

TRANSFER OF IMPLICIT SEMI-FORMAL TEXTUAL LOCATION DESCRIPTIONS IN THREE-DIMENSIONAL MODEL CONTEXTS

Anne Göbels, Fabian Rivadeneyra, and Jakob Beetz

Chair of Design Computation, RWTH Aachen University, Germany
goebels@dc.rwth-aachen.de

Abstract

To create comprehensive digital twins for maintaining existing bridges, it is essential to link newly acquired model data with legacy inspection information. However, the current bridge inspection documentation in Germany is only text-based. Thus, the locations of damages and components are defined in implicit, semi-formal textual descriptions, which depend on the viewing direction of the inspecting engineer. This paper presents an approach to automatically transfer these location descriptions into explicit three-dimensional model elements and conceptual damage representations. The approach works with rule-based transformations encoded in dedicated algorithms and achieves unambiguous model-based localisation and geometrical representation of the documented inspection data. Concurrently, the process links legacy information with new model data, contributing to an efficient implementation of digital twins.

Introduction

The maintenance of bridges is essential to preserve a working infrastructure. However, most bridge maintenance systems (BMSs) are based on textual documentation of damage detected during a manual inspection on-site. This maintenance approach requires high staffing expenditure and disregards state-of-the-art capabilities for recording, managing, and analysing data with digital solutions. In Germany, 45% of all bridges were built before 1980 and calculated for a lower traffic load than required today (Bundesanstalt für Straßenwesen, 2021). As this stock of bridges (will) need increased maintenance effort, there is a demand to modernise and optimise the existing maintenance processes.

Thus, Building Information Modelling (BIM) and Digital Twin (DT) approaches are being explored to enable improved, predictive maintenance using efficient data acquisition methods, model-based information management systems, and computer-supported analysis (Saback et al., 2022). This work is part of the research project *Twins-Gen*, funded by the German Ministry for Digital and Transport, which attempts to generate DTs of existing bridges for maintenance using automated methods.

The aim is to create processes that only need low manual input/acquisition work and can be applied to many bridges with manageable resources. Besides creating 3D bridge models, an essential part of the project is integrating and processing existing BMS data. This data is crucial

to develop DTs that can access the fundamental construction data, the maintenance history, and the bridge's current condition.

The German BMS is *SIB-Bauwerke*¹, a proprietary database application that determines an overall bridge condition grade based on manually entered damage documentation. It works text-based and offers no geometrical representation of the bridge (3D model, point cloud, etc.). Thus, the assignment of construction information to individual bridge components, and the localisation of damage at specific areas, is expressed by textual descriptions. These descriptions can be selected from a set of predefined natural language-based terms (e.g., front abutment, at the end of the superstructure, xx. beam from the left, right side panel, etc.) that are easy to use by humans but depend on the viewing direction and subjective perception of the inspector.

Thus, the current procedure leads to implicit and partly ambiguous assignment and localisation of information, which impedes accurate / neutral communication about damage/bridge components and can cause failures in the maintenance process, e.g. if the subsequent inspector cannot find a described damage location. Moreover, the textual expressions are not machine-readable/interpretable, preventing computer-aided (geometrical) queries or analysis.

This paper demonstrates how the implicit textual location descriptions can be automatically transferred to three-dimensional (IFC) model contexts to identify explicit model elements and zones. The localisation approach works with rule-based transformations based on expert rules encoded in dedicated algorithms. The process results are the successful determination of a model element and the identification and creation of damaged areas and their 3D representation based on the textual description. In addition, the process stores the correlations of the BMS and model data sets as links in the Linked Data format.

In contrast to stochastic machine learning (ML) methods, this approach can be applied to single data sets. In our use case example (a straight, multi-span girder bridge with 40 damages), the process achieved to transfer 80% of the textual location descriptions. Since 76% of the 56,000 bridges (state: 2020) stored in *SIB-Bauwerke* have a similar structure to our use case example, the approach can potentially be applied to about 43,000 bridges. The re-

¹<https://sib-bauwerke.de/>

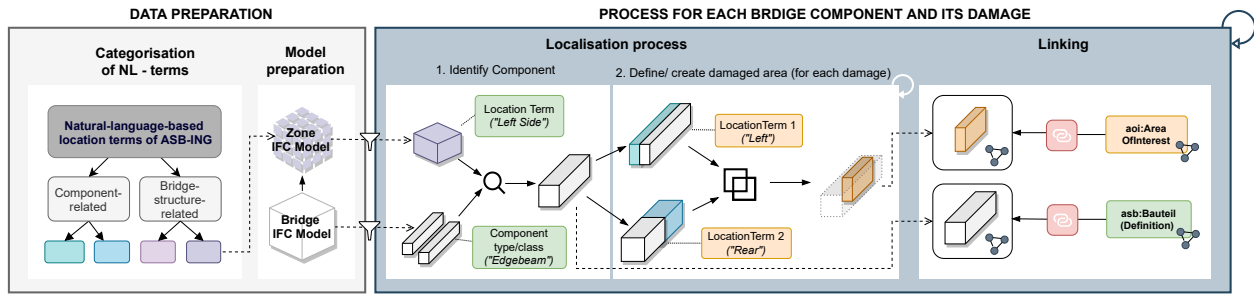


Figure 1: Schematic illustration of the transfer process.

sulting annotated/interlinked datasets could then be used as training data for advanced ML methods for processing natural-language location descriptions in the context of 3D models/DTs.

