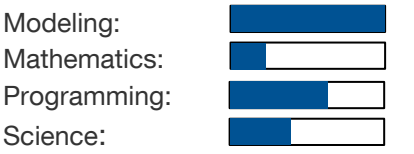
Chair of Computational Modeling and Simulation
TUM School of Engineering and Design
Technical University of Munich



Software Lab:

RDF-based catalogue system for construction monitoring data

Description

In the area of big data, data management aspects become increasingly important to find and access data entries quickly. In the context of construction sites, it becomes common that large amounts of data are collected to monitor the progress of the site. This can include, for example, images, point clouds, measurements of temperature sensors, and many others. Catalogue systems help create and manage metadata about data entries to speed up data queries and access. The team shall use an RDF graph to manage such a catalogue of raw monitoring data. RDF (Resource Description Framework) is a technology commonly used on the web to describe knowledge graphs with clear semantic meaning. It shall be used to create and manage a graph that holds metadata about data captured on a construction site, e.g., containing information about the date of all images that were captured by a particular camera, the number of sensors used on the construction site, or the location where a point cloud from result from a laser scan is stored. While the RDF graphs hold only information about metadata, there should be dedicated databases that store the actual monitoring data.

Main Tasks:

- Set up an RDF graph database and other DBs to host the data catalogue and the monitoring data.
- Investigate existing data models for raw data storage suitable for construction monitoring data (e.g. SOSA/SSN, DCAT, etc.) → decide on a schema or a combination of schemata that defines your RDF catalogue will be structured.
- Develop automated and manual methods to feed data records into the catalogue
    - Create a User Interface (UI) for manual data upload (e.g. a web application). The manual upload should result in storing the raw data in its dedicated DB and adding a metadata entry to the RDF graph.
    - Automated integration of live data streams (e.g. sensors that send data frequently)
- Provide functionalities for quick catalogue querying

It is of utmost importance that a wide range of different types of raw data and raw data formats are taken into account for the developed catalogue systems.

Supervisor

Jonas Schlenger, TUM Chair of Computational Modelling and Simulation, jonas.schlenger@tum.de

References

[1] R. Cyganiak, D. Wood, and M. Lanthaler (2014). RDF 1.1 Concepts and Abstract Syntax. Accessible: https://www.w3.org/TR/rdf11-concepts/

[2] H. Dibowski et al.(2020). Using Semantic Technologies to Manage a Data Lake: Data Catalog, Provenance and Access Control. 13th International Workshop on Scalable Semantic Web Knowledge Base Systems.

[3] A. Haller et al. (2017). Semantic Sensor Network Ontology. Accessible: https://www.w3.org/TR/vocab-ssn/