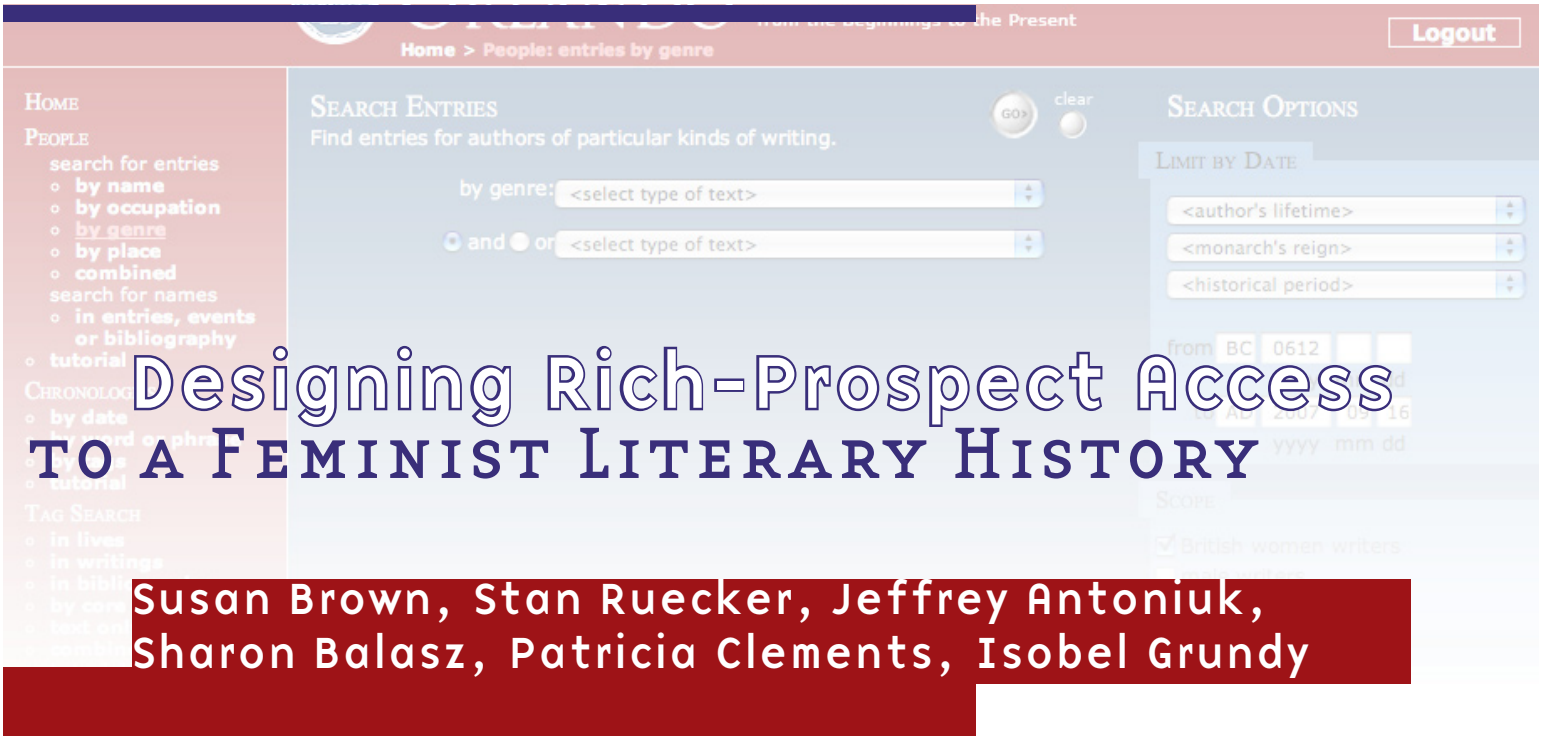


Designing Rich-Prospect Access to a Feminist Literary History

Susan Brown, Stan Ruecker, Jeffrey Antoniuk, Sharon Balasz, Patricia Clements, Isobel Grundy

