

PRISMS: a knowledge graph in support of open scholarship

Alexander Huber¹ and Emma Huber¹

¹Huber Digital, Oxford, UK, digital@hubers.org.uk

1 Introduction and problem statement

This paper presents the motivation for and realisation of a new open-access open scholarship platform (currently in public beta) based on the Semantic Web, named PRISMS.¹

The aim of the PRISMS Open Scholarship platform is two-fold:

1. It offers a publication platform for digital scholarly editions, with full-text (preferably encoded in TEI) and facsimiles, and any accompanying materials, such as introduction, editorial statement, critical apparatus and results of analytical processes, contextual source materials, bibliography, and indices;
2. It facilitates the semantic annotation of these editions and their related scholarship (in any format) by enabling easy-to-perform formal ontological modelling (based on the CIDOC-CRM family of ontologies²).

PRISMS was born out of the realization that digital editions do not break with the historicity or materiality of the sources they organize and present, but instead remediate and extend them in ways that enable new forms of access, engagement, presentation, and analysis. PRISMS conceptualizes digital editions as living entities that perform rather than merely document the remediation they engage in.

The scholarship that underpins each digital edition provides the essential context for these remediation processes, and collectively they sustain the knowledge network that supports all scholarly engagement with the texts from any disciplinary viewpoint. PRISMS is designed to allow for the collaborative and collective modelling of this continuum of digital editions and scholarship by placing digital editions, their material and contextual basis, and the resulting academic engagement in a linked context, building on the standards and tools provided by the Semantic Web.

We believe that this type of formalization is beneficial for the purposes of this project in at least three ways: firstly, ontologies encourage us to be as precise and explicit as possible, thus reducing the reliance on implicit knowledge. Secondly, formal models encourage collaboration as they can be shared, re-used, adapted (forked), enhanced, aggregated, and developed collaboratively. Thirdly, as a form of knowledge

¹ <https://www.prisms.digital/>

² <http://www.cidoc-crm.org/>

representation, visualization, and preservation, formal models support computational processing and ultimately reasoning, and can develop alongside the mental models and human reasoning we engage in as scholars.

2 Approach and implementation

PRISMS can be used as a delivery platform for the tens of thousands of texts that are included. It also supports the creation and publication of new editions. An edition can either be created as a standalone entity and subsequently published on PRISMS, a process for which free online training is provided³, or an edition can be created inside PRISMS, e.g. by importing a IIIF facsimile of a text and by starting transcribing it either in an XML-aware or a standard text-editor.

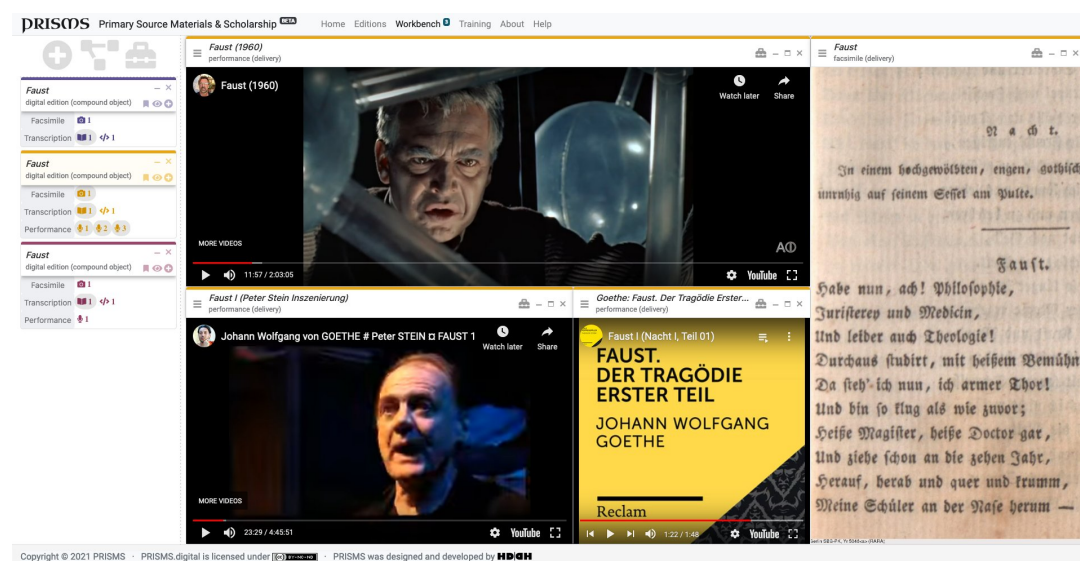


Figure 1: A view of the PRISMS workbench, with three editions of Faust loaded, and an aggregation of primary and research materials in support of a performance analysis, with a facsimile, two videos, and an audio book.

