FRAMEWORK FOR SEMANTIC INTEGRATION OF IOT, BIM AND GIS OF TRANSPORT INFRASTRUCTURE

IM-SAFE.EU

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Document information

Name	Remark
Document Name	Framework for semantic integration of IoT, BIM and GIS of transport infrastructure
Version No.	V1.0
Due date Annex I	15/07/2020
Report date	30/04/2022 (last updated in 28/04/2022)
Number of pages	58 including Appendix
Lead Author	Matthias Weise (AEC3)
Other Authors	Rizal Sebastian (TNO), Nicolas Mitsch (AEC3), Ana Sánchez Rodriguez (UVIGO)
Dissemination level	Public

Document history

Ver.	Date	Remark	Author	Checked by
Rev. 0.1	01/05/2021	Creation of the document	M.Weise, N.Mitsch	
Rev. 0.2	14/04/2022	First draft for revision	All	M.Weise
Rev. 0.3	19/04/2022	Amendments from reviewer	All	M. Bohms
Rev. 0.4	25/04/2022	Amendments from reviewer	All	M. Rigotti
Rev. 0.5	26/04/2022	Amendments from reviewer	All	A. Strauss
Rev. 0.6	28/04/2022	Document for approval	All	F. Cinquini and A. Bigaj-van Vliet
Rev. 1.0	30/04/2022	Document Uploaded	All	A. Bigaj-van Vliet

Document approval

Ver.	Date	Approval	Project's Role	Beneficiary
Rev. 1.0	30/04/2022	A. Bigaj-van Vliet	Coordinator	TNO

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Executive summary

This deliverable is a result of work package 4, which is dealing with topics related to digitalisation as enabling technology for monitoring and predictive maintenance as being addressed by the IM-SAFE project. Deliverable 4.2 is focusing on semantic integration of data sources being relevant in context of asset management, in particular covering IoT, BIM and GIS. It gives insights into latest developments in open standards for interoperability and describes possible implementation using W3C Linked Data/Semantic Web technology including further standardization work related to its use. Results of this work will be used in task 5.3 to draw recommendations for further standardization work related to digitalisation in monitoring and asset management.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Table of Content

DOCUMENT INFORMATION 2
DOCUMENT HISTORY 2
DOCUMENT APPROVAL 2
EXECUTIVE SUMMARY 3
TABLE OF CONTENT 4
TABLE OF FIGURES 6
TABLE OF TABLES 7

1. INTRODUCTION 8

1.1 OBJECTIVE 8
1.2 METHODOLOGY AND STRUCTURE OF THE DELIVERABLE 8
1.3 LINK TO OTHER WORK PACKAGES 10
1.4 CHALLENGES AND DATA SHARING PRINCIPLES 10

2. MODELLING – DATA INTEGRATION WORKFLOW 12

2.1 OVERALL WORKFLOW FROM DATA ACQUISITION TILL DECISION SUPPORT 12
2.2 EXAMPLES AND STUDIES 14
2.3 STANDARDS RELATED TO PROCESS AND WORKFLOW DEFINITIONS 15
2.3.1 RIGHT LEVEL OF ABSTRACTION WITH FOCUS ON DATA SHARING 15
2.3.2 DEALING WITH DATA DEPENDENCIES 15
2.3.3 PROCESS SPECIFICATIONS FOR BIM-BASED COLLABORATION 16
2.3.4 STANDARDS FOR WORKFLOW AND PROCESS DEFINITIONS 17
2.3.5 ACCOMPANYING STANDARDS DEALING WITH SEMANTICS 18
2.4 CONCLUSION 19

3. MODELS – OPEN STANDARDS FOR INTEROPERABILITY 20

3.1 INTERNET OF THINGS (IOT) 21
3.1.1 SCOPE 21
3.1.2 TECHNOLOGY 21
3.1.3 STATUS AND FUTURE DEVELOPMENTS 22
3.2 BUILDING INFORMATION MODEL (BIM), DIGITALISATION OF THE AEC INDUSTRY 24
3.2.1 SCOPE 25
3.2.2 TECHNOLOGY 27
3.2.3 STATUS AND FUTURE DEVELOPMENTS 28
3.3 GEOSPATIAL INFORMATION SYSTEM (GIS) 29
3.3.1 SCOPE 29
3.3.2 TECHNOLOGY 30
3.3.3 STATUS AND FUTURE DEVELOPMENTS 32

4. MANAGEMENT – DATA INTEROPERABILITY FRAMEWORK 33

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

4.1 OVERALL GOALS BY EC **33**
4.2 CHALLENGES AND LINKED DATA PRINCIPLES **33**
4.3 LINKED ASSET DATA AND SEMANTIC WEB TECHNOLOGY **35**
4.4 CEDR-INTERLINK APPROACH **36**
4.5 SEMANTIC MODELLING AND LINKING **40**
4.6 EMERGING DIGITAL TWIN CONCEPT **42**
4.6.1 PROGRESS OF DIGITAL TWIN BEYOND BIM 42
4.6.2 RECOMMENDED PROGRESS TOWARDS SEMANTIC DIGITAL TWINS (SDT) 43

5. CONCLUSION AND RECOMMENDATIONS **45**

6. BIBLIOGRAPHY **47**

APPENDIX A - FURTHER RESOURCES **50**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Table of Figures

Figure 1.1 – Overview about Work packages and their relationship from DoW. 10

Figure 1.2 – FAIR principles adopted to IM-SAFE topics related to data interoperability. 11

Figure 1.3 – Interoperability framework from EIR, including Semantic Interoperability. 11

Figure 2.1 – Generalized monitoring process from on-site measurements to decisions support for maintenance. 12

Figure 2.2 – General assessment of existing structures according to ISO 13822 16

Figure 2.3 – Relationship between ISO 19650, ISO 29481 and EN 17412 as describe in EN 17412. 18

Figure 2.4 – Recommendation related to documentation of business processes from the EIF report. 19

Figure 3.1 – Three main areas for data interoperability with some overlapping scope that is discussed in more detail in chapter 3. The figure shows the principle of overlapping information only. 20

Figure 3.2 – Publisher subscriber concept of MQTT 22

Figure 3.3 – SAREF device 23

Figure 3.4 – Sensor Data Integration Approach 24

Figure 3.5 – Backbone structure of IFC from T. Liebich/buildingSMART 25

Figure 3.6 – Extract from FM-related use cases discussed in the IfcBridge project. 26

Figure 3.7 – Assessment of BIM-based use cases from the IfcBridge project 27

Figure 3.8 – Digital Terrain Model and Digital Surface Model 29

Figure 3.9 – Point cloud of a train station and rails 31

Figure 4.1 – Challenge of having various datasets about the same asset and the idea of linking those datasets together (from Bakker et al. 2019) 34

Figure 4.2 – Overall approach to link all relevant datasets including e.g. BIM (from Bakker et al. 2019) 34

Figure 4.3 – Framework developed by INTERLINK for a European Road OTL (from www.roadotl.eu) 36

Figure 4.4 – Principle of data modelling using a neutral representation of the reality that can later be “typed” according to specific use cases such as the 3 traffic lights red, yellow, green. 39

Figure 4.5 – Adoption of the modelling principle to a bridge use case with BIM/IFC being the neutral representation of the building that is further classified by AM4INFRA, the German ASB-ING and the Swedish CoClass. 39

Figure 4.6 – W3C Linked Data/Semantic Web tech supporting FAIRness for data 40

Figure 4.7 – SML Top Level ontology 41

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

Table of Tables

Table 1.1 – Most relevant standardization bodies in context of BIM-based work. 9
Table 3.1 – Common formats for GIS data 30
Table 4.1 – List of ontologies being part of the EUROTL framework 38

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

1. Introduction

1.1 Objective

The goal of Task 4.2 within work package 4 is to address the relation between the future standards in monitoring with the open IT standards that are already known or in use in construction, especially on the **interoperability** of Internet of Things (IoT), Building Information Model (BIM), Geographical Information System (GIS), and Linked Data/Semantic Web (LD/SW). Goals mentioned in the DoW have been:

- Create comprehensive insight into the latest developments in open standards for interoperability
- Clarifying the advantages of the standardized Linked Data approach and 'Semantic Web'-based ontologies
- Outlining the possible implementation of the CEDR-INTERLINK 'Modelling and Linking Guides' to link different types of data

This deliverable summarizes the results of Task 4.2 and addresses the following topics related to data interoperability:

1 Use case support: Show the concept of a data integration approach adopted for monitoring and predictive maintenance using existing BIM, GIS and IoT standards; such approach first must identify relevant data being used and, more importantly when dealing with data interoperability, shared between different stakeholders.
2 Standards for data exchange and sharing: Review of existing standards from BIM, GIS and IoT considering requirements derived from monitoring and predictive maintenance use cases.
3 Common framework and technology: Recommendations for an interoperability framework initially based on CEDR-INTERLINK approach and semantic web technology.

It has been very clear from the beginning that a framework for semantic integration must deal with different domains and thus different types of data exchange standards, which however need a common ground to be able to work together. Also, it is not only the data about the building or a measurement itself that matters. It is also about the information that adds context to such data such as for instance how data was produced (which sensor, which data analytics) and who is the owner of the data. Such additional "meta-data" is needed to better understand measurements and to draw the right conclusion for predictive maintenance or in case of an emergency.

1.2 Methodology and Structure of the Deliverable

Work on Task 4.2 was following a same logic as known from introducing BIM-based project collaboration. This logic can be best described by the three meanings of M in BIM, namely:

1. **M**odel - meaning the *data* being shared between stakeholders,
2. **M**odelling - meaning the *processes* behind the data and
3. **M**anagement - meaning the *implementation and integration* making data exchange and sharing happen.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

All three types are equally important when dealing with interoperability topics. Applying this logic was leading to the following activities and main structure of the deliverable:

1. Draft of typical ICT-related workflows to be used as a reference to identify parts for further standardisation.
2. Identification of main standards from BIM, GIS and IoT domain and state-of-the-art analysis in order to identify the status and potential gaps for monitoring and predictive maintenance use cases.
3. Generic standards and frameworks dealing with implementation of data sharing solutions.

Work on Task 4.2 was executed by different partners with different expertise and background knowledge. We also tried to include results from other IM-SAFE work packages as well as feedback collected from CoP members. The focus is clearly on existing standards from ISO, CEN and other standardization bodies or organisation because they are already agreed within a community. An overview about relevant standardization bodies is given in Table 1.1.

Table 1.1 – Most relevant standardization bodies in context of BIM-based work (see Bibliography).

Standardization body	Description	Standards
buildingSMART International	A non-profit organization defining standards for BIM-based working, bSI standards are also published as ISO, CEN and national standards	IFC (ISO 16739), bSDD, BCF, mvdXML, ifcXML, IDS
ISO	International standardization body also covering BIM and related standards	ISO 12006 (basis for bSDD) ISO 19650, ISO 23386/23387 ISO 29481 (IDM)
CEN	European standardization body	EN 17412 (LOIN), preEN 17632 (SML)
W3C	Standardization body for all internet-related standards including semantic-web and model-linking approaches	XML/XSD, OWL/RDF/Turtle, JSON-LD, SPARQL/SHACL
Green Building Consortium	Industry driven standard for energy-efficiency related applications	gbXML
OGC	Open Geospatial Consortium, International voluntary consensus standards organisation for geospatial content and services.	(city)GML, GeoSPARQL

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

1.3 Link to other Work Packages

As shown in Figure 1, Task 4.2 connects the findings of WP2 and WP3 through digitalisation as enabling technology. As such the sensors used and data transmitted as described in Task 2.1 Appendix A as well as data from reported tests and devices have been taken into consideration when evaluating the scope of current technologies available in the construction industry. The data collected is among other possible use cases to be used in predictive maintenance and therefore needs to be stored and processed in a way that building safety and maintenance strategies can be evaluated. The findings of WP4 will be used as input for WP5 to propose further standardisation efforts.

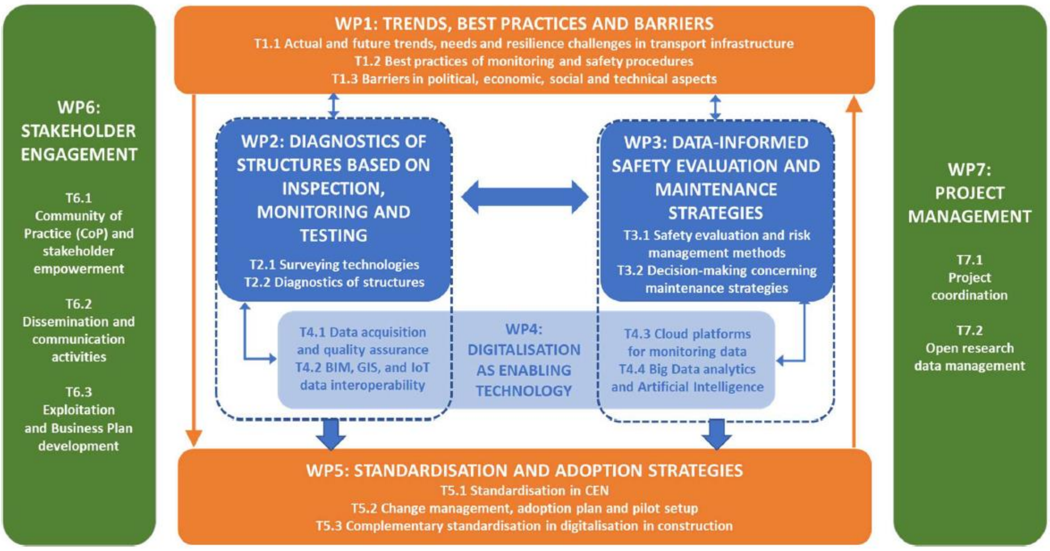


Figure 1.1 – Overview about Work packages and their relationship from DoW.

1.4 Challenges and Data Sharing Principles

Our infrastructure are long-lasting investments that shall be functional for many decades. Compared to the lifespan of such assets ICT solutions in terms of software and data models, and recently the arrival of BIM and digital twins, are a still young, rapidly evolving technology. Accordingly, it is crucial to establish procedures to manage digital data about our infrastructure so that all necessary information is available, usable, and up to date.

The research community is faced with similar challenges and recently came up with the so called FAIR[1] guiding principles, which is a very memorable acronym and could be applied to data sharing in engineering applications as well.

[1] https://www.go-fair.org/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

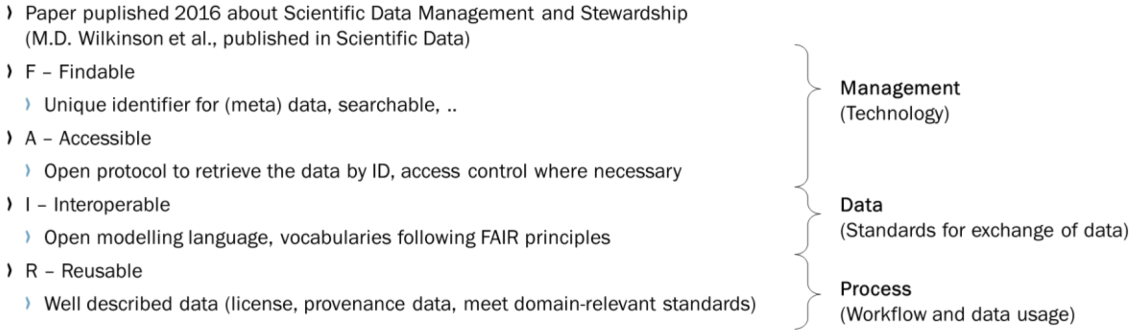


Figure 1.2 – FAIR principles adopted to IM-SAFE topics related to data interoperability.

The acronym FAIR stands for data that must be **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. It describes core challenges in a decentralized, distributed environment with multiple actors and different types of domain data. This principle can be linked to the topics and structure of this deliverable as shown in Figure 1.2.

Other relevant developments and guidelines regarding interoperability are coming from the New European Interoperability Framework (EIF)[2], published by the European Union in 2017. Beside a set of general interoperability principles, it also an interoperability model with four layers, where semantic interoperability is one of them (Figure 1.3).

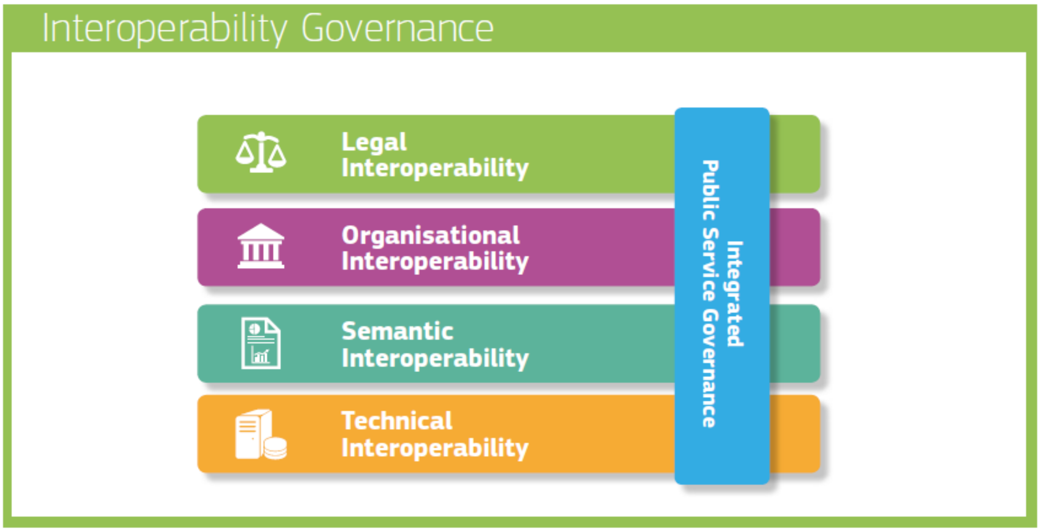


Figure 1.3 – Interoperability framework from EIR, including Semantic Interoperability.

