BIM and Semantic Web: Key Enablers for Nuclear Knowledge Management?

Pascal Mosler[1] , Ahmed Han Karapinar[2] , Benedikt Malte Hauke Broich[2] and Theo Kastner-Guhl[2]

[1]Institute of Numerical Methods and Informatics in Civil Engineering, Technical University of Darmstadt, Darmstadt, Germany

[2]Department of Computer Science, Technical University of Darmstadt, Darmstadt, Germany

E-mail(s): mosler@iib.tu-darmstadt.de

Abstract: Knowledge management plays a crucial role in nuclear facilities. Due to the diverse business processes to be mapped across individual life cycle phases and the heterogeneous needs of various stakeholders, no uniform platform standard for knowledge management in the field of nuclear technology has yet been established. This paper presents a concept for a software architecture designed to enable digital knowledge management in this context. It is based on an analysis of existing approaches for linking 3D geometries with semantic information and documents, in particular, based on Building Information Modeling, as well as existing information models in nuclear technology. Both disciplines are increasingly relying on semantic web technologies to enable end-to-end interoperability of data across the entire life cycle of a building. Semantic web technologies such as RDF and SPARQL also play a key part in the proposed software architecture. Finally, the paper describes the framework conditions under which the architecture can be further developed into a mature information system. A central feature is the use of extended reality, which enables on-site and collaborative use of the system in multiplayer mode.

Keywords: Knowledge Management, Nuclear Facility, BIM, Semantic Web, Extended Reality

1 Introduction

A nuclear facility is *"a facility (including associated buildings and equipment) in which nuclear material is produced, processed, used, handled, stored or disposed of"* [1]. In addition to nuclear power plants (NPPs), these facilities include interim storage facilities, repositories, nuclear reprocessing plants, and research facilities. Assuring the safe and efficient operating phase (operating & maintenance) is crucial for the economic success of such a facility. The workforce of an NPP is large and comprises a wide range of disciplines and qualification levels. A large number of stakeholders are involved in the environment of nuclear facilities, including owners and operators, engineering construction contractors, regulatory bodies, and others [2].

According to the International Atomic Energy Agency (IAEA), knowledge management (KM) is *"an integrated, systematic approach to identifying, managing and sharing an organisation's knowledge [. . .]"* [3]. Knowledge can be divided into explicit knowledge, implicit knowledge, and tacit knowledge. In the case of nuclear facilities, knowledge is continuously accumulated during the individual life cycle phases (planning, construction, operation, decommissioning). Document management is not only an important component of KM but also a crucial part of overall project management [4]. Handling different data formats and attempts to structure data represent a major challenge. Decommissioning is particularly complex for nuclear facilities due to radiation protection and must be carefully planned and carried out. This requires access to the whole operating knowledge of the facility.

2 Technical Basics

2.1 Building Information Modeling

Building Information Modeling (BIM) describes the consistent and continuous use of digital information over the entire life cycle of a built facility—including its planning, construction, and operation [5]. A BIM *model* describes a comprehensive digital image of a building. It contains the three-dimensional geometry of the individual elements of the building as well as non-physical objects such as rooms and floors. The properties and hierarchical dependency of these objects are described semantically [5]. The Industry Foundation Classes (IFC) are a standardized data format or data model for the manufacturer-neutral exchange of digital building models [5]. IFC establishes a strict separation between geometry and semantics. This separation makes it possible to link a semantic object with one (or more than one) geometric representation.

