Editorial Introduction to the Fifth Issue

Tobias Blanke, Laurent Romary

HAL Id: hal-00922223

https://hal.inria.fr/hal-00922223

Submitted on 25 Dec 2013

Tobias Blanke and Laurent Romary

Editorial Introduction to the Fifth Issue

Tobias Blanke and Laurent Romary

Editorial Introduction to the Fifth Issue

1   In recent years, European governments and funders, universities and academic societies have increasingly discovered the digital humanities as a new and exciting field that promises new discoveries in humanities research. The funded projects are, however, often ad hoc experiments and stand in isolation from other national and international work. What is lacking is an infrastructure to leverage these pioneering projects into systematic investigations, with methods and technical environments that can be taken up by others. The editors of this special issue are both directors of Digital Research Infrastructure for the Arts and Humanities (DARIAH), which aims to set up a virtual bridge between the many digital arts and humanities projects across Europe. DARIAH is being developed with an understanding that infrastructure does not need to imply a seemingly generic and neutral set of technologies, but should be based upon communities.

2   DARIAH can learn from the successes and failures of such communities. The Text Encoding Initiative (TEI) has been one of the successes. It has emerged as the de facto standard for online critical scholarly editions as well as a format that promotes interoperability and exchange. Furthermore, the TEI as a community brings people together in a concentrated dialogue about particular problems and challenges they face and projects they can build together, such as an infrastructure to support TEI scholarship, comprising workflows, support for collaborative encoding, and distributed annotation. The TEI is therefore the perfect point of orientation for DARIAH to develop long-term outlook European communities not just about scholarly editions but around people, data, and tools. The papers in this special issue very much reflect this community spirit and demonstrate how new approaches and services can emerge through collaboration.

3   The first paper, by Timothy Allen, Clovis Gladstone and Richard Whaling ("PhiloLogic4: An Abstract TEI Query System"), presents the method used by PhiloLogic, a reference search engine for encoded textual documents, to adapt to the great variety of syntax and semantics found in TEI documents. Based upon an internal abstract model for representing structural levels within XML documents, it provides an easy-to-customize environment for users to extend the scope and complexity of document handling.

4   A more TEI-centric approach is provided by the TAPAS project, presented by Julia Flanders and Scott Hamlin ("TAPAS: Building a TEI Publishing and Repository Service"). The idea is to offer a generic repository service that is based on a core number of generic publishing scenarios identified within the TEI user community. From an infrastructural point of view, TAPAS wants to meet the needs of small projects that often do not have the means to deploy their own repository environment.

5   An additional level of complexity is offered in the next paper, "Islandora and TEI: Current and Emerging Applications/Approaches," where Kirsta Stapelfeldt and Donald Moses show how a Drupal- and Fedora-based digital object management system can be adapted to deal with TEI documents and how a whole workflow, from digitization to complex inline annotations, can be created within a coherent architecture. The authors reflect on one of the key issues in processing TEI documents: managing precise encodings with the specifics of OCR or NLP software (in this case, GATE).

6   Stepping back from an analysis of particular software, Quinn Dombrowski and Seth Denbo ("TEI and Project Bamboo") analyze the impact that the widely deployed TEI Guidelines and the corresponding user community have had on the development of Project Bamboo. Conceived as a digital infrastructure project for the humanities, Bamboo addressed various forms of services, from technical components to generic guidelines for data curation. By reviewing several experiments and activities carried out within the initial phase of the Bamboo project, the authors demonstrate how the TEI Guidelines have been seminal in understanding interoperability issues and therefore will continue to serve as an essential basis for future text-based infrastructural initiatives in the humanities.

7   In the next paper, Mark Hedges, together with several colleagues representing the TextGrid, TEXTvre, and DARIAH initiatives, discusses how infrastructures for textual scholarship can be made sustainable and how the TEI, itself seen as an infrastructure, can contribute to this idea. TextGrid was based on the TEI Guidelines from the start, followed by TEXTvre; both from the beginning did not follow the model of generic repository infrastructures but emphasized identifying the community characteristics of dedicated text-based virtual research environments. As examples of how to work together on tools for editing, annotation, etc., these projects have inspired the European DARIAH infrastructure to pay particular attention to how the TEI can offer services to the wider community of digitally based research in the humanities.

8   *TEI as an infrastructure* is indeed the best way to summarize the final paper of this special issue ("The evolution of the Text Encoding Initiative: from research project to research infrastructure"), where Lou Burnard provides a general overview of the historical, editorial, and technical developments of the TEI Guidelines over the last 25 years. Lou Burnard shows in particular how the early customization principles included in the TEI Guidelines, followed by the invention of the ODD specification language, helped popularize data modeling in support of one's research objectives in the humanities.

Cite this article

Electronic reference

Authors

Tobias Blanke
Tobias Blanke is the director of the MA in Digital Asset and Media Management. His academic background is in philosophy and computer science. Tobias Blanke has played a leading role on several multi-disciplinary research projects, mainly concerned with digital libraries and archives, as well as infrastructures for research, particularly in the arts and humanities. Recently completed research includes work on open-source optical character recognition; open linked data and semantically enriched, intelligent digital content; scholarly primitives; and document mining and information extraction for research. Tobias Blanke also continues earlier research on new logic-based approaches to information science. He is on several international committees. Most notably, he is one of the transition directors of DARIAH. He also leads the joint research work for EHRI, a pan-European consortium to build a European Holocaust Research Infrastructure.

Laurent Romary
Laurent Romary is *Directeur de Recherche* for INRIA (France) and guest scientist at Humboldt University (Berlin, Germany). He carries out research on the modeling of semi-structured documents, with a specific emphasis on texts and linguistic resources. He received a PhD degree in computational linguistics in 1989 and his *Habilitation* in 1999. He launched and directed the *Langue et Dialogue* team at Loria (Nancy, France) and participated in several national and international projects related to the representation and dissemination of language resources and to man-machine interaction, coordinating the MLIS/DHYDRO, IST/MIAMM, and eContent/Lirics projects. He has been the editor of ISO standard 16642 (TMF – Terminological Markup Framework) and is the chairman of ISO committee TC 37/SC 4 on Language Resource Management, as well as member (2001–2007) then chair (2008–2011) of the TEI Council. In the recent years, he led the Scientific Information directorate at CNRS (2005–2006) and established the Max-Planck Digital Library (Sept. 2006–Dec. 2008). He currently contributes to the establishment and coordination of the DARIAH infrastructure in Europe as a transitional director.

Copyright