Connecting Object Data in Natural Stone Renovation Projects

Citation for published version (APA):

Document status and date:
Accepted/In press: 01/01/2024

Document Version:
Typeset version in publisher’s lay-out, without final page, issue and volume numbers

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 18. Aug. 2024
Abstract

Industry experts working on renovation projects rely on a wide range of information. Standards to connect this information are lacking, as well as tools to communicate the data with the industry expert. This study researches the application of semantic web technologies to connect heterogeneous data in renovation projects. Based on this connected data, we present a web-based linked building data viewer through which practitioners can ask relevant questions about a project. Close collaboration with industry experts led to time improvements and a positive attitude towards data-driven support systems.

Introduction

Historical buildings hold significant cultural and economic value and are key elements of a country's cultural heritage. Governments implement policies and regulations to safeguard and preserve those buildings. To do so, practitioners rely on a lot of information about the monument. This information is, however, often documented using various standards, data formats, and software systems, resulting in a situation of fragmented information that is difficult to manage.

The construction industry is increasingly adopting Building Information Modeling (BIM) as a method to digitize information in collaborative workflows. The Industry Foundation Classes (IFC) is a widely recognized standard to store information about objects in BIM models. Common Data Environments (CDE) provide ways to collaboratively use IFC-based BIM models and other information. However, the capabilities of those CDEs to provide users with an integrated overview of information are often limited, as the lack of data integration standards and the vendor lock-in of these systems make it difficult to link related information across different systems and stakeholders.

Semantic web technologies promise to alleviate some of those issues, as they enable connecting and structuring the currently fragmented datasets from various domains and stakeholders (Pauwels et al., 2017), and applying those technologies to integrate BIM models with built heritage information could help in structuring, validating, visualizing, and sharing information in monumental restoration and preservation projects. Efforts in this direction have already been made by other researchers. Hamdan et al. (2019) developed a modular ontology for defining damaged objects and their topology. Bonduel (2021) contributed to the development of a network of new and existing modular ontologies to integrate heterogeneous built heritage information. These works both conclude that more research is needed to apply linked data in real projects and that a better translation from real-world processes to semantic web technologies is necessary.

Next to the necessary development of more data integration standards, Hamdan et al. (2019) suggest that those standards should also be linked to practical actions so that integrated data can actually answer the questions of end users. Bonduel (2021) suggested the development of demo applications to translate the integrated data to end users. Various initiatives to build web-based linked building data viewers exist, such as LBDviz (Donkers et al., 2023), LD-BIM1, RUB-IP (Hagedorn et al., 2023), BIM-SIM (Chamari et al., 2022), and ConSolid (Werbruck et al., 2024). However, the current functionalities of those viewers are limited and use in a heritage context requires specific functionalities that are related to the requirements of end users in this specific domain (see Figure 2).

Other research has already made efforts that could help in structuring, validating, visualizing, and sharing data in the context of restoration projects. However, in practice, there still exists a research gap in structuring and visualizing such data in the construction phase of heritage projects. The created solution should use those new techniques while being integrated with existing processes and tools on site. This paper investigates the capabilities of applying semantic web technologies to integrate information from BIM models with other files and information used by practitioners. This work hereby focuses on the integration of data that is not exclusively restricted to RDF graphs. Instead, a software solution is relied upon that combines multiple datasets of multiple types, sometimes referred to here as 'connected data'. This setup and approach is not much different from several recent digital twin implementation initiatives.

This paper presents an ontology and a software tool to integrate data related to built heritage and make this data accessible to practitioners. Section 2 first presents state-of-the-art built heritage data integration, with a specific focus on natural stone renovation projects in the Netherlands. The processes in natural stone renovations, including the relevant information streams, are then
analyzed, including the relevant end users and their requirements. Those requirements are then translated into a software architecture. The information requirements are also translated into a network of ontologies, combining both existing ontologies and a newly developed extension. These methods are then finally used to develop a web-based tool that enables different end users to interact with the integrated information. The tool is validated in a real-world case study at the Dam Square in Amsterdam, The Netherlands.

**Related Work**

**Conservation of Natural Stone Built Heritage**

Historical buildings in the Netherlands hold cultural and economic value, attracting tourists and preserving the country's heritage. The conservation of these buildings is regulated through self-regulation and guidelines established by the Dutch Cultural Heritage Agency (RCE) and the Foundation ERM (Naldini & Hunen, 2019). The ERM guidelines provide technical specifications and quality standards for restoration work, ensuring the preservation of historical value. The process involves inspections, damage assessment, identification of cultural-historical value, determination of appropriate materials and techniques, and decision-making using the “Conservation ladder” (Lubelli et al., 2021). The whole conservation cycle should be documented to provide valuable information for future restoration projects. For natural stone restoration projects, aspects such as cultural-historical value, damage assessment, damage identification, material type, and shape need to be documented (ERM, 2020). Documenting these aspects is done using a wide range of standards and data formats, creating a situation that is difficult to manage.

