**PREBOR — Ontopedia**

<2 images>

**Towards an Ontopedia for Post-Medieval Hebrew Manuscripts**

Prebor, G. and Zhitomirsky-Geffet, M.

Digitization of the national cultural heritage is a rapidly expanding field essential for preserving and maintaining the historical data and leveraging its future research. Historical handwritten Hebrew manuscripts are one of the most unique and authentic evidences of the Jewish culture and thought that survived through the centuries. Scholars from various fields increasingly study these manuscripts to reveal historical, linguistic, religious, philosophical, and social aspects of the Jewish life in different times and places (Sirat, 2002; Richler, 2010–2011). Currently, the only available digital representation of these manuscripts’ metadata is library catalogs. The largest collection of Hebrew manuscripts metadata is offered by the catalog of the Institute of Microfilmed Hebrew Manuscripts in the National Library of Israel (http://web.nli.org.il/). This catalog is accompanied with a search engine to retrieve records by a limited number of parameters, such as author, title, date, and subject, while most of the data still remain unsearchable and thus undiscovered. Many important research questions cannot be answered by searching the existing manuscript catalogs, e.g.: How many works on specified subjects were composed in this period in different countries? What people were involved in their creation and distribution? Are these people related to each other, and how? What historical events could influence these people and their works? Are these original works or copies of older works from other manuscripts or even from printed books? Hence, to enable a systematic research of the knowledge embedded in the manuscripts there is a need for a formal conceptual data model with a high level of semantic granularity, an ontology. To the best of our knowledge, there is no formal ontology for the realm of historical handwritten Hebrew manuscripts.

Therefore, in this research we design an ontological model to reflect all the cultural riches stored in post-medieval Hebrew manuscripts. We focus on the post-medieval period (16th century and later), because tens of thousands of works that belong to this period are underexplored in the research literature (Prebor, 2015). The underlying philosophical approach behind the proposed ontology is to view a manuscript as a ‘living entity’ and develop a data model of its story. This model includes stages and milestones in its biography (e.g., creation, printing, or acquisition), and its influence and interactions with other manuscripts, people, places, and historical and cultural events (see Figure 1). A sequence of events and places (as Jewish writers were spread over the world) constitutes a timeline of history against which manuscripts, people, and their relationships can be placed. Apparently, this ontology creation requires researching the most appropriate existing ontologies for the domain of cultural heritage (such as CIDOC-CRM, a result of a 10-year project [Doerr, 2003] and Europeana Data Model [Winer, 2014]).

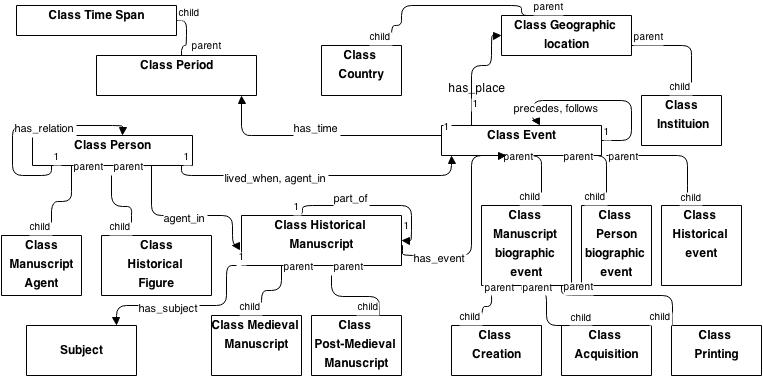


Figure 1. A proposed conceptual model for the upper-level historical manuscript ontology. The square boxes represent the top-level classes and their subclasses; the properties are displayed by single-line arrows, while the names of properties appear near the corresponding arrows. Child-parent arrows represent IS-A/subclass relationships.

Finally, the ontopedia (online ontology-based encyclopedia) for the Hebrew historical manuscripts will be constructed as a culmination of this multidisciplinary study that comprises and integrates research in various fields of science. Thus, computer scientists are involved in automatic mapping of catalog fields to ontology and user interface design, information experts contribute to the conceptual model development and evaluation, librarians sort and analyze catalog records, while researchers of Hebrew paleography complete the missing features of the manuscripts and build query sets for extending the ontopedia.

