Back-end development with web crawling / scraping for real estate valuation

Background
Real estate valuation has so far been carried out using cumbersome and lengthy processes. The appraisers spend more than 50% of their time gathering information about the market and the location of the property, formatting tables and diagrams and writing a 100-page report with the same content. The digitization of these processes leads to enormous time savings. Therefore, the aim of this IDP is to develop an application that automates the gathering of information about the market and the location of the property.

Task
In this project you will work on the following tasks:

- Identification and selection of suitable tools, programming languages and frameworks (e.g. tools from AWS; Python, Java; etc.)
- Design and development of data pipelines from web crawling / scraping to storage.
  - Obtain data from structured and unstructured data sources.
  - Extraction of data (text, numbers, images) from several webpages and documents in different formats.
  - Processing and saving the data in a suitable database.
- Cooperation and close coordination with another IDP, which is based on the data generated.
Requirements

- Understanding of back-end developments.
- Experience with web crawling / scraping and databases.
- Systematic and independent way of working and the ability to work in a team.

Contact

Please do not hesitate to contact me if you have any questions regarding this project:

Dr.-Ing. Christina Mauer
Tel.: +49 89 289-22483
E-Mail: c.mauer@tum.de

M.Sc. Maximilian Reiser
Tel.:+49 89 289-22483
E-Mail: m.reiser@tum.de