Related Research

Improving bridge maintenance processes and technologies is subject to many research works. On a standardisation level, the IFC Bridge extension gives a basis to handle bridge models properly in IFC, enabling interoperable BIM Models for bridge maintenance (Borrmann et al., 2019).

Artus and Koch (2022) proposed a method for detailed modelling of damages in IFC to enable an object-oriented BIM-supported inspection process and damage analysis. However, processing and locating existing damage data in an IFC model is not in the scope of the approach. The research of Liu et al. (2022) addresses the integration of IFC-based inspection data into the existing BMS, but their focus is on the compatibility of newly acquired data with the BMS. They do not discuss how the existing data can be accessed and located in the opposite direction in the IFC model. Nevertheless, the approach could be well integrated into the later use of the DT of the TwinGen project. In the field of semantic web technologies, several ontologies were developed for representing (bridge) models and maintenance data in Linked Data. The ifcOWL Ontology enables the representation of IFC models in the Resource Description Format² (RDF) (Beetz et al., 2009). Al-Hakam et al. developed the Damage Topology Ontology (DOT) and the Area of Interest Ontology (AOI) (Al-Hakam et al., 2019; Al-Hakam and Scherer, 2020). The DOT Ontology allows defining and linking damage (patterns) to components, and the AOI Ontology can further specify the damaged area on a component.

Thus, there are many developments to support future inspections and enable improved data management using BIM or Linked Data approaches. However, the automatic integration of existing BMS data to the new proposed models/methods, instead of a (manual) re-recording, is not addressed. Therefore, the approach presented in this paper can fill a significant gap by showing how the existing data can be integrated into new digital solutions and benefit

from their advantages.

Method

The presented method aims to automatically transfer textual location descriptions of legacy BMS data to three-dimensional model contexts. The approach enables component information from the BMS to be assigned to corresponding model elements and damage information to be localised and visualised in a model.

Figure 1 illustrates a brief overview of the transfer process. In a preliminary step, we analysed the predefined natural language terms of SIB-Bauwerke used to describe the location of bridge data. Further, the IFC model was prepared for the localisation process. In the main part of the method, all component and damage location descriptions of a bridge are processed, identifying specific model components and areas. In the last step, the IFC elements found or created in the process are linked to the BMS data through Linked Data.

As the resources and preconditions for this work resulted from previous work steps within the TwinGen project, these are briefly described in the following subsection. Next, the sub-tasks of the method are explained and finally demonstrated in a use case.

Previous Work

Within the TwinGen project, processes are developed to create DTs automatically based on point clouds and existing data sets (construction plans, inspection information etc.). One central aspect is automatically creating a 3D model derived from a point cloud. Another focus is on the semantic interlinking of the created and existing data sources to implement a comprehensive DT model. Within this work step, special attention is given to integrating existing BMS data sets, as they represent a rich data source for the semantic enhancement of the newly created, solely geometric model.

In the German BMS *SIB-Bauwerke*, each database table represents a different aspect of the bridge (components, material, management, etc.) or the inspection process (damages, measures, inspection metadata, etc.). Besides a few exceptions (personal data, free text field, comments), each input into the database (construction documentation, inspection results, etc.) is made using pre-

²<https://www.w3.org/TR/rdf-concepts/>

Table 1: Extract of a converted damage documentation of SIB-Bauwerke

Property	Value
rdf:type	asb:Schaden (Damage)
asb:PruefungUeberwachung_Pruefjahr (Inspection year)	2019
asb:Schaden_Schaden-ID-Nummer (Damage ID)	21
asb:Schaden_Schaden (Damage type)	asbkey:ArtSchaden_Wasserschaden (Water damage)
asb:ASBING13_Bauteil (Component)	asbkey13:130011910000000_Widerlager (Abutment)
asb:ASBING13_Bauteilergaenzung (Supplement)	asbkey13:130031100000000_Beton (Concrete)
asb:ASBING13_Bauteilgruppe (Component group)	asbkey13:390021200000000_BauteilgruppeUnterbau (component group: Substructure)
asb:Schaden_Ortsangabe (Location description)	asbkey:Ortsangabe_WiderlagerVorn (Abument, Front) asbkey13:130115000000000_Unten (Bottom) asbkey13:130101100000000_Links (Left)