Susan Brown, *of the University of Guelph and the University of Alberta, is a co-founder and Project Co-investigator of the Orlando Project.*

Stan Ruecker *is an Assistant Professor of Humanities Computing and English and Film Studies at the University of Alberta. He is a Project Co-investigator on the Orlando Project.*

Jeffrey Antoniuk *holds an Master of Science in Computing Science from the University of Alberta. He is the Systems Analyst on the Orlando Project.*

Sharon Balasz *holds a Master of Arts in Russian and a Master of Library and Information Science from the University of Alberta. She is the Textbase Manager on the Orlando Project.*

Patricia Clements *is a Professor Emeritus of English and Film Studies at the University of Alberta. She is a co-founder and the Project Director of the Orlando Project.*

Isobel Grundy *is a Professor Emeritus of English and Film Studies at the University of Alberta. She is a co-founder and Project Co-investigator of the Orlando Project.*

The Politics of Category Work

Much is at stake in the development of new technologies for writing and reading women's literary history. The Orlando Project has been engaged in this endeavour for more than a decade: our efforts to date have emerged in the online publication of *Orlando: Women's Writing in the British Isles from the Beginnings to the Present* in 2006 (orlando.cambridge.org; see also http://www.cambridge.org/online/orlandoonline). This complex textbase, which at first glance resembles a conventional online reference work, is actually a complexly structured experiment in new ways of doing and presenting feminist literary history. This experiment continues not only with regular updating and expansion but also with new work by the project team towards the next phase of the Orlando history: the integration of sustained discursive history into the textbase. This paper surveys some of the major concerns associated with producing feminist literary history electronically and outlines some strategies we are developing for providing an interface for the discursive portions of the history.

As John Seely Brown and Paul Duguid argued in *The Social Life of Information*, digital technologies are caught up in powerful social networks, communities, support systems (or lack of them), and institutions. When a project is a feminist one, there are valid concerns regarding the digital medium itself. One is access: in North America, women equal or (for younger women) outstrip men in internet participation, but elsewhere this is not the case. If not in the academy or in business, in the home context it seems probable that women as a group own or have access to less powerful and less up-to-date computers than do men as a group. Even in Canada and the US, women are under-represented in science and technology education, as presenters at new technology conferences, and lag behind men as early adopters of technology, suggesting that there may be something of a disconnect between an experimental feminist digital project and its target user base. Even if this does not follow, since many feminists have found in digital media a hospitable space for exploring new modes of representation, expression, creativity, communication, and activism, social factors nevertheless have a profound impact (see Fernandez et al). What in a digital resource is readable, knowable, learnable and by whom is in large part a result of the ways in which the digital is socially embedded, as well as the particulars of the resource.

Technologies shape the academic world in large part via category work. As numerous scholars from Geoffrey C. Bowker and Susan Leigh Star to Hope Olson have demonstrated, categories and classificatory systems have profound effects, whether we are talking about library cataloguing

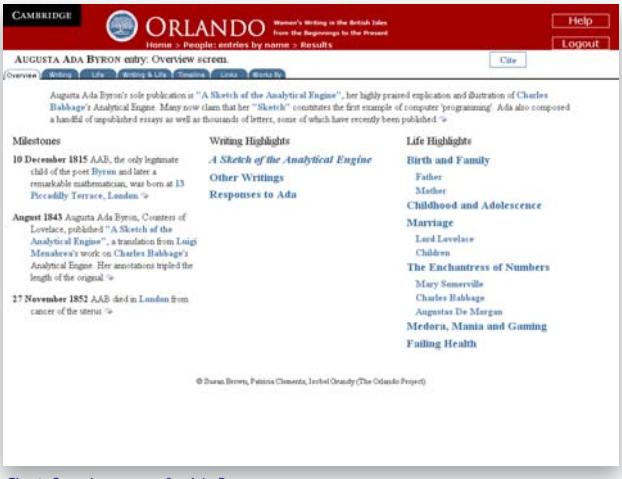


Fig. 1: Overview screen for Ada Byron entry.

systems of apartheid or the conceptual organization of nursing as a discipline. Categories are never more powerful or potentially pernicious than when invisible and naturalized. Preparing material for digital presentation involves implicit or explicit category work that has a major impact on how a system looks and works, and what someone can do with it. The interface between a computer user and a system's underlying design becomes crucial, then, if we agree with Andrew Feenberg that the technical codes of our culture are not neutral, and tend to naturalize and make routine "the pursuit of power and advantage by a dominant hegemony" (15). Since feminist scholarship aims to work against that hegemony, interface issues become vital.

