Geospatial Research Innovations and Development (GRID)
UNSW Built Environment, L4 R.4036, Red Centre West Wing
Sydney, NSW 2052 Australia

RG202877

Liveable City Digital Twin

Analytics for agile decision making

Lessons Learned and Recommendations

DELIVERABLE 10

Jack Barton, Abdoulaye Diakité, Michael Rigby, Masoud Rahimi, Sisi Zlatanova

Contents

Liveable City Digital Twins....3
Initiatives for establishing Digital Twins in Australia....4
The Liveable City Digital Twin project....8
Lessons learned and Recommendations....8
1. Manual data cleaning and joining....8
2. (Spatial) Semantic enrichment....9
3. Data integration....9
4. Database storage....11
5. Sensors and real-time data....12
6. Maintenance and update of DT content....13
7. Visualisation, UI/UX....13
Industry stakeholder review....14
Conclusion....18
References....20
APPENDIX 1....21

List of Figures

Figure 1: What three words define liveability for you?....3
Figure 2: Digital Twin Functional Capabilities: Source. SCCANZ, 2020....3
Figure 3: ANZLIC Digital Twin Principles. Source: ANZLIC, 2019 cit.in SCCANZ, 2020....6
Figure 4: Data sharing stages (Griffith 2021)....6
Figure 5: Screenshot of the 'IoT devices register' table from the Liverpool City Council IoT and Open Data platform....12
Figure 5: Which is your primary sector?....15
Figure 7: What is your primary role?....15
Figure 8: Which organisation are you with?....15
Figure 9: Tech spectrum self-assessment....15
Figure 10: What data do you imagine would benefit a Liveable City Digital Twin?....15
Figure 11: Data integration is the biggest headache for Digital Twins?....15
Figure 12: Metadata alone is insufficient?....16
Figure 13: Standards are important for a Digital Twin?....16
Figure 14: Semantics are important for Digital Twin?....18

List of Tables

Table 1: FAIR Principles. See https://www.go-fair.org/fair-principles/ and http://www.nature.com/articles/sdata201618....5
Table 2: Initial input data used for the project....10
Table 3: List of 3DCityDB Tables (left) and final list of data (right)....10
Table 4: What three words define liveability for you?....17

Table 5:What data do you imagine would benefit a digital twin? 17

List of Acronyms

Term	Description
AURIN	Australian Urban Research Infrastructure Networks
BIM	Building Information Modelling
DEM	Digital Elevation Model
CESIUM/ TerriaJS	3D Geospatial visualization platform for the Web
3DCityDB	3D City Database The database schema implements the CityGML standard with semantically rich and multi-scale urban objects facilitating complex analysis tasks, far beyond visualization. 3DCityDB extends the capabilities of PostgreSQL/PostGIS
CityGML	CityGML is an open data model and XML-based data exchange format that can be used to describe urban and landscape objects
DBMS	DataBase Management System
GeoJSON	GeoJSON is an open standard format designed for representing simple geographical features, along with their non-spatial attributes
GRID	Geospatial Research Innovation and Development lab, based at UNSW
IFC	Industry Foundation Classes, a standard file format as used in BIM software
PMS	Pavement Management System
PostgreSQL/PostGIS	PostgreSQL is an open-source relational database. PostGIS is an extension enabling geographic operability on top of the standard database
UNSW	University of New South Wales
LCC	Liverpool City Council

This project is supported by FrontierSI, AURIN, UNSW Research Infrastructure and in collaboration with NSW Spatial Services, Liverpool City Council, Data61, QLD DNRME, Astrolabe and Spatial Vision.

Liveable City Digital Twins

This project is about liveable Digital Twins. But what exactly is 'liveable'? Australian Curriculum defines liveability as "An assessment of what a place is like to live in, using particular criteria, for example, environmental quality, crime and safety, education and health provision, access to shops and services, recreational facilities and cultural activities."[1] The term has been used widely to evaluate different characteristics (e.g., environment or climate) of a city. Several magazines such as the Economist Intelligence Units (EUI), Monocle and Global Finance or consulting firms such as Mercer publish world ranking for most liveable cities. All of them consider a range of parameters starting from *affordability*, *happiness*, *healthcare*, *hygiene* and ending with *clime rates* and *consumer prices*. Despite the varying criteria, certain cities such as Vienna, Zurich, Vancouver, Sydney, Melbourne, Auckland, Amsterdam, Copenhagen, Tokyo and Stockholm are generally in the top ten.

We asked our stakeholders to offer three words to define liveability. Figure 1 shows responses from the project's stakeholder workshop at the conclusion of the project. These responses align with the definition above. In our project, we used the term *liveable* quite loosely to indicate that we develop a virtual replica of the real world and supporting tools to monitor certain aspects of city life. We have concentrated on bringing together weather temperature, people's movements and building shadowing in an appropriate 3D virtual environment.

Figure 1: What three words define liveability for you? 2020

Function	Capability
Connect	There must be a 'live' connection between the digital replica and the physical world. This connection allows various disparate information and data from the physical world to come into a unified virtual environment.
Integrate	Intelligently checks and links relevant data from different sources (and across sectors) to effectively enable meaningful analysis to those who see the value.
Visualise	To display real time multisource data to the user. This allows access to the information users need, precisely when they need it, across the whole project and asset operation lifecycle.
Analyse	Federated data sets from various sources can be processed, modelled, analysed and simulated to bring business objectives to life.
Secure	Having a security minded management approach to data and information by applying relevant technical security and privacy standards.

Figure 2: Digital Twin Functional Capabilities: Source. SCCANZ,

The range of words combined with the key notion of liveable being *connected* provide indications that a digital framework should enhance decision making through connectivity, integration, visualisation and analysis of federated datasets to ensure secure mechanisms. These capabilities are the minimum requirements that the Australia/New Zealand Smart Cities Council (SCCANZ) uses to define a Digital Twin, as described in **Error! Reference source not found.**.

The discussion on Digital Twins is broad and spans geometric components to simulation and modelling capabilities that Digital Twins should support. The definitions also vary:

- ISO/TC 184 (1) Digital Twin manufacturing framework – Part 1 Overview and general principles starts: "*A Digital Twin is a digital model of a particular physical element or a process with data connections that enable convergence between the physical and virtual states at an appropriate rate of synchronisation".*

[1] See Australian Curriculum, Assessment and Reporting Authority https://www.australiancurriculum.edu.au/f-10-curriculum/humanities-and-social-sciences/hass/Glossary/?term=liveability last accessed 16th Oct 2022

- Draft technical report of IEC/TC 65 ISO/TC 84 JWG 21 on Smart Manufacturing Reference Models defines Digital Twin as a '*Digital representation of physical individuals as well as of virtual entities in an information framework that interconnects traditionally separated elements and provides an integrated view throughout life cycles (digital twins and digital thread)*."