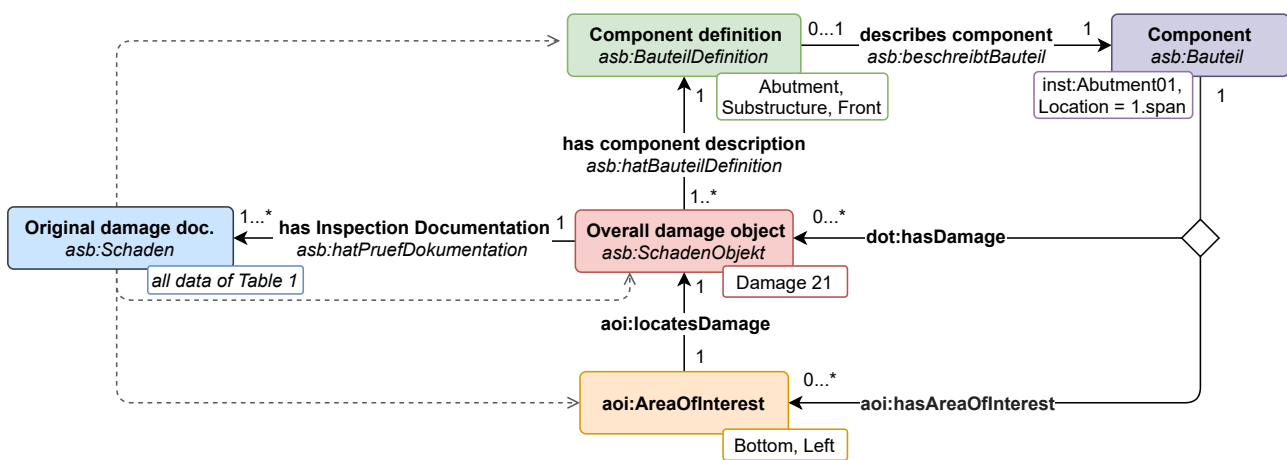


Figure 2: Newly implemented object-oriented structure of the converted SIB-Bauwerke RDF Graph.

defined terms/text blocks to ensure uniform expressions. The database structure and the terms are defined in the German *Instruction for the Road Information Database – Subsystem construction data (ASB-ING)*³ (Bundesministerium für Verkehr, Bau und Stadtentwicklung, Abteilung Straßenbau., 2013). The ASB-ING data model contains 120 classes, 80 data types, 500 attributes and 3000 predefined terms/ text blocks to document bridge (inspection) data (Göbels and Beetz, 2021).

To integrate the legacy BMS data into the DT, the data was extracted from the proprietary, closed database application. As Linked Data is used in the TwinGen approach to interlink the heterogeneous data sets, we also converted the exported BMS database tables into the Linked Data format (Resource Description Format (RDF) graphs).

The first step of the "SIB-Bauwerke to RDF" conversion was to transform the underlying ASB-ING structure, including the predefined terms, into an ontology (ASB-ING Ontology⁴) (Göbels and Beetz, 2021). Based on that, an automatic process was implemented that transfers the data of the extracted database tables into RDF graphs (Göbels,

³Original title: "Anweisung Straßeninformationsbank, Teilsystem Bauwerksdaten"

⁴prefix asb: www.w3id.org/asbingowl/core

2021).

However, the conversion only changed the format of the BMS data but not the inner structure and quality. The information assignment to components and the damage locations were still textual expressions. Table 1 shows a simplified extract of a converted damage entry. Rows five to eih describe the affected component, and rows 9 and 10 represent the damage location, respectively, the damaged area of the element. The individual bridge components are also stored in SIB-Bauwerke but are not related to the damage entries. Thus, in the RDF graph, they are separate RDF subjects and not linked to the damages.

To prepare the BMS data for the transfer process to the 3D model, the content of the generated RDF graphs was internally transformed into an object-oriented structure. The properties of the original damage documentation describing the component are aggregated in a *component definition* object. The properties defining the damage location/area are stored as an *area of interest* object. The information of the *component definition* object is used to identify the separately documented component. In brief, the damage and components previously recorded independently are now explicitly interlinked. Thus, it can be easily identified which individual bridge element has which dam-

