Topic Proposal for a Master's thesis at the Information Systems Group

Knowledge Graph for Works of Art

Knowledge Base Construction with Deep Learning Techniques

This topic proposal for a Master's thesis focuses on knowledge discovery and semantic text.

Problem

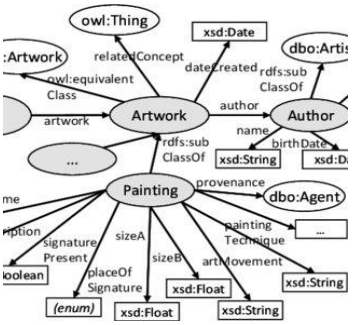
- Textual data on cultural heritage data such as art catalogs, paintings etc. in digital format.
- It is challenging to identify and connect the important entities to each other.
- Lack of domain specific information makes this problem harder.
- Efficient means to represent this is needed in order to mine further knowledge.

Approach

- Design of a schema for the art domain based on important entities.
- Machine learning and neural networks based techniques for natural language processing tasks such as named entity recognition and relation extraction.
- Knowledge graph construction and refinement for structured representation will be explored.

About the Project

It is non-trivial to understand the relationships among different entities in noisy data. A knowledge graph is a standard choice for structural representation of important concepts that is capable of showing the relationships and attributes of these entities. The construction of knowledge graphs is an elaborate process, involving many linguistic and data oriented tasks. The lack of standard benchmarks and annotated data for the art domain makes these tasks even more complex. We aim to leverage the classical statistical approaches as well the recent neural networks based techniques for these tasks and attempt to adapt them for the art domain.



This Master's thesis will focus on large scale text analytics and natural language processing tasks such as named entity recognition, relationship mining and word sense disambiguation. State-of the-art neural network approaches will be explored for these tasks. Familiarity with deep learning architecture and data intensive tasks would be preferred.

Contact

If you are interested in this topic, I would be happy to discuss further. Feel free to stop by at my office F-2.07 or send me an e-mail: Nitisha.Jain@hpi.de.

Design IT. Create Knowledge.

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam