

Conversion of legacy domain models into ontologies for infrastructure maintenance

Anne Göbels¹[0000-0002-6175-6499] and Jakob Beetz¹[0000-0002-9975-9206]

Design Computation, Faculty of Architecture, RWTH Aachen University,
52062 Aachen, Germany

Abstract. This paper introduces an approach to automatically convert the German data model for documenting infrastructure inspections (ASB-ING) into a semantic ontology. Thus, an extensive data model is created that enables the representation of infrastructure and maintenance-related facts in Linked Data and is compatible with the German standards and guidelines.

The national standard "ASB-ING" contains about 120 classes and 500 attributes to describe a bridge or tunnel and its damages. It is the basis for the relational database program "SIB-Bauwerke", in which most bridges in Germany are recorded over the last 20 years. By converting the ASB-ING model into an Ontology, compatibility with this extensive available data set is ensured, enabling a direct translation of the existing asset data into Linked Data models.

Concurrently, it is possible to relate the classes of this national data model with classes of other national or international ontologies. By mapping the ASB-ING Ontology to ifcOWL, the formerly only tabular data can be linked to standardised geometry models without requiring additionally developed geometric representation. Interlinking the ASB-ING Ontology with other international and national infrastructure data models can improve asset information exchange of cross-border projects.

Keywords: Legacy Infrastructure models · Infrastructure Ontology · Data Conversion · ASB-ING · SIB-Bauwerke .

1 Introduction

Since infrastructure maintenance is expensive and inspections are still predominantly done manually, [3] digital solutions utilising Building Information Modelling (BIM) or Digital Twins (DT) promise to support and improve this process. An essential aspect in this regard is the inclusion of existing legacy data that has been accumulated over long periods of time into information models. This data includes technical information about the structure, planning and building phase documentation, and semi-structured inspection data from the last decades. The acquisition and maintenance of this information are carried out in individual, disconnected, distributed data sets by the respective federal states or local authorities. However, all these heterogeneous data sets need to be taken into account to obtain a solid basis for analysis and necessary maintenance measures.

To achieve comprehensive information models that interlink all these data sets, Linked Data can be utilised. Linked Data is a part of the Semantic Web Stack that has been specifically built to enable interoperability between diverse data sets from different sources. A prerequisite for representing the existing data in Linked Data models are schemas or structured vocabularies, including ontologies, defining the semantic framework for these data sets.

The documentation of infrastructure facility inspections is standardised and has to follow domain-specific data models and conventions. These underlying structures can serve as a foundation for developing ontologies compatible with the existing inspection data. Applying this approach has the advantage of creating extensive ontologies whose base structure is already tested in the field and created by domain experts. Using proven and agreed upon existing vocabularies eliminates the need to develop entirely new data structures and dependencies, which requires much effort and expert knowledge.

In research documented in this paper, we apply this approach to the "ASB-ING" - the German standard data model to describe infrastructural artefacts in many use cases, including the documentation of bridge inspections[5]. It defines the data structure of the software "SIB-Bauwerke", which engineers use to document and store inspection results.

The ASB-ING data model gets directly converted into the ASB-ING Ontology via an automated process. Thus, the ASB-ING Ontology contains the same content as the original data model but provides a Linked Data-compliant format, extending its functionality and scope of application. For example, the ASB-ING Ontology can be linked to other infrastructure ontologies and classifications to create crosscutting data models or to ifcOWL to add geometric entities. Moreover, it becomes possible to represent the existing data sets of SIB-Bauwerke in Linked Data models since the data sets are compliant with the ASB-ING Ontology.

The paper is structured as follows: In section 2, related research activities are described and compared to our approach. Section 3 provides a brief overview of the standard landscape regarding bridge inspections and introduces the "ASB-ING" data model in more detail. Section 4 comprises the main part of this research and focuses on the conversion process of the domain model into a semantic ontology. Moreover, we present briefly how the proposed ontology can be used to transfer the existing inspection data of "SIB-Bauwerke" into a Linked Data Model in a further step. In the concluding chapter 5, we summarise the results of our work and identify unresolved issues and future research approaches.

