

Universal Collaborative Annotations with Thin Clients – Supporting User Feedback to the Atlas of Living Australia

Ron Chernich¹, Steve Crawley², Jane Hunter³

eResearch Lab, The University of Queensland, Brisbane, Australia,

¹chernich@itee.uq.edu.au; ²scrawley@itee.uq.edu.au; ³jane@itee.uq.edu.au

INTRODUCTION

Although Web publishing has become ubiquitous in modern research as a means of presenting and sharing information, to date this media has been essentially one-way. To address this limitation, numerous systems have proposed and demonstrated mechanisms by which readers can share mark-up of on-line documents [1]. Although studies have demonstrated the value to collaboration of shared annotations on Web resources [2], there is still no wide-spread, universally-accepted, standardized approach to annotation services. Moreover there are a number of distinct, intrinsic limitations within the World Wide Web's underlying technologies that prevent the wide scale implementation of interoperable, shared annotations. This presentation will examine these limitations and show how our *Danno* research project [3] has managed to overcome many of the inherent technological barriers to wide-spread Web site annotation.

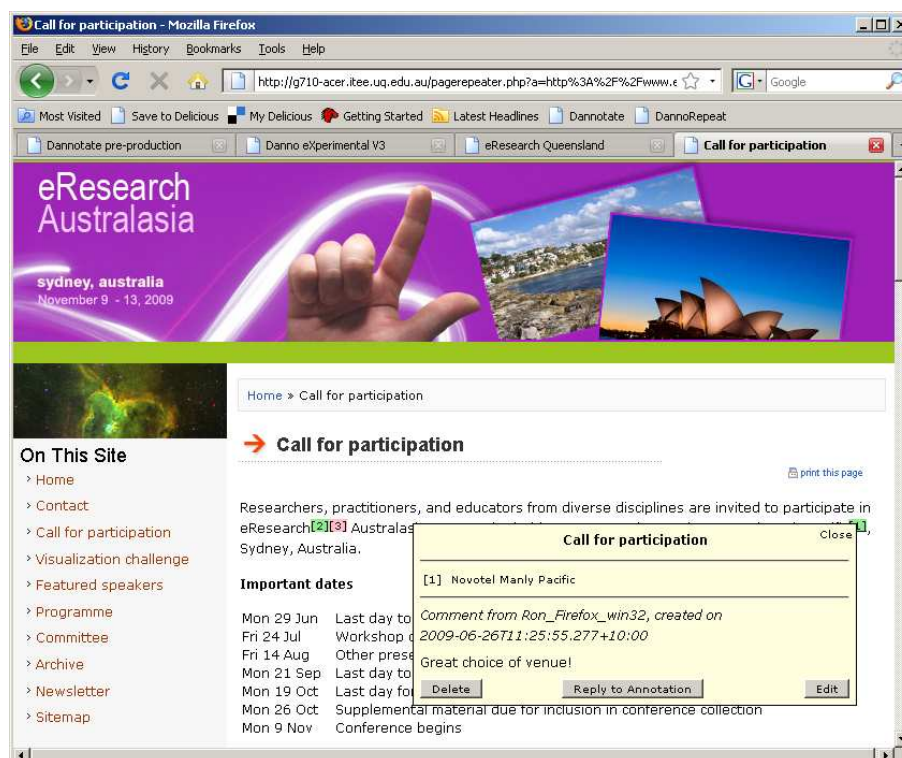


Figure 1: Shared annotations inserted into an arbitrary Web resource

ANNOTATION TYPES

In the context of Web content, annotations can take many forms and formats, depending on the intended audience. For a single researcher, the ability to quickly and easily highlight areas of a document and attach notes, while they are reading it online, can be of great value [4]. This is analogous to reading a paper copy with a highlighter pen in hand. The reader's flow of concentration is not broken and the persistent markings allow the researcher to easily return to sections of interest on later reading. Even though annotations made like this may never be shared, they still provide value.

More usefully, a researcher will highlight a section of page text and enter a comment relating to it. The annotation may be categorized as: Comment, Keywords, Question, See Also, Ranking, Feedback etc. Sharing such annotations within a community permits others to respond with new annotations or replies to existing annotations. The World Wide Web Consortium (W3C) developed a recommended approach via the *Annotea* project [5]. This defines a protocol based on Resource Definition Framework (RDF) that provides a well-structured, extensible, shared annotation schema. Being RDF, the schema is easy extended should a more specialized structure be required. Client applications to create and display *Annotea* annotations such as *Annozilla* [6] have shown the practicality of this text-only approach, but limit their wider adoption by being architected as browser vendor specific plug-in code.