Orlando

Feminist literary history investigates and seeks to elucidate the uneven historical operations of the contested category of gender, particularly in the field of women's writing. It has important corrective work to do, since the stories told about literary history are still often unsatisfactory. To take just a few of the most egregious and persistent claims: women don't write satire; women don't write ambitious forms like epic; women are expressing their feelings, rather than taking out insurance policy, in grovelling prefaces; women have been unaware of or hostile to female predecessors. Accounts (or disregard) of women writers are often still shaped by the assumption that women write not by training or learned skill but by instinct, leading to the conclusion that therefore they are more spontaneous than men, are not helped but sometimes spoiled by teaching or influences, and their earliest work is often the best. Further, biographical discussions of women writers often exhibit a confusion between using knowledge of the material circumstances of composition to deepen understanding of finished work, and using this knowledge for special pleading, or to suggest that women need special forbearance in criticism of their work. These sorts of misconceptions, along with our desire to re-centre literary history on what has generally been treated as marginal or absent, drive our interest in fact and context, in filling out and complicating the picture in a careful and nuanced fashion. In an electronic environment, such a project must devise a means of doing so while making its own category work and structural principles as explicit as possible.

Orlando has done this in producing an extensive literary historical textbase. It is organized around a core set of biographical and critical entries on individual women writers, and in order to make it as user-friendly as possible, the interface resembles a standard literary reference work, as can be seen from a sample entry screen (Fig. 1).

Orlando's text, organized as entries, timeline materials, and bibliographical information, can be sliced and diced into differently ordered or selected chunks, and navigated according to hypertext links. So, for instance, you can move from the Ada Byron entry front screen (Fig. 1) into the screen discussing her life, which contains a hyperlinked mention of her mother, born Anne Isabella Milbanke, whose much-contested name is sometimes given as Lady Noel Byron. Clicking on Lady Byron's name takes you to a set of links offering, among other options, a timeline of Lady Byron's life whose eight items are derived from the entries for her daughter Ada, her husband George Gordon Lord Byron, as well as those for Anna Jameson, and Harriet Beecher Stowe. *Orlando* can be searched and reformatted dynamically into sets of excerpts based on various category markers embedded in the prose. These tags are XML (Extensible Markup Language) markup; they label not only the formal structure of *Orlando* materials (as does HTML or Hypertext Markup Language) but their semantic content. For instance, biographical entries are sectioned invisibly by the tagging with sixteen major topics (and more specific tags within those tagsets); discussion of writings draw on many more tags in the three areas of production, reception and textual features (Fig. 2). Such semantic tags allow for dynamic user-directed searching of the textbase according to a huge variety of criteria with varying degrees of specificity. You can survey all discussions of violence within the "Life" sections of entries, or you can gather together mentions of a particular publisher, say Smith, Elder, in the context of writers' relationships with their publishers. Most inquiries will take multiple tacks, for instance, inquiry into the place of anthologies in women's literary history. Suppose you wish to investigate the part played by anthologies in the production and circulation of writing by women. You might begin by searching