As such, PRISMS supports both the discovery or creation, and the presentation of full-text editions. More importantly, PRISMS enables the scholarly aggregation of relevant primary and secondary sources materials necessary to support the scholarly exploration of the materials from a given disciplinary or methodological point of view (see e.g. Figure 1). The purpose of this aggregation process is two-fold:

1. It supports the entire research process from the early stages of discovery and collection right through to publication and dissemination, sometimes referred to as scholarly primitives;⁴
2. It provides a set of tools in support of subsequent analytical and modelling steps on the level of both the individual edition and across editions.

³ <https://www.prisms.digital/training/>

⁴ <https://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html>

2.1 Tools

PRISMS hosts a variety of tools alongside the digital editions, which can be categorized as component tools (such as text-based tools, image-based tools, XML-based tools, etc.) and workbench tools (those available across document types and editions). The former category of tools is useful for any type of close scholarly work, and in PRISMS these tools include a bookmarking tool, an annotation tool for both texts and images, the ability to keep research records or transcriptions in the form of notebooks, and integration of Voyant Tools⁵ for statistical analysis and a variety of visualisations of texts. The latter category includes the ability to participate in the shaping of the knowledge graph by modelling concepts and relationships, an easy way to organize your research materials on the workbench, and the ability to download, share, and publish contributions for the benefit of all.

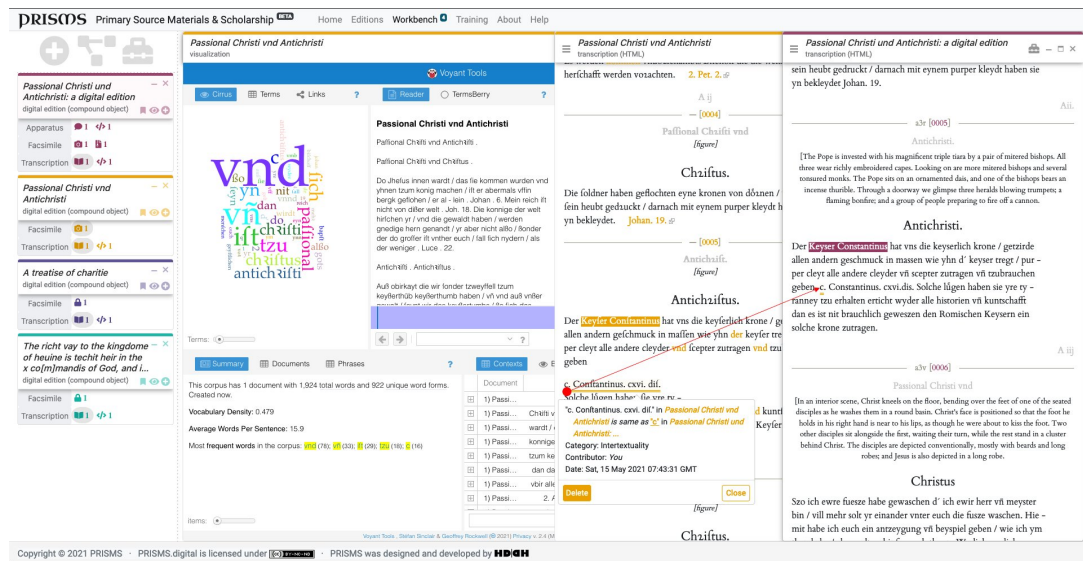


Figure 2: Some of the built-in analysis and visualization tools in PRISMS: Voyant Tools is shown alongside a relation being made between two editions, and annotations highlighted.

Figure 2 shows a snapshot of some of the component tools available in PRISMS. Component tools support textual work, particularly mark-up of entities, the creation of annotations and connections between textual elements either within or across texts. All interactions with the texts are stored in open standards, such as the W3C's Web Annotation Data Model. PRISMS has been launched with corpora from the EEBO-TCP⁶, ECCO-TCP⁶, EVANS-TCP⁶, the DTA extended core corpus⁷, and the Taylor Editions⁸ scholarly editions platform, and as mentioned before new texts and collections can be added at any time.

2.2 The graph

Every edition published in PRISMS has received basic modelling of both their material and digital manifestations based on the available components (e.g. facsimiles,

⁵ <https://voyant-tools.org/>
⁶ <https://textcreationpartnership.org/tcp-texts/>
⁷ <https://www.deutschestextarchiv.de/>
⁸ <https://editions.mml.ox.ac.uk/>

apparatus, and secondary materials) and metadata at ingest into the platform. These models are expressed using CIDOC-CRM (particularly the Core and FRBR/LRM models, supplemented by provenance and digital extension ontologies) and are stored in the RDF-native Blazegraph database⁹ that underpins PRISMS.

