

Draft 25-01-2008 Peter Boot pboot@xs4all.nl

This section will discuss the present state of annotation facilities in the digital edition. I'm interested here in online interactive facilities, not in annotation facilities available with e.g. CD-rom editions such as (De Smedt and Vanhoutte 1999) or (Daly and Young 2002). Neither am I interested in external annotation tools that might be used in conjunction with the digital edition, such as Google Notebook, Diigo or Notefish. What I am looking for is an integrated presentation of text and annotation tools, with the annotations displayed alongside the edited text. It should be possible to annotate not just at the level of the text, but at the level of text constituents – pages, paragraphs or other. Annotations may be private, shared or public. Adding annotations may require registration, perhaps open only to experts, or may be open to every visitor. The annotations may be subjected to a review process. They may consist of free text comments, of a simple tag, or have a more complex structure. They may be meant for the purpose of explanation, discussion, transcription, translation, or for other purposes. I am specifically interested in annotation in scholarly text editions, but it is hard to define these terms ('scholarly', 'text', and 'edition') precisely, and there is no need to do so here.

Wikisource and wikibooks

The first annotation platform that I want to discuss here is that provided by wiki software. The reason for this is not that wiki software should provide a platform that is very suitable for serious annotation, but it is because wiki is what has captured the imagination of many. When I try to explain to non-scholars what annotation is and why it is important, the wiki is their point of reference – 'ah, so what you want is a kind of wiki!' And then I find myself explaining that, yes, I want a wiki, but what I really want is something that can do other things beside.

A wiki is a publication system where multiple people can collaborate in publishing web pages. A simple markup-system, sometimes hidden by a wysiwig editing system, allows non-technical persons to create and modify pages. Typically, wiki pages are edited collaboratively, allowing editors to extend earlier content of a page. The best known wiki is Wikipedia, but wiki systems are used as knowledge management systems in many organisations. In the field of digital humanities, e.g. the TEI, Digital medievalist and Digital classicist communities have their own wiki's.

A second reason why I feel that a discussion of wiki software should open this section is the vitality and energy of the wiki community. Academic use of annotation facilities tends to be experimental and very careful and seems more worried by potential risks than excited about possibilities. The wiki community, predominantly non-academic, tends to act rather than to worry, and may help academics imagine the kind of environment that would help us do our work.

Among the sister projects of Wikipedia are Wikisource and Wikibooks. Wikisource describes itself as 'an online library of free content publications collected and maintained by the community', Wikibooks is 'community for creating a free library of educational textbooks that anyone can edit'. Both Wikisource and Wikibooks can accommodate editions with annotations. In the case of Wikisource, the primary object will be the edited text, to which the annotations are supplementary material. In the case of Wikibooks, the new intellectual content will be the annotation, to which the annotated text can be provided as reference material. The usual case in what Wikibooks names 'annotated books', however, is that the annotated text is

only present as a reference.¹ Such is the case in, e.g., the annotations to *The Complete Peanuts*.² Annotations to a number of chapters of *The Brothers Karamazow* in contrary are displayed next to the texts.³

On wikisource, the focus is on providing editions of existing texts. These editions may be annotated, a mechanism that seems most popular in the German Wikisource site.

Sometimes notes from an existing edition are included, as in the *Zimmerische Chronik*,⁴ but often both transcription and notes are original, as with *Von abtuhung der Bylder*.⁵

Wiki software can be used outside of the realm of the Wikimedia foundation, the organisation that operates Wikipedia and its sister projects. There is an extensive wiki site devoted to the annotations of novelist Thomas Pynchon's works, as there are sites devoted to Terry Pratchett, to Eco's *The Mysterious Flame of Queen Loana*,⁶ and no doubt to many other authors and works. As these sites do not contain the text being annotated, they are somewhat outside the scope of this section, but much of this work is serious and inspiring.

Weblogging and CommentPress

Weblogs are publication systems that are typically used to publish series of short, diary-like articles, usually presented in reverse chronological order. Most weblog systems have a comment-facility that readers may use to comment, communicate with the author, etcetera. While weblogs are not usually employed as an edition platform,⁷ weblogs are home to a significant volume of today's primary writing. Weblog comments themselves are, in a different sense of the word, more 'primary' than the annotations one expects from a scholarly editor. Still, the ubiquity of the comment form, not just on blogs but on a plethora of news sites,⁸ cannot but influence what we will expect of the future's digital edition.

A weblog plugin specifically created to facilitate paragraph-by-paragraph signed comments is CommentPress (Fitzpatrick 2007). CommentPress is a 'theme' for the popular blog tool WordPress. To each section and paragraph it adds an icon that readers may use in order to add comments. In the right sidebar CommentPress adds a display of recent comments and a menu showing other options for browsing comments, such as 'by commenter' and 'by section'. CommentPress has been used as a discussion platform, as in (Scholarly Publishing Office 2007). It is being used in an educational setting for the annotation of Melville's *Pierre; or, The Ambiguities*.⁹ Martin Luther King's *Letter From Birmingham Jail* has been published and analysed using the CommentPress platform.¹⁰

Scholarly edition projects

Unlike the wiki or weblog approaches, the projects that I will discuss in this sub-section function within a more traditional editorial background.