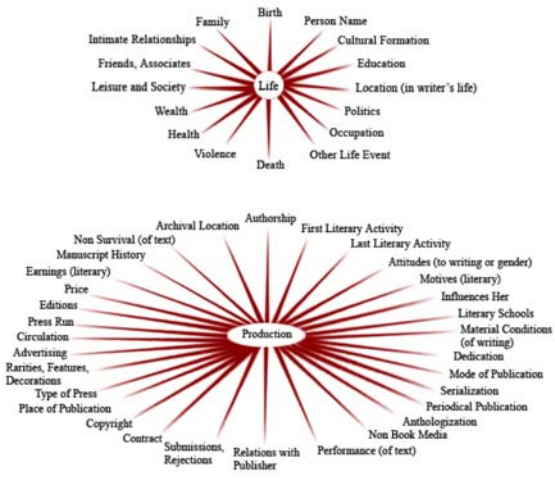


Fig. 2: Orlando life tagset and the writing (production) tagset

for writers whose entries are tagged for anthologization. This produces a list of 116 women whose entries mention the inclusion of their work in anthologies, more than 10% of the total (currently 870) of women with entries in the textbase. Looking at these results in the form of excerpts reveals 198 mentions: not every instance of the anthologization of work by a writer included in the textbase, but a wealth of reference to anthologies, some famous, some canon-making, some quirky, some specialized, etc. A search on "anthology" as an attribute of the tag <genre> yields 257 hits (again delivered in the form of excerpts) about writers who compiled anthologies (fewer than 257 writers, since some produced several anthologies). A search on anthology in the Author Summary tag whittles down those 116 plus 257 writers to 53 whose association with anthologies (nearly all as compilers rather than those included) was important enough to merit mention in a brief outline of their significance. Finally, a search on the word "anthology" within the Genre Issue tag gives a much more selective set of results, discussing such matters as what is important or desirable in an anthology, whether a lavish use of quotation can be regarded as equivalent of anthologizing, etc. This is just one quite simple example of what the semantic markup can do: other possible searches include London performances; writers politically active in particular causes or organizations; writers who quote other writers or are influenced by particular writers; books set in a particular place, and so on.

The Orlando team is now turning from the production of the more granular material in the form of entries (although we continue to add another 50 writers per year along with updating existing materials) to writing several books of continuous literary historical prose that are planned for both print and electronic dissemination. In their electronic form, they will be integrated with the rest of the *Orlando* textbase, similarly tagged to reflect the content, and they will interact dynamically with the other materials in terms of hyperlinking and search and retrieval. The remainder of this essay focuses on the interface strategies we are developing for these books.

Interface

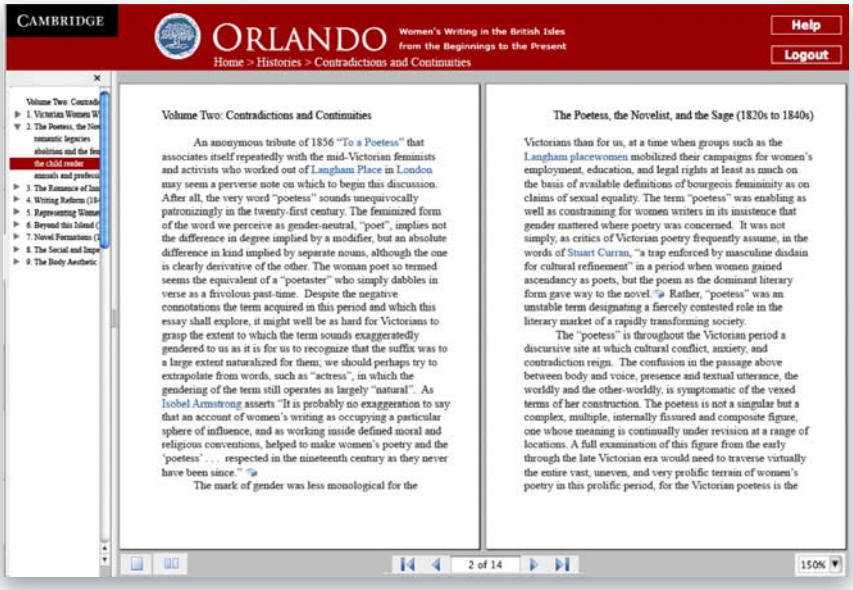
The user interface is "what we see on the computer screen," the representation of the system to the reader or user (Bell et al). It is a crucial aspect of a system, translating the underlying architecture and internal operations into something that communicates and permits interaction with the user. Interfaces must provide navigational and knowledge management assistance crucial to the usability of complex and extensive electronic texts such as *Orlando.*