**Connected Data**

Several initiatives have aimed to develop web-based building data. One of the first initiatives was the creation of ifcOWL, which translated the IFC schema into an RDF-based schema using semantic web technologies (Beetz et al., 2009; Pauwels & Terkaj, 2016). However, as it mirrors the IFC EXPRESS schema, it inherits the same complex structure and size. The Linked Building Data Community Group of the World Wide Web Consortium (W3C LBD CG) was created to identify and align ontology development initiatives for building data. The approach they took was to create a set of interoperable, flexible, and open standards covering different aspects and domains (Rasmussen et al., 2020). Rasmussen et al. (2017) took the first step in creating the Building Topology Ontology (BOT), which is a flexible and lightweight ontology to define the relationship between elements within a building. This is a basic ontology that describes building structures using the classes bot:Building, bot:Storey, bot:Space, and bot:Element. Implementing this ontology would result in a shift from file-based collaboration to data-centered collaboration using RDF, directly linking data instead of referring to other documents.

While the use of RDF graphs is immensely powerful to create a semantically rich cloud of linked data, several other types of data exist that typically do not fit this RDF graph principle. Several hybrid methods have therefore been proposed and deployed in the last few years that combine semantic web technologies with 3D models, pictures, point clouds, and so forth, e.g. Werbrouck et al. (2020). Another example is the use of semantic web technologies to describe information that does not fit into the BIM exchange schema IFC, for instance describing information about damages on objects while linking to the actual object in the IFC model (Hamdan et al., 2019). Also smart building implementations, e.g. based on Brick, adopt this approach to connect to telemetric data (Chamari et al., 2023). The adoption of semantic web technologies for the description of heritage buildings could help in integrating the different data sources into a full virtual representation. These hybrid methods, which integrate semantic web technologies with BIM models and other files, are considered a highly promising route for handling the data heterogeneity challenges in the built heritage domain.

**Existing Ontologies Related to Natural Stone Restoration**

Researchers proposed various ontologies that enable the representation, classification, and relationship modeling of historical buildings, their components, damages, inspections, interventions, geometry, and additional properties. Bonduel (2021) proposes the Construction Tasks Ontology (CTO), as an extension to BOT, providing terminology for construction tasks. The Damage Topology Ontology (DOT) contains classes and properties that define topological relations between the damage and the affected building components (Hamdan et al., 2019). Based on the same principles as BOT, DOT, and CTO also support the creation of extensions in a specific domain. Bonduel (2021) used this principle when converting the MDCS damage atlas into an ontology that extends DOT. Another example is the Natural Stone Damage Ontology (NSD), which is based on a different damage atlas (Hamdan, 2023; Saeed & Hamdan, 2019). Saeed & Hamdan (2019) have also developed the Stone Component Ontology (SCO), as an extension ontology of BOT, for semantically modeling natural stone facades. Furthermore, the Ontology for Managing Geometry (OMG) provides generic concepts for managing geometry descriptions in a semantic web context (Wagner et al., 2019). It focuses on general-purpose functionalities that allow the handling of single, multiple, and versioned geometry descriptions. OMG can be extended by the File Ontology for Geometry formats (FOG), which extends the OMG to describe file format, versioning, and object identifier information for geometry data (Bonduel et al., 2019). All these ontologies were developed using the principles of the W3C LBD CG, creating modular data structures that can be used following a plug-and-play.
principle to represent information in, for example, a natural stone restoration.

These evolutions also happened in practice. The Dutch Cultural Heritage Agency (RCE) has published its registry of national monuments as a large RDF graph or knowledge graph. This dataset is also linked with a registry of all buildings and addresses in the Netherlands as a large RDF graph by the Dutch cadaster. The RCE has also published thesauri terms as linked data using SKOS (CHT). This includes, for example, different types of damages and materials.