2 Related Research Activities

Several research and development efforts have recently proposed solution approaches for digital representation and processing of infrastructural assets.

The ongoing IFC Infra initiative [9] within buildingSMART extends the IFC Schema with infrastructure specific entities for bridges, roads, tunnels and rails. It is still under development, and new entities have to pass a long approval

phase before being included in the schema. Since it is an international top-down development process, the IFC Schema cannot include all needed classes or attributes for domain-specific or national data structures. Instead, it will serve as a common superset to allow interoperable implementation into software tools.

Hamdan and Scherer developed the Bridge Ontology "BROT" [7], which is derived from the "Building Topology Ontology" (BOT) [12]. "BROT" provides classes of spatial zones and bridge components. Moreover, the authors create sub-ontologies extending the the core "BROT" ontology with definitions for component details, construction specifications, building materials and structural analysis. The research proposes how the "BROT" ontology can be aligned with ifcOWL [10]. Moreover, a process is drafted to generate an ontological model from an IFC model automatically. The core ontology does not include maintenance-related vocabulary, but it is possible to extend it with external terminologies due to its modular structure.

R. Helmerich developed an approach called "BRONTEX" [8] in cooperation with the Federal Institute for Materials Research and Testing (BAM). It proposes a Linked Data-based platform to store and query information about material sciences and testing related to existing bridges. The ontology "BRONTEX" contains just very generic concepts of structural information but detailed data sets about materials, deterioration processes, defects, and causes. Subclass relations provide a specialisation-generalisation classification of the ontology, but the classes do not have other properties. Because the ontology is developed alongside a use-case scenario of old iron steel bridges, it mainly focuses on information and knowledge about this specific bridge type.

All the approaches mentioned above have in common that they were developed manually and often in the scope of one specific use case. While representing functional structures in their respective use cases and applications, their breadth and depth are often too limited to be useful in other practical applications. Moreover, the previous approaches do not have a solution for how existing data and processes can be aligned with their new models and concepts.

The European CEDR-INTERLINK Project, introduced in a paper from Luiten et al. [11], is a Linked Data-based management system regarding civil infrastructure assets in Europe. Within that project, they develop an international European Object Type Library (OTL), which enables interoperability between the National Road Authorities. The OTL Ontology contains standardised objects types with geometric data and metadata, properties and relations. It is linked to national and international standard data structures for infrastructure management and enables direct information exchange of assets, leading to improved processes of cross-border projects.

Related to this project is the development of the German data model OK-STRAwL [2]. It is also created by converting existing data structures into ontologies. Here the basis is the "Object catalogue for the road and traffic system in Germany" (OKSTRA), provided in UML/XMI, which is translated into a Web Ontology.

The INTERLINK project also includes test use cases to demonstrate and test typical data transfer processes in national environments. One of these test projects is the “Handover of BIM Linked Data”. It is located in Germany and uses BIM processes during the design phase of a bridge and for the data handover after the construction phase. For this purpose, two test ontologies of the ASB-ING model are developed to interlink asset entities defined in OTL with national classifications [1], neither ontology contains the whole data structure but only focuses on one aspect. The “ASB-ING Condition Ontology” contains classes to describe the condition of bridges, following the terms defined in the ASB-ING. The “ASB-ING Building Element Ontology” provides the building elements of the ASB-ING standard as ontological classes. In the next step, they are mapped to ifcOWL classes via shape definitions.

In the context of these previous developments, the automatic conversion of the complete ASB-ING data model has several improvements and could be beneficial for projects like CEDR- INTERLINK.

