

Data, Dialogue, and Discovery

A Controlled Workflow for AI-Supported Historical Research and Analysis with nodegoat

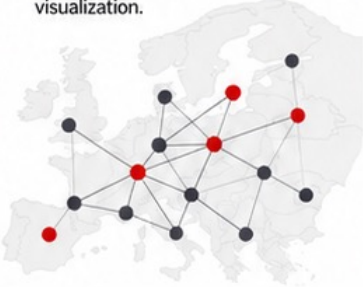
Kaspar Gubler

Historical Institute, University of Bern



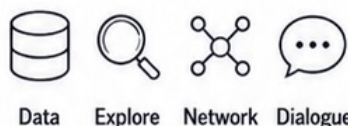
1. Context

- Developed in the context of the Repertorium Academicum (REPAC) project at the Historical Institute, University of Bern.
- REPAC contains large biographical datasets on European academics between 1250 and 1550.
- The data include temporal and spatial attributes that allow the reconstruction of academic networks, mobility patterns, and knowledge spaces.
- **nodegoat** functions as the research environment for data management, exploration, and visualization.



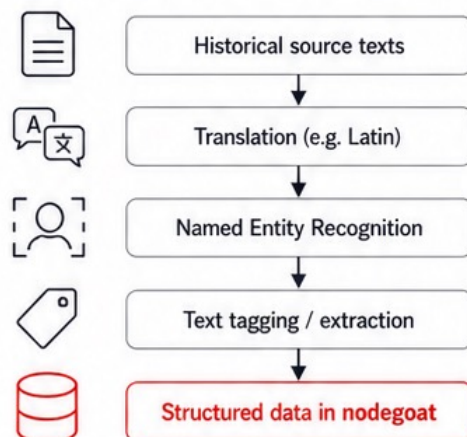
2. Why nodegoat?

- Structured research data in a single environment
- Integration of AI-supported workflows
- Transparent links between sources, data, and analysis
- Exploratory visualizations and contextual querying



3. Workflow 1: Information retrieval and data enrichment

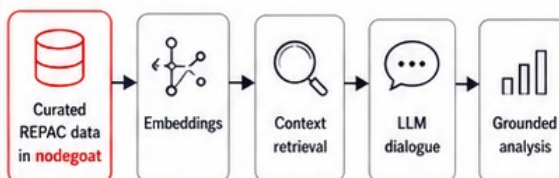
AI-supported modules extend **nodegoat**-based workflows for source processing and data entry.



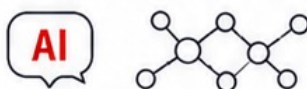
- Supports multilingual information detection
- Connects linguistic processing with historical data management
- Enables faster yet controlled enrichment of research data

4. Workflow 2: Retrieval-augmented dialogue with research data

Curated structured data from **nodegoat** are combined with large language models in a transparent RAG pipeline.



- Relevant information is retrieved before model interaction
- Project-specific context grounds interpretation
- Reduces arbitrariness and helps limit hallucinations
- Strengthens scholarly control over the analytical process



5. Research value

- Links data management, analysis, and dialogue
- Makes AI use more transparent and verifiable
- Supports semantic exploration of historical research data
- Opens new possibilities for discovery in digital history

“ Retrieve first, interpret second. ”

6. Critical reflection



- AI-assisted workflows can accelerate research and analysis.
- However, students may reach results faster while deepening their subject knowledge less thoroughly.
- Greater speed may also reduce serendipitous discoveries in the research process.
- Controlled and reflective use of AI remains essential in historical scholarship.

nodegoat-based AI workflows show how research data, contextual retrieval, and dialogic analysis can be combined without giving up scholarly transparency.