An example of a more advanced annotation service permits the user to select a region of an image or a segment of a multimedia file as the target for an annotation (e.g. comment, question, feedback, tags). Moreover, the annotation itself, may not always be textual, but could be an image, audio or video file. The technology to perform all these types of operations exists and services providing the capability have been in limited use for some time [6]. They require either specially authored Web pages (able to be annotated through embedded Javascript and controls), specific client software installation, or both. These browser-specific approaches and the additional effort they require of annotation authors and viewers, limit the widespread adoption of annotation services. The challenge is in making them more widely available, preferably on *any* Web resource, using *any* common Web Browser as the client.

BARRIERS TO UNIVERSAL ANNOTATIONS

The first barrier to such a universal, shared annotation service using thin-client technology is security. As the Internet has evolved, Web browser designers have had to implement some stringent rules in order to make the desktop Web browser a safe tool through which to transact business. One of the key foundations of this security is the *Same Origin Policy* [7] which prevents the browser from exchanging information with a third party; that is, no communications are permitted to a remote host other than the one which provided the page currently being viewed. However, in order to share annotations, we require a central annotation repository, or set of repositories that can be read and written from the context of any arbitrary Web page. This prime requirement would seem to be inconsistent with established and rigidly enforced browser security. Our presentation will describe the ways in which this restriction can be overcome and the attendant restrictions and risks imposed.

The next hurdle is delivering the client side presentation and functionality to the broadest possible audience, with the least possible restriction and impact on the user's chosen environment. While standards (such as DOM) exist that define how a Web page is to be modeled within a Web browser, these are interpreted with subtle differences by some browsers, and totally ignored by others. The problem is compounded by the number of software versions that are commonly encountered from individual browser vendors and the operating systems a user may be running them under. Our current client code supports the most commonly encountered Web browsers, namely Firefox, Safari, Opera, Google Chrome, and Microsoft Internet Explorer 8. The presentation will give an overview of the problems encountered and the solutions we identified in creating a *universal annotation client*.

THEORY INTO PRACTICE

The Atlas of Living Australia (ALA) project [8] will provide a single biodiversity reference for Australian flora and fauna. It is being built with five year funding under the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS) [9] by the CSIRO, the University of Queensland, and others. A vast amount of Australian biodiversity data already exists. The ALA will not replicate this data, rather it will act as a discovery mechanism allowing amateur and professional researchers, government bodies, and the general public to access and contribute to it. It is known that not all of the data are correct, nor are all the data problems known. The ALA portal will incorporate an annotation service that enables users to query, question, and suggest corrections to entries. The quality of the data will thus be increased through collaboration and harvesting of corrections using OAI-PMH that can then be fed back to the data owners. As the profile of the anticipated user base for the ALA is huge, devising an annotation service that can accommodate the diverse user environments has posed a technological and conceptual challenge. We will highlight some of the problems encountered and demonstrate the solutions developed to this point.

REFERENCES

1. Heck, R.M., Luebke, S.M., Obermark, C.H.: *A Survey of Web Annotation Systems*. Available from: http://www.math.grin.edu/~rebelsky/Blazers/Annotations/Summer1999/Papers/survey_paper.html, accessed 25 June 2009.
2. Hunter, J.: *Collaborative, Semantic Tagging and Annotation Systems*. Chapter for "Annual Review of Information Science and Technology", American Society for Information Science & Technology, Volume 43, 2009.
3. *Danno*. Available from: <https://metadata-net.svn.sourceforge.net/svnroot/metadata-net/>, accessed 20 June, 2009.
4. Hemminger, B.: *NeoNote: Suggestions for a Global Shared Scholarly Annotation Service*. D-Lib Magazine, May/June 2009. Available from: <http://www.dlib.org/dlib/may09/hemminger/05hemminger.html#17>, accessed 29 May 2009.
5. *Annotea*. Available from: <http://annotea.org>, accessed: 26 June 2009.
6. Schroeter, R., Hunter, J., Guerin, J., Kahn, M.I., Henderson, M.: *A Synchronous Multimedia Annotation System for Secure Collaboratories*. Available from: <http://www.itee.uq.edu.au/~eresearch/papers/2006/eScience2006.pdf>, accessed 25 June 2009.
7. Same Origin Policy. Available from: http://en.wikipedia.org/wiki/Same_origin_policy, accessed 26 June 2009.
8. *Atlas of Living Australia*. Available from: <http://www.ala.org.au>, accessed 26 June, 2009.
9. *National Collaborative Infrastructure Strategy*. Available from: <http://ncris.innovation.gov.au>, accessed 26 June 2009.