Using ICDD for BIM and GIS Integration in Infrastructure

Judith Krischler[1*], Paul-Christian Schuler[1], Jakob Taraben[1] and Christian Koch[1]

[1] *Chair of Intelligent Technical Design, Bauhaus-Universität Weimar, Germany*

Abstract
Planning of infrastructure projects usually involves a thorough knowledge about the built and natural environment and involved assets, making it necessary to consider heterogeneous data represented in heterogeneous forms and shapes. Due to the large spatial extend and effects of such projects, planners need an accumulated overview from as early as possible to make well-informed decisions, making it necessary to consider both data from the domains of Geoinformation Systems (GIS) and Building Information Modeling (BIM). Model container approaches, such as the internationally standardized Information Container for Linked Document Delivery (ICDD), hold the potential to create links between building data and their geolocated context in one environment using GIS-services. A Linked Data approach contributes to a BIM and GIS integration that allows those domains to stay independent yet connected, making transitions between data models redundant. This paper proposes a methodology for deriving the geolocation from different ICDDs containing heterogeneous data to aggregate them together with relevant GIS data specifically for the use case of (early) infrastructure planning. The result includes a newly aggregated ICDD including both BIM and GIS data. The methodology is implemented in an opensource framework, and an academic case study is furtherly presented. Eventually, results are discussed, and conclusions are drawn.

Keywords
ICDD, BIM/GIS Integration, Infrastructure, Georeferencing

1. Introduction

Infrastructure projects often contain many different assets, such as bridges, tunnels, or crossroads structures, among more specific assets like railway stations and other building constructions. When starting a new project, one of many necessities is to identify all affected existing buildings within the respective project area and to collect data about them to create a solid basis for future informed decision-making along the planning process. The collected data might contain all sorts of files, including documents, 2D drawings, 3D Models or data from Geographic Information Systems (GIS) of the surroundings. The final data set must be curated carefully for the current use cases, which includes assembling an individual selection of the required files and excluding all data that is not considered.

LDAC 2024: 12th Linked Data in Architecture and Construction Workshop, June 13–14, 2024, Bochum, Germany
* Corresponding author.
Judith.krischler@uni-weimar.de (J. Krischler); paul-christian.schuler@uni-weimar.de (P.-C. Schuler); jakob.taraben@uni-weimar.de (J. Taraben); c.koch@uni-weimar.de (C. Koch)
0000-0003-2107-9558 (J. Krischler); 0000-0003-3758-1286 (P.-C. Schuler); 0000-0003-2397-5452 (J. Taraben); 0000-0002-6170-7611 (C.Koch)

Linking drawings and documents to building information models has been subject to numerous research, but especially the full integration of BIM and GIS data still faces many obstacles and could profit from a Linked Data approach. Meanwhile both domains play a crucial role in infrastructure planning, they operate on different scales, use different data standards and usually different software solutions. Attempts to convert between data models often lead to the loss of information, caused by the lack of matching between classes or problems to unify topologies. A linking concept would allow both domains to remain separated yet connected, without forcing a unification [1]. Furthermore, an integrated view on both BIM and GIS data in the same environment would enable many trans-sectoral strengths that are so far limited to the respective technical scope.

The goal of the present article is to explore different technical solutions for linking BIM and GIS data in a common software environment. For this purpose, first, requirements are evaluated that arise from using heterogeneous data in a geolocated context. The review of relevant research depicts different approaches and technologies of linking both domains within the Linked Data context. Secondly, a methodology for connecting geolocated BIM and GIS data is described. Subsequently, an implementation and a case study are presented to test the formerly developed methodology. Eventually, the results are discussed, and conclusions are drawn.

2. Related Work

This chapter focuses on the aspects of a BIM and GIS integration within the Linked Data context. To achieve this, requirements towards a full integration of BIM and GIS with respect to infrastructure projects are identified. Next, Linked Data concepts for BIM and GIS integration are reviewed and a possible solution, the Information Container for Linked Document Delivery (ICDD) is introduced. Furthermore, an explanation is given to why the ICDD might be a promising candidate, taking into account the identified requirements of the first part of this chapter. In general, it has to be distinguished between requirements that have to be met when authoring ICDDs and requirements that matter when reading ICDD. Within this paper, the aim lies overall on the ICDD-authoring side, as it focuses on how to link BIM and GIS data for a specific purpose.

2.1. Requirements for BIM and GIS Integration in Infrastructure

The meaning of the term BIM and GIS integration can be interpreted from different points of view. The authors of [2] classify BIM and GIS integration approaches into data, process and application level integration, which is a definition that has been adopted by many subsequent studies. The Technical Report ISO/TR 23262, however, defines only two levels: data level and service level and therefore does not consider a process level in particular [3].

