Continuous Assessment 1

Knowledge Graphs

Prof. Dr. Kai Eckert, Tobias Malmsheimer

April 16, 2026

1 Task

Use knowledge graph techniques and data integration methods learned in the lecture to construct a knowledge graph from multiple heterogeneous data sources. Document your approach, design decisions, and results.

1.1 Topic

The topic is for you to choose. Use the lecture to discuss ideas. Think about the domain you want to model, which data sources you will integrate, and what kind of queries or insights your knowledge graph should enable.

1.2 Project submission (graded)

The submission consists of:

- a written documentation of your approach and the results
- code or scripts used for data processing and graph construction
- the resulting knowledge graph (e.g. as RDF/Turtle, JSON-LD, or similar)

- data (if the data is more than 10MB), provide stripped down example data that can be used to demonstrate the approach

Submission via Moodle in a ZIP File. groupname.zip:

- /groupname
  - groupname-documentation.pdf ← Your written documentation
  - /code ← Scripts, pipelines, and data
  - /code/readme.txt ← Anything I need to know to run the code (e.g. dependencies, how to load the graph)

1.2.1 Documentation

The documentation focuses on the design and results of your knowledge graph. Minimal contents are:

- Research question or concise description of the knowledge graph's purpose
- Description and justification of the data sources (at least 4), including provenance, licensing, and structure
- Description of the selection process: why these sources, how were they assessed for quality and relevance?
- Description of the integration approach: schema alignment, entity resolution, linking strategy
- Description of the pipeline / implementation
- Evaluation of the resulting knowledge graph: coverage, consistency, quality, and example queries or use cases demonstrating its value
- Possible improvements or extensions

The documentation is limited to 6 pages (Font size 11, without Figures, TOC, . . .).

1.2.2 Code and Data

- Everything needed to reproduce and load your knowledge graph.
- Exact information about required dependencies and how to run your pipeline (readme.txt).
- The resulting knowledge graph file(s) and example/full data as appropriate.
- Links to all original data sources.
- Make sure that you keep a backup of the full data in case of questions and to reproduce the results.

1.3 Project presentation (graded)

- Present your results in class (7-10 min, strictly enforced).
- Focus:
  - Brief summary of your idea and domain
  - Which data sources did you use (at least 4) and how did you obtain them?
  - Your integration approach: how did you align and link the sources? What challenges arose. your first presentation?
  - Results: structure and size of the graph, example queries, evaluation
  - Structure and potential use cases of your knowledge graph, example queries, avaluation.