The Beginning, the Middle, and the End: New Tools for the Scholarly Edition

Stan Ruecker
ITT Institute of Design

Geoffrey Rockwell, Daniel Sondheim, Mihaela Ilovan, Jennifer Windsor, Mark Bieber, Luciano Frizzera, Omar Rodriguez, Kamal Ranaweera, & Carlos Fiorentino
University of Alberta

Stéfan Sinclair
McMaster University

Milena Radzikowska
Mount Royal University

Teresa Dobson
University of British Columbia

Ann Blandford, Sarah Faisal, & Alejandro Giacometti
University College London

Susan Brown
University of Alberta, University of Guelph

Brent Nelson
University of Saskatchewan

Piotr Michura
Academy of Fine Arts, Krakow, Poland

The INKE Research Team

Stan Ruecker is Associate Professor at the Institute of Design, Illinois Institute of Technology. Stan holds advanced degrees in English, Humanities Computing, and Design, and has expertise in the design of experimental interfaces to support online browsing tasks. Email: sruecker@id.iit.edu .

Geoffrey Rockwell is a Professor of Philosophy and Humanities Computing, University of Alberta. He is Director of the Canadian Institute for Research in Computing and the Arts (CIRCA). Email: geoffrey .rockwell@ualberta.ca .

Daniel Sondheim is a graduate student in Humanities Computing and Library and Information Studies at the University of Alberta. He is a graduate research assistant for the interface design team of INKE. Email: sondheim@ ualberta.ca .

Mihaela Ilovan is a graduate in Classic Archaeology and studied Humanities Computing and Library and Information Studies at the University of Alberta. She is currently working as a graduate research assistant for the interface design team of INKE. Email: ilovan@ualberta .ca .

Abstract

This article discusses a set of prototypes currently being designed and created by the Interface Design Team of the Implementing New Knowledge Environments (INKE) project. These prototypes attempt to supplement the user experience in reading digital scholarly editions, by supporting a set of tasks that are straightforward in a digital environment but in a print edition would be sufficiently more difficult as to be prohibitive. We therefore offer these experimental prototypes as a collection of new affordances for the scholarly edition, although they may reasonably be extended, with some variation, to other kinds of digital text.

Keywords

Human-computer interaction: Digital scholarly editions; Experimental interface design

Jennifer Windsor is a graduate student in the Humanities Computing Program at the University of Alberta. She is a research assistant for the interface design team of INKE. Email: jwindsor@ualberta.ca .

Mark Bieber is a graduate of the Department of Computer Science at the University of Alberta. Email: bieber@ualberta.ca .

Luciano Frizzera is a graduate student in the Humanities Computing Program at the University of Alberta. He is a research assistant for the interface design team of INKE. Email: dosreisf@ualberta.ca .

Omar Rodriguez is a research computing programmer/analyst in the Arts Resource Centre at the University of Alberta. His research interest is in the area of physically accurate animation and real-time rendering. Email: omar.rodriguez@ualberta.ca .

Kamal Ranaweera is a senior analyst in the Arts Resource Centre at the University of Alberta. Email: kamal.ranaweera@ualberta.ca .

The INKE Research Group comprises over 35 researchers (and their research assistants and postdoctoral fellows) at more than 20 universities in Canada, England, the United States, and Ireland, and across 20 partners in the public and private sectors. INKE is a large-scale, long-term, interdisciplinary project to study the future of books and reading, supported by the Social Sciences and Humanities Research Council of Canada as well as contributions from participating universities and partners, and bringing together activities associated with book history and textual scholarship; user experience studies; interface design; and prototyping of digital reading environments.

Ruecker, Stan et. al. (2012). The Beginning, the Middle, and the End: New Tools for the Scholarly Edition. *Scholarly and Research Communication, 3*(4): 040153, 7 pp.

