OPEN ANNOTATION COLLABORATION (OAC):
ENABLING INTEROPERABLE ANNOTATION OF SCHOLARLY DIGITAL RESOURCES
CNI Spring 2009 Task Force Meeting -- Project Briefing

Annotating is a pervasive element of scholarly practice for both the humanist and the scientist. It is a method by which scholars organize existing knowledge and facilitate the creation and sharing of new knowledge. It is used by individual scholars when reading as an aid to memory, to add commentary, and to classify. It can facilitate shared editing, scholarly collaboration, and pedagogy. Over time annotations can have scholarly value in their own right. Yet scholars remain dissatisfied with the options available for annotating digital resources. Scholars wanting to annotate have to learn different annotation clients for different content repositories, have no easy way to integrate annotations made on different systems or created by colleagues using other tools, and are often limited to simplistic and constrained models of annotation. The importance of annotating as a scholarly practice coupled with the real-world limitations of existing practices and tools supporting annotation of digital content has had a retarding effect on the growth of digital scholarship and the level of digital resource use by scholars. While progress has been made on individual tools and in the context of isolated content repositories, lack of a shared data model of annotation and a consensus on an interoperability standard for the exchange of annotation records has undercut the impact of such efforts.

Annotations in the Web Architecture

In order to guarantee the optimal integration of an interoperable annotation environment with both the Web and the Data Web, any annotation data model and interoperability specification must be based on the primitives of the Architecture of the World Wide Web (Resource, URI, Representation), and principles from the Semantic Web and Linked Data efforts. Where appropriate, existing standards (e.g. W3C, IETF, ISO, OASIS, OAI) should be leveraged.
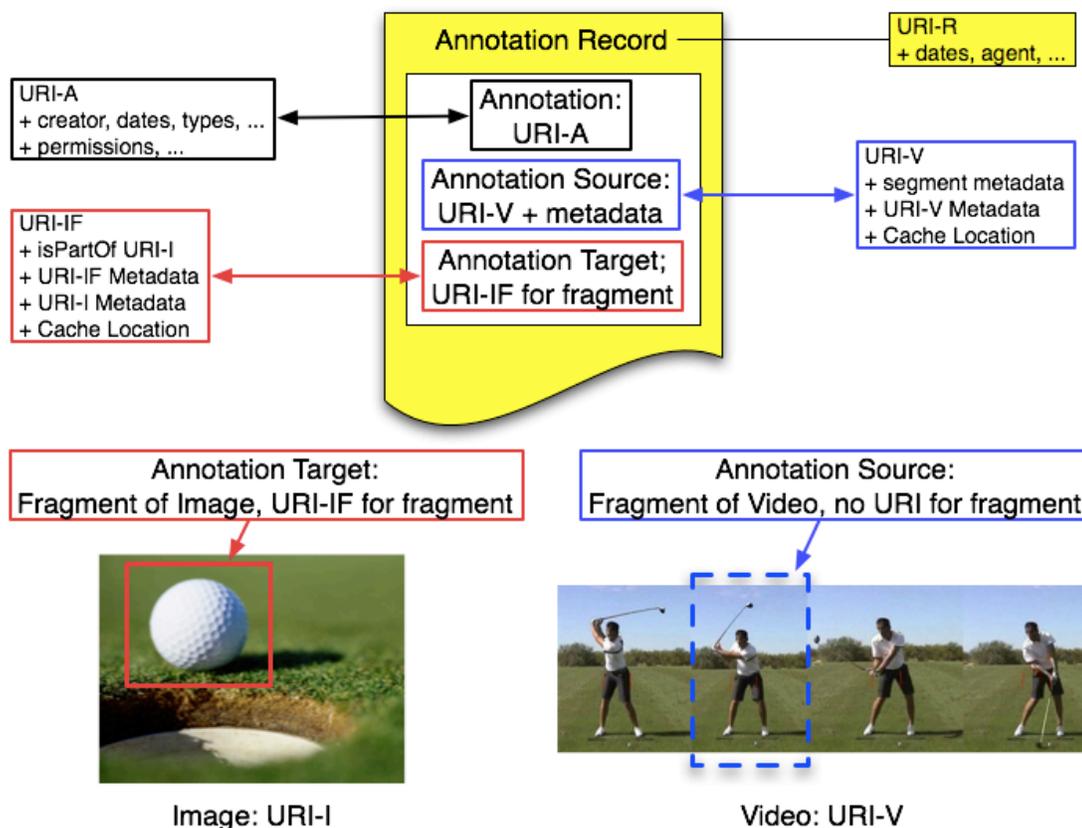