The manufacturing world provides a quite clear and well-defined view on the digital aspect of the twin (Lehner et al, 2022). Generally, the geometric components needed for the twin are clearly outlined. The focus is then on providing the link to the real world and developing monitoring, estimation, simulation and prediction components. The Digital Twins for cities are much more comprehensive, interrelated and diverse. The complexity comes from two major factors:

1) Large amounts of heterogenous spatial data are already available and maintained by different data vendors with various characteristics. This includes similar datasets with different descriptions, accuracy, quality, access etc.
2) The scope of the Digital Twin is hard to identify. It might be related to a specific application (e.g., health, energy, flood, urban heat island) or the tasks of a specific group of people (e.g., emergency responders, urban planners), which can span large areas (e.g., precinct, council, state country), use a variety of sensor information that is typically extremely large (from monitoring weather to human movements), with potential dynamic components differing significantly (from hour changes like in flood or shadowing).

Therefore, the Digital Twins focussed on cities and urban regions tend to discuss issues related availability, accessibility and integration of spatial data. For example:

- Data61 specifies Digital Twin as '*A secure way of allowing industries, businesses, individuals and other areas of Government to see a full picture of accurate, up to date datasets in a user-friendly format'(link).*
- Victoria State Government: '*The Digital Twin Victoria (DTV) platform is the most comprehensive digital model ever assembled for Victoria. It brings together masses of 2D, 3D and live data in a single online place open for everyone to use'. (link).*

Clearly, to fully represent the built environment, digital twins need to accommodate various types of data, such as geospatial reference data, asset attributes (e.g., natural, physical, social, economic), management data, real-time asset performance, utilisation data, which further have to be processed (i.e., ETL) then enhanced with appropriate analytics, simulations and predictions. These simulation and predictions are a critical part of Digital Twins that would facilitate understanding real world phenomena and allow measures to be taken towards making the living environment more sustainable and endurable.

Initiatives for establishing Digital Twins in Australia

Each state in Australia is working on developing a Digital Twin platform, which has to support the establishment of Digital Twins for city and urban applications. While it is still too early to determine whether this will be realised, the probability is high that only a set of a key Digital twins may be included. As mentioned above, the first large group of challenges to be addressed is finding the spatial data needed and integrating this into one valid, fit-for-purpose Digital Model. The issue is multi-faceted covering technical administrative and legal aspects. Griffith and Truelove (2021) specify the key components of a Digital Twin ecosystem should include as:

- Agreeing on rules, protocols and standards to discover, share and access data, services and capability including agreed approaches for authentication of user identity and role, authorisation to access particular data (or levels of detail/granularity of data), and access conditions (such as access costs, licence and use restrictions).
- Ensuring data custodianship and authority remains with the contributing organisation so that custodians can maintain control over shared data, monitor access to and use of their data through role-based access, and maintain authority for their respective functions and data.
- Defining digital twin-compatible data (i.e., interoperable, compatible, cross-platform and platform-agnostic) that will allow for digital twins in different sectors and government jurisdictions to mature at different rates and levels of complexity.
- Seamless integration of government and non-government data to realise the benefits of combining industry, government, research and community sector data.
- Maintaining data over time so that it accurately reflects the current environment and changes in the environment, where responsibility for ensuring data is up-to-date, accurate, quality assured and compatible would reside with data custodians and/or digital twin operators.
- Customisable, user-driven access to data in a form that can leverage new technologies and adapt to user needs, such as connecting digital twins in a particular region or city for localised insights

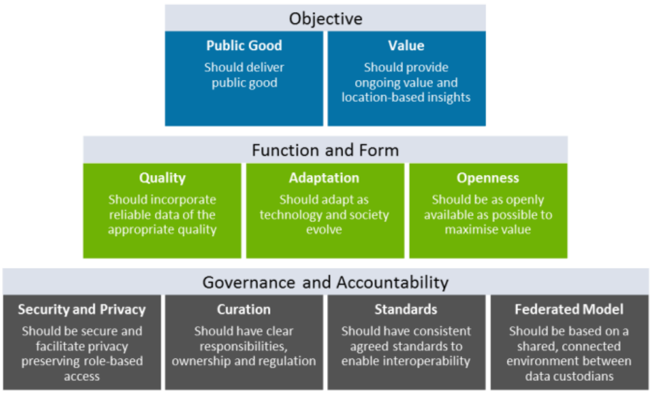
To support sharing and interoperability of data at a foundational level, DT data needs to be compliant with the *FAIR Guiding Principles for scientific data management and stewardship*, which seeks to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets, as described in Table 1:

Table 1: FAIR Principles. See https://www.go-fair.org/fair-principles/ and http://www.nature.com/articles/sdata201618

Findable The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services F1. (Meta)data are assigned a globally unique and persistent identifier F2. Data are described with rich metadata (defined by R1 below) F3. Metadata clearly and explicitly include the identifier of the data they describe F4. (Meta)data are registered or indexed in a searchable resource
Accessible Once the user finds the required data, she/he/they need to know how they can be accessed, possibly including authentication and authorisation. A1. (Meta)data are retrievable by their identifier using a standardised communications protocol A1.1 The protocol is open, free, and universally implementable A1.2 The protocol allows for an authentication and authorisation procedure, where necessary A2. Metadata are accessible, even when the data are no longer available
Interoperable The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing. I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. I2. (Meta)data use vocabularies that follow FAIR principles

I3. (Meta)data include qualified references to other (meta)data
Reusable The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings. R1. (Meta)data are richly described with a plurality of accurate and relevant attributes R1.1. (Meta)data are released with a clear and accessible data usage license R1.2. (Meta)data are associated with detailed provenance R1.3. (Meta)data meet domain-relevant community standards

The discussion on data can be framed in different ways. For example, the Australia/New Zealand Land Information Council (ANZLIC) has developed a hierarchical organisation of principles for building Digital Twins at three levels (Figure 3): governance, function and objective. It should be highlighted that while a Digital Twin is not yet another foundational data set in itself, it is comprised from them and should deliver public good and ongoing value. It is also emphasized that standards and federated models should guide organisation of data and responsibilities with ownership information clearly specified. The very important issue of open data is also included in the hierarchy; however it is clear to that not all data can be made open, e.g., human or commercial sensitivity.