While the focus of this task is on semantic interoperability, it also addresses organisational and technical topics as for instance covered in chapters 2 and 4.

[2] https://op.europa.eu/en/publication-detail/-/publication/bca40dde-deee-11e7-9749-01aa75ed71a1/language-en

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

2. Modelling – Data integration workflow

Conceptually, there are a lot of similarities between different implementations of BIM as mentioned in the introduction. The definition of a BIM implementation typically starts with analysing collaboration processes with the goal to derive all information requirements according to agreed project goals and use cases.

This first main chapter is therefore looking into typical monitoring as well as predictive and preventive maintenance processes to better understand what data is needed, when and for what purposes, and who is responsible for that data. It shows the typical data flow and explains the types of involved standards.

Looking into such data-driven processes will also help to answer the questions regarding what data must be permanently stored and what data must be shared with other stakeholders. It is not meant to be an all-encompassing analysis, because such agreements cannot be generalized for all monitoring and maintenance processes. However, the used methodology and therefore used standards should show how data interoperability topics need to be addressed from use case driven view.

2.1 Overall workflow from data acquisition till decision support

In this section, a brief description of the main steps in the workflow from data acquisition till decision support will be provided in the context of inspection, monitoring, maintenance, and safety of transport infrastructures. Subsequently, the description will be assessed against the typical workflows as found from literature. Finally, the workflows as supported by existing software solutions are shown to analyse the current extent and the need for workflow standardisation.

The following scheme illustrates the overall workflow.

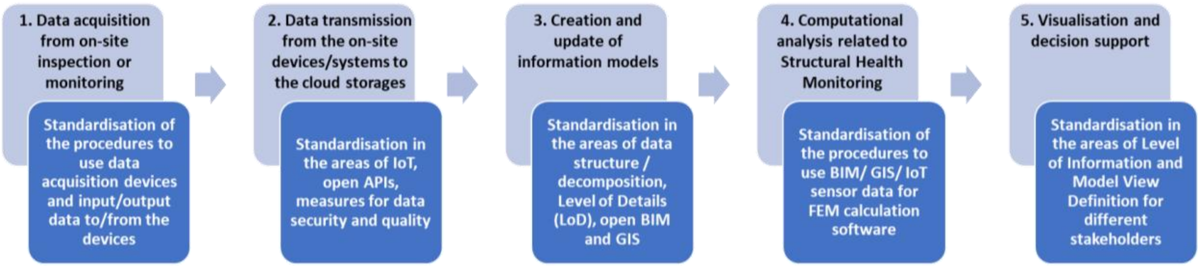


Figure 2.1 – Generalized monitoring process from on-site measurements to decisions support for maintenance.

For each step in this workflow, a brief description is provided about: what data is collected or generated; what the data type/format is and where it is stored; how it is published or shared; and which relevant standards apply to the data and the workflow. The overall workflow consists of five steps:

1. **Data acquisition from on-site inspection or monitoring.**

   At this step, data is generated by among others, acoustic and optical measurement devices, attached and remote sensors, and static and moving 3D scanners. The types of data at this step vary from acoustic signals, geometric data and images to photogrammetry, aerial and polygon meshes from 3D scans.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

There is still a lack of standard procedures to use the devices/systems for the data acquisition, for instance with regard to positioning of the data capture devices/systems, acoustic and optical properties, and the quality of output data. These aspects are studied in Work Package 2 and Task 4.1 of the IM-SAFE project. Related report published by the IM-SAFE project is the D4.1.

2. **Data transmission from the on-site devices/systems to the cloud storages.**

   Nowadays, more and more measurement, sensing and scanning devices have internet and GPS connectivity. Therefore, the captured or generated data, including the GPS positions, can be transmitted directly through the internet from the infrastructure asset's location to the cloud platforms.
   Standardisation is relevant in the areas of Internet of Things (IoT) – this subject will be elaborated in Chapter 3 of this report– along with the interfaces (API) between the measurement; sensing and scanning devices and the cloud platforms; the measures to prevent transmission loss and cloud data loss; and the design and the workflow handling on the cloud platforms – this subject is addressed in Tasks 4.3 and 4.4 of the IM-SAFE project. Related reports published by the IM-SAFE project are D4.3 and D4.4.

3. **Creation and update of information models.**

   The acquired information from inspection and monitoring can be used to create, reconstruct, or update the information models of the infrastructure assets, such as the BIM or GIS models (data). In this regard, standardisation is needed in the aggregation of the information models, the data structure/decomposition, and the Level of Details (LoD). There are existing BIM and GIS open standards, such as IFC by buildingSMART and GML by Open Geospatial Consortium. These standards are further discussed in Chapter 3 of this report.

4. **Computational analysis related to Structural Health Monitoring.**

   Received measurements typically require further data processing to draw conclusions related to safety and building maintenance. Especially for analysing maintenance and safety of transport infrastructures, it is important that the information from inspection and monitoring, either or not embedded in and retrieved from BIM or GIS models, is eligible for analytical use for Structural Health Monitoring (SHM). As such, this information usually needs to be fed into a Finite Element Model (FEM) and the software computational tools. A relevant area for standardisation is related with the conversion of the information from BIM (or other information models) to FEM, commonly known as BIM-to-FEM. Standardisation can be concerned with automating the BIM-to-FEM process, both in terms of time efficiency by reducing or minimising the manual effort by data scientist and civil engineers, as well as in terms of computational accuracy and result quality depending on the compatibility of the software tools.

5. **Visualisation and decision support.**

   Both for engineers and asset owners or managers who are making decisions regarding maintenance and safety of the transport infrastructure, it is important that the data, information models and analytical results can be presented in a clear and coherent way for unbiased decision-making. Different platforms and software tools have different visualisation modules and techniques. The relevant area for standardisation regards the Level of Information shown for different purposes; the Model View Definition (MVD) for different disciplines (e.g., architects, structural engineers, MEP engineers) and different stakeholders; and the structure of decision-support dashboard.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

2.2 Examples and studies

The aforementioned steps and the overall workflow are supported by a number of studies as reported in scientific literature:

- With regards to road infrastructures, Justo et al.[3] presented a study on the automatic generation of a complete BIM IFC (Industry Foundation Classes) instance model of a road from point cloud data. The study followed a top-down approach where the model definition was analysed to elaborate a list of geometric and semantic parameters that need to be extracted from the point cloud data to effectively build the model. The approach comprised two main steps. First, the data required to model the IFC entities were obtained from both the point cloud processing and external sources that described the standardized geometric characteristics of road elements. An automatic IFC model generation procedure was then performed, and it resulted in an IFC-compliant file containing the road alignment, guardrails, vertical signs, and semantics in the form of property sets. Second, point cloud processing was performed to generate a classification of guardrails in road scenarios, along with other elements such as vertical signs and road markings. These contributed to a set of geometric and semantic parameters for the definition of a 3D model.
- With regard to railway infrastructures, Soilan et al.[4] stated the similarities and differences in the methods for processing 3D point cloud data to BIM IFC. The IFC modelling procedures are similar both for roads and railways. However, there are significant differences between the geometries and elements to be defined. Hence there is a differentiation at an upper level of abstraction in the spatial structure (IfcSpatialStructureElement), which defines a hierarchy where a project is divided into sites and facilities (e.g., road or railway, road segments), as well as in the positioning of most elements in the infrastructure model as guided by the alignment (IfcAlignment / IfcAlignmentCurve). IFC Road and IFC Rail have not yet been published as final standards at the time of writing of this deliverable. It is expected that future research and standardisation on the extent of the generation of IFC Road and IFC Rail models from 3D point cloud data will lead to automation in the digitalization of transport infrastructures.
- With regard to bridges, Sánchez-Rodríguez et al.[5] presented a methodology consisting of three main parts: (i) point cloud classification; (ii) point cloud-to-mesh conversion; (iii) mesh-to-IFC conversion. To overcome the gaps in the existing point cloud data capturing, the methodology used advanced geometric reconstruction techniques and mapped the segmented assets to the latest IFC schema. During the processing, two Levels of Details (LoDs) were considered to make the resulting model lightweight and easy to consume in various use case scenarios. This study further examined whether and to what extent bridges can be better described using the latest schema extension proposal IFC4x2. IFC

[3] Justo, A., Soilan, M., Sanchez-Rodríguez, A. and Riveiro, B. (2021). Scan-to-BIM for the infrastructure domain: Generation of IFC-compliant models of road infrastructure assets and semantics using 3D point cloud data. In: Automation in Construction 127 (2021). https://doi.org/10.1016/j.autcon.2021.103703.

[4] Soilan, M., Justo, A., Sanchez-Rodríguez, A., Lamas, D. and Riveiro, B. (2021). 3D point cloud data processing and infrastructure information models: methods and findings from SAFEWAY project. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2021 XXIV ISPRS Congress.

[5] Sánchez-Rodríguez, A., Esser, S., Abualdenien, J., Borrmann, A. and Riveiro, B. (2020). From point cloud to IFC: A masonry arch bridge case study. In: Proceedings of 27th International Workshop on Intelligent Computing in Engineering. Berlin, Germany.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

4x2 is used for the data model since it covered requirements for detailed geometric representations paired with specific product classifications for each component. This will avoid the need for re-implement a way of combining geometry with semantics in a formal way. The study referred to the IfcBridge extension by buildingSMART International (bSI) that included new classes and types to enable a more detailed description of bridge structures. ifcBridge is an ongoing development by bSI that is relevant for transport infrastructures along with IfcRoad, IfcPortsAndWaterways, IfcTunnel, IfcRail, etc.

2.3 Standards related to process and workflow definitions

The discussion in previous chapters explains the typical data flow from a single measurement to decision support about maintenance. This data flow can easily be adopted to different monitoring scenarios, even including manual inspection representing the traditional source of measurement that is manually uploaded to an asset management system or database.

2.3.1 Right level of abstraction with focus on data sharing

The steps in the discussed monitoring workflow should be documented and properly specified so that data processing can be implemented according to the goals specified for a particular monitoring activity. A generic workflow with various activities, decision points and data reports is for instance part of the ISO 13822, and a good example for such kind of agreements (see Figure 2.2). Technically, data processing can mean a lot of different things. For one scenario it could mean that relevant data is extracted from a wider array of gathered data, maybe using Artificial Intelligence or other sophisticated algorithms. At the same time, additional data from external sources can be added to the database and aligned with existing time series data. On more abstract workflow level data processing can be seen as an activity that uses available data (input data) to produce more meaningful results (output data).

In many cases data processing is hidden in a closed system (black box) and, if no further communication is needed, does not have to be further broken down in such data flow diagram. The activity itself is thus described but typically not further specified. Such documentation and moreover the use of open standards is therefore primarily needed if different systems or users must work together and consequently must be able to share required data. Accordingly, the overall goal for such documentation is to focus on data sharing or exchange requirements.

2.3.2 Dealing with data dependencies

A data flow furthermore shows dependencies between subsequent data processing steps and thus can be used to trigger data update processes or just enables to see if stored data is up-to-date or not. A simple timestamp without knowing about data dependencies will not be sufficient for proper data management in such heterogeneous, multi-source environments. There are more reasons to agree on and document data flows, which in fact have a lot of similarities to agreements being done for BIM-based project collaboration and therefore might be reused in context of monitoring scenarios.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

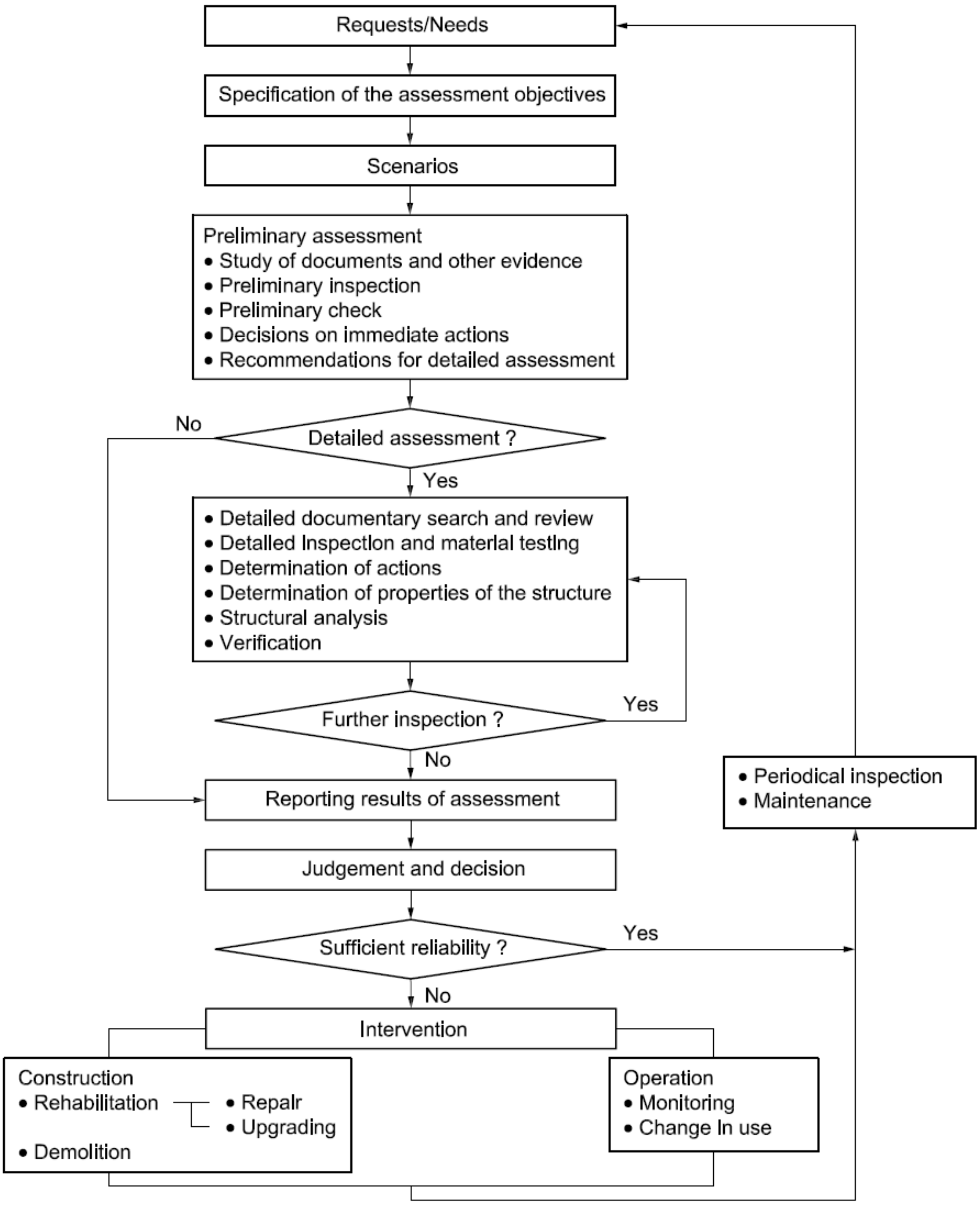


Figure 2.2 – General assessment of existing structures according to ISO 13822

2.3.3 Process specifications for BIM-based collaboration

While standards for describing the building have been in focus in the very beginning (see next chapter about BIM and IFC), accompanying standards about the use of BIM in a collaborative environment were soon recognized and nowadays process data are just as important as the data about the building itself.

Focusing on questions related to data sharing, agreements about the process shall answer five main questions:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

- **Why** is the requested data needed?
  It typically links to an activity or a particular goal that should be achieved. Such activity could be to classify identified cracks in a building according to some damage classification.

- **What** data is needed?
  It describes the input data from a domain expert point of view. It includes types of objects with all relevant properties. For example, a “crack” with further information about the position including the link to the affected building element and the crack dimensions. We may have additional information about the crack like a timestamp, which however is not needed to do the damage classification and thus is not required.

- **Who** is responsible to deliver the data?
  This should identify the source of information, that might be linked to a responsible actor or a system like an autonomous drone survey as described in D4.3.

- **When** is the data needed?
  Design and construction projects are typically linked to phases and milestones. This information will assign the requirement to some sort of a project timeline. Monitoring is however a recurring activity that normally follows a particular inspection interval or recurrent pattern. Accordingly, it is expected that this question is less relevant for building monitoring.

- **How** to describe and share the data?
  This question should provide all ICT-related specifications that enables to implement described activities in software services. Required data must be linked with a data exchange standard so that for instance database queries can be coded in a SQL, SPARQL or other API-based query and data manipulation language.

2.3.4 Standards for workflow and process definitions

Quite some work has been put into the specification of workflows to enable the use of BIM in design, construction and hand-over to facilities management scenarios. The following standards including a short evaluation should be mentioned:

- **ISO 19650** [2]: This standard outlines the concepts and principles for BIM information management. It “*provides recommendations for a framework to manage information including exchanging, recording, versioning and organizing for all actors. It is applicable to the whole life cycle of any built asset, including strategic planning, initial design, engineering, development, documentation and construction, day-to-day operation, maintenance, refurbishment, repair and end-of-life.”*
  This standard is less technical and provides an overall framework.

- **ISO 29481** [7]**:** This standard is also known as Information Delivery Manual (IDM). It focuses on a machine readable, applicable and transferable data schema and corresponding code to facilitate information exchange between project participants. The methodology set out by this standard is intended to help describe data requirements and purpose, and latest part 3 was finally adding a standardized approach for electronic-based exchange through an XML format. While this standard enables to specify workflows and process as described above, it does not specify How data shall be shared or exchanged.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

- **EN 17412** [2a]**:** This standard was recently published and is also known as Level Of Information Needs (LOIN). It somehow overlaps with ISO 29481 and provides further details about the required data being linked to a process specification (see Figure 6). While the principles specified in part 1 are already published, a standardized approach for electronic-based exchange is still missing and is currently developed as part 3. It is not yet clear if part 3 will enable to specify How data shall be shared using an open data format like IFC, CityGML or others.

