Digitization of local land use plans (Bebauungspläne): A RAG-approach

Laia Domenech Burin | Ludwig Maximilian University of Munich | Germany Hide details

Authors:

- Laia Domenech Burin | Ludwig Maximilian University of Munich | Germany
- Dr. Phil. Felicitas Sommer | Technische Universität München | Germany
- Dr. Deepika Mann | Bundeswehr University Munich
- Sebastian Riznik | Germany
- Hope Dellaeyram Ewudor | Ghana

Local land use planning provides insights into local social, economic, and environmental contexts, making it a valuable data source for experimental regional statistics. Yet, the accessibility and analysis of relevant data are hindered by its dispersion across various online platforms in non-standardized, unstructured PDF formats. We present an open-source data extraction pipeline designed to systematically extract and analyze data from German municipal land use plans (Bebauungspläne). The pipeline identifies land utilization ordinances (BauNVO) in the building plans, flooding protection measures, and measures of maximal floor coverage height in the plans. Our methodology, applied to the regions of North Rhine-Westphalia and Bavaria, involves a four-step process: (1) scraping links to building plans from regional geoservice Web portals, (2) bulk downloading and converting PDF documents into raw text, (3) employing keyword searches to identify key data points, and (4) using a language model to extract numerical values related to building area, height, and presence of built-up surface. The extracted data is stored in a structured database that can be used for efficient data retrieval and advanced statistical analysis. By developing an open-source data extraction pipeline, we aim to bridge the existing gaps in data accessibility and usability.