Doctoral Symposium on Research on Online Databases in History (RODBH 2019)

April 4th-5th 2019, Leipzig

Thomas Riechert,[1] Francesco Beretta,[2] George Bruseker,[3] Edgard Marx,[4] Jennifer Blanke,[5] Vincent Alamercery,[6] Tracy Hoffmann,[7] Natanael Arndt[8]

Abstract: This editorial provides an introduction to the field of research of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019) which was collocated with the 3rd Data for History workshop. The workshop series is situated in the field of digital humanities and targets the interconnection of subjects of historical research, knowledge engineering, and information science. The common interlink of this disciplines is the use of research data, data management, and all accompanying activities as well as the organization of collaborative community processes.

1   Introduction

Finding interconnections and relations between historical artifacts is an aim of historical research. The aim of knowledge engineering is to formalize and manage sets of interconnected data. Historical research covers interdisciplinary domains e.g. from geography to social science on identifying historical places, sites, and settlements and its relation to historical events and persons. This endeavor of working with cross domain interdisciplinary data involves a lot of material, often in hardly manually to manage quantities.

Historians collect data and collaborate to include external data sets to their own research data. A methodological model to support this collaboration is described by the Heloise Common Research Model (HCRM) [RB16], a service based layer model. Services for the collaboration on data sets in historical research are provided in three different layers: repository layer, application layer, and research interface layer. Within the HCRM, the

---

[1] Leipzig University of Applied Sciences, Germany; thomas.riechert@htwk-leipzig.de

[2] CNRS, Université de Lyon, Laboratoire de recherche historique Rhône-Alpes, France; francesco.beretta@cnrs.fr

[3] Foundation for Research and Technology - Hellas, Institute of Computer Science, CCI, Greece; bruseker@ics.forth.gr

[4] Leipzig University of Applied Sciences, Germany; edgard.marx@htwk-leipzig.de

[5] Herzog-August Library, Wolfenbüttel, Germany; blanke@hab.de

[6] École normale supérieure de Lyon, Université de Lyon, Laboratoire de recherche historique Rhône-Alpes, France; vincent.alamercery@ens-lyon.fr

[7] Leipzig University Library, Leipzig, Germany; tracy.hoffmann@uni-leipzig.de

[8] Institute for Applied Informatics, Leipzig, Germany; arndt@informatik.uni-leipzig.de

repository layer handles the data access and storage while the application layer provides services to interlink data sets and deals with the resulting complexity to enable research interfaces that are accessible to historians.

The heterogeneous and highly collaborative setup of projects in historical research involves a high need for communication, driving the necessity of making the data provenance and evolution accessible. In the context of historical prosopographical research, the provenance of the data is relevant to evaluate their credibility and to consider the influence of their origin. The same applies to the terminology evolution, it allows the interpretation and understanding of the past in the present. Several approaches exist to improve the efficiency of the terminology evolution process by involving the community in a collaborative process employing an Ontology Management Environment such as OntoME[9] [Be18].

In this editorial, we present an overview of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019)[10] collocated with the 3rd Data for History workshop[11], held from April 4th to 5th 2019 in Leipzig. The doctoral symposium was hosted by the Leipzig University of Applied Science, Faculty of Computer Science and Media (HTWK) as part of the Research project *Early Modern Professorial Career Patterns - Methodological research on online databases of academic history*, in cooperation with the European Consortium on Data for History (DfH). A general meeting of the DfH Consortium was part of the event. It included the discussion of the OntoME application, an ontology management environment, as a platform to design vocabularies and ontologies such as those contributed by the workshop participants.

The remainder of this editorial is organized as follows. Section 2 presents a summary of the contributions presented in the symposium. Section 3 concludes giving an outlook on the future work.

2   Summary of the RODBH 2019 Workshop

The doctoral symposium dealt with research questions related to online databases in the academic field of History. It gathered doctoral students from the Humanities, Computer Science, and Digital Humanities engaged in the application of Semantic Web technologies in their research and concerned with ontology building in their projects.