Since the translation process requires a minimum of manual input, we do not have to focus on specific parts of the data model but can convert the whole structure at once. Thus, an extensive ontology is created, including a hierarchical structure and logical dependencies and relations, to describe structural and maintenance related facts about bridges or tunnels. Moreover, an automatic translation of existing inspection data into Linked Data Models is possible due to the compatibility with SIB-Bauwerke, which stores the maintenance data for most infrastructure artefacts in Germany for more than 20 years.

3 ASB-ING data model

3.1 Context and field of application

The “Anweisung Straßeninformationsbank, Teilsystem Bauwerksdaten“ (ASB-ING) [engl.Instruction for Road Information Database, subsystem structural data] [5] is Germany’s authoritative data model for bridge and tunnel maintenance documentation. It is the basis for a database application called “Straßeninformationsbank – Bauwerke” (SIB-BW) [engl.: Roaddatabase – Structures] [13] used for the inspection documentation of infrastructure assets of the German federal government.

The structure and content of ASB-ING follow the specifications made in the “Guideline for the uniform collection, assessment, recording and evaluation of results of structural tests according to DIN 1076” (RI-EBW-PRÜF) [4]. The mentioned DIN 1076 [6] defines the monitoring and inspection processes and rules for infrastructure facilities in Germany.

The ASB-ING is updated periodically to reflect new developments in the field of infrastructure and maintenance. Currently, a new version of the ASB-ING standard is under development, modelled in UML. It is a hierarchical data structure containing about 120 classes, more than 500 attributes and 3000 predefined value assignments, all related by restricted relations representing mandatory

rules and standards. We have been given access to this new model for research purposes, but it has not yet been published. If we mention the ASB-ING model in the following sections, we refer to the new version from 2020.

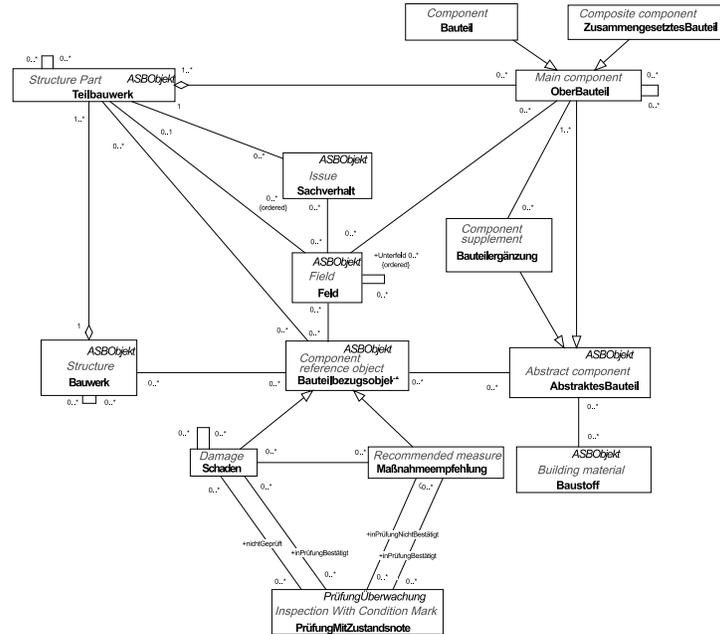


Fig. 1. Overview diagram of the ASB-ING model (Translations of main terms in grey) [Source: Landesamt für Straßenbau und Verkehr M-V]

3.2 Structure

The ASB-ING data model covers general information, like the metadata or administrative data and structure related aspects, containing the building elements and their type and material specification. The building elements are related by aggregations to represent the modular building structure. Additionally, the ASB-ING model defined classes to document the monitoring processes and the resulting structure assessments with determined damages and related (recommended) measures (see figure 1). The classes of the different domains are related by restricted associations to display the dependencies and logical or rule-based constraints (see diagram 1 and 2).