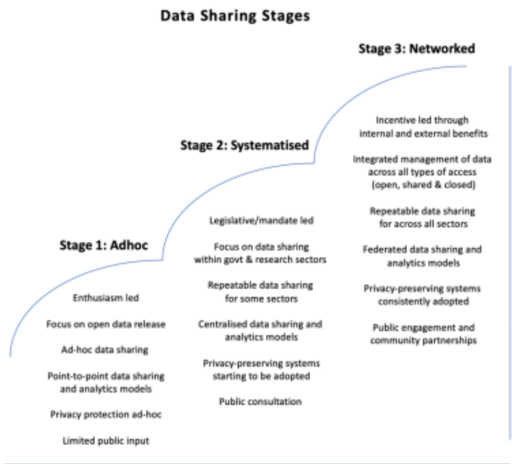


Figure 3: ANZLIC Digital Twin Principles. (Source: ANZLIC, 2019 cit.in SCCANZ, 2020); *Figure 4: Data sharing stages (Griffith 2021)*

This hierarchy denotes the importance of a solid foundation in supporting higher level outcomes. It is important to realise that preparing Digital Twins to mature over time, as discussed by Griffith and Truelove 2021 (Figure 4), requires a careful approach to design and development. This also indicates that a Digital Twin can have varying levels of maturity, as described in Figure 5, where each tier is interrelated to the others; an improvement in one area can induce an improvement in another.

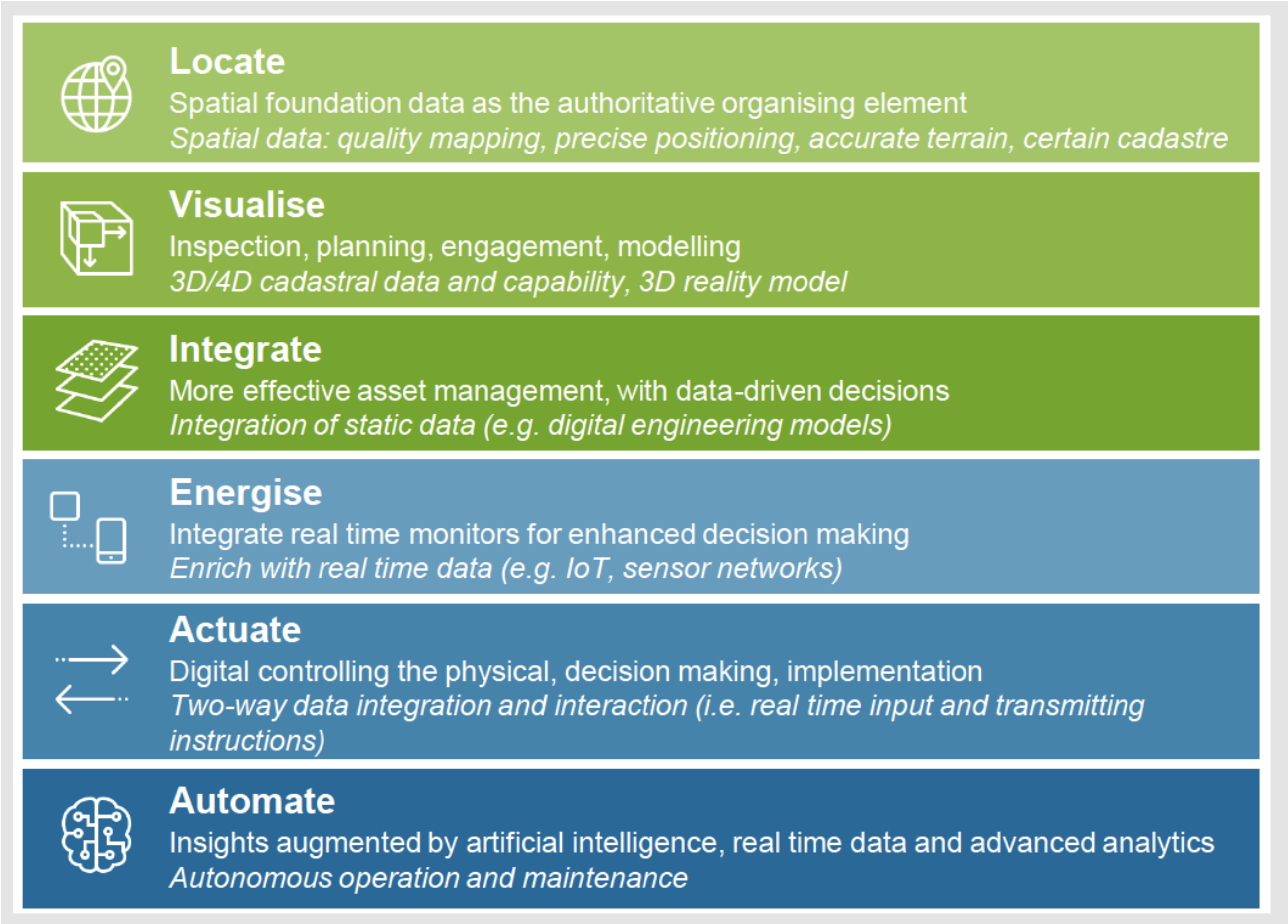


Figure 5: ANZLIC Spatially enabled digital twin maturity model (Source: ANZLIC, 2019 cit.in SCCANZ, 2020);

In this project we have developed a working model demonstrating how all these concepts can be applied to a fully functional example. During the course of this work, numerous challenges were faced and addressed either in an ad-hoc manner or using international frameworks, concepts and knowledge.

This report presents lessons learned during the project and provides recommendations with respect to data cleaning, semantic enrichment, data integration, database storage, sensors and real-time data, maintenance and update of DT content, visualisation and UI/UX. An Industry stakeholder review is also presented covering issues of data integration, management and standardisation.

The Liveable City Digital Twin project

This project focussed on challenges related to developing a full-stack Digital Twin (DT) using open-source components. It aimed to be standards compliant, database driven, and accessible online or via diverse software front ends.

The system was developed using contemporary data sources in the Australian context, providing as close to a real-time DT as practicable. The DT has the options to access federated data networks for foundational data and add value to these data by mapping to a semantically rich models (i.e., CityGML) that serves to integrate, harmonise and update data with a bidirectional user interface.

The Liverpool precinct in Sydney, NSW was used as a case-study with a focus on applied technology to assist in the management of urban heat islands, street shading and walkability. Phases of this project have included four parts: Building of Digital Twin, Use of Digital Twin, Update of Digital Twin and Industry Stakeholder Review. The sections below elaborate on challenges and solutions and providing recommendations for future stakeholders building Digital Twins (Infrastructure Australia, 2018).

Lessons learned and Recommendations

An overview of the project developments is available in Diakite et al. (2022).

1. Manual data cleaning and joining

Data are created and maintained for different purposes and vary in granularity (levels of detail), representation (line, polygon, polyhedron), attributes and quality. The obtained data must be inspected, and relevant features/attributes must be identified.

In this project

To maximise the potential for interoperability, one of the main goals was to use a database management system (DBMS) with a standardised data model, able to support geometric and semantic data. Hence the choice of 3DCityDB, adopted for our experiments as its database schema are designed to align with the OGC CityGML standard. It is available for implementation with PostgreSQL/PostGIS. We used 3DCityDB as a proof of concept to demonstrate that all project's components can be stored in a standardised format and attached to 3D city objects.

