
Wikisource and the Scholarly Book

**Scholarly and Research
Communication**

VOLUME 3 / ISSUE 4 / 2012

Christian Vandendorpe
University of Ottawa

Abstract

Wikisource, a project of the Wikimedia foundation, is a growing online library aiming to provide well-edited texts thanks to an army of volunteers. In this article, the author attempts to assess the strengths of this project as well as its shortcomings, and makes some suggestions that, if adopted, would make Wikisource the ultimate platform for reading and editing scholarly books.

Christian Vandendorpe is professor emeritus at the Department of Lettres françaises of the University of Ottawa. Email: Christian.Vandendorpe@gmail.com .

Keywords

Reading tools; Interface; Layout; Annotations

CCSP Press

Scholarly and Research Communication

Volume 3, Issue 4, Article ID 040145, 8 pages

Journal URL: www.src-online.ca

Received March 22, 2012, Accepted March 22, 2012, Published April 10, 2013

Vandendorpe, Christian. (2012). Wikisource and the Scholarly Book. *Scholarly and Research Communication*, 3(4): 040145, 8 pp.

© 2012 Christian Vandendorpe. This Open Access article is distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc-nd/2.5/ca>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.


Last summer, as a thought experiment, I began to lay down the principles that should guide the creation of the ultimate digital book – or, since the term “book” may be a bit confusing these days, I was wondering what should be the main characteristics of the new knowledge environment prototype that the INKE (Implementing New Knowledge Environments) group is working on. I finally set down a few very general principles:

- Civilization is a cumulative process and relies on the art of nurturing the legacy of the past. As a consequence, we have to ensure that books of profound relevance for our culture will be available in digital form and fully integrated into the cultural ecosystem of the Web.
- New metaphors should be adopted in order to think about the digital book. In place of the container metaphor, we should put the emphasis on the book as a set of paths in an ever-growing forest of knowledge.
- The reader should be at the centre of the new digital culture and should have the possibility to fully integrate texts from the past to their present experience.
- The publishing platform must be open, versatile, and robust. It should be indefinitely expandable.

The remarkable achievements of Wikisource

In the midst of this reflexion, it came to me that some years ago the Wikimedia Foundation started a project called Wikisource, and I decided to revisit it. I was actually quite surprised to discover how much this project has grown and matured since its beginning in November 2003. According to Wikipedia: “The original concept for Wikisource was as storage for useful or important historical texts. These texts were intended to support Wikipedia articles, by providing primary evidence and original source texts, and as an archive in its own right” (Wikipedia, 2012a, para. 3). The project soon evolved to become “a general-content library” (Wikipedia, 2012a, para. 1). Today, the results are impressive, not only by the quantity of texts already available (about 250,000 in English), but also by the breadth of the project and its preoccupation for being a reliable source.

QUALITY EDITING

Texts added since 2008 are properly sourced and their status is clearly indicated. A text must have undergone a revision by two editors before being made available in reading format for the public. If you navigate Wikisource by clicking on the link “Random page,” you will rapidly come to a page that has not yet been proofread (see http://en.wikisource.org/w/index.php?title=Page:Fables_of_Aesop_and_other_important_mythologists.djvu/294&action=edit&redlink=1). If you click on the symbol , you arrive at the index of the same book, showing the state of the proofreading process (see http://en.wikisource.org/wiki/Index:Fables_of_Aesop_and_other_important_mythologists.djvu), denoted by a series of colour-coded radio buttons. The default state is “not proofread” in red; the yellow button means that the page has been proofread by one contributor; when a second contributor has proofread the same page, clicking on the yellow button changes its value to green, meaning that the page has been validated.

This system ensures a high quality text. A colour-coded bar at the top of the page (see http://en.wikisource.org/wiki/Fables_of_%C3%86sop_and_Other_Eminent_

Mythologists/The_Life_of_%C3%86sop/Chapter_I) under the title, shows the state of advancement of the proofreading process.

DISPLAY OPTIONS

At the beginning, books were formatted in a very basic layout (for example, http://en.wikisource.org/wiki/Oliver_Twist/Chapter_1), following Wikipedia's crude model of aligning text at full window width. However, many pages added since 2008 offer the reader the option of changing the display (for example, see http://en.wikisource.org/wiki/Personal_Recollections_of_Joan_of_Arc/Book_I/Chapter_2), notably to opt for the layout 1, or column format, which allows for a more comfortable reading experience. Unfortunately, there are still some members of the community who strongly oppose this innovation, hampering its general implementation.

