

Final Report

RDM-Workflows for Construction Engineering and Architecture

1. General Information

- Proposer: Prof. Achim Menges, Dr. Tobias Schwinn & Hana Svatos-Raznjevic (ICD), Gerald Jagusch & Roger Winkler (ULB)
- Affiliation: Institute for Computational Design and Construction (ICD), University of Stuttgart & University and State Library, Technical University of Darmstadt
- Project title: RDM-Workflows for construction engineering and architecture – Development and integration of exemplary application profiles for timber construction data
- Project runtime: 04/2023–04/2024

2. Executive Summary

The "RDM-Workflows for Construction Engineering and Architecture" project aimed to address the development of domain-specific infrastructure and low publication rates of research data within architecture and civil engineering. The project focused on enhancing data interoperability and accessibility by developing workflows that transform tabular research data into enriched knowledge graphs, utilizing application profiles created with the NFDI4ING Metadata Profile Service.

Key achievements include the creation of application profiles for datasets related to multi-storey timber buildings and stakeholders, the development of the CSV-RDF-Mapper for converting tabular data into knowledge graphs, and the publication of project data and tools on GitHub. These efforts were complemented by engagement with the architectural research community to assess needs and demonstrate the workflow's value.

The project also identified significant challenges faced by researchers, such as fragmented data environments, the need for improved data literacy, and the complexities of using ontologies for adequate data representation. Collaborative efforts with the NFDI4Ing consortium highlighted the potential for broader applications of the developed tools and workflows beyond architecture, suggesting adaptability to other disciplines.

The outcomes of this project have laid the groundwork for a follow-up initiative under the DFG funding program, aimed at further automating the workflow for data collection, semantic enrichment, and publication. This future project aspires to empower architectural researchers to produce FAIR Digital Objects, enhancing the

accessibility and usability of research data within the NFDi4ing framework. The project's findings and tools were disseminated through various academic and professional platforms, contributing to the ongoing dialogue on research data management in the built environment.

3. Work Report and Results

Contemporary research on the built environment is increasingly interdisciplinary, with ecology and sustainability driving current debates. Emerging methods like computational design and engineering, which involve digitally integrated design, planning, and manufacturing processes, offer innovative solutions by leveraging data in various states. However, the researchers' awareness and domain specific infrastructure for publishing research data in architecture and civil engineering are still underdeveloped. While data mostly remains inaccessible as data is barely published in structured forms, published data often lacks detailed, domain-specific metadata, making it difficult for researchers to effectively search, find, and reuse data according to the FAIR [1] principles.

The main goal of the project “RDM-Workflows for construction engineering and architecture” [2] developed into showcasing a workflow that improves the interoperability of not only a dataset's metadata, as seen in existing solutions like COSCINE or the Data Ingest Service within NFDi4ING, but the data itself, utilizing application profiles created with the NFDi4ING Metadata Profile Service [3]. Existing, representative datasets on multistorey timber buildings and stakeholders [4-8] (Fig. 2), created and published by the ICD, served as a basis for the technical development. At the same time they were used for the engagement with other researchers in the field as throughout the project it became clear that efforts were needed to explore why research data publications in architectural research are still low and underrepresented in the addressed communities.



Fig. 1: Published dataset in DARUS by the ICD

To create application profiles for describing tabular research data, a comprehensive list of vocabularies and ontologies relevant to the architectural domain with a focus on buildings and stakeholders was collected and evaluated. A subset of these were subsequently added to the NFDi4ING Terminology Service, as some were not listed yet (WP1).

Using the column headers of the original tabular data and terms from the collected vocabularies, exemplary application profiles for buildings and stakeholders were modeled using the NFDi4ING Metadata Profile Service. In this context, also a list of data presentation websites was created, in order to compare and discuss the schemas as data representations. As the drawings dataset turned out to be more complex and further discussion and involvement with the community is needed, it became a part of the follow up funding application (WP 2).