The symposium provided students an interdisciplinary and international forum as a scientific environment to present and discuss their research. The symposium specifically aimed to raise issues concerned with the research and methodology in historical research.

In the Data for History meeting, students were able to learn from experts in an atmosphere of collaborative feedback and suggestions both from peers and experienced researchers,

---

[9] Homepage of the OntoME tool: `https://ontome.dataforhistory.org/`

[10] The RODBH 2019 workshop homepage: `https://pcp-on-web.htwk-leipzig.de/event/RODBH2019`

[11] The Data for History workshop homepage: `http://dataforhistory.org/3rd-data-for-history`

giving the opportunity to obtain an insight into directions for research on online databases in History and an opportunity to network with peers and future colleagues.

Eight contributions [Po19, Na19, Pi19, Ho19, Ra19, Od, Co, Je] were accepted for presentation, out of which five papers were included in the proceedings [Ri19] which we describe in brief.

The paper *Digital Edition Publishing Cooperative for Historical Accounts and the Book-keeping Ontology* by Pollin [Po19] is concerned with historical bookkeeping. It explains the formalization of historical financial records within the Bookkeeping Ontology. The author defines a workflow to publish RDF data, as part of the digital editions of historical financial records, as Linked Open Data.

The paper *Data Modeling of Complex Historical Information* by Nasarek [Na19] focuses on ontology engineering based on occupational articles from early modern encyclopedias and modern classification schemes. The paper describes a workflow to implement information from unstructured text into a labelled property graph and gives modeling guidelines for dealing with the multidimensionality of ontological and process-oriented data.

In the paper *Definition of the life cycle of cultural property, the concept of stratum* Pineau [Pi19] proposes a life cycle model of cultural property that allows the representation of different key moments and trends in the life of a cultural property. By adding strata to the life cycle, it is possible to isolate specific analyses of certain problems in the life of the cultural property.

With the paper *Developing a Mediated Vocabulary for Video Game Research* by Hoffmann [Ho19] we take a look on modern cultural resources to identify possibilities to develop and adapt data management principles to be jointly applied to modern and historic data. The paper describes an approach to a video game vocabulary capable of describing video games for multiple contexts and of integrating heterogeneous video game databases.

Radisch [Ra19] looks at geographic information in the paper *Automated Georesolution of Place Names in Serial Sources*. This paper presents a algorithm-based solution to automate georesolution of place names in historical serial sources.

In addition to the doctoral presentations, two invited project presentations were delivered to the symposium. Pierre Vernus from Université de Lyon presented the H2020 SilkNow project,[12] a project concerned with the understanding, conservation and dissemination of European silk heritage. Meanwhile, Sebastian Hellmann from the Leipzig University/Institute for Applied Informatics, Leipzig presented recent developments in the DBpedia data bus[13], an infrastructure to support improvements to the management of the data source and accessibility by agents.

---

[12] Homepage of the SilkNow project: http://silknow.eu/
[13] Homepage of the DBpedia data bus project: https://databus.dbpedia.org/

The further schedule of the workshop was taken up by the work of the Data for History consortium. The sessions were concerned with the discussion of four broad topics:

- The governance of the consortium,

- Geographic modelling,

- Metadata organization, and

- Modelling questions.

The discussion on governance was led by Charles van den Heuvel (Huygens ING - Amsterdam) and Francesco Beretta (CNRS - Université de Lyon). Specifically different models of governance and obtaining official status were introduced and debated. In particular, the possibilities to become an official DARIAH working group were explored. Further, the experiences of the TEI community were examined as an inspiration for Data for History.

The topic of prosopography was dealt with in two presentations: *Recording people* by Richard Light (Free UK Genealogy) and *Creating new CIDOC CRM classes and properties to describe silk-related artefacts* by Marie Puren (Université de Lyon).