2.2 Semantic Web

The concept of the Semantic Web was introduced by Tim Berners-Lee et al. in 2001 [6]. Based on the World Wide Web (WWW), the aim was to create a machine-searchable system with various machine-readable information and information endpoints. In their conception, they already envisaged the possible linkage between information and physical objects using ontologies [6]. On the one hand, this characteristic makes the Semantic Web interesting in the context of BIM [7]. On the other hand, the Semantic Web plays an important role in knowledge modeling, a subfield of artificial intelligence, where the aim is to achieve machine-readable formalizations [5]. The World Wide Web Consortium (W3C) published the RDF (Resource Description Framework) standard, with which information is formally expressed in triples. A triple is a structural unit that forms a directed, labeled graph, with the edges indicating a named relationship between two resources represented as nodes [8]. Ontologies are largely based on these RDF triples. In analogy to the Structured Query Language (SQL) for relational databases, the W3C has established SPARQL as a query language for graphs.

2.3 (Knowledge-centric) Plant Information Models

In 2009, the American research organization EPRI (Electric Power Research Institute) started to develop a standardized information model, the Plant Information Model (PIM). The background to the development of the PIM was the realization that "silo thinking", i.e. distributed data silos within an organization, leads to unnecessary cost increases due to inefficient collaboration. The PIM was to be used for the construction of new nuclear facilities, enable interoperability, and map the entire plant life cycle. KM has also already been recognized as a use case as well as the advantages of

linking documents and data with a 3D model of the facility based on an object taxonomy. Technically, XML (Extensible Markup Language) was used to enable an object relationship model. In 2014, the development was suspended without the PIM having been completed [9].

In 2020, the IAEA published its concept of a knowledge-centric PIM and defined it as *"a semantically organized set of information describing plant structures, systems and components [...] within a knowledge framework [...]"* [8]. This concept combines the characteristics of a PIM and semantic web technologies. The focus is no longer on mere data, but on contextualized information, and enables more efficient KM internally and in exchange with stakeholders [8].

3 Existing Approaches for Linking Geometry with Semantic Information and Documents

The relationship between geometry and semantics plays a major role in KM. This chapter examines this interplay in the AEC sector and in the nuclear industry.

3.1 AEC Sector

The idea of linking project documents to a model-based information system was already being discussed before BIM was established [10]. An early example is that of Chau et al. [11], in which the Work Breakdown Structure (WBS) was proposed as a link to 3D CAD. In 2019, the early approaches of using the WBS to link objects and information/documents were taken up again by Cha and Lee [12]. In their BIM-based document management system, documents are compiled by the respective owner and organized using a breakdown structure. The documents are stored in a database and their metadata is linked to the BIM [12].

Motawa and Almarshad [13] proposed a knowledge-based BIM for building maintenance in which specific semantic information is entered in the Autodesk Revit authoring software and exported with the IFC schema. A more comprehensive proposal was presented by Opitz et al. [4], who proposed technologies such as RDF and XMP (Extensible Metadata Platform) to link documents and their metadata. In their exemplary application BIMdox, a mapping takes place between the document metadata and the IFC schema. The documents are stored in a database and searched using SQL. In addition, they categorized basic options for creating a link between documents and objects in a project management environment: a) directly in the building model (using IFC), b) external link model, c) as part of an ontological model specification, and d) within the document meta-model.

Quattrini et al. [14] use the example of HBIM (Heritage BIM) to present a Semantic Web-related linking approach in which IFC data is stored in an RDF triplestore. In the Revit authoring software, the user can semantically enrich a BIM via IFC properties and export it in the IFC schema. Using an IFC-to-RDF tool, the information can be converted into triples and entered into a triplestore. Converters such as IFC to RDF play a decisive role in such approaches. For example, Bonduel et al. [15] developed an IFCtoLBD converter based on ifcOWL and the IFC-to-RDF converter, which focuses on linked (building) data graphs. This converter is also used by Jäkel et al. [16] to convert IFC models into RDF graphs. They go on to use Information Container for Linked Document Delivery (ICDD) for validation.

