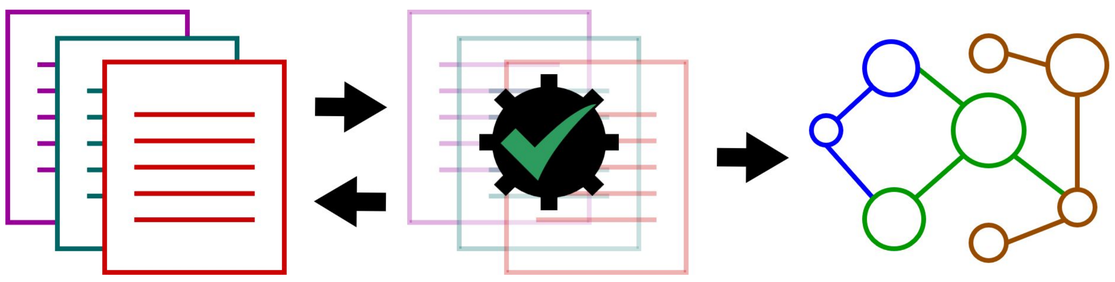
Spatio-Temporal Knowledge Graph Construction for Ecological Systems from Heterogeneous Data Using Generative AI

Philipp Plamper, Anika Groß

Context

- Global change strongly impacts biodiversity and ecosystem functioning in agricultural landscapes and therefore requires the development of measures to restore ecosystems and landscapes.
- **AgriRestore**[1] is an interdisciplinary research project addressing key knowledge gaps through real-world observations, long-term experimental monitoring and knowledge synthesis.
- Given diverse subprojects involving multiple study sites and experiments, a **Knowledge Graph** could serve as a central hub for literature research and knowledge synthesis.

Conceptual Approach



Technical Resources

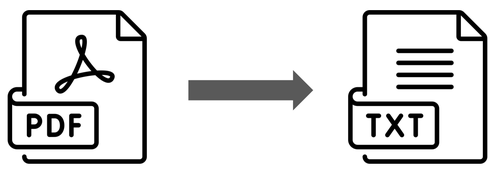
- RAM: 16 GB
- GPU: NVIDIA RTX A6000
- CPU: Intel® Core™ i7-1185G7
- LLM: qwen2.5:32b-instruct-q4_0[2]

PyMuPDF

Rapid Fuzz

Workflow

1. Transform Text

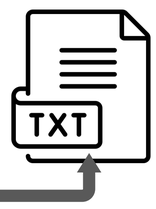


2. Extract Data using LLM

Prompt
- extract information for each ‘Label’
  - [PAPER]
  - [STUDY LOCATION]
  - [YEAR OF SAMPLING]
- indicate source of extracted information
- save extracted information as triples
  - **[LABEL] | [DATA] | [SOURCE]**

3. Evaluate Extracted Data

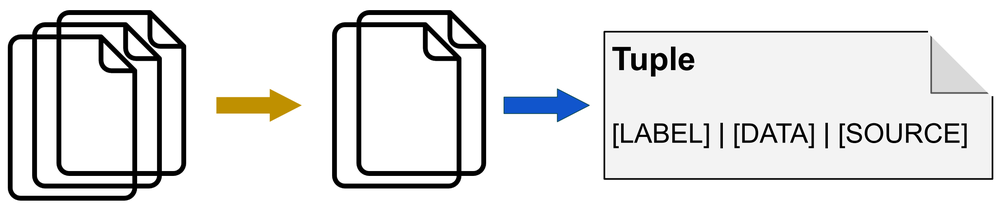
Results
...
[LABEL] | **[DATA]** | **[SOURCE]**
...



4. Create Knowledge Graph

Results

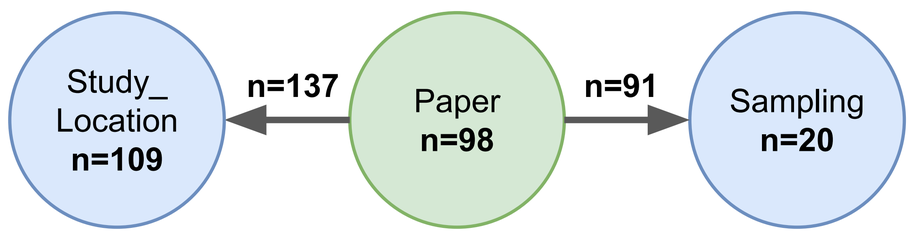
Processing



- exc. context length: -17
- exc. tuple length: -6
- wrong instruction: -2

- insufficient similarity: -161
- year mismatch: -21
- location mismatch: -20

Knowledge Graph



Improvements

Processing	Knowledge Graph
• computing power • llm model selection • llm parameter tuning • prompt engineering • evaluation process	• graph modeling[3] • spatial/temporal integration • expandability with additional data, e.g. authors, objectives, methods • graph analysis

[1] AgriRestore Website, https://agrirestore.de/, 02.03.2026
[2] Ollama qwen2.5:32b-instruct-q4_0, https://ollama.com/library/qwen2.5:32b-instruct-q4_0, 02.03.2026
[3] Plamper, P., Köpcke, H., & Groß, A. (2025). A Survey on Spatio-Temporal Knowledge Graph Models. arXiv preprint arXiv:2512.16487

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 528485254 - FIP 16