In-house software was developed to select attributes and link spatial datasets from different sources, convert data types and validate objects, with interfaces developed to access environmental data about air quality and other environmental information (heat stress index, ozone, brightness, etc.) from the council servers. The following major challenges were identified:

- Although 34 sensors (people and vehicle counting) were made available to the project, the sensor data did not provide enough context for monitoring temperature change (only two weather stations located at the center of the study area were available) or perform specific analysis (i.e., the people counting devices would only provide numbers without spatial context that would allow the overlay of trajectories with shadow polygons, for example).
- Initially many hurdles were encountered relating to ambiguous or missing data. For example, it was challenging to connect the *Device name, Descriptio*n, *Address* of a traffic sensor to a road using 'towards', 'facing' or 'Cnr' (corner) attribute descriptions, to the road network with elements split and stored per centreline. This required the road objects to be picked manually.

- Although it is understood some BIM data existed within private firms, it was not possible to acquire it within the scope of the project.

Recommendations:
R1.1: Standardised Automated tools need to be developed as a sharable layer on top of openly available API/data to promote replicable and scalable methods, in accordance with FAIR principles.
R1.2: Current levels of detail (both lines and polygons) should be provided for streets and footpaths. Datasets with topology need to be generated.

2. (Spatial) Semantic enrichment

Semantic information is a critical component of spatial data (Oosterom and Zlatanova, 2008). However, some input datasets may not provide semantic information, which may result in under-performing models. This highlights the many significant opportunities that standards, such as CityGML and IFC, and semantic enrichment can play in developing solutions.

In this project

- Semantic information was associated with many of the 2D and 3D datasets, but it was largely insufficient in the input data.
- The building dataset that we used corresponds to the CityGML LoD2. However, there was no distinction between walls and roof features in the input data. Therefore, the faces of the building had to be classified in 3 main semantic classes: RoofSurface, GroundSurface and WallSurface. An In-house routine was developed to distinguish between the faces: vertical faces are considered as walls; the rest is considered as roof if they lie above a height threshold and ground otherwise. This gives us a richer dataset where component such as roofs can be specifically queried, which could be useful for applications such as solar potential estimation on roofs.

Recommendations:
R2.1: Foundational spatial datasets should be semantically rich and delivered via conventional spatial data formats, such as Web services, in GML, GeoJSON or SHAPE file format and following FAIR principle. Interfaces for CityGML and IFC data should be developed as well.
R2.2: Standard-oriented DBMS data models such as 3DCityDB should be leveraged more to facilitate interoperability.
R2.3: Semantic enrichment should be automated in well-defined and standard-based workflow and fed back into the compatible DBMS (e.g. 3DCityDB).
R2.4: Open-source tools must be developed for most commonly used spatial schemas (such as CityGML and IFC).

3. Data integration

Data integration has been long-discussed topic, specifically when considering 3D data (Zlatanova and Prosperi, 2006). Currently, national coverage open- and vendor-based datasets typically require an amount of preparation to make them fit-for-purpose to the specific application.

Several spatial datasets have been obtained for the test area of City of Liverpool, which correspond to four classes of CityGML: Building, Transportation, Relief and Waterbody (see Table 2). As these all have different sources, data structures, objects, and descriptions, a critical task for their integration and storage was mapping classes and attributes to CityGML and importing them into 3DCityDB. Through this course of work, a number of new data sets were discovered and added.

Building	**Relief**	**Transportation**	**WaterBody**
Buildings3D (Spatial Services)	**DEM (grid)** (Spatial Services)	**Railway3D** **RoadSegment3D** (Spatial Services) **PMS (2D)** **Road Cadastre (2D)** (Liverpool City Council & AURIN)	**Hydrolines3D** (Spatial Services) **Polygons (2D)** (Geoscience Australia)

Table 2: Initial input data used for the project.

The final set of 3DCityDB tables and data sets is listed in Table 3:

3DcityDB table	Data set
BUILDING	Buildings
CITYOBJECT	All data
CITYOBJECT_GENERICATTRIB	All data
GENERIC_CITYOBJECT	IoT Sensors
SURFACE_GEOMETRY	Buildings
THEMATIC_SURFACE	Buildings
TRANSPORTATION_COMPLEX	Railways Roads
WATERBODY	Water bodies

Data set	Geometry type	Format
Buildings	MultiPolygonZ	GDB
Roads	MultiLineStringZ MultiPolygon	GDB & SHP
Railways	MultiLineStringZ	GDB
Water bodies	MultiLineStringZ MultiPolygon	GDB & SHP
Vegetation	MultiPolygon PointZ	GDB & SHP
Terrain (DEM)	Raster (grid)	GeoTIFF
IoT Sensors	Point	GeoJSON

Table 3: List of 3DCityDB Tables (left) and final list of data (right)

To optimise these data for ingestion, the following pre-processing were performed:

- **Buildings**: semantic information was reconstructed; however, the data was already a polygonal mesh that required not further processing for integration into a TIN.
- **Road polygons**: Polygonal data was preferred (over centrelines) to label polygons according to their corresponding feature class and use the ground polygons for pedestrian related analysis. These polygons were simplified, resulting in a 98.4% reduction of points. Overlapping/intersecting parts were removed using the 'Difference' Boolean operation in QGIS and then the points were converted to 3D by DEM-based sampling to recover their elevation. Using an in-house python script for QGIS.
- **Railway polygons**: Delivered as 2D line geometry, it was not necessary to seek a polygonal version as width is consistent. Pre-processing included buffering, point simplification and elevation sampling.
- **WaterBody**: Hydroline data are only represented using LineStrings, which is not optimal for our 3D model. A hydrology polygon dataset was preferred, although it did not include temporary waterbodies like creeks and drains.
- **Terrain/Relief**: A TIN was generated directly from the raw DEM incorporating the footprints of the buildings. As this TIN contained ~ 5.8 million triangles it is not efficient to work with and most applications. We generated a simpler TIN minimising the number of polygons while including all the features of importance (buildings, roads, railways, and waterbodies) using the 3D constrained Delaunay triangulation implementation of PostGIS. We used 3D boundary lines of the features as constraints for the triangulation. The resulting TIN count of 28K

polygons makes it suitable for an efficient use in spatial analysis, while preserving good elevation information for all features. 20

- **IoT Sensor data: H**eight values of the sensors were missing in the datasets, a value of zero is temporarily used as the Z dimension.

Approaches for 3D integration are discussed in Yan et al (2019) and Li et al (2020).

