

# Semantisation of Rules for Automated Compliance Checking

Edlira Vakaj<sup>1</sup>, Maxime Lefrançois<sup>2</sup>, Amna Dridi<sup>1</sup>, Thomas Beach<sup>3</sup>, Mohamed Gaber<sup>1</sup>, Gonçal Costa Jutglar<sup>4</sup> and He Tan<sup>5</sup>

<sup>1</sup> Faculty of Computing, Engineering and Built Environment, Birmingham City University, Birmingham, UK

<sup>2</sup> Mines Saint-Etienne, Univ Clermont Auvergne, INP Clermont Auvergne, CNRS, UMR 6158 LIMOS, F - 42023 Saint-Etienne France

<sup>3</sup> School of Engineering, Cardiff University, Cardiff, UK

<sup>4</sup> IAR – Integrated Architecture Research, La Salle – Ramon Llull Univeristy, 08022 Barcelona, Catalonia, Spain

<sup>5</sup> Department of Computer Science and Informatics, School of Engineering, Jönköping University, Box 1026, 551 11, Jönköping, Sweden

## Abstract

The Architecture, Engineering, and Construction (AEC) industry is subject to numerous regulations and standards that govern the design, construction, and maintenance of buildings and infrastructure. These regulations often involve complex language and technical jargon, which can be difficult to understand and apply in practice. Semantisation, or the process of transforming natural language into machine-readable data with explicit meaning, can address this challenge by creating structured representations of regulations that can be exchanged and processed by computers. This can enable automated compliance checking, facilitate communication between stakeholders, and improve the efficiency and effectiveness of regulatory enforcement.

Semantic Web (SW) provides a framework for integrating and processing data from various sources, which can be used to automate compliance checking for buildings. Ontologies, Linked Data, Reasoners and rule-based systems are some of the SW technologies that can be applied to building compliance checking. Jointly, other Artificial Intelligence (AI) methods such as Natural Language Processing (NLP) can be used to extract and analyse compliance requirements from natural language text in an automated way. However, there are several challenges in using SW technologies and NLP in building compliance checking, such as the need for accurate and complete data, the complexity of modelling compliance requirements, and the difficulty of interpreting natural language text.

Automated Compliance Checks for Construction, Renovation or Demolition Works (ACCORD1) is a Horizon Europe project that aims to digitalise permitting and compliance processes. ACCORD will use BIM and other data sources to improve the productivity and quality of design and construction processes. It will, then, aim at contributing to the development of a semantic framework for European digital building permitting processes, regulations, data and tools.

In this talk, we will discuss the opportunity and challenges that SW and AI bring to rule formalisation and standardisation in the AEC domain. This has important implications for the construction industry, where compliance is critical for ensuring safety, quality, and sustainability of construction projects. This work is carried under ACCORD's Work Package (WP) 2 "Semantisation of regulations and open format for machine-readable rules". This WP's objective is to develop an ontology that models building compliance requirements, including laws, regulations, processes, and documentation. Developing such ontology that captures the overall domain is an enormous task. Therefore, we are investigating the potential alignment with existing ontologies and data models. The ontology requirements will be derived from the rule formalisation methodology and an extensive literature review conducted in two separate tasks. The methodology defines different pathways (manual, semi-automatic and automatic) on digitising and/or formalising rules, and subsequently instantiate the ontologies of rules. The extracted rules can be encoded as RDF to create a knowledge graph that can be used for automated compliance checking. This knowledge graph can be further enriched by incorporating the rules extracted from the textual data such as building codes, standards, and regulations. By leveraging the power of Machine Learning and NLP, the process of extracting rules from building regulation text can be automated, thereby reducing the time and effort required for compliance checking. A whole task is dedicated to the automatic extraction of

rules from regulatory text. Despite the challenges of natural language, NLP and more specifically Large Language Models seem a very promising solutions to automate rule extraction. Whereas, the manual and semi-manual approach will be facilitated by the Graphical Rule Formalisation tool developed on a separate task. The outputs of this WP will be used internally in the ACCORD compliance checking orchestration platform and shared with the community for further development and use..

### **Keywords**

Machine Learning, Ontologies and Building Data Standards (e.g. IFC), Knowledge graphs, Rule/compliance checking

---

Proceedings LDAC2023 – 11th Linked Data in Architecture and Construction, June 15–16, 2023, Matera, Italy

EMAIL: edlira.vakaj@bcu.ac.uk (E. Vakaj); maxime.lefrancois@emse.fr (M. Lefrançois); amna.dridi@bcu.ac.uk (A. Dridi); beachth@cardiff.ac.uk (T. Beach); mohamed.gaber@bcu.ac.uk (M. Gaber); goncal.costa@salle.url.edu (G. Costa Jutglar); he.tan@ju.se (H. Tan)

ORCID: 0000-0002-0712-0959 (E. Vakaj); 0000-0001-9814-8991 (M. Lefrançois); 0000-0003-0339-4474 (M. Gaber); 0000-0001-9168-680X (G. Costa Jutglar); 0000-0003-1303-9791 (H. Tan)



© 2023 Copyright for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)