# Data Integration and Analytics

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

#### Chair: TBA

## **Talks**

### Cultivating IoT Success - How Growing Bonsais Revealed the Future of Industrial IoT

Today, the Internet of Things (IoT) touches every corner of the globe. This demo illustrates how knowledge graphs and semantic reasoning are revolutionizing practices such as predictive maintenance and component management for companies worldwide—from the world’s largest buildings, logistics networks, offshore platforms and industrial plants, to one of nature’s smallest—the bonsai tree.

Bonsai trees are tough to look after with relatively complex and specific needs, so maintaining a garden of 20 (including different species, each with their own preferences no less) is no small feat. Critical to their survival is the question: ‘When should I water?’—asked by their owners daily. Too much water suffocates the tree and rots the roots, too little and it’s not long before years of effort are lost. Even harder is the question of what to do when you’re away from home.

The solution lies in soil sensors that monitor moisture levels for a selection of the trees; reporting their readings every minute. Along with weather data, these readings are processed on a central hub—a Raspberry Pi running the knowledge graph and semantic reasoning engine RDFox. Based on this data, the hub can issue alerts to the owner and control a series of pumps, each of which irrigates a different group of trees. The network of sensors, devices, and pumps, along with the trees themselves, are represented as a graph in RDFox, capturing functional properties and their relationships. By incorporating real-time sensor data and applying semantic reasoning, the system intelligently determines when individual trees need watering, activating the appropriate pump and giving the thirsty plants the drink they need.

Although clearly simplified in this scenario, the same principles can be scaled up to the vast and complex situations found in industry and nature. By using semantic reasoning in local hubs, IoT capabilities, industrial control and maintenance, and predictive analytics can be pushed out to the edge, bringing several benefits to real-world situations such as a complex of buildings, ships, or processing plants. By optimizing the device network, upfront capital can be reduced alongside ongoing operational costs. The same can be said for the system being monitored by improving the efficiency of the system itself and increasing the lifespan of its assets through reduced wear.

With this demo, we’ll show how this can all be achieved with a knowledge graph.

| Peter CrockerCEO and Co-Founder at Oxford Semantic Technologies[Affiliation page](https://www.oxfordsemantic.tech/) | Mikkel H. BrynildsenChief Data Scientist |
| --- | --- |

### Knowledge Models as Silver Bullet for Quality Intelligence - Pantopix [SP]

In many highly regulated industries, in particular the pharmaceutical industry, controlled vocabularies are state of the art. In addition, companies typically provide their own terminology. Internal as well as standard vocabularies may refer to different concepts and are subject to changes over time.

This situation very often makes it a herculean task to compare labels and/or concepts, update vocabularies, trace changes, and, after all, make use of controlled vocabularies for downstream processes such as information retrieval and information delivery. For instance, a content management system may contain tags from previous versions of the vocabulary, and users need to de-reference those labels or trace their life cycle.

In this talk, we illustrate how we solved these challenges for an industry partner in the pharmaceutical industry with its high standards for data quality and system/process validation. We show how we combined public and custom-made ontologies, integrated vocabularies from various sources and integrated mechanisms for versioning based on PROV-O in our knowledge model. In addition, we demonstrate how a querying module is integrated into the customer’s workflows, providing the basis for future applications such as autoclassification, information retrieval and intelligent information delivery.

| Martin LeyPantopix [Affiliation page](https://www.pantopix.com/) | Johann WagnerPantopix [Affiliation page](https://www.pantopix.com/) |
| --- | --- |

### Leveraging Semantic Technologies for Enhanced Enterprise Knowledge & Governance: A Real-World Use Case )

Initial Situation:

Large organizations often face a critical challenge: connecting their enterprise architecture, data catalog, data lineage, and compliance systems. These systems, crucial for smooth business operations, are often disjoint, operating in isolation or 'silos'. This separation hampers not just data accessibility, but also broader aspects such as compliance tracking and architectural coherence. Our financial services client confronted this precise issue, highlighting the urgent need for a solution that could seamlessly integrate these key areas.

Approach and IT-Solution:

Our solution's distinctive edge lies in its capacity to unify enterprise architecture, data cataloging, and compliance systems using the RDF graph model. By transforming diverse data into RDF triples, and employing ontologies, rules, and reasoning, we transform disjointed data into an interconnected information landscape.

In addition to data integration, we have developed a user-friendly frontend that offers faceted search and interactive exploration of the graph model. This interface enhances user engagement with the system, encouraging data discovery and insights generation.

Our deployment process, underpinned by CI/CD principles, handles various data formats, enhancing data pipeline management. Our frontend will be released as open-source software, enabling other organizations to benefit from our approach.

Success Criteria for / Benefit of the Semantic Solution:

Our solution's success is best demonstrated by the substantial benefits experienced by our financial industry customer. By integrating metadata from diverse sources, the system allowed them to feed metadata daily from over 400 databases containing more than 42,000 tables into a comprehensive knowledge graph. Additionally, we were able to unify disparate but crucial organizational components like the agile service delivery organization (including arena, feature team, release train), employee data, domain architecture information, and business processes.

This approach has enhanced visibility and accessibility of information for the 60+ recurring users of our system. With over 17+ million triples contained, users can now traverse data from all sources in milliseconds, illustrating the system's enhanced performance and scalability.

Prospects and Recommendations:

Semantic technologies offer great potential to unify enterprise architecture, data catalogs, and compliance systems. For businesses planning to use such technologies, we recommend starting with the most problematic data silos, gradually expanding the integration.

We've seen that complex organizations can reduce complexity by adopting a data-centric approach with ontologies as a schema layer. Our experience suggests that even complex enterprises might need only a few hundred concepts as their core model, leading to a complexity reduction of 10-100 times.

As future steps, we aim to handle more data and connect more organization components, making semantic technologies benefits more widespread.

Demo:

In our presentation, we will provide a short demonstration of our system, showcasing the real-world application of semantic technologies and the transformative effect they have on an organization's enterprise knowledge and governance.

| Adrian Gschwend | Michael Rauch |
| --- | --- |