In this project:

- The semantics of different data sets needed to be harmonised to align with CityGML nomenclature,
- Obtained roads datasets were represented by centrelines, but polygons were needed,
- Ground-level and heights needed to be established for subsequent integration of building footprints, roads, paths and waterbodies, and
- Vector terrain data needed to be generated from DEM and simplified

Often this process is conducted in an ad-hoc fashion, however, within the urban modelling context, there is an opportunity for developing standardised automation routines.

Recommendations:
R3.1: To avoid file-based systems and optimise potential for automated workflows, well-structured, standard compliant data sources should be offered with option to connect via API.
R3.2: Data should be published in accordance with the FAIR Principles (See Table 1)
R3.3: Metadata should comply with AS/NZS ISO 9115.1:2015 so users may assess appropriateness/reliability for purpose
R3.4: DT project functional requirements and data agreements must be explicit, defining a motivational problem to solve.

4. Database storage

DataBase Management Systems (DBMS) provide a series of benefits over ad-hoc file-based systems: they facilitate improved sharing, security/privacy, integration, timeliness, consistently, reliability, productivity, and ultimately decision-making and value add at higher levels in a DT workflow.

In this project:
3D City Database (3DCityDB) was chosen to store all project components attached to 3D city objects in a standardised spatial schema. This database schema implements CityGML, which maintains semantically rich, multi-scale urban objects. This project used open source DBMS PostgreSQL/PostGIS, which has grown to offer powerful query and analysis capacity. 3DCityDB is under 3rd party development and in the proof-of-concept phase. The following lessons were learned:

- 3DCityDB importer/exporter is only for CityGML files
- A critical task for integration and storage of data is mapping the classes and attributes of collected data to CityGML and importing in 3DCityDB.
- Several datasets couldn't be mapped to CityGML classes.
- Inclusion of a 3D DBMS greatly helped ease maintainability and scalability by facilitating standardisation and data structure.

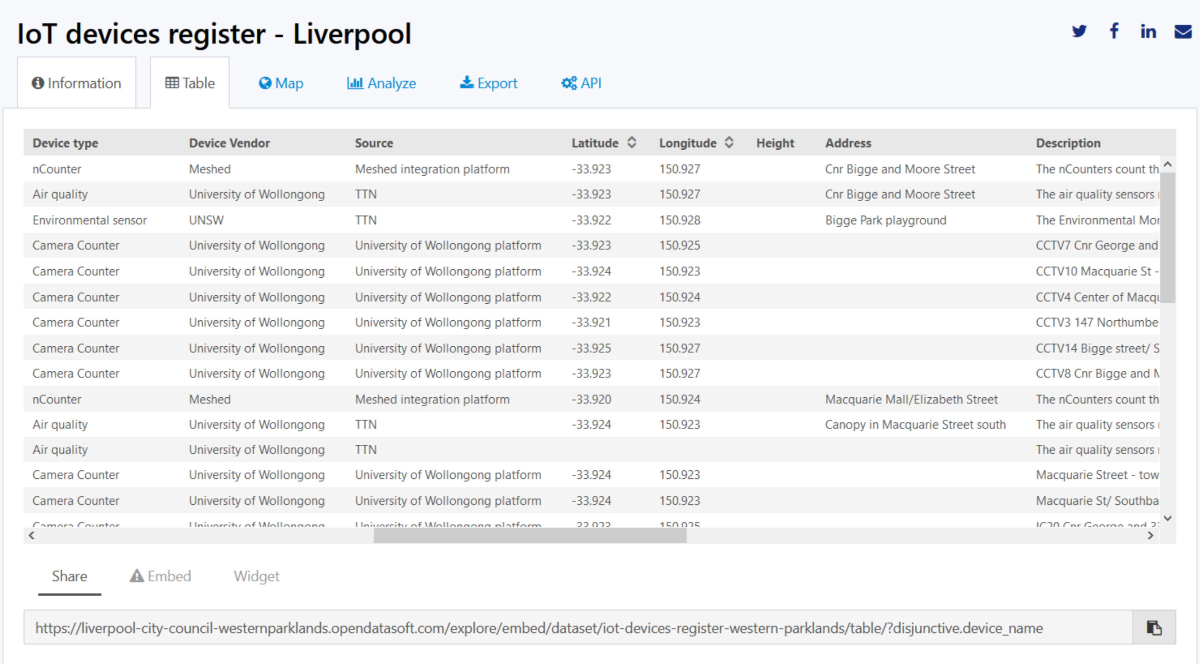
In federated datasets, each respective custodian is charged with keeping their repositories valid and up to date; a remote user cannot (should not) be able to edit this data. Data consumers may ingest a localised feed of federated data and enhance this with a layer of semantic information particular to the DT implementation. If this process is standardised, this enriched data may then be accessed by downstream users or fed-back into the original source.

Recommendations:
R4.1: Data agreements can help API data to be value-added and re-distributed. Licence permitting, data repositories can provide APIs to find and access/extract data. Upon specific authorisation data repositories should facilitate updates.
R4.2: Linkable schemas should be maintained to permit semantic-web operability.
R4.3: Solutions should reduce risks from proprietary lock-in, walled gardens and non-standard formats reducing interoperability, with APIs provided to convert internal data schemas to standard schemas (national or international).
R4.4 Adopted database spatial schemas should offer a mechanism for extension.

5. Sensors and real-time data

A primary functionality of a Digital Twin is the capacity to harness *live* data feeds. These might update at different increments, and some sensors can even provide up-to-the-minute data that is approaching a real-time information flow.

In this project



Device type	Device Vendor	Source	Latitude	Longitude	Height	Address	Description
nCounter	Meshed	Meshed integration platform	-33.923	150.927		Cnr Bigge and Moore Street	The nCounters count th
Air quality	University of Wollongong	TTN	-33.923	150.927		Cnr Bigge and Moore Street	The air quality sensors
Environmental sensor	UNSW	TTN	-33.922	150.928		Bigge Park playground	The Environmental Mo
Camera Counter	University of Wollongong	University of Wollongong platform	-33.923	150.925			CCTV7 Cnr George and
Camera Counter	University of Wollongong	University of Wollongong platform	-33.924	150.923			CCTV10 Macquarie St -
Camera Counter	University of Wollongong	University of Wollongong platform	-33.922	150.924			CCTV4 Center of Macq
Camera Counter	University of Wollongong	University of Wollongong platform	-33.921	150.923			CCTV3 147 Northumbe
Camera Counter	University of Wollongong	University of Wollongong platform	-33.925	150.927			CCTV14 Bigge street/ S
Camera Counter	University of Wollongong	University of Wollongong platform	-33.923	150.927			CCTV8 Cnr Bigge and
nCounter	Meshed	Meshed integration platform	-33.920	150.924		Macquarie Mall/Elizabeth Street	The nCounters count th
Air quality	University of Wollongong	TTN	-33.924	150.923		Canopy in Macquarie Street south	The air quality sensors
Air quality	University of Wollongong	TTN					The air quality sensors
Camera Counter	University of Wollongong	University of Wollongong platform	-33.924	150.923			Macquarie Street - tow
Camera Counter	University of Wollongong	University of Wollongong platform	-33.924	150.923			Macquarie St/ Southba