Development of the interface we describe here is a significant undertaking. As Michael Best argues, while the Text Encoding Initiative guidelines for XML have become established standards for the encoding of text in the humanities, "the actual method of displaying texts thus marked up has not yet reached any kind of predictability" (6). This posed major challenges for the design of *Orlando*'s existing delivery system (Orlando Team, 2006). The further challenge, as we turn to the books, of translating lengthy scholarly argumentation into digital media is underscored by some of the findings of online usability studies. Print doesn't translate unproblematically to a screen. For instance, studies by Jakob Nielsen indicate that pdf files, the closest emulation we have on screen of print pages, decrease usability by 300%. We need more usability studies of academic users, who are more accustomed to reading lengthy and complex documents than the general public, whose professional lives involve more and more reading on-line, and who may be more strongly motivated to "dig" for information than the users typical of studies focused on commercial sites. However, it is clear, even now, that projects such as Orlando must explore the modes of textuality enabled by digital interfaces, so as to allow users to leverage critical interpretive systems such as semantic tagging comfortably and effectively. Searches on the semantic tagging will draw on the discursive literary histories along with the rest of the textbase materials, so the books must be "chunkable" while also making substantial arguments. Because the books won't be easily sorted by name- or date-organized entries, chronological lists, or bibliographical materials, these lengthy texts will require a different set of navigational tools if the online version is to deliver a reading and usage experience in any way comparable in ease and efficiency to the printed version.

Interface principles

Both in the Orlando Project and more generally in a number of interface experiments, we have adopted a small set of guiding principles. First, we think it is important that the reader has more than one path into the text, and more than one form of the text available at the end of the path. For instance, *Orlando* contains three distinct groups of entry points, for searches directed respectively at finding out about people and chronologies, and for searches making expert use of the tagging to dynamically assemble subsets of the textbase which cut across individual entries. Most of

Fig. 3: This sketch of the Orlando Reading View is intended to simulate the facing pages of a book, with the added advantage that the text can be resized within the text block.



Orlando is available at what is essentially the paragraph, or several-paragraph, level of granularity, so that material can be read as usual in the context of longer documents, or else extracted from the longer documents and reassembled for reading as a series of paragraph-sized pieces that weren't originally written together.

Secondly, rather than obfuscate or naturalize the category work performed by the system, we try to make the user interface transparent without being overwhelming. Our category work and the principles that inform it are available in a number of forms, so that users can assess how the system generates results and on that basis make informed choices and readings of the textbase.

Our third set of principles derives from previous work by Ruecker (2006) on rich-prospect browsing interfaces, where some meaningful representation of every item in a collection is combined with emergent tools for manipulating the display. In addition, the representations should provide access to further information. The dynamic construction of chronologies within the current *Orlando* is an example of this kind of manipulation. For the proposed interfaces that provide tag browsing and index browsing, the original meaningful representation consists of the items in the table of contents, and the emergent tools are the lists of tags and index entries that can be used to dynamically insert the contents of tags into the table of contents (TOC). Since each line in the TOC will allow access to the volume, we have also met the condition of the representation serving as a means of accessing further information.

Reading view

Our work on the Orlando Project offered the opportunity to explore the design of dynamic tables of contents. As mentioned above, the next phase of Orlando will involve writing and publishing a series of three volumes of literary history, intended for broad coverage of three historical periods and written for continuous extended reading. The volumes will be published in print form, but will also be available online for subscribers to the project. Our goal in the reading view (Fig. 3) is to provide an interface that resembles facing pages, with a number of additional features. The reader can, for instance, dynamically resize the text to a fontsize more convenient for reading, and an optional navigation feature based on the table of contents is provided on the left side of the screen, allowing the reader to move easily between chapters or other sections of the book. Readers can access this reading screen in several different ways, including directly from the list of volumes, indirectly through the dynamic table of contents, and with another degree of indirection through the index browser or tag browser, after identifying tag content of interest and clicking on it. Finally, these reading pages can also be accessed through other entry points to the project, many of which will offer the possibility of including the "histories" in the searches.