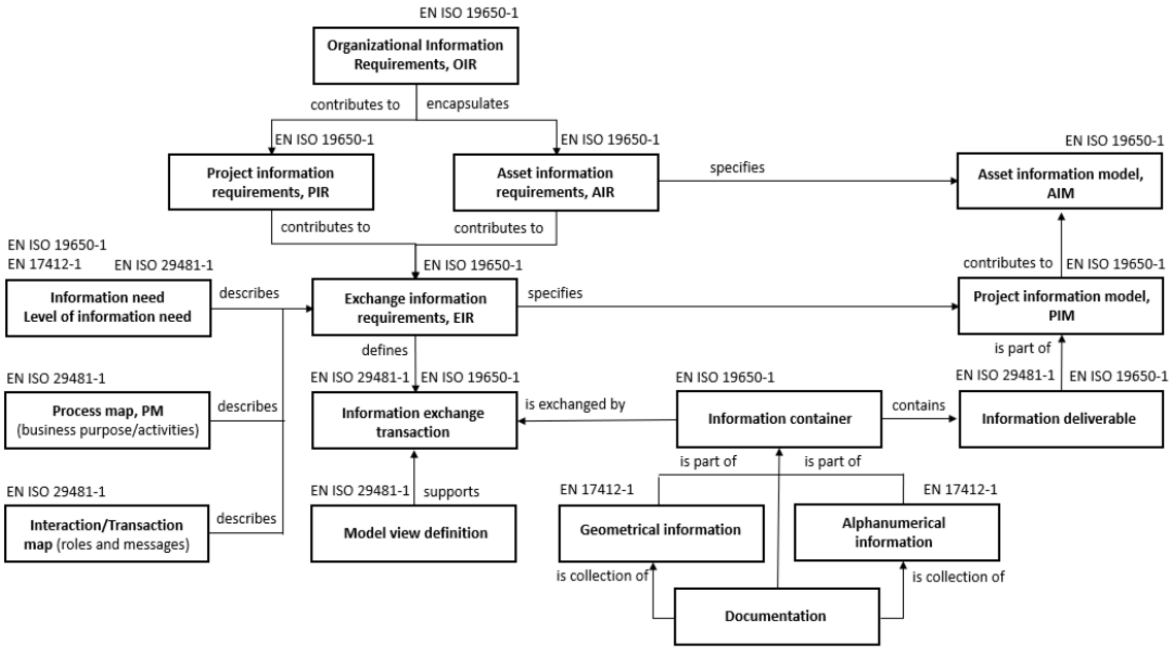


Figure 2.3 – Relationship between ISO 19650, ISO 29481 and EN 17412 as describe in EN 17412.

While there is a good coverage for capturing processes including domain requirements, the mapping to a technical specification like IFC and other standards is still not fully solved. Although there are solutions and ongoing developments within buildingSMART that can cover such agreements (MVD with mvdXML and latest Information Delivery Specification – IDS), they are both focused on IFC support and thus do not cover all possible data sources as expected in a heterogeneous asset management or digital twin environment.

2.3.5 Accompanying standards dealing with semantics

There are further ICT standards related to interoperability and process specification, which both are dealing with semantics and how to specify terms and properties. They are basically used to encode for instance classification systems being used for many purposes and are established as national standards in many countries. It is a generic way to provide a machine-readably representation of national classifications or dictionaries of terms representing classes of objects including their hierarchy and characteristics (or properties).

- **ISO 12006** [1]**:** This set of standards, in particular part 2 and 3, are about classification systems and object-oriented information. It also provides the background for the buildingSMART Data Dictionary (bSDD, [10]), which is an implementation but also a service for hosting AEC related specifications, and moreover is trying to link captured content to the IFC standard. bSDD went through several revisions and lately is also trying

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

to harmonize with the ISO 23386 standard [5]. While the aim to interrelated various classification systems together did not work out in practice, also due to conceptual constraints, the system is now focusing on technical services for classification systems that optionally are linked with the IFC standard [9, 13].

- **ISO 23386 and 23387** [5, 6]**:** These standards have been developed to agree on product data focusing on properties and their grouping into class specific configurations. The conceptual basis is quite generic and while having a lot in common with ISO 12006 major differences are first availability of additional meta-data for properties supporting the management of such agreements, essentially covering the whole lifespan starting from first proposal, to review processes, its use in product catalogues and finally setting it out of order or replacement by newer versions. Second, while a link between different classification systems or a data structure like IFC is technically possible, it is clearly not in focus and thus not further specified.
- **preEN17632 (SML)** [8]**:** This standard gives recommendations about semantic data modelling based on Linked Data/Semantic Web (LD/SW) and therefore should be mentioned in this context as well. Further information is given in chapter 4.5.

Above mentioned standards are defined for the AEC industry, although being generic to some extent and thus usable for other industries as well. Tools developed for those standards will therefore at least partially compete with other standards for semantic specifications, in particular the Semantic Web Technology including additional (standardized) ontologies covering various aspects, including property and knowledge representation (e.g. QUDT, SKOS), provenance data (PROV-O) and many more. This aspect with focus on data management in distributed, heterogeneous environments is discussed in more detail in chapter 4.

2.4 Conclusion

Although not being in focus when dealing with data interoperability this chapter showed that there is much more than just the data about the building (bridge or any other construction) including the measurements from monitoring systems. These topics have been addressed already by presented standards and need to be considered when dealing with ICT support for monitoring and predictive maintenance use cases. Experiences with BIM show that it is necessary to include meta-data[6] about the data, which supports data management in a distributed, heterogeneous environment. This finding is also in line with recommendation 28 (regarding business process alignment) in EIF report [18] from the European Unition.

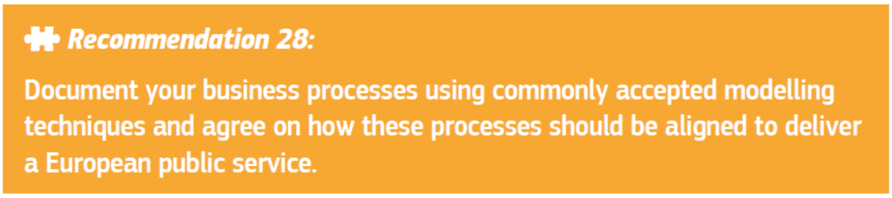


Figure 2.4 – Recommendation related to documentation of business processes from the EIF report.

[6] Where are different types or levels of „meta-data“. It can be data about the data or about the data models.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

3. Models – Open standards for interoperability

When looking into monitoring scenarios for infrastructure, or in general the Digital Twin concept being a digital representation of the physical asset, three main areas can be identified for representing the as-is situation of a building:

- Measurements describing the condition of a building in a particular point in time.
- Physical and functional representation of the building itself including all relevant (static) properties.
- Its location and interaction with the environment and infrastructure network.

These areas are linked with following standardization activities being discussed in more detail in this chapter:

- Internet of Things – IoT
- Building Information Model(ing) – BIM
- Geographic Information Systems – GIS

While the Digital Twin concept is the focus of much research and many developments, this chapter will give an overview mainly about standards being available today, either published by official standardization bodies, non-profit organisations or being an (open) industry standard (see also Table 1.1 in chapter 1).

The three areas partially overlap in scope as illustrated in Figure 8. While such overlap, if seen as a confedereded database, duplicates data and thus adds the risk of inconsistencies, it enables to have different views to the physical world that can be linked together by their duplicates like for instance an object representing the building in different data sources. Discussing the scope of each area will highlight such overlaps.

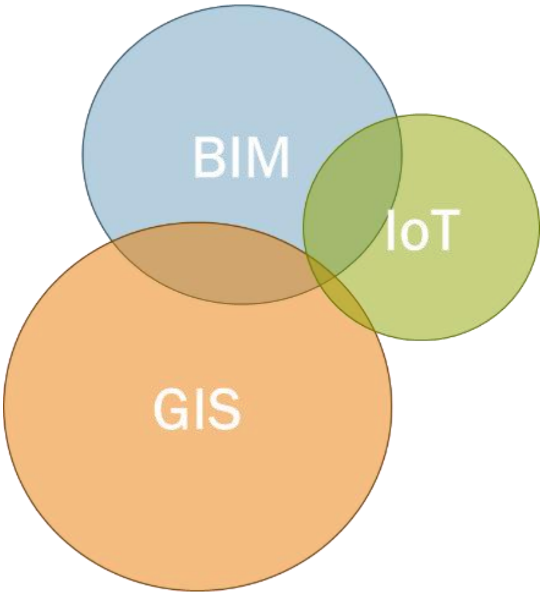


Figure 3.1 – Three main areas for data interoperability with some overlapping scope that is discussed in more detail in chapter 3. The figure shows the principle of overlapping information only.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171. © 2021

Standards from presented areas have been developed by different bodies and thus have a different history, partially dating back to modelling technologies that have been introduced in the 80ies like the STEP family of standards (ISO 10303, [9a]). It is not necessarily a bad thing to use mature technology even in ICT, but such differences will make a joined use more challenging. Accordingly, each standard is also characterized by the used technology for its specification as well as the data representation (or serialization).

Last not least, discussion of standards will finish with a short assessment of its current state and expected future developments.

3.1 Internet of Things (IoT)

The Internet of things describes technologies that enable physical objects to send or receive data. Such objects are sensors, hubs or actuators that are able to gather, transfer and process data through the internet or other communication technologies in order to trigger actions from the actuators. Since these physical objects are fulfilling different functions and are produced by different manufacturers, they use different ways to communicate with each other and transfer information. Trying to connect entities of different manufacturers or application areas can be a difficult task, which is why standardisation in IoT is needed.

3.1.1 Scope

At this point in time, the Internet of Things has entered most areas of life. From small computers in wristwatches or other so-called wearables to mobile devices and household items such as refrigerators or vacuum cleaners, most objects we encounter in everyday life now have the ability to communicate with other objects. This ability to connect with each other creates new opportunities to collect, process and use data. In the construction industry the IoT can be used to monitor buildings regarding their performance, energy balance, structural integrity and more. Sensor data can also be used to automate different functions within a building such as automatic window opening when certain CO_2 thresholds are hit. This monitoring data can then be combined with geometric data to create a digital twin. A digital twin contains geometric data as well as information to represent a building as closely as possible in a virtual environment.

The collected data from buildings can not only be used for monitoring the current status of buildings but also to predict their behaviour. These predictions can be either used to assess real building performance compared to the expected building performance but also to implement predictive maintenance.

3.1.2 Technology

The data transmission between sensors and data collection hubs or databases takes place using data transfer protocols such as Zigbee, Z-Wave, MQTT or DDS. Depending on the application. Common criteria for choosing the right protocol for a certain application are distance, amount of transmitted data and whether the data is collected in one centralized hub or distributed to multiple destinations.

The gathered data gets processed and stored in a database where it can be aligned with other sources of data such as weather services or traffic data.

When it comes to integrating IoT sensors or actuators into a digital twin concept, the creation of a digital twin relies heavily on proprietary software solutions that don't support a vendor

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

neutral data integration. To integrate IoT objects and BIM-data, the position of every IoT object needs to be known and aligned with the geometric data of the BIM-model.

3.1.3 Status and future developments

Due to the relative novelty and the increased use only in recent years, driven by the improvements in wireless network technology, official standardization bodies have not yet dealt with IoT standardisation, which is why national or international standards cannot be used for the implementation of IoT systems. However, since customers might want to use sensors from different manufacturers, a clear consensus when it comes to data transmission, data processing and the underlying information model is required.

Although some protocols have become established as unofficial standards in the field of IoT, there are no fixed specifications as to which protocols must be used and when. So far, only those are selected that are best suited for the respective application. For example, Z-Wave or Zigbee are generally used for building automation because they consume little power. The lower data rate compared to Bluetooth Low Energy is not relevant in the field of building automation, as the data volumes involved are manageable.

For IoT applications in the infrastructure sector, the data volumes involved are significantly higher, which is why protocols such as Zigbee or Z-Wave are not the best choice. Also, infrastructure monitoring doesn't require communication between devices as much as it is relevant in the building automation sector, which is why the advantages of Zigbee for example shouldn't be used (Telkonet, 2022, [31]). The distances between IoT objects and data collection points are also greater than in a typical building automation application, which is why solutions such as Bluetooth or Bluetooth Low Energy would not be optimal. In this case, the Data Distribution Service (DDS) or MQTT (Palmieri et al. 2019, [26]) standards, for example, which are designed for high data throughput while using low bandwidth (Hernández-Moral et al. 2021, [22]), would be suitable. The implementation in the form of a publisher-subscriber concept makes it possible to distribute the accruing data depending on the respective use case. An object only receives data if it is subscribed to a topic. This topic is filled with data from objects that publish their information as shown in Figure 9.

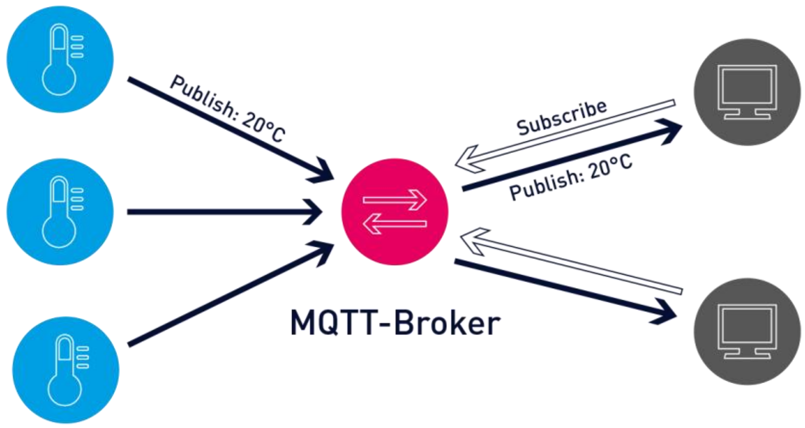


Figure 3.2 – Publisher subscriber concept of MQTT[7]

[7] https://hlassets.paessler.com/common/files/infographics/mqtt-architecture.png (14.04.2022)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171. © 2021
IM-SAFE-Project | TNO, The Netherlands, All right reserved |

However, not only the transfer of data needs to be looked at when it comes to standardisation in IoT. Other aspects such as the information models behind the data need to be considered as well. Just because IoT objects are able to communicate with each other in theory because they use the same transmission protocol does not mean that they understand each other. Just like humans use different languages to communicate, IoT objects rely on different information models to convey information as opposed to just data. For example, different manufacturers can define sensing ranges, measurement intervals, or measurement units in different ways.

There are standardisation approaches such as SAREF [38a] that aim to provide a standardized semantic model for IoT systems. The basic data structure is defined in the SAREF core ontology. To increase the applicability to further areas, so-called extensions have been developed. These include:

- SAREF4ENER, an extension for the energy sector
- SAREF4BLDG, for buildings
- SAREF4CITY for the application in so-called smart cities

While SAREF and its extensions allow the description and integration of many different types of sensors, such as the ones described in Deliverable 2.1 Appendix A, there are no reference implementations for most sensors which leads to potential misunderstandings when using SAREF as underlying information model. The information needed to implement a SAREF device such as sensors or actuators is shown in Figure 10. It shows that SAREF devices can be described without information on the devices location which makes it necessary to add this information later in the process of data integration.

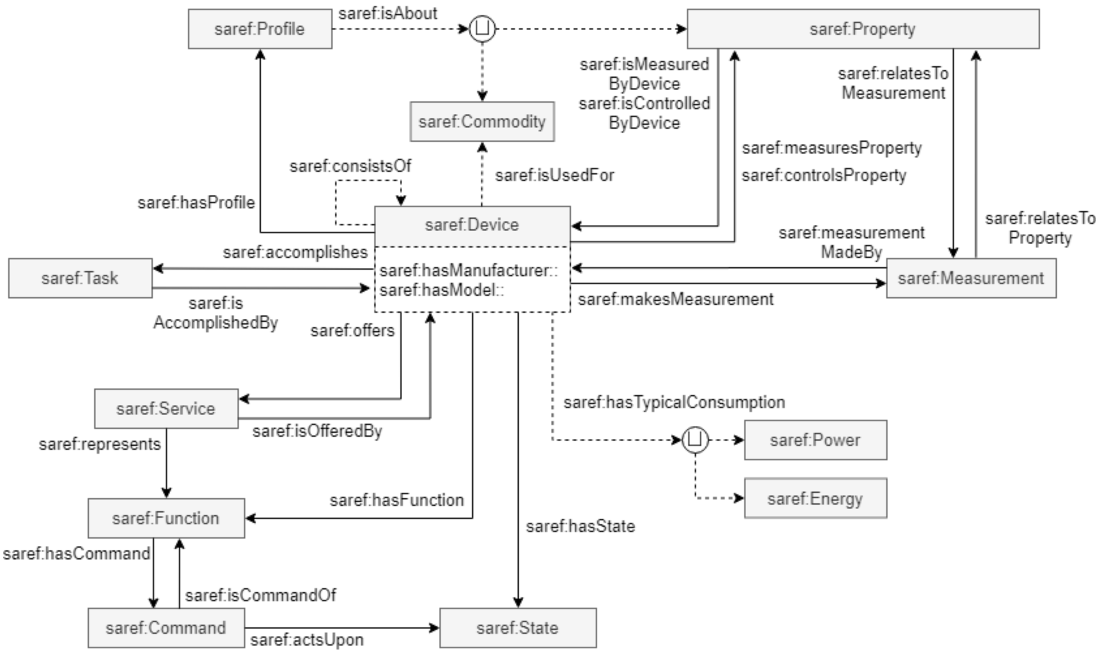


Figure 3.3 – SAREF device[8]

[8] https://saref.etsi.org/core/v3.1.1/diagrams/Device.png (14.04.2022)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

The SAREF core ontology and the extensions are available in the formats JSON-LD, N3, N-Triples, RDF/XML and Turtle

Other approaches include the industrial data spaces (IDS), an information model that tries to provide a framework for the design of IoT infrastructures in an industry environment. The framework can be implemented using open-source software (Alonso et al. 2018). As a driver of web technology standardisation, the W3C also published ontologies aiming to provide a standardized approach of sensor description. The Semantic Sensor Network (SSN) and the SOSA (Sensor, Observation, Sample, and Actuator) ontologies cover a wide range of applications including large-scale scientific monitoring and industrial and household infrastructures (W3C, 2017, [32]).

