

Linked Open Data in the Humanities: The LODinG Project

Background

The Linked Open Data (LOD) paradigm enables the linking of multidisciplinary research outcomes using a coherent ontological classification. LOD employs the interoperable formats and allows heterogeneous data to be formally modeled and linked into a common body of knowledge.

The LODinG project aims to explore the potential of the LOD at the intersection of qualitative and quantitative studies in the humanities and promote the transparent, interdisciplinary reasoning supported by state-of-the-art data management infrastructure and knowledge networks.

Tools & technologies: RDF, OWL, CIDOC-CRM, SPARQL

Keywords: Linked Open Data, Digital Humanities, Knowledge Graphs, Network Analyses, Multilingualism

Central Research Focus

Effective methods of collecting, modeling, linking, releasing and analyzing machine-readable information relevant to digital humanities in the form of Linked Open Data.

Research Area 2

(Digital Editing, German Studies & Romance Studies)

RA 2 applies the translational methods to compare standardized botanical and medical terms in a multilingual, historical perspective.

Methods: OCR, LLOD, translation-equivalent analyses

Materials: *Nomina propria* from early modern period (multilingual)

Research Area 4 (Sinology & Computer Science)

The fourth RA focuses on Modern Standard Chinese scientific literature and converts the findings into synthetic LOD statements.

Methods: Entity extraction, LOD statement synthesis

Materials: Selected corpus of scientific texts

Research Area 6 (Jurisprudence & Digital Humanities)

This RA analyzes legal multilingual sources and aims to identify differences in legally-binding terminology among all official languages of the EU using linguistic LOD.

Methods: Conceptual indexing, multilingual corpus annotation

Materials: Parallel corpora of legal texts

Infrastructural Support

(Trier Center for Digital Humanities & Trier University Library)

This work area focuses on developing technical solutions to achieve the project's goals. It aims to create the interfaces for unstructured data, enabling non-standard formats to adapt to the LOD framework.

Research Area 1 (Digital Lexicography & German Studies)

This RA emphasizes the importance of the representation of dictionary entries in the LOD framework and provides support for the integration of lexicons into the Semantic Web.

Methods: Semantics by reference, Semantic Web

Materials: Pandemic-related neologisms (diachronic perspective)

Research Area 3 (Digital Humanities & Computer Science)

This RA focuses on content retrieval. It applies the methods of information extraction to transform, e.g., abstracts and keywords into machine-readable LOD statements.

Methods: OpenAlex (in part), Semantic publishing

Materials: Selection of works across the disciplines

Research Area 5 (Cultural Studies & Computational Linguistics)

This RA employs multimodal data. It combines text and image analyses and indexing to build more robust knowledge networks.

Methods: LLMs, text and image integration, knowledge networks

Materials: Wine labels, postcards, maps, book illustrations

Research Area 7 (Integration, Federation & LOD resources)

This cross-sectional RA aims to integrate the standardized entries into a modular ontology that supports federated queries.

Methods: Enhancement of KGs, Wikidata identifiers

Materials: Domain- and discipline-specific entities

Anticipated Results

- Developing a modular cross-domain data model for the humanities using the Linked Open Data paradigm
- Publishing best practices guidelines for employing the Linked Open Data in interdisciplinary research


Funding

Forschungsinitiative des Landes Rheinland-Pfalz



Rheinland-Pfalz

MINISTERIUM FÜR
WISSENSCHAFT
UND GESUNDHEIT

Project website

<https://tcdh.uni-trier.de/en/projekt/loding>

