Swimming in (Linked) Data: Using the Semantic Web to make sense of Spatial Humanities

Rebecca Kahn[*1], Piraye Hacigüzeller[*2,3], and Patricia Murrieta-Flores[4]

[1]Alexander von Humboldt Institute for Internet and Society (HIIG) – Französische Straße 9 10117 Berlin, Germany
[2]University of Oxford – United Kingdom
[3]Univeristy of Leuven – Belgium
[4]Lancaster University – Bailrigg, Lancaster. UK LA1 4YW, United Kingdom

Abstract

The availability of born digital data as well as digitised collections, is changing the way we study and understand the humanities. This amount of information has even greater potential for research when semantic links can be established, and relationships between entities highlighted. The work of several projects presented in the workshop have shown that connecting historical data according to their common reference to places (expressed via URIs stored in gazetteers) is a particularly powerful approach to investigating digital textual and image-based sources. The power of linked data is that it enables search and query beyond one's own projects, and allows leverage of the quantity of relationships between semantic connections that can be made and discovered.

This workshop is designed to show that working with semantic data searches is not meant to render specific results, but rather to highlight how, just as humanities scholarship is about interpretation, association and occasional serendipity, digital humanities scholarship can be used to discover distributions, anomalies, or patterns, rather than specific answers.

This workshop will introduce a range of tools which, when used as part of a pipeline, can enable the extraction, linking and visualisation of semantic data found in historical sources. These will include tools for creating annotations (Recogito) conceptualising linked open data (a TripleStore) and sharing linked data (Wikidata and Peripleo).
Recogito an award winning tool developed by Pelagios that enables annotation of place references in text, images and data through a user-friendly online platform. The principal function of Recogito is the ability to allow non-experts to produce semantic annotations, while at the same time allowing the user to export the data produced as valid RDF, XML, GeoJSON and TEI formats. Participants will be able to explore Recogito's newer features, such as collaborative annotation and controlled vocabularies, as they upload and annotate text and images and then download their annotations in the available data formats. By exploring sources and discussing the related challenges of using a wide range of sources, participants will be able to discuss how Recogito might be used to support their own research. Attendees will also be invited to explore Peripleo, a map-based visualisation for exploring Linked Open Data relationships. Working with over 8 million objects, across almost 2 million years, users will be able to search across a linked data ecosystem for places, documents,

---

[*]Speaker

objects and keywords.

Wikidata is increasingly being used by humanities and GLAM researchers as both a source of linked open data, and as a mechanism for connecting other collections of data (in particular GLAM sources) to each other. This workshop will give a brief introduction to how researchers can contribute their data to Wikidata, thus adding to the web of Linked Open Data, and also use it as an interconnection format, to create hybrid data sources.

We would recommend running these explorations across 2 workshop sessions, in order to give attendees enough time to get to grips with the tools. While there will be short introductions and time for discussion, most of the time will be used for hands-on, practical experimentation.

This workshop would also serve to form part of the stated activities of Pelagios' 2018 DARIAH Theme grants agreement, which aims to establish the use of Linked Data within scholarly practice by establishing its use within the DARIAH ecosystem of tools and services, and to bring DARIAH community members into the broader LOD community by formalising existing, informal collaborations with DARIAH Working Groups.