VogonWeb and Quadriga - Creating and Showcasing Quadruple Networks Online

Erick Peirson
Arizona State University
bpeirson@asu.edu

Julia Damerow
Arizona State University
jdamerow@asu.edu

Manfred Laubichler
Arizona State University, Santa Fe Institute, Marine Biological Laboratory, Woods Hole; KLI,
Klosterneuburg Austria
Manfred.Laubichler@asu.edu

Research in the humanities is based on close reading and analysis of sources. As Kirschenbaum observes "[w]hatever else they might do, humanities scholars read" (Kirschenbaum 2007). To support this activity, software tools are being developed to facilitate what Moratti (2000) called "distant reading" (Moretti 2000). These include applications that allow collaborative annotation of documents (see for example MIT HyperStudio 2014), and large-scale analysis of texts (e.g. using topic modeling, as described by Blei 2012). Moreover, there is a growing body of tools and practices for sharing existing datasets, such as publication metadata or document annotations, using RDF and Linked Data technologies. A major challenge in adopting those technologies, however, is the difficulty of expressing provenance and contextual information (for discussions and proposed solutions of this issue see for example Eckert 2013 or Oeckelon et al. 2013).

We demonstrate a suite of tools under development at ASU for collaboratively annotating digitized texts with historical networks. At the center of the tool suite are two applications: "VogonWeb," a web application for text annotation, and "Quadriga," a platform for storing, sharing, and analyzing network annotations. Both applications are based on the "Quadruple" model. Quadruples are historical and conceptual relationships in the form of nested contextualized triples that allow capturing temporal and provenance information and link individual network components to specific evidence in digitized texts. We show how the Quadruple model and attendant technologies begin to address problems of provenance and contextuality in large collaboratively-produced digital datasets.

Quadruples can be transformed into "simple" triples to be discoverable by Linked Open Data technologies. Transformed Quadruples can be understood as what Barbera called "views" onto the underlying data (Barbera 2013). However, the creation of such views in many cases leads to a loss of information in regards to contextual information. We will demonstrate the benefits of Quadruples but also this issue of information loss.

VogonWeb is the successor of a desktop application called "Vogon" (also developed at ASU) that can be used to annotate texts with Quadruples. It abstracts the process of Quadruple creation by providing an easy to use web interface that avoids technical jargon and graph theory terminology. Quadruple networks can be submitted to Quadriga, which provides functionality to manage quadruple annotation projects and allows showcasing networks datasets. Quadriga serves as a publishing and sharing platform for quadruple networks generated with VogonWeb.

While both applications are still in their beta stage, a few groups have started to use them in their research projects. We will showcase research results from those projects, and discuss plans for future development of the platform.

References

Barbera, M. (2013) 'Linked (open) data at web scale: research, social and engineering challenges in the digital humanities', *JLIS.it,* vol. 4, no. 1, pp. 91-101.

Blei, D. M. (2012) 'Topic Modeling and Digital Humanities', *Journal of Digital Humanities,* vol. 2, no. 1. Available from: http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/ [9 February 2016].

Eckert, K. (2013) 'Provenance and Annotations for Linked Data', in *DC-2013 – The Lisbon Proceedings*, Lisbon, Portugal, 2-6 September 2013. Available from: http://dcpapers.dublincore.org/pubs/article/view/3669 [9 February 2015].

Kirschenbaum, M. G. (2007) 'The Remaking of Reading: Data Mining and Digital Humanities', *NGDM 07*, National Science Foundation. 12 October 2007. Available from: http://www.csee.umbc.edu/~hillol/NGDM07/abstracts/talks/MKirschenbaum.pdf [25 January 2016].

MIT HyperStudio (2014) 'Annotation Studio'. Available from: http://www.annotationstudio.org/ [25 January 2016].

Moretti, F. (2000) 'Conjectures on World Literature' *New Left Review*, vol. 1, pp. 54-68.

Ockeloen, N., Fokkens, A., Braake, S. T., Vossen, P., De Boer, V., Schreiber, G., and Legêne, S. (2013) 'BiographyNet: Managing provenance at multiple levels and from different perspectives', in *Proceedings of the 3rd International Workshop on Linked Science 2013*, pp. 59-71.