Issues related to geographical data and historical places were discussed in the presentations *Settlement and administrative units types* by Bogumił Szady (Katolicki Uniwersytet Lubelski Jana Pawła II - Lublin) and *Publishing historical gazetteers & alignments to ontologies* by Carmen Brando (EHESS - Paris).

The organization of metadata and general modelling were discussed in the presentations *How to facilitate multi-perspective data exploration?* by Lodewijk Petram and Sebastiaan Derks (Huygens ING - Amsterdam), *CIDOC CRM Based Provenance Metadata for Interoperability of Cultural and Research data* by Regina Varnienė-Janssen (Vilniaus universitetas), and *Of creators and copyists. Modeling storylines of artists, objects and their replicas. Authenticity, provenance and validation* by Charles van den Heuvel (Huygens ING - Amsterdam).

Following the formal presentation, the symposium participants split into four different working groups in order to discuss the individual topics in detail.

During the course of the second day of the symposium, a hands-on tutorial was provided to the participants in order to learn how to align ontologies using the OntoME tool. The presentation of general data alignment workflow brought up a wide discussion about the issue of the collaborative ontology management in a distributed research consortium.

The workshop was closed by an open discussion of lessons learned and potential directions involving the gathered participants.

3   Conclusion

In the Doctoral Symposium on Research on Online Databases in History (RODBH 2019) we had a diverse range of works that focus on representation and modeling historical data through RDF ontologies as well as on reasoning on this data. Several subjects were highlighted in the doctoral symposium such as:

- The usage of a common repository and platform to share and consume historical data,

- The viability to apply a common research ontology,

- Collaborative ontology management and development in a distributed research consortium, and

- The issues of modeling events, accounts, and geographical historical data.

The access to current available approaches for storing, exploring, and sharing historical data needs to be adapted and well defined to be used by historians (the research interface layer in the HCRM model). More specifically, there is a need to facilitate the finding, exploration, and visualization of ontologies and data sets. In a community process a common understanding of data is necessary. The Data for History online forum[14] can be used to establish a better communication within the community to guide future research and development directions.

Acknowledgements

---

[14] The Data for History Forum: `http://forum.dataforhistory.org/`

References

[Be18]   Beretta, Francesco: OntoME. An ontology management environment for aligning data models and producing FAIR data. In: Time Machine Conference 2018. Lausanne, Switzerland, October 2018.

[Co]   Cosson, Julie: "Refuge huguenot" database: Digital Humanities applied to a 17th century diaspora. Oral Presentation.

[Ho19]   Hoffmann, Tracy: Developing a Mediated Vocabulary for Video Game Research. In: Proceedings of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019). 2019.

[Je]   Jeanson, Loïc: A first draft on a tool for ontology alignment tolerance. Oral Presentation.

[Na19]   Nasarek, Robert: Data Modeling of Complex Historical Information. In: Proceedings of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019). 2019.

[Od]   Odore, Angelo: Marseille during the French Revolution, 1789-1794. Historical mapping test from GIS. Oral Presentation.

[Pi19]   Pineau, Karl: Definition of the life cycle of cultural property, the concept of stratum. In: Proceedings of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019). 2019.

[Po19]   Pollin, Christopher: Digital Edition Publishing Cooperative for Historical Accounts and the Bookkeeping Ontology. In: Proceedings of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019). 2019.

[Ra19]   Radisch, Erik: Automated Geo-resolution of Place Names in Serial Sources. In: Proceedings of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019). 2019.

[RB16]   Riechert, Thomas; Beretta, Francesco: Collaborative Research on Academic History using Linked Open Data: A Proposal for the Heloise Common Research Model. CIAN-Revista de Historia de las Universidades, 19(0), 2016.

[Ri19]   Riechert, Thomas; Beretta, Francesco; Bruseker, George; Marx, Edgard; Blanke, Jennifer; Alamercery, Vincent; Hoffmann, Tracy; Arndt, Natanael, eds. Proceedings of the Doctoral Symposium on Research on Online Databases in History (RODBH 2019). CEUR Workshop Proceedings, April 2019.