

An OAC-Compliant Toolbox Enabling Consistent Annotation Functionality Across Repository User Interfaces

Shannon Bradshaw¹ and Martin Foys²
Drew University
{sbradsha, mfoys}@drew.edu

Background

In 2009-2010, funded by an NEH Digital Humanities Startup Grant, the Digital Mappaemundi Project (DM) developed a suite of tools to permit region-of-interest selection and annotation in medieval maps and transcriptions of geographic texts. During this time, DM was invited to participate in a set of use-case studies on interoperability with the Parker on the Web (PW) project at Stanford. The DM-PW use-case exposed our project to the annotation needs of resource repositories more broadly. Sparked by this use case, our thinking on the tools we had developed for DM evolved beyond maps. In June 2010, we refocused our project to target the development of a general-purpose suite of annotation tools for images. We are now collaborating in the Mellon-funded DMSTech project at Stanford which is focused on developing a number of repository interoperability standards and technologies.

Following 18 months of research and development, we have identified a number of functional requirements for annotation user interfaces. It is our hope that through the OAC workshop we might refine this model through an exploration of the case studies brought by others. Likewise, we hope to contribute our own understanding of a wide range of annotation needs.

Annotation System Overview

Our annotation system consists of two major components:

1. A broadly-usable JavaScript library (toolbox) for integrating annotation functionality with web-accessible user interfaces to image content anywhere on the Web, though with an initial focus on digital manuscript repositories.
2. A repository for annotation data with a rich web services interface. All annotations are exposed as first class web resources with stable URIs. Search capabilities enable clients to filter by keyword, annotator, and other features and will enable filtering by creation date (including before/after). The web service also provides “search suggest” functionality so that clients may query for keywords or tags based on a small number of characters (e.g., “ba” returns “babylon”, “babel”, etc). Clients (i.e. interface tools, other repositories) may store annotations through the web service, using URIs to identify resources and annotations.

We are implementing this technology so that either major component of the system is usable independent of the other. For example, a site could choose to use only the JavaScript library. Developers may easily swap the use of one backend web service for another. The user interface library will also be modular with regard to individual annotation tools.

¹Department of Mathematics and Computer Science

²Department of English

Through research into the annotation needs of users in many areas of study, we have identified the following functional requirements for a broadly usable annotation system.

- OAC as the data exchange language in which the web service makes annotation data available to clients and in which the user interface library communicates annotation events to a backend system.
- A self-contained windowing environment that enables users to interact with multiple resources and tools open at the same time.
- Users may define regions of interest within images using arbitrary points, polylines, and polygons. Points are necessary to “pin” descriptions, etc. to specific locations within images. Polylines are necessary in order to string together multiple regions of interest in an image in a particular order. Polygons enable users to define arbitrary regions of interest.
- Entire manuscripts, folia, or regions of interest may serve as targets of annotations.
- The system must provide a customizable schema for text used as the body of annotations. We are currently experimenting with an interface that enables users to define their own schemas as needed using a Twitter-like tagging mechanism using hash (#) characters. Using this mechanism, scholars can create field name/value pairs following any flat schema.
- Annotations may link together any combination of targets (including other annotations) as part of a single annotation.
- The ability for annotation creators to constrain the presentation of annotations. (e.g., define the viewing order for a series of regions of interest for the purpose of telling a specific story).
- Users may create annotations that span multiple repositories and other sites.
- Access control at the user level for resources and annotations.
- Faceted search on user (annotation creator), keyword, tag, etc.

Relevance to Workshop

If a reusable set of user interface annotation tools such as those we envision can be developed, individual sites will be spared the burden and cost of developing annotation tools of their own. Thus there is incentive for their adoption. Furthermore, a consistent set of tools across repositories means users need learn to use annotation tools only once. They will therefore be more likely to make regular use of annotation functionality, creating more annotations that enrich content and improve search within repositories. Since the reusable annotation tools we envision use OAC as a data exchange language, sites that employ them will adopt OAC with little or no additional effort, thereby enabling exchange of annotation data and other forms of interoperability between sites more easily.

Challenges

The challenges for the Annotation Toolbox project are as follows:

- Ensuring our interface library meets the annotation needs of users broadly.
- Ensuring a consistent interpretation and presentation of OAC data in user interfaces across sites.
- Ensuring access to annotation data is protected as defined by users and resource providers (e.g., credentials issues)

- Sustainability of annotation data over time.
- Technical issues involved in sharing annotation data among sites including: consistency for redundant data stores, notification of new and updated annotations for all interested parties, notifications to users of one site regarding relevant resources in other sites.

Attendee

Shannon Bradshaw, Associate Professor of Computer Science, Drew University. Co-Director and Technical Lead for the Digital Mappaemundi and Annotation Toolbox projects.

Projects

Digital Mappaemundi Project: Funded in 2009/2010 under an NEH Digital Humanities Startup Grant. The project is currently funded through a generous gift by a private donor. Finally, we await word on an invited proposal for additional funding.