Figure 6: Screenshot of the IoT devices register *table from the Liverpool City Council IoT and Open Data platform.*[2]

Figure 6 shows an exhaustive list of all the IoT sensors installed on the test site, along with their relevant metadata (e.g., id, description, device name, location). These data were integrated in the 3D model, enabling management of explicit information about sensor devices, their properties and spatial location. This ensured direct and efficient spatial analysis with the sensed data. Every sensor device was registered in the CITYOBJECT table and is provided with a unique *cityobject_id*. The same cityobject_id is used as *id* in the GENERIC_CITYOBJECT table, recorded using the longitude and latitude values of the sensor. The experienced challenges were as follows:

[2] See: https://data.liverpool.nsw.gov.au/explore/dataset/iot-devices-register-western-parklands/table/?disjunctive.device_name&disjunctive.device_type Last accessed 18/11/22

- Height values of the sensors are missing in the datasets, an altitude of 0 is temporarily used as the Z coordinate of the points.
- There was no class that provides direct compatibility with IoT sensors as features in the current version of CityGML (2.0), although this is expected to change in the coming version CityGML 3.0.
- Sensor data records does not need to be explicitly stored in the DT database, unless necessary for archiving and postprocessing, as it might be resource consuming.

Recommendations:
R5.1: Where possible, sensor data should have accurate 3D location.
R5.2 Sensor data must be standardised and anonymised at the point of creation.
R5.3 Standards such as SensorML, OGC Measurements and Observations should be accorded to.

6. Maintenance and update of DT content

A key benefit of a spatial digital twin is the ability to share and re-integrate data with a network of other geographically spread clients. A standardised foundation is critical, along with a dynamic system for harmonisation and value-adding of information as successive participants interoperate with the system. An important feature of our DT is to optimise the potential for information re-usability by complying with the FAIR principles. Further, beyond a simple one-way data flow, a GUI has been designed to permit a user to directly edit/add information via established connection between the UI and the DBMS. This allows live interaction with the model, ensuring real-time information update, assisting in large scale data curation, quality assurance, validity, update and maintainability of the data foundation.

In this project

- We concentrated only on attribute update, i.e., updating information without changing the spatial schema.
- Geometric updates were not considered as they require a more elaborated UI and validity check functionality.

Recommendations:
R6.1: Data agreements should accommodate bi-directional exchange of spatial schema, geometry and attributes.
R6.2: APIs for update with security, locking and validity functions. Options for server-based and client-based validation should be explored.
R6.3: Secure authorisation systems need to be in place to facilitate both open access and privilege-based access.

7. Visualisation, UI/UX

A visualisation platform is a critical component of Digital Twins. A focus of this project was to deliver an accessible frontend interface for users to perform intuitive visual analysis. As shadow casting plays a key role in urban cooling (important for liveability), generated 3D Shadow Polygons were made available for analysis and cross-platform shareability, as well as to be able to precisely determine the amount of shadow on 3D features integrated into the terrain.

Depending on the application, a desktop or web-based platform can be established with a range of interaction and immersion (AR/VR) possibilities, utilising GIS (ArcGIS, QGIS), CAD (SketchUp, Autodesk), BIM (Revit, Bentley Systems), game engines (Unity, Unreal, NVIDIA) or global virtual environment (Cesium, Google Earth).

In this project

- QGIS and Cesium as open-source enablers were used as interfaces for visualisation of 3D data.
- *Cesium ion* can ingest a wide variety of formats, but the process of generating optimised 3D tiles through it implies a file transfer process hindering updates of models.
- Cesium provides tools to develop user interfaces, but updating geometric data would require more elaborated platforms.
- Cesium has limited support for WFS and currently no support for direct access to DBMS.
- GeoJSON was used to connect to Postgres/PostGIS, however this came at a significant overhead in terms of loading time and usability.
- QGIS has direct compatibility with Postgres/PostGIS data and has a wide range of powerful tools allowing manipulation and analysis. However, 3D visualisation is limited.
- New systems and platforms are progressively being released, which offer more extensive visualisation and interaction options.

Recommendations:
R7.1: Consider the development of modular components to handle several front-ends, in line with OGC W3C web services.
R7.2: Consider game engines such as Unity3D, Nvidia Omniverse, which offer extended 3D data modelling and visualisation capabilities.

Industry stakeholder review

An industry stakeholder workshop was run in the final phase of the project with the aim of demonstrating the development work to users with a professional interest and expertise in Digital Twins, and to solicit feedback and opinion with regard to deployment in urban environments.
Participants were presented with an overview of the working Liveable City Digital Twin and via an online survey (Mentimeter.com) prompted with a series of questions. The questions were designed to glean an understanding of what area the participants worked in, how they self-identified in terms of technological aptitude, what they defined as 'liveable' and what resources would be of benefit to a Liveable City Digital Twin. A number of questions prompted feedback on data integration, standards, metadata and semantics (See Appendix 1)

The total number of respondents (19) were from Local Government (6, 32%), Industry (4, 21%), State Government (4, 21%), Academia (3, 16%) and other (2, 10%). There were no participants directly employed by the Federal Government.

Figure 7: Which is your primary sector?

Figure 8: What is your primary role?

Figure 9: Tech spectrum self-assessment

Figure 10: Which organisation are you with?

Generally, respondents were from organisations with a direct interest in digital twins, with expert-level roles, as illustrated in Figure 8 and Figure 9*: Tech spectrum self-assessment* *Figure 10*. The question illustrated in Figure 1 gauged the self-assessed technical capacity of the participants (n=21) by prompting input along a spectrum from 0-5, with the average result being 4.1.

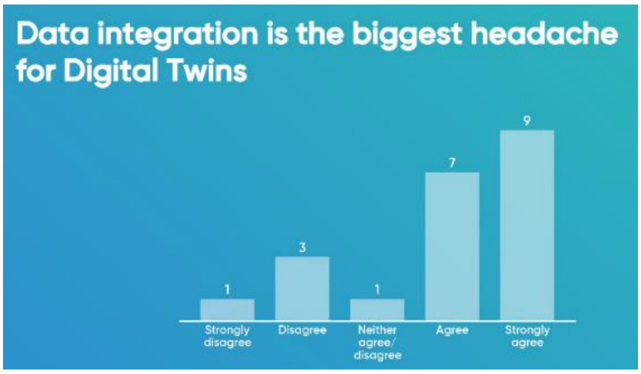


Figure 11: What data do you imagine would benefit a Liveable City Digital Twin?

Figure 12: Data integration is the biggest headache for Digital Twins?

