MANAGING URBAN DIGITAL TWINS WITH AN EXTENDED CATALOG SERVICE

M. Knezevic*[1], A. Donaubauer [1], M. Moshrefzadeh[1,2], T. H. Kolbe [1]

[1] Chair of Geoinformatics, Technical University of Munich, Arcisstr. 21, Munich, Germany
(marija.knezevic, andreas.donaubauer, thomas.kolbe)@tum.de
[2] City of Munich, Local Government Department, GeodatenService, Denisstr. 2, Munich, Germany
(mandana.moshrefzadeh@muenchen.de)

Commission IV, WG IV/9

KEY WORDS: Smart City, Urban Digital Twin (UDT), Urban Digital Twin Management, data infrastructure, Catalog Service.

ABSTRACT: A city is a complex system and involves many stakeholders with different interests and information about the physical city objects. Information integration from distributed sources is therefore a common challenge in many Smart City projects, as well as in projects for setting up Urban Digital Twins (UDT). In the context of distributed data infrastructures, catalog systems provide a significant contribution to solving the data integration challenge. However, neither existing catalogs for managing urban data nor catalogs used in Spatial Data Infrastructures (SDI) are tailored for managing UDTs. In response to this gap, our research focuses on the development of an extended catalog service including a metadata model specifically designed to manage distributed UDTs. The requirements for the metadata model and the functionality of the catalog were obtained from several smart city projects. As a proof of concept, the paper shows how the catalog implemented on the basis of the CKAN software package is used for managing the UDT of an urban redevelopment project in the city of Munich, Germany.

1. INTRODUCTION

The term Digital Twin (DT) is used in many domains and for numerous applications (Füller et al. 2020, Kritzinger et al. 2018). In the context of product life-cycle management in the early 2000's the term Digital Twin was first coined (Jones et al. 2020). The DT concept involves a physical entity of interest being represented by a digital counterpart, the DT, which is maintained throughout the lifecycle of the real entity. A DT is connected to its real, physical counterpart via sensors, and, depending on the type of DT, also actuators. The sensors not only observe the state of the physical entity of interest, e.g. the current rotation speed of an engine, but also the state of the physical environment in which the entity is located, e.g. the current weather conditions (Jones et al., 2020). Several architectures have been developed for building DTs in different application domains. In addition to the physical entity, the virtual entity and the interconnection between them (IoT sensors and actuators, data flows), simulation models and algorithms are generally recognized as key components of a DT (Kritzinger et al., 2018).

A defined methodology and terms, for establishing a DT, are also applicable to the Smart City domain (Füller et al. 2020, Shahat et al. 2021). It is necessary to organize and manage it to provide the exchange and management of information between the participants who are involved in the Smart City. All types of DTs aim to improve the performance of physical entities through the use of information technology (Jones et al., 2020). Cities need a comprehensive framework for organizing the various objects that require registration. The Urban Digital Twin (UDT) is a collection of digital resources of the city. It digitally represents all aspects of the real world required for a specific purpose and makes them accessible, analyzable, and visualizable for applications and users. A complex object (real or digital) consists of components that are made up of other components. It means that every object can be scaled down to the smallest possible detail. During the creation of an UDT, it is important that all needed registered objects have their unique identity, thematic properties, and metadata that uniquely characterizes them. It should be considered that objects have static as well as dynamic components. According to Jones et al. (2020) data ownership is a challenge in the DT context and integration between virtual entities is seen as a non-trivial future field of research. While Jones et al. (2020) refer to DTs in general, the challenge of data ownership is particularly compelling for UDTs where data is distributed across a wide range of stakeholders such as different departments of the municipality and enterprises to individual citizens. The approach presented in this paper shows how a catalog system contributes to managing an UDT under these complex conditions.

2. RELATED WORK

2.1 Digital Twins in the context of Smart Cities - Urban Digital Twins

Batty (2018) examines UDTs from the point of view of urban planning and doubts whether DT in the sense of an exact mirror image of a city will ever be feasible. However, he considers the basic idea of DT to be highly relevant from the perspective of urban planning. As an example of the first step in this direction, he describes linking a semantic 3D city model with real-time IoT sensors. Detailed semantic 3D city models enriched with thematic and real-time information are the backbone of Urban Digital Twins, which have already been implemented in cities like Helsinki (Ruohomäki et al., 2018), Rotterdam (Butot et al., 2020), Zurich (Schrotter & Hürzeler, 2020), Singapore (Niculescu & Wadhwa, 2015) and Atlanta with a special focus on human-infrastructure-technology interactions (Mohammadi & Taylor, 2017).

In addition, the Architecture, Engineering and Construction (AEC) domain is also addressing the issue of DTs for physical objects in the city. The main focus here is on the highly detailed semantic, three-dimensional representation of individual buildings and the maintenance of a DT over the entire life cycle of the buildings. Deng et al. (2021) describe the evolution from Building Information Modeling (BIM) to DT. They define a taxonomy consisting of five levels, ranging from conventional BIM to the coupling of BIM and simulation, BIM and IoT, BIM and artificial intelligence (AI), and the ideal of the DT. While earlier BIM concepts assumed a common integrated digital model for all stakeholders involved, current concepts deal with the establishment of so-called Common Data Environments (CDE), which are supposed to make it possible to keep several digital representations of the physical or planned object in

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W3-2022-119-2022 |

parallel and to link the heterogeneous data of the different representations (Preidel et al., 2016).
Shahat et al. 2021 state that using a DT in smart cities will improve its sustainability. In their work they present the potential of the Digital Twin city. The issues and differences between 3D models and DTs of the Smart City are discussed and a DT is seen as an opportunity to improve city design and operability. The large-scale model of a city proves to be a serious issue, which brings challenges in data handling. Based on the DT literature, five themes are addressed: data management, visualization, situational awareness, planning and prediction, as well as integration and collaboration.