VERSIONS

Versions are enabled by a mechanism of *transclusion*, i.e., "the inclusion of a document or part of a document into another document by reference" (Wikipedia, 2012b, para. 1). This feature is useful not only for illustrations but also for toggling between variants of a text, such as the original spelling and the modern one (see text by La Fontaine in http://fr.wikisource.org/wiki/Le_Philosophe_Scythe). This is a huge improvement on the former model where notes are in a raw format, like in this page created in 2005: http://en.wikisource.org/wiki/Prologue_to_the_Aeneids, or even with the traditional critical edition, like in this page (see http://en.wikisource.org/wiki/Piers_Plowman). One might expect that a parallel display will be adopted for these kinds of pages in the future.

Some pages of Wikisource already offer in parallel two versions of the same text, like this page by Caxton (1481) showing the original text and a *respelled* version: http://en.wikisource.org/wiki/The_Historye_of_Reynard_the_Foxe. This page of *The Romance of the three kingdoms*, an ancient Chinese novel, juxtaposes the original text in Chinese and the English translation: http://en.wikisource.org/wiki/Romance_of_the_Three_Kingdoms/Chapter_1. By contrast this page of Beowulf shows only the original text in old English, while the translation has to be found is on another page, like the Beowulf (Gummere): http://en.wikisource.org/wiki/The_Oldest_English_Epic/Chapter_1. Wikisource is a work in progress!

LINKS TO TRANSLATIONS

Following the lead of Wikipedia, Wikisource adopted a multilingual portal in the summer of 2005 and is now open in 58 languages and active in 50 languages. Some popular works are already linked to their translations in foreign languages. This feature would be quite useful for translations of ancient works. One can imagine a multi-language edition of Aristotle's *Poetics*. For the moment, *Poetics* is available in five languages: an English version from 1790 (see [http://en.wikisource.org/wiki/Poetics_\(Bywater\)](http://en.wikisource.org/wiki/Poetics_(Bywater))), a Spanish one from 1798, a French one from 1922, a Russian one from 1927, and the Greek version. This is a good beginning, although we are far from having all the versions needed in order to track the reception of this work across time and space.

Suggestions to make Wikisource better

There is no doubt that Wikisource is a serious project aimed at offering readers well-edited texts. In order to improve it, I would offer a few suggestions.

COMMENTS AND ANNOTATIONS

According to Wikipedia, “Wikisource has the capacity for annotated editions of texts” (Wikisource, 2012a, para. 3). The help page of Wikisource also states that it “applies the wiki approach to building a free library by letting volunteer contributors collect, organize, proofread, illustrate and *annotate* English-language source texts ” (Wikisource, 2012b, para. 2, emphasis added). Indeed, there are a few annotated pages, notably on *Doctor Jekyll* (see http://en.wikisource.org/wiki/The_Annotated_Strange_Case_Of_Dr_Jekyll_And_Mr_Hyde), *Travels with a Donkey* (see http://en.wikisource.org/wiki/The_Annotated_%27Travels_with_a_Donkey_in_the_Cévennes%27), and *Mishnah* (see http://en.wikisource.org/wiki/Mishnah/Seder_Zeraim/Tractate_Berakhot/Chapter_1/1). However, I did not find anything in the Wikisource layout that would support annotations other than in the form of footnotes. The reason was given on a Wikisource page in June 2011: “There is currently no community consensus on whether or not the English Wikisource should host user-generated content such as source text annotations. This page has been blanked pending resolution” (Wikisource, 2011, para. 1). By browsing through the history of that page, one can follow the discussions and appreciate the pro and con arguments.

This decision is unfortunate and may weaken the interest in this project not only amongst the general public but also within the academic community. The most important books of our culture have been heavily commented on, illustrating the multiple paths offered to interpretation by a same text. Ideally, those comments should be made available next to the work they refer to. If Newton, for example, did comment on a verse of the Bible, or Borges on Shakespeare, it would be useful to find these comments in a pop-up window next to the original verses. If annotations were tagged according to the date they were originally written, and automatically ordered in reverse chronological order, a reader could follow the evolution of interpretations across eras. This would be precious information for humanists.

