Annotation evolution: how Web 2.0 technologies are enabling a change in annotation practice

Simone Sacchi
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
501 E. Daniel Street, MC-493, Champaign, IL 61820-6211
sacchi1@illinois.edu

ABSTRACT

Are Web 2.0 tools and technologies changing how and why scholars annotate their research sources? We begin to answer this question by assessing current technology and tools that support new functions for one of the most common scholarly research activity: taking notes. The results suggest a new approach to personalized information retrieval.

Keywords

Annotations, Annotation Practice, Annotation Function, Annotation Purpose, Scholarly Communication, Information Retrieval.

1. INTRODUCTION

Taking notes has always been one of the most common activities performed by scholars, students, and readers during any research-related work. The idea of adding notes to written paper goes back to the Medieval Age when Roman law was taught at the universities. Notes enriched the books with the purpose of marking, selecting and giving specific emphasis to a particular portion of the text [1]. Since then, paper-based annotation practices have changed in the kind of marks applied – from glosses in the margins to sticky notes, from the use of the nib strokes to highlighting with colors – but have not changed significantly in the function they perform. Essentially, the creators' intention has not changed much over the centuries. In trying to answer the question "what is the purpose/function of annotation?" we could say: *drawing attention to a particular portion of the text, and attaching personal context/content to the text itself*.

However, is it still true that annotation functions are basically unchanged when we change the medium from fixed paper to more fluid digital resourses? Moreover, is the possibility of annotating content in a Web 2.0 networked environment changing the nature the annotation practice itself?

This poster outlines the preliminary results of an analysis – from a functional perspective – of some of the most advanced annotation tools that exploit the Web 2.0 "web as a platform" paradigm [10] and suggests that, not only these technological shifts have enhanced the traditional annotation practices, they also allow new and previously unpredicted annotation functions and purposes.

iConference 2011, February 8-11, 2011, Seattle, WA, USA

2. RESEARCH FRAMEWORK

There have been attempts to classify annotation from a functional perspective that provide insight on the users' perspective and the patterns used in the practice itself [7]. However, the first systematic study was by Renear at al. with their Functional Taxonomy of Annotation [12]. Renear's study was intended to support the development of digital annotation able to fulfill at best the functions that were already established on the paper medium. Their analysis produced a framework of six macro-categories of annotation functions – A - *Recording and Scheduling Reading*; B - *Basic Highlighting*; C - *Commentary*; D - *Classification*; E - *Copyediting / Editing / Joint Authoring*; F - *Speech Acts* – but they were aware that "new media, […], new functionalities […] and new market domains will almost certainly yield at least some new functions that we will need to accommodate " [12].

Based on a preliminary review of the literature on system design, information representation, and users' annotation behaviors, this research takes on Renear's work, extending it to the Web 2.0 environment. Five different annotation tools that comply with the "web as a platform" paradigm were analyzed: most of them are developed to extend the functionalities of the browser itself and are implemented via the Firefox Web Browser extension system (Diigo, SparTag), as a Greasemonkey script (ShiftSpace) or as a Bookmarlet (Diigo, SharedCopy); only one (A.nnotate) is a full-featured server-side web application that requires no components other than the browser.. The results outline that Web 2.0 tools are enabling new annotation functions, extending at least Renear's *A-4 Trail* and the *D-2 Subject Classification* categories.

3. OBSERVATIONS

Several observations can be made immediately:

1. The technological shift from paper to the digital medium that allowed the logical modeling [2, 3] and deconstruction of the document unit, supports the reshaping of reading paths through annotation trails;
2. Annotations can be considered themselves digital objects.

These observations had been anticipated by the hypertext community and by Renear et al. in their taxonomy, however:

3. Now all annotation objects can be enriched with descriptive metadata, tagged, shared, retrieved, aggregated, and clustered independently from the target resource;

4. Annotations can themselves be the target of further annotations [6], therefore the annotation network is multi-layer;

5. Web 2.0 environment annotation systems share features with the already established social bookmarking and social tagging systems [4, 5, 14] integrating them as scholarly communication tools.

Moreover, not enough attention has been paid to some implications:

6. Annotation objects form a distinct network where links are established on the basis of shared properties of the annotations.

7. Every annotation object now support multiple trail function

8. Folksonomy over shared annotations can extend the annotation network at a community level

These relations can be – and in fact are – exploited to support new annotation functions.