3.2   Nuclear Industry

In 2021, the French energy supplier EDF (Électricité de France) launched the plant life cycle management (PLiM) project. The aim is to improve information management—especially for new reactors. Overall, the central tasks of KM in nuclear facilities are to be addressed, including the internal exchange of knowledge, the exchange with external stakeholders, and the transfer of knowledge to future generations. Technologically, a kind of digital twin shall be created in which data can be consistently generated, validated, stored, and retrieved [8]. Akinyemi et al. [17] developed a data integration framework for offshore decommissioning waste management in the oil and gas domain. There are numerous cross-references between the oil and gas domain and the nuclear industry, which result from similarities in the complexity of the plants and the same safety culture. They analyzed the architecture of existing research approaches in the field of waste management for building assets and identified in the work of Bilal et al. [18] a very advanced approach with the use of Neo4j for data storage and ifcOWL for knowledge representation. In their own proposal, the documents are stored in a relational database, while the meta information is stored in RDF format in a triplestore. This information can be searched using SPARQL. To input the data, they define a spreadsheet template.

Pallas is a reactor for the production of medical isotopes that is currently under construction. Extensive efforts are being made here to test semantic interoperability in practice. The data exchange processes within a Common Data Environment (CDE) should be carried out following the standards RDF(S) and SPARQL. Breakdown structures and the IFC model of the plant form the backbone of the CDE [19].

4   Analysis

This chapter highlights the differences and similarities between the approaches to linking geometry with semantic information and documents in the AEC sector and the nuclear industry.

4.1   Life Cycle Phases and Challenges in the Knowledge Transfer

It is noticeable that KM is closely related to data and information management in practical implementation. Proximity to BIM has only recently been systematically established in nuclear technology through research projects such as Pallas [19]. Semantic web technologies are becoming increasingly important in this context. The IAEA also makes an explicit recommendation for their use in KM and data exchange [20]. In the AEC sector, the life cycle phases are weighted differently than in nuclear facilities. The most striking difference to the phase definition of the IAEA [8] is that decommissioning hardly plays a role in the AEC sector. For nuclear facilities, on the other hand, decommissioning is a decades-long, labor-intensive, and costly process with high strategic relevance. The experiences in the construction industry and in the nuclear industry are comparable in terms of the build-up of knowledge and, in particular, its loss in transition phases.

4.2   BIM and Spatial Anchoring of Knowledge and Documents

BIM is still considered a relatively "young" methodology in nuclear engineering. Here, KM methods are not primarily based on a BIM model, but on considerations of traditional document management. Accordingly, little attention is paid to the interdependencies of geometry and semantics in BIM. Instead, BIM is generally understood as "geometry". Is BIM-based KM therefore a suitable approach for nuclear facilities? A closer look shows that there are considerable implementation challenges, particularly in the operational phase. Current research and concepts of digital twins are increasingly linking BIM

to the operation phase. Cha and Lee [12] emphasize that information management is particularly important in this phase, but that BIM in its current implementation cannot be the sole anchor point for information management in this phase. Linking documents and knowledge with BIM elements ultimately results in the spatial allocation of this information to physical objects. Mandatory component identification systems are already used in NPPs, which serve as a reference for the analog storage of documents. Particularly during the operating phase, with regular maintenance and revisions to be carried out, the identification system enables the unique allocation of physical components —not only for analog document management purposes but perspectively also for digital KM.

4.3 Ontologies and Data Structures for Integrating Explicit and Implicit Knowledge

In the research and practical approaches considered in this work, there is little differentiation between documents, data, and knowledge. In classical KM, however, clear boundaries are drawn. The mere attribution of a BIM model in authoring software, as in Motawa and Almarshad [13] and Deshpande et al. [21] may or may not be equivalent to the input of knowledge [22]. The basic principle of the Semantic Web of mapping and exchanging knowledge at a granular level can also be used for mere information or data. Implicit knowledge plays a prominent role in the economic success of an organization [22]. However, current documentation requirements in both the construction and nuclear industries refer to explicit knowledge. Sobhkhiz and El-Diraby [23] point out that documented implicit knowledge often exists in unstructured data form, is difficult to integrate into existing structures with predominantly explicit knowledge and therefore remains predominantly in the heads of the knowledge carriers. Hu et al. [22] point out that even with highly developed knowledge organization systems such as ontologies and other technologies such as knowledge graphs, the hurdles to mapping experiential/implicit knowledge are high. For Sobhkhiz and El-Diraby [23] however, ontologies alone are no guarantee that unstructured data can be better linked. They point to the maintenance and further development effort for such structures, especially if relations between entities change in the future.