Methodology

This paper starts with natural stone renovation projects and investigates the use of the presented technologies and workflows (connecting data) for this data challenge. This paper first reviews the existing processes in natural stone renovation projects. Expert interviews with a project leader, planner, BIM coordinator, and BIM manager resulted in a list of competency questions which is translated to a use case diagram. Based on those competency questions and the processes in natural stone renovation projects, a network of ontologies is created that can support integrating the available data. The use case diagram is then translated into a system architecture that enables the end users to interact with the various types of information through a web application: Heritage LBDViz. A prototype of this application is tested using a real-world renovation project of the Dam Square in Amsterdam, The Netherlands, and validated against the competency questions in collaboration with the industry professionals. The development and validation of this tool was an iterative process with industry professionals where different functionalities were created and validated through three iterations: definition of the competency questions and use case diagram, development of the user interface, and validation.

Natural Stone Restoration Processes

The process of natural stone facade restoration consists of a series of tasks depending on the type of intervention that needs to be performed. To visualize this process, a BPMN schema was created, see Figure 1. The ERM guidelines (ERM, 2020) specify that information should be documented throughout the conservation cycle of a heritage property. The data column in Figure 1 shows when this data is collected. This gives an insight into when information is generated so that it can be associated with specific parts of the process. The schema does not show every possible workflow. However, it gives a clear overview of the main steps that are part of the process and the different types of workflows that are possible.

Firstly, an inspection of each object in the building is carried out in two steps. The first is aimed mainly at gathering information about the object. This information is often gathered in a spreadsheet. In the second step, it is decided what intervention is needed. This information is then used to produce a cost estimation, including the type of replacement material, which creates an additional spreadsheet of information. Once the interventions and costs are known, the actual work on the construction site begins. Three types of workflows are used in the execution of the interventions. Some interventions can be carried out in situ, creating a simple process with only one task: the repair. However, it is often necessary to carry out an intervention in the masonry workshop, which results in a series of additional tasks. In this case, the object must be dismantled from the building and then transported to the masonry workshop. Once the object has been repaired or reconstructed, it can be transported back to the site where it is reassembled in the building. Dismantling an object can require some of the surrounding objects to be dismantled as well. These are then transported to a location for temporary storage and, when the time is right, transported back to the site and used in a reassembly task.

Progress is recorded at different steps by taking photos of the object. When all the work has been carried out and all the objects are back in place, a check is performed to ensure that the result meets the desired quality. Throughout the process, the multiple sources of data are passed on to different stakeholders to meet their information needs. This creates a situation where multiple sources of data exist in different places.

End User Requirements

During interviews with the industry experts about what information they need throughout the process, several competency questions were identified. Below are some of the most pressing ones:

CQ1 What is the location of the object in the building?
CQ2 What photos are available of a specific object?
CQ3 What is the volume, material, and shape of the object?
CQ4 What interventions are conducted on an object?
CQ5 How to give an overview of all relevant information on an object?
CQ6 What is the height of the object relative to ground level?
CQ7 What is the progress status of every object in the building?

The various competency questions were translated into use cases to be implemented as functionalities in the system (Fig. 2). They thus inform the normal user requirement definition stage of a regular software development process. These functionalities can then be developed during the different iterations of the prototype.

\[1 \text{https://linkeddata.cultureelerfgoed.nl/}
\[2 \text{https://labs.kadaster.nl/cases/bag-ld}
\[3 \text{https://linkeddata.cultureelerfgoed.nl/}
\[4 \text{https://data.cultureelerfgoed.nl/term/id/cht.html} \text{CQ1 What is the location of the object in the building?}
\[5 \text{CQ2 What photos are available of a specific object?}
\[6 \text{CQ3 What is the volume, material, and shape of the object?}
\[7 \text{CQ4 What interventions are conducted on an object?}
\[8 \text{CQ5 How to give an overview of all relevant information on an object?}
\[9 \text{CQ6 What is the height of the object relative to ground level?}
\[10 \text{CQ7 What is the progress status of every object in the building?} \text{CQ1 What is the location of the object in the building?}
\]
development. Each use case either supports another use case or provides an answer to one of the competency questions. Multiple types of users are considered in the use cases being developed. The use cases for the beginner user focus on recognizable output, extended with new possibilities that introduce the user to the system. A more advanced user is provided with functionalities to find traditional information through the system. Finally, a pro user is also provided with the ability to create queries based on their extensive knowledge of the data structure.