The Virtual Humanities Lab at Brown's University (Zafrin and Armstrong 2005) is a site where scholars of the Italian Renaissance collaborate on a number of important texts from

¹ Statement based on inspection of mainly the English, Dutch and German versions of Wikibooks in December 2007.

² http://en.wikibooks.org/wiki/Annotations_of_The_Complete_Peanuts

³ http://en.wikibooks.org/wiki/The_Grand_Inquisitor

⁴ http://de.wikisource.org/wiki/Zimmerische_Chronik

⁵ http://de.wikisource.org/wiki/Von_abtuhung_der_Bylder

⁶ <http://pynchonwiki.com/>, <http://wiki.lspace.org/wiki/Annotations>, <http://queenloana.wikispaces.com/>.

⁷ They may be: see for an example Bert Van Raemdonck (ed.), *Virginie Loveling. In Oorlogsnood. Virginie Lovelings Dagboek 1914-1918. Tekstkritische Editie Door Bert Van Raemdonck* (Gent: CTB, 2004).

⁸ General purpose sites such as Topix (<http://www.topix.com>) or specialized sites like Slashdot (<http://slashdot.org/>).

⁹ <http://pierre.edublogs.org/>.

¹⁰ <http://www.futureofthebook.org/natestearns/letterfrombirminghamjail/>

the period.¹¹ The project conceives of itself as a ‘virtual editing house’ where scholars will have the possibility to add semantic encoding and annotation to the edited text. ‘Semantic encoding’ includes things like persons, terms, and themes; Zafrin and Armstrong also mention structural elements, quotations, rhetorical devices, etc. It should be possible to annotate the text as well as other annotations, possibly even in a forum. Even though this is clearly an experimental site, it is traditional in the sense that the annotations can only be made by established scholars with privileged access. The annotations they create are accessible to the public.

Another annotation system was developed for the Online Chopin Variorum Edition (OCVE) (Bradley and Vetch 2007).¹² The OCVE will house facsimiles of multiple editions and manuscripts of Chopin’s works. Students can attach annotations to points on the facsimile. In order to provide some support for exploratory annotation that is organised into a coherent whole at a later stage, annotations can be assigned categories and grouped into trees. Annotations can be either private or public. Like the system developed for the Virtual Humanities Lab, the OCVE’s system is still at an experimental stage. It was developed with an eye towards the edition’s end users, but, as in the VHL’s case, the edition has not yet reached the stage of end user usability.

The publication platform Telma (Bertrand 2006) was developed by the French IRHT and the École nationale des chartes, specifically for the publication of medieval charters.¹³ Telma contains limited facilities for annotation and bookmarking. The facilities are open to all registered users, and anyone can register. All annotations are private. The annotations are not shown with the annotated texts, but can only be accessed through a separate menu. There is no possibility for further organisation or categorisation of the notes.

Digital library projects

The number of scholarly edition projects that offer annotation facilities is surprisingly, and disappointingly, low. The digital library community however hardly does better. A number of sites offer annotation facilities at document level,¹⁴ which is not enough for our purposes. (Arko et al. 2006) describe a system for shared annotation of items in the Digital Library for Earth System Education (DLESE). The DLESE system is mainly oriented towards recommending items to other users of the library.

A test version of the new system for Gallica, the French National Library’s digital library offers bookmarking and tagging at page level.¹⁵ The bookmarks can be grouped into folders. The tags can be used to retrieve the corresponding pages. This goes some way towards developing the digital library into a research tool. All bookmarks and tags are private however. They cannot be shared with other researchers.¹⁶

The annotation system employed in the digital collections of the University of California (UCLA) is more powerful.¹⁷ The reader can create proper notes, not just tags, that are shown next to the page. The annotations and bookmarks can be stored as virtual collections, that can be either private or public. It is even possible to allow others to edit the same collection. A researcher could use a virtual collection to store pointers to the locations in

¹¹ <http://golf.services.brown.edu/projects/VHL/index.php>

¹² <http://www.ocve.org.uk/>

¹³ <http://www.cn-telma.fr/>

¹⁴ Public ‘reviews’ in the Internet Archive (<http://www.archive.org/details/texts>), private labels and reviews in the ‘My Library’-facility in Google Books (<http://books.google.com/>).

¹⁵ <http://gallica2.bnf.fr/>

¹⁶ Digital collections that are created using ContentDM software contain a facility for bookmarking pages. The collected bookmarks are private. See for an example the edition of Phillis Wheatley’s *Poems on various subjects, religious and moral*, at <http://www.sc.edu/library/spcoll/wheatley/wheatleyp.html>.

¹⁷ See for an example the Canon Law Digital Edition at <http://digital.library.ucla.edu/canonlaw/>.

a book that are relevant to an article he or she has written. A teacher could add clarifications to a number of pages he has assigned to a class. A researcher might even use the facility for in-depth annotation of an entire work, although page level annotations may not be the best means for doing that.

Collaboratories

Depending on one's strictness in defining scholarly text editions, a number of other sites might be mentioned here. Among these are a number of 'collaboratories', sites that facilitate collaboration among scholars in the study of, usually, documents. Annotation is an essential facility for collaboration, although successful collaboration requires more. Examples are Collate, a film collaborator (Frommholz et al. 2003) and IPSA, a collaboratory for the study of botanical illustrations (Agosti et al. 2003). One of the most ambitious collaboratories is no doubt HyperNietzsche (D'Iorio 2000). Annotation is here only an ingredient in a site facilitating genetic study of Nietzsche's manuscripts.

A very remarkable site is Footnote.¹⁸ Footnote contains historic material, mostly from the US National Archives and other agencies. Visitors to the site can add what is called 'annotations' – basically transcriptions of names, places, dates and other text, and free-form comments. They can create 'spotlights' – essentially a way of highlighting and commenting images or image regions. References to spotlights can be embedded in external web pages. Another way for visitors to add information to the collection is to connect images. Connections can be typed: the system provides a number of standard types, but the visitor can create his or her own connection type, and can even further annotate the connection. Visitors can also upload new material. Perhaps the most interesting way that visitors can contribute value is by creating story pages: pages that tell a story around images from the Footnote archives and possibly other, uploaded images. Footnote does many things that would be useful in digital editions. Footnote is, however, a commercial venture, and not all of the content on the site is freely accessible.

Summing up

It seems self-evident that an integrated annotation facility, even if the annotations are only accessible to the annotator himself, can be very useful. In spite of this, the number of digital editions, however one defines them, that offer annotation facilities is surprisingly low. Where sites provide annotation facilities, they often hide the annotations in separate windows, making them hard to access and use.

The question why annotation facilities are so rare is hard to answer. From the perspective of the researcher, who would entrust his research data to an unknown party, reservations are maybe understandable. The provider of the digital edition, however, should have no such qualms. No-one seems to have argued that annotation facilities are pointless. The answer must be that budgets are short and annotation facilities feel like an 'extra', something nice to have but not essential when a deadline looms.

Agosti, Maristella, Benfante, Lucio, and Orio, Nicola (2003), 'IPSA: A Digital Archive of Herbals to Support Scientific Research', *Digital Libraries: Technology and Management of Indigenous Knowledge for Global Access* (Lecture Notes in Computer Science, 2911; Heidelberg: Springer), 253-64.

Arko, Robert A., et al. (2006), 'Using Annotations to Add Value to a Digital Library for Education', *D-Lib Magazine*, 12 (5).

¹⁸ <http://www.footnote.com/>.

- Bertrand, Paul (2006), 'Telma. Naissance d'un centre de digital humanities en France', *L'IRHT, avenir d'une tradition* (Paris: IRHT).
- Bradley, John and Vetch, Paul (2007), 'Supporting Annotation as a Scholarly Tool: Experiences from the Online Chopin Variorum Edition', *Literary and Linguistic Computing*, 22 (2), 225-41.
- D'Iorio, Paolo (ed.), (2000), *HyperNietzsche. Modèle d'un hypertexte savant sur Internet pour la recherche en sciences humaines. Questions philosophiques, problèmes juridiques, outils informatiques*. (Paris: Presses Universitaires de France).
- Daly, Peter M. and Young, Alan R. (2002), *The Emblems of George Wither and Gabriel Rollenhagen. A Collection of Emblems*, eds Antonio Bernat Vistarini and Tamás Sajó (Studiolum).
- De Smedt, M. and Vanhoutte, E. (1999), 'Stijn Streuvels. De teleurgang van den Waterhoek. Elektronisch-kritische editie/electronic-critical edition', (Amsterdam: Amsterdam University Press (CD-ROM)).
- Fitzpatrick, Kathleen (2007), 'CommentPress: New (Social) Structures for New (Networked) Texts', *Journal of Electronic Publishing*, 10 (3). <<http://hdl.handle.net/2027/spo.3336451.0010.305>>, accessed Fall 2007.
- Frommholz, I., et al. (2003), 'Document-Centered Collaboration for Scholars in the Humanities—The COLLATE System', *Proc. 7th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2003)*, 434-45.
- Scholarly Publishing Office 'University Publishing in a Digital Age. Social Commentary on the Ithaka Report', <<http://scholarlypublishing.org/ithakareport/>>.
- Van Raemdonck, Bert (ed.), (2004), *Virginie Loveling. In Oorlogsnood. Virginie Lovelings dagboek 1914-1918. Tekstkritische editie door Bert Van Raemdonck* (Gent: CTB).
- Zafrin, Vika and Armstrong, Guyda (2005), 'Towards the electronic Esposizioni: the challenges of the online commentary', *Digital Medievalist*, 1 (1). <<http://www.digitalmedievalist.org/article.cfm?RecID=1>>, accessed 2007-12-28.