Consumers as Creators: Understanding the annotation needs of the scientific community through the domain of botany

Statement of need

The Missouri Botanical Garden (MOBOT), with partners at Saint Louis University (SLU), propose a $49,926 Planning Grant to analyze Web annotation needs of the scientific community and develop a prototype of how those needs may be met within a digital library platform. The *New Media Consortium Horizon report: 2015 Library Edition*, identifies the Semantic Web and Linked Data as key technologies that will significantly impact academic and research libraries in the next two to three years. Libraries increasingly understand it is insufficient to simply provide online collections access; users want integrated Semantic Web tools among library site services. Annotating (i.e., making comments on a resource) is an important part of the vision for the Semantic Web. While some existing annotation tools can be re-purposed by libraries, most fall within a proprietary environment for particular groups, are not well-suited to general audiences' needs, and do not allow easy sharing and discovery of annotations across the Web. This conflicts with Semantic Web principles and limits access and value. With the recent release of recommendations from the *World Wide Web Consortium (W3C) Web Annotation Working Group* regarding a standardized data model, vocabulary, and protocol, opportunity exists to create a single, common, Resource Description Framework (RDF) based specification for annotating digital resources.

This project will advance the annotation needs of the scientific community in its broadest sense, with the goal of developing methods that are expandable to other communities. Our prototype will focus on a particular group of scientists, namely systematic botanists, that has advanced its creation of vocabularies and definitions of data model elements that leverage adoption of annotations. These examples can be applied to demonstrate their potential impact in scientific research, humanities and other scholarly arenas. The staff of the Center for Biodiversity Informatics (CBI) at MOBOT is well-acquainted with this researcher community, with a decade-long history of developing content repositories to access biodiversity literature, plant specimen data, and living collections information. Through close collaboration with botanists CBI has identified valuable use cases for developing in-depth user assessments of annotation needs. With this grant, CBI will perform a landscape review of existing tools and test the applicability of one of those tools within a platform called *Botanicus* (http://www.botanicus.org/) - the freely accessible portal to historic botanical literature from the Peter H. Raven Library at MOBOT.

History of annotation in print and online publications

Scientists have annotated books and journals for centuries. Hand-written annotations have served many purposes since the earliest printing of books, including to indicate ownership, respond to the text, or record mottoes and proverbs. A few scientists' preserved, personal libraries provide a wealth of information about the influence of their contemporaries on themselves, and their hand-written marginal notes offer personal reflections on the theories these books contain.

Web annotations constitute a re-creation and extension of these age-old functionalities as a new, interactive mode built on and linked through Web technology. The move from print to online publications has made it easier to create and share annotations. Online annotations were possible as far back as 1993 with an early version of the *Mosaic* web browser, but during that era, Web 1.0 users were mostly viewed as mere content consumers. Few platforms supported content creation until the move into a Web 2.0 environment. Consumers became creators, and the more recent spread of social media throughout the Web has led to a new understanding of its value as a place to connect, build and share data. As we reach the era of Web 3.0, the original vision of the Semantic Web is being realized although little Web content currently exploits all possibilities. Libraries rarely incorporate annotation functionality into their interfaces with a few exceptions (e.g. *Trove* and *Europeana*).

Relevant annotation tools, projects and standards

This project will focus on open source annotation tools that follow established standards such as the Web Annotation Data Model and/or the International Image Interoperability Framework (IIIF). At least four tools

meet these criteria: RERUM, Hypothes.is, digilib, and Annotorious. RERUM will be used as a prototype for testing with data found in the *Botanicus* platform. Its ability to create and store Web annotations as well as store IIIF documents gives it an advantage over other annotation tools. RERUM developers from the Walter J. Ong, S.J. Center for Digital Humanities at SLU will participate in this project by providing technical support on its setup. CBI will also seek input from outside experts and organizations such as the Annotating All Knowledge (AAK) Coalition - a group of key scholarly publishers, libraries, and technologists.

Relevance to curating collections and national impact

Curation is the act of selecting and interpreting content - a role previously limited to content providers. As Web users move from consumers to creators, they actively engage in the curation role and their annotations offer additional access points beyond traditional bibliographic information that libraries provide. User-added annotations can lead to a richer dialog and broader context around curated collections than is otherwise had by restricting the curation role to traditional curators, thereby increasing the collections' impact, value and reach.

Annotations can better facilitate discovery and extraction of knowledge from scientific literature. Until recently, the idea of an integrated standardized annotation as "a unit of conversation built into the very fabric of the Web" was a far reaching objective. Today it is becoming a high priority need that libraries must address as part of their services in order to impact the future practice of research. The AAK Coalition posits this in turn will "transform scholarship" by enabling "personal note taking, peer review, copy editing, post publication discussion, journal clubs, classroom uses, automated classification, deep linking, and much more". CBI will build upon current strategic initiatives in the implementation of the Semantic Web to demonstrate the importance of annotation functionality that comments on, characterizes, extends or links entities from different realms in science. We will demonstrate through a prototype in *Botanicus*, how annotations can produce far-reaching impacts across virtual libraries of any type of cultural heritage institution.

Work plan

The following activities will be conducted within one year, beginning May, 2018:

- **Analyze annotation needs** of ten users of a botanical virtual library from five institutions (months 1-3)
- **Develop a feasibility study** for how existing annotation tools could satisfy botanists' needs (months 4-6)
- **Install a prototype** of one of the tools within a digital library platform (months 7-11)
- **Assess outcomes** from this project to determine activities needed in a full-scale Project Plan (month 12)

Performance goals and outcomes

The intended outcome of the proposed project is to illuminate literature annotation needs of scientific and other research communities by honing in on the annotation needs of a well-defined user group in systematic botany. Assessment of the practicality of an existing tool to satisfy the annotation needs of botanical users, including technical, economic, and operational considerations, will inform developers on best practices to integrate an annotation tool within a virtual library. Ultimately, a list of planning activities and partner commitments needed for a robust project proposal will result. MOBOT uses a variety of publishing channels, including traditional print media, conference presentations and emerging social media, to promote its content, services, and activities. These will be used to disseminate project results to the biodiversity and broader scientific communities.

Budget

Total Planning Grant cost is $49,926. Two CBI researchers, Trish Rose-Sandler and William Ulate, will gather and analyze necessary information about scientific annotation needs, handle project management, and implement and test a prototype. Salaries & wages for the year are $36,900, plus $8,487 (23% of salary) included as fringe benefits, plus indirect costs of $4,539 (10%). MOBOT will provide hardware, software and server support staff in-kind and SLU will provide technical consultancy support in-kind.