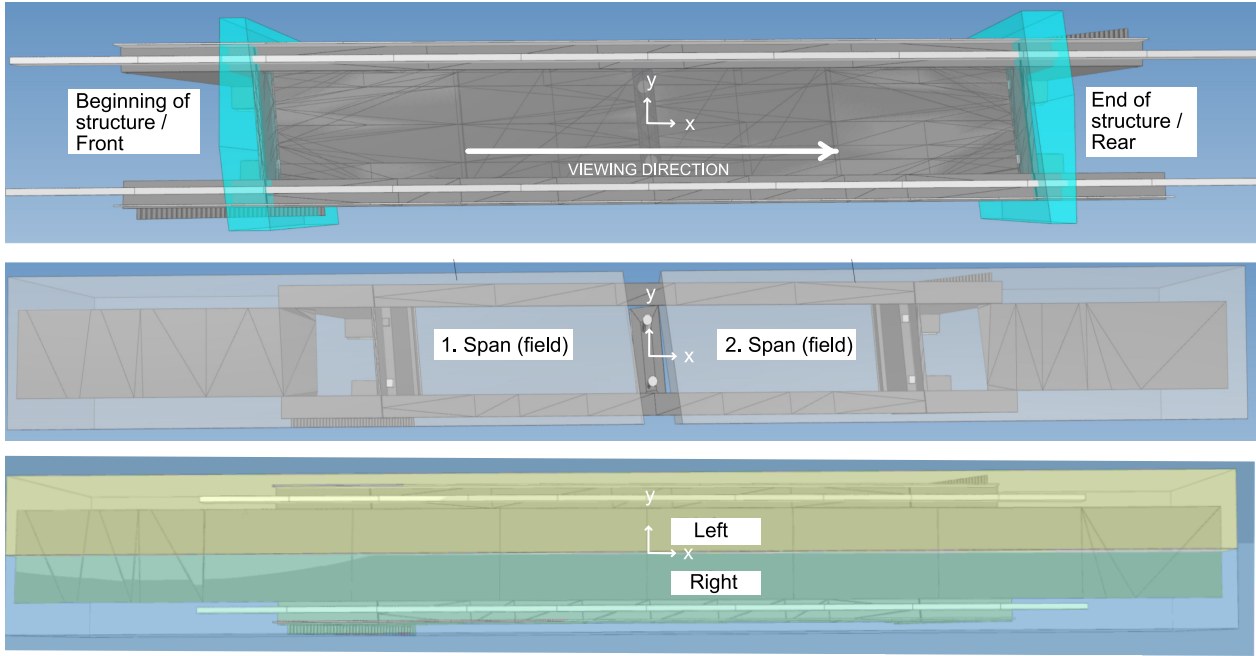


Figure 3: Creation of bridge zones according to the bridge-structure-related location terms

age(s). Figure 2 shows the implemented object-oriented structure, which is compatible with the 3D model, enabling easier processing and linking in the following work.

Analysis of Natural-Language-Based Location Terms

The first step of the transfer process was the analysis of the predefined location terms according to ASB-ING. We identified three main characteristics of the terms: They use two different ways of referencing the context, each location can be described by a variety of interchangeable terms (e.g. "General right" and "Right side"), and about 60% (129 of 206) of terms are dependent on the viewing direction.

The two context-referencing categories are *component-related* and *bridge-structure-related*. Component-related terms refer to a logical subpart of an individual element (top, bottom, right, left, front, rear, etc.) and are used for damage localisation. Bridge-structure-related terms refer to an area of the bridge structure in its entirety (e.g., beginning/end of the bridge structure, 2. span, right/left side of the bridge, area of pillars, 3. component from left, etc.). They are used for both component and damage localisation.

The terms within the categories are further aggregated to limit the variety of expressions and facilitate processing. Component-related terms are assigned to the internal bridge axes, where X is the primary (longitudinal) axis, Y is the secondary axis, and Z is the vertical axis. For example, "General Right" and "Right Side" are assigned to the category *Y-Right*; "Upper Side" and "On the top" are assigned to the category *Z-Top*. Bridge-structure-related terms are summarised into zone categories which are most used in the sample BMS datasets: *Zone-Beginning/End of bridge*, *Zone-Right/Left side of bridge*, and *Zone-x. span*.

Model Preparation

Regarding the model data, further requirements and preparations have to be addressed to implement the transfer process. As the process works on the object level, a prerequisite is a mapping table between the terms/classes of components according to the ASB-ING (Ontology) and the IFC (Bridge) schema. The mapping table can be found in the GitHub repository of the implementation process (see Section "Use Case Example").

Viewing Direction Alignment

Another preliminary task is the transfer of the recorded viewing direction (e.g., "From West to East" or "From City A to B") of the BMS data to the internal model coordinate system to interpret the viewing-dependent location expression (left, right, front, etc.) correctly. In this prototypical approach, we adjusted the provided model manually inside a modelling software, following two requirements for easy automatic processing: the primary axis of the bridge must be parallel to the internal model coordinate system; the internal model coordinate system must correspond with the externally defined viewing direction. In the presented workflow, the x-axis was defined as the bridge's primary/longitudinal axis, and the origin of the internal coordinate system was set at the horizontal centre of the bridge. Thus, negative x-values correspond to "front", and positive x-values to "rear" (viewing direction goes from "front" to "rear"). Negative y-values are interpreted as "right", and positive y-values as "left" (see. Fig. 3).

