NLP for Ontology Development: A use case in spacecraft parts domain

Abstract:

Ontologies can help in the digitalization of engineering design processes. We focus on the design and development of spacecraft, in particular on archiving and discovery of spacecraft part specifications. We created an ontology that is used as a skeleton for creating database schemas and data exchange APIs as well as populating a knowledge graph. However, the development of an ontology from scratch is a rather laborious task. We have to consider various sources of information, such as existing space-system standards, product specifications provided by manufacturers, and actual consumers, i.e. engineers involved in the design process - each with their own perspective. Furthermore, these documents and demands are changing more and more frequently due to new disruptive technologies and an increasing competition in the space business. We aim to close the gap between the tedious process of manual ontology development and fast-paced spacecraft parts development. Due to the lack of machine readable descriptions in this area, we apply NLP techniques to extract technical information from these documents. This information can be used in two ways: first, to create and improve our spacecraft parts ontology and second, to populate a knowledge graph.

Description of talk:

Ontologies have recently gained more attention in industry and the engineering domain. However, the number of actual implementations is rather stagnant on a comparatively low level. One major obstacle is the effort needed to create a domain specific ontology. So far, this task is done manually by ontology engineers and reviewed by domain experts, such as system engineers, who also have to revise the created ontology regularly to keep up with emerging new technologies. However, this requires a tremendous effort and is highly biased by the experiences and opinions of the involved personal.

In this talk, we outline our motivation and plans to use ontologies to assist the spacecraft design process, e.g., by providing a basis for cross-platform data exchange. Then, we dive into the challenges we are facing: the creation of an ontology and keeping it up to date. Considering the various sources of information needed for developing a spacecraft parts ontology, we utilize NLP techniques to explore and extract information from data sheets provided by different manufacturers. We present solutions to extract information from unstructured data, e.g. technical documents, as well as how to use the extracted information for creating an ontology. To keep the created ontology up-to-date, the same solutions, combined with the information retrieved from other general purpose knowledge bases, can be used to enrich the ontology over time. The ontology can also be used to improve the information extraction process itself. This repeated feedback loop allows for continuous improvement of both the ontology itself and the information extraction process in general. Nevertheless, to ensure the accuracy of extracted information, we include a human-in-the-loop step in our solutions. Finally, we will demonstrate the initial results of our application and show our spacecraft parts ontology developed using a semi-automatic data extraction and discuss areas for further improvement.

Our solution for ontology development presented in this talk is not limited to

the domain of space systems, but can be applied to other engineering domains, where valuable knowledge is buried in unstructured text. We suggest the potential usage, since the extracted information is not only empowering ontologies, but it can also populate databases or knowledge graphs.