4. DISCUSSION

Understanding the notes-to-notes relations supported by these new tools suggests extending Renear's taxonomy with at least one new functional category: *Selecting and Clustering*:

1. Scholars take advantage of the annotation metadata – in particular the note-level tagging system – to instantiate relations among annotations;

2. The overlaying annotation graph reflects a connection between the annotated portions of texts;

3. The established relations are exploited to cluster and retrieve selected logical components of the text over the network.

Summarizing, clustering annotations enables the subsequent grouping and retrieval of selected fragments of text, creating new, transversal, and semantically enriched reading-paths across documents. This is a new approach to personalized information retrieval, and can be considered a first level of strategic reading, i.e. move rapidly through the literature to assess and exploit content with as little actual reading as possible [11, 13].

5. FUTURE DIRECTIONS

In the current development trend the implementation of annotation features is not entirely delegated to a dedicated application: the browser itself is exploited as an interactive annotation tool. Unfortunately, web tools are still not adequate to prompt the experience of a full desktop application, and desktop applications are barely adequate to capture the fluidity of annotation forms that we see on paper [9]. More than ten years have passed since Marshall's wondering about the future of digital annotation [7, 8] and technologies like AJAX and HTML5 have successfully pushed the web-interface development in new interactive directions. However, we still face questions about the requirements for a device to be *the* annotation tool – a tool able to supports new annotation functions in an intuitive way familiar from the paper-and-pen environment. We are still far from a true annotation *revolution*, but we are experiencing some non-preexisting possibilities that – we hope – allow us to use the term *evolution*.

6. ACKNOWLEDGMENTS

This study would not have been possible without the support, suggestions, and feedback from Carole L. Palmer and Allen H. Renear. I want also to thank Tim Cole and the Mellon Foundation funded Open Annotation Collaboration (OAC) research group at the University of Illinois for the many inspirations.

7. REFERENCES

[1] Agosti, M., Bonfiglio-Dosio, G., and Ferro, N. 2007. A historical and contemporary study on annotations to derive key features for systems design. *International Journal on Digital Libraries*, 8(1), 1–19.

[2] Bradley, J., and Vetch, P. 2007. Supporting Annotation as a Scholarly Tool. *Literary and Linguistic Computing*. 22(2). 225-241.

[3] Bradley, H. 2009. Suggestions for a Global Shared Scholarly Annotation System. *D-Lib Magazine*, 15(5/6), 1082–9873.

[4] Cattuto, C., Barrat, A., Baldassarri, A., Schehr, G., and Loreto, V. 2009. Collective dynamics of social annotation. *Proceedings of the National Academy of Sciences*, 106(26).

[5] Gazan, R. 2008. Social Annotations in Digital Library Collections. *D-Lib Magazine*, 14(11/12), 1082–9873.

[6] Hunter, J., Cole, T., Sanderson, R., Van de Sompel, H., 2010 The Open Annotation Collaboration: A Data Model to Support Sharing and Interoperability of Scholarly Annotations. Paper presented at *Digital Humanities 2010* (London, UK, July 07 – 10, 2010)

[7] Marshall, C. C. 1997. Annotation: from paper books to the digital library. In *Proceedings of the second ACM international conference on Digital libraries,* 131–140.

[8] Marshall, C.C. 1998. Toward an Ecology of Hypertext Annotation. *Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia* (Hypertext 98), 40-48.

[9] Obendorf, H. 2003. Simplifying annotation support for real-world-settings: a comparative study of active reading. In *Proceedings of the fourteenth ACM conference on Hypertext and hypermedia,* 120–121.

[10] O'Reilly, T., and Battelle, J. 2009. Web Squared: Web 2.0 Five Years On. Paper presented at the Web 2.0 Summit 2010 (San Francisco, CA, October, 20 – 22, 2009).

[11] Palmer, C. L., Teffeau, L. C., and Pirmann, C. M. (2009). Scholarly Information Practices in the Online Environment. Report commissioned by OCLC Research. Published online at: www.oclc.org/programs/publications/reports/2009-02.pdf

[12] Renear, A., DeRose, S., Mylonas, E., and van Dam, A1999. An Outline for a Functional Taxonomy of Annotation. White Paper presented at Microsoft Research (Redmond, WA, 1999) Retrieved from: http://hdl.handle.net/2142/9098

[13] Renear, A. H., and Palmer, C. L. 2009. Strategic reading, ontologies, and the future of scientific publishing. *Science*, *325*(5942), 828-832.

[14] Yang, S. J. H., Chen, I. Y. L., and Su, A. Y. S. 2007. Personalized Annotation Management: A Web 2.0 Social Software for Enhancing Knowledge Sharing in Communities of Practice. In *Seventh IEEE International Conference on Advanced Learning Technologies*, 625–627.