# Linked Data II

##### Session 2.1 (SEMANTiCS)

#### Time: Wednesday, September 18, 2024 - 13:00 to 14:30

#### Chair: Marvin Hofer

## **Talks**

### UniPart: Optimizing Streaming Graph Partitioning towards Universal Adaption in RDF Triple Stores

With increasing size of Resource Description Framework (RDF) graphs, the resulting graph structures can become too large to be managed on a single compute node, lacking the necessary resources to execute a partitioning of the graph -- in particular, when the partitioning method relies on global graph information for which the entire graph has to be loaded into the main memory. This paper introduces a window-based streaming partitioning technique to obtain distributed RDF graphs, overcoming the memory limitations of traditional partitioning methods. We evaluated our approach, UniPart, by comparing it with established graph partitioning algorithms such as METIS, LDG, and WStream. The comparison focused on key metrics, including the proportion of edge cuts.

|  |  |  |
| --- | --- | --- |
| Wenhui Yang | Ahmed Al-Ghezi | Lena Wiese |

### TriplyDB: Unlock your data’s potential [SP]

Linked data has the potential to transform and replace traditionally proprietary database paradigms. Unfortunately, linked data infrastructure is often too immature to run comfortably in production systems, and first-time users get stuck in setting up disconnected tools in the command-line.

Triply wants to make linked data accessible to everyone. With a user experience that welcomes first-time users, and an infrastructure that can be deployed in large companies.

Operating from the vibrant campus of the VU University Amsterdam, Triply builds a freemium triple store that runs completely online, scales horizontally, and is certified to run in secure and mission-critical environments.

|  |
| --- |
| Wouter BeekTriply |

### Querying and Ingesting Linked Open Data about the Decision-making Process of the Flemish Government through Themis

In recent years, Flanders has been investing heavily in the openness of government.   
Until 2019, decisions of the Flemish government were published in purely human-interpretable form, on a Drupal website coupled to their document management system. However, the lack of semantics made it challenging for applications to consume this data.  
  
For this purpose, redpencil has developed the Themis open data portal (<https://themis.vlaanderen.be/>), where information regarding the decision-making process of the Flemish Government is disseminated in the form of Linked Open Data. More specifically, every non-confidential decision that is made during the weekly council of ministers in Flanders is automatically exported from the government's platform and published to Themis in the form of RDF. To maximize interoperability, the data is structured according to the OSLO data model, an ontology created specifically for this purpose with the cooperation of all parties involved with the decision-making process.  
  
While the intention of Themis is to stimulate reuse in semantic applications, the data is accessible for a wide range of consumers. The main consumer of this data today is the governments own press release website which provides human-readable information about the decisions made by the government.

There are two main ways to access the data (other than through the browser). On the one hand, it can be ingested through DCAT data catalogs, which is useful for applications that want all the data, all at once, for example to analyze it in bulk. On the other hand, the data can be queried using SPARQL, which is useful for applications that are interested in specific data, about specific decisions or topics. I will briefly introduce the technicalities behind both methods, and provide real-world examples of where this approach is used. To conclude, I will showcase a number of useful scenarios and third-party integrations that exist today.

|  |  |
| --- | --- |
| Tom De Nies | Alvin Demeyer |

### 

### Enabling Delayed-Answer Auctions for RDF Knowledge Graphs Monetisation

Traditionally, querying knowledge graphs is free of charge, however, ensuring availability of data and service incurs costs to knowledge graphs providers. The Delayed-Answer Auction (DAA) model has been proposed to fund the maintenance of knowledge graphs endpoints, by allowing customers to sponsor entities in the Knowledge Graph so query results that include them are delivered in priority. However, implementing DAA with a time to first results acceptable for data consumers is challenging because it requires reordering results according to bid values. In this paper, we present an approach to enable DAA with low impact on query execution performance. Our approach relies on (i) reindex sponsored entities by bid value to ensure they are processed first (ii) Web preemption to ensure delayed answering. Experimental results demonstrate that our approach significantly outperforms a baseline execution in terms of time to deliver the first results.

|  |  |
| --- | --- |
| Hala Skaf-Molli | Pascal Molli |
| Luis-Daniel Ibanez | Abraham Bernstein |

### 