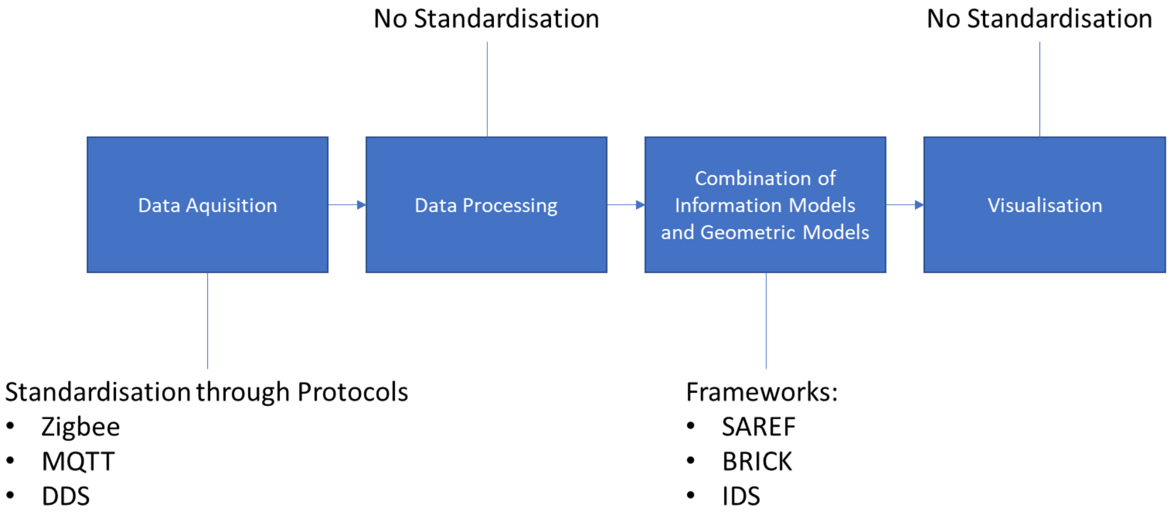


Figure 3.4 – Sensor Data Integration Approach

Figure 11 aims to provide a brief overview of the necessary steps for IoT data integration in BIM models and how heavily each step is standardized. The listed standardisation approaches are only a selection of the standards available. This shows that several steps of the process chain that is necessary to integrate IoT data with BIM are not yet standardized and therefore can't profit from the multiple advantages of standardisation. The aspect of computational analysis of data for structural health monitoring as described in chapter 2 has been deliberately left out of the workflow diagram shown in Figure 11 since the analysis of data is not a necessary requirement for IoT data integration in BIM models.

3.2 Building Information Model (BIM), digitalisation of the AEC industry

BIM became a synonym for digitalisation and modernization of design and construction processes in the AEC industry. While its roots of the technical development go back to the 90ies, initiated by the private sector, it is meanwhile being promoted by the government and public clients. Various national roadmaps along with the start of BIM pilot projects supported by research projects have been created in recent years. These efforts, started in around 2016 with the big UK initiatives, are now starting to enter daily practice. Switching to BIM-based processes will surely take some time, in particular when thinking about remodeling the existing building stock. Nevertheless, BIM will be a main pillar for improving efficiency and quality in the whole building industry including asset management.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171. © 2021

Due to the novelty of BIM for many stakeholders there is still a lot of confusion about its application. While some will argue that existing building stock is not available as a BIM representation and thus the use of BIM will be limited to new buildings only, there are meanwhile new solutions to create a BIM with help of surveying technologies. For instance, chapter 2.2 gives an example for using laser-scan technology and new data processing algorithms for converting point clouds into more meaningful building elements and properties. Accordingly, it is expected that BIM will play an important role in asset management for new as well as existing buildings. As pointed out in chapter 2.3, the main question again is what “BIM” data is needed to support monitoring and predictive maintenance use cases.

3.2.1 Scope

The scope of BIM is quite extensive as it is supposed to be a digital representation of the building covering the whole life cycle, i.e., from first sketches, design, construction, asset management till refurbishment or dismantling. While the term BIM basically describes a new technology and way of working (in future) it also needs to be implemented in standards and software. Today, beside national or use case specific standards the main open BIM standard is IFC (ISO 16739, [9]), which is developed by the buildingSMART non-profit organization and went through various versions. The latest version is IFC4.3 [13], which will be released in 2022 and will also be published as an ISO standard. IFC4.3 was extended by infrastructure domains such as bridge, road, rail, and ports & waterways[9]. Tunnel is currently development and will most likely be added to the next release.

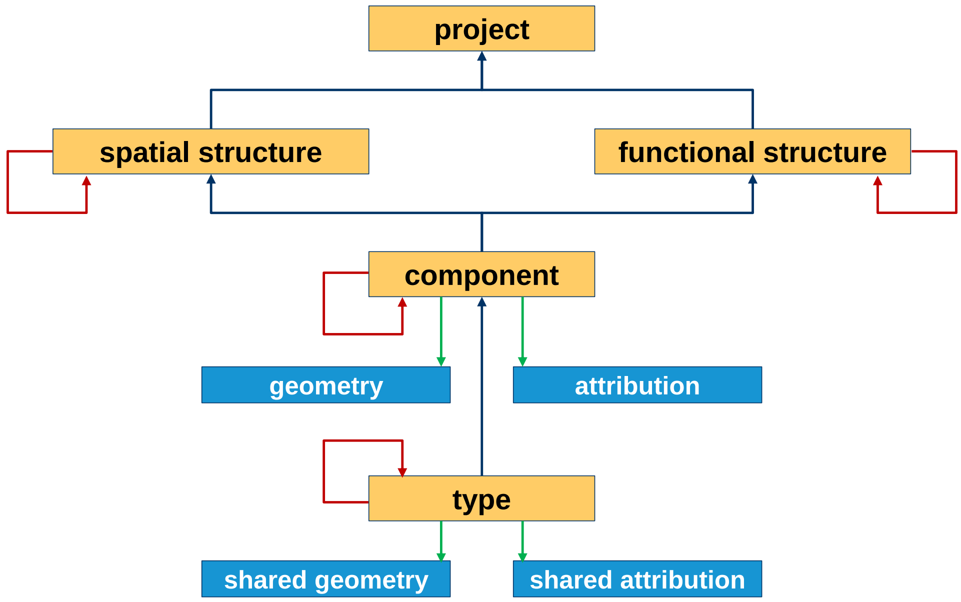


Figure 3.5 – Backbone structure of IFC from T. Liebich/buildingSMART

IFC first of all and foremost covers the physical representation of a building and is subdivided into typical building elements and equipment (components) like beams, columns, walls, radiator, burner etc. It also includes non-physical concepts like spaces and zones, as well as functions, systems, and other breakdown structures or physical assemblies. Other than that, the IFC schema also supports the specification of construction processes including work resources and schedules, which however is not being in focus of current use and software

[9] https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

implementation. Each component, besides having interrelationships to other components as described above, can have additional properties that are either predefined by the IFC standard or specified on a project basis being agreed in use case specifications as described in chapter 2.3.

The main structure of IFC is shown in Figure 3.5, which in fact is very flexible and powerful. Missing components can be defined as a proxy element, like user defined properties. They can further be classified by properties or a reference to an external classification system.

Although BIM and IFC are supposed to cover the whole lifecycle, the focus has been mainly on design processes. Further use cases are construction and hand-over of design or as-built data to facilities management. BIM-based use cases within the asset management phase itself have not been in focus so far. Once all relevant building data has been imported into the asset management system further BIM-based data exchange is rarely happening, because the building data itself typically does not change as long as no major refurbishment is carried out.

No	Use case	Description	Purpose	IFC exchange scenario	Required geometry representation	Required semantic information	Priority	Complexity	MVD
9	Handover to asset management	use the model to support operation and maintenance of the bridge,	use the model for inspection, damage detection, condition rating, condition prediction, maintenance planning	Design application to asset management system	Faceted BRep, Sweep Geometry where suitable (Deck, Rebar, Boring Piles etc)	Classification Material Maintenance information	high	medium	Bridge Asset Management View
10	Handover to GIS for spatial analysis	Handover the bridge design to GIS for environmental analysis and/or asset mgmt.	GIS systems provide functionality for environmental analysis and can be used for asset management	Design application to GIS system	Faceted BRep, Sweep Geometry where suitable (Deck, Rebar, Boring Piles etc), potentially based on alignment	Major design attributes	high	low	Alignment-based Bridge Reference View

Figure 3.6 – Extract from FM-related use cases discussed in the IfcBridge project.

While the IFC extension project for the bridge domain confirms this choice of use cases (see Figure 3.6), it also shows that the scope of IFC with its support for explicit geometry is a good basis for "Operation and Maintenance" (see Figure 3.73).

Beside geometrical data it also enables to define basic asset management data, see for instance *Pset_MaintenanceStrategy*[10], *Pset_Condition*[11] and element grouping by *IfcAsset*[12], the positioning of sensors (see definition of *IfcSensor*[13] as physical object with further properties) as well as to specify constraints[14] to values may representing limiting states that should trigger further control or maintenance activities. It also supports time-stamped data entries, time series, for instance to specify the performance history of various elements.

[10] http://ifc43-docs.standards.buildingsmart.org/IFC/RELEASE/IFC4x3/HTML/lexical/Pset_MaintenanceStrategy.htm

[11] http://ifc43-docs.standards.buildingsmart.org/IFC/RELEASE/IFC4x3/HTML/lexical/Pset_Condition.htm

[12] http://ifc43-docs.standards.buildingsmart.org/IFC/RELEASE/IFC4x3/HTML/lexical/IfcAsset.htm

[13] http://ifc43-docs.standards.buildingsmart.org/IFC/RELEASE/IFC4x3/HTML/lexical/IfcSensor.htm

[14] http://ifc43-docs.standards.buildingsmart.org/IFC/RELEASE/IFC4x3/HTML/lexical/IfcConstraint.htm

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

3.2.1.1 Limitations

While IFC can be used as a reference structure for operation and maintenance, it clearly has its limitations when it comes to the representation of the building condition. Although there is the possibility to specify the condition of components, including the use of external classification codes, it does not enable to specify damages in more detail, like for instance cracks or corrosion of elements, nor does it allow to deal with uncertainties. Gaps are also seen when it comes to monitoring and maintenance plans.

A

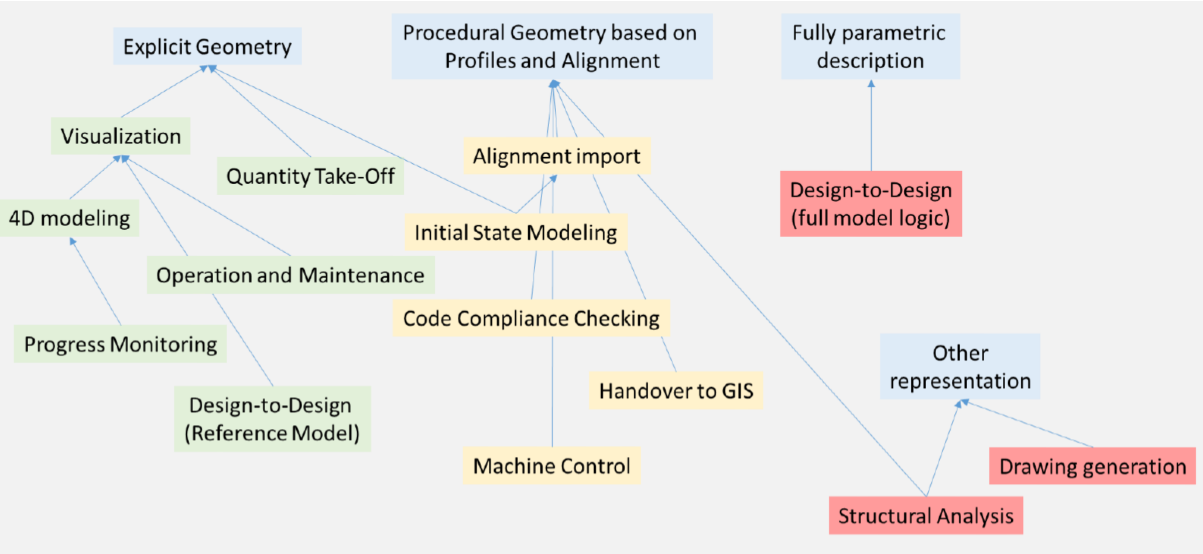


Figure 3.7 – Assessment of BIM-based use cases from the IfcBridge project

3.2.1.2 Overlap with IoT and GIS

BIM/IFC can be seen as a reference to the building and its physical components, including installed sensors that may even be grouped into a monitoring system. It also supports the use of geo references for proper positioning of the building, thus providing a link to GIS.

All components will also get a globally unique identifier, which however follows a specific coding scheme that is not compatible with the more flexible URI approach know from the web technology. Such identifiers enable to uniquely identify the asset and individual components and thus can be used for linking different data sources as for instance done in the ICDD standard (ISO 21597).

3.2.2 Technology

IFC follows the object-oriented modelling paradigm. Its main specification is based on the STEP family of standards (ISO 10303), which includes the EXPRESS modelling language (part 11) and the STEP physical file format (part 21) for data serialization. IFC has also adopted various definitions like for instance parts for the geometrical and topological representation.

The schema specification is using a quite deep inheritance structure as well as objectified relationships making the use and implementation of IFC more challenging. It also specifies quite a few consistency constraints, which have been softened from release to release but still are a noteworthy difference to many other standards. Those constraints include simple settings like about mandatory attributes, cardinality of references or inverse relationships but

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171. © 2021

also more complex rules and even functional description. Such specification should ensure the consistency of the data set, but also increase implementation efforts and complexity of the standard. Many of those constraints cannot be represented in other modelling languages like for instance XSD or OWL and thus they offer a less rich representation than IFC.

IFC is also published as XSD and OWL schema representation with associated XML and RDF serialization formats. They enable alternative implementations and use cases, also using more common implementation tools that are in particular interesting for light-weight downstream applications trying to provide simple and cost-efficient services. The research community was also pushing the development of an ifcOWL representation, which again enable a new kind of use cases.

3.2.3 Status and future developments

IFC started in 1996 and since then has gone a long way meanwhile being the only open BIM standard with international use and a variety of IFC implementations. Latest version of IFC at the time of this writing is IFC 4.3 with above mentioned extensions for the infrastructure domain, including support for bridges. Support for tunnels will follow most likely with the next release of IFC.

IFC provides a good basis for quite a number of typical AEC use cases, although it may not cover all features supported by a specialized application. It thus focuses on most important data, which in many cases will be enriched by additional agreements to support national or project specific requirements. This is a very powerful feature of IFC and enables to adopt to specific markets, including the risk of having different IFC dialects at the end of the day. Discussion about the future of IFC will most likely further strengthen current flexibility but combined with a common approach based on the bSDD technology for dealing with national classification codes and own properties. This will mitigate the risk of having incompatible IFC solutions, which however is still seen as one of the biggest challenges for the future.

Existing implementations of IFC do not cover all functionalities embedded in the schema. Scheduling, constraint management and structural analysis are for instance not yet supported by software tools. In general, quality of IFC implementations and compatibility between different software tools is also an important topic when it comes to implementation of BIM-based use cases. In fact, there is a bunch of aspects like used IFC version, version 2x3 is for instance still heavily used in practice, used software tools and releases for BIM authoring but also data quality control and data storage, experiences of involved users and last not least the quality of project agreements as explained in chapter 2.3. Many BIM-related developments are happening now in the AEC industry so that BIM and IFC will fulfil an important role most likely in nearly all coming digitalization scenarios.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

3.3 **Geospatial Information System (GIS)**

3.3.1 Scope

GIS supports topics such as design and planning of buildings, spatial understanding, integration into the environment, visualization, and public relations. Using various GIS data, BIM projects can be extended to include the geographical context, such as: properties, protection zones, infrastructure, terrain, and 3D city models. Within the field of infrastructure projects, the most relevant GIS models are the digital terrain model, protection zones, administrative borders, digital elevation model, the digital landscape model and the digital surface model. An example of a digital terrain relief and surface model is shown in figure 15.

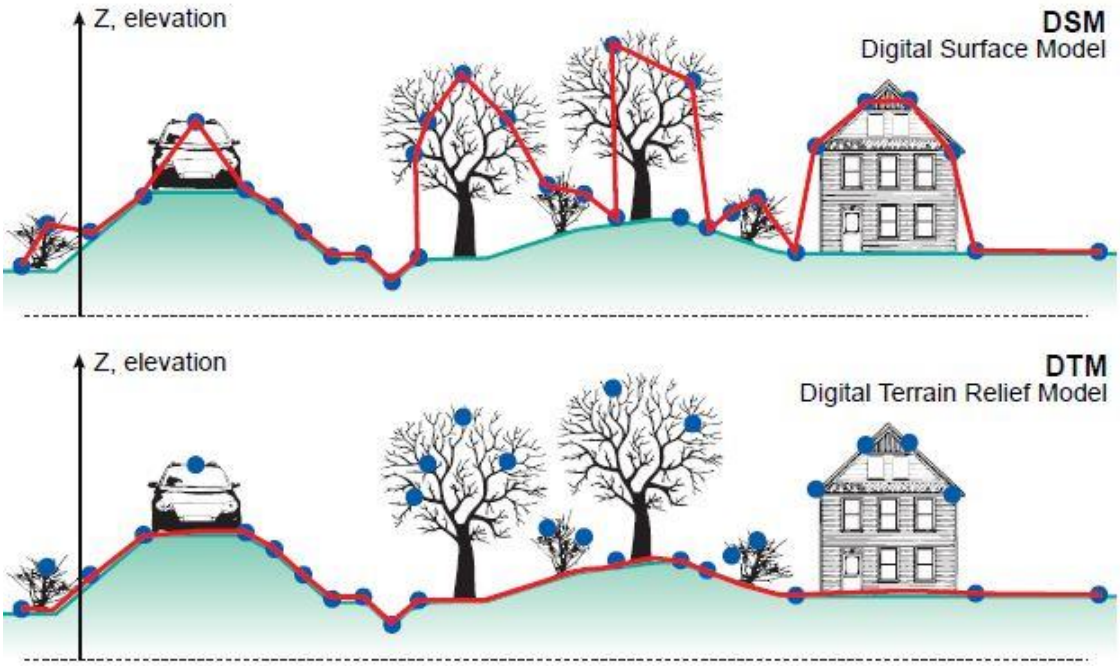


Figure 3.8 – Digital Terrain Model and Digital Surface Model[15]

The GIS area focuses on semantic objects in a spatial and thematic context. Infrastructure projects that are to be constructed in the European union must adhere to the INSPIRE directives for data exchanges (Noardo et al, 2019).

