NeoNote: User Centered Design Suggestions for a Global Shared Scholarly Annotation System

Author: Brad Hemminger, UNC-CH School of Information and Library Science

Short Abstract for Information Seeking Support Systems NSF Workshop 2008

Introduction

Significant changes are occurring to scholarly communications due to technological innovations, primarily the advent of computers and the internet. Compared to thirty years ago, scholars use computers both to do their work and write their papers. They publish in journals still, but the articles are available and accessed more in digital format than in print format. Scholars increasing share their work with others by putting their work on their websites, in institutional repositories, or emailing to others. They use collaborative tools for writing papers, or creating shared content on wikis. Scholars are beginning to compile large digital collections of research papers instead of print collections. They are beginning to annotate these papers electronically, and to share these annotations. They save citation information digitally in citation managers and automatically incorporate this information when writing their papers. They start their searches more often with search engines that traditional library catalog resources, and they spend most of their time searching for and looking for information in web browsers.

There has been a proliferation of tools developed to help support scholars in performing this myriad of activities, but in most all cases, the tools are designed for only a specific task or environment. As a result, scholars are forced to utilize many different incompatible tools to perform their scholarly work and communication. This paper looks at the problem from the scholar's perspective and proposes a user interface well suited to scholarly work practices. It also proposes a paradigm for how scholarly annotations could be captured, stored, searched and re-used on both a global scale, and at the level of an individual research laboratory or group. It is our hope that the paradigms proposed in this paper will provoke discussions in the digital library community about shared global designs for digital library content, including annotations.

As part of a user centered design process to develop a global shared annotation system, information from surveys [Hemminger ASIST 2007], our interviews with scientists at UNC, and feedback from users of scholarly annotation tools were analyzed to generate a list of features scholarly researchers required. Currently, there are no systems designed to provide a single comprehensive system to address all user's needs. There are, however, many individual applications that provide excellent support for one or more features needed by the scholarly researcher. This paper describes the desired features of a single comprehensive system, gives examples of current tools that support these features, and then describes the architecture of our resulting design (NeoNote, [REF URL]). Overall, we believe the most effective way such an interface could be delivered, would be as part of a next generation interface to the web, closer to the original hypertext systems first proposed like Xanadu [http://en.wikipedia.org/wiki/Project_Xanadu]. The important change would be that annotation and sharing would be integrated into the interface and user experience, in the same way that browsing is in today's web browser interface to the web. Alternatively, a smaller incremental step would be to have today's browser incorporate such capabilities through plugins, such as are common in Firefox. To test these ideas, we have taken the latter course, and implemented these features as an integrated set of tools through a plugin to the Firefox browser. Our implementation, NeoNote, is focused on global shared annotations, and is part of the more general NeoRef [URL] scholarly communications project at UNC. A video introducing the NeoNote system which visually

shows how the paradigm works is available on YouTube [http://www.youtube.com/watch?v=PUn09--HRaw] and at a higher resolution on the author's website [http://www.ils.unc.edu/bmh/pubs/NeoNote-For-YouTube.wmv]. Additionally, current challenges facing such global scholarly annotation repositories are discussed.

In this paper, for simplicity, discussion and examples about the information (content items) that scholars identify and want to save, will be scholarly journal articles and annotations on them. This is because this paradigm is familiar to most readers, currently better understood by researchers, and there are more examples of applications working with scholarly journal articles. However, this discussion should apply to all types of content items (multimedia (such as audio, pictures, video), data analysis, datasets (such as genetic sequences, observations from astronomy, weather, geography) etc. The proposed (NeoNote) design does support all data types.

I. User Interface Features Required for a Successful Global Shared Annotation System

Selection from Web Page. Searching and selection of content items should be done from within the single interface to the web that the users utilizes (currently this is web browsers). While in the past many literature database searches were performed on specific interfaces like Dialog, ISI (?), nowadays searching occurs of these same databases via web pages, and even more frequently via web-based search engines. As a result, researchers need to be able to add content items identified on a web page to their "personal digital library" with a single click.
Examples: Zotero, Connotea, RefWorks. Add descriptions and links. The search features should be an integrated part of web browser (or whatever future application provides the user access to the web), rather than distinct applications.

Multiple levels of annotation complexity should be supported. Annotations are done for multiple reasons. Some annotations are complex (examples, comment, tags, link) , some are very simple (save item, save item plus highlighted text). All kinds of annotations should be very easy to create and the user should be to specify them directly from the web browser environment.

Simple Item capture with one click. For simple types of capturing an item (e.g. to a citation database, or to highlight text as annotation), the user should be able to do it directly from the web browser page with a single click (to save). Examples: Google Notebook (although first have to select notebook). Fleck (but requires turning on Fleck for page, then saving bookmark).

Simple Item Highlighting. A users should be able to highlight text on a content item, and have this automatically be captured as an annotation on that content item. Multiple highlights of an article should be able to be done successively without requiring any other interactions (for instance to again select a highlight tool). For instance, this models how students highlight textbooks with a highlighter—they pick it up and then continue to highlight until they are finished.

Complex Item Annotation and Capture. In some cases, users will want to capture additional information about an item. They may want to add a text note, tags, links to other items, etc. This should be supported, easily and conveniently, while not complicating the ability to do simple item captures or highlights.

Concept of Current Folder or Notebook. This is critically important as users frequently are working on a particular topic, and they simply want all items they continue to select to all be captured and "stored" in the same "group". The concept of the group is common, and supported as folders in most browser bookmarks, or as a "notebook" in Google Notebooks.
Examples: Google Notebook. Clipping an item places it in the current notebook (which is the most recently used one).

Organizing and Searching. Users expressed two preferences. First for research purposes they tended to prefer to organize materials by placing them in a folder. Second, they wished to be able to tag content items with tag words or phrases. They wanted to search by any of the tags including the folder label, by the full-text of the content item, and potentially by other metadata of the item.
Examples: Social Network Sites de.li.cious, Flickr, etc (Wikipedia entry).

Automatic capture of citation. When the user selects an item for capture, the citation information (all relevant metadata) should be automatically captured into the annotation, and available for future use.
Examples: Zotero, Connotea, RefWorks

Automatic capture of the item, not just the link. While for many web pages, users are content to just capture a link to the original resource, most scholars prefer to maintain their own copy of research articles. At some point in the future when content items are all easily retrieved via universally unique identifiers from long term stable archives, it may no longer be necessary to keep private copies. From our data, though, scholars currently clearly prefer to keep their own copies, just as they have made print or Xerox copies in the past.

Saved content and annotations should be universally accessible. Most users operate from many locations, and many computers. Users want to have their annotations and content items available on th web from any computer at any location.
Google Notebooks, Google Docs, Connotea, social bookmarking tools (many examples, see Wikipedia page http://en.wikipedia.org/wiki/List_of_social_software#Social_bookmarking).

Annotations and Content should be sharable. Scholars want to be able to share their annotations and content items as they see fit. In some cases this means with the world, but in most cases they wish to share it with their research group, department or collaborators.
Examples: Connotea, Social Networking Sites (Wikipedia page, URL form above)

Collaborative Sharing. Scholars wish to be able to save papers to shared digital library, collaboratively edit annotations, revise papers, work on documents. The documents need to be available from any computer at any location, and include provisions for handling multiple write accesses at the same time.
Examples: GoogleDocs, Web based File Systems (WebDAV, commercial systems (SharePoint, etc)).