As the projects FID BAUdigital and the services of NFDi4ing evolved during the project's runtime, the focus shifted from merely developing an application profile that describes datasets to creating application profiles that transform disconnected tabular data into linked and enriched knowledge graphs. Therefore, no new dataset has been published in the NFDi4ING Data Ingest Service, the former FID BAUdigital Workbench. Instead, the CSV-RDF-Mapper was developed to facilitate the transformation of the original tabular data into a new knowledge graph utilizing the application profiles (WP 3).

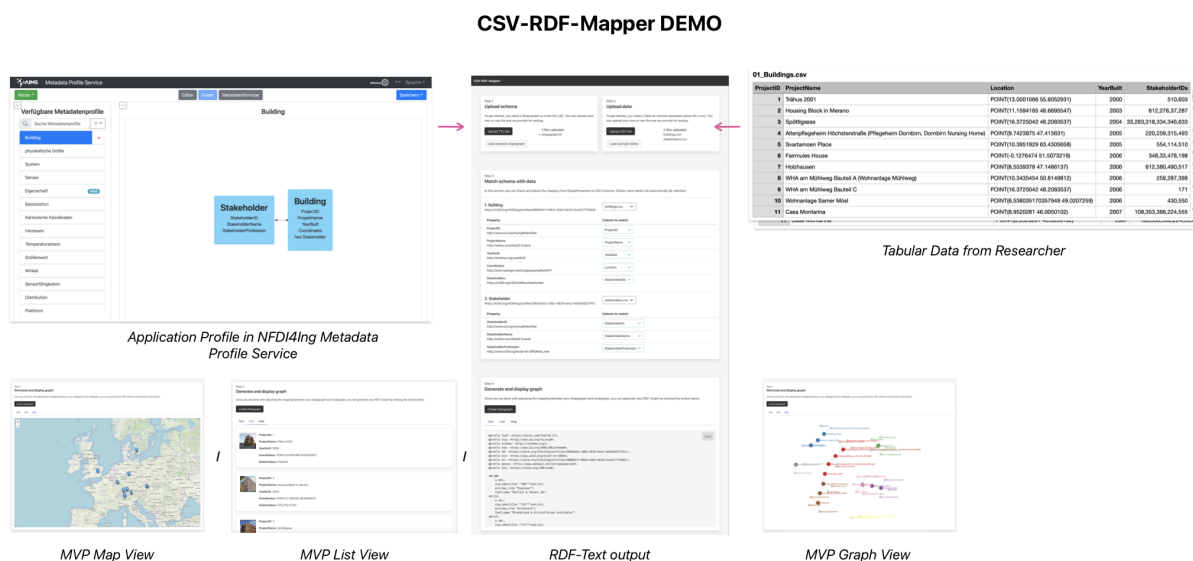


Fig. 2: Application profile in the NFDi4ING Metadata Profile Service (upper left), exemplary tabular data on buildings and stakeholders (upper right), the CSV-RDF-Mapper Demo (center) with example data visualizations and presentations as map, list, text or graph (bottom)

The developed workflow and application profiles were then presented to representatives of the architectural research community, practicing architects as well as relevant stakeholders from the infrastructural domain in diverse formats, including interviews, but also public presentations such as the SIGIDUS Seminar Series by the University of Stuttgart in June 2024. As we aimed to demonstrate the feasibility and value of the workflow and to evaluate the community's interest in further project development, this engagement helped to assess the specific needs of the community in more detail and now serve as a basis for the follow up funding application (WP 4).

