



August 26, 2025 – CoRDI 2025, Knowledge Graphs session

Indexing Historical Research Data: MemO and the NFDI4Memory Knowledge Graph

**Sarah Rebecca Ondraszek, Tabea Tietz, Etienne Posthumus,
Heike Fliegl, Jörg Waitelonis, and Harald Sack**



NFDI4Memory and associated challenges

- Part of the National Research Data Infrastructure (NFDI) with a focus on the **historically oriented sciences**.
- Highly **diverse disciplines, data formats, vocabularies** (OAI-PMH, RDF/XML, ...).
- **Fragmented**, heterogeneous historical research data (EHRI, DDB, FactGrid, ...).
- Multiple, **unconnected access points** (also via **data aggregators**).
- **Limited resources** – workforce, time, infrastructures, etc.



Indexing, federating, unifying

- The goal is to **connect** these **distributed, heterogeneous** data and enable FAIR data sharing.
- Single **point of entry** for research data and services in the **NFDI4Memory Data Space**.
- MemO knowledge graph (KG) as **central index** connecting **decentralized sources without duplicating data**.
- Semantic linking via shared concepts (people, places, events).
- A corresponding **formalization** through the **MemO** (ontology), based on NFDIcore.

NFDIcore & NFDI4Culture ontology as a baseline

- **NFDIcore: mid-level ontology** (aligned with BFO), models people, organizations, projects, datasets, services.
- **NFDI4Culture (CTO):** domain extension for **cultural heritage**, covering artefacts, events, media, etc.
- As a modular extension, **MemO** adds concepts relevant for the **indexing of historical science data**.



Further related works

- **ArCo**: Italian cultural heritage KG for object-level integration (art works, literary works, ...).
- **Europeana**: multilingual cultural heritage metadata aggregation via linked entities.
- **Sampo**: the FindSampo ontology for indexing all Sampos across Finland.
- MemO builds on these principles with a **focus on federation across data** from different repos in the historical sciences and in relation to **other NFDI endeavors**.



MemO and the MemO knowledge graph

- Two subgraphs:
 - **Research Information Graph (RIG)** → researchers, institutions, services.
 - **Research Data Graph (RDG)** → metadata on historical sources & provenance.
- Competency-question-driven design (**discoverability, provenance, contextual links**), e.g.:
 - Which data contains information on specific persons, institutions, events, etc.?
 - Who created or published the document or data?



MemO concepts and modules

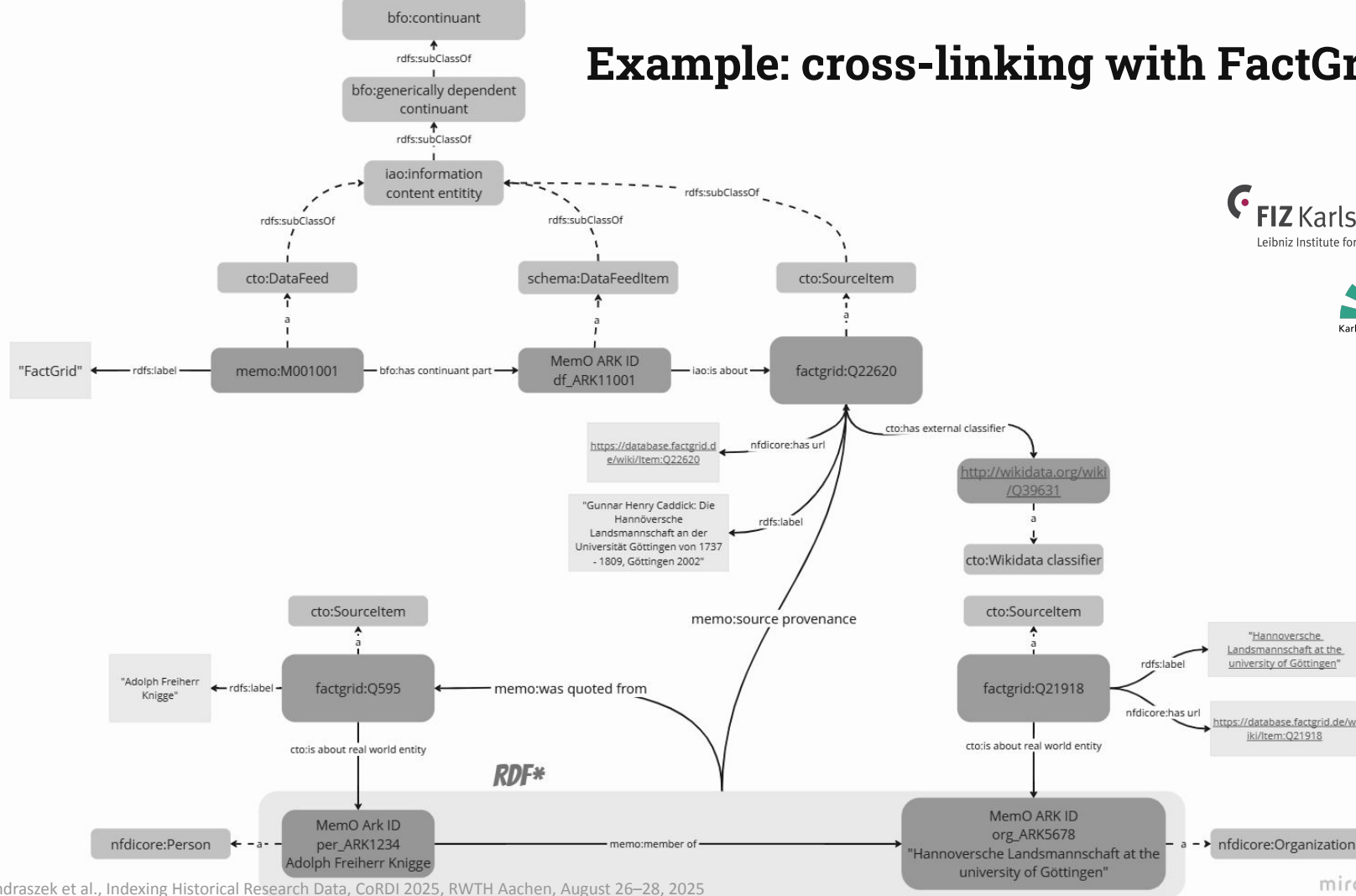
- MemO introduces modular extensions for **provenance** and extended **document representation**.
- Hierarchical relations (e.g., in the context of archival items and collections), as well as ...
- ... subject categories.