In parallel to these developments driven by the geospatial, urban planning and AEC domains, smart cities in general, and the development of DTs for smart cities in particular is also in the focus of the general ICT domain. While the Urban Digital Twin approaches described above consider 3D city models as an essential component, the potential of 3D city models and of geospatial data in general is often not recognized in approaches of general ICT. For example, Farsi et al. (2020) discuss in their book on DT technologies and smart cities IoT, Machine Learning, and Cloud Computing as enablers for UDTs. However, 3D semantic and geometric representations of the physical environment of the city are not considered and location data is only mentioned in connection with consistency assurance for IoT sensors and privacy issues in the localization and tracking of people.

While the aforementioned research explores methodological and technical issues for the creation of UDT, further research is dedicated to investigating non-technical aspects of UDTs, for example from the perspective of public policymaking and human aspects of UDTs (Papyshev & Yarime, 2021).

2.2 Management of urban information resources using catalog services

Catalog systems have been used for the management of urban information resources for some time. Two main fields of application have emerged: Catalog systems for the management of information resources (data, services) in spatial data infrastructures (SDI) and catalog systems as the basis of Open Data platforms.

Access to structured metadata, which describes geospatial data and services is at the core of SDIs. SDI regulations, such as INSPIRE in the European Union, request the provision of metadata as descriptive data for easy thematic and spatial search for official geodata and services as well as for their evaluation (Bernard et al. 2014). SDI catalogs, which are often connected to so-called geoportals (e.g. the INSPIRE geoportal[1]) provide access to these metadata. The data and services described by the metadata are typically distributed over the systems of the governmental data providers. The interoperability of these systems is achieved via standardized Web service interfaces. SDIs follow the publish-find-bind paradigm. This allows software applications supporting the standardized Web service interfaces to directly use information resources registered in the SDI catalog without any manual data conversion. SDI is well established at various administrative levels, down to individual municipalities. However, the scope of SDI is limited to geospatial information resources and they are focussing on users and use cases connected to the geospatial domain.

Open data platforms (e.g., the London Data Store[2]) have a different scope. They intend to allow access to any open data produced by governmental agencies, often without standardising the access interfaces or formally describing data.

An overview of open data catalogs can be found in Lisowska (2016). The author highlighted the importance of open data and listed some Open Data Portals around the world. In this paper Lisowska faced CKAN as one of the most well-known software and DCAT as one of the most used metadata standards for the establishment of Open Data Portals. Although DCAT enables the linking of metadata records for information resources, which is a key function in our approach to managing UDTs, linking is often not used in practice in open data platforms. For each record that is to be made available as Open Data, a separate entry is created in the catalog that is not linked to other records. However, Lopez et al. in 2012 present "Querio City", the prototype of a Linked Open Data platform for Urban Information Modeling, where the main problem is how to represent and manage urban data as an information resource in a practical and consumable way.

3. MANAGING URBAN DIGITAL TWINS WITH AN EXTENDED CATALOG SERVICE

The capabilities of SDI catalogs on the one hand and Open Data catalogs on the other hand, as described above, need to be extended to manage UDTs composed of distributed heterogeneous information resources since their requirements do not fully cover all needs for management of the UDTs. The following sections first show our research methodology and then how the catalog fits into a larger concept for information management in smart city projects. We present the requirements and a metadata schema for a catalog for managing UDT. This metadata schema extends the capabilities of catalogs used today for urban data management to address the needs of UDT management.

3.1 Research methodology

Based on the gained experience and the literature review, a first approach was to conduct a requirements analysis, identifying the available information and data models that existing Open Data as well as Spatial Data Infrastructure catalogs are providing. The existing catalogs were compared in order to find gaps in the developed models and approaches. The aim of the development and the implementation of a new information model is to close the identified gaps of the previously developed models. Adoption and expansion of a new established metadata model was to support missing categories, relationships, and attributes. The final challenge was to map the new data model to an existing catalog service and test its functionality.

3.2 The Urban Digital Twin – a compilation of digital resources of the city

From a technical perspective, an UDT is a compilation of the digital resources of the city. The compilation is specific and is determined by the purpose or task of the UDT. It digitally represents all aspects of the real world required for its purpose and makes them accessible, analyzable and visualizable for applications and users. The aim is to be able to gain essential insights into the current state of the city and its development on the basis of the observation and analysis of the UDT and thus support planning and decision-making in the best possible way.

[1] https://inspire-geoportal.ec.europa.eu/

[2] https://data.london.gov.uk/

Our involvement in smart city projects of several European cities showed that the digital resources in this context can be classified into various types, such as digital models of the physical environment, real-time data, static data sets, urban analyses, simulations and visualizations.
The representation of the physical city objects in digital form often includes their 2 and/or 3-dimensional geometry, on different levels of detail and in different formats. Digital models of the physical environment include a wide range of possible spatial models, information models such as semantic 3D city models or models from Building Information Modeling (BIM), official geobase data, as well as various purely geometric models such as 3D meshes or 3D point clouds. These data can be provided as static files or via Web services/Web APIs.

Real-time data from the various thematic areas of the city comprise all available dynamic information including in-situ (e.g. air quality sensors) and remote observations (e.g. drone imagery). A significant feature of the UDT is its continuous synchronization with the real world where a key feature of UDT is the continuous synchronization with the real world achieved by using (near) real-time data. Their observations are provided using dedicated Web services or platforms with Web APIs.

Static data and collections from the various thematic areas of the city (including mobility, environment, energy, etc.) are typically provided as static files or via Web services/Web APIs. Analysis functionalities are a collection of modeling, analysis and simulation tools that use the collected information and data for the UDT, sensor data or other sources of data and produce results that can be again stored as additional data.