System Architecture

Based on the existing processes and the end user requirements, a 5-layer conceptual system architecture was established for combining the RDF graphs with 3D models and other files, see Figure 3, following earlier found best practices (Donkers et al., 2023). The resources layer contains the heterogeneous data produced in the data layer in Figure 1, including BIM models, photos, and states. The data is structured using a network of ontologies. An ETL procedure then transforms some of the source information to RDF and loads the data into the RDF graph. The Connected Data layer combines an RDF graph in the backend with links to the 3D BIM model and other files. As said, this is similar to recurring Digital Twin development projects. A services layer provides capabilities to query data from the connected data layer and ask questions about this information. Users can interact with this information via user interfaces. This paper presents a 3D model viewer - Heritage LBDviz - and a spreadsheet exporter as user interfaces that implement functionalities based on the described use cases. The architecture defines the components independently of the selected solutions, focusing on the responsibilities of these different layers in the system. Implementation of the system architecture must include considerations for communicating between different layers.
Results

After development, the prototype was validated in a real-world setting. The restoration of the National Monument on Dam Square was chosen as a case study. The monument consists of a pylon with statues and a memorial wall with reliefs. In front of the monument, two lion figures stand as sentinels. The twenty-two-meter-high pylon, that is made of concrete, the memorial wall and the lions are cladded in travertine, a marble-like porous limestone from Tuscany, Italy. The last restoration took place in 1997. In 2022, 25 years later, the National Monument needed to be repaired or restored. During the restoration, the travertine elements were dismantled and transported to the stonemasonry workshop, where they were cleaned, repaired, and replaced where necessary, minimizing disruption to the Dam Square.

The dataset provided does not include all the data described in the process schema, but the project offers a realistic representation of the current way in which practitioners operate and collect and manage data. The following data sources were available for this research: a Revit model and PDF elevations to give an insight into the geometry of the building objects; the Autodesk Build system used to carry out the inspection, which contains data on object properties and photographs; the actual set of photographs taken during the process; and finally some data documented in several spreadsheets, such as the interventions carried out on the elements and the dimensions of all the objects.

Network of Ontologies

Based on all the available data, a mapping was done to find all the classes for the case project data in the BOT ontology and all its extensions, such as CTO, DOT, and OMG. However, some of the data available in this case study could not be described using these existing ontologies. The Restoration Intervention Information (RII) ontology was created to define a set of classes and relationships to fill this gap. The ontology focuses on the interventions that are carried out on the heritage assets and the different properties of elements that are used as information in this process. The ontology describes classes of interventions based on the conservation ladder, defining the class rii:Intervention and its subclasses rii:Preservation, rii:Repair, rii:Reconstruction (Fig 4). These interventions are then linked to the cto:InspectionTask for which they were selected and the dot:Damage they are intended to cover. Finally, the cto:hasTaskContext is used to group the various cto:RepairTask instances that are performed to carry out the rii:Intervention. In addition to interventions, the ontology is also used to add external resources to the graph, specifically photos. The DOT ontology already defines a dot:ExternalResource, but this is specifically described as documentation that should only be damage-related. As not all information is damage-related, it was decided to use the class rii:ExternalResource with the subclass rii:ExternalImage, which is used in combination with the data property rii:filePath to link to externally stored images (Fig. 4).
The ontology also includes an OWL property chain based on cto:isSubjectOfTask and rii:capturedDuringTask, creating the relation rii:hasDescriptionInResource with the inverse rii:describesElement. The RII ontology is available at https://marijn.janssensteenberg.github.io/rii#.

Heritage LBDviz

To demonstrate the possibilities of our method in the context of restoration projects, a web application has been developed, serving as an easily accessible user interface that practitioners can visit through their browsers. This application - Heritage LBDviz - is an extension of the already existing LBDviz (Donkers et al. 2023). The tool (Figure 5) is a single-page web application combining an IFC viewer with several menu windows. IFC files are loaded through the IFC.js library, a JavaScript library that allows users to load, view, and edit IFC models in the browser. It also parses the IFC file to JSON so that one can interact with it through JavaScript. Behind this IFC file is a backend with the data, including RDF graphs and non-RDF data. This data can be accessed via Comunica.js, a meta query engine that can be used to execute SPARQL queries against multiple SPARQL endpoints or RDF files simultaneously. Heritage LBDviz5 and a guide to recreate the tool6 are available on GitHub.