In this context, [1] see BIM and GIS integration as a matter of interoperability according to ISO/IEC 2382:2015, so integration on the data level, and therefore their defined requirement for an integration includes the exchange of information between different software. As mentioned by [2], research in the direction of a linked methodology also falls within the data level category, as the work of [3] shows exemplarily. Some approaches show a mix of integration level, such as [4]. The authors created horizontal federation (mappings) of the BIM ontology ifcOWL and the several GIS ontologies defined by ISO/TC 211. The mapping is conveyed on four different layers: Metamodel level, conceptual (abstract) schema level, conceptual (application) schemas level and implementation schemas level [4], and therefore includes both data level and application/service level integration. Integration on the process level includes bringing together domain-specific workflows, as stated by [2].

This paper considers the aspects presented in the following chapters taking into account the data and service levels defined in the Technical Report ISO/TR 23262 and does not consider the process level, as it represents the latest agreement stated in the standardization work. The following paragraphs present requirements affecting both data and service level, offering a specialized view on infrastructure projects.

Topological Dimensionality

When integrating BIM and GIS in infrastructure, heterogeneous data with different dimensionalities has to be linked across domains. Dimensionalities refer to the topological dimensionalities as defined for example by DIN EN ISO 19107. This standard distinguished between 0-dimensional objects (e. g. points), 1-dimensional objects (e. g. curves), 2-dimensional objects (e. g. surfaces) and 3-dimensional objects (e. g. solids).

The relevance of considering the aspect of topological or geometric representation is currently a matter of discussion, according to [5]. However, diverse data of differing data models and/or topological dimensionality must be addressed differently for the purpose of integration or additional processing due to its differing nature. For example, point cloud data, as an accumulation of points in a cartesian coordinate system but without further semantics, hold different potentials and challenges for usage compared to engineering drawings, that might hold more semantics, yet are only represented by lines, points and planes in a two-dimensional space. Furthermore, it matters when evaluating data quality, suitability for the usage within specific use cases and the expressiveness towards certain aspects, such as accuracy. Differing modelling paradigms might cause integration trammels as well but are not considered within this paper.

Georeferencing and Accuracy

As a matter of fact, one aspect towards full integration covers not only the technological methods but also the consistent exchange of georeferencing among applications and formats [6]. Meanwhile the authors of [6] investigated georeferencing in the context of IFC and CityGML conversion, the geolocation itself can function as a link between different objects, features and assets overall. Absolute georeferencing usually refers to, among other parameters, a homogeneous coordinate reference system (CRS), based on a common geodetical and vertical datum as well as a common map projection [7]. To create an aggregated set of data which is linked towards its location, it has to be transformed to the target CRS first. Relative georeferencing can be used to create an individual, project-specific CRS, as explained by [8]. Without providing a properly defined absolute or relative geolocation in both domains, an integration is not feasible. Especially the application of Geoinformation Systems function is based on spatial relationships. Without a shared placement, spatial operations do not lead to valid results. Hence, a concept for linking both concepts must include at least relatively georeferenced data.

The aspects of geodetic accuracy of the data to be linked is tightly connected to the aforementioned georeferencing. Linking data with differing accuracies could exacerbate problems of integration, especially when using spatial operations and/or the linking of instances that are not in the same location. [8] defined a so-called Level of Georeferencing (LoGeoRef) as a quality indicator for geodetic placement and how to apply it to the openBIM format IFC.

Besides the georeferencing quality, the Level of Accuracy (LOA) describes the standard deviation of a document, e. g. an engineering drawing or a 3D model, in relation to the ground truth and categorises this accuracy in 5 levels, as described by [9]. When planning infrastructure projects, the accuracy and correct geolocation of input data is highly relevant as the need for accuracy raises significantly with the ongoing planning process. A visualization of linked data with mixed LOAs could imply a higher accuracy than what is really given. This makes it necessary to clarify it clearly when integrating.

Data Quality and Information Requirements

To evaluate the data quality of geographic information, DIN EN ISO 19157 defines, depending on the scope or use case, parameters such as completeness, logical consistency, positional accuracy, temporal quality, thematic quality and meta quality elements. The BIM-specific series of standards, DIN EN ISO 19650, refers to ISO 8000 when it comes to data quality, which includes the aforementioned aspects. This touches the concept of data quality as mentioned by DIN EN ISO 19157 as well as by the means of, for example, Employer's Information Requirements (EIR). [5] mentions the concept of so-called GeoEIR, which is an adaptation of the common perception in BIM of defining the information need of the employing party that has to be delivered [10]. The ongoing standardization of the agnostic Level of Information Need (LOIN) principles, as laid out by DIN EN 17412, has opened the possibility of adapting the concept as well for the GIS environment, giving the chance to reconcile the BIM-specific Level of Development (LoD) and the GIS-specific Level of Detail (LOD). Information requirements and data quality, following the precept of logical consistency, lead to the need for a structured way of linking both domains, allowing the overarching application of operations and queries.

