

ORD-Xplore: Bridging Open Research Data Collections through Modality Abstractions



Universität Zürich^{UZH} Madhav Sachdeva¹, Michael Blum¹, Yann Stricker², Tobias Schreck³, Rudolf Mumenthaler^{2,4}, and Jürgen Bernard^{1,4}

¹University of Zurich, Switzerland; ²University Library Zurich, Switzerland; ³Graz University of Technology, Austria; ⁴Digital Society Initiative, Switzerland



Digital Society Initiative

Library

University of Zurich



Problem statement

Not necessarily every digital edition adheres to meta(data) standards to enable interoperability, re-use and long-term archiving [1]. A high degree of manual effort is required from stakeholders for data curation and standardization [2].

Approach

We characterize data, users, and tasks to derive requirements and guidelines for ORD-Xplore to support digital librarians through ideating meaningful abstractions [3] by using modalities.

Modalities

We refer to modalities as abstract data characteristics that offer opportunities for digital librarians to gain a high-level perspective on unifiable aspects across digital editions.

Text - The textual content in a document, can be subdivided into categories (Handwritten, Correspondence, Annotations)

Spatial - Describes geographical alignment of editions, and enables the analysis of spatial relationships across editions

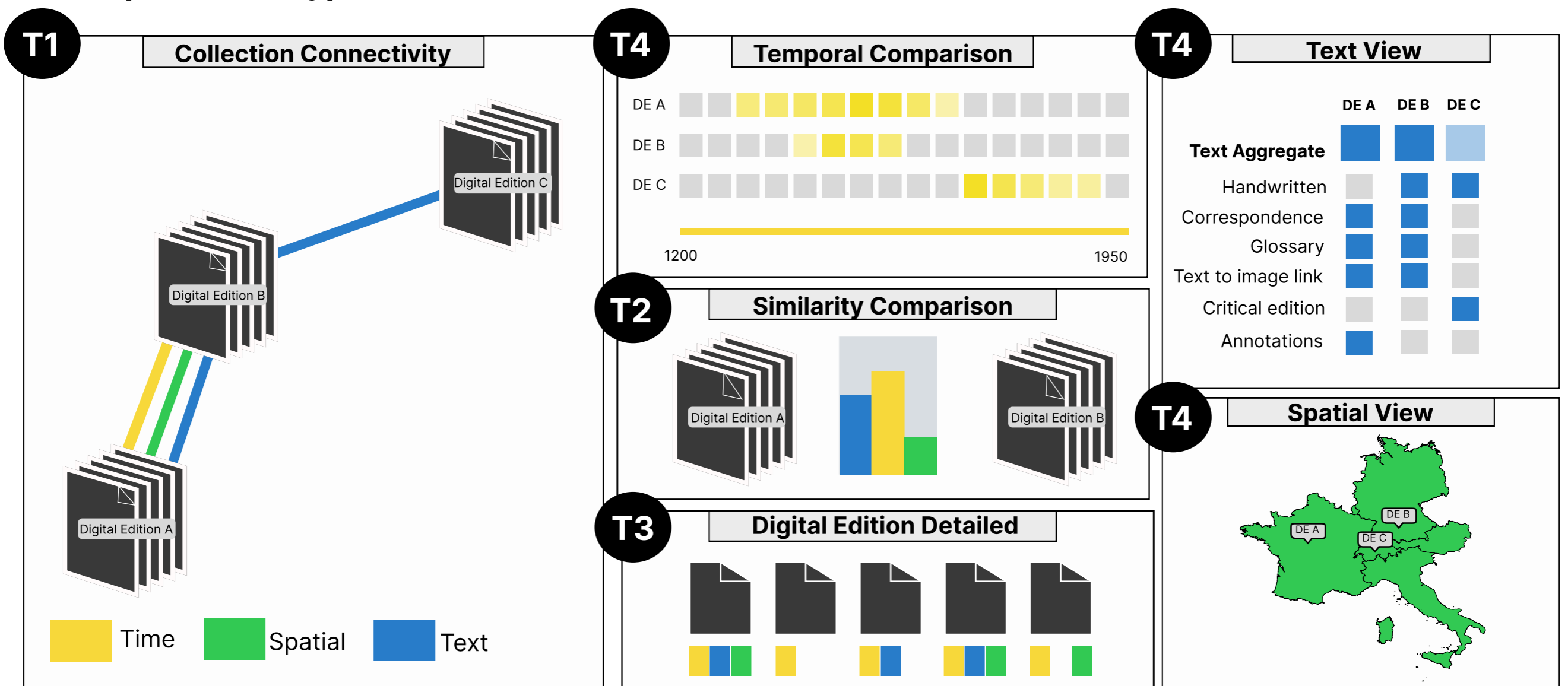
Temporal - Describes the time-orientation of digital editions and serves as external primary key to align documents within and across editions

Heterogeneity of Digital Editions

We derive modalities from the digital editions and their tags.

	date	placeName	country	location	sourceDesc	sic	idno	language
Atharvaveda	Yellow	Grey	Grey	Grey	Blue	Grey	Light Blue	Grey
Bullinger	Yellow	Green	Green	Green	Blue	Grey	Blue	Blue
Heinrich Wölfflin	Light Yellow	Light Green	Grey	Grey	Light Blue	Blue	Grey	Light Blue
Königsfelden	Yellow	Green	Grey	Grey	Blue	Blue	Blue	Grey
Robert Walser	Yellow	Grey	Grey	Grey	Blue	Blue	Light Blue	Grey
iurisprudencia	Grey	Grey	Grey	Grey	Blue	Blue	Blue	Grey

ORD-Xplore Prototype



Tasks: We draw four analysis tasks for supporting digital librarians in the standardization and unification process

Task 1: Connectivity Overview: explore by which modalities digital editions may be unified

Task 2: Pairwise edition comparison: compare unification possibilities between two digital editions

Task 3: Edition details: analyze details of modalities for a single digital edition in focus

Task 4: Modality Exploration: explore a modality across digital editions; provide support for all ideated modalities

Workshop Outcome

In a workshop with 22 digital librarians and other domain experts, we made the following findings:

- Digital editions can be unified through time, spatial and text modalities which are found across digital editions
- Not necessarily every edition adheres to the same metadata standards within and across digital editions
- Besides the 3 modalities, there can exist other modalities or divisions of existing modalities that may be helpful such as people/entities
- Editions often are heterogeneous in tag usage, which calls for more fine-grained bottom-up analysis at the granularity of tags

Future work

We plan to also investigate bottom-up tag-based analysis, possibly augmented with content-based edition abstractions given the limitation of remaining at a structural level with the current prototype. The bottom-up analysis is also reinforced by domain experts and digital librarians from our workshop. In addition, a stronger focus on data curators, as a second user group will help towards solving the gaps in metadata needed for unification efforts.

References

- [1] Bleier R., Bürgermeister M., Klug H. W., Neuber F., Schneider G.: Digital Scholarly Editions as Interfaces, vol. 12. BoD—Books on Demand, 2018. 1, 2
- [2] Wilkinson M. D., Dumontier M., Aalbersberg I. J., Appleton G., Axton M., Baak A., Blomberg N., Boiten J.-W., da Silva Santos L. B., Bourne P. E., et al.: The fair guiding principles for scientific data management and stewardship. Scientific data 3,1 (2016), 1–9. 1
- [3] Munzner T.: A nested process model for visualization design and validation. Transactions on Computer Graphics (TVCG) 15, 6(2009), 921–928. doi:10.1109/TVCG.2009.111. 1