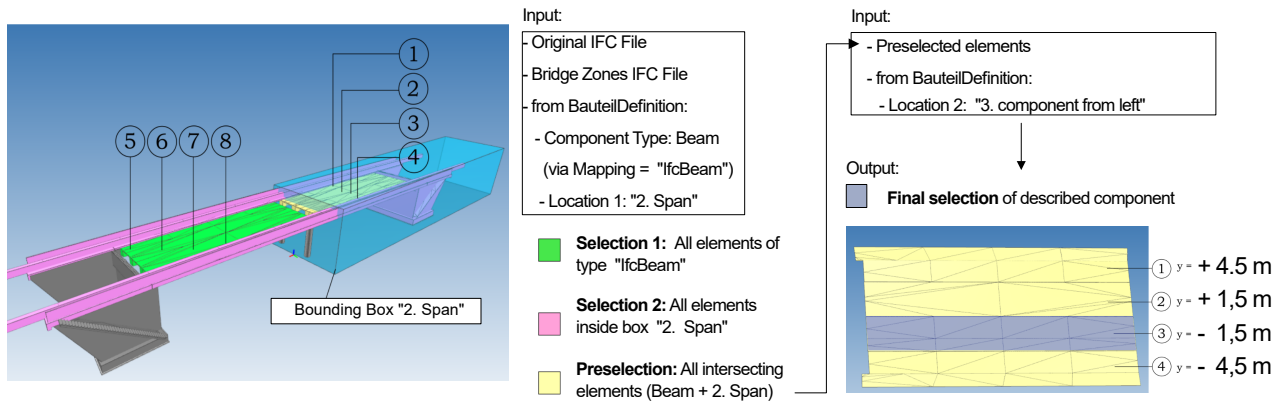


Figure 4: Localisation process of a component (3rd Beam from left, at the 2. Span).

Creation of Zone Model

To transfer the bridge-structure-related location terms, the corresponding zones inside the IFC model must be defined. Therefore, the aligned IFC model is automatically divided into the previously identified zone categories (e.g. "Beginning", "1. span", etc.) (see. Fig. 3). According to ASB-ING, a bridge's beginning and end are marked by the two roadway junctions, and the bridge's spans are defined by the abutments and piers. Thus, the IFC model is filtered for these elements using the created ASB-ING to IFC mapping. The beginning and end zones are bounding boxes with an offset to each side of the road junction. The offset can be set according to the dimension of the bridge. The bridge span zone is represented by a bounding box extending from the front to the rear abutment. If the bridge has more than one span, the position of the piers is used to divide the bounding box into separate spans, named in ascending order from the beginning of the bridge. The zones *Right side of the bridge* and *Left side of the bridge* are created by splitting a bounding box along the primary axis into halves, which encompasses the whole bridge. All zone-representing bounding boxes are named according to their previously defined category names and stored as *IfcBuildingElementProxy* elements in a separate IFC model, which will be referred to in the upcoming process as the *zone model*.

Localisation Process

The localisation process works component-based; thus, first, each bridge component of the BMS data graph is mapped to its corresponding model element based on the textual location description and class assignment. If the component is damaged, the location/area for each of its damages is identified, and geometric representations of them are created.

Locating Components

The components' class and location information is stored in the associated *component definition* (see Figure 2). Using the ASB-ING to IFC mapping, all IFC elements according to the component class are preselected from the model (e.g. ASB-ING: Longitudinal beam = *IfcBeam*).

Then, this preselection gets intersected with the bounding box representing the (first) location term (e.g., "2. Span"), leading to a reduced preselection of elements. If a second location term exists (e.g. "right side), this process is repeated with the corresponding zone and the previously determined preselection.

The aim of the process is to determine one specific component by applying the bounding boxes as spatial filters to the preselection based on the class type.

Figure 4 shows an example with a component of the class "Longitudinal beam", with the location description "2. Span, 3. Component from left". This example also shows another type of localisation based on a defined order of elements. Here, the preselected elements get sorted by their y-values in descending order (as positive y-values are "left"), and the third element is selected.

Locating Damage

The damage localisation process starts if the component is damaged (= is linked to damage objects in the BMS graph). We decided to conduct this process not with the IFC element itself but with a bounding box that encases/represents it. Bounding boxes are easier to process and can adapt the concept to components that are not perfectly rectangular, with a sufficient degree of accuracy for our approach.

For each damage of the component, its location information is retrieved, which is stored as an *Area of Interest* (see Figure 2). Since we previously assigned all location terms to an axis (*Y-Right*, *Z-Top*, *X-Front*, etc.), the terms can be used to divide the component along this axis. Figure 5 demonstrates the general concept of dividing an element into three areas per axis. If the damage location is defined with only one term (e.g. "Right Side"), the area is picked from these predefined areas. If the damage is located using multiple axis-references, the intersection of the corresponding areas is calculated. Figure 6 shows an example of this process with the location information "Top", "Right", and "Front".