[15]http://www.charim.net/datamanagement/3214.04.2022)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

3.3.2 Technology

In the GIS area, formats are standardized by the "Open Geospatial Consortium (OGC)" and the "World Wide Web Consortium (W3C)". The defined open standards of the OGC and W3C are extended by defacto standards used in the industry. Here it is possible to distinguish between vector data and raster data (GISGeography, 2022, [37]). GIS data, such as aerial and satellite images, digital elevation models, 3D city models, landscape models and extracts from the real estate cadastre are standard products of the geodata-holding agencies (ADV, 2022, [35]).

A differentiation of the data sets into vector and raster formats must be made, since these are analyzed, processed, and presented differently. For example, aerial imaging data can’t be saved as xml files.

The most common file formats for GIS data are shown in table 2

Table 3.1 – Common formats for GIS data

Vector formats	Raster formats
.dxf - Drawing Interchange Format (open)	.tif - Tagged Image File Format (open)
.kml - Keyhole Markup Language (open)	.xyz - ASCII grid (open)
.gml - (City) Geography Markup Language (open)	.jpg - JPEG File Interchange Format (open)
.shp – Shapefiles (proprietary)	
.geojson – GeoJSON (open)	
.xml - Extensible Markup Language (open)e	

More detailed information regarding the different file formats used for GIS applications can be found in Chapter 7.2

GIS data is often also made available via standardized services, such as "WCS - Web Coverage Service", "WFS - Web Feature Service" or "WMS - Web Map Service”. Thus, information can be queried, processed, modified, or presented in real time. In BIM authoring software, however, the possibility of connecting a corresponding service is usually missing. For this reason, the use of additional GIS software is still necessary.

The creation of point clouds is part of the step of “data acquisition” as described in chapter 2. A conventional point cloud, generated by a terrestrial laser scanner, is composed of several different raw scan data. These scan data are initially available without reference to the structure or to a superordinate system and are, at best, pre-registered. Only through registration do they receive an assignment between the scan points. There are various procedures for this, such as: Cloud2Cloud and layer matching. From now on, one speaks of a registered point cloud, which can now be assigned to a local reference system (e.g.: plant engineering) or a superordinate coordinate system (geodetic coordinate system) by means of control points. This then corresponds to a georeferenced point cloud.

Typical application areas for the use of point clouds are:

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

- As-built documentation ("as-is" or "as-built") - for the derivation of geometries & models. Figure 16 shows a point cloud suited for the documentation of a train station.
- Construction progress documentation ("as-performed") - for documentation and monitoring of construction progress
- Construction monitoring - monitoring of engineering structures

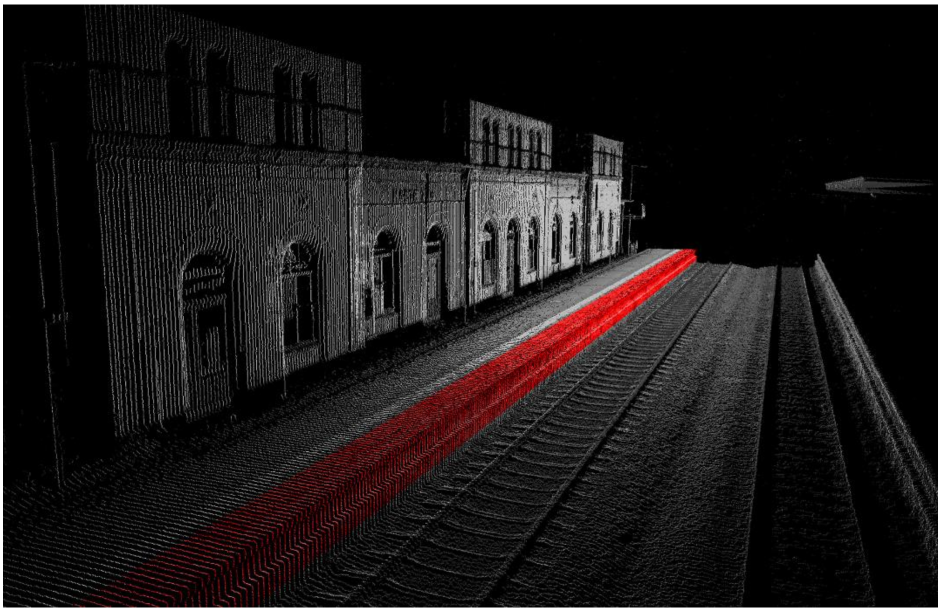


Figure 3.9 – Point cloud of a train station and rails[16]

Possible sources for point clouds are:

- Surveying and engineering offices (terrestrial laser scanning, UAV - aerial photography, mobile mapping)
- Geodata holding agencies (e.g., The Federal Agency for Cartography and Geodesy, Germany).

Point clouds can be stored in different formats. This usually affects the processing time and the required storage capacity. The following point cloud formats are particularly frequently used:

- E57 -vendor-neutral point cloud format
- ASCII (.xyz, .csv, .txt) - simple data structure, but high storage requirements (without intensity or colors)
- ASCII (.svy) - with additional intensity (IR) or color (RGB) values if necessary
- ASPRS standard LiDAR point cloud format (.las)
- Proprietary formats: Leica (.pts, .ptx), FARO (.fls, .fws), Riegl (.3dd), Z+F (.zfc, .zfs)

[16] https://www.ipm.fraunhofer.de/de/gf/objekterfassung-laserscanning/anw/bahnmesstechnik/lichtraumprofil/jcr:content/stage/stageParsys/stage_slide_61835180/image.img.jpg/1611831209859/3D-punktwolke-bahnhof-2000px-72dpi.jpg (14.04.2022)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

3.3.3 Status and future developments

Since the standardisation of GIS data is already well advanced, only a small selection of currently used standards will be listed in this chapter. A more detailed list of GIS standards can be found in Chapter 7.2 and on the OGC website (https://www.ogc.org/docs/is).

The INSPIRE directive aims to facilitate data exchange between actors of different European countries and cover 34 spatial data themes required for different environmental application (EU INSPIRE, 2022, [36]). A common spatial data infrastructure is needed since infrastructure projects often cross borders and therefore rely on GIS data from different sources, such as elevation models or cadastral data.

Open standards such as LandInfra by the OGC attempt to bridge the gap between the BIM and GIS world. However, because of its complexity, the knowledge of the LandInfra is still very limited. Further research is needed to fully explore its potential to connect BIM and GIS (Sebastian et al. 2020, [25]).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

4. Management – Data interoperability framework

The previous chapters gave an overview on standardization related to data flows and the meaning of shared data. Those agreements being published as an open standard are the basis for data interoperability and semantic integration. Looking into monitoring scenarios it is also clear that different stakeholders and tools are involved that all need to work together. However, they also have different views on the asset that is to be monitored and maintained.

It is commonly recognized that there will not be a single standard covering all aspects related to the lifecycle of our buildings and its environment. Consequently, established standards coming from different domains, being maintained by different standardization bodies, and additionally being based on different technologies must work together in a heterogeneous and distributed environment. Well-known ICT organizations, namely IEEE, ISOC, W3C, IETF and IAB, recently published the "OpenStand Principles" addressing topics related to standardization and its joined use. Further information about these principles is given in the Appendix A.

This chapter is addressing the question how all views about a building can be brought together in an ICT environment that enables data interoperability and semantic integration.

4.1 Overall goals by EC

As envisioned by the European Commission (EC), in a foreseeable future, all building and civil infrastructure assets in the EU would have a digital logbook. The Digital Building Logbooks (DBLs) keep up-to-date information of the assets throughout their lifecycle, and enable the asset owners, end-users and authorities to access, enrich and use the information for monitoring the assets' structural health conditions, minimizing their carbon footprints and energy consumptions, and optimising their performance towards the users' needs. The EU-wide harmonized set of logbook data protocols would enable interoperability and inclusion of external databases.

The development of a semantic data model and a data management plan as well as of a standardized approach for data collection, data management and interoperability including its implementation framework, are key to achieve the ambition as defined by the EC (GROW/2021/OP/0014: Technical study for the development and implementation of digital building logbooks).

4.2 Challenges and linked data principles

Various research and development projects dealt with asset management topics. Bakker et al. (2019, [39]) and De Kleijn et al. (2018, [20]) very well described the challenges for Life Cycle Management of infrastructural assets, which in general is highly dependent on reliable data. Consequently, to make the right decisions the data needs to be findable, up-to-date and understandable. They for instance conclude about the current situation: *".. in practice the data about infrastructural assets is seldomly stored in one place and structured following a interoperable standard. Data about these assets has been gathered over long periods of time according to different protocols for various purposes. Much of the data is unstructured and only available as text, without structured semantics. Even if data is structured and semantics are properly defined, there can still be different semantic conventions in different datasets. Geometric representations vary, system boundaries are often not uniformly used, and even unique coding systems can be inconsistent over the datasets. Attempts to re-design data*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

structures to a "one fits all" solution often fail, because individual processes have their own standardized semantics throughout the sector."

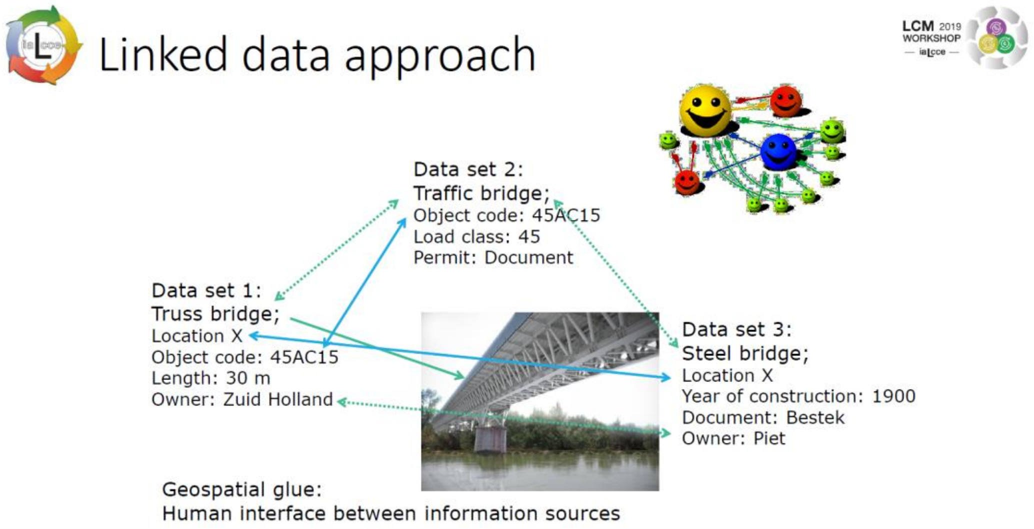


Figure 4.1 – Challenge of having various datasets about the same asset and the idea of linking those datasets together (from Bakker et al. 2019, [39])

Regarding a possible solution they also describe practical constraints coming from software implementation and then follow a realistic approach that is trying to deal with grown systems and imperfections that exists in practice. They further explain: ".. *software systems often dictate how data is stored. Most software packages are not per se designed to correspond with the specific data the organization needs. Integrating data structures therefore often result in replacing existing software systems. Choosing a "one fits all" approach might theoretically sound ideal, but in practice creates a huge barrier to make the step to Life Cycle Management. An alternative methodology is to accept the world as it is, and deal with it in a structured manner."*

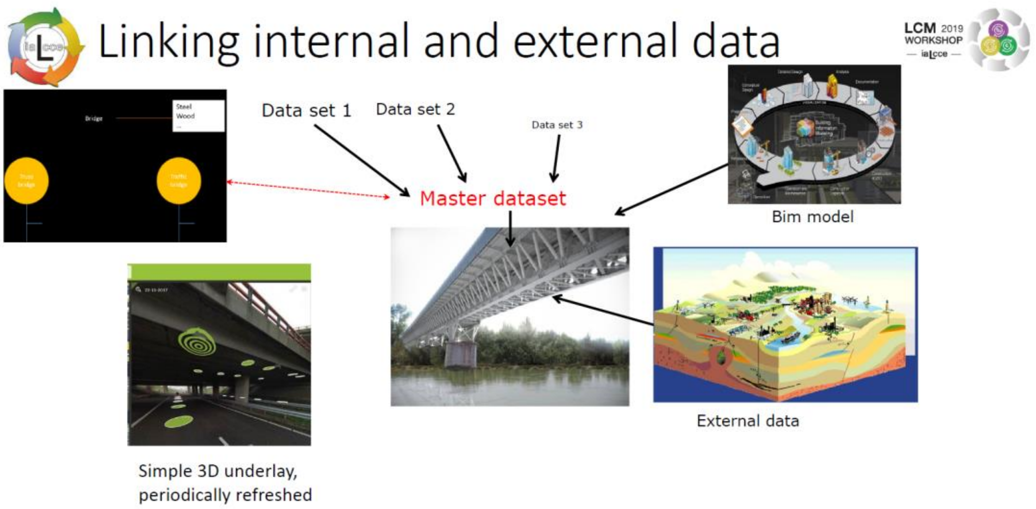


Figure 4.2 – Overall approach to link all relevant datasets including e.g. BIM (from Bakker et al. 2019)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

A main challenge is therefore to find a solution that enables to integrate internal and external datasets, also considering their use. Similar to many other researchers they propose to use linked data technology, which in their case was combined with an easy-to-use 3D viewer and machine learning algorithms, also showing the potential of having access to all relevant data in same manner. Their work showed that new knowledge can be generated by combining the individual datasets. Finally, they published an interesting list of conclusions regarding linked data and further applications on top of it:

- *The "brown field approach", taking the world "as is" as starting point for further development, can be applied well to data science;*
- *Linked data can help to combine data sources to one data cloud that can be questioned and analysed without changing the initial data sets;*
- *A 2D 3D interface can be an effective user interface to question the data cloud*
- *Linked Data is a good starting point for data integration*
- *3D Linked Data principles can help to combine data from internal and external datasets*
- *Machine learning principles can be used to:*
    - *to check and improve datasets (flaw detection);*
    - *support data integration;*
    - *inform the user about the likelihood of the data being correct;*
    - *combine the available Linked Data to new information*

4.3 Linked asset data and semantic web technology

The principle of model linking has been extensively discussed in the AEC industry. A noteworthy approach is the Multi-Model Container (MMC)[17] that was also proposed as an international standard published by buildingSMART. Underlying technology is XSD/XML and its focus has been on linking datasets that are stored in different file formats. The specification itself is generic and does not provide further link semantics for specific use cases. This is to be agreed in separate specifications or standards like for instance the DIN SPEC 91350[18], which covers the link from BIM to the bill of quantities. While no further constraints are specified regarding its implementation a simple and very pragmatic implementation was to tie everything together in a ZIP package that includes all datasets and additional link data.

Work on MMC was later combined with a similar effort based on Semantic Web Technology. The result of that effort is the Information Container for Data Delivery (ICDD, ISO 21597, [3, 4]). There two parts of this standard:

- Part 1: Container with payload of linked documents, meta data as LD
- Part 2: Provision of nine semantic document links (potentially deep links)

While ICDD is based on Semantic Web Technology, both specification of ontologies and serialization of the datasets, it is still a container format for the exchange of files. This however will limit its use in distributed, web-based environments, where web-enabled solutions based

[17] https://github.com/BuildingSMART/MMC

[18] https://www.din.de/resource/blob/333980/e807743c462a0b9e4c6d24704b8fb837/broschuere-din-spec-91350-data.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

on Semantic Web Technology can speed-up its implementation and use. Such containers however might be embedded into the Linked Data Platform[19] thus solving such constraints.

The most important requirement for linking multiple databases is the existence of identifiers or primary keys in each of the databases so the information of interest from multiple sources can be selected in one query. Semantic Web Technology, beside allowing to describe the meaning of data thanks to the Web Ontology Language (OWL, [14]) and to serialize dataset for instance in the RDF format, includes a data query and manipulation language (SPARQL, [16]) as well as solutions for defining constraints to the dataset (SHACL, [17]). Those standards are published by the W3C organization, which is setting the standards for the internet, like for instance the Hypertext Markup Language (HTML) or the HTTP protocol (together with IETF). Accordingly, those W3C standards dealing with semantic integration are expected to play a crucial role for any distributed, web-enabled solution. Accordingly, this set of standards was chosen by many approaches for this type of application scenarios.