Each class can have several attributes. For attributes concerning numeric information, like geometry or quantity data, the values have defined data types but can be filled individually (e.g. "Length" must be given in metres, the value can be any decimal number). There are also "union data types" that combine

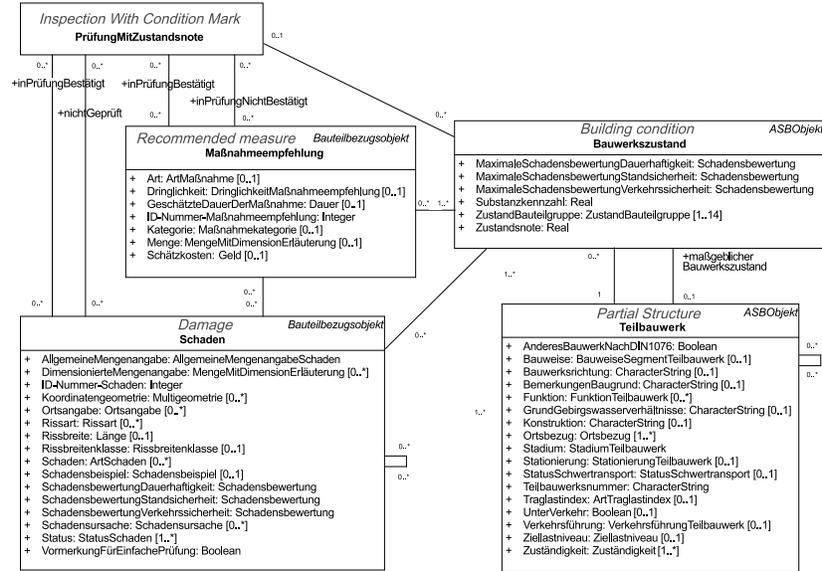


Fig. 2. Inspection and Maintenance diagram of the ASB-ING model (Translations of main terms in grey) [Source: Landesamt für Straßenbau und Verkehr M-V]

various simple attributes. For example, the data type "duration" consists of several individual attributes like "Years", "Months", "Days", and so on.

For attributes describing semantic information, the value assignment is managed via predefined terms stored in "key table"/dictionary classes. Such dictionaries are similar to the standardise PSets in IFC, the ADEs in CityGML or the dictionaries in OKSTRA. In this way, a standardised description and definition of entities are ensured. Each "key table" class covers one topic, e.g. "damage type" or "building element type". Each topic has specifications, structured hierarchically and stored with an ordering number (the key). For example, the "damage type"-key table has the specification "4.2 – Surface damage metal" and "4.2.1.4 – Rusted". The ASB-ING model has about 240 of those "dictionaries" and more than 3000 defined terms.

The ASB-ING is a detailed data model covering a significant proportion of entities needed for a comprehensive digital information model of bridges and tunnels and their maintenance data. Since relations and dependencies are modelled, rule-based queries can be derived from them. Due to its UML format, the content is easy to access for further processing.

For general purposes, not related to German inspection processes, the data structure could be too complex and should be broken down to its elementary structure. The value-assignment is not straightforward, but the standardised terms could be an entry point for automatic processing. What is missing are geometric model entities representing their semantic counterpart.

4 Data Conversion

4.1 Conversion of the ASB-ING data model into a semantic Ontology

To develop the ASB-ING Ontology, we use the Web Ontology Language OWL. Since UML, in which the ASB-ING model is provided, and OWL are similarly structured, the conversion process is uncomplicated. To implement the ontology, we used the following name spaces:

- asb: <https://w3id.org/asbingowl/core#>
- asbkey: <https://w3id.org/asbingowl/keys#>
- asbkey13: <https://w3id.org/asbingowl/keys/2013#>
- owl: <http://www.w3.org/2002/07/owl#>
- rdfs: <http://www.w3.org/2000/01/rdf-schema#>
- xsd: <http://www.w3.org/2001/XMLSchema#>

Classes and Properties The conversion of the UML entities follows the comparison displayed in table 1. The ontology classes get relation definitions to other classes if required and additional information like a label and an explanatory comment if available.