The project generated a variety of data, all of which was collected and published in a GitHub repository.¹ This repository includes the list of domain-specific terminologies, showcase application profiles and datasets for buildings and stakeholders, the software code of the CSV-RDF-Mapper with a demo application (via GitHub Pages), as well as example RDF graph data results. In addition, a detailed documentation of the workflow and a list of relevant websites that present architectural data are attached to the GitHub repository as a wiki² (WP 5)

This project in general served and enabled the applicants to develop a subsequent project within the DFG funding program "e-Research Technologies." This upcoming project aims to further develop the workflow into an automated pipeline for data collection, semantic lifting, enrichment, and publication. The ultimate goal is to enable architectural researchers to create enhanced FAIR Digital Objects based on the NFDI4ING Services framework. The project allowed the team to identify and explore the main challenges and barriers that architectural researchers face when attempting to publish their research data. These challenges include connecting currently fragmented data collection and processing environments, improving researchers' data literacy and research data management (RDM) knowledge, preparing data so that it is understandable, traceable, and trustworthy, and effectively using ontologies and authority data to achieve adequate representation.

Collaboration and exchange with the NFDI4Ing consortium, specifically with the base service measure S-3 (metadata & terminology services) and the Archetype Alex and Ellen teams, revealed that the developed tools and workflows have potential applications beyond the discipline of architecture. The workflow, which semantically transforms tabular data into knowledge graphs, could be adapted to other fields and further enriched. Archetype Alex, in particular, saw potential for integrating semi-structured formats like HDF5 into the workflow. Additionally, the S-3 measure identified that the developed tool might be suitable for harvesting and semantifying data and metadata, thus enhancing its usability within the larger NFDI4Ing framework.

Acknowledgement

"The authors would like to thank the Federal Government and the Heads of Government of the Länder, as well as the Joint Science Conference (GWK), for their funding and support within the framework of the NFDI4Ing consortium. Funded by the German Research Foundation (DFG) - project number 442146713."

Literature and Links

- [1] Wilkinson, M. D. et al. (2016). 'The FAIR Guiding Principles for scientific data management and stewardship', Sci-entific data 3:160018. DOI: <https://doi.org/10.1038/sdata.2016.18>.
- [2] Winkler, R., & Svatoš-Ražnjević, H. (2023). NFDI4Ing Seed Fund Proposal 2022: RDM-Workflows for construction engineering and architecture. Zenodo. <https://doi.org/10.5281/zenodo.7802981>.
- [3] Nils Preuß, Matthias Bodenbenner, Benedikt Heinrichs, Jürgen Windeck, Mario Moser und Marc Fuhrmans. Creating application-specific metadata profiles while improving interoperability and consistency of research data for

¹ <https://github.com/ULB-Darmstadt/csv-rdf-mapper>, 13.08.2024

² <https://github.com/ULB-Darmstadt/csv-rdf-mapper/wiki>, 13.08.2024



the engineering sciences. Techn. Ber. Submitted paper to IEEE Transactions on Knowledge and Data Engineering. Sep. 2023. <https://doi.org/10.26083/tuprints-00024573>.

- [4] Svatoš-Ražnjević, Hana; Menges, Achim, 2022, "Multi-storey Timber Buildings Data: Architectural and Structural Data on 350 Mass-Timber Projects from 2000-2021", <https://doi.org/10.18419/darus-2733>, V1
- [5] Svatoš-Ražnjević, H.; Orozco, L.; Menges, A. Advanced Timber Construction Industry: A Review of 350 Multi-Storey Timber Projects from 2000–2021. Buildings 2022, 12, 404. <https://doi.org/10.3390/buildings12040404>.
- [6] Orozco, Luis; Svatoš-Ražnjević, Hana; Menges, Achim, 2022, "Stakeholders in Multi-storey Timber Data", <https://doi.org/10.18419/darus-2740>, DaRUS, V2
- [7] Orozco, L.; Svatoš-Ražnjević, H.; Wagner, H.J.; Abdelaal, M.; Amtsberg, F.; Weiskopf, D.; Menges, A. Advanced Timber Construction Industry: A Quantitative Review of 646 Global Design and Construction Stakeholders. Buildings 2023, 13, 2287. <https://doi.org/10.3390/buildings13092287>.
- [8] Svatoš-Ražnjević, Hana; Menges, Achim, 2024, "Multi-storey Timber Buildings and Design and Construction Stakeholder Constellations Data: 99 DACH Projects", <https://doi.org/10.18419/darus-4442>, DaRUS, V1