The question of deciding if contemporary readers should also be able to add their own comments to any passage of a text poses a difficult problem since the cloak of perceived anonymity may in some cases give free rein to very bad manners. But this is not always the case, as documented by Disqus, a company providing a commenting platform to more than 400,000 websites: analyzing 500,000 comments, they found that anonymous users commented 4.7 times more than commenters using a real identity and that 61% of these anonymous comments were positive (MacArthur, 2012). On Wikisource, the presence of comments would make visible the relevance of a text to readers across time and space: in some cases, the comments of a particular reader could be even more interesting than the text to which they relate.

Although apparently mundane, the question of annotations is – like the proverbial tip of the iceberg – a pointer to the tectonic changes happening in the way people today relate to reading and writing. The default model of reading in the nineteenth century epitomized by the novel and its continuous thread of words has given way to the “starry” path we follow on a screen, going from one piece of information to another one and following our own associations in place of those put forward by a writer. The possibilities offered to the readers for expressing their opinions have also exploded. While the printed novel offered a minimum space for readers’ annotations, the screen has been encouraging readers’

feedback and interaction since its very beginning, by asking them either to evaluate their overall interest in an article, or to react to it. This trend has gained in strength with the advent of a large number of social tools, allowing readers to exchange their impressions and make recommendations. As a result, readers are ever more interested in adding their own comments to a text or in reading other readers' comments.

In 2008, a project put forward by *if:book*, London and funded by the Arts Council England invited seven women to comment on *The Golden Notebook* and put online the entirety of the original text in a codex-like layout and in parallel with the annotations (see <http://thegoldennotebook.org/book/p1>). The public can still participate in a forum, although some entries are plagued by spam. In spite of these minor inconveniences due to a lack of safeguards, I agree with Sam Anderson: "What I Really Want Is Someone Rolling Around in the Text" (Anderson, 2011). Bob Stein, who created the website Comment Press and its remarkable tools, was certainly right when he wrote that "the future of marginalia is bright" (Stein, 2011, para. 1). Proofs of this trend are everywhere.

Since 2009, Amazon allocates a personal space to readers of its Kindle books where they can copy and paste selected quotations and write annotations on particular pages, a space that the commenters can make public if they wish. When reading a book on the Kindle, one can also see how many readers have highlighted a particular sentence or paragraph. Reading is thus no longer a very private act, but part of the social sphere and a way of sharing our thoughts and enthusiasms with others. The giant library has clearly found a financial incentive for this initiative. There could even be a business model for including annotations in the books themselves. In August 2011, the COPIA website (<http://www.thecopia.com/home/index.html>) began to sell books specially annotated for a niche public: around Halloween, for example, the website homepage displayed the following advertisement for Bram Stoker's book: "Read Dracula like never before, as two vampire experts offer more than 200 annotations on eroticism, Twilight, Transylvanian folklore and more." Possibilities are endless.

In the discussions leading to the lack of support for annotations in Wikisource, some contributors argued that if Wikisource were open to annotations and comments, it would cause overlap with other projects, such as Wikipedia (if annotations were encyclopaedic) or Wikibooks, a collection of textbooks (if annotations were didactic). This argument does not make much sense, because if annotations were of either one of these types, they should be just hyperlinks, which are already possible in Wikisource. In fact, the main reason why the community could not reach a consensus on this question is because it would conflict with two overarching principles governing Wikipedia. The first one is the exclusion of original works; the second one is the obligation of adopting a neutral point of view, which is antithetic to the essence of a personal comment. Even if the efficiency of these two principles has been proven in the redaction of the encyclopaedia, they are clearly not adequate for a reading space like Wikisource where a large part of the interest stems from the juxtaposition of various points of view on subjective matters, like love, liberty, and human affairs in general. People read literature not in search of a neutral point of view but in order to connect with an individual voice and better decipher their own psyche through the lens provided by someone else.

If comments and annotations were allowed, however, the real problem for the administrators of Wikisource would be in attempting to filter derogatory comments without incurring the accusation of censorship. Maybe comments and annotations should be allowed only for non-anonymous users who have passed a vetting process and who are made aware that their entire collection of comments could be erased if they were found in contempt of an ethic chart.

Given the importance of this question in the context of Web 2.0 and the ever-greater integration of writing and reading, it is quite important to create a standard method of creating and accessing comments, in order for the authors to keep their comments even if they move from a reading platform to another one. In July 2011, the Niso Group got a grant from the Mellon Foundation for that purpose and is working on a project called “Open annotation collaboration.” One can find the slides of a presentation made by this group in a San Francisco conference in October 2011 online (see <http://www.slideshare.net/azaroth42/nisointernet-archive-meeting-on-social-bookmarking-and-annotation>).

