices rather than the evolution of integration strategies over time.

5. Discussion

This work summarizes and analyzes the survey responses on data integration levels and their practical applications. Although the survey was distributed to over 160 experts in the digital twin field, most responses were derived from academia. This highlights the need to extend the survey and enrich the respondent base by including more representatives from industry, government, and policy-making sectors. This will enable, for example, comparisons between practitioners to examine how perspectives on data integration and updating strategies vary across these sectors. Furthermore, this work serves as a foundational framework to support ongoing efforts in standardization and interoperability, as well as to better understand how existing practices align with established data integration frameworks and standards. For instance, initiatives led by the Open Geospatial Consortium (OGC), such as the Urban Digital Twin Interoperability Pilot (UDTIP) and Testbed-18: Building Energy Data Interoperability, emphasize the importance of a robust data integration framework. These initiatives focus on combining diverse datasets required for use cases in energy domains, infrastructure planning, and monitoring, while also addressing the repetitive and non-standardized methods traditionally applied for the harmonization, mapping, and integration of these datasets. These initiatives leverage different levels of data integration (Energy ADE, enrichment of the data model and processing service and client-side integration). The need for standardized approaches to improve the interoperability of data across various sectors is highlighted in these OGC projects, which aim to create integrated, efficient, and scalable solutions for urban digital twins and encourage the adoption of OGC standards.

6. Conclusion

The survey findings indicate that data integration plays an important role throughout the lifecycle of UDTs. Data integration is present from the creation phase, where geospatial data is integrated and aligned with urban information models, to the use phase, where it supports application-specific needs, as well as during updates and change-detection processes. In this work, we aimed to understand how digital twin practitioners address the technical challenges of data integration. The practitioners also identified the integration levels that were either applied or adapted to suit their project needs. Through a survey-based approach, this study highlights the interrelationships between the most advanced current spatial scales of digital twins, draws comparisons between integration levels, spatial scale, lifecycle phases, and the tools and methodologies used to achieve effective integration. While our categorization of data integration levels appears to cover most current implementations, it is likely that new levels will emerge, particularly with the integration of Generative Artificial Intelligence (GenAI) models into the design, construction, and operation of UDTs. The integration levels reported in this study are primarily based on practitioners' self-assessment. To support a consistent interpretation, the questionnaire included definitions of the proposed integration levels to facilitate respondents' understanding of the classification. Following the survey, the reported integration levels were reviewed for consistency against the project descriptions and the reported integration methods to assess whether the selected classifications aligned with the characteristics of each UDT project (see Appendix). Nevertheless, differences in interpretation may have influenced some classifications, as reflected by the responses discussed in Section 4.2.4. Future work could further strengthen this validation process through independent assessments of documented UDT implementations. It is important to mention some limitations regarding the representativeness and the distribution of the survey. Most respondents are from academia (24 out of 35), with only a small number representing government or policymaking organizations (7) or the private sector (4). From a geographical perspective, most reported UDT projects are based in Europe, which reflects where much of the current UDT research activity is concentrated. While this is not surprising given how the survey was distributed, it does mean that the findings are more representative of academic and European perspectives than of the broader UDT community. Insights from industry practitioners, municipal governments, or projects outside Europe may differ in significant ways, for instance, in how integration decisions are driven by commercial constraints or different regulatory environments. Expanding the respondent base in future work to include more diverse stakeholders and geographic contexts would strengthen the generalizability of these findings.

References

Abdelrahman, M., Macatulad, E., Lei, B., Quintana, M., Miller, C., Biljecki, F., 2025. What is a Digital Twin anyway? Deriving the definition for the built environment from over 15,000 scientific publications. *Building and Environment*, 274, 112748. doi.org/10.1016/j.buildenv.2025.112748.

AlBalkhy, W., Karmaoui, D., Ducoulombier, L., Lafhaj, Z., Linner, T., 2024. Digital twins in the built environment: Definition, applications, and challenges. *Automation in Construction*, 162, 105368. doi.org/10.1016/j.autcon.2024.105368.

Alva, P., Mosteiro-Romero, M., Stouffs, R., 2024. Estimating Operational Greenhouse Gas Emissions in the Built Environment using an Urban Digital Twin: Sustainable City Management Tool for Decarbonisation of Cities. *Proceedings of the 29th International Conference of the Association for Computer Aided Architectural Design Research in Asia (CAADRIA) 2024*, 365–374. doi.org/10.52842/conf.caadria.2024.2.365.

Diakite, A.A., Ng, L., Barton, J., Rigby, M., Williams, K., Barr, S., Zlatanova, S., 2022. LIVEABLE CITY DIGITAL TWIN: A PILOT PROJECT FOR THE CITY OF LIVERPOOL (NSW, AUSTRALIA). *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W2-2022, 45–52. doi.org/10.5194/isprs-annals-X-4-W2-2022-45-2022.