The digital resources are provided by various suppliers and can be used by other stakeholders for their applications. Therefore, in order to understand which digital resources are offered by which parties and in which form, as well as the associated usage conditions, a metadata catalog with a user interface is required. The catalog serves for querying, but also for interactively creating and updating catalog entries. This must not only allow the input of entries by one stakeholder (or a group of stakeholders, such as the public administration), but also by private companies as well as citizens and their organizations (associations, initiatives).
Despite of the heterogeneity of the digital resources, they can be classified into different types and described by metadata in order to make them searchable. With the UDT being a specific compilation/selection of digital resources, each instance of an UDT is represented by a catalog entry of the category *Digital Twin*. The catalog entry does not contain any own resources or records but refers to the digital resources belonging to it via links. Since each catalog entry can belong to a different owner, this means that each actor can define (create in the catalog) their own UDTs for a different purpose.

However, the catalog not only manages the distributed information resources of the UDT, it can also provide additional important values, such as providing information about the involved stakeholders, vendors and users as well as documenting further aspects of urban data management processes (e.g. which method is implemented by which software and which data sets have been successfully combined to fulfill a specific task). Based on the metadata and the links between the datasets, it is possible to collect information from previously registered objects that can help in gathering knowledge relevant to new projects, e.g., what kind of data has been used for a particular project, which methods and software have been used to achieve certain results, who is a responsible stakeholder for the particular data. Such information is not provided by spatial data infrastructures (SDI) and Open Data Platform (ODP) catalogs so far.

The requirements a catalog needs to meet are described in the following section.

3.3 Requirements for managing an Urban Digital Twin and their consideration by existing catalogs

The purpose of a catalog is primarily to manage the information about distributed resources and provide an overview of these resources to potential users and clients in an interoperable manner rather than centrally store and manage all information resources and physical data. From the users' perspective, the catalog plays the role of an interface through which they can efficiently find the desired information that is distributed over different locations.
To provide such efficiency and at the same time respect the distribution and heterogeneity of resources, the catalog should follow certain principles and meet several requirements:

1. Requirements imposed by the distributed nature of information, resources and data: Information about an object in the real world is spread across several stakeholders from different domains. During the planning and execution of urban transformations, many stakeholders are affected and will have to work together (Kolbe et al., 2020). This requires the combination and mutual usage of information resources from different partners in the development, which is a challenging aspect of managing UDTs. Hence, such a catalog should make it conceptually and technically possible for the stakeholders to individually register and publish the information about their resources, access permission, take care of licensing, etc.

In addition, a catalog functionality allows users to select the catalog entries to create a certain UDT instance, facilitating the creation of UDTs.

2. Representing individual objects by individual catalog entries: The digital representation of physical real-world objects (such as buildings or streets) is an important element in distributed UDTs. The catalog should not only be able to register records or data services that provide information about all (or a set of) objects. It should also be possible to register objects individually, i.e. a separate catalog entry is created for each building or street object, which also provides the object with a unique ID (the record ID of the catalog).

The unique (stable) ID ensures that the digital object remains the same over the entire lifespan of its real-world counterpart, and can be identified uniquely. Each object should be consistent and up-to-date. Hence, for each object there is a need for continuous check-ups to ensure that the object is aligned and updated with the changes in its original source and with the changes in its counterpart in the real world. Therefore, versioning and data virtualization is very important.

This makes it possible to register object-specific information resources and link them to the catalog entry of the individual object. Each of these information resources can come from a different stakeholder and be managed under their user account. For example, the city administration (land registry office) can create a building as a catalog entry, and the owner can advertise their existing data on the building (e.g. a document specifying the energy demand) as their own catalog entries and link them to the catalog entry of the building. In addition, the municipal utilities could link the entry for the building's electricity meter to the building entry. Since only metadata are registered and linked, this does not automatically mean access to the registered information resources.

3. Support of various resource types: In an environment with multiple disciplines and with complex interconnections of various domains, we face varied requirements. In terms of resource information management, representing each type of resource requires specific information. Resources are not only datasets or Web services and APIs as is typically represented in catalogs of spatial data infrastructures (SDI) or open data platforms, but also other frequently used resources in the context of smart cities such as sensor devices, software, projects, etc.

4. Distinction between instances and prototypes: Apart from the type of resource, it is notable that in the smart city context, it is necessary to make a distinction between products/prototypes and instances. Prototypes do not belong to the UDT. Only the instances of these are managed in the UDT. Nevertheless, it is interesting to know with which software system from which manufacturer (or from which open source project) was used (e.g. Web Feature Service for access to the Virtual District Model with Web address at which this service can be reached). If instances of these products/prototypes are then actually operated or used for a city or in a project, they are also registered in the catalog and linked to their corresponding products/prototypes. This allows learning from the data management process and using the findings in new processes, in particular in relation to distributed UDTs.

5. Semantic relations: As mentioned earlier, an UDT is a compilation of digital resources of the city, and both the UDT and the digital resources it contains need to be represented by individual catalog entries. In order to express the connection between the UDT and its digital resources, the catalog needs to have the ability to establish semantic relations between catalog entries. In addition to the relations between an UDT and its digital resources, there are also relationships among the digital resources. Expressing these relationships is even more important when the connections are not obvious. These relationships are not limited to the resources. They are also an expression of the connections between the organizations that own or provide these resources. Therefore, a relationship is used not only to link information resources but also implicitly to show how organizations are connected.

Furthermore, for automatic information retrieval and analytics a catalog should be able to represent different types of relationships, for example, to express whether resources can exist independently and are connected only for a specific purpose, or whether they depend on each other, or if there is an aggregation hierarchy between multiple resources.