When asked about data needed for Liveable City DT, responses highlighted an interest in diverse sources (*Figure 11*) with a general focus on transport, environmental and population data. Responders were inclined to link liveability to mobility as many aspects in a dynamic city environment are related to transportation ('20 min city', pollution, Urban heat Island). Despite the fact that many 3D models are perceived a DT, no indication was given that 3D data sets are needed. Live, real-time data was identified to be of central benefit to a DT. As a core functionality and minimum requirement, a DT must deal with real time, up-to-date data. **Error! Reference source not found.** and **Error! Reference source not found.** present the raw data from the Menti survey and aggregates responses into key themes.

It was generally agreed that data integration was the 'biggest headache' for digital twins, as illustrated in Figure 12. The one respondent who strongly disagreed cited data licencing issues as more problematic than data integration. Data integration has many aspects, not just technical integration but agreements on access and use of data need to be negotiated and enacted. This is often a primary step in a project and is prone to delays affecting the critical path of a development.

Responders also agreed that metadata alone was insufficient for digital twins, with two respondents disagreeing and two neutral as shown in *Figure 13*. It should be noted metadata standards are quite elaborated and consist of: ISO 19115-1:2014 Geographic information – Metadata – Part 1: Fundamentals ISO 19115-2:2019 Geographic information - Metadata – Part 2: Extension for acquisition and processing and ISO/TS 19115-3:2016 Geographic information – Metadata – Part 3: XML Schema implementation for fundamental concepts (currently under development).

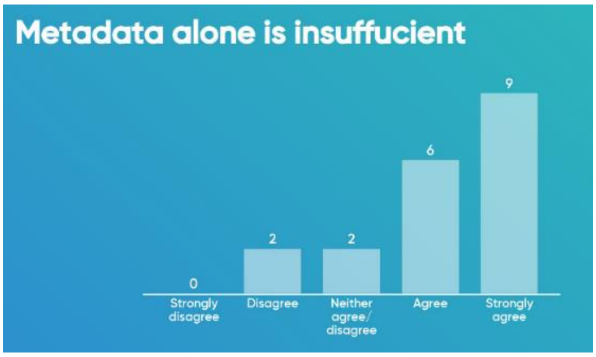


Figure 13: Metadata alone is insufficient?

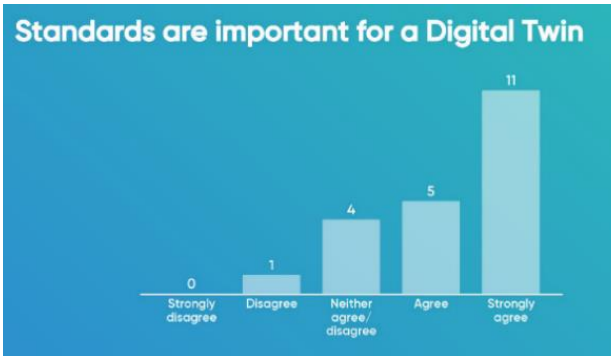


Figure 14: Standards are important for a Digital Twin?

Table 4: What three words define liveability for you?

Connectivity		**10**
	connected	6
	connected systems	1
	connectivity	2
	Integrated	1
Safety		**10**
	safe	3
	Safety	6
	secure	1
Quality of life		**7**
	Comfortable	1
	Fancy	1
	pleasant	1
	positive	1
	Quality of life	1
	vibrancy	1
	vibrant	1
Amenity		**6**
	amenity	1
	no floods	1
	open spaces	1
	services	1
	Shadow	1
	Sunlight	1
Greenery		**5**
	Green	3
	green spaces	1
	Greenspaces	1
Healthy		**4**
	Health	1
	healthy	2
	Heath	1
Sustainability		**4**
	Sustainability	2
	sustainable	2
Transport		**4**
	ease of transport options	1
	Efficient transport	1
	transport	1
	walkability	1
Community		**4**
	Community	2
	Community engagement	2
Accessibility		**3**
	Access	1
	Accessibility	1
	Accessibility	1
Data		**2**
	Any data	2
Other		**2**
	Aware	1
	diverse	1
Affordability		**1**
	Affordable	1

Table 5:What data do you imagine would benefit a digital twin?

Data		**13**
	CCTV data	1
	data from insurance	1
	elevation data	1
	future simulation	1
	hazard data	1
	live data	5
	Real-time	1
	sensors	1
	statistical	1
Amenity		**8**
	accessible	1
	aesthetics	1
	diverse application	1
	mobility	1
	services	1
	services and shops	1
	streetscape features	1
	walkability	1
Environment		**7**
	Climate	1
	Climate emergency	1
	contamination	1
	Flood	1
	Flood data	1
	soil type	1
	surface type	1
Transport		**7**
	Transport	2
	Transport data	2
	Transportation	3
Population		**6**
	CALD	1
	Health	1
	people	1
	population	1
	Population density	1
	sociodemographics	1
Pedestrian		**3**
	crowds	1
	pedestrian and traffic	1
	route scoring	1
Space/land use		**2**
	space-usage	1
	Zoning data	1
VR		**1**
	Immersive	1

A dedicated standard ISO19157 on data quality is also in development with two parts 1 on general requirements and part 3 on data quality measures register. In principle, these are both on 'data quality metadata' as the result of quality evaluation (as proposed in these standards) are typically reported as ISO 19115 compliant metadata.

Having well-described metadata data assists in achieving FAIR compliance. A major advantage of having this structure is the down-stream value add. For instance, an API connection could permit extraction of only a subset of the classes and attributes representing only the data needed for a DT. Being able to do this increases efficiency by having ready-to-go data, less transfer volume and computation and the potential for more accessible (and cheaper) data availability.

Figure 14 shows it was generally strongly agreed (11/21) that standards were important for digital twins, with only one respondent disagreeing and four neutral. Despite the overall support on standards, it was not discussed which spatial standards might of interest for DT. The general impression in the course of the project is that limited data sets are provided in accordance with standards. The notion of a standard is appreciated but the use of specific technical standards for exchange of data sets is very limited.

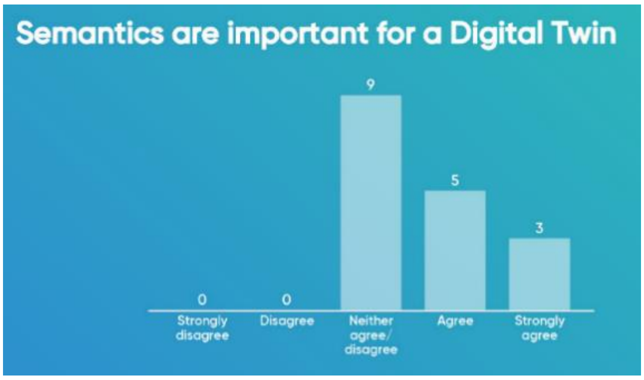


Figure 15: Semantics are important for Digital Twin?