Diaz-Sarachaga, J.M., 2025. Developing an assessment governance framework for urban digital twins: Insights from smart cities. *Cities*, 156, 105558. doi.org/10.1016/j.cities.2024.105558.

Ellul, C., Stoter, J., Bucher, B., 2022. LOCATION-ENABLED DIGITAL TWINS – UNDERSTANDING THE ROLE OF NMCAS IN A EUROPEAN CONTEXT. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W2-2022, 53–60. doi.org/10.5194/isprs-annals-X-4-W2-2022-53-2022.

Ellul, C., Stoter, J., Bucher, B., Olsson, P.-O., Billen, R., DeLathouwer, B., 2024. Towards National Connected Digital Twins - A Geospatial Perspective. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4-W5-2024, 147–154. doi.org/10.5194/isprs-annals-X-4-W5-2024-147-2024.

Gil, J., Petrova-Antonova, D., Kemp, G.J., 2024. Redefining urban digital twins for the federated data spaces ecosystem: A perspective. *Environment and Planning B*, 23998083241302578. doi.org/10.1177/23998083241302578.

Jeddoub, I., Nys, G.-A., Hajji, R., Billen, R., 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. doi.org/10.1016/j.jag.2023.103440.

Jeddoub, I., Nys, G.-A., Hajji, R., Billen, R., 2024. Data integration across urban digital twin lifecycle: a comprehensive review of current initiatives. *Annals of GIS*, 1–20. doi.org/10.1080/19475683.2024.2416135.

Jeddoub, I., 2026. Urban Digital Twins Levels of Integration: From Conceptualization to Technical Implementation. Doctoral thesis, University of Liège, Liège, Belgium. hdl.handle.net/2268/343029.

Krause, R.M., Fatemi, S.M., Nguyen Long, L.A., Arnold, G., Hofmeyer, S.L., 2024. What is the Future of Survey-Based Data Collection for Local Government Research? Trends, Strategies, and Recommendations. *Urban Affairs Review*, 60(3), 1094–1115. doi.org/10.1177/10780874231175837.

Lehner, H., Kordasch, S.L., Glatz, C., Agugiaro, G., 2024. Digital geoTwin: A CityGML-Based Data Model for the Virtual Replica of the City of Vienna. In: Kolbe, T.H., Donaubauer, A., Beil, C. (Eds.), *Recent Advances in 3D Geoinformation Science*. Springer Nature Switzerland, Cham, 517–541. doi.org/10.1007/978-3-031-43699-4_32.

Lei, B., Janssen, P., Stoter, J., Biljecki, F., 2023. Challenges of urban digital twins: A systematic review and a Delphi expert survey. *Automation in Construction*, 147, 104716. doi.org/10.1016/j.autcon.2022.104716.

Noardo, F., 2022. Multisource spatial data integration for use cases applications. *Transactions in GIS*, 26(7), 2874–2913. doi.org/10.1111/tgis.12987.

Santhanavanich, T., Padsala, R., Würstle, P., Coors, V., 2022. THE SPATIAL DATA INFRASTRUCTURE OF AN URBAN DIGITAL TWIN IN THE BUILDING ENERGY DOMAIN USING OGC STANDARDS. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W2-2022, 249–256. doi.org/10.5194/isprs-annals-X-4-W2-2022-249-2022.

Stoter, J.E., Ohori, G.A., Dukai, B., Labetski, A., Kavisha, K., Vitalis, S., Ledoux, H., 2020. State of the art in 3D city modelling: Six challenges facing 3D data as a platform. *GIM International: the worldwide magazine for geomatics*, 34.

Weil, C., Bibri, S.E., Longchamp, R., Golay, F., Alahi, A., 2023. Urban Digital Twin Challenges: A Systematic Review and Perspectives for Sustainable Smart Cities. *Sustainable Cities and Society*, 99, 104862. doi.org/10.1016/j.scs.2023.104862.

Xia, H., Liu, Z., Efremochkina, M., Liu, X., Lin, C., 2022. Study on city digital twin technologies for sustainable smart city design: A review and bibliometric analysis of geographic information system and building information modeling integration. *Sustainable Cities and Society*, 84, 104009. doi.org/10.1016/j.scs.2022.104009.

Xu, H., Omitaomu, O., Sabri, S., Zlatanova, S., Li, X., Song, Y., 2024. Leveraging generative AI for urban digital twins: a scoping review on the autonomous generation of urban data, scenarios, designs, and 3D city models for smart city advancement. *Urban Informatics*, 3. doi.org/10.1007/s44212-024-00060-w.

Yan, J., Lu, Q., Li, N., Chen, L., Pitt, M., 2025. Common data environment for digital twins from building to city levels. *Automation in Construction*, 174, 106131. doi.org/10.1016/j.autcon.2025.106131.

Appendix

The surveyed UDT projects are publicly available in a GitHub repository(https://github.com/JEDDOUB/UDT_Projects_Data_Integration_Analysis-/blob/main/Appendix.pdf). Although not exhaustive, the repository will be continuously expanded as new UDT initiatives are documented.