# Knowledge Graphs

#### Time: Thursday, September 21, 2023 - 10:45 to 13:00

#### Chair: TBA

## **Talks**

### Methodology for Creating a Knowledge Graph with RDF Reification

Purpose: This paper presents the construction of a Knowledge Graph (KG) of Educational Resources (ER), where RDF reification is essential. The ERs are described based on the subjects they cover considering their relevance. RDF reification is used to incorporate this subject's relevance. Multiple reification models with distinct syntax and performance implications for storage and query processing exist. This study aims to experimentally compare four statement-based reification models on four triplestores to determine the most pertinent choice for our KG.

Methodology: The authors built four versions of the KG. Each version has a distinct reification model, namely standard reification, singleton properties, named graphs, and RDF-star, which were obtained using RML mappings. The KG consists of 45,000 ERs, 13,000 authors, 135,000 subjects, and 8,250,000 statements linking the ERs to their subjects. Each of the four triplestores (Virtuoso, Jena, Oxigraph, and GraphDB) was set up four times (except for Virtuoso, which does not support RDF-star), and seven different SPARQL queries were experimentally evaluated.

Findings: In general, this study shows that standard reification and named graphs lead to good performance. It also shows that Virtuoso outperforms Jena, GraphDB, and Oxigraph in most queries, in the particular context of the used KG.

Value: The recent specification of RDF-star and SPARQL-star sheds light on statement-level annotations. The empirical study reported in this paper contributes to the efforts toward the efficient usage of RDF reification. In addition, this paper shares the pipeline of the KG construction using standard semantic web technologies.

| name | CV |
| --- | --- |

### 

### Create Enterprise AI assistants based on LLM in minutes! [SP]

LLM understands language to an unprecedented level. In this talk, we present how we can combine enterprise data like PDFs, Website and Knowledge Graphs (KGs) with LLM in order to construct AI assistants in enterprise settings. These ranges from search to chatbot interfaces that: 1) are easy to setup, 2) answer questions in an unprecedented way! We will show how you can answer questions easily on an uploaded PDF, on a entire website domain or an RDF Knowledge Graph.

| Dennis Diefenbach, Founder & CEO / CTO QA Company | QA Company  <https://the-qa-company.com/> |
| --- | --- |

### 

### Why an Event-Native Mindset is Essential for Knowledge Graphs

Gartner has noted that it is not enough for businesses today to be “ready to change,” but instead, companies need to be “ready to act” in real-time and understand the context of the action. Simply being prepared to respond to change is not sufficient in today’s fast-paced and constantly evolving business environment. Organizations need to be able to anticipate and proactively respond to changes in order to stay competitive.

But how can a business create a data architecture that supports ‘ready to act’ applications and systems? Gartner suggests embracing an Event–Native Mindset and with an Event-Driven Architecture that delivers continuous intelligence and keeps the business always ready.

Applying an Event–Native Mindset to Data Modeling

Consider for a moment that everything that happens within a business environment is an event, and every event impacts an entity or is carried out by an entity. An entity in this context is a core business concept like a customer, patient or product. Everything a patient does – getting diagnosed, visiting a specialist, being discharged or receiving a prescription – is an event. Anything that happens to a business’ customer, from making purchases, returns or calling for support, is an event. When products are created, tested, and updated, these activities are also events.

By adopting an Event–Native Mindset in data modeling, organizations can develop more agile and responsive data architectures that are better suited to handle complex and rapidly changing data environments. This approach can help organizations to more quickly identify and respond to changes in data patterns, and to more effectively leverage the insights and value that can be derived from event-driven data models.

During this presentation we will discuss best practices for building Event-Driven Knowledge Graphs based on several industry projects in Pharmaceuticals, Manufacturing and Consumer Service Centers.

| [Jans Aasman](https://franz.com/) | CEO at Franz Inc |
| --- | --- |
| [Sheng-Chuan Wu](https://franz.com/) |  |

### 

### Towards a large-scale federated knowledge graph for government information

Dutch government organizations have been actively involved in creating and publishing RDF knowledge graphs since 2012. Some initiatives have proven quite successful through the years. Currently, many RDF datasets are available, including RDF datasets containing information registered in so-called base registrations, which by law are assigned to provide an authoritative description of certain aspects of the world. Well-known cases are datasets from the bureau of land registry (Kadaster), but there are many more.