4.4 CEDR-INTERLINK approach

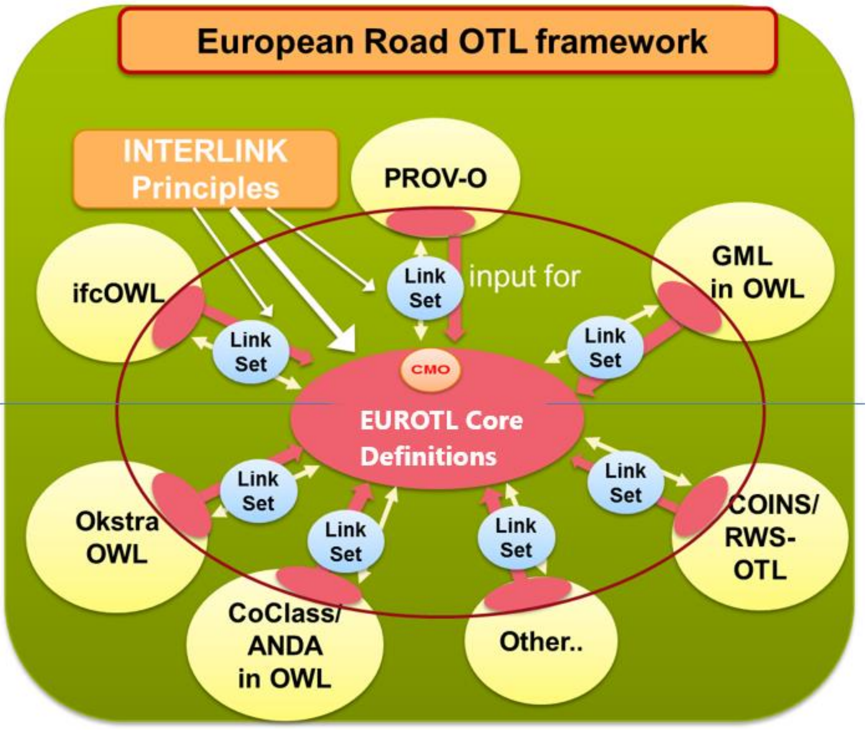


Figure 4.3 – Framework developed by INTERLINK for a European Road OTL (from www.roadotl.eu)

The INTERLINK[20] project, initiated by the Conference of European Directors of Roads (CEDR[21]), was dealing with asset management of roads and was faced with quite similar questions as IM-SAFE, namely the integration of different data sources. The use of Semantic Web Technology was part of the project proposal, but all potential data sources had to be combined into a *European Road Object Type Library* framework (Figure 4.39).

[19] https://www.w3.org/TR/ldp/

[20] INformation managemenT for European Roads using LINKed data

[21] https://www.cedr.eu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

The framework developed by INTERLINK is following a hybrid approach, that on one hand is trying to harmonize semantic agreements by providing a common core covering basic definitions supplemented by common modelling guidelines that should be followed when extending the framework. On the other hand, it enables to integrate and reuse existing ontologies that are for instance national standards. The framework itself is divided into four parts (see also https://www.roadotl.eu/static/eurotl-ontologies/index.html):

- **European Road OTL Core Definitions**
  "This is an ontology that serves as a hub to which more specific domain ontologies may be linked. This ontology covers highly reusable definitions such as provenance, quantities and units, temporal and spatial locations, transport networks, basic support for asset lifecycle and also main asset types and properties as needed for sharing asset lifecycle data."

- **Domain ontologies**
  "which are ontologies that already exist in some form and that covers specific needs for specific use cases. These ontologies are linked with the core definitions using Linking Rule Sets."

- **Linking Rule Sets (LRS)**
  "which are ontologies or mapping descriptions with the only purpose to provide the relationships between elements in ontologies (e.g. between the domain ontologies and the European Road OTL Core Definitions) in a machine-readable way"

- **Modelling & linking guide (MLG)**
  "as provided by INTERLINK which contains recommendations for how to model the above ontologies to enable a smooth integration into the framework". This proposal was letter brought into the Semantic Modelling and Linking Guide (preEN 17632)

The object type library at the time of finishing the project was defined by the ontologies listed in Table 4.1. Those ontologies have been integrated by using linking rule sets being itself defined in a separate ontology and thus follows the idea of model linking as defined by the MMC and ICDD approach.

The INTERLINK framework is quite easily extendable as it enables to add further *rdf:type* links in order to classify data as required by specific use cases or national codes. Such "dynamic typing" while runtime is not limited to single classification, and might be established manually by the user or, preferably, by linking rule sets that encode the classification criteria based on the object characteristics such as for instance material type, assignment to damages or others. Such dynamic typing, which may even include equivalence (or close match) relationships between different classification systems, is shown in Figure 4.521. The modelling principle itself is shown in Figure 4.421.

The INTERLINK framework was designed to be extendable. This is necessary to support new use cases that may not yet be obvious and surely will arise in the future. To avoid conflicts when integrating or linking new datasets a modelling guideline was defined that should help to make the right modelling decisions. This guideline was then further developed and is now available as the standard preEN 17632, which is discussed in more detail in the next chapter.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Table 4.1 – List of ontologies being part of the EUROTL framework

Prefix	Name
eurotl	EUROTL: The European Road Ontology
prov-o	PROV-O: The Provenance Ontology
skos	SKOS: Simple Knowledge Organization System
am4infra	AM4INFRA: Asset Management Approach for Transport Infrastructure Networks
am4infra--eurotl	AM4INFRA--EUROTL: Linking Rule Set between AM4INFRA and EUROTL
geo	GeoSPARQL: A Geographic Query Language for RDF data
geosparql--eurotl	GEOSPARQL--EUROTL:LinkingRule Set between GeoSPARQL and EUROTL
IFC4x1_Final	IFC4X1_Final: ifcOWL provides a Web Ontology Language (OWL) representation of the Industry Foundation Classes (IFC) schema
IFC4x1_Final--eurotl	IFC4x1_Final--EUROTL:Linking Rule Set between IFC4x1_Final and EUROTL
rd:,tn:,net	rd:INSPIRE data specification for Road networks tn:INSPIRE data specification for Transport networks net:INSPIRE Generic Network Model
inspire--eurotl	INSPIRE--EUROTL:Linking Rule Set between INSPIRE transport networks and EUROTL
iso19148	ISO19148:Linear Referencing
iso19148--eurotl	ISO19148--EUROTL:Linking Rule Set between ISO19148 transport networks and EUROTL
library	LIBRARY:Ontology metadata

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

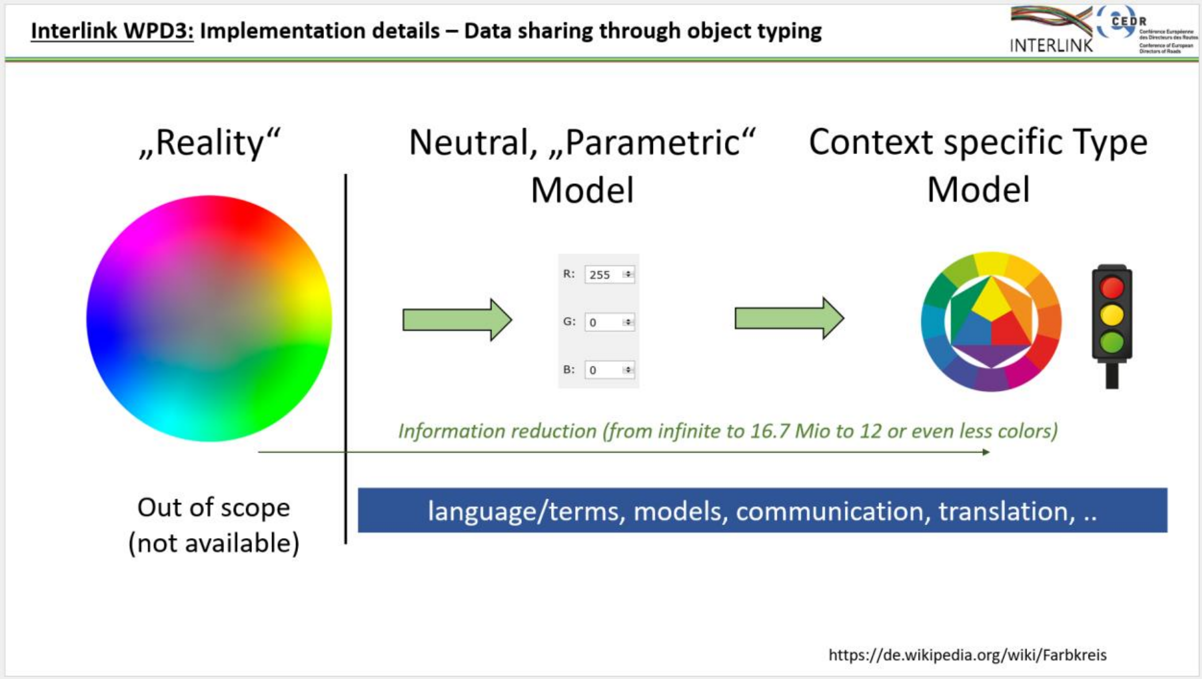


Figure 4.4 – Principle of data modelling using a neutral representation of the reality that can later be “typed” according to specific use cases such as the 3 traffic lights red, yellow, green.

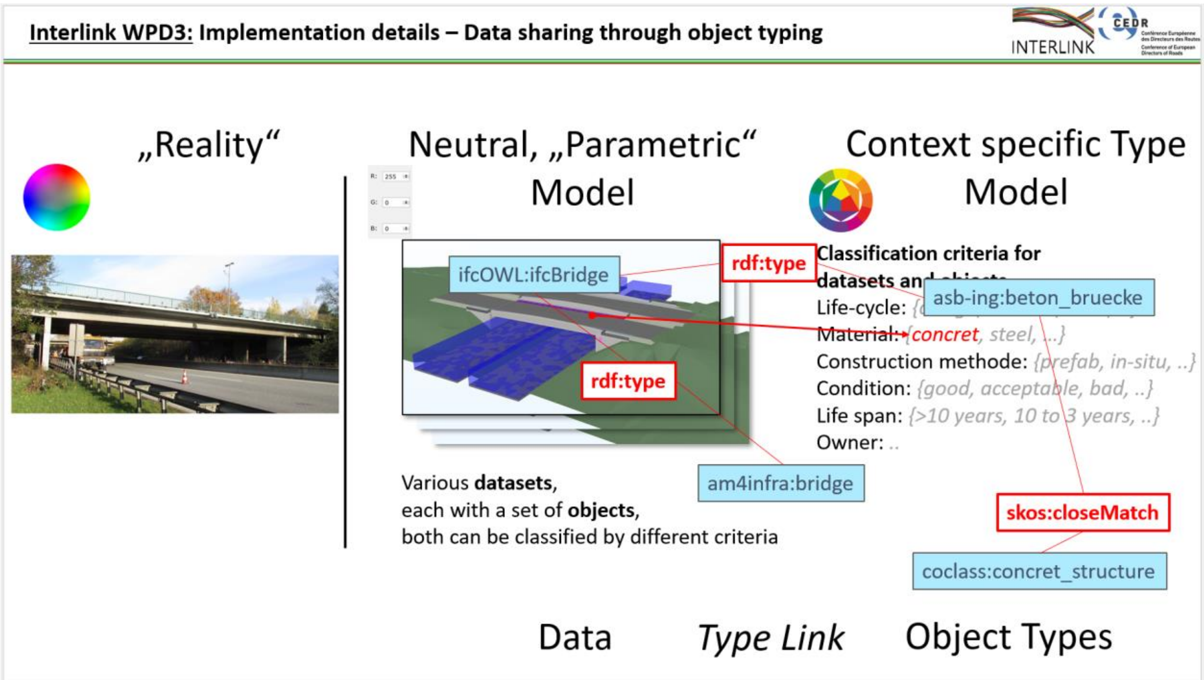


Figure 4.5 – Adoption of the modelling principle to a bridge use case with BIM/IFC being the neutral representation of the building that is further classified by AM4INFRA, the German ASB-ING and the Swedish CoClass.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

4.5 Semantic Modelling and Linking

Semantic Modelling and Linking, or short SML, is a European standard being prepared by CEN TC442/WG4/TG3 referred to as prEN 17632:2022. The standard is expected to become final in 2022. This standard provides both abstract and concrete modelling patterns applying W3C Linked Data/Semantic Web technology in the Built Environment.

Linked Data/Semantic Web technology is seen as key mechanism to provide FAIR as introduced in 1.4 and shown in relation to Semantic Web technology in Figure 4.62. The FA follows from its ‘webbased-ness’, the I from its standard formats and query language (RDF-XML, Turtle, JSON-LD and SPARQL) and the R coming from its ontologies that are again defined by standard languages (RDF, RDFS, OWL and SHACL).

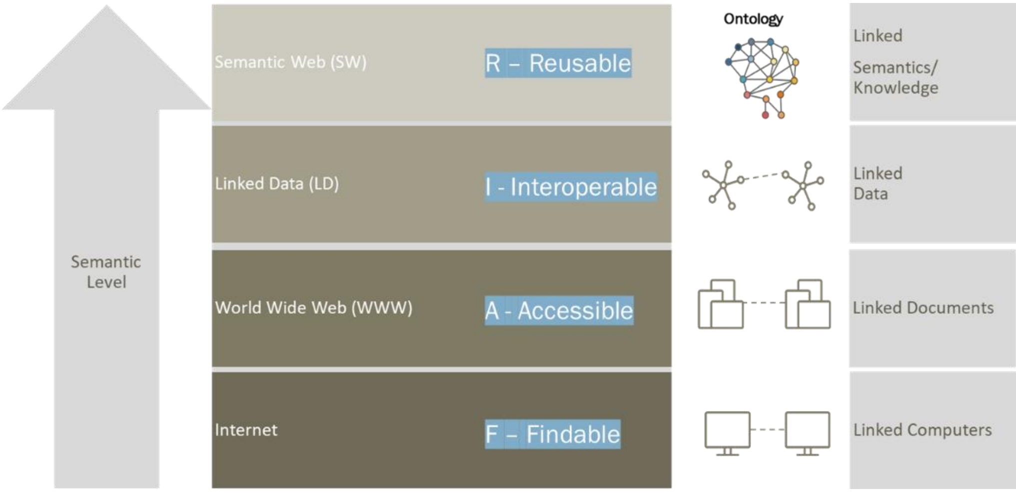


Figure 4.6 – W3C Linked Data/Semantic Web tech supporting FAIRness for data

The SML approach is strongly related to work going on in:

- W3C Linked Building Data (LBD) community group (SML is based on variants of the OPM-Ontology for Property Management); and
- buildingSmart International (bSI) related to running ISO 12006-3 and JSON-based buildingSmart Data Dictionary (bSDD) experiments.

SML first provides a common data language that gives ‘bindings’ towards SKOS, RDFS, OWL and SHACL used, depending on the modelling purpose (or in combination). SML also provide essential modelling patterns in the form of a standard top level ontology with predefined relations including decomposition (“the hasPart” relationship) (Figure 4.73).

Furthermore, SML recommends modelling patterns for generic modelling issues like:

- Identification (URI-strategy),
- Naming & Annotation,
- Enumeration Types (like for code lists),
- Typical decomposition (via constraints) obtaining meronomies,
- Quantity kinds & Units, and
- Complex Properties.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

And finally patterns for supporting basic "systems engineering" capabilities like:

- Planned versus Realized; and
- Functional ("roles") versus Technical ("kinds").

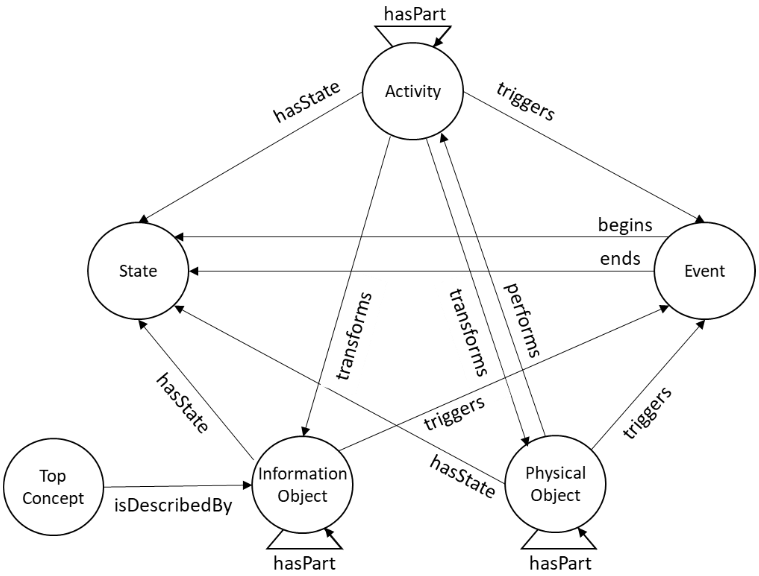


Figure 4.7 – SML Top Level ontology

Recently, a New Work Item (NWI) has been proposed for an SML Part 2 dealing with more domain-specific extensions including:

- Support for distinction between spatial regions and real ("tangible") objects, the latter being discrete or continuous ("bulk matter").
- Support for the materialization of physical objects, adding generic chemistry aspects directly relevant for the built environment dealing with concrete, steel and asphalt.
- Support for the interaction between objects including connections, interfaces and ports (parts of objects where such interaction can take place). Interaction being defined as activities where material, information, energy/forces are transferred.
- Support for the definition of requirements, unstructured and structured, coming from client needs, laws and regulations or sector recommendations.
- Support for implicit groups having no explicit members (to model situations like "all main girders of some steel bridge").
- Support for the explicit modelling of measurements reusing the existing W3C Semantic Sensor Network (SSN)/Sensor, Observation, Sample, and Actuator (SOSA) ontology, incl. extended QUDT support, and
- Support for geospatial geometry (OGC GeoSPARQL/wgs84).

Many of these modelling aspects are highly relevant for monitoring and asset management. Thus, SML provides not only modelling guidelines but also basic semantic agreements to be used in IM-SAFE related use cases. It is a generic, future proof method, based on W3C linked data standards, providing a base for Asset Life cycle Information Management (ALIM)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

modelling and linking the many relevant data sources along asset's life cycle and supply chain as shown for instance in the example shortly described in the appendix A.

