AI is increasingly used in life-changing applications, e.g.
- Automatic Grading
- Job applications
- Business decision making

However:
- Most AI models are black boxes
  - We only see the output but not the decision process

This Project: Learn how to extract a descriptive graph of entities and their relations for a document and use it for explainable AI decisions (e.g. via a graph neural network)

Prerequisites
- Enrollment in Data Engineering or ITSE
- Basic knowledge of deep learning and natural language processing
- Software development skills
  - Experience with Python is essential
  - One person with front-end development skills would be helpful
- Interest in academic research
  - We intend to write a research paper about this work

Useful Resources
- Background on Natural Language Processing
  - Stanford NLP Course
  - HuggingFace Transformers Library
  - Speech & Language Processing book by Jurafsky & Martin
- Libraries and Datasets:
  - Project Gutenberg
  - BookSum Dataset
- Google Colab for computing (we will later also use HPI GPU servers)
  - https://colab.research.google.com/