Granularity and Mapping

Both domains, BIM and GIS, have fundamentally different scopes. Meanwhile in BIM, the asset and furthermore the components of the building are in the center of attention, GIS provides a scope on the feature level and does not include individual components. Consequently, linking both domains means also linking specific building objects with spatial features or even abstract concepts, like land boundaries, which might not be represented in the physical world. Often, there are not even corresponding counterparts to be matched so that the BIM component might not even exist within the GIS world. Furthermore, as stated by [5]: "While in IFC relationships are objectified, in GIS relationships are expressed more as an association." However, both standardization organizations behind both specialist domains, OGC and buildingSMART, are working on harmonizing their data models, e. g. CityGML and IFC, to ensure compatibility on a data level [11]. The 3D city modelling marks a transition zone between both domains as the Level of Detail (LOD), respectively the geometric granularity, reaches component level in LOD3 and crosses therefore the boundary to the BIM world [12]. However, it is not yet a seamless solution, as it is still based on the idea of conversion, yet nonetheless it paves to way for a Linked Data approach.

The challenge of differing granularity and scope is not specific to the infrastructure sector. Nonetheless, infrastructure projects depend heavily on GIS data and are particularly faced with the need of bringing both worlds together. On top of that, one specific of infrastructure planning like roads and railways, is that building components such as signals are mainly bound to the road or rail alignment, so to speak relatively placed during planning, which is a parametric dependency [13]. Linking therefore could also just mean creating mixed relationships between both entities that, on the one hand, allows inheritance and hierarchy and on the other hand

GIS-specific spatial relationships such as "includes" or "intersect", as defined by the standard DE-9IM [14].

Provision of Data

Infrastructure projects have to include a lot of GIS data from different backgrounds. In addition to surveying data, this also includes environmental planning data such as geology, vegetation structure or nature and species protection as well as basic geodata such as cadasters, topographical maps or digital elevation models. Due to the long period from planning to construction of the respective infrastructure, this data is prone to changes and therefore, GIS services are beneficial. Meanwhile services are a common concept in the GIS domain, the concept of services as by the GIS definition is mostly irrelevant among the BIM domain [5]. This causes an area of conflict as it means that a dynamic, respectively "on-demand", concept is confronted with a static concept. To maintain the coherence among both domains, a full integration must allow both concepts to coexist on an equal footing.

2.2. Linked Data Concepts for BIM and GIS Integration

Latest research has identified the potential of using Linked Data concepts for BIM and GIS integration. The authors of [1] reviewed different BIM and GIS integration approaches and distinguished between three main approaches: 1. converting between open standards, 2. unification/combination of open standards and 3. linking open standards through linked data and semantic web technology. Eventually it was stated the third approach to be the most promising one for future research towards a full integration[1]. Supporting those findings, the authors of [15] found a growing interest in unification attempts and furthermore, a trend of unidirectional approaches for the integration of BIM to GIS. [16] states that: "[...] a full harmonization would not be appropriate, and further improvements of the semantic interoperability must be achieved by linking and mapping heterogeneous information models.".

Within the Linked Data context, different approaches have been stressed and evaluated with regard to effectiveness of integration. Furthermore, several works include the application of Linked Data concepts for specific use cases.

The authors of [3, 4, 17–19] describe contributions to the BIM and GIS integration using and connecting ontologies. [3] used the ifcOWL ontology, extended by tunnel and infrastructure elements as well as CityGML models, that were converted to RDF. Because of the complexity and size of ifcOWL, the authors of [20] introduced a Building Topology Ontology (BOT) to enable also the combination with other ontologies and therefore hold the potential to also serve the BIM and GIS integration. Connected to that, [17] used an ontology for the surrounding environment of a building which is semantically compatible to be linked to the BOT of [20]. [21] use ontology databases are used within tunnelling context to link and assess heterogeneous data in the geospatial context.

The use of Linked Data concepts imply also challenges, besides handling complex ontologies like ifcOWL. The authors of [22] raise important questions on identity relations for linking heterogeneous objects from the BIM and GIS domains. They point out the lack of biunique meaning of an identity relation, leaving it to the risk of misinterpretation and therefore suggest the development of alternative vocabularies, the categorization of information and the enrichment of links with meta- or contextual data as possible overcoming [22].

2.3. The Information Container for Linked Document Delivery (ICDD)

The ICDD, which is based on COINS developed in the Netherlands, provides an open framework for exchanging heterogeneous and distributed building data [23], allowing for cross-domain integration. It is subject to the international standard DIN EN ISO 21597 and utilizes the compressed ZIP format (*.icdd). At its core, an ICDD contains an index.rdf file, which describes and lists all the subjects to be linked. Furthermore, an ICDD container leverages two Ontology Resources, expressed as Resource Description Framework (RDF):

- Container Ontology: Defines the structure and metadata associated with the container itself and describes the documents it contains.
- Linkset Ontology: Defines the semantic relationships and links between the various documents and data pieces embedded within the container.

The *Payload documents* contain the actual files and documents to be linked, such as model files or bill of quantities. The *Payload triples* contain the actual links, following the specification from Ontology Resources. Beyond established file-level referencing, ICDD boasts the ability to link on a sub-document level. This granular approach allows, for example, referencing of individual object attributes such as building component within an IFC file.