Example: cross-linking with FactGrid

- Ingestion of information about documents, organizations, persons, events.
- Central aspects: Provenance information (for claims, for provided information, etc.).
- Challenges: Entity vs. metadata about an entity (when it comes to persons, for example).

Example: cross-linking with FactGrid



Example: cross-linking with FactGrid

1

Researcher with a specific research question or request.

Which documents mention Wilhelm Bernhard Trommsdorff?



2

Researcher can look for the name, the occupation, the time-span, or further details with a SPARQL query or using a facettted (keyword) search in the Data Space.



'physician', 'Erfurt',
'Wilhelm Bernhard
Trommsdorff'

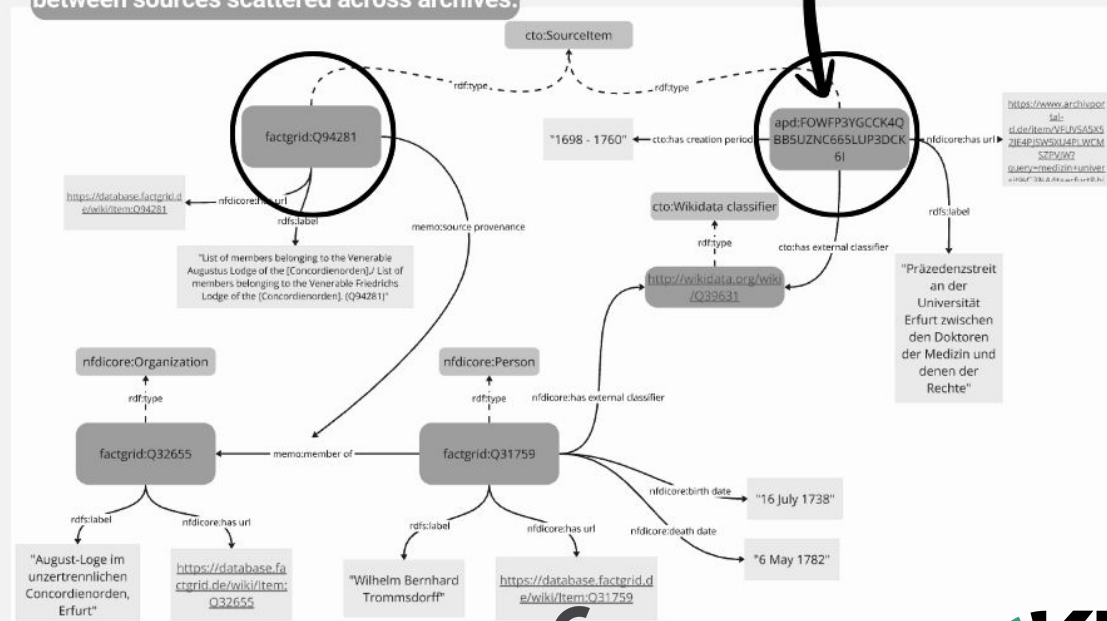
```
SELECT DISTINCT document.DocumentLabel 'sourceEntity','sourceLabel'
WHERE {
  document a cto:SourceItem
}
UNION
{
  document memo:source _provenance ?sourceEntity .
  FILTER (?sourceEntity = fact:grnd.Q31799)
}
UNION
{
  document cto:hasExternalClassifier wd:Q39631 .
  BIND(wd:Q39631 AS ?sourceEntity)
}
OPTIONAL { document rdf:label 'DocumentLabel' . }
OPTIONAL { document rdfs:label 'sourceLabel' . }
```

3

Using external, shared classifiers and standardized, interconnective relations, MemO enables uncovering relationships between sources scattered across archives.

4.

Matches via the MemO KG in FactGrid and Archivportal-D.





THANK YOU!

Sarah Rebecca Ondraszek et al.
FIZ Karlsruhe – Leibniz Institute for Information
Infrastructure
Karlsruhe Institute for Technology (KIT)

Private Mastodon: @[sourisnumerique](#)

FIZ ISE Mastodon: @[fizise](#)

Mail: sarah-rebecca.ondraszek@fiz-karlsruhe.de