

DATA MODELING AS A HIGH-WIRE ACT

Balancing Requirements, Juggling Vocabularies, and not Falling (Short of Established Best Practice)

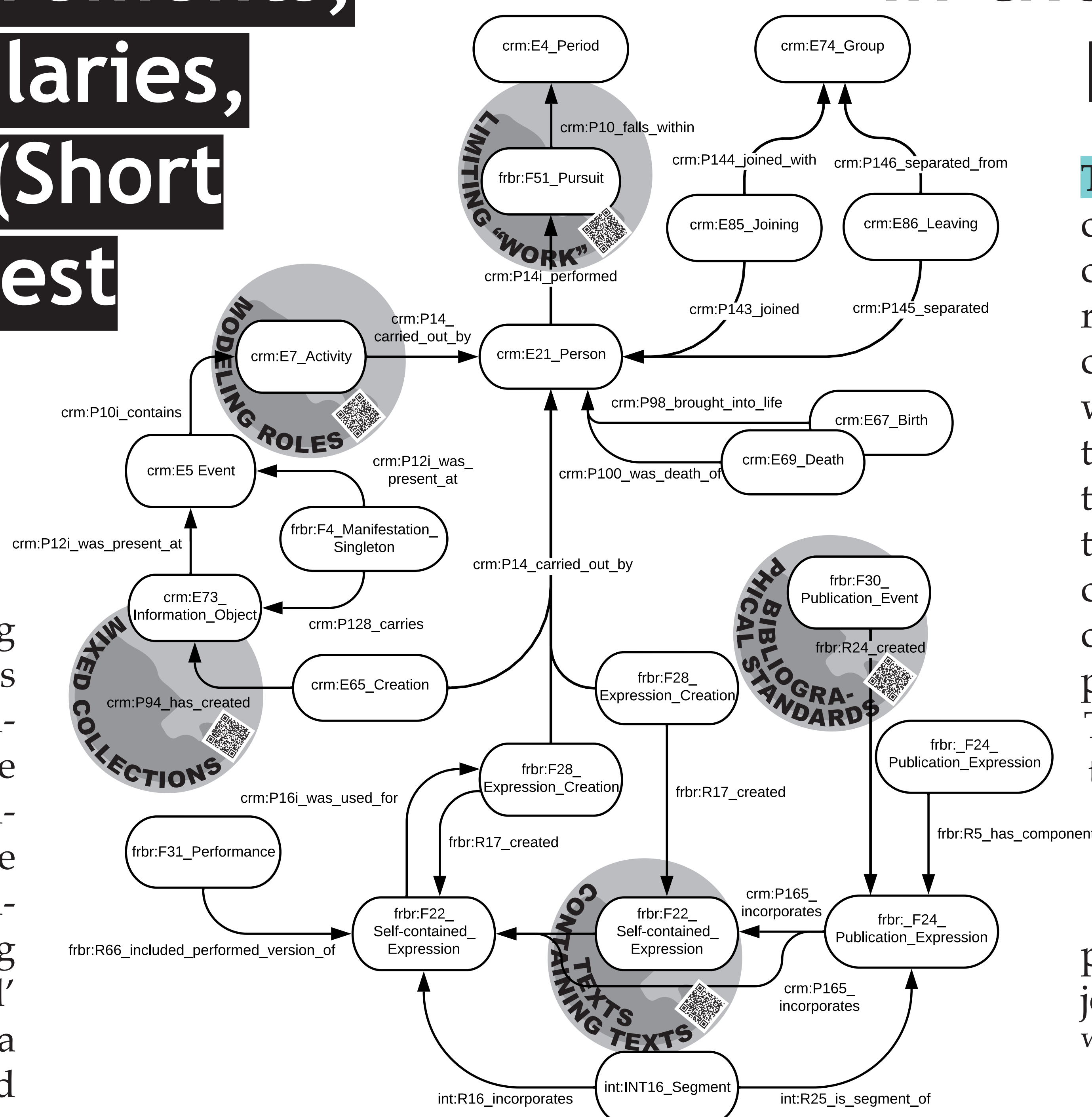
in the SemanticKraus Project Context

The SemanticKraus project is dedicated to providing a reference resource containing research data from several projects on the writer, satirist, and cultural critic Karl Kraus (1874–1936), who shaped Austrian culture from the fin de siècle to the end of the Austrian First Republic. The project aims to compile and transform biographical and bibliographical index data to create a comprehensive network of people, texts, and their interrelations. This data is currently being converted to RDF and will be accessible via a SPARQL endpoint and a UI built with ResearchSpace.

The project is open to modular expansion through future research projects in this area.

www.oeaw.ac.at/acdh/projects/semantickraus

Creating a data model for the Semantic Web with the purpose of integrating several pre-existing data sources means balancing various requirements: creating consistency across datasets while accounting for differences in respective source data, achieving a high degree of precision while considering feasibility, adhering not only to modeling standards but also those of ‘traditional’ academic practice, and more. The data model for SemanticKraus was created with these requirements in mind; various aspects of the presented model exemplify the results of these considerations.



The model above gives an overview of the modeling of biographical data (top section) – including, e.g., party affiliation and occupation – and bibliographical data: The lower section is mainly concerned with published (e.g. literary, journalistic) texts, the left section with mixed collections of Karl Kraus's court documents.