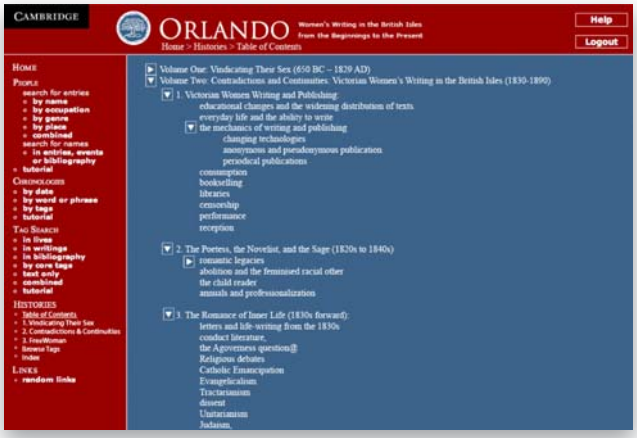


Fig. 4: The Dynamic TOC sketch shows a simple but interactive version of the standard table of contents, where items can be expanded or collapsed as in a typical GUI interface.

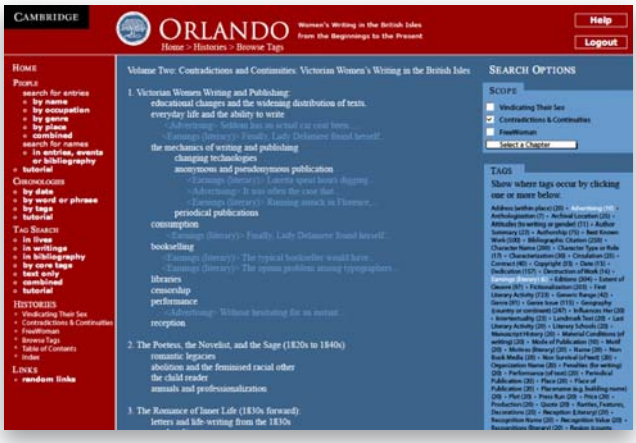
other readers may wish to dip into the books by using the table of contents. In this case, the proposed TOC is in effect a series of links that can be collapsed or expanded as the reader scans through for a chapter or section of interest (Fig. 4). Clicking on one of the expanded items will open the appropriate volume at the location indicated.

The use of a dynamic table of contents is not an innovation in itself, but represents what may be a necessary step in the process of encouraging the reader toward more interactive approaches to accessing online material intended for continuous extended reading. Such approaches are represented in our Orlando sketches by the Tag Browser and Index Browser.

Tag Browser

The proposed Tag Browser provides a degree of interactivity with the online table of contents that has not been widely explored, although the possibility for this new affordance is present in any collection with semantic encoding (Ruecker 2005). Our sketch for this interface (Fig. 5) shows the list of available tags on the right, with the table of contents for Volume Two in the central panel. The conventional TOC entries are in white, and in light blue are the first few words of content that has been tagged by the collection authors with the tag chosen by the user. In this case, the tags are "advertising," which occurs three times in the document, and "earnings (literary)," which occurs six times. By clicking on any of the lines of text in the table of contents, the reader can choose to be linked either to that point in the document, or to a dynamically-assembled collection of paragraphs that contain the tag of interest.

Fig. 5: The proposed Tag Browser is a more sophisticated design than the Dynamic TOC, since the Tag Browser allows users to add or subtract content items from the view, according to the choices made by clicking on items on the right panel of the screen.



Index Browser

While the Tag Browser provides access to spans of texts enclosed by tags, the Index Browser adds the concept of identifying locations of interest as points in the text rather than as ranges (Fig. 6). A typical index entry does not enclose a body of text but instead identifies a single location. This proposed interface combines both kinds of material in a single list, which has been moved here to a central location in order to better accommodate the greater number of items. The TOC here, as in the proposed Tag Browser, serves as the basis for showing prospect on the entire document, with tagged items or index entries dynamically inserted as the reader selects them.