If a damage has more than one location specification on the same axis (e.g. "Right and Left"), the localisation process is repeated for each of them, resulting in two separate dam-

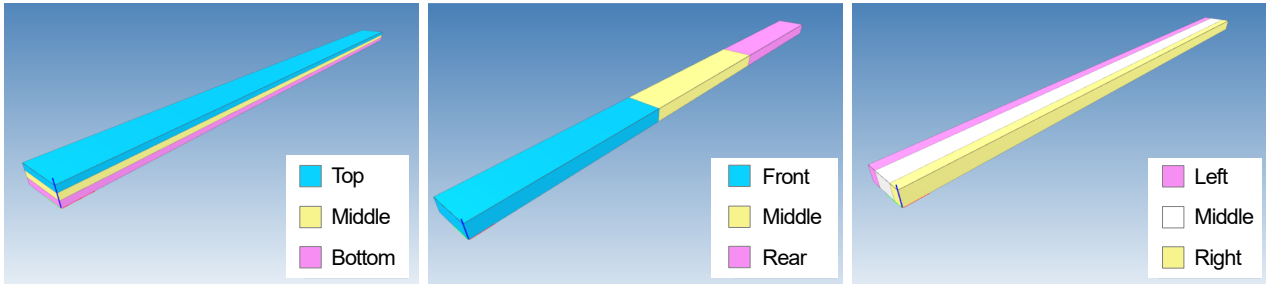


Figure 5: Subdivision of a component into three areas per axis direction.

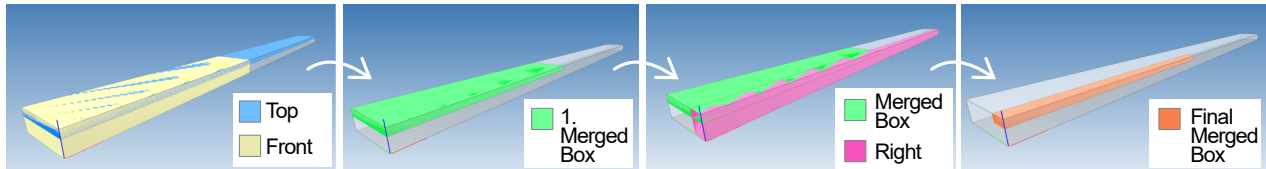


Figure 6: Locating a damage area by intersecting component areas.

age locations. If damage is located by a bridge-structure-related location term, the zone model is taken to intersect the component with the corresponding zone(s) (e.g. damage on the roadway at the "1. Span").

The technical result of the intersection/localisation process for each damage is a subset of the component-encasing bounding box. All resulting subsets, considered as 3D representations of the damages, are stored as *IfcBuildingElementProxy* objects in a separate IFC file.

Linking

In the last step of the process, the correlations found between BMS and IFC data are stored as links. For this purpose, the IFC model gets also converted into RDF using the IFCtoLBD Converter⁵. The BMS components are then linked to the corresponding IFC graph elements at the object level via *asb:hasModelRepresentation*. Also, the IFC file containing the damage (area) representing bounding boxes (Proxies) gets converted, enabling to link the *damage-* and *area of interest-*objects of the BMS graph to their geometrical counterpart.

Use Case Example

The use case example is a straight, multi-span girder bridge from 2002 with 40 damages. The SIB-Bauwerke dataset of the bridge was provided by the roadway authority, and the IFC model was created manually based on a point cloud. Since the IFC Bridge extension classes were not yet implemented in the modelling tool, we used user-defined IFC element types (e.g. ASB-ING Abutment = *IfcWall* with *IfcWallTypeEnum* "USERDEFINED": "ABUTMENT") to work already with the terms. The scripts and the data sets can be found on GitHub⁶.

Applying the approach to the use case example resulted in

⁵<https://github.com/Design-Computation-RWTH/IFCtoLBD>

⁶<https://github.com/AnneGoebels/DamageLocation>

a successful localisation and linking of a total of 72% of the documented damages and components. As a benchmark/ground truth served the manual localisation and assignment of damages and components documented in the SIB-Bauwerke data by an expert.

In detail, 82% of the damages (33/40) could be successfully located, of which 57% (19/33) could be assigned to a specific component area. The remaining 43% (14/33) could only be assigned to a component as a whole.

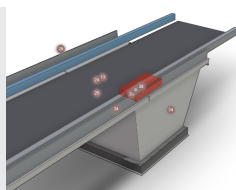
The success rate for assigning SIB-Bauwerke components to the corresponding model elements reached 60% (20/33). The non-assigned components, however, were mainly secondary equipment elements (railings, drainage components, etc.), from which only 35% (6/17) could be located, whereas 87% (14/16) of the main components (abutments, superstructure, caps, etc.) could be successfully located.