5 Concept

In the following, a software architecture for a fully digital KM platform for nuclear facilities is outlined. In earlier research [24], extended reality (XR) was discussed as a visualization technology in nuclear technology and was seen as a useful solution for enabling direct KM on-site. With the help of XR, spatially linked knowledge can be viewed on-site (AR), but a location-independent experience (VR) is also possible. The implementation in software is possible with the Unity Game Engine. A suitable headset that is both AR- and VR-capable is the Meta Quest 3. To view knowledge—meaning documents and granular knowledge elements linked to physical components reflected in a BIM model—together, a multiplayer mode must be established. This also enables verbal exchange, which is of great importance in analog KM settings. Meta provides corresponding functions for the Meta Quest 3 with the Colocation SDK, which are supplemented by the Photon Unity multiplayer network.

Given these constraints, how could data handling be organized? The lowest common denominator of the considered Semantic Web approaches is that they use RDF as their data model. The concept by Akinyemi et al. [17] considered in Section 3.2 proposes a sensible connection and storage of semantic data with regard to information management. This includes the outsourcing of binary files, a triplestore

for storing metadata, and a document meta-model for formalizing knowledge. Such a model with defined attributes can be used as a spreadsheet template for entering existing data, but also later as an "interface" for the user to insert implicit knowledge elements within the application.
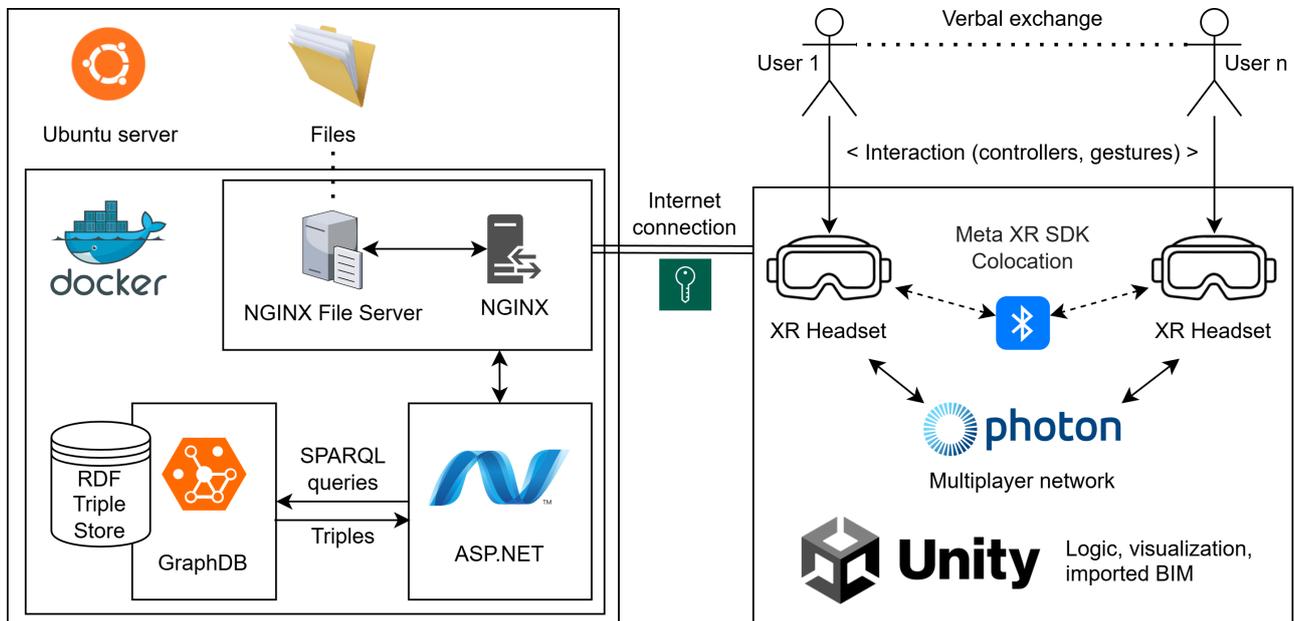