In the context of spatial data management, SDI, and Open Data Platform (ODP), the existing catalogs and their metadata models are limited in supporting many of the above-mentioned requirements. The focus of SDI and ODP catalogs are more on managing the datasets and Web services, while for the UDT, it is necessary to have information about a wide range of the city's digital resources such as devices, methods, software, etc. as mentioned above.
To the best of the authors' knowledge, there is no concept or implementation for SDI or ODP catalog which distinguishes between prototypes and instances for the digital resource types mentioned above.
An interesting aspect discussed above is the semantic relationships between different resources. The existing catalogs in the SDI and ODP either lack this information or the relationships are defined as optional attributes in the metadata model, making it difficult to discover the meaning of relationships and further possible connections such as relationships between the owners or providers of linked resources.

In the SDI or ODP catalog the focus is not on individual objects. Rather, whole datasets are registered and offered either as a Web service or with a downloadable link covering a collection of objects. This, to a great extent, limits the establishment of semantic relationships on the scale of individual objects.

The main focus of SDI and ODP catalogs are on offering an overview of registered distributed data and information, however the UDT catalog should offer an overview of all applied registered resources and the way they are connected and used in a specific UDT. This includes also the resource owners and their role in an UDT. So, in fact, in contrast to the SDI and ODP, the UDT catalog should provide more than a simple overview. The UDT catalog should in addition offer a backbone for a comprehensive information flow analysis of all registered information (incl. digital resources and involved stakeholders).

The comparison between the requirements of an UDT catalog and existing SDI and ODP catalogs and their metadata models convinced the authors of this paper, that there is a need for a concept for the UDT catalog. This implies further adaptation of the existing metadata models and extension of catalog functionalities. The details of these extended concepts are discussed in the next sections.

3.4 Metadata schema for a Catalog to manage Urban Digital Twins

The stakeholders have different information about the same physical object. They typically do not have their data accessible to other stakeholders, the data is not standardized or not interoperable with other platforms and other stakeholders, or has a different identifier for the same object.

All information required to locate and use a digital resource are fully described and accessible in a metadata catalog. To describe the digital resources and the semantic links between them, the catalog must be based on an appropriately expressive metadata model. Figure 1 shows the UML class diagram of the metadata schema we developed for managing UDTs with a catalog. The most important class of the schema is *InformationResource*. It allows deciding if the specific information is public or private (*ResAccessRight*) within an organization. *InformationResource* has nine instantiable sub-classes for classifying the digital resources (*Project, Online Service, Dataset, GeoObject, Device, Online Application, DigitalTwin, Software* and *Method*). The subclasses *RealWorldInstance*, *Prototype*, and *PhysicalThing* are abstract classes that are only introduced to semantically group some of the sub-classes. In the catalog instance the nine instantiable sub-classes are used as *main categories* (Figure 2). The main group *Dataset* generally contains documents and information about any city information resource stored as files either in the catalog itself or somewhere else in the Web. A distinction must be made between *Online Service* and *Online Application*. The *Online Service* contains data which are provided via an API and have no end user interface (e.g. Web Feature Service providing access to geospatial data), while the *Online Application* has a user interface (e.g. for the exploration of online 3D simulations). The class *Method* represents the functions or algorithms developed by an expert to accomplish a certain task (e.g. a solar potential analysis algorithm).

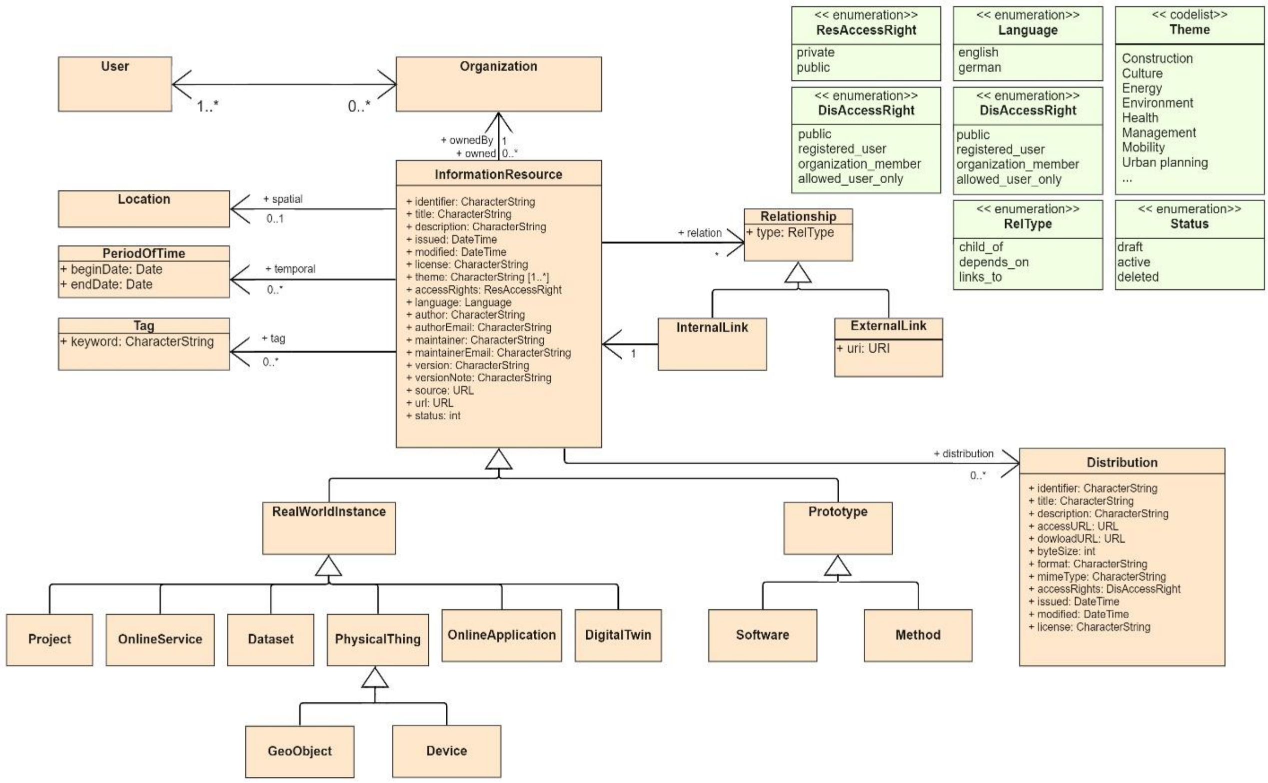

Figure 1: UML Class diagram - Generalized overview of the metadata model of the Urban Digital Twin. This is an adapted version from (Moshrefzadeh et al. 2020) extended by the subclasses *RealWorldInstance*, *Prototype*, and *DigitalTwin*.

The class *Software* is used to represent a specific tool or system, e.g. the FROST Server for the OGC Server Things API[3] or some specific libraries like citygml4j[4]. A *GeoObject* is an object that has a geographic component that is clearly defined (e.g. building, street lamp, bench). The category *Device/Item* is assigned to all physical devices, such as sensors and actuators. *Digital Twin* is a collection of different digital resources defined for a subject topic or a geographic area. It is comparable to a shopping cart in an online shop, where all information resources for a specific use case are collected. There can be more than one Digital Twin, which can be defined by different stakeholders. To make browsing and searching within the catalog even more intuitive, an overview allows navigation within the record at the main category level. The *Distribution* allows the control of the data accessibility of the digital resource. The classes *User* and *Organization* represent stakeholders in the catalog. The instances of these classes represent tenants in the catalog. This enables distributed maintenance of catalog entries. Users can be part of many organizations. But one *InformationResource* can be owned just by one organization. *Relationships* define semantic links between information resources. It is possible to establish a relation to another resource registered in the catalog (*InternalLink*) or a relation to an external resource (*External-Link*). There are three predefined semantic relationship types (*RelType*): *links_to*, *depends_on*, and *child_of*. For connecting independent resources, the *links_to* type of relation is used. The *depends_on* type expresses that an information resource can only be interpreted in the context of another information resource (e. g. the energy certificate for a building has to be seen in the context of the building). For expressing relations like project and sub-projects, the relation *child_of* should be applied. All relationship types can be used to relate catalog entries from different owners. The spatial and temporal extent of a digital resource is represented using the class *Location* and *PeriodOfTime* respectively (Moshrefzadeh et al., 2020).

In addition to the main categories, there are also topics that can be selected for each catalog entry in order to semantically annotate a specific digital resource. Examples for such topics are mobility, environment, energy, finance etc. While each catalog entry can have any number of topics, it must be assigned exactly one main category. The catalog is an important source of information, data storage and a place of information and knowledge sources. All these information makes it easier to search the catalog. In contrast to SDI and open data catalogs, where a catalog entry typically describes an entire dataset, a catalog managing an UDT should be able to also represent individual physical objects (categories *GeoObject* and Device).