COLLATING VERSIONS

Ideally, the different editions of a same work should be referenced on the main page with the possibility of displaying various versions in parallel. Since ancient books have known many versions and many translations, it would be very useful to give the user the possibility of comparing them in the same working space. The synoptic display of two, three, or four versions of a work should be made possible by a function that would allow the transclusion *at will* of other pages in the main window. A French user could thus display in parallel a play by Shakespeare along with two or three translations. Such a system would be of tremendous interest for the study of comparative literature and for an effective dialogue of cultures.

TIMELINE AND CONTEXT

In order to help users make a meaningful choice, it would be interesting to display books visually. It could be on a timeline or a map of the world where background events could be displayed through applications like the one found at Conflict History (<http://wwwconflicthistory.com>). The Perseus website, explored by Sondheim in this volume (“Interfacing the Collection”), offers very rich ways of interfacing the contents.

There are many other ways to give a new life to old books. The website *smalldemons* proposes original ways of creating a “storyverse”:

Suppose someone took every meaningful detail from all the books you love. Every song mentioned, every person, every food or place or movie title. ... Together they create something vast, wonderful and entirely new. A Storyverse. A place where details touch, overlap and lead you further. To new music to listen to. New movies to watch. Places to visit. People to know. And of course, new books to read. (Document2Small Demons, 2011, para. 1)

The various examples are very illuminative and quite in tune with our new mediasphere where texts are intertwined with images and sounds.

CITATIONS: HISTORY MATTERS

In a bid to gain approval from students and teachers, Wikisource produces for each work a way to make a citation in various formats (MLA, APA ...). According to the system in place, the English version of Aristotle's *Poetics* should be referenced in the following manner:

“The Poetics translated by Bywater/1.” Wikisource, The Free Library.
9 Jan 2009, 13:05 UTC. Wikimedia Foundation, Inc. 24 Oct 2011 </
en.wikisource.org/w/index.php?title=The_Poetics_translated_by_
Bywater/1&oldid=946293>. (Wikisource, n.d.)

Although the general idea is good, I wonder why, after having taken so much care in being faithful to the book scanned, the citation mentions neither the place nor the date of the original book. One would think that this information is more important for referencing a book than the date of its inclusion in Wikisource.

MULTIPLE FORMATS

The “book” should be available on various platforms, ensuring ubiquitous access as well as the possibility for readers to continue their reading path through the text from where it was left at the previous session. Presently, a page can be saved in PDF, Open Document Text, and OpenZim. Being by default in letter format size, the PDF is not ideal for reading on a tablet or on a phone. The endnotes, moreover, are not hyperlinked in this PDF, which is unfortunate. It would be very useful if there were support for saving a page or a collection of pages in epub format.

Conclusion

Thanks to its community of volunteers, Wikisource has been growing at a steady pace. At this moment, it has reached a critical mass of texts that makes it well equipped to become a platform of choice for reading and editing texts. It is compliant with most of the principles that I put forward at the beginning of this article. It should evolve on a few fronts, however, if it wants to keep its lead, notably against such behemoths as Amazon and Google Books. First, there should be a real embrace of the new reading environments offered by tablets and mobile devices as well as column and multi-column layout because text is not just a string of characters but a space offered to the eye for the delight and the reflection of the reader. Also, if Wikisource would open up in a sensible and easy way to comments and annotations, it would offer a wealth of possibilities to researchers and students: small teams of graduate students could work on critical editions of texts, compare translations in various languages, or add comments illuminating the meaning of a paragraph. In short, the whole spectrum of text productions from the last five millennia could become readable and hyperlinked in meaningful ways and open new ways of pondering our own place in the evolution of humanity.