Heritage LBDviz aims to fulfill the competency questions suggested by practitioners by adding the functionalities as shown in the use case diagram (Figure 2). First, beginner users can open the URI of a specific object, which will automatically open the viewer and highlight the specific object (CQ1). One can then use the various query functionalities to ask for information about this object, either by manually creating SPARQL queries or by using the predefined SPARQL queries hidden behind buttons in the user interface. This allows users to query information from the RDF graph, such as an object’s material, size, and other properties (CQ2) and current and historical interventions on the object (CQ4), as well as non-RDF data such as photos (CQ3). All photos linked to an object can be queried and will be printed in the user interface (Figure 5). Next to printing results of queries in the user interface, Heritage LBDviz also allows users to export the results into spreadsheets. Those spreadsheets replace the traditionally manually created spreadsheets, saving time, while also allowing interaction between the spreadsheet and the viewer through hyperlinking cells with the URIs of objects (CQ5). Finally, users can interact with the IFC model in the viewer, for example by measuring distances in the 3D model (CQ6) and by highlighting objects. Using the ontology, Heritage LBDviz can also perform live reasoning and infer the current status of all elements in an IFC model after which all elements are highlighted using associated colors (CQ7).

Heritage LBDviz was validated using the Dam Square monument and the system was tested together with industry experts. Firstly, the tool seems to significantly save time in finding information on a specific object. Construction workers now don’t need to go down to the workmen’s shelter to find information on a block, and the normally manually created spreadsheets are now automatically generated. The fact that heterogeneous information can be accessed via one user interface is also expected to reduce errors. Inconsistent information regularly causes elements to contain errors after repair, delaying the process and increasing failure costs. The vendor-neutral approach makes projects less reliant on specific software, especially in multi-stakeholder projects. Finally, using the tool enhanced the awareness of the available data and its value, and industry experts could easily come up with suggestions on innovative functionalities of the tool.

---

5 https://github.com/MarijnJanssenSteenberg/Heritage-LBDviz
6 https://github.com/AlexDonkers/Frontends-and-LBD
Discussion

This paper discusses an approach to integrating data related to renovation projects and making this data available to industry professionals. This paper is not the first paper to present data integration methods related to the topic of renovation projects. This project builds on these earlier projects and re-uses several existing ontologies. Furthermore, this paper presented an extension in the form of the RII ontology to define knowledge on interventions in restoration projects and specific properties that are commonly used in practice. The ontology follows best practices that suggest a modular approach and extends already existing ontologies. Although this ontology is tested in a natural stone renovation case study, its modularity should enable extensions toward other types of projects.

Secondly, this project confirms the feasibility and usefulness of making different types of data available through the appropriate use of semantic web technologies (RDF graphs) and state-of-the-art web development practices (service-oriented architecture). Moreover, it can be experienced that this approach to data integration reduces the gap between developers and data on the one hand, and practitioners on the other. As suggested in Donkers et al. (2023), co-creating the Heritage LDBviz environment together with practitioners significantly reduces barriers to using these data fittingly in practice. Validating the tool together with industry experts triggered ideas for more complex functionalities. This shows future opportunities for using such integrated data in restoration projects, and it shows that co-creating such tools in iterative cycles together with industry experts can create more awareness and engagement toward data-driven decision support systems.

Further to the above contributions, two points for further development were identified. First, Heritage LDBviz mainly focuses on visualizing data to answer the competency questions of users. However, validating the quality of the data is essential to make sure that the information is reliable. Future research in quality control methods, such as using SHACL, is needed to improve the reliability of the tool. Second, it is believed that access to a complete dataset of past restorations can create many opportunities for subsequent cycles of conservation. This may be better in reach using the proposed way of working, as it has a high level of modularity and scalability. More research is needed to understand how such long-term data preservation can be organized.

Conclusion

Conserving built heritage is a complex process that involves a lot of information. This research identifies two important challenges: we lack methods to integrate this information, especially beyond the RDF formatting, and we lack tools to make this information accessible and useful to practitioners working on renovation projects. This paper applies semantic web technologies and novel web development techniques to solve these challenges in close collaboration with the industry experts. After a thorough review of existing processes in natural stone renovations, using both literature and discussion sessions with industry experts, seven competency questions were defined. These led to the development of a network of ontologies, extended with the newly developed RII ontology. The knowledge from the review of renovation processes was translated to a use case diagram, based on which a system architecture for a web app was created. This web-based BIM viewer - Heritage LDBviz - aims to help the industry experts in answering their competency questions. The tool was validated together with those industry experts in a case study at the Dam Square. This validation showed that Heritage LDBviz was able to answer the competency questions and that it had the positive side effect of creating more engagement and awareness of such data-driven support systems.
The developed network of ontologies shows that it is possible to integrate heterogeneous data related to restoration projects. The web application shows that communicating this data via a simple user interface can help industry experts in their daily work and improve their understanding of the value of these data for future use cases. While the methods presented in this paper are specifically tested on natural stone renovation projects, we believe that the system architecture can be used in a much wider range of contexts, as long as there is a need to connect, structure, and communicate data.

References