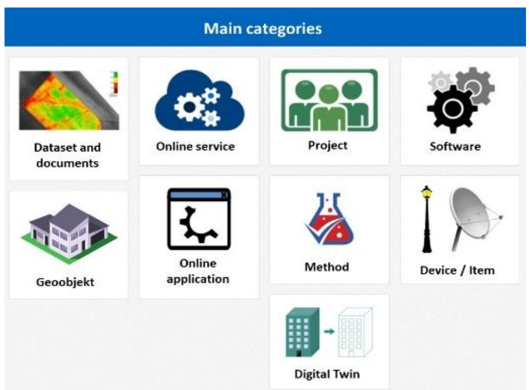

Figure 2: Main categories of information resources

[3] https://www.iosb.fraunhofer.de/en/projects-and-products/frost-server.html

[4] https://github.com/citygml4j

https://doi.org/10.5194/isprs-annals-X-4-W3-2022-119-2022

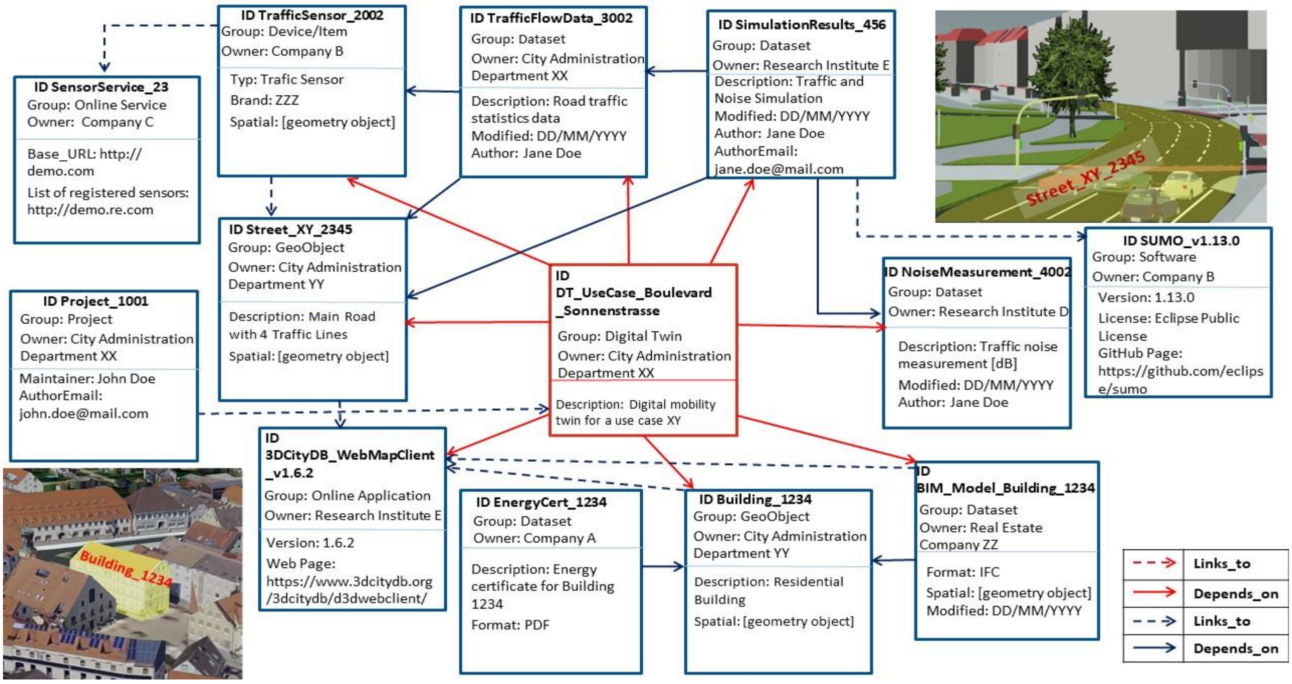

Figure 3: Example for a Distributed Digital Twin represented by connected catalog entries of different types and owned by different stakeholders. Note, that each catalog entry has many further metadata attributes.

Figure 3 shows that the metadata schema makes it possible to manage distributed data owned by different organizations about specific physical objects by creating relationships (see Requirement 2, section 3.2). The connection between the physical object *TrafficSensor_2002* can provide real-time information about the physical object *Street_XY_2345*. The sensor data platform *SensorService_23* provides the possibility to manage data from several sensors. *Street_XY_2345* and *Building_1234* are registered in the catalog as a *GeoObject* and this information resource (not the building and street) is owned by *City Administration Department YY*. For the same objects, metadata has been added to better describe the them. The energy certificate of the building (*EnergyCert_4567*) and the BIM Model (*BIM_Model_Building_1234*) are provided by different organizations (see Requirement 1 and Requirement 3, section 3.2). These two objects are defined in the catalog as distinct dataset entries assigned to *Building_1234* with the *depends_on* relationship (see Requirement 5, section 3.2). A physical sensor (device) *TrafficSensor_2002* is mounted to collect information about the current traffic situation for a *Street_XY_2345* and is connected via a *links_to* relationship. Simulation results (*SimulationResults _456*) are dependent on the traffic flow (*TrafficFlowData_3002*) and noise (*Noise Measurement_4002*) measurements. The simulation (*SimulationResults_456)* is generated in the software tool (*Sumo_1.13.0*) where the *links_to* connection is defined. The visualizations of the objects with geo-location are given in *OnlineApplication* (*3DCityDB_WebMap Client)* and connected with *links_to* relationship (see Requirement 4, section 3.2). All information resources belonging to an UDT are represented within the catalog just like in this example.The UDT itself is also a catalog entry (*DT_UseCase _Boulevard_Sonnenstrasse)* - an instance of the class *DigitalTwin* - which is related to the instances of the classes *GeoObject*, *Device*, *Dataset* etc. by relations of type *depends_on*. Connection between the project (*Project_1001*) and the UDT (*DT_UseCase_ Boulevard_X)* is defined with the *links_to* relation. Each project may contain more than one DT and the DT may be associated with more than one project.

Thus, an arbitrary number of UDTs can be defined by different stakeholders. As a proof of concept, the metadata schema has been implemented using the open source catalog system CKAN (see 4.2). However, we would like to point out that the schema can also be mapped onto the DCAT2 standard (Moshrefzadeh et al., 2020), ensuring interoperability with other catalog systems as well.

4. USE CASE BOULEVARD SONNENSTRASSE IN MUNICH

4.1 Digital resources for Boulevard Sonnenstrasse

In order to illustrate and evaluate this concept a use case in the City of Munich has been implemented. The City of Munich is working on improving multiple aspects of city processes including mobility and traffic, air quality, and energy demands. In this context information on the current situation as well as planned scenarios are available. As a test case, we have chosen "Boulevard Sonnenstrasse", a project by the City of Munich developing future scenarios for the large-scale redesign of a part of an inner-city ring road. The project involves a wide range of urban digital resources, from geospatial data sets to macroscopic and microscopic traffic simulations and VR visualization.
The following list of digital resources of type *Dataset* illustrates the heterogeneity of the resources that represent the UDT of the project "Boulevard Sonnenstrasse". There are a number of data sources with different data formats, which are maintained by different municipal departments. This includes city models (.citygml), topographic data (.gdb, .shp), site plans (.pdf, .dwg), meshes (.obj), point clouds (.las), visualization formats (.czml, .udatasmith) but also simulation results (.csv, .xml). Realtime data is provided by sensor Web services (implementations of the OGC SensorThings API). In addition, with applications using this data as well as a number of different stakeholders, this results in a complex and distributed UDT.

A combined visualization of several data sources is illustrated in Figure 4. This includes the city model (CityGML LoD2 buildings), streets generated from topographic data, and vegetation. This data is converted to visualization formats such as GLTF/CZML, and visualized within the 3DCityDB Web-Map Client. Additionally, point cloud data of the same area is shown.