Websites

Beowulf (Gummere). URL: http://en.wikisource.org/wiki/The_Oldest_English_Epic/Chapter_1
Creating Page:Fables of Aesop and other eminent mythologists.djvu/294. URL: http://en.wikisource.org/w/index.php?title=Page:Fables_of_Aesop_and_other_important_mythologists.djvu/294&action=edit&redlink=1
Conflict History. URL: <http://www.conflicthistory.com>

- COPIA. URL: <http://www.thecopia.com/home/index.html>
- Fables of Æsop and Other Eminent Mythologists/The Life of Æsop/Chapter I. URL: http://en.wikisource.org/wiki/Fables_of_%C3%86sop_and_Other_Eminent_Mythologists/The_Life_of_%C3%86sop/Chapter_I
- Index:Fables of Aesop and other eminent mythologists.djvu. URL: http://en.wikisource.org/wiki/Index:Fables_of_Aesop_and_other_eminent_mythologists.djvu
- Le Philosophe Scythe. URL: http://fr.wikisource.org/wiki/Le_Philosophe_Scythe
- Mishnah/Seder Zeraim/Tractate Berakhot/Chapter 1/1. URL: http://en.wikisource.org/wiki/Mishnah/Seder_Zeraim/Tractate_Berakhot/Chapter_1/1
- Oliver Twist/Chapter 1. URL: http://en.wikisource.org/wiki/Oliver_Twist/Chapter_1
- Open Annotation Collaboration Review. URL: <http://www.slideshare.net/azaroth42/nisointernet-archive-meeting-on-social-bookmarking-and-annotation>
- Personal Recollections of Joan of Arc/Book I/Chapter 2. URL: http://en.wikisource.org/wiki/Personal_Recollections_of_Joan_of_Arc/Book_I/Chapter_2
- Piers Plowman. URL: http://en.wikisource.org/wiki/Piers_Plowman
- Poetics (Bywater). URL: [http://en.wikisource.org/wiki/Poetics_\(Bywater\)](http://en.wikisource.org/wiki/Poetics_(Bywater))
- Prologue to the Aeneids. URL: http://en.wikisource.org/wiki/Prologue_to_the_Aeneids
- Romance of the Three Kingdoms/Chapter 1. URL: http://en.wikisource.org/wiki/Romance_of_the_Three_Kingdoms/Chapter_1
- The Annotated Strange Case Of Dr Jekyll And Mr Hyde. URL: http://en.wikisource.org/wiki/The_Annotated_Strange_Case_Of_Dr_Jekyll_And_Mr_Hyde
- The Annotated ‘Travels with a Donkey in the Cévennes’. URL: http://en.wikisource.org/wiki/The_Annotated_%27Travels_with_a_Donkey_in_the_C%C3%A9vennes%27
- The Golden Notebook. <http://thegoldennotebook.org/book/p1>
- The History of Reynard the Foxe. URL: http://en.wikisource.org/wiki/The_History_of_Reynard_the_Foxe

References

- Anderson, Sam. (2011, March 4). What I really want is someone rolling around in the text. *The New York Times*. URL: http://www.nytimes.com/2011/03/06/magazine/06Riff-t.html?_r=1&pagewanted=all [June 16, 2012].
- MacArthur, Amber. (2012, January 18). Anonymity is toxic to online comments, except when it's not. *The Globe and Mail*. URL: <http://www.theglobeandmail.com/technology/digital-culture/amber-mac/anonymity-is-toxic-to-online-comments-except-when-its-not/article620877> [June 16, 2012].
- Small Demons. (2011). *Welcome to the Storyverse*. URL: <https://www.smalldemons.com> [June 16, 2012].
- Stein, Bob. (2011). The future of marginalia is bright (not dim). *if:book*. URL: http://www.futureofthebook.org/blog/archives/2011/02/the_future_of_marginalia_is_br.html [June 16, 2012].
- Wikipedia. (2012b). *Wikisource*. URL: <http://en.wikipedia.org/wiki/Wikisource> [June 16, 2012].
- Wikipedia. (2012b). *Transclusion*. URL: <http://en.wikipedia.org/wiki/Transclusion> [June 16, 2012].
- Wikisource. (2011). *Wikisource:Annotations*. URL: <http://en.wikisource.org/wiki/Wikisource:Annotations> [June 16, 2012].
- Wikisource. (2012a). *Wikisource*. URL: <http://en.wikipedia.org/wiki/Wikisource#Library> [June 16, 2012].
- Wikisource. (2012b). *Help:Introduction*. URL: <http://en.wikisource.org/wiki/Help:Introduction> [June 16, 2012].
- Wikisource. (n.d.). *Cite*. URL: http://en.wikisource.org/w/index.php?title=Special:Cite&page=The_Poetics_translated_by_Bywater/1&id=946293 [June 16, 2012].