The value restriction of an attribute, defined in the ASB-ING model, determines if this attribute is converted into a OWL Datatype or Object Property of the ontology. Attributes having textual or numeric values are defined as Datatype Properties. Their datatype classes from the UML model (e.g. "Character string" or "Length") are additionally mapped to a suitable XSD datatype. Attributes pointing to "key table"- or union datatype classes are implemented as Object Properties in the ontology since they refer to other classes. All properties have a defined domain and range, and if available, a label, a comment or a sub-property relation.

The diagrams in figure 3 show a part of the initial UML model compared to the ontology. As can be seen, the general structure and content of the domain model are not changed but merely transferred to the respective OWL types. A current version of the complete ontology can be found via its URI an at GitHub ¹.

Table 1. Comparison of the entities in the ASB-ING UML model and in the ontology.

UML	Ontology
Class	owl:Class
Attribute	owl:DatatypeProperty / owl:ObjectProperty
Generalisation	rdfs:subClassOf
Aggregation	asb:isPartOf
Association	asb:associatedWith

¹ https://github.com/AnneGoebels/ASB-ING_Ontology

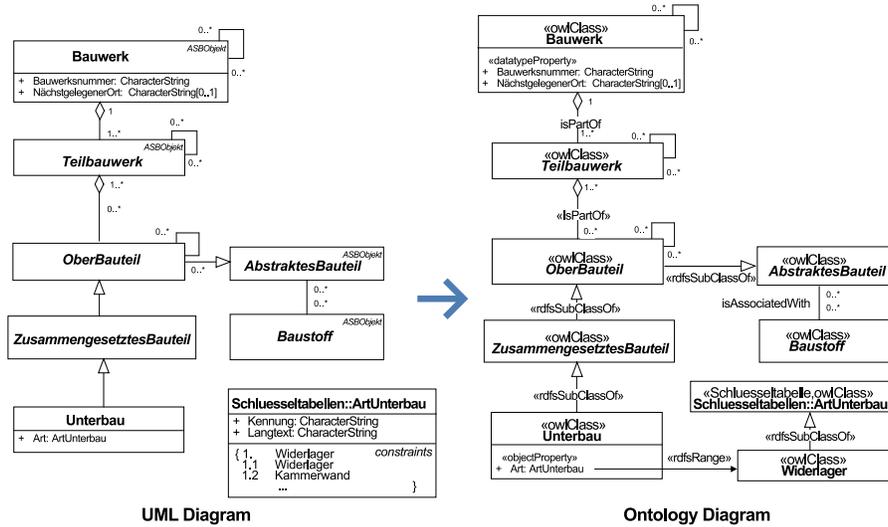


Fig. 3. Comparison of the ASB-ING as UML model and as an Ontology

Key tables The translation of the key tables requires some more attention. Each key table contains a list of key-value pairs. The original concept of the ASB-ING model is to access them via the attributes "Text" and "Key" of the key class. However, we decided to transform each key-value pair into an individual class. A class is more expressive as a simple textual value, facilitates the value-assignment, can have several semantic annotations, and is easier to access via queries. Each "key-value-pair" class is named by combining the individual value name and its super-class (dictionary) name. The order number is stored using the attribute "Kennung". As the value-key pairs are structured hierarchically, the corresponding classes are related to each other via sub-class relations. Additionally, all "key-value-pair"-classes are sub-classes of their key table class. Figure 4 displays the key tables in the ASB-ING UML model and their conversion to OWL classes.

Since the value-key pairs are initially modelled as value options, their semantic meaning is different from the main classes of the ontology. Thus, we decided to create a sub-ontology/vocabulary for these classes. It uses a distinct namespace (with the prefix "asbkey") to distinguish between these different kinds of classes.