The importance of semantics was lightly recognised. Figure 15 shows a slight majority of respondents (9/17) were neutral considering the importance of semantics with regard to digital twins, however just under half (8/17) either agreed or strongly agreed as to the importance of semantic content. No respondents disagreed. This might highlight the need to illustrate the value of semantics and related services, such as vocabularies, as they contribute to interoperability and increasing FAIRness.

Conclusion

Spatial data and related data services are foundational components of a DT. Integrated spatial data enables the creation of new datasets and insightful information that helps us work towards better liveability in our increasingly complex environments. Interoperability is critical in this process. Improvements in data sharing, open data release policies, data integration and standardisation initiatives like the FAIR data principles remove barriers for both data custodians, application developers and data users. Mobilising these data has important flow-on effects for establishing a common data strategy and shared foundation of services.

Certainly, issues exist regarding data completeness and availability, and as such, a DT should be able to handle incomplete data and ingest higher quality data as they become available. Spatial DBMSs are powerful and robust, and this project has demonstrated the capacity to structure, query and allow update of attributes from an accessible web interface, directly interacting with the DBMS. This presents powerful bi-directional functionality, creating dynamic information with real-time update and semantic enrichment for vocabulary, geometry, topology, and attributes. It should be noted that data for DT do not need to be centrally stored, but a federated standardised spatial schema should be utilised though the whole data.

Being deployed on the web, access privileges are able to be implemented, ensuring trust in the use of private and confidential data with processes and protocols to ensure it is protected and only shared with authorised users. Again, this is only possible to be achieved if well-defined data schemas are available.

During the project, it was confirmed that acquiring some type of data (e.g., BIM) can be difficult, harmonising different schemas can be time consuming and different sources / schemas / semantic and metadata repositories often provide partial information. Integration of data to create a valid model, that is, one without interactions, gaps, disconnected features, flying above or sinking 3D buildings, is a critical and complex process. This may require a range of tasks, such as the development of tools and APIs, to address these aspects.

We also revealed some specific system architecture constraints, for example, direct query to 3DCityDB from Cesium is slow (due to the necessary GeoJSON conversion and the slow ingestion of such format from CesiumJS), QGIS 3D visualization is limited and the 3DCityDB importer/exporter only deals with CityGML files (or KML/Collada). Commercial ETL (extract, transform, load) tools that can pull data form different sources into a data lake or cloud object storage have to be further explored. 3DCityDB (or other spatial schemas) can be part of the data lakes to be visualised in virtual environments, such as Cesium, game engines as Unity and Omniverse, or DT platforms, such as Microsoft Azure or Eclipse ecosystem. These ‘connectivity’ tools will be inseparable from future DT implementations by providing shared infrastructure services for different applications. It stands to note that capabilities in less mature and resourced organisations should be taken into account and scalable solutions prioritised.

A common understanding about DT is essential to help all stakeholders realise shared benefits. New governance arrangements, sustainable resourcing, shared infrastructure and improved data services are critical for supporting digital twins (Griffith and Truelove, 2021). National frameworks, such as AURIN, FrontierSI, SSSI, SCCANZ, ANZLIC, Geoscience Australia and the like, provide valuable resources to bridge between emerging DT initiatives, and with good practice in place, ensure many opportunities not otherwise available with ad-hoc methods.

References

ANZLIC, 2019, Principles for Spatially Enabled Digital Twins of the Built and Natural Environment in Australia. URL: https://www.anzlic.gov.au/resources/principles-spatially-enabled-digital-twins-built-and-natural-environment-australia

Bolton, A., Butler, L., Dabson, I., Enzer, M., Evans, M., Fenemore, T., Harradence, F., et al., 2018. The Gemini Principles. Report of Cambridge: Centre for Digital Built Britain. URL: http://www.cdbb.cam.ac.uk/DFTG/GeminiPrinciples.

Diakite, A.A, L. Ng, J. Barton, M. Rigby, K. Williams, S. Barr, and S. Zlatanova, 2022, Liveable City Digital Twin: a pilot project for the city of Liverpool (NSW, Australia), ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-4/W2-2022, 45–52

Griffith, C., Truelove, M., 2021 Framework for Spatially Enabled Digital Twins: Information Paper, CSIRO Data61, Australia. URL: https://www.anzlic.gov.au/sites/default/files/files/Data61%20Digital%20Twin%20Framework%20Information%20Paper%20-%20February%202021.pdf

Infrastructure Australia, 2018, Planning Liveable Cities: A place-based approach to sequencing infrastructure and growth. Available at https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLIII-B4-2022/295/2022/

Lehner, D., J. Pfeiffer, E-F. Tinsel, M.M. Strljic, S. Sint, M. Vierhauser, A. Wortmann and M. Wimmer, 2022, Digital Twin Platforms: Requirements, Capabilities, and Future Prospects, IEEE Software, Vol 39 (2), pp 53-61.

Li, W., S. Zlatanova, A. A. Diakite, M. Aleksandrov and J. Yan, 2020, Towards Integrating Heterogeneous Data: A Spatial DBMS Solution from a CRC-LCL Project in Australia, ISPRS Int. J. Geo-Inf. 2020, 9(2), 63

Oosterom, P. van and S. Zlatanova, 2008, Creating Spatial Information Infrastructures: Towards Spatial Semantic Web, ISBN 978-1-4200-7068-2 (hardback), CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 185 p.

SCCANZ, 2020. Data Leadership Guidance Note – Digital Twin. Centre for Data Leadership. URL: https://www.smartcitiescouncil.com/sites/default/files/anz_smartcitiescouncil_com/public_resources/guidance_note_digital_twin_issue_0.pdf

Wilkinson. M.D et al, 2016, The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data, 15March 2016, 9 pages.

Yan, J., S. Zlatanova, M. Aleksandrov, A. Diakite. C. Pettit, 2019, Integration of 3D Objects and Terrain for 3D Modelling Supporting the Digital Twin. ISPRS Annals of Photogrammetry, Remote Sens and Spatial Inf. Sci. 2019 IV-4/W8, 147–154

Zlatanova, S. and D. Prosperi, 2006, Large-scale 3D Data Integration: Challenges and Opportunities, 2006, ISBN 0-8493-9898-3, CRC Press, Tailor & Francis Group, Boca Raton, FL, 245 p.

APPENDIX 1

Menti interview questions:

1. Which is your primary sector?
2. Which organisation are you with?
3. What is your primary role?
4. Where are you on the Tech spectrum?
5. What three words define liveability for you?
6. What data do you imagine would benefit a Liveable City Digital Twin?
7. Data integration is the biggest headache for Digital Twins
8. Standards are important for a Digital Twin
9. Metadata alone is insufficient
10. Semantics are important for a Digital Twin