Latest research has applied the concept of ICDD to the context of Common Data Environments [24–26] as well as highlighting the needs for unleashing more potential by using hybrid linksets and to develop software prototypes that are able to write ICDDs using the Shapes Constraint Language (SHACL) to validate the created ICDDS [27]. So far, the ICDDs has not been subject to research actively using ICDD for linking BIM and GIS data, even though the heterogeneous nature of an integration of both domains makes it a considerable technology and the potential has been identified by other authors, as well, such as [5].

2.4. Research Question

This paper explores the possibility of linking GIS and BIM data using a model container approach like ICDD. To achieve this, requirements for a successful BIM and GIS integration have been collected from related research and standards. The developed methodology and implementation describe a framework to create ICDDs from GIS environments in consideration of BIM data and how to eventually assemble it. A case study is carried out to test the identified requirements and to conclude on how the ICDD standard can or cannot fulfill them in its current form.

3. Methodology

This chapter describes a methodology to assemble BIM and GIS data within the ICDD standard for the purpose of providing relevant data for planning or modifying infrastructure assets within a defined project area. When planning linear infrastructure projects, due to the network character, larger parts of the traffic network have to be considered to make sure that effects on adjacent traffic junctions are taken into account. Furthermore, all structures need to be identified, that might be affected by potential building or maintenance activities. To prevent redundant work, the existing documentation referring relevant structures must be reduced to the necessary information for the projected use case and/or extended by necessary information that was not yet included, e. g. information not considering the asset, but the entire project area.

Figure 1 gives an overview of the methodology. Based on a user-defined region of interest in a defined CRS, relevant data needs to be collected. The region of interest can be represented by a georeferenced polygon.

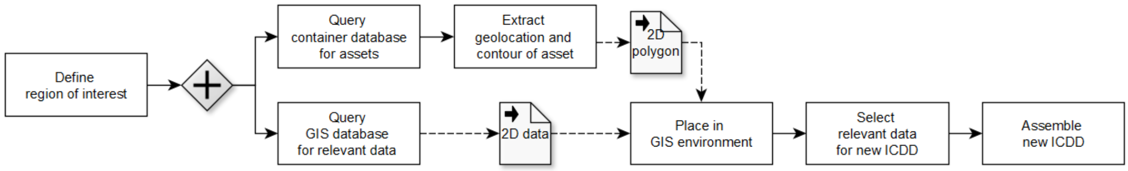

Figure 1: General workflow of proposed framework

Relevant data means on the one hand the integration of inventory data, meaning data about built assets, e. g. a bridge, which has already been collected and assembled in ICDDs within a database. On the other hand, relevant data refers to environmental information, such as adjacent buildings, cadaster data, flora and fauna, etc., which lays within the region of interest and can be accessed by querying GIS databases.

While GIS data is already georeferenced and can therefore be selectively recalled with the help of spatial operations, ICDDs itself are not georeferenced, which makes it necessary to first extract the geolocation and the contour of the asset from the browsed ICDDs. Based on the resulting georeferenced 2D polygon, spatial operations can now be applied as well. Only data which lies within the defined region of interest is furthermore considered relevant to the use case. The resulting georeferenced 2D polygon can then, together with the extracted GIS data, be placed within a GIS environment to visualize all relevant built assets of the area, among the relevant GIS data.

Assuming that the earlier identified ICDDs do not only contain information relevant to the use case, the user can now decide which data, among with the selected GIS data, shall be linked and assembled within the new ICDD. This selection process is carried out only considering the assets that lie within the region of interest, which prevents including information which are of no use for the targeted use case.

4. Implementation

This chapter describes the implementation that was carried out to use ICDDs containing, among others, BIM data from built structures within the GIS environment QGIS. The implementation targets to query for asset data within a certain geographical area, filter it and to newly aggregate and link it within an ICDD for the purpose of planning. The implementation was carried out using the QGIS Python API and different openly accessible Python libraries such as geoPandas and IfcOpenShell.

4.1. Extraction of Geolocation from ICDDs

Firstly, a region of interest within a target coordinate reference system is defined, for which planning actions shall be projected. This region of interest is expressed as a 2D globally georeferenced polygon and can be either defined directly within the open-source software QGIS using onboard-operations or imported within the GIS-format geoJSON using the QGIS Python API. The developed plug-in then queries for those ICDDs that are located within, respectively overlapping with, the formerly defined region of interest.

Figure 2 top left depicts the subsequent process. Once the database query got toggled, the developed QGIS plug-in searches the database with ICDDs and browses the index.rdf (idx) for IFC files. If there are IFC files linked within the browsed ICDD, the geolocation of each bridge can be extracted from them using the parameters of IfcProjectedCRS and IfcMapConversion. The parameters from IfcMapConversion are used to transform the local IfcProjectedCRS to the global one of the region of interest.