All unlocated components (12) and 42% of the unlocated damages (3/7) could not be located because the searched components have not been modelled in the initially simplified 3D model (e.g. drainage system, bird screens, parking bracket, etc.). Due to the modelling technique, two damages on the cross beam could not be located, as it was modelled in conjunction with the superstructure. Another two damages could not be found due to missing information in the BMS. The monolithic modelling method also causes slightly incorrect localisation of damage to abutments, as they consist of three joint walls. In this case, damage with the location term "right" was located on the right wing wall, but the right side of the front wall was intended (see, e.g. figure 7, 3rd example).

That no specific component area was found for 14 damages is primarily due to the fact that the damage area was described as "continuous", and thus the whole component was meant (see, e.g. figure 7, 2nd example). For six of the damages, there was no indication of the location in SIB-Bauwerke. Here, the damage type sometimes implicitly

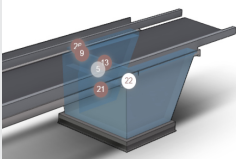
Location/ Representation of damage area with 3 location terms

Cap, surface, [28]
Multiple,
Spalled,
Rear wing wall,
Right,
Upper side,
only in the area of the right rear wing,



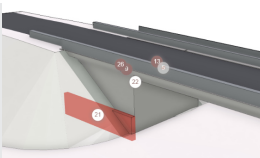
Damage assigned to whole component

Abutment, concrete surface, isolated, cracked, [22]
abutment front,
in the area of abutment-wall and wing,
horizontal and vertical



Wrongly located damage, caused by monolithic modelling of abutment

Abutment wall, Concrete, One spot, [21]
Wet spot,
Front abutment,
Right,
Bottom,
Drip spout drains to front wall,



Too large damage area in comparison to size given in free text field

Cap, Concrete, Isolated, Spalled, [9]
Quantity: 2 spot(s),
At end of structure,
Right,
L = 10, 20 CM

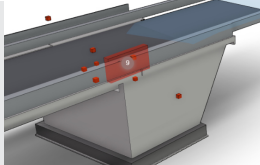


Figure 7: Comparison of the original damage documentation of SIB Bauwerke (left) and the created geometric representation and localisation in the IFC model (right). [Bold: location description terms, Italic: Free text field information]

defines the location, which requires expert knowledge.

The damage areas/boxes created reflect the accuracy of the location description made in SIB-Bauwerke using the property "Location". Further information from the free text field and the property "Damage size" was not processed in the current method. As a result, the damage boxes are sometimes inaccurate / too large (see, e.g. figure 7, 4th example). In addition, the process always creates three-dimensional boxes, which do not correctly represent the often "only" superficial types of damage.

In total, 68% of the non-localised cases are caused by missing or incorrectly modelled model elements. The remaining 24% are due to missing information in SIB-Bauwerke. Thus, by better aligning the model with the structure/content of the BMS data, a significant improvement in the success rate can be achieved. By extending the transfer process to handle also free text, damage type and size information of SIB-Bauwerke, the missing location information could be compensated to locate/represent more damages (more accurately) in the 3D model.

However, the accuracy achieved in the use case example is sufficient to find the damage on site, talk about it remotely, and make geometric/spatial queries about mostly damaged components/areas. In addition, the concurrent linking of

the data contributes to the creation of the DT, which can subsequently analyse the semantic and geometric data in combination.

The results of the use case example can be viewed in a prototypical viewer online⁷.

Result Analysis and Discussion

This paper demonstrated an approach to automatically transfer implicit textual location descriptions into explicit locations and areas of a 3D model. The results of the use case example show that the proposed method can identify specific model elements based on a textual description and achieve three-dimensional representations of described damaged areas. The successful localisation enables the generation of object-specific links between the BMS and 3D model data, contributing to an automated digital twin implementation.

The results further illustrate that the key factor for a successful transfer process is the adequate modelling of the 3D model. It should reflect the perspective of the maintenance domain, i.e. all elements should be modelled as individual components, even if they are compound structures from a construction point of view. This way, the assignment of information and damage to elements like separate beams of the superstructure or specific walls of the abutment is possible. To ensure accurate location, ancillary components should also be modelled, at least schematically. Therefore, the mapping between the IFC and ASB-ING components must be extended accordingly.

However, the use case has shown that 72% of the textual descriptions can already be transferred with the current process. This saves already much time in finding the damage during an inspection and generates a large amount of training data considering the possible use cases of 43,000 bridges. Assuming that each bridge has 70 descriptions (a low estimate), this would result in over 2 million processed and localised component and damage location descriptions, given a success rate of about 70%.

Still, the costs of creating a 3D model (with or without the requirements mentioned above) remain very high and would hardly be carried out for all 43,000 bridges. Therefore, in future work steps, it will be attempted to replace the IFC model with a schematic, parametric model, which represents the BMS/SIB-Bauwerke components by bounding boxes. These can be adapted to the respective bridge using bridge-specific composition/construction rules and the documented component dimensions. Based on this schematic model, the damage localisation process of the demonstrated approach can be carried out.