4.6 Emerging Digital Twin concept

Digital Twin is essentially describing a concept for having a digital, up-to-date representation of the physical building or asset. While the idea is clear and compelling, the implementation is still subject of research activities and essentially is facing above mentioned data integration challenges.

Developments related to infrastructure are for instance reported in Sebastian et al. 2021 [21], who describes the vision of Digital Twins for transport infrastructure and proposes a conceptual recommendation on standardisation of open interoperability in Digital Twins. In this work the Digital Twin concept includes the use of Cyber-Physical Systems (CPS), which allows ".. *remotely sensing, real-time monitoring and controlling of devices; and therefore, provide a direct integration and synchronization between the physical and virtual worlds. A Digital Twin leverages the synchronicity of the cyber-physical bi-directional dataflows, and it can represent the physical reality at a level of accuracy suited to its purpose.* "

Although a common definition is yet to be established, there is a shared understanding that very well fits to the requirements as defined in the IM-SAFE project. *"A Digital Twin is a virtual model that is connected to the physical object – allowing for monitoring, analysing, predicting and controlling the condition, behaviour and performance of the object."* Sebastian et al. furthermore describes two ways of its relationship with simulations:

a) "Digital Twin as the model that represents a system upon which varieties of simulations can be based; and
b) a Digital Twin as the simulation of the system itself.

As such, a Digital Twin is more than just a virtual model that contains static and dynamic data. It also integrates algorithms that describe the physical counterpart and can support decision-making on actions in the real world based on processed data and simulations."

The three main aspects of Digital Twins within the context of IM-SAFE concerning inspection, monitoring and maintenance of transport infrastructures are:

1. Data management, which addresses quality, interoperability, sharing and processing of the data on which Digital Twins are based.
2. Integration of expert models, which addresses fidelity of the algorithms in the knowledge domain where Digital Twins are used.
3. Simulation-based decision-making, which addresses visualization of the data and objects along with decision support for the experts and end-users, as well as automated decision-making and feedback to the physical objects through actuators.

This deliverable focuses on the first, especially the standardisation for the interoperability of data on which Digital Twins are based.

4.6.1 Progress of Digital Twin beyond BIM

Typically, BIM consists of a three-dimensional model of a building or civil infrastructure asset, containing the information of the properties of the objects or elements in the model. A BIM model as semantically rich digital representation of a physical object provides a good basis for setting up a Digital Twin. Relevant BIM-based information to be integrated and enriched in a

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Digital Twin can include, for instance, geometric models and changes in the building layout; monitoring data over the condition and degradation of structural components; and occupancy, usage, operational and performance information of the building or civil infrastructure (see also chapter 3.2).

Extending BIM to Digital Twin requires the means and the solutions for capturing and processing real-world data and feeding it back into the model to create a circular information loop. An important aspect here is the management of information throughout the asset's lifecycle. Thanks to the bi-directional dataflows, Digital Twins can overcome the limitations of BIM to capture dynamic information, update the digital model automatically, and perform simulations for decision-making throughout the lifecycle of a building or civil infrastructure asset. In some industrial sectors, like airplane engine manufacturing, the Digital Twin of the engine is updated immediately because the engineers need to know the exact condition of an engine in real-time. However, with civil infrastructures, the crucial changes on the assets usually occur in a much longer period. The update rate and frequency of Digital Twins should, therefore, be adjusted to suit its application purpose.

4.6.2 Recommended progress towards Semantic Digital Twins (SDT)

Semantic Digital Twins (SDT) goes beyond BIM-based Digital Twins and enables a broader data interoperability of Digital Twins. SDT breaks through the constraints of static and closed data with recursive interoperability issues and opens the way towards a Linked Data paradigm. It facilitates linking the data derived from Internet of Things (IoT) sensors with semantic data as input for Artificial Intelligence (AI) algorithms to predict the performance of civil infrastructures. By relying on Semantic Web and Linked Data solutions, SDT can derive the data from Internet of Things (IoT) sensors, comprise the whole asset lifecycle information, and make it machine-interpretable for AI.

With Semantic Web and Linked Data technologies, various data can be combined by representing the information in structured graphs. This approach allows for efficiently linking and sharing of information of entirely different natures, for example BIM data, Geographic Information System (GIS) data, Asset Management (AM) data, material repositories, regulation data, cadastre data, and urban data. Moreover, with the increased applications of sensing technologies in the built environment, Semantic Web and Linked Data technologies are bringing an added value to enrich the existing information models with sensing and monitoring data.

Semantic Web introduces new ways of managing data and metadata and maintaining a higher order of logical and conceptual schemas. Properties and values of assets can be defined by shared schema or ontology. An ontology is the explicit formal specification of concepts and their relations in a domain. When used in data management, ontologies can guide a Digital Twin to validate the domain data models by allowing interactions between various data which is held in different formats. As ontologies are semantically richer than databases, a Digital Twin ontology model will maintain the semantics of data and the concept definition throughout the asset's lifecycle.

An SDT ontology model needs to contain the conceptual knowledge from a certain domain of application. In the built environment, until now the role of Semantic Web has only been considered as complementary or supporting to BIM. This paradigm will change with the development of an SDT where an ontology approach is considered more suited for the future compared to older standard file formats. The Semantic Web based approach has, therefore, become part of the UK's government strategy for defining and developing BIM Level 3 and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

beyond. As such, it is logical to adopt the Semantic Web approach to progress beyond BIM towards SDT.

To facilitate the widescale adoption of SDT, this approach is demanding the FAIR principle as shortly described in the introduction, i.e., the information in an SDT must be findable, accessible, interoperable, and reusable (FAIR) by relevant users and organizations in compliance with the EU General Data Protection Regulation (GDPR).

Sebastian et al. (2021, [21]) also describes the implementation on a bridge, the Ijssel bridge in The Netherlands. Further details about this example are added in the appendix.

Latest development within CEN is the proposal for a new working group within TC442, WG9 on “Digital Twins”. Beyond data aspects there will also be standardization of software functionalities and monitoring and control aspects.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

5. Conclusion and recommendations

This deliverable provides an overview on data integration and interoperability topics related to monitoring and asset management use cases of infrastructure buildings. It addresses in particular interoperability between Internet of Things (IoT) – delivering the measurements that describe the current state of the building –, Building Information Models (BIM) – being the digital representation of the as-designed or as-built asset – and Geographic Information Systems (GIS) – being the digital representation of the infrastructure network. Thereby, the focus is on available open standards. The overview about IoT, BIM and GIS presented in this deliverable makes clear that:

- There is not a single standard that covers all data being of interest for monitoring scenarios. It is thus expected that a combination of ICT standards will be needed to support semantic data integration.
- Likewise, there is not a single technology being used by those standards. Thus, implementation needs to be based on a neutral solution being able to integrate and manage all relevant data.
- Existing standards like IFC, GML, SAREF etc. provide a good basis for semantic data integration, but do not yet cover all aspects related to monitoring and maintenance scenarios. Thus, existing standards need to be extended according to requirements that must be derived from monitoring use cases.
- Further classification systems are likely to be needed to deal with national or project specific requirements. Thus, a solution is needed that can further classify the neutral and standardized representation of the building.

Beside standards for describing the data about a building and its condition, the deliverable also explains the use of that data and how to describe work- and dataflows in monitoring and maintenance scenarios. Such information provides context information about the use of the data and is seen as equally important for data management, which in general is faced with the challenge to combine distributed and heterogeneous databases. The deliverable describes existing standards covering those aspects in context of BIM developments. Main conclusions are:

- Semantic integration requires the documentation of monitoring processes in order to identify the data that needs to be managed and shared between users and systems. Accordingly, a good ICT-related documentation should be the basis for further standardization work.
- Many efforts have been put into proper documentation of BIM-based working, and additional standards have been developed. Those standards are generic to great extent and thus should ideally be reused for documentation of monitoring processes.

Last not least, implementation of data integration remains a big technical challenge due to the diversity, heterogeneity, and distribution of the data. The main conclusions related to implementation are:

- A single system or solution will not be able to deal with all monitoring and asset management use cases. Accordingly, implementation must be based on open, robust and easily extendable solutions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

- Linked Data and Semantic Web technology provide a proper answer to identified challenges. Its applicability has been demonstrated in research and pilot projects. Because of its generic nature as well as its expected importance for the management of all kinds of web-based data this approach is a promising basis for further standardization.
- The use of Semantic Web technology is very flexible and offers different ways for implementation. While such flexibility is a good thing in general, it can unnecessarily complicate data integration due to harmonized modelling approaches. Accordingly, additional guidelines such as the SML can help to simplify data integration by harmonization of modelling styles and, as far as possible, reuse of existing basic ontologies.
- Further standardization should be based on existing developments and findings. Worth mentioning are the CEDR-INTERLINK framework and ongoing research related to the Semantic Digital Twin concept.

While this deliverable should give an overview about current situation, further, more detailed recommendations relation to ICT and data interoperability are subject of WP5.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

6. Bibliography

[1] ISO 12006-2:2015, *Building construction — Organization of information about construction works — Part 2: Framework for classification*

[2] EN ISO 19650:2018, *Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles, Part 2: Delivery phase of the assets.*

[2a] EN 17412-1:2020, *Building Information Modelling - Level of Information Need - Part 1: Concepts and principles.*

[3] ISO 21597-1:2020 (ICDD), *Information container for linked document delivery — Exchange specification — Part 1: Container*

[4] ISO 21597-2:2020, *Information container for linked document delivery — Exchange specification — Part 2: Link types*

[5] EN ISO 23386:2020, *Building information modelling and other digital processes used in construction — Methodology to describe, author and maintain properties in interconnected data dictionaries*

[6] EN ISO 23387:2020, *Building information modelling (BIM) -- Data templates for construction objects used in the life cycle of any built asset -- Concepts and principles*

[7] EN ISO 29481-1:2016 (IDM), *Building information models — Information delivery manual — Part 1: Methodology and format*

[8] CEN/TC 442 prEN 17632:2021, *Building information modelling (BIM) – Semantic modelling and linking (SML).*

[9] ISO 16739-1, *Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema.*

[9a] ISO 10303, *Industrial automation systems and integration — Product data representation and exchange*

[10] bSDD:http://bsdd.buildingsmart.org/

[11] mvdXML: https://github.com/buildingSMART/mvdXML

[12] IDS: https://github.com/buildingSMART/IDS

[13] IFC (link to all versions): https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/

[14] OWL: https://www.w3.org/OWL/

[15] RDF: https://www.w3.org/RDF/

[16] SPARQL: https://www.w3.org/TR/sparql11-query/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

[17] SHACL: https://www.w3.org/TR/shacl/

[18] New European Interoperability Framework (EIF): https://op.europa.eu/en/publication-detail/-/publication/bca40dde-deee-11e7-9749-01aa75ed71a1/language-en

[19] Sebastian R., Böhms M. & Luiten B. (2020), *Exploring the opportunity and feasibility for developing and implementing Asset Lifecycle Information Management (ALIM) in the Building, Maritime and Offshore sectors.* By TNO report R12088.

[20] De Kleijn M., Wilcke X & Lucas C. (2018), *Innovatie data management asset management Rijkswaterstaat: een strategische verkenning naar 3d linked data en machine learning in het kader van het air-lab Linked data & 3d technology ten behoeve van toekomstig asset management Rijkswaterstaat: een strategische verkenning in het kader van het air-lab, SPINlab VU & Rijkswaterstaat* (document available in Dutch and upon request), *Spinlab Vrije Universiteit Amsterdam 2018. GITHUB - https://github.com/SPINLa*

[21] Sebastian R., Böhms H.M., Luiten G.T. (2021), *Developing a standardized approach of Asset Lifecycle Information Modelling for Semantic Digital Twin in the built Environment.* In: Proceedings of the Joint Conference of CIB W78 and Linked Data in Architecture and Construction, Luxembourg, 11-15 October 2021.

[22] Hernández-Moral, G.; Mulero-Palencia, S.; Serna-González, V.I.; Rodríguez-Alonso, C.; Sanz-Jimeno, R.; Marinakis, V.; Dimitropoulos, N.; Mylona, Z.; Antonucci, D.; Doukas, H. *Big Data Value Chain: Multiple Perspectives for the Built Environment*. Energies 2021, 14, 4624. https://doi.org/10.3390/en14154624

[23] Alonso, A., Pozo, A., Cantera, J.M., Vega, F. de la, Hierro, J.J. (2018), *Industrial Data Space Architecture Implementation Using FIWARE*, in: Sensors (Basel). 2018 Jul; 18(7): 2226).

[24] Noardo F, Arroyo Ohori K, Biljecki F, et al. *Reference study of CityGML software support: The GeoBIM benchmark 2019—Part II.* Transactions in GIS. 2021;25:842–868. https://doi.org/10.1111/tgis.12710

[25] Sebastian, R., Böhms, M., Luiten, B*. Exploring the opportunity and feasibility for developing and implementing Asset Lifecycle Information Management (ALIM) in the Building, Maritime and Offshore sectors.* TNO report R12088 Date: 6 January 2020.

[26] A. Palmieri, P. Prem, S. Ranise, U. Morelli, T. Ahmad, *"MQTTSA: A Tool for Automatically Assisting the Secure Deployments of MQTT Brokers"*, in 2019 IEEE World Congress on Services (SERVICES), Milan, Italy, 2019, pp. 47–53.

[27] PROV-O: https://www.w3.org/TR/prov-o/

[28] SKOS: https://www.w3.org/TR/skos-reference/

[29] QUDT: http://www.qudt.org/

[30] FAIR: https://www.go-fair.org/

[31] Zigbee: https://www.telkonet.com/what-is-zigbee/

[32] W3C-SSN/SOSA: https://www.w3.org/TR/vocab-ssn/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

[33] W3C-Spatial: https://www.w3.org/2015/spatial/wiki/Main_Page

[34] OGC: https://www.ogc.org

[35] ADV: https://www.adv-online.de/AdV-Produkte/Liegenschaftskataster/ALKIS/

[36] INSPIRE: https://inspire.ec.europa.eu

[37] GISGeography: https://gisgeography.com/spatial-data-types-vector-raster/

[38] Object Property Management (OPM): https://w3c-lbd-cg.github.io/opm/

[38a] SAREF: https://www.saref.etsi.org

Presentations:

[39] Bakker J., Boter J., de Kleijn M., van 't Veer R., Wilcke X, Lucas C. & Scholten H. (2019), *3D Linked Data and BIM for Life Cycle Information Management.* Presentation at IALCCE LCM Workshop, 27-29 October 2019, Rotterdam, The Netherlands.

Project Deliverables:

[40] Köhler J., Liljefors F.; IM-Safe Deliverable 1.2, Background Documentation – Online best practice guide- ,2022

[41] Sánchez Rodríguez A. et al.; IM-Safe Deliverable 2.1, Online Interactive Catalogue of Surveying Technologies for Transport Infrastructure, 2022

[42] Darò P. et al.; IM-Safe Deliverable 3.2, Background Materials for Implementation of Decision-Making Regarding Maintenance Strategies, 2022

[43] Rigotti M. et al.; IM-Safe Deliverable 4.3, Design of IT Platforms for Monitoring Data of Transport Infrastructure, 2022

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Appendix A - Further Resources

OpenStand Principles

From Sebastian et al. 2020 [19]:

On August 12, 2012, the Institute for Electrical and Electronics Engineers (IEEE), Internet Society (ISOC), World Wide Web Consortium (W3C), Internet Engineering Task Force (IETF) and Internet Architecture Board (IAB), jointly affirmed a set of principles which have contributed to the exponential growth of the Internet and related technologies. The "OpenStand Principles" define open standards and establish the building blocks for innovation. Standards developed using the OpenStand principles are developed through an open, participatory process, support interoperability, foster global competition, are voluntarily adopted on a global level and serve as building blocks for products and services targeted to meet the needs of markets and consumers. This drives innovation which, in turn, contributes to the creation of new markets and the growth and expansion of existing markets. There are five, key OpenStand Principles, as outlined below:

1. *Cooperation Respectful cooperation between standards organizations, whereby each respects the autonomy, integrity, processes, and intellectual property rules of the others.*
2. *Adherence to Principles - Adherence to the five fundamental principles of standards development, namely:*
   - *Due process: Decisions are made with equity and fairness among participants. No one party dominates or guides standards development. Standards processes are transparent, and opportunities exist to appeal decisions. Processes for periodic standards review and updating are well defined.*
   - *Broad consensus: Processes allow for all views to be considered and addressed, such that agreement can be found across a range of interests.*
   - *Transparency: Standards organizations provide advance public notice of proposed standards development activities, the scope of work to be undertaken, and conditions for participation. Easily accessible records of decisions and the materials used in reaching those decisions are provided. Public comment periods are provided before final standards approval and adoption.*
   - *Balance: Standards activities are not exclusively dominated by any particular person, company or interest group.*
   - *Openness: Standards processes are open to all interested and informed parties.*
3. *Collective Empowerment Commitment by affirming standards organizations and their participants to collective empowerment by striving for standards that:*
   - *are chosen and defined based on technical merit, as judged by the contributed expertise of each participant;*
   - *provide global interoperability, scalability, stability, and resiliency;*
   - *enable global competition;*
   - *serve as building blocks for further innovation; and*
   - *contribute to the creation of global communities, benefiting humanity.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

4. *Availability Standards specifications are made accessible to all for implementation and deployment. Affirming standards organizations have defined procedures to develop specifications that can be implemented under fair terms. Given market diversity, fair terms may vary from royalty-free to fair, reasonable, and non-discriminatory terms (FRAND).*
5. *Voluntary Adoption Standards are voluntarily adopted, and the success is determined by the market.*

Reference Study of CityGML Software Support

From Noardo et al. 2019 [24]

Interoperability through open standards is critical for the effective reuse and exchange of data and it is essential for reciprocal integration of different types of data. The integration of 3D city models with building information models (BIMs) has become a widely discussed topic in recent research. Two open standard data models considered for accomplishing such an integration are the Open Geospatial Consortium CityGML[22] for 3D city models, and buildingSMART Industry Foundation Classes (IFC[23]) for BIM models.