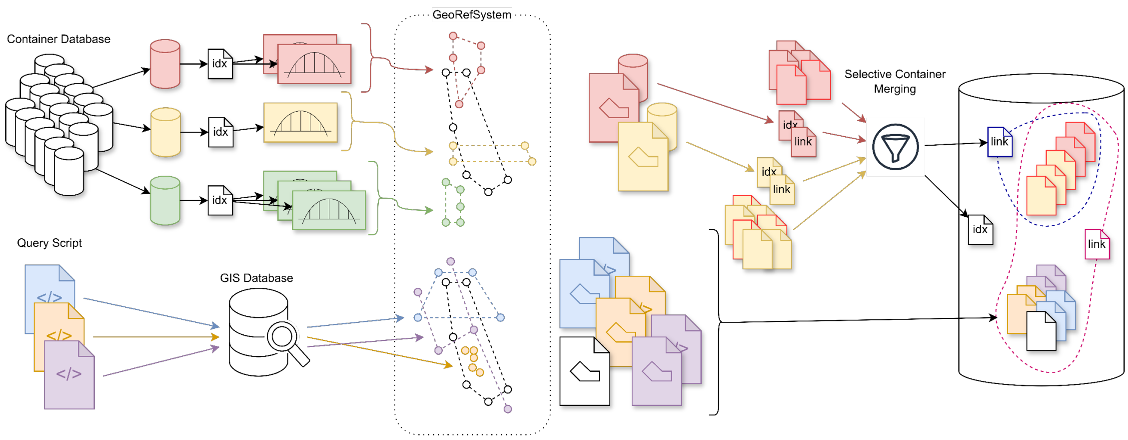

Figure 2: Complete process for creating a new ICDD from relevant data.

4.2. Deriving the Concave Hull of the IFC File

In a next step, the extent of the model is derived as a georeferenced polygon and can be compared with the region of interest. In this case, it is assumed that the ICDDs contain an IFC file of the respective asset. In order to compare the geolocation of the structure with the region of interest, the spatial extent of the structure must be extracted. To achieve this, three options for the spatial representation were considered: Deriving a bounding box, deriving a convex hull or deriving a concave hull. Figure 3 depicts the differences between the approaches.

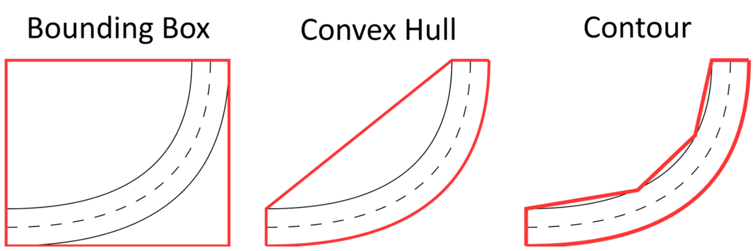

Figure 3: Possible approximation methods.

A bounding box encloses the building with a rectangle, which is a relatively simple calculation, yet gives imprecise results that include a lot of irrelevant area. A convex hull refers to the smallest area that contains all points as well as all connections between the points and gives therefore more precise results than the bounding box [28]. However, in case of cantilevered objects, a lot of surface area is incorporated that does not belong to the structure. The concave hull only includes the area that contains all points and is as small as possible. The area is dependent on the maximum permissible edge length and therefore delivers precise results even with cantilevered objects [29]. To achieve precise results, the concave hull approach was chosen and extracted as a 2D georeferenced polygon, which can now be compared to the polygon of the region of interest.

4.3. Enrichment with GIS Data

Parallel to the extraction of the geolocation, the Overpass API is used to query, in this case, OpenStreetMap (OSM) for maintenance-relevant GIS data within the region of interest as it is shown in Figure 2 bottom left. The user decides on which GIS data shall be added when using the plug-in and creating the Overpass Query. Then, spatial operations are again performed to spatially limit this data to the project area. The Overpass Query can be performed either directly in QGIS using plug-ins like QuickOSM or expressed using Overpass XML or Overpass QL within the QGIS Python API. The result from this step are new GIS layers of necessary information that serve the defined use case(s).

4.4. Aggregation of new ICDD according to Use Case

In a final step, the needed data from the bridge ICDD as well as the queried GIS data can be aggregated within a new ICDD that represents the new dataset for the specified use case. To achieve this, the user is selecting the documents that shall be taken from the relevant ICDDs to the new ICDD. The triples of the chosen data of interest are taken to the new index.rdf. The chosen data is filed to the *Payload documents* folder.

The relevant triples of the chosen data from ICDD are extracted from the index.rdf of the respective containers and merged within the index.rdf of the new container. The URIs of the transferred triples are also transferred, provided they remain unique. If there are conflicts between two or more objects, new URIs are assigned. The *Payload documents* folder of the newly created ICDD is structured in compliance with the different bridge assets and the documents are filed accordingly. Existing links of the selected payload documents are as well copied to the new *Payload triple* folder if the triple is still relevant, meaning if the object that was linked to the selected data still exists within the new ICDD.