Figure 1: An initial perspective on a Resource-centric annotation model

This Web and Resource centric perspective on annotation interoperability yields the following initial view of a possible Annotation Data Model (Figure 1):

- *An Annotation is considered an autonomous Resource that groups a Source Resource and one or more Target Resources by means of an annotation relationship.*

- *The author of the Annotation, of the Source and of the Target(s) may be different. The Source may exist before the annotation relationship is created, or may be created at the same time as the relationship.*

- *Resources of any media type can be Sources or Targets. The Annotation is a conceptual Resource and has no media type.*

- *Annotations, as Resources with URIs, can be annotated with further Annotations.*

Goals & objectives of the project

The overarching goals of this project (consisting of multiple phases) are:

- *To facilitate the emergence of a Web and Resource-centric interoperable annotation environment that allows leveraging annotations across the boundaries of annotation clients, annotation servers, and content collections. To this end, interoperability specifications will be devised.*

- *To demonstrate through implementations an interoperable annotation environment enabled by the interoperability specifications in settings characterized by a variety of annotation client/server environments, content collections, and scholarly use cases.*

- *To seed widespread adoption by deploying robust, production-quality applications conformant with the interoperable annotation environment in ubiquitous and specialized services, tools, and content used by scholars -- e.g.: Zotero, AXE, LORE, Co-Annotea, Pliny; JSTOR, AustLit, MONK.*

Phase 1 of this project, May 2009 - July 2010, will focus on: an analysis of existing annotation models, systems, and architectures and scholarly practices regarding annotations; the creation and public release of an alpha-level annotation interoperability specification consisting of an Annotation Data Model and Read Annotation Interfaces; integration of Ajax XML Encoder (AXE) code libraries for annotation into Zotero, providing Zotero with a new, embedded scholarly annotation tool; and a minimal reference implementation, demonstrating proof-of-concept feasibility and laying the foundation for subsequent work. Phase 1 objectives reinforce each other. For instance, the AXE-Zotero integration will inform development of data model and specification, and vice versa. As a means to help ensure concreteness and applicability of data model and interoperability specification,, we have preliminarily identified seven potential annotation use case scenarios that once further elaborated will help guide our work. These scenarios will be refined, augmented and periodically revisited over the course of phase 1 of this project. An application to fund phase 1 of this project has been made to the Andrew W. Mellon Foundation and is currently pending.

The Collaboration

- **Herbert Van de Sompel (co-PI)** -- *Research Library, Los Alamos National Laboratory*
- **Tim Cole (co-PI), Tom Habing, Carole Palmer, Allen Renear** -- *University Library & Center for Informatics Research in Science and Scholarship, University of Illinois at Urbana-Champaign*
- **Neil Fraistat (co-PI), Douglas Reside** -- *Maryland Institute for Technology in the Humanities, University of Maryland*
- **Jane Hunter (co-PI), Anna Gerber** -- *eResearch Laboratory, School of Information Technology & Electrical Engineering, University of Queensland*
- **Daniel Cohen, Sean Takats** -- *Center for History and New Media, George Mason University*
- **Robert Sanderson** -- *Department of Computer Science, University of Liverpool*
- **John Burns, Clare Llewellyn** -- *JSTOR*