As mentioned in section 3.1, we work with the new, unpublished version of the ASB-ING model. Nevertheless, the current documentation of bridges and tunnels is based on the previous version, which also has key tables with about 20.000 key-value pairs. As the mapping of the old keys to the new keys is not yet finished, we decided to transform also the "old" key-value pairs into OWL classes to enable the processing of existing data without the mapping. These "old" key-value-pair classes form another sub-ontology/dictionary, having the

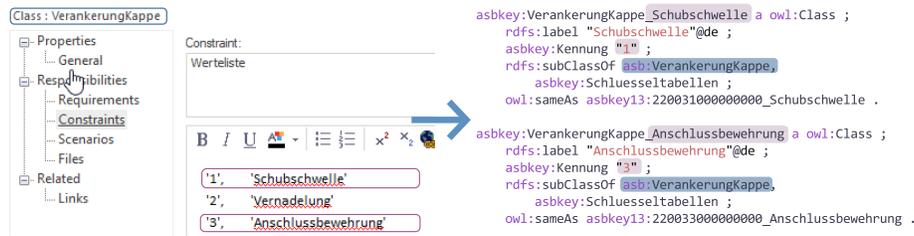


Fig. 4. Conversion Concept of the key classes (left: Screenshot of key class constraints of the UML model, right: converted OWL key classes (formatted in turtle))

transitional prefix “asbkey13”. Figure 5 shows that each old key-value-pair class is named by combining the key number and its text. The hierarchical structure is represented via sub-class relations, and an ”rdfs comment” alternatively presents all superordinate terms to ensure a quick human interpretation of the data.

The two key-class sub-ontologies are not entirely independent modules but rather a structured collection/vocabulary of all classes that can be used as the value of an Object Property of the core ASB-ING Ontology.

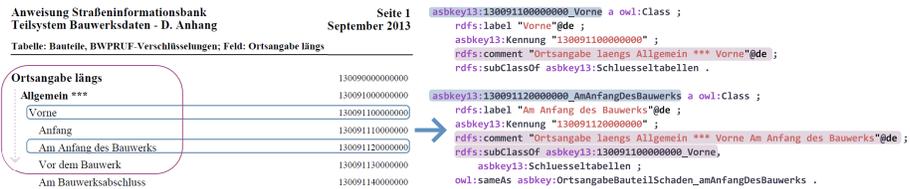


Fig. 5. Conversion Concept of the old key classes (ASB-ING 2013) (left: Screenshot of a key table PDF, right: converted OWL key classes (formatted in turtle))

Implementation Contrary to the previous ASB-ING model or other domain standards, the new ASB-ING model is provided in UML. This feature is what enables the implementation of the conversion approach in the first place.

In essence, the UML model is exported as an XMI file, from which the classes, attributes, value data types and relations are retrieved. They are then defined as the respective ontological entities. The old keys are provided as an Excel sheet, from which each key (= each row) is fetched and defined as an OWL class. The complete code for automatically converting the UML model into a semantic ontology and the codes for creating the sub-ontologies are on GitHub ².

² https://github.com/AnneGoebels/ASB-ING_Ontology

4.2 Concept for converting existing inspection data into Linked Data models

Inspection data of most bridges and tunnels in Germany is stored in SIB-Bauwerke, which is based on the ASB-ING data model. Currently, reading, writing, and processing this data is only possible within this commercial software. Exporting data from SIB-Bauwerke delivers a zip file containing database files and a folder with images and 2D plans. Each database file contains one table covering one main topic (e.g. “building materials” or “damages”). The data of one bridge is divided into 61 tables, having altogether about 1500 columns.

Manual inspection or even further processing of the data in this format is complicated and challenging. The column names are illegible, and key table values are only represented via their key number. But using the ASB-ING Ontology, including the key-table sub-ontologies and a mapping table, these extensive data sets are processed and converted into RDF within minutes. Thus the data is accessible for Linked-Data methods and interpretable for external applications. The detailed process will be presented in a forthcoming paper, but a brief description of the process is provided to demonstrate the application of the ontology. Figure 6 displays a graphical overview of it.