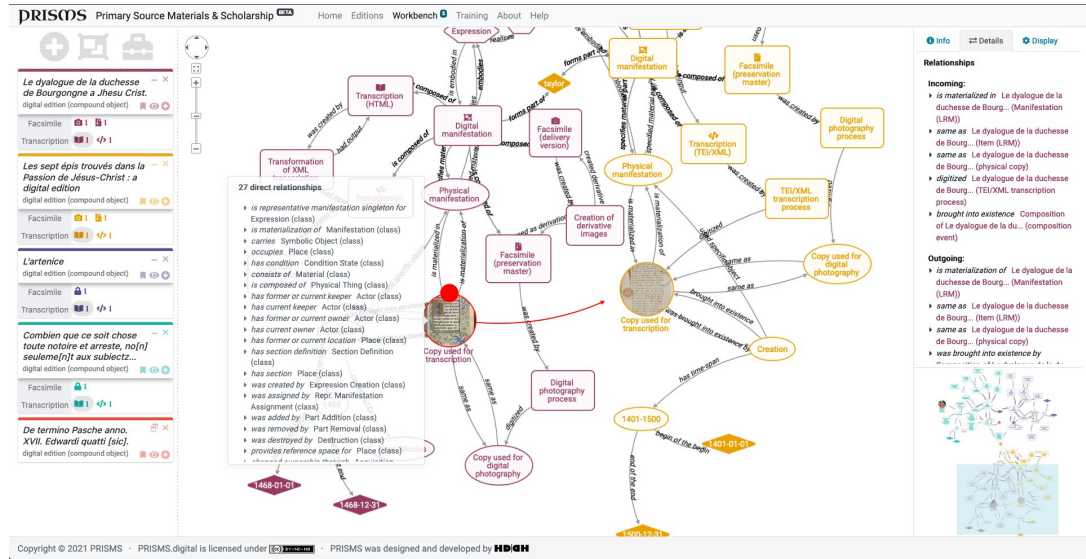


Figure 3: Modelling both the material legacies and digital remediation processes with Linked Data technologies and Cytoscape.js

Figure 3 shows a snapshot of the visual modelling interface created by PRISMS that facilitates easy-to-perform formal ontological modelling using the CIDOC vocabularies and a Cytoscape.js-powered frontend for visualizing user interaction.¹⁰ The graph is queried and updated in real-time with every user interaction using SPARQL and SPARQL Update. The application is powered by JSON-LD created from the SPARQL queries the user triggers in the interface when engaging with the platform. All interactions with the platform are stored in the user's private graph for easy retrieval and display. These include any bookmarks, annotations or relations made, any semantic modelling that has been performed, or snapshots of the workbench, and are all stored in the graph database. By default, everything in PRISMS is private. Everything contributors add is immediately visible to them (in their private graph), and they can conduct their scholarship in complete privacy, without delay or interference. Only when the user decides to publish their contributions will they be made available to everyone. Depending on the type of contribution made, there are options to share, download, and/or publish them.

3 Contribution and further work

PRISMS is about connecting primary and secondary resources semantically. A prerequisite for this process is the ability to add resources to PRISMS in an easy and flexible way. This includes anything from a digital facsimile in IIIF, a transcription of a source text, a PDF of a journal article, a video of a theatrical performance, an audio book, an image or a 3D-model of a sculpture mentioned in a text, etc. We have

⁹ https://github.com/blazegraph/database/wiki/Main_Page

¹⁰ PRISMS also supports a more text-centric input method that is based on HTML forms.

found the graph underpinning PRISMS, and the modelling it enables in real time, to be attractive to a range of users, from those interested in digital Book History to those looking for open alternatives to paywalled subscription databases for use in the classroom and beyond.

Moving forward, we will continue to work on integrating the digital research and tools important to PRISMS' users (e.g. reference manager, images taken in reading rooms, items deposited in institutional repositories). We envision that over time, PRISMS will evolve both as a powerful discovery tool and a personal research tool.

References

- Ciotti, F. (2015). Digital methods for literary criticism. Lecture slides. University of Rome Tor Vergata.
- Ciula, A. and Ø. Eide (2017). Modelling in digital humanities: Signs in context. *Digital Scholarship in the Humanities* 32, i33–i46.
- Ciula, A., Ø. Eide, C. Marras, and P. Sahle (2018). Models and Modelling between Digital and Humanities — A Multidisciplinary Perspective. *Historical Social Research Supplement* 31.
- Ciula, A. and C. Marras (2016). Circling around texts and language: towards 'pragmatic modelling' in Digital Humanities. *Digital Humanities Quarterly* 10(3).
- Eide, Ø. (2015). *Media Boundaries and Conceptual Modelling: Between Texts and Maps*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
- Kräutli, F. and M. Valleriani (2018). CorpusTracer: A CIDOC database for tracing knowledge networks. *Digital Scholarship in the Humanities* 33(2), 336–346.
- Oldman, D., M. Doerr, and S. Gradmann (2016). Zen and the art of linked data: New strategies for a semantic web of humanist knowledge. In S. Schreibman, R. Siemens, and J. Unsworth (Eds.), *A New Companion to Digital Humanities*, pp. 251–273. Malden, MA: Wiley Blackwell.