Most of these RDF knowledge graphs have been developed in a bottom-up fashion. Without much overall guidance, the datasets are predominantly expressed in terms of their own locally designed ontologies. Over the years, there have been initiatives to design shared ontologies as standards to be used in certain domains. When considering individual cases, results seem to be mixed. The overall trend, however, is that significant efforts are made to connect existing datasets, creating a large-scale federated knowledge graph for government information.

This presentation will discuss this general trend, focusing on a specific case. Currently, there is a dataset containing geometries describing administrative areas. Another dataset identifies government organizations, their properties and their history. The two datasets are maintained by different organizations, ultimately residing under the Ministry of the Interior and Kingdom Relations. The case illustrates what the impact of such an undertaking is on the locally designed ontologies. We discuss some recurrent modelling mistakes and how to deal with them. It is shown how a formal ontological based on Guizzardi’s work on UFO and OntoUML supports effective ontology alignment.

Disclaimer: the author is involved in the work reported in this presentation. This work is commissioned by KOOP, the publications office of the Dutch government. All content is presented in the author's personal capacity. KOOP does not endorse any parts of the presentation.

| Jan Voskuil CEO Taxonic & Ontologist | Taxonic  <https://www.taxonic.com/> |
| --- | --- |

### 

### Responsible AI and LLMs

After ChatGPT was made available to the public in November 2022, it became clear to even the biggest AI skeptics that a new era has now dawned and that AI will eventually enter many areas of life in the coming years.

Can AI become responsible?

ChatGPT attracted a lot of attention with its detailed answers covering a wide range of knowledge areas, but at the same time the call for responsible and explainable AI became louder again. Indeed, a notable drawback of generative AI, and large language models (LLMs) in particular, is their tendency to often generate superficial and inaccurate information that, moreover, does not provide any provenance information.

While it has become clear in recent months that generative AI and especially LLMs, on which ChatGPT is also based, are arguably fundamental building blocks of an enterprise-grade AI architecture, this needs to be complemented by other technologies and measures in order to speak of responsible AI. In particular, governance models and legal frameworks have yet to be put in place, as will be mandated, for example, by the [EU in its forthcoming AI Act](https://artificialintelligenceact.eu/), to provide sufficient assurances to all stakeholders (investors, companies, citizens and consumers, etc.).

In this talk, Andreas Blumauer (Semantic Web Company) discuss the merging of [LLMs and semantic technologies](https://www.poolparty.biz/blogposts/how-chat-gpt-works-non-explainable-ai), in particular how knowledge graphs can be used in combination with services like ChatGPT to develop applications that combine the best of both worlds to lead to responsible, explainable generative AI.

Questions identified as critical, particularly for regulated industries, will focus on:

* + What is the role of high-quality, well-structured training data?
  + How should the ‘human-in-the-loop (HITL)’ design principle be evaluated in the context of LLMs?
  + What AI applications for regulated industries, e.g. around ESG standards, are realistic in the immediate future?

Andreas have been working on the fusion of different AI, content and knowledge technologies for several decades, and at the end of this talk still dare to look into the future to discuss and contrast different scenarios with more or less usable, respectively responsible AI.

| Andreas Blumauer, CEO at Semantic Web Company | Semantic Web Company  <https://semantic-web.com/>  Andreas Blumauer is CEO and co-founder of Semantic Web Company (SWC), the provider and developer of the PoolParty Semantic Platform. With headquarters in Vienna, Austria, but operating globally, SWC has worked with over 200 commercial, government, and non-profit organizations to deliver AI and semantic search solutions, knowledge platforms, content hubs, and related data modeling and integration services. SWC was named to KMWorld’s prestigious list of “100 Companies that Matter in Knowledge Management” from 2016 to 2021 and has been named multiple times in Gartner’s Magic Quadrant for Metadata Management Solutions and as a Sample Vendor in their Hype Cycle for Natural Language Technologies.  In his role as CEO, Andreas is responsible for both the strategic growth of the company and its organizational evolution toward a highly focused customer orientation. SWC has grown every year since its inception under his leadership, and has been able to develop a cutting-edge and unique software platform that is ISO 27001 certified, and deployed globally across a number of key industries.  A high profile partner network has been built around this platform, implementing scalable semantic AI platforms for his clients, many of whom are among the largest and most innovative companies in their industries. |
| --- | --- |

### 

### 