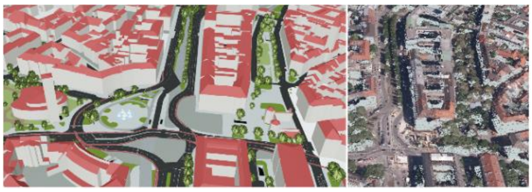

Figure 4: Interactive visualization of different data within a common 3DCityDB Web-Map Client and corresponding coloured point cloud data

4.2 Implementation of the Use Case in the CKAN Catalog

The open source catalog software CKAN has been used for establishing many data management systems. It is used by national and regional government organizations worldwide, where users can publish, search and visualize data. Every catalog entry in CKAN is called a *Dataset* – no matter if it represents a project, an online service or a geoobject (c. f. section 3.3). This is equivalent to the class *InformationResource* in Figure 1. *Resources* in CKAN are equivalent to the class *Distribution* in Figure 2. A CKAN dataset consists of information (metadata) and an arbitrary number of *Resources* that can have different formats. Authorizations to create and edit data are regulated by users associated with organizations. CKAN is used though mainly as a data catalog with an integrated distribution solution. It has been used by the Australian government, the Canadian government, the Singapore government, and found its implementation in the City of Munich, Germany (CKAN Association). CKAN provides numerous search functions and it is allowed to define relations between catalog entries (*Datasets*). In the remainder of this section, we will demonstrate how we use CKAN to implement our approach for managing UDTs with a catalog[5]. The realization of the relations between connected Datasets in the Web interface of the running catalog is shown in Figure 5. The metadata entry *Digital Twin Use Case Boulevard* Sonnenstrasse is a digital resource of type *DigitalTwin* and therefore uses *DigitalTwin* as main category. The graph in Figure 5 shows that nine other catalog entries assigned to different main categories and belonging to different organizations are linked to the catalog entry representing the UDT. The main categories are represented by different colors for the nodes of the graph. The green symbol represents *Dataset and documents*, the grey symbol refers to the

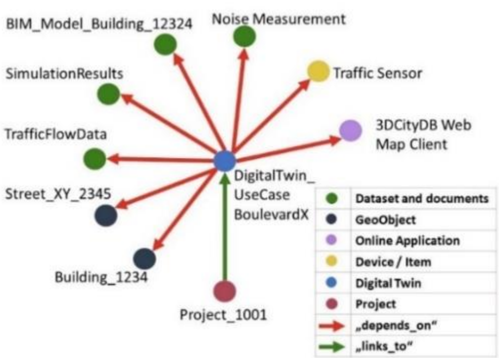

Figure 5: Visual exploration of an UDT represented by linked catalog entries in the CKAN user interface

[5] the extended CKAN software is available as Open Source and can be downloaded from the following repository: https://github.com/tum-gis/SDDI-CKAN-Docker

category *GeoObject*, the purple symbol represents the *OnlineApplication* category, the yellow symbol has been chosen for *Device/Item*, the blue symbol is for *Digital Twin* and the dark red symbol indicates the *Project* category. Clicking on one of the symbols displays the corresponding catalog entry and the arrows indicate the type of links. The red arrow represents the *depends_on* and the green arrow indicates the *links_to* relationship. Figure 2 shows the realization of the main categories (named groups in CKAN) in the Web interface of the catalog.

5. FUTURE WORK: FROM DIGITAL TWINS TO DIGITAL TRIPLETS

The goal of setting up an UDT is to gain significant insight about the city's current condition, thus providing the best possible data basis for planning and decision-making.
However, if an UDT represents the current state of the city, the question arises how to realize *what-if* scenarios that virtually change the state of the city. Our answer to this challenge is a clone of the UDT, an Urban Digital Triplet. The term Digital Triplet occurs in the manufacturing industry, where a third component is added to the Digital Twin. Using data stored in the Digital Twin and/or its parts it would allow the creation of different scenarios with expertise from different human knowledge (Umeda et al., 2019, Umeda et al., 2021). Realization of cloning of the UDT is planned to be implemented and demonstrated within the CKAN catalog. A new catalog entry will be created by cloning the DT entry with all the metadata. The copies of the metadata entries will be modifiable (e.g. change of ownership) and the associated digital resource needs to be cloned and will become modifiable as well. A Digital Triplet is derived from a DT with a cloning process where the references from the DT to the included information resources are mostly staying the same. Only information resources that need to be modified to play *what-if* scenarios must be cloned and the Digital Triplet then links to the cloned instead of the original data. The cloned resources can then safely be modified by the owner of the cloned resources without compromising the original data. For this purpose, it is necessary to create a suitable extension for CKAN and to develop a library (Clone lib) that provides functions for the cloning process, such as cloning on information resources level.

6. CONCLUSION

The establishment of an Urban Digital Twin brings numerous advantages. All collected documentation and data on the physical objects can be used to monitor and control the physical objects. The same data can be used to create *what-if* scenarios and display various simulations that can form the base for the future implementation of the designed scenario and to make important decisions at the city level. However, due to the distributed nature of urban data, establishing and maintaining an UDT is challenging from a data management and governance perspective. Using the proposed catalog as a central component for data management, distributed systems and multiple stakeholders can participate with their existing resources, tools, and applications. This solution provides easy information sharing both internal and external to the organization, benefiting everyone. There is no loss of information between users and objects and assignments /responsibilities can be clearly defined.

The developed methodology allows, in the process of registration of a new information resource, the selection of the main category

with its metadata that best fits the attributes which are going to be registered. This approach differs from the approach we encounter in INSPIRE and the OGC catalog system.
If a Digital Twin is properly defined, it is possible to track past events and register current events, and using such information may provide useful information for plans (Shahat et al. 2021a).
Umeda et al. (2019) see that any Digital Twin concept can be improved, saying that it is impossible to create a DT that is perfectly identical to the physical world. For Industry 4.0 Umeda introduces the "Digital Triplet", which aims to help users to be able to create different scenarios during the product life cycle of the object. It considers the creation primarily as the creation of various scenarios, the results of which can significantly impact the life of a physical object. Since the UDT represents the current state of the real city, and as such should not be changed, the concept of cloning the UDT is required in order to create Digital Triplets. Digital Triplets can be defined by any user and they are clones of a DT to be modified by the user. It has the purpose to define urban development scenarios which need to be used across all applications and analytical tools.

