

Semantic Web for Pre-Modern Scientific Data

1. Context

Since the mid-1970s classicists and historians have developed textual databases, intending mostly to gather and explore large amounts of primary source materials. For a long time, they mainly focused on text digitization and markup. They only recently decided to try to explore the possibility of transferring some analytical processes they previously thought incompatible with automation to knowledge engineering systems, thus taking advantage of the growing set of tools and techniques based on the languages and standards of the semantic Web, such as linked data, ontologies, and automated reasoning. On the other hand, Semantic Web researchers are willing to take up more ambitious challenges than those arising in the native context of the Web in terms of anthropological complexity, addressing meta-semantic problems of flexible, pluralist or evolutionary ontologies, sources heterogeneity, hermeneutic and rhetoric dimensions. Thus the opportunity for a fruitful encounter of knowledge engineers with computer-savvy historians and classicists has come. This encounter may be inscribed within the more general context of digital humanities, a research area at the intersection of computing and the humanities disciplines which is gaining an ever-increasing momentum and where the Linked Open Data is playing an increasingly prominent role.

2. Goals and motivations

The purpose of the workshop is to provide a forum for discussion about the methodological approaches to the specificity of annotating “scientific” texts (in the wide sense of the term, including disciplines such as history, architecture, or rhetoric), and to support a collaborative reflection, on possible guidelines or specific models for building historical ontologies. A key goal of the workshop, focusing on research issues related to pre-modern scientific texts, is to emphasize, through precise projects and up-to-date investigation in digital humanities, the benefit of a multidisciplinary research to create and operate on relevantly structured data.

The workshop organizers all belong to the Zoomathia¹ international research network funded by the French National Scientific Research Center (CNRS). This network gathers French, Italian, German and English researchers and aims to study the formation and transmission of ancient zoological knowledge over a long period, with an historical, literary and epistemological approach, and create open knowledge sources on classical zoology to be published on the web of linked data. This workshop is an opportunity to present the activity of the network, to enlarge the network with interested participants of the workshop and to benefit from the results of related research projects.

3. Particular interest of the topic

Beyond the significance of semantic Web technology and formalization to the humanities (and vice versa), and the advantage for researchers to bridge the gap between traditional practices of textual analysis and knowledge engineering practices afforded by Semantic Web models and technologies, one of the principal interest of the very topic of pre-modern historical data management lies in *historical semantics*, and the opportunity to jointly consider how to identify and express lexical, theoretical and material evolutions. Dealing with historical texts, a major problem is indeed to handle the discrepancy of the historical terminology compared to the modern one, and, in the case of massive, diachronic data, to take into account the contextual and theoretical meaning of words and sentences (that changes according to the genre, the date, and the general origin of the texts), and their semantics (in other words, what they refer to, which is also liable to change). In that respect the iconographic data, which are relevant in history of science and arise similar problematic could be addressed as well and offer suggestive insights for a global methodology for diverse media.

¹ <http://www.cepam.cnrs.fr/spip.php?rubrique229>.

3. Audience

This highly interdisciplinary workshop aims to bring together members of often disconnected or distant communities (philology, history, philosophy, and computer science) whose interests often intersect in the area of digital humanities and to favor cross-fertilization and exchange of ideas. Several members of the Zoomathia network will surely attend the workshop. Expected number of participants: 25.

4. Workshop format

A full-day workshop organized into several sessions where papers selected by the workshop program committee will be presented. Each session will be chaired with the aim of provoking discussion among all the participants. One special session will focus on the specific research topics of the Zoomathia research network. The workshop will close with a panel discussion, which will address the issue of defining a road map for the adoption of semantic-Web related technologies for the study of pre-modern scientific texts and of the history of the ideas and their transmission.

5. Program Committee

Here is a provisory composition of the PC, which will be extended up to twenty members.