TagCloud Creator

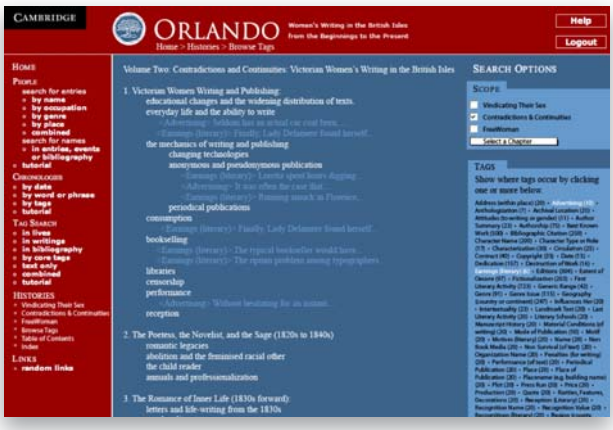
In the case of *Orlando*, the project members have the opportunity to provide translation of XML tag names into forms that are hopefully understandable by people outside the project. For other digital collections, it will be necessary to create an intermediate technology that can allow librarians to generate a meaningful selection of tags and rename them for general use (Fig. 7). Once the tags have been selected and renamed, they can be made available for an interface that allows readers to add or subtract tagged content from an existing table of contents.

Conclusions

Each of the interface sketches we've described here is at the beginning of our research life cycle, which moves from static sketches to kinetic versions to interactive prototypes, and finally into development systems available online to the larger community of readers. At each stage, we are able to carry out some form of user studies involving discussions about the functionality and the design, and in later phases we can measure actual use through screen captures and analysis of log files. The strategies that we are pursuing build on our previous understanding of the need to provide the reader with a range of possibilities, some of which are tightly coupled to approaches inherited from print, while others emerge from the new opportunities offered by working in a digital environment. The stages in the process, however, are also iterative, meaning that responses from prospective readers who act as our study participants will influence the direction taken in the next round of designs.

Given that our user group is predominantly but by no means exclusively female, we must ask also whether gender considerations in interface development extend beyond what we've outlined here, which are hardly gender-specific. For instance, studies have found that women tend to engage more in communication-oriented activities rather than information-seeking ones: might these sorts of factors have interface design implications? To historians with a strong sense of the changing constructions of gender, however, the question is a complicated one. It seems entirely possible that hypotheses about gendered interface design may reflect rather than trouble gender assumptions (Paasonen). These are complex issues that require extensive research into the myriad ways in which "technologies operate as a site for the production of gendered knowledge and knowledge of gender" (Wajcman 45). In the meantime, the Orlando Project's expanding literary history will give priority to rich prospect and the transparent deployment of categories in our development of new interfaces for reading and engaging in feminist scholarship.

Fig. 6: While the sketch of the Tag Browser provides access only through the tagged content, the Index Browser adds another level of sophistication, by including both tags and index items.



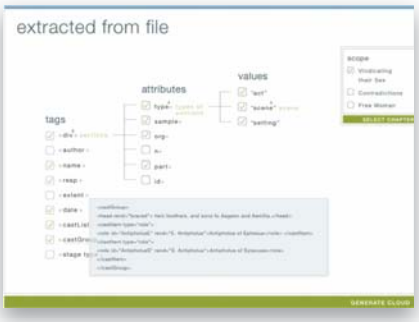


Fig. 7: The TagCloud creator is shown here in three parts, beginning at the top. It is intended to generalize the approach we've been developing for the Orlando volumes, by allowing the curator of any digital document tagged in XML to create an equivalent of the Tag Browser, which appears as a tag cloud and an interactive table of contents. A working prototype has recently been developed by Stéfan Sinclair at McMaster University (Ruecker and Sinclair 2007).