A general technical limitation of the transfer process is due to the implementation using bounding boxes. The advantage of easy processing of bounding boxes only applies to straight bridges. If they are curved, the creation and intersection of bounding boxes would be much more complex and require more computational effort. However, an

⁷www.design-computation-rwth.github.io/LinkedDataViewerPublic/

analysis of the 43,000 bridges to which the approach could be applied shows that 77% of them are "simple", straight bridges. Thus, the method could still be applied to about 33,000 bridges in Germany.

Conclusion

The presented approach and its results have shown that the automatic transfer of implicit textual location descriptions of legacy BMS data into three-dimensional model contexts is possible. Based on a component class mapping and the intersection of bounding boxes, the method presented has achieved a success rate of 72%. This rate could be increased significantly with the improvement of the model data basis presented in the Result Analysis section.

The resulting geometrical representations of damage areas improve access and management of bridge data and facilitate the inspection process. Finding the damage on-site and communicating about them remotely is easier and independent of subjective interpretation and pre-knowledge. Moreover, the achieved interlinking of the BMS and IFC data enables computer-aided analyses of their geometrical and semantic correlations.

The automated approach contributes to an efficient implementation of DTs for existing structures and a predictive maintenance strategy without an immense workload of repeated data acquisition. It also proves that it is possible to use the existing BMS data already together with new digital solutions without (waiting for) modernised BMSs. In future work, the presented approach can serve as a basis for further augmentation of 3D models. For example, damage pictures can be placed at the automatically localised damage locations in the correct scale, or surface models of the damage can be added. Moreover, the results of the approach can be a basis for ML methods that use stochastic Natural Language Processing (NLP) algorithms, opening the approach up to unstructured text data outside SIB-Bauwerke.

Acknowledgements

We thank the German Federal Ministry for Digital and Transport (BMDV) for their support in funding the Twin-Gen research project and the Autobahn GmbH Südbayern for providing the data set of the use case example bridge. We also want to thank the Geodetic Institute and Chair of Construction Informatics & Geoinformation Systems of RWTH Aachen for their contribution to the implementation of this work.

References

- Al-Hakam, H., Bonduel, M., and Scherer, R. J. (2019). An ontological model for the representation of damage to constructions. In *Proceedings of the 7th Linked Data in Architecture and Construction Workshop - (LDAC)*, pages 64–77, Lisbon. CEUR-WS.
- Al-Hakam, H. and Scherer, R. J. (2020). Areas of Interest - Semantic description of component locations for damage assessment. In *EG-ICE 2020 Proceedings: Workshop on Intelligent Computing in Engineering*, Berlin.
- Artus, M. and Koch, C. (2022). Object-Oriented Damage Information Modeling Concepts and Implementation for Bridge Inspection. *Journal of Computing in Civil Engineering*, 36(6):04022029–1 – 04022029–21.
- Beetz, J., van Leeuwen, J., and de Vries, B. (2009). IfcOWL: A case of transforming EXPRESS schemas into ontologies. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 23(1):89–101.
- Borrmann, A., Muhic, S., Hyvärinen, J., Chipman, T., Jaud, S., Castaing, C., Dumoulin, C., Liebich, T., and Mol, L. (2019). The IFC-Bridge project – Extending the IFC standard to enable high-quality exchange of bridge information models. In *Proceedings of the 2019 European Conference on Computing in Construction*.
- Bundesanstalt für Straßenwesen (2021). Brückenstatistik. <https://www.bast.de/DE/Statistik/Bruecken/Brueckenstatistik.pdf>? 2022-12-08.
- Bundesministerium für Verkehr, Bau und Stadtentwicklung, Abteilung Straßenbau. (2013). Anweisung Straßeninformationsbank Segment Bauwerksdaten.
- Göbels, A. (2021). Conversion of infrastructure inspection data into linked data models. In *Proceedings of the 32. Forum Bauinformatik 2021*.
- Göbels, A. and Beetz, J. (2021). Conversion of legacy domain models into ontologies for infrastructure maintenance. In *Proceedings of the 9th Linked Data in Architecture and Construction Workshop - LDAC 2021*, volume 3081, page 12, Luxembourg. CEUR-WS.
- Liu, L., Hagedorn, P., and König, M. (2022). BIM-Based Organization of Inspection Data Using Semantic Web Technology for Infrastructure Asset Management. In *Proceedings of the 1st Conference of the European Association on Quality Control of Bridges and Structures*, volume 200, pages 1117–1126.
- Saback, V., Popescu, C., Blanksvärd, T., and Täljsten, B. (2022). Asset Management of Existing Concrete Bridges Using Digital Twins and BIM: a State-of-the-Art Literature Review. *Nordic Concrete Research*, 66(1):91–111.