Figure 1: Proposed software architecture for an XR-based KM platform for nuclear facilities

In the software architecture presented here, an RDF triplestore is connected to Unity. The dotNetRDF framework is used to query and manipulate RDF data. A BIM model can either be deposited in Unity or linked to it. Depending on the method chosen, the component attributes are retained in the IFC schema and can be read out in Unity. The management of knowledge elements and binary data as well as their linking to geometries is carried out on an Ubuntu server using the document meta-model. The server uses Docker as an abstraction and security layer. There are three Docker containers: NGINX (file server and reverse proxy), ASP.NET (web framework for creating a REST API), and GraphDB (RDF triplestore). Figure 1 shows the software architecture developed. RDF is used in numerous ontologies and vocabularies, is flexible and expandable. Therefore, due to the RDF data structure, ontologies can be added as required without having to change the basic model.

6 Discussion and Conclusion

In this contribution, an XR-based KM platform for use in nuclear facilities was proposed incorporating established principles from BIM and the Semantic Web. In future research, this software architecture is to be further developed into a prototypical information system and tested in a suitable environment. The overall goal is to establish a higher-level approach that can be adapted and refined based on existing geometric and semantic data, depending on the specific use case. Hu et al. [22] see potential in the higher-level field of knowledge science through technologies such as IoT, AI, cloud computing, and blockchain. The platform presented here can generally be expanded using these technologies. In particular, AI functionalities such as RAG (Retrieval-Augmented Generation) could usefully supplement the classic querying of information from the graph database and binary data.

References

[1] International Atomic Energy Agency, *IAEA Nuclear Safety and Security Glossary: Terminology Used in Nuclear Safety, Nuclear Security, Radiation Protection and Emergency Preparedness and Response* (Non-Serial Publication), 1st ed. Vienna: International Atomic Energy Agency, 2022.

[2] International Atomic Energy Agency, *Initiating nuclear power programmes: Responsibilities and capabilities of owners and operators: Iaea nuclear energy series no. ng-t-3.1 (rev. 1)*, 2020.

[3] International Atomic Energy Agency, *Guide to knowledge management strategies and approaches in nuclear energy organizations and facilities: Iaea nuclear energy series no. ng-g-6.1*, Vienna, 2023.

[4] F. Opitz, R. Windisch, and R. J. Scherer, "Integration of document- and model-based building information for project management support", *Procedia Engineering*, vol. 85, pp. 403–411, 2014. DOI: 10.1016/j.proeng.2014.10.566

[5] A. Borrmann, M. König, C. Koch, and J. Beetz, editors, *Building information modeling: Technologische Grundlagen und industrielle Praxis* (VDI-Buch), 2., aktualisierte Auflage. Wiesbaden and Heidelberg: Springer Vieweg, 2021.

[6] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities", *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001. DOI: 10.1038/scientificamerican0501-34

[7] P. Hagedorn et al., "Bim-enabled infrastructure asset management using information containers and semantic web", *Journal of Computing in Civil Engineering*, vol. 37, no. 1, 2023. DOI: 10.1061/(ASCE)CP.1943-5487.0001051

[8] International Atomic Energy Agency, *Application of plant information models to manage design knowledge through the nuclear power plant life cycle* (IAEA TECDOC Series). Vienna: IAEA, 2020, vol. no. 1919.