Works Cited

Bell, David, Brian D. Loader, Nicholas Pleace and Douglas Schuler. *Cyberculture: The Key Concepts*. London and New York: Routledge, 2004.

Best, Michael. "Forswearing Thin Potations: The Creation of Rich Texts Online." *Mind Technologies: Humanities Computing and the Canadian Academic Community*. Ed. Raymond Siemens and David Moorman. Calgary: University of Calgary Press, 2006. 1-14.

Continued on next page . . .

Bowker, Geoffrey C. and Susan Leigh Star. *Sorting Things Out: Classification and its Consequences*. Cambridge, Mass. and London: MIT Press, 1999.

Brown, John Seely and Paul Duguid. *The Social Life of Information*. Boston: Harvard Business School Press, 2002.

Brown, Susan, Patricia Clements, and Isobel Grundy, ed. *Orlando: Women's Writing in the British Isles*. Cambridge: Cambridge UP Online, 2006. <http://orlando.cambridge.org>.

Dholakia, Ruby Roy et al. "Gender and Internet Usage." 2003. <http://ritim.cba.uri.edu/wp2003/pdf_format/Wiley-Encycl-Internet-Usage-Gender-Final.pdf>.

Feenberg, Andrew. *Transforming Technology: A Critical Theory* Revised. Oxford: Oxford UP, 2002.

Fernandez, Maria, Faith Wilding, and Michelle M. Wright, ed. *Domain Errors!: Cyberfeminist Practices*. Brooklyn: Autonomedia, 2006.

Kottke, Jason. "Gender Diversity at Web Conferences" Blog entry 22 February 2007. <http://www.kottke.org/07/02/gender-diversity-at-web-conferences>.

Nielsen on pdfs <http://www.useit.com/alertbox/20010610.html>.

Noguchi,Yuki. "Women Narrow the Internet Gender Gap, Survey Finds." *Washington Post*. Thursday, December 29, 2005; Page D01. <http://www.washingtonpost.com/wp-dyn/content/article/2005/12/28/AR2005122801403.html>.

Marcus, Aaron. "Cultural Dimensions and Global Web User-Interface Design: What? So What? Now What?" *Proceedings of the 6th Conference on Human Factors and the Web in Austin, Texas, 19 June 2000*. Preprint. <http://www.amanda.com/resources/hfweb2000/hfweb00.marcus.html>.

Olson, Hope A. "The Power to Name: Representation in Library Catalogs." Signs 26.3 (2001): 639-668.

Orlando Team. "Between Markup and Delivery; or Tomorrow's Electronic Text Today." *Mind Technologies: Humanities Computing and the Canadian Academic Community*. Ed. Raymond Siemens and David Moorman. Calgary: U of Calgary P, 2006. 15-32.

Paasonen, Susanna. "The Woman Question: Addressing Women as Internet Users." *Domain Errors: Cyberfeminist Practices*. Ed. Maria Fernandez, Faith Wilding, and Michelle M. Wright. Brooklyn: Autonomedia, 2006. 89-108.

Ruecker, Stan. "The Electronic Book Table of Contents as a Research Tool." Paper presented at the President's Panel. Congress of the Humanities and Social Sciences: Consortium for Computers in the Humanities / Consortium pour Ordinateurs en Sciences Humaines (COCH/COSH) Annual Conference. London, Ontario. May 27- 30, 2005.

---. "Experimental Interfaces Involving Visual Grouping During Browsing." *Partnership: the Canadian Journal of Library and Information Practice and Research* 1(1). 2006.

Ruecker, Stan and Stéfan Sinclair. "Multilevel Displays and Document Blueprinting: Dynamic Browsing Using XML Structures and Text Features." Paper presented at the Digital Humanities 2007 conference. University of Illinois at Urbana-Champaign. June 4-7, 2007.

UNIFEM Central and East European Office. *Bridging the Gender Digital Divide*. <http://web.undp.sk/uploads/Gender%20and%20ICT%20reg_rep_eng.pdf>.

Wajcman, Judy. *Technofeminism*. Cambridge: Polity, 2004.