- Annetta Alexandridis, University of Cornell, USA (confirmed)
- Valentina Ascitti, King's College, Great Britain
- Aurélien Berra, Université of Paris X, Nanterre, France (confirmed)
- Arianna Betti, University of Amsterdam, Netherlands (confirmed)
- Alice Choyke, Central European University, Hungary
- Gregory Crane, Tufts University, USA (confirmed)
- Kai Eckert, Universität Mannheim, Germany (confirmed)
- Elizabeth Mc Donald, project manager of Equipex Biblissima, Paris
- Albert Meroño Peñuela, VU University Amsterdam and KNAW, Netherlands (confirmed)
- Andrea Tettamanzi, Université Nice Sophia Antipolis, France (confirmed)
- Sabine Obermaier, Universität Mainz, Germany

7. Related workshops and conferences

- 5th International Conference of the ESHS on *Ontologies and semantic web: New topics of research for historians of science and technology*, Athens, November 2012³
- DHBenelux conference, The Hague, June 2014⁴
- A general annual forum of discussion on Digital Humanities is the Digital Humanities conference (DH), whose next edition will be held in Sidney, June 2015⁵
- Conference on *Humanités numériques: l'exemple de l'Antiquité* to be held in Grenoble, September 2015⁶
- Conference on *Humanités numériques: Identités, pratiques et théories* to be held in Montreal, August 2015⁷

8. Names and contact information of the workshop organisers/chair(s)

Arnaud Zucker, Université Nice Sophia Antipolis (France), e-mail: zucker@unice.fr

Isabelle Draelants, CNRS, IRHT (France), e-mail: isabelle.draelants@irht.cnrs.fr

Catherine Faron Zucker, Université Nice Sophia Antipolis (France), e-mail: faron@unice.fr

Alexandre Monnin, INRIA (France), e-mail: alexandre.monnin@inria.fr

³ <http://5eshs.hpdst.gr/abstracts/442>

⁴ <http://dhenelux.org/>

⁵ <http://dh2015.org/>

⁶ <http://www.irht.cnrs.fr/fr/actualites/appel-communication-colloque-humanites-numeriques-l-exemple-de-l-antiquite>

⁷ <http://hn2015.org/>

Arnaud Zucker is full professor of Classics (Greek Literature) at University Nice Sophia Antipolis since 2005. His research domain is the transmission of zoological knowledge and sciences in ancient Greece. He published several books and papers on ancient zoology and astronomy, and was twice awarded from the French Association for Greek Studies. He is Deputy Director of the CEPAM laboratory gathering archeologists, chemists, historians and philologists. He founded and is director of the electronic journal *Rursus* on rewriting process in antiquity, and is member of the editorial board of other journals or series (*Anthropozoologica*, *Blityri*, *Polymnia*, *Spicae*). He participated to the French national project SourcEncyMe on an annotated corpus of medieval encyclopaedias, and since 2014 he coordinates the international and interdisciplinary scientific network Zoomathia on the formation and transmission of ancient zoological knowledge. Since 2010 he organizes each year several international conferences and scientific meetings (3 in 2014).
<http://physiologos.org>

Isabelle Draelants is Senior researcher in Medieval History at Centre national de la recherche scientifique, in the IRHT institute in Paris. She completed her PhD in Medieval history in 2001 in Université catholique de Louvain-la-Neuve and defended her Habilitation thesis at the Paris-IV Sorbonne University in 2008 on Natural philosophy, encyclopedias and transfer of rationality before 1300. Since 2003, she is responsible of the Atelier Vincent de Beauvais, specialized in the study of medieval encyclopedism; she organized several related workshops and led a project called SourcEncyMe (Sources of medieval encyclopedias), aiming at developing an annotated online corpus of encyclopedic medieval texts, with a collaborative platform allowing to tag and identify the textual sources. She founded an electronic journal called *Spicae* focusing of medieval compilations of knowledge. She is the scientific referent of IRHT for the Zoomathia network.
<http://www.irht.cnrs.fr/fr/annuaire/draelants-isabelle>

Catherine Faron Zucker is Assistant professor of Computer Science at University Nice Sophia Antipolis since 2002. She is vice-head of the Wimmics joint research team between Inria, CNRS and UNS. Her research domains are on Graph based Knowledge Representation and Reasoning, Ontologies, Semantic Web and Social Web. She participates to several national and international R&D projects on these domains, among which she is the scientific referent of I3S for the Zoomathia network. She is the scientific organizer and chair of a session of the annual CrEDIBLE workshops on federating distributed biomedical imaging data and knowledge since 2012. She co-organised a French workshop on Artificial Intelligence and e-Education in 2013. She was the general and scientific chair of the French National conference on Knowledge Engineering in 2014.
<http://www.i3s.unice.fr/~faron/>