Two kinds of 3D information systems have been developed, studied and used in recent times, revealing their potential in related fields:

- 3D city models, which are used to represent city objects in three dimensions and advance previous 2D maps and other cartographic products, in order to support city analysis and management, city planning, navigation, and so on;
- building information models, which are used in the architecture, engineering and construction fields to design and manage buildings, infrastructure and other construction works, and which also have features useful to project and asset management.

Several international standards exist to govern the representation of the built environment in a shared manner, to foster interoperability and cross-border analysis and, consequently, actions, or to reuse tools, analysis methods and data themselves for research and, possibly, government. Some examples of international standards are: the European Directive for an infrastructure for spatial information in Europe (INSPIRE) (https://inspire.ec.europa.eu), aimed at the representation of cross-border territories in Europe, for common environmental analysis; the Land and Infrastructure standard (LandInfra, https://www.ogc.org/standards/landinfra), by the Open Geospatial Consortium (OGC), aimed at land and civil engineering infrastructure facilities representation; and the green building data model (gbXML, https://www.gbxml.org) aimed at the representation of buildings for energy analysis.

Nonetheless, the two dominant reference open standards for those two models are CityGML, by the OGC, focusing on urban-scale representation of the built environment, and the Industry Foundation Classes (IFC, ISO 16739), by buildingSMART, aimed at the very detailed representation of buildings and other construction works for design and construction objectives, first, but also intended to enable project management throughout the process, and asset and facility management in a following phase. Those standards are both intended to be very comprehensive and are therefore very wide and articulated. They both use complex data

[22] http://www.citygmlwiki.org

[23] https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

models allowing for a wide variety of models using object-oriented representations, even if that comes at the cost of slower and more inconsistent implementations.

Due to the overlapping interests in both fields (meeting in the building-level representation), increasing attention is being paid to 3D city model-BIM integration (GeoBIM), where the exchange of information between geospatial (3D city models) and BIM sources enables the reciprocal enrichment of the two kinds of information with advantages for both fields, e.g., automatic updates of 3D city models with high-level-of-detail features, automatic representation of BIM in their context, automated tests of the design, and so on.

The GeoBIM benchmark project (https://3d.bk.tudelft.nl/projects/geobim-benchmark/) studies the two standards involved in the GeoBIM integration (IFC and CityGML) with the aim of understanding whether one of the two offered more effective solutions that could be possibly borrowed by the other one in future developments. The aim of the benchmark was to get a better picture of the state of software support for the two open standards (IFC and CityGML) and the conversions between them, in order to formulate recommendations for further development of the standards and the software that implements them. In addition, we tested two known major technical issues related to GeoBIM integration and which are known to be solved only partially in practice: the ability of tools and methods to georeference IFC and the conversion procedures between IFC and CityGML.

The GeoBIM subject can be divided into several sub-issues:

- The harmonization of data themselves, which have to concretely fit together, with similar (or harmonizable) features (e.g., accuracy, kind of geometry, amount of detail, kind of semantics, georeferencing).
- Interoperability, which is a fundamental key in the integration. It is important to note here, that before enabling the interoperability among different formats (e.g., GIS and BIM), which is the theme of point 3 below, the interoperability GIS-to-GIS and BIM-to-BIM itself is essential. That means that the formats of data have to be understood and correctly interpreted uniquely by both any person and any supporting software. Moreover, an interoperable data set is supposed to remain altogether unchanged when going through a potentially infinite number of imports and exports by software tools, possibly converting it to their specific native formats and exporting it back.
- The effective conversion among different formats, that is, transforming one data set in a (standardized) format to another one in compliance with the end format specifications and features.
- The procedures employing 3D city models and the ones based on BIM should be changed in order to obtain better advantages by the use of both, integrated, since those systems enable processes which are usually more complex than just the simple representations.

CityGML (citygmlwiki.org and citygml.org), by Open Geospatial Consortium, is the most internationally widespread standard to store and exchange 3D city models with semantics in the geospatial domain. It establishes a structured way to describe the geometry and semantics of city objects. CityGML 2.0 (current version, considered in this project) contains classes structured into 12 modules, each of them extending the core module, containing the most general classes in the data model, with city object-specific classifications, (e.g., Building, Bridge, WaterBody, CityFurniture, LandUse, Relief, Transportation, Tunnel, Vegetation). These modules contain one or more classes representing specific types of objects, which differ in the way they are structured into smaller parts and the attributes that are expected for each. The most developed and most used module in practice is the Building module.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

CityGML as a data format is implemented as an application schema for the Geography Markup Language (GML) (CityGML uses version 3.1.1 of GML). It is an open format and human readable, which means that the information could potentially be retrieved even if losing backward compatibility in software. However, GML presents many issues from a software developer point of view, since, for example, too many alternatives[24] are allowed even for simple objects, and a supporting application is supposed to foresee all possible combinations of them. The result of this complexity is that few software programs completely support all possible combinations, and most of its richness and power is lost. An additional consequence of the kind of storage of such models is about their computational requirements: usually very large and complex files are produced, and it can be time- and resource-intensive to manage them properly in software.

As a possible solution to those issues, CityJSON[25] version 1.0.0 was recently released, providing a JSON encoding for a subset of the CityGML 2.0.0 data model. CityJSON follows the philosophy of another (non-standardized but working) encoding of CityGML: 3DCityDB: to store the models efficiently and allow practitioners to access features and their geometries easily. The deep hierarchies of the CityGML data model are replaced by a simpler representation. Furthermore, some more restrictions are applied and one and only one way is allowed to represent the semantics and the geometries of a specific feature. CityJSON is in the process of becoming an OGC community standard.

Research finding: Limited interoperability of the software tools

This study was designed to point out and provide evidence on the support and issues of available software for standardized information in CityGML version 2.0. Interoperability is essential for a number of use cases, and even for merely exchange and reuse data. Standards are supposed to be enabling such interoperability and standardization is the essential premise for the development of any integration, including that of GeoBIM. In fact, the potential integration with other kinds of data, including BIM and respective standards, requires structures and formats to be respected reliably within data sets. Data which are compliant with a standard are not supposed to change or vary when produced or used by different tools, otherwise, it would be almost impossible to plan, design and implement effective solutions for mapping, conversions and object transformation to other formats.

In particular, the four topics investigated are:

- What support is there for IFC within BIM (and other) software?
- What options for georeferencing BIM data are available?
- What support is there for CityGML within GIS (and other) tools?
- What options for conversion (software and procedural) (both IFC to CityGML and CityGML to IFC) are available?

For this purpose, a set of representative IFC and CityGML data sets were provided and used by external, voluntary participants in the software they would like to test in order to check the support in it for the open standard considered. Full details about the software tested and a full list of participants can be found in the respective pages of the benchmark website. The significant number of participants, balance in skills, fields of work, levels of confidence about

[24] http://erouault.blogspot.com/2014/04/gml-madness.html

[25] https://www.cityjson.org

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

the software tested (asked them to be declared) offered the possibility to limit the bias in the results.

Results were reported for 15 software packages, including both bespoke CityGML viewers but also generic GIS tools. Although we expected clearer patterns in the results, which would make it possible to better understand the remaining problems of interoperability in CityGML models, the only clear result is that very little interoperability is actually achieved. There are very few tools able to read the standardized data sets correctly and even fewer that are able to export them consistently. The ability to uniquely interpret the models and to leave them consistent through the import–export phases is absolutely essential for interoperability and what it enables (data exchange, data reuse, etc.).

From the experience gained, we can see how probably some difficulties in the implementation lie in the standard itself. The management of semantics is sometimes problematic (e.g., loss, change), with the main issues related to the management of hierarchies and other relationships. Despite being one of the most exciting possibilities in models of this kind, the lack of suitable support should compel a rethink of complex relationships to make them simpler and effectively manageable. A clear definition of how to structure entities, priorities and limits is also needed. Similar points can be made for geometries: constrained validity rules would help a homogeneous implementation, resulting in homogeneous and consistent models.

A general recommendation is therefore a collaboration of software developers with standardization institutions, starting from the needs of practitioners and users.

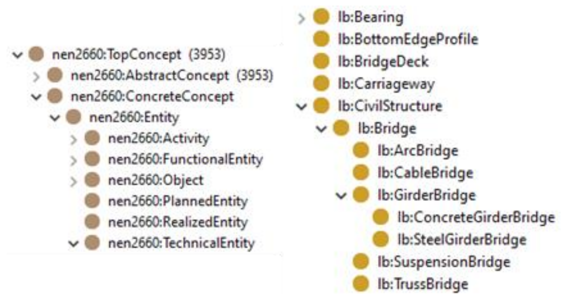
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

Example for use of Semantic Web/Linked Data

Using an exploratory case study of a highway bridge in the Netherlands, the Semantic Web / Linked Data approach was applied and a bridge ontology was developed according to the Dutch (NEN 2660)) and European (CEN TC442 SMLS) draft standards. The broader objective of the case study was to develop a Digital Twin prototype of the bridge for Structural Health Monitoring (SHM), predictive maintenance simulations, and impact analysis of the changing traffic loads and vehicle configurations (such as truck platooning where two or more trucks move in convoy using connectivity technology and automated driving support systems).

The case methodology contained the following steps: 1) Choosing the appropriate linked data language and format (also called ‘serialization’); 2) identifying and model the relevant concepts; 3) relating those concepts in a specialization hierarchy (also known as assigning the taxonomy); 4) identifying and modelling the relevant attributes; 5) identifying and modelling the relevant (inter)relations; 6) identifying and modelling the relevant constraints especially the decomposition constraints (giving the meronomy); and 7) populating the ontology for the real bridge. In this case study, it was decided to use Ontology Web Language (OWL) as linked data modelling language, and Turtle as the human-friendly format. The preliminary results from the case study are summarized in following Table.

Appendix A 1 - Example showing the ontology specification

<table>
<tr><th>Result representation / example</th><th>Result description</th></tr>
<tr><td>
</td><td>A small fragment of the defined girder bridge ontology is shown here; as an example, the SteelGirderBridge is a subclass of respectively a GirderBridge, a Bridge and a CivilStructure. The CivilStructure itself is a specialization of the NEN 2660 top level model concept ‘TechnicalEntity’.</td></tr>
<tr><td>lb:designLifespan
  rdf:type owl:ObjectProperty ;
  rdfs:range nen2660:QuantityValue ;
  nen2660:hasQuantityKind quantitykind:Time ;
  nen2660:hasUnit unit:YR .</td><td>All relevant attributes and relations are defined; an OWL code example for a designLifespan (the planned lifetime of the bridge) is shown here.</td></tr>
<tr><td>https://w3id.org/liggerbrug/def/designLifespan
or, in abbreviated from: ‘lb:designLifespan’.

The prefix ‘lb’ stands for: <https://w3id.org/liggerbrug/def#> making the total id/name for the property designLifespan.</td><td>The Turtle code is a convenient shorthand for a set of triples of the form ‘subject-predicate-object’. The predicate object part is shared for the same object. Each component of the triple is an Internet/WWW URI consisting of a ‘name space URI’ represented by a prefix and an actual id/name.</td></tr>
</table>

The implementation of open standards in the Digital Twin ontology modelling can be summarized as follows. The draft Dutch NEN2660 standard (NEN 2021) aligned with the European TC442 Semantic Modelling and Linking (SML) standard (CEN/TC 442 2021) were

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.

used for defining both the ontology and its data instances. The modelling patterns were applied for:

- modelling the quantities with their quantity kinds and units according to Quantities, Units, Dimensions and dataTypes (QUDT - public repository, https://github.com/qudt/qudt-public-repo);
- modelling the enumeration types with allowed items;
- defining the approach for identifiers (including an URI-strategy) and annotations like labels and definitions.

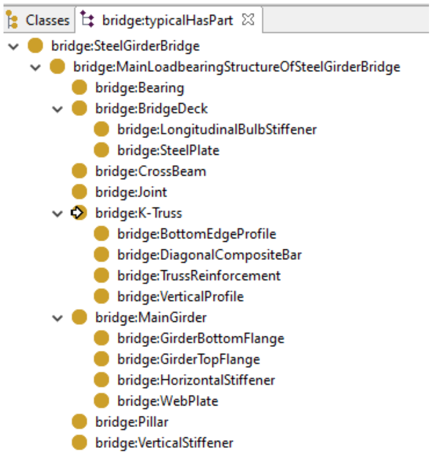
The languages from W3C were applied according to the approach envisaged by NEN and CEN. In the case study, this approach was applied for measurement and other datasets, and it resulted in standardized data for the bridge, comprising a standard format (Turtle) and standard semantics (bridge ontology) specified in standard languages (OWL) according to NEN2660. This standardized data is accessible through the standard W3C Linked Data Query Language (SPARQL) and can be used for various Digital Twin functionalities. The attribute modelling was done according to NEN 2660 as a relation (or in terms of OWL: an object property). Its quantity-kind was ‘Time’ and its measuring unit was a year (‘YR’). Quantity kinds and units were fully reused from the most recent QUDT ontology. The preliminary results are summarized in the following Table.

Appendix A 2 - Example showing attribute specification.

<table>
<tr><th>Result representation / example</th><th>Result description</th></tr>
<tr><td><pre>
prob:StochasticValue
  a owl:Class ;
  rdfs:subClassOf [
      a owl:Restriction ;
      owl:cardinality "0"^^xsd:nonNegativeInteger ;
      owl:onProperty prob:posteriorMean ;
    ] ;
  rdfs:subClassOf [
      a owl:Restriction ;
      owl:cardinality "0"^^xsd:nonNegativeInteger ;
      owl:onProperty prob:posteriorStdev ;
    ] ;
  rdfs:subClassOf [
      a owl:Restriction ;
      owl:cardinality "1"^^xsd:nonNegativeInteger ;
      owl:onProperty prob:priorDistributionType ;
    ] ;
  rdfs:subClassOf [
      a owl:Restriction ;
      owl:cardinality "1"^^xsd:nonNegativeInteger ;
      owl:onProperty prob:priorMean ;
    ] ;
  rdfs:subClassOf [
      a owl:Restriction ;
      owl:cardinality "1"^^xsd:nonNegativeInteger ;
      owl:onProperty prob:priorStdev ;
    ] ;
  skos:prefLabel "StochasticValue"@en .
</pre></td><td>The range of this attribute is a QuantityValue (defining a complex value that can hold an actual value plus extra metadata). Because of the application of Bayesian methods, a neutral stochastic value is also defined and used to show how some constraints look like in in this instance. A stochastic value is a complex value involving prior and posterior mean-values with standard deviations.

The various constraints are modelled in OWL by making a concept a subclass of the set of concepts having the constraint. So, for instance, a StochasticValue is a subclass of the class of things having exactly one prior mean value (in bold lines shown here).</td></tr>
</table>

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958171.



When applying such a constraint to decomposition relations, a typical decomposition structure also referred to as a meronomy can be modelled and visualized. The taxonomy and the meronomy are together often seen as the 'backbone' of the ontology.

With a simple SPARQL-query, explicit typical decomposition relations can be inferred from the instance decomposition constraints resulting the hierarchy shown here.

```
ib:IJsselbridge_1-main_detailed
  rdf:type lb:SteelGirderBridge ;
  lb:amountOfSpans 5 ;
  lb:constructionMethod lb:WithGirders ;
  lb:designLifespan [
      rdf:type nen2660:QuantityValue ;
      rdf:value 80.0 ;
      nen2660:hasUnit unit:YR ;
    ] ;
  lb:length [
      rdf:value "295.0"^^xsd:float ;
      nen2660:hasUnit unit:M ;
    ] ;
  lb:materialType lb:Steel ;
  lb:residualLifespan [
      rdf:value "20.0"^^xsd:float ;
      nen2660:hasUnit unit:YR ;
    ] ;
  nen2660:hasPart ib:MainLoadbearingStructureOfIJsselBridge .
```

When all concept, attributes, relations and constraints are modelled, one can instantiate the ontology for instances of bridges, called individuals in 'linked data speak'. Here is some global data as example that is fully compliant (i.e. logically consistent) to the ontology with all its constraints.

The concept or ontology design decision that was taken in order to implement and maintain the live data connection between the digital and physical twins can be summarized as follows:

- A dataset was selected to describe the design breakdown, attributes and relations of the elements in the semantic digital twin case study of the steel bridge. Anomaly data values were then generated to contain the deviations between the real conditions and the design data. These values were based on the comparison between the prior assumptions and the posterior values derived from the repeated finite element analyses after the actual data was taken into account. Both prior and posterior data as well as the mean and the standard deviations were stored in the digital twin datasets. For the Bayesian 'learning' ('updating of the knowledge') the actual measurements under specific traffic load conditions were used.
- An example of the ontology design in the case study for the analysis for 'estimating the thickness of the bottom flanch of the steel bridge that has been reduced by corrosion' is shown below:

```
aib:ThicknessReduction_1
  a anomaly:ThicknessReductionOfBottomFlange ;
  anomaly:forPhysicalObject ib:GirderBottomFlange_west ;
  anomaly:globalEffect false ;
  anomaly:hasCause aib:Corrosion_1 ;
  anomaly:parameterType anomaly:Deterministic ;
  anomaly:thickenssReductionBottomFlange [
      prob:posteriorMean [
          rdf:value "8.2E5"^^xsd:float ;
        ] ;
      prob:posteriorStdev [
          rdf:value "1.1E5"^^xsd:float ;
        ] ;
```


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.

```
        prob:priorDistributionType prob:NormalDistribution ;
        prob:priorMean [
            rdf:value "0.03"^^xsd:float ;
          ] ;
        prob:priorStdev [
            rdf:value "0.0006"^^xsd:float ;
          ] ;
      ] ;
.
```

- In a intended next phase of the case study, the sensing data from different measurement points at the steel bridge will be stored as linked data according to the standard ontologies like W3C's SSN/SOSA.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958171.