Similar to the aforementioned process, the GIS data originating from OpenStreetMap, as well as the expression that led to query for them, is stored to *Payload documents* in a suitable data format, such as geoJSON, compliant with the defined target CRS. The links for including the GIS data are of the Elaboration type (compliant with DIN EN ISO 21597-2) and connect the GIS layers with the bridge assets as well as with the executed query expressions. Figure 2 shows the complete process as described above.

5. Case Study

To validate the presented implementation, a case study was carried out. For this, three different bridge assets were modelled and exported as georeferenced IFC files, containing the said geolocation and conversion parameters within IfcProjectedCRS and IfcMapConversion. The IfcProjectedCRS contains the name of the target CRS, a description of the same, along with further information like geodetic datum. The IfcMapConversion refers to a defined source CRS and the target CRS (as defined in IfcProjectedCRS) and gives the coordinates of the project origin to be transformed [30]. The bridges themselves differed in their overall width and/or length so that the contour would show differing results. Those IFC files were assembled in exemplary ICDDs, linked only coarsely to other sample files to reenact the described use case, and stored in a database. In the next step, the project area was imported to the QGIS plug-in using a 2D polygon. The database of the sample ICDDs was then connected.

Now, the index.rdf files of the contained ICDDs were browsed for IFC files and the contour of all three bridges was extracted as georeferenced polygons and placed within the QGIS environment, as can be seen in Figure 4. The extracted polygons are shown in red within the area of interest in blue. The polygons were added as layers to QGIS. It is visible that all three bridges either lie within the project area or intersect it partly, which is why they are considered relevant.

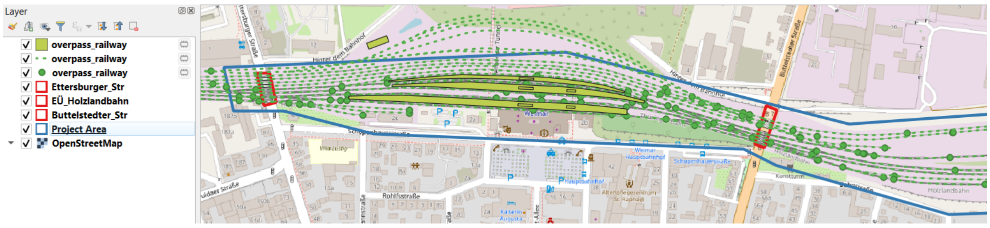

Figure 4: Results of Overpass query (in green) and concave hulls (in red), displayed in QGIS.

Now, additional relevant GIS data is queried using the Overpass API within QGIS. In this case, a query is defined which searches for features with the key "railway" within a polygon, which is in this case the project area layer. Listing 1 illustrates this query as semantic triple, as these queries are instantiated as external document in index.rdf.

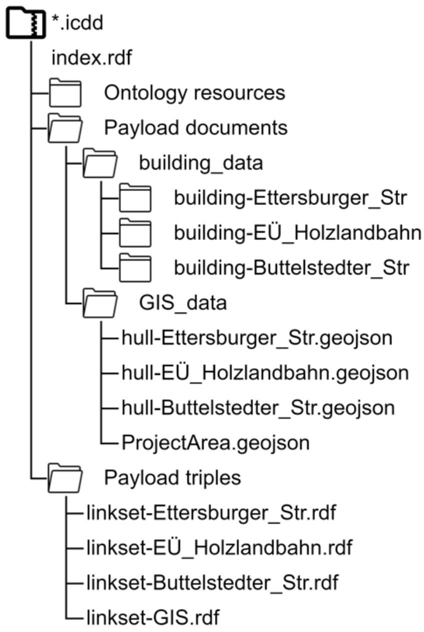

Figure 5: Final Structure of ICDD

```
<ct:containsDocument>
  <ct:ExternalDocument rdf:about="http://BUW-ITD.org/
      index#ExternalDocument5daa3aec-af11-4747-b1de-c2b068abb920">
    <ct:name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
      OverpassRequest_ProjectArea_Railway</ct:name>
    <ct:url rdf:datatype=http://www.w3.org/2001/XMLSchema#string>
      http://overpass-
      api.de/api/interpreter?data=(way[railway](poly:"50.991860489939953
      11.320502657243983 50.992076906735186 11.326896656259354 50.991261454779242
      11.332208848134721 50.991536137555677 11.337688120413073 50.991324474906357
      11.338787767276999 50.990336625579104 11.338619021909103 50.990208399410228
      11.334092246307851 50.990766677763922 11.330855315799738 50.990955426595818
      11.330243365953681 50.991094980300325 11.323316330457848 50.991354345113621
      11.321626988272911 50.991348922215856 11.320692893840727 50.991860489939953
      11.320502657243983"););(. ;>;);out;
    </ct:url>
  </ct:ExternalDocument>
</ct:containsDocument>
```

Listing 1: Overpass query as an instance in index.rdf.

The results of the query are added as layers to QGIS. Figure 4 shows the QGIS surface with additionally added layers, displayed as green points, dashed green lines and green polygons.