Alexandre Monnin is researcher at Inria in the Wimmics team where he previously initiated the French DBpedia project. He holds a PhD in Philosophy from Paris 1 Panthéon-Sorbonne on the Philosophy of the Web. He was previously Senior Consultant for UNESCO (IIEP, Paris), Fellow at IKKM (*Bauhaus Universität*) and Head of Web Research at the *Institut de Recherche et d'Innovation* of the Centre Pompidou. Since 2013, he is senior expert for Etalab, the Prime Minister Open Data agency of the French government. He co-founded the W3C "PhiloWeb" Community Group in 2012 with Harry Halpin and Henry Story. He supervised the "PHILOWEB: Philosophy of the Web" FP7 Project and launched Linked Data projects for the French Ministry of Culture. He organizes the PhiloWeb conference since 2010 and was one of the program chairs of the ACM WebScience 2013 conference in Paris. <http://web-and-philosophy.org/>

9. A one-page draft of the Call for Papers (see next page)

Call for Papers: Workshop on Semantic Web for Pre-Modern Scientific Data

The first **Workshop on Semantic Web for Pre-Modern Scientific Data** will be held in conjunction with the **12th ESWC 2015 Conference** which takes place between **31st May and 4th June** in **Portoroz, Slovenia** and will provide a leading international and interdisciplinary forum for disseminating the latest research in the field of Semantic Web for the study of pre-modern scientific texts and of the history of the ideas and their transmission.

Classicists and historians are interested in developing textual databases, in order to gather and explore large amounts of primary source materials. For a long time, they mainly focused on text digitization and markup. They only recently decided to try to explore the possibility of transferring some analytical processes they previously thought incompatible with automation to knowledge engineering systems, thus taking advantage of the growing set of tools and techniques based on the languages and standards of the semantic Web, such as linked data, ontologies, and automated reasoning. On the other hand, Semantic Web researchers are willing to take up more ambitious challenges than those arising in the native context of the Web in terms of anthropological complexity, addressing meta-semantic problems of flexible, pluralist or evolutionary ontologies, sources heterogeneity, hermeneutic and rhetoric dimensions. Thus the opportunity for a fruitful encounter of knowledge engineers with computer-savvy historians and classicists has come. This encounter may be inscribed within the more general context of digital humanities, a research area at the intersection of computing and the humanities disciplines which is gaining an ever-increasing momentum and where the Linked Open Data is playing an increasingly prominent role. The purpose of the workshop is to provide a forum for discussion about the methodological approaches to the specificity of annotating “scientific” texts (in the wide sense of the term, including disciplines such as history, architecture, or rhetoric), and to support a collaborative reflection, on possible guidelines or specific models for building historical ontologies. The iconographic data, which are also relevant in history of science and arise similar problematic could be addressed as well and offer suggestive insights for a global methodology for diverse media. A key goal of the workshop, focusing on research issues related to pre-modern scientific texts, is to emphasize, through precise projects and up-to-date investigation in digital humanities, the benefit of a multidisciplinary research to create and operate on relevantly structured data. One of the main interests of the very topic of pre-modern historical data management lies in *historical semantics*, and the opportunity to jointly consider how to identify and express lexical, theoretical and material evolutions. Dealing with historical texts, a major problem is indeed to handle the discrepancy of the historical terminology compared to the modern one, and, in the case of massive, diachronic data, to take into account the contextual and theoretical meaning of words and sentences and their semantics. Papers on ancient and medieval biological science and zoology are particularly welcome.

Topics covered by the workshop include the following:

- Ontologies and vocabularies in Ancient Science
- Semantic annotation of ancient and medieval scientific texts
- Information/knowledge extraction from archaeological objects and texts
- Semantic integration of heterogeneous and contradicting knowledge
- Representation of the historical dimension of Scientific Knowledge
- Impact of Semantic Web technologies on Digital Humanities
- Knowledge Engineering for ancient zoological science and literature
- Social Web, collaborative systems, tagging, and user feedback

Paper Submission:

We invite short position papers (max. 6 pages) and regular research papers (max. 12 pages) describing innovative ideas covering the topics of the workshop. Submissions must be written in English and follow the LNCS guidelines. Each submission will be reviewed by at least 3 Program Committee members. Papers must be submitted via EasyChair. Accepted papers will be published in the CEUR workshop proceedings series.

Schedule

- Due date for full workshop papers submission: March 25, 2015
- Notification of paper acceptance to authors: April 15th, 2015
- Camera-ready of accepted papers: April 25th, 2015
- Workshop: June 1st, 2015