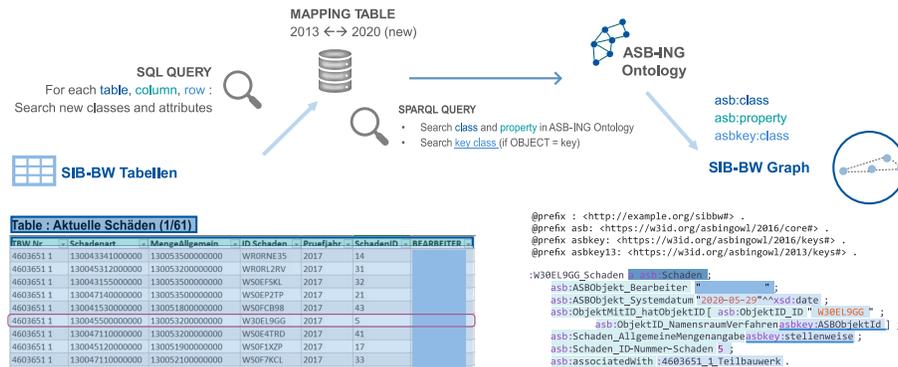


Fig. 6. Conversion process of database tables from SIB-Bauwerke to an RDF Graph [Source Table: ZPP Ingenieure AG]

Broadly defined, the tables are the classes of the ontology, and their columns are the properties. The cells contain the actual values as textual or numerical input or as an encrypted key number. For the latter, the respective key class is searched in the sub-ontologies and used instead of the number.

Two mapping processes are required as those tables are structured based on the previous version of the ASB-ING. First, the old table and column names are mapped to their respective counterpart of the new ASB-ING model. We got access to this predefined mapping table from the responsible authority. In the next

step, we extend this mapping table with the new developed Ontology classes and properties. Finally, we use this table as a reference to directly get the ontological entity of the tables and columns of SIB-Bauwerke. In this way, each data input from the inspection documentation gets processed and transformed into its Linked Data representation. Altogether, they form an RDF-Graph, representing the maintenance information of one infrastructure facility.

This graph serves as a starting point for further processing of the information. For example, the entities from the inspection data can be linked to external geometric representations, technical drawings or related pictures. One issue that can benefit significantly from interlinked Data is the localisation of damage, which is currently only done via textual description. By a visual representation of damage in 2D plans or 3D models, the inspection process can be made more efficient and spatial analysis of weak areas can be performed.

5 Conclusion and Future Work

The presented approach succeeds in translating the existing data model for infrastructure inspections into an ontology. Thus, with a minimum of manual data creation, an extensive ontology is derived mostly automated, corresponding to the established data structures used in practice. The ontology enables converting existing maintenance data into Linked Data Models, which are the basis for additional use cases requiring interoperable digital processing and analysis such as structural health monitoring, predictive maintenance etc.

The next step is to link and map the ASB-ING Ontology to other domain-relevant ontologies and structures, like OTL, BOT, BROT, ifcOWL etc. This way, infrastructure and inspection information captured with the ASB-ING Ontology can be enriched with external data and used in other contexts and applications.

For example, links between the converted inspection data graph (as outlined in section 4.2) and an IFC model by mapping the structural component classes of the ASB-ING Ontology with ifcOWL classes enables a mutual enrichment of both. The inspection data gets visual representation and geometric entities, and the IFC model is semantically enriched with infrastructure-specific classes without an additional extension of the IFC schema.

The inclusion of the ASB-ING Ontology into the OTL framework of the INTERLINK Project [11] can add an essential national data structure and will increase international information exchange. Due to the possibility of a mostly effortless conversion of existing asset data into Linked Data Models, the practical applicability and implementation of the INTERLINK approach in Germany is significantly simplified. All bridges and tunnels documented in SIB-Bauwerke can be converted following the same automatic process and thus be a significant contribution to the digital infrastructure ecosystem of Europe.