In a final step, the new ICDD with both GIS and BIM data can be assembled (as can be seen in Figure 5**Error! Reference source not found.**). To do so, the bridge data to be transferred is selected by the user. The new ICDD is now ready to be assembled and includes both the queried GIS data and the formerly selected bridge models, which are filed in separate folders among the

Payload documents. The *Payload triples* contain now the still relevant old linksets of the bridge assets and the new linkset of the GIS data.

6. Discussion

This paper introduced the application of the container concept, namely ICDD, to the problem area of BIM and GIS integration, for which the ICDD standard has been identified to hold the potential to support a BIM and GIS integration. This chapter revisits the requirements for BIM and GIS integration in infrastructure, which were earlier described, and gives a concluding overlook about how they can be met by the ICDD standard and how future improvements could look like.

Topological Dimensionality

The ICDD standard is agnostic, which means no obstacle to connecting data of differing topological dimensionality. However, a mutual visualization of both BIM and GIS data in the QGIS environment made it necessary to reduce the dimensionality of the BIM models by deriving a 2D concave hull, which could then be used to identify relevant bridge assets. This is considered more an obstacle to a future ICDD-reading software environment or viewer, which is aiming at displaying both domains equally, yet the ICDD standard would be perfectly able to serve the purpose of providing data of different topological dimensionality.

Georeferencing

The global position was in this case derived from the queried IFC files and saved implicitly within the newly created ICDD, as the GIS data was coherently georeferenced in the target CRS. This is a possible solution and remains within the ICDD standard.

An optimal standard-compliant solution could imply using ontologies that displays the necessary georeferencing information, independent from the source. One conceivable option would be to use the DCAT to describe polygons in semantic triples. This offers the possibility to integrate information in the WKT format [31]. Including meta data could offer further information about the included documents, with the geographic location being one of them. However, the question arises on whether information, which is usually remaining within the ICDD creator software, such as rules and decisions on how or why something got linked, should be also part of the ICDD standard as it carries implicit (user) knowledge that, without the ICDD creator, cannot be reproduced. It became apparent that many decisions and information *about* payload documents (e. g. their georeferencing) are being made in the *creating software* and remain there, which makes it hard to reproduce information about *why* or *how* links were created, information chosen or even *what* a container encompasses. This means that the container itself is only the result of a decision-making process yet does not necessarily contain sufficient information on how to contextualize it. To use ICDDs for their initial purpose, future work should consider including meta data in a standardized way, containing crucial information *about* the included documents.

Data Quality and Information Requirements

In this case, three solutions were possible to select the data from the asset ICDDs: 1. Take the entire container without selectively considering the content, 2. let the user select the relevant data to be aggregated, and 3. consolidate the date based on pre-defined information requirements. In the demonstrated case, the direct user interaction (2.) replaced the consideration of information requirements that might origin from structured concepts like LOIN (3.). The user here queried GIS services with directly inputted keywords and selected manually relevant data from the identified asset ICDDs. Including information requirements in the process of assembling use-case-specific ICDDs could point towards a full automation of the present methodology, as it could hold requirements towards the questions of which data should eventually be included and even how it shall be linked. Standards like DIN EN 17412 could be stressed more to evaluate possible solutions.

Granularity and Mapping

Mapping GIS feature with BIM objects appears to be meaningful when sharing a similar granularity, e. g. to map a polyline which represents a railway line to its corresponding BIM model. In the case of spatial relations, like defined by DE-9IM the standard [14], linking on file level considering geolocation might be beneficial, as it does not require the clashing Level of Detail between GIS and BIM to be harmonized. In this case, QGIS was used to derive such spatial relations, yet they were stored only implicitly within the ICDD. Future usage of the ICDD could include those relations directly.

Provision of Data

In the case of the demonstrated use case, GIS data was included within the ICDD. A further step towards integration could also mean to externally link GIS services instead of including them in a static form, such as shown in the use case, and link the queries to be performed internally within the ICDD. This would reduce the amount of redundant data.

Acknowledgements

This research was partially funded by grant number 19F2217F through the mFUND initiative of the German Federal Ministry for Digital and Transport (BMDV). The content of this paper, including all statements and findings, is the sole responsibility of the authors and does not represent the views of the funding institution.

References

[1] E. Guyo, T. Hartmann und L. Ungureanu, "Interoperability between BIM and GIS through open data standards: An overview of current literature" in *Proceedings of the 9th Linked Data in Architecture and Construction Workshop*, 2021.

[2] S. Amirebrahimi, A. Rajabifard, P. Mendis und T. D. Ngo, "A BIM-GIS integration method in support of the assessment and 3D visualisation of flood damage to a building", *Trans. in GIS*, Jg. 61, 2016.

[3] S. Vilgertshofer, J. Amann, B. Willenborg, A. Borrmann und T. H. Kolbe, "Linking BIM and GIS Models in Infrastructure by Example of IFC and CityGML" in *ASCE International Workshop on Computing in Civil Engineering 2017*, 2017.

[4] E. Hbeich und A. Roxin, "Linking BIM and GIS Standard Ontologies with Linked Data" in *Proceedings of the 8th Linked Data in Architecture and Construction Workshop*, 2020.

