# **Session 3.4 Ontologies**

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

#### Chair: TBA

## **Talks**

### Use of healthcare ontologies at scale in IQVIA

In IQVIA we process and integrate healthcare data at scale (100B+ medical records per year). In our operations, we make a varied use of ontologies: to develop metadata schema to annotate data quality; to standardise terminologies and code lists to enable scalable analytics (e.g.: as per OMOP standard for Real World Evidence); or to standardise data for regulatory submissions (e.g.: as in CDISC for FDA clinical trials submissions).

In this talk, we would like to showcase some of the peculiarities and learnings of working with ontologies at such scale (note that in the healthcare space, “ontologies” encompass a broad range of artefacts from code lists to DL-based representations).

First, we stratify our ontologies in three different layers:

• “standard ontologies”: public or not, these are the standards in use, that form the basis for interoperability.

• “actual ontologies”: as a result of the broad observation of the healthcare system that we perform, we collect a variety of terms in use that don’t directly map to standard ontologies (>44M). These can correspond to different names, language or regional variations, new concepts not yet standardised (e.g.: a newly released covid self-test) or simply concepts not captured by standards (e.g.: laboratory consumables).

• “curated ontologies”: for many of our analytics need, we need extended reference data that is typically not found in the public space (e.g.: comprehensive normalisation of products across markets, or identification of healthcare organizations).

Standard ontologies pose different order of challenges, like different variations and multiple coexisting versions. Often standard ontologies “in use” are also affected by some amount of noise (e.g.: accidental mix of ontologies to diagnose an indication). We intend to showcase some of the internal tooling that we have developed to address such problems.

Actual ontologies present a view of what of an ontology is practically in use: (e.g.: not all terms of branches of ontologies are used, some probably representing theoretical concepts not existing in clinical practice). A question to foster a discussion with the audience is to what extent the gap between “standard ontologies” (a-priori) or “actual ontologies” (a-posteriori) is at the basis of interoperability issues.

Finally, we note that our “curated ontologies” (missing ontologies we developed to support analytics), rather present information typical of MDM system. We would then conclude with a framework we are using to merge the MDM and the ontology mindsets.

---

I am currently a full time employee (Principal) in IQVIA (An SP500 healthcare data company), focusing on internal optimization and product development in the area of ontology and metadata. I also organise http://swat4ls.org (conference on semantic technologies in life sciences) as a pro-bono activity.

In terms of expertise and background:

My PhD thesis was on the use of ontologies, and I have published a few papers on ontologies, semantics and FAIR: https://orcid.org/0000-0002-3201-9617

Before joining IQVIA I was (in reverse order):

Director of data strategy in Novartis (corporate): proposing KGs to link data across the enterprise

Associate director information systems at Novartis: supervising an ontology provision system and working on data curation semi-automation

Independent consultant: semantic POCs for clinical trials, nutrition, meteorology (WHO) and more

Staff scientist: data integration based on ontologies

Postdoc in medical terminologies

I know mostly focus on ontology business cases, technology development, change management and communication.

| Andrea Splendiani |  |
| --- | --- |
|  |  |

### Semantic Models - The brains behind explainable and trustworthy data-driven business decisions [SP]

Semantic models, or ontologies, are essential for building knowledge graphs that enable explainable and reliable business decisions. They enrich data with context and meaning, facilitating comprehensive analyses and knowledge discovery. These models drive symbolic AI solutions and can augment data-driven AI with a layer of trust. This talk will introduce metaphactory's visual interface as the gold standard for semantic knowledge modeling and will highlight how metaphactory empowers SMEs and business users to actively participate in and drive the modeling process. It will also cover the integration of semantic models with additional assets, such as vocabularies, enabling data publishing, data curation, semantic search, and use cases like enterprise architecture modeling and business knowledge modeling.

| Sebastian Schmidt, CEO, metaphacts | metaphacts  <https://metaphacts.com/> |
| --- | --- |
|  |  |

#### Ontology and Knowledge Graphs for API-Based Search in an Analytics Self-Service App Store

The transformation towards a data-driven organization has resulted in the rapid growth of analytics self-service tools at Daimler Truck. Discovering and utilizing analytics self-service tools within the app store poses challenges due to the lack of effective search functionality. This presentation introduces a solution that leverages an ontology and knowledge graphs to implement a powerful API-based search in an app store for analytics self-service tools.

We'll begin by explaining the importance of using ontologies in structuring knowledge and explore the use of knowledge graphs as a means to represent and store ontology data. The domain-specific ontology is designed and developed by Daimler Truck captures essential concepts, relationships, and attributes in the analytics self-service tool domain. The knowledge graph integrates app store tools and metadata based on the defined ontology. The implemented search functionality utilizes APIs, allowing users to query the app store based on specific requirements. Semantic reasoning enhances the search results, providing accurate and context-aware tool recommendations.

Finally, we'll discuss the benefits of using ontologies and knowledge graphs for search functionality, including enriched metadata and adaptability for adding new tools and features. Infrastructure and setup will be presented as well.

In conclusion, this presentation showcases the utilization of ontologies and knowledge graphs to implement API-based search in an analytics self-service tools app store. The proposed approach enhances tool discovery, facilitates informed decision-making, and ensures adaptability to the evolving analytics landscape.

| Jannik Wiessler | Data Scientist bei Daimler Truck AG |
| --- | --- |
|  |  |