# Session 2.4

##### Session 2.4 (SEMANTiCS)

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

#### Chair: TBA

## **Talks**

### Investigate the Impact of Contextual Information on LLMs for Taxonomy Expansion

This paper presents an exploratory study that investigates the use of various Large Language Models (LLMs) for the task of taxonomy expansion.

Our objective is to enhance the taxonomical structure by querying LLMs for (1) child taxons and (2) alternative labels of existing taxons. Beginning with an incomplete taxonomy, we explore the most effective ways to prompt LLMs exploiting explicit and shared knowledge captured in manually curated taxonomies to provide context for the task at hand. We experiment with different prompting templates, well-recognized taxonomies (EuroVoc, STW, UNESCO), and popular language models (Claude, Claude3, Llama2).

Our results suggest feasibility of solving of the proposed task with the modern LLMs and human oversight. Moreover, we observe certain patterns and trends in the performance of the models, noting that it was not possible to identify a single best configuration that would fit all models.

| Artem Revenko | Anna Breit | Salma Mahmoud |
| --- | --- | --- |
| Mark Szabo | Tomas Knap |  |

### Oxford Semantic Technologies [SP]

| TBAOxford Semantic Technologies |
| --- |

### Towards Digital Sustainability Reporting: An Ontology for Mapping of Indicators in GRI and ESRS

Sustainability reporting by Small and Medium Enterprises (SMEs) is gaining importance. SMEs form the backbone of European industries, and their customers rely on them to ensure regulatory compliance. In preparing sustainability reports, a combination of standards is commonly used, which encompasses overlapping yet distinct requirements on sustainability indicators. Different standards categorize shared indicators under varying topics, while they also mandate unique indicators to assess identical sustainability phenomena. This poses challenges for SMEs in reporting against multiple standards. Considerable human efforts are demanded to determine the interconnected requirements across different standards. Additionally, reporting on overlapping indicators for new standards results in significant redundant work. Mapping of indicators between different standards allows the semantic interoperability of standards by indicating matching and distinct requirements, aiding in addressing these challenges. Therefore, this paper focuses on developing an ontology for mapping indicators from two significant standards, GRI and ESRS. We introduce the Sustainability Reporting Standards Ontology (RSO). RSO formally represents environmental indicators in GRI and ESRS, emphasizing indicator requirements such as unit, quantity, and measurement variables. RSO is implemented in the RDF/OWL format and will be made available online. Furthermore, we provide an ontology-based mapping between indicators, supported by concrete examples that illustrate the interconnections between them.

| Yuchen Zhou | Yuan Cao |
| --- | --- |
| Alexander Perzylo |  |

### Exploring Prompt Generation Utilizing Graph Search Algorithms for Ontology Matching

The interoperability of domain ontologies, developed by domain experts, necessitates their alignment before attempting to match them. Within these ontologies, defined concepts often encounter an ambiguity problem stemming from the use of natural language. This interoperability issue raises the underlying ontology matching (OM) challenge. OM might be defined as the identification of correspondences or relationships between two or more entities, such as classes or properties among two or more ontologies. Rule-based ontology matching approaches, e.g., LogMap and AML have not outperformed machine learning based matchers on the Ontology Alignment Evaluation Initiative (OAEI) benchmark datasets, especially on the OAEI Conference Track since 2020. Supervised machine or deep learning approaches produce the best results but require labeled training datasets. In the era of Large Language Models (LLMs), robust zero-shot prompting of LLMs can also return convincing responses. While prompt generation requires prompt template engineering by domain experts, contextual information about the concepts to be aligned can be retrieved by leveraging graph search algorithms. In this work, we explore how graph search algorithms, namely (i) random walk and (ii) tree traversal can be utilized to retrieve the contextual information to be incorporated into prompt templates. Through these algorithms, our approach refrains from considering all triples connected with a concept to be aligned in its contextual information creation. Our experiments show that including the retrieved contextual information in prompt templates improves the matcher's performance. Additionally, our approach outperforms previous works leveraging zero-shot prompting.

| Julian Sampels | Sefika Efeoglu |
| --- | --- |
| Sonja Schimmler |  |