# SDM

##### Session 3.1 (SEMANTiCS)

#### Time: Friday, September 22, 2023 - 10:30 to 12:00

#### Chair: Tassilo Pellegrini, Co-Director, Institute of Innovation Systems

## **Talks**

### Semantifying the PlanQK Platform and Ecosystem for Quantum Applications

Purpose: The goal of this paper is to report about an application of a knowledge semantification approach for the “platform and ecosystem for quantum applications” (PlanQK). Specifically, we describe how the quantum computing knowledge, which is submitted on the platform by scientific community and industry actors, is incorporated into the platform knowledge graph (KG) and semantically annotated. Furthermore, we delve into the KG-based semantic search we developed for the platform.

Methodology: To achieve the data transformation from a relational database (RDB) into a machine-readable Resource Description Framework (RDF) format, we define a mapping between the database and ontology schemes, handle the data semantics and store the instance data in a unified KG. Moreover, we discuss a mechanism for a continuous synchronization of the knowledge in the graph. The semantic search we present further in this paper incorporates both native and faceted search techniques, leveraging the indexed parts of the KG.

Findings: The paper shows an on-the-fly solution that seamlessly integrates ontology and a web platform data stored in a relation database into a KG. It outlines the semantic search over the graph, which is not restricted to a single graph but can be also executed across an ecosystem of the connected graphs.

Value: RDB to RDF mapping and semantic search are complex tasks that require an architecture, which supports an effortless data integration and evolution over time. The presented solution is not limited to the PlanQK platform and quantum computing domain but can be applied in other applications as well. In further work, the PlanQK KG could be made available as Linked Data and connected with related graphs to build a quantum computing knowledge system.

| Darya MartyniukFraunhofer Fokus | Researcher at Data Analytics Center (DANA), Fraunhofer FOKUS. Research interests and project tasks are focused on knowledge representation and Semantic Web technologies as well as quantum machine learning. |
| --- | --- |

### Streamlining Data Intelligence: A Unified Platform for Managing and Analyzing Process Data in Industrial Settings

This conference submission presents a cutting-edge platform designed to address the challenges associated with managing and analyzing process and process-analytical data in industrial settings. Such data exists in diverse formats, including instrument data and text data, which makes it crucial to have a unified approach for data access and analysis. Our proposed platform serves as a centralized hub, providing a single point of access for searching, contextualizing, visualizing, and analyzing all types of process data.

The platform incorporates a robust semantic layer and semantic search capabilities, which enable users to effectively navigate and explore the data landscape. By leveraging semantic technologies, the platform goes beyond basic keyword-based search, allowing for more intuitive and context-driven data discovery. This ensures that users can easily locate and access the information they need, regardless of its format or source.

Furthermore, our platform offers both self-service and predefined data objects, catering to the diverse needs of stakeholders across different departments. This flexibility empowers users to generate customized analyses and reports, promoting data-driven decision-making within their respective domains.

To ensure seamless integration and utilization, our platform is carefully instrumented and consumed by stakeholders. We provide a comprehensive overview of the implementation process, highlighting the steps involved in setting up and configuring the platform to align with specific organizational requirements. Additionally, we showcase real-world use cases and success stories from multiple departments, illustrating how the platform has revolutionized their data-driven practices.

| Sebastian GablerSemantic Web Company |  |
| --- | --- |

### Managing the metadata of a diverse digital media archive as a Knowledge Graph

Meemoo manages a large quantity of mainly audiovisual material from more than 170 partners in cultural heritage and media. More than 6 million objects are currently stored, ranging from digitised newspapers, photos, videos, and audio. In addition, a number of access platforms make the digitised content available to specific target groups, including teachers, students, professional re-users, or the public.

Metadata is a key element in all of meemoo’s processes. An important part of our activities is to collect, integrate, manage, and search a large variety of heterogeneous metadata across the archived content. The scale of this has increased enormously, so a good and integrated approach is needed to deal with the amount of metadata, its need for flexibility, and how easy it is to find. One of the specific challenges is modelling and storing data from machine learning algorithms (speech recognition, face detection and entity recognition) for reuse.

In this talk, we will discuss the key points and lessons learned from implementing the new metadata roadmap at meemoo (<https://meemoo.be/en/publications/metadata-roadmap-the-route-to-an-improved-metadata-infrastructure>), which is focused on a semantic Knowledge Graph-based infrastructure. The goal of the roadmap is to establish a better data practice within the organization and offer application-independent, uniform access to (meta)data that is spread across various systems and formats. The implementation consists of extensive data modeling, using RDF, SKOS, SHACL to construct a uniform view over all relevant data, and using SPARQL in combination with GraphQL for establishing efficient access by platforms.

| Miel Vander SandeFlemish Institute for Archives |  |
| --- | --- |