[9] EPRI, "Advanced nuclear technology: New nuclear power plant information turnover guide (revision 1)",

[10] C. H. Caldas, L. Soibelman, and L. Gasser, "Methodology for the integration of project documents in model-based information systems", *Journal of Computing in Civil Engineering*, vol. 19, no. 1, pp. 25–33, 2005. DOI: 10.1061/(ASCE)0887-3801(2005)19:1(25)

[11] K. W. Chau, M. Anson, and J. P. Zhang, "4d dynamic based management and visualization software: 1. development", *Automation in Construction*, vol. 14, no. 4, pp. 512–524, 2005.

[12] H. Cha and D. Lee, "Framework based on building information modelling for information management by linking construction documents to design objects", *Journal of Asian Architecture and Building Engineering*, vol. 17, no. 2, pp. 329–336, 2018. DOI: 10.3130/jaabe.17.329

[13] I. Motawa and A. Almarshad, "A knowledge-based bim system for building maintenance", *Automation in Construction*, vol. 29, pp. 173–182, 2013. DOI: 10.1016/j.autcon.2012.09.008

[14] R. Quattrini, R. Pierdicca, and C. Morbidoni, "Knowledge-based data enrichment for hbim: Exploring high-quality models using the semantic-web", *Journal of Cultural Heritage*, vol. 28, pp. 129–139, 2017. DOI: 10.1016/j.culher.2017.05.004

[15] M. Bonduel, J. Oraskari, P. Pauwels, M. Vergauwen, and R. Klein, "The ifc to linked building data converter : Current status", 2018.

[16] J.-I. Jäkel, E. Heinlein, H. Morgenstern, H. Kim, and K. Klemt-Albert, "System- and data-integrated linking of digital 3d models of existing bridge structures with knowledge graphs of non-destructive diagnostic methods", *Journal of Information Technology in Construction*, vol. 30, pp. 603–630, 2025. DOI: 10.36680/j.itcon.2025.025

[17] A. G. Akinyemi, M. Sun, and A. Gray, "Data integration for offshore decommissioning waste management", *Automation in Construction*, vol. 109, p. 103 010, 2020. DOI: 10.1016/j.autcon.2019.103010

[18] M. Bilal et al., "Big data architecture for construction waste analytics (cwa): A conceptual framework", *Journal of Building Engineering*, vol. 6, pp. 144–156, 2016. DOI: 10.1016/j.jobe.2016.03.002

[19] L. van Ruijven, *The pallas project - a practical experience of semantic interoperability: Complexity and capabilities of skos, rdf and owl/shacl*, 2023.

[20] International Atomic Energy Agency, *Exploring Semantic Technologies and Their Application to Nuclear Knowledge Management* (IAEA Nuclear Energy Series). Vienna: IAEA, 2021.

[21] A. Deshpande, S. Azhar, and S. Amireddy, "A framework for a bim-based knowledge management system", *Procedia Engineering*, vol. 85, pp. 113–122, 2014. DOI: 10.1016/j.proeng.2014.10.535

[22] Z.-Z. Hu, S. Leng, J.-R. Lin, S.-W. Li, and Y.-Q. Xiao, "Knowledge extraction and discovery based on bim: A critical review and future directions", *Archives of Computational Methods in Engineering*, vol. 29, no. 1, pp. 335–356, 2022. DOI: 10.1007/s11831-021-09576-9

[23] S. Sobhkhiz and T. El-Diraby, "Dynamic integration of unstructured data with bim using a no-model approach based on machine learning and concept networks", *Automation in Construction*, vol. 150, p. 104 859, 2023. DOI: 10.1016/j.autcon.2023.104859

[24] P. Mosler and U. Rüppel, "Designing a knowledge-centric plant information model by integrating tacit knowledge", *31st International Conference on Nuclear Engineering*, 2024. DOI: 10.1115/ICONE31-135585