[5] C. Clemen, "Trends in BIM and GIS Standardization: Report from the Joint ISO/TC59/SC13-ISO/TC211 WG: GIS-BIM", *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2022.

[6] F. Noardo *et al.*, "Tools for BIM-GIS Integration (IFC Georeferencing and Conversions): Results from the GeoBIM Benchmark 2019", *IJGI*, 2020.

[7] buildingSMART Australasia, *User Guide for Geo-referencing in IFC: How to Setup Geo-referencing in a Building or Linear Infrastructure Model.*
[8] C. Clemen und H. Görne, "Level of Georeferencing (LoGeoRef) using IFC for BIM", *Journal of Geodesy, Cartography and Cadastre*, 2018.
[9] U. S. Institute of Building Documentation, *USIBD Level of Accuracy (LOA) Specification Guide.*
[10] A. Borrmann, M. König, C. Koch und J. Beetz, Hg., *Building Information Modeling: Technology Foundations and Industry Practice*, 2018.
[11] T. Gilbert *et al.,* "Built environment data standards and their integration: an analysis of IFC, CityGML and LandInfra" in 2020.
[12] *OGC City Geography Markup Language (CityGML) Encoding Standard,* Open Geospatial Consortium, 2021.
[13] U. Maschek, *Sicherung des Schienenverkehrs: Grundlagen und Planung der Leit- und Sicherungstechnik*. Germany: Springer, 2022.
[14] C. Strobl, "Dimensionally Extended Nine-Intersection Model (DE-9IM)" in *Encyclopedia of GIS*, S. Shekhar und H. Xiong, Hg., 2008.
[15] J. Krischler, M. Mellenthin Filardo und C. Koch, "GIS and BIM Integration Approaches for Early Railway Planning Phases" in *Proceedings of the 2022 EC3*, 2022.
[16] K. Jetlund, "Harmonizing and linking conceptual models of geospatial information: Technologies for information modelling in GIS, ITS and BIM" Dissertation, 2021.
[17] M. Daneshfar, T. Hartmann und J. Rabe, "A GIS-based Ontology for Representing the Surrounding Environment of Buildings to Support Building Renovation" in *Proceedings of the 8th Linked Data in Architecture and Construction Workshop*, 2020.
[18] A.-H. Hor, G. Sohn, P. Claudio, M. Jadidi und A. Afnan, "A Semantic Graph Database for BIM-GIS Integrated Information Model for an Intelligent Urban Mobility Web Application", *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2018.
[19] E. P. Karan, J. Irizarry und J. Haymaker, "BIM and GIS Integration and Interoperability Based on Semantic Web Technology", *Journal of Computing in Civil Engineering*, 2016.
[20] M. H. Rasmussen, M. Lefrançois, G. F. Schneider und P. Pauwels, "BOT: The building topology ontology of the W3C linked building data group", *SW*, 2020.
[21] M. Stepien, A. Jodehl, A. Vonthron, M. König und M. Thewes, "An approach for cross-data querying and spatial reasoning of tunnel alignments", *Advanced Engineering Informatics*, 2022.
[22] F. Beck, J. Abualdenien und A. Borrmann, "An evaluation of the strict meaning of owl:sameAs in the field of BIM GIS Integration" in *Proceedings of the 9th Linked Data in Architecture and Construction Workshop*, 2021.
[23] S. van Nederveen, R. Beheshti und P. Willems, "Building Information Modelling in the Netherlands: A Status Report" in *CIB World Congress: Building a Better World*, 2010.
[24] M. Senthilvel, J. Oraskari und J. Beetz, "Common Data Environments for the Information Container for linked Document Delivery" in *Proceedings of the 28th EG-ICE*, 2021.
[25] M. Senthilvel, J. Oraskari und J. Beetz, "Implementing Information Container for linked Document Delivery (ICDD) as a micro-service" in *Proceedings of the 8th Linked Data in Architecture and Construction Workshop*, 2020.
[26] M. Senthilvel und J. Beetz, "Conceptualizing Decentralized Information Containers for Common Data Environments using Linked Data" in *Proceedings of the CIB W78*, 2021.

[27] P. Hagedorn, M. Senthilvel, H. Schevers und L. Verhelst, "Towards usable ICDD containers for ontology-driven data linking and link validation" in *Proceedings of the 11th Linked Data in Architecture and Construction Workshop*, 2023.
[28] C. B. Barber, D. P. Dobkin und H. Huhdanpaa, "The quickhull algorithm for convex hulls", *ACM Trans. Math. Softw.*, 1996.
[29] Z. Yahya, R. W. Rahmat, F. Khalid, A. Rizaan und A. Rizal, "A Concave Hull Based Algorithm for Object Shape Reconstruction", *IJITCS*, 2017.
[30] buildingSMART International, *Standards Server: Version 4.2 bSI Candidate Standard.*
[31] *Data Catalog Vocabulary (DCAT) - Version 3*, 2024.