However, since the ASB-ING Ontology is quite complex, further investigations are necessary to determine the relevant subsets of classes and properties needed for alignments or mappings to the respective external data structures.

6 Acknowledgements

This research is part of the "TwinGen" project funded by the Federal Ministry of Transport and digital Infrastructure (BMVI). We would like to thank the Landesamt für Straßenbau und Verkehr Mecklenburg-Vorpommern for supporting us with the new ASB-ING model, as well as the Regierungspräsidium Tübingen for giving us access to the mapping concepts and tables.

References

1. ASB-ING Test Ontologies, <https://roadotl.eu/static/eurotl-ontologies/testontologies/index.html>. Last accessed 17 Apr 2021
2. Beetz, J., Amann, J., Borrmann, A.: Analyse von Einsatzmöglichkeiten von verbundenen Informationen (Linked Data) und Ontologien und damit befassten Technologien (Semantic Web) im Bereich des Straßenwesens. (2018) https://www.okstra.de/forschung/linked-data_DE.html. Last accessed 13 Apr 2021
3. Borrmann, A., Fischer, O., Dori, G., and Wild, M.: Intelligente Brücke - Konzeption eines modular aufgebauten Brückenmodells und Systemanalyse. Bundesanstalt für Straßenwesen, Bergisch Gladbach (2014)
4. Bundesministerium für Verkehr und digitale Infrastruktur : RI-EBW-PRÜF Richtlinie zur einheitlichen Erfassung, Bewertung, Aufzeichnung und Auswertung von Ergebnissen der Bauwerksprüfungen nach DIN 1076. Bundesanstalt für Straßenwesen, Bergisch Gladbach (2017).
5. Bundesministerium für Verkehr, Bau und Stadtentwicklung, Abteilung Straßenbau. : ASB-ING 2013 Anweisung Straßeninformationsbank Segment Bauwerksdaten. Bundesanstalt für Straßenwesen, Bergisch Gladbach (2013).
6. DIN Deutsches Institut für Normung e.V., 1999. DIN 1076 - Ingenieurbauwerke im Zuge von Straßen und Wegen Überwachung und Prüfung. Berlin: Beuth Verlag GmbH, pp.1–9.
7. Hamdan, A., Scherer, R.: Integration of BIM-related bridge information in an ontological knowledgebase. In: Proceedings of the 8th Linked Data in Architecture and Construction Workshop - (LDAC), pp. 77–90. CEUR-WS, (2020)
8. Helmerich, R.: "Knowledge representation system about existing bridges In: Bridge Maintenance, Safety, Management, Resilience and Sustainability, Biondini and Frangopol, pp. 245–252. Eds. Taylor and Francis Group, London (2012)
9. IFC Infra Website, <https://ifcinfra.de/>. Last accessed 13 Apr 2021
10. ifcOWL, <https://technical.buildingsmart.org/standards/ifc/ifc-formats/ifcowl/>. Last accessed 13 Apr 2021
11. Luiten, B., Böhms, M., Alsem, D., O’Keeffe, A.: Asset information management using Linked Data for the life-cycle of Roads. In: Life Cycle Analysis and Assessment in Civil Engineering: Towards an Integrated Vision: Proceedings of the Sixth International Symposium on Life-Cycle Civil Engineering (IALCCE 2018), CRC Press, Ghent (2018). <https://doi.org/10.1201/9781315228914>
12. Rasmussen, M., Pauwels, P., Hviid, C., Karlshøj, J.: Proposing a Central AEC Ontology That Allows for Domain Specific Extensions. In: Lean and Computing in Construction Congress - Volume 1: Proceedings of the Joint Conference on Computing in Construction, pp. 237–244. Heriot-Watt University, Heraklion (2017). <https://doi.org/10.24928/JC3-2017/0153>
13. WPM-Ingenieure GmbH, 2020. SIB-BAUWERKE. Neunkirchen-Heinitz: Bundesanstalt für Straßenwesen (BASt).