ACKNOWLEDGEMENTS

We thank the City of Munich for their cooperation, funding, and for providing the data used in the Digital Twin Munich project funded by the Federal Ministry for Digital and Transport and in the project Connected Urban Twins (CUT) funded by the Federal Ministry of the Interior and Community .
Many thanks to Christof Beil and Bruno Willenborg for the constructive discussions and their help with the description and images provided for the use case Boulevard Sonnenstrasse.

REFERENCES

Batty, M., (2018). *Digital twins.* SAGE Journals, Environment and Planning B: Urban Analytics and City Science 2018, Volume 45(5), 817 – 820.

Butot, V., Bayerl, P. S., Jacobs, G., de Haan, F., (2020). *Citizen repertoires of smart urban safety: Perspectives from Rotterdam, the Netherlands*, Technological Forecasting and Social Change, Volume 158.

Bernard, L., Mäs, S., Müller, M., Henzen, C., Brauner, J., (2014). *Scientific geodata infrastructures: challenges, approaches and directions.* International Journal of Digital Earth, Volume 7, Issue 7, 613 – 633.

CKAN Association: CKAN. https://ckan.org/ (Accessed 30.03.2022)

Deng, T., Zhang, K., Shen, Z.-J. M., (2021). *A systematic review of a digital twin city: A new pattern of urban governance toward smart cities.* Journal of Management Science and Engineering, Volume 6, 125 -134.

Farsi, M., Daneshkhah, A., Hosseinian-Far, A., Jahankhani H., (2020). *Digital Twin Technologies and Smart Cities.* Springer. ISBN 978-3-030-18734-7.

Füller, A., Fan, Z., Day, C., Barlow, C., (2020). *Digital Twin: Enabling Technologies, Challenges and Open Research*, IEEE Access, Volume 8, 108952-108971.

Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B., (2020). *Characterising the Digital Twin: A systematic literature review.* CIRP Journal of Manufacturing Science and Technology, Volume 29, Part A, 36-52.

Kolbe, T. H., Moshrefzadeh, M., Chaturvedi, K., Donaubauer, A., (2020). *The Data Integration Challenge in Smart City Projects*. Chair of Geoinformatics, Technical University of Munich, https://mediatum.ub.tum.de/doc/1554725/1554725.pdf, Accessed: 04.03.2022.

Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., (2018). *Digital twin in manufacturing: A categorical literature review and classification.* IFAC-PapersOnLine, Volume 51, No. 11, 1016–1022, https://doi.org/10.1016/j.ifacol.2018.08.474.

Lisowska, B. (2016). Metadata for the open data portals. Discussion Paper No.6, Joined-up Data Standards Project. http://devinit.org/wp-content/uploads/2018/01/Metadata-for-open-data-portals.pdf

Lopez, V., Kotoulas, S., Sbodio, M. L., Stephenson, M., Gkoulalas-Divanis, A., Aonghusa P. M., (2012). *QuerioCity: A Linked Data Platform for Urban Information Management*. The Semantic Web – ISWC 2012, Springer Berlin Heidelberg, 148 – 163.

Mohammadi, N. & Taylor, J. (2017). *Smart city digital twins.*

IEEE Symposium Series on Computational Intelligence (SSCI), doi: 10.1109/SSCI.2017.8285439.

Moshrefzadeh, M., Machl, T., Gackstetter, D., Donaubauer, A., Kolbe, T. H., (2020). Towards a Distributed Digital Twin of the Agricultural Landscape. Journal of Digital Landscape Architecture (5), 173 – 186.

Niculescu, A. I. & Wadhwa, B., (2015). *Smart cities in South East Asia: Singapore concepts - an HCI4D perspective*. ASEAN CHI Symposium'15, 20 – 23, doi: 10.1145/2776888.2780362.

Papyshev, G. & Yarime, M. (2021). *Exploring city digital twins as policy tools: A task-based approach to generating synthetic data on urban mobility.* Data & Policy, 3, E16.

Preidel, C., Bormann, A., Oberender, C., Tretheway, M., (2016) *Seamless Integration of Common Data Environment Access into BIM Authoring Applications: the BIM Integration Framework.* European Conference on Product and Process Modelling, Volume 11.

Ruohomäki, T., Airaksinen, E., Huuska, P., Kesäniemi, O., Martikka, M., Suomisto, J. (2018). *Smart city platform enabling digital twin.* In 2018 International Conference on Intelligent Systems (IS), 155-161, IEEE.

Schrotter, G. & Hürzeler, G., (2020). *The Digital Twin of the City of Zurich for Urban Planning.* PFG – Journal of Photogrammetry Remote Sensing and Geoinformation Science 88(5), 99 – 112.

Shahat, E., Hyun, C.T., Yeom, C., (2021). *City Digital Twin Potentials: A Review and Research Agenda. Sustainability* 13, no. 6: 3386.

Umeda, Y., Goto, J., Hongo, Y., Shirafuji, S., Yamakawa, H., Kim, D., Ota, J., Matsuzawa, H., Sukekawa, T., Kojima, F., Saito, M., (2021). *Developing a Digital Twin Learning Factory of Automated Assembly Based on 'digital Triplet' Concept*. Proceedings of the 11th Conference on Learning Factories (CLF).

Umeda, Y., Ota, J., Kojima, F., Saito, M., Matsuzawa, H., Sukekawa, T., Takeuchi, A., Makida, K., Shirafuji, S., (2019). *Development of an education program for digital manufacturing system engineers based on 'Digital Triplet' concept*. Procedia Manufacturing. 31. 363-369.