The results of our project will greatly contribute to the study of Hebrew manuscripts and cultural heritage. It will enable posing queries and cross-referencing data from various vocabularies in the Semantic Web (as in Figure 2). Large-scale automated reasoning will also enable a comparison of the effect of time and place on qualitative characteristics and quantitative distribution of manuscripts.

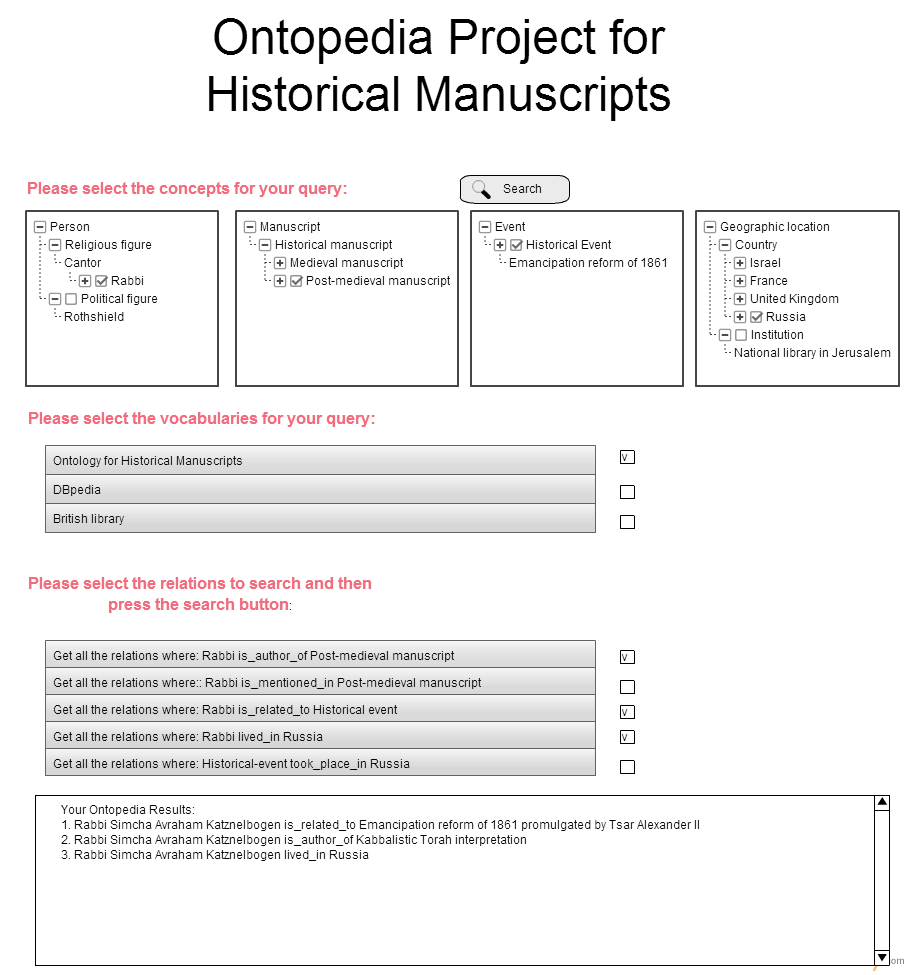


Figure 2. Sketch of the user interface for query construction for selected ontological relations by browsing the ontology. A sample query is: ‘All the Rabbinic figures who authored post-medieval manuscripts related to some historical event in Russia’.

**References**

**Doerr, M.** (2003). The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine,* **24**(3): 75–92.

**Prebor, G.** (2015). The Manuscript Collection of the Séminaire Israélite de France. Towards a New Catalog. *La Bibliofilia*, CXVII.

**Richler, B.** (2010–2011). Lezione Dottorale, *Materia giudaica*, XV–XVI, 14–16.28.

**Sirat, C.** (2002). *Hebrew Manuscripts of the Middle Ages*. Cambridge University Press, Cambridge.

**Winer, D.** (2014). Judaica Europeana: An Infrastructure for Aggregating Jewish Content. *Judaica Librarianship,* **18**(1): 88–115.