The model below operates at the level of texts and text passages: It shows the interrelation between the two and their relation to textual references to persons as well as to other texts.

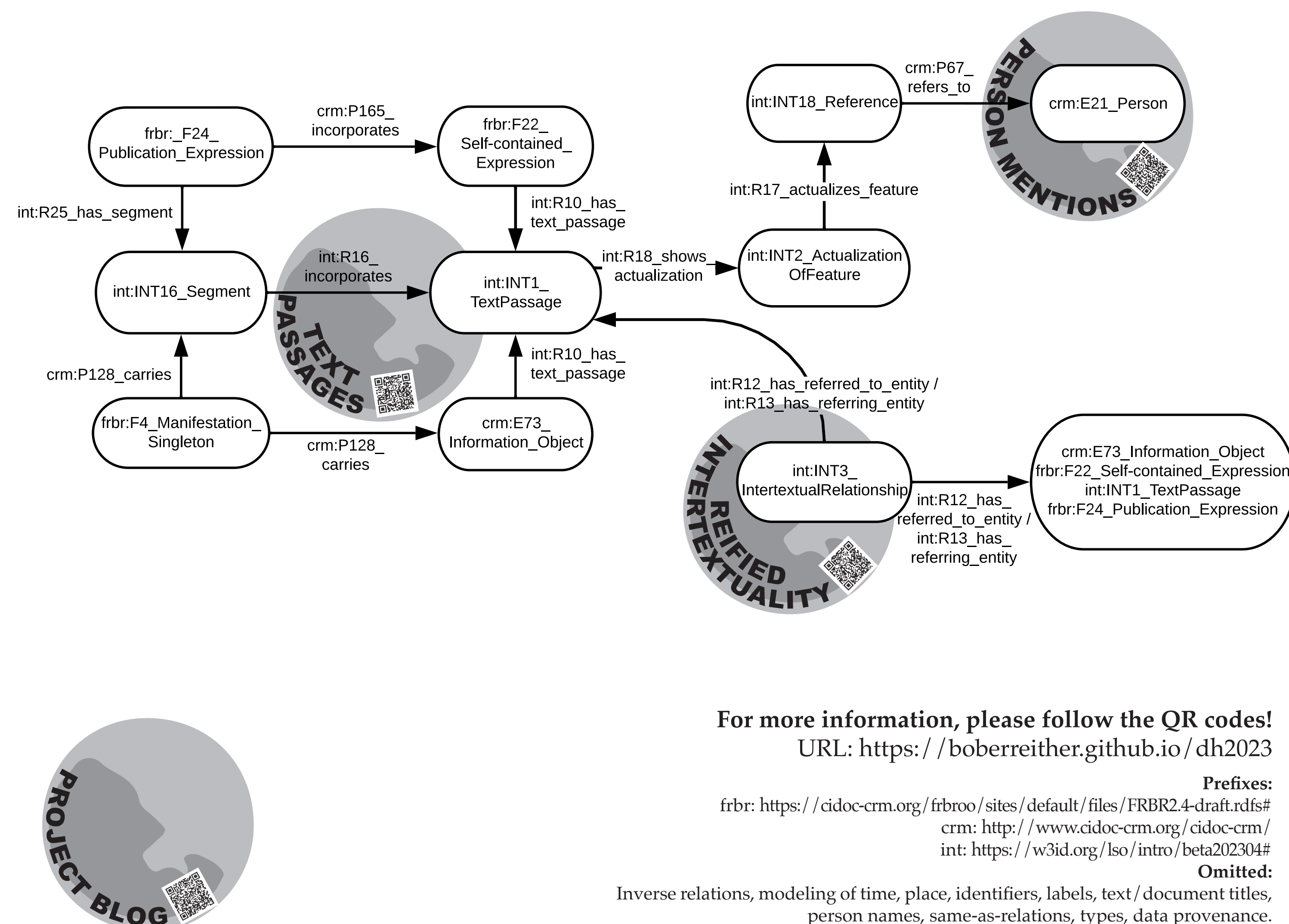
The SemanticKraus project data originate from three different Kraus-related edition projects:

Die Fackel online (fackel.oeaw.ac.at) – 16,900 persons in a biographical data base (MySQL) with 129,000 mentions in the transcriptions of *Die Fackel* (22,000 XML files); a text index (XML) listing 13,000 texts in *Die Fackel*.

Dritte Walpurgisnacht (kraus1933.ace.oeaw.

ac.at) – an index file containing 400 biographical and 1,900 bibliographical entries, referenced in 3,000 passages of a transcription of the 1933 text *Dritte Walpurgisnacht* (both TEI-XML).

Karl Kraus Rechtsakten (kraus.wienbibliothek.at) – 4,000 transcripts of legal documents (TEI-XML) with references to three index files (TEI-XML) containing 1,100 bibliographical and 1,700 biographical entries.



For more information, please follow the QR codes!
URL: <https://boberreither.github.io/dh2023>

Prefixes:
frbr: <https://cidoc-crm.org/frbroo/sites/default/files/FRBR2.4-draft.rdfs#>
crm: <http://www.cidoc-crm.org/cidoc-crm/>
int: <https://w3id.org/lso/intro/beta202304#>

Inverse relations, modeling of time, place, identifiers, labels, text/ document titles, person names, same-as-relations, types, data provenance.