Introduction

In enhancing the scholarly edition, we mean that we are designing experimental prototypes for a reading environment consisting of a primary text supplemented with scholarly apparatus. In particular, we have framed our work in new digital tools for this kind of edition by pointing out that they are respectively situated within the reading experience at the beginning of the text (the table of contents), the end of the text (the citations), and in the middle of the text. One consequence of this approach is that for this category of design, we are treating the scholarly edition as having a fixed sequence, rather than being comprised, for example, of a set of passages that can be reorganized dynamically — as one might change the alphabetical order of encyclopaedia entries by subject to instead represent alphabetical order by first author's last name, or by some predefined taxonomy of knowledge. That said, several of the tools we are working with do allow the user to dynamically reorganize the material, but always from the basis of a reading environment that continues to support the display of the primary sequential text.

In the first area — near the beginning of the book — we are extending our earlier prototype on the dynamic table of contexts or TOC (Ruecker, Brown, Radzikowska, Sinclair, Nelson, Clements, Grundy, Balasz, & Antoniuk, 2009) to support not only the dynamic insertion by the reader of XML-marked passages, but also the provision of related visualizations such as Bubblelines or frequency graphs (Figure 1). We are also exploring the interactive reorganization of the TOC itself, in accordance with the ideas proposed in Nelson (2011). Finally, we are experimenting with embedding the dynamic TOC within other open source reading environments. The goal of the dynamic TOC is to provide a flexible form of prospect and internal linking that can be tailored by the user.

Figure 1. The dynamic TOC with a Bubblelines panel lower left.
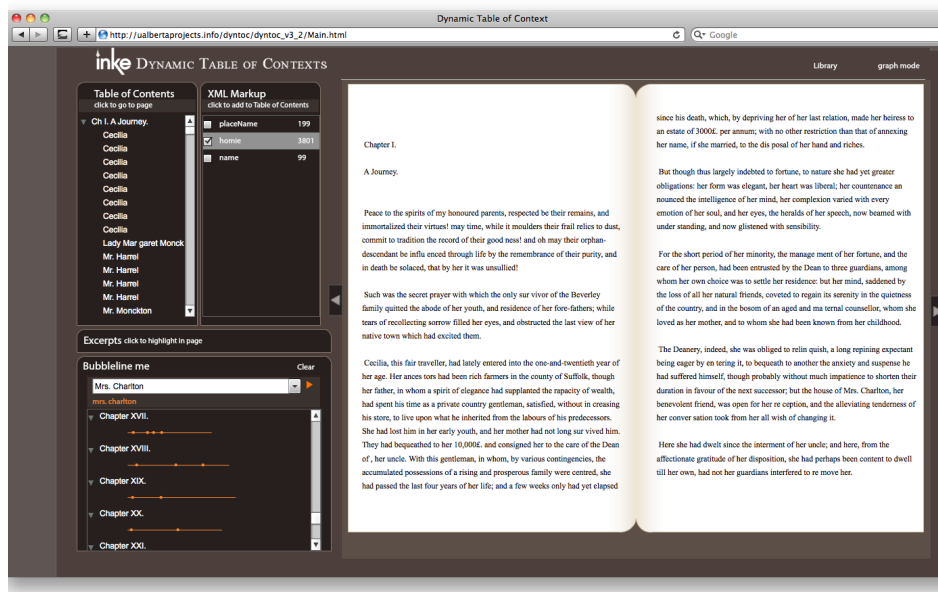
Carlos Fiorentino is a design instructor at the University of Alberta and MacEwan University. He is an interface – data visualization designer and researcher, and co-founder and consultant at pix design+web+multimedia studio, based in Buenos Aires, Argentina. Email: carlosf@ualberta.ca .

Stéfan Sinclair is an Associate Professor of Digital Humanities at McGill University. His primary area of research is in the design, development, usage and theorization of tools for the digital humanities, especially for text analysis and visualization. Email: sgsinclair@gmail.com .

Milena Radzikowska is an Associate Professor in Information Design, Faculty of Communication Studies, Mount Royal University. Email: mradzikowska@gmail.com .

Teresa Dobson is an Associate Professor and the Director at the Digital Literacy Centre at the University of British Columbia. Email: teresa.dobson@ubc.ca .

Ann Blandford is a Professor of Human–Computer Interaction at University College London. Email: a.blandford@ucl.ac.uk .

Sarah Faisal is an Honorary Research Associate at the University College London Interaction Center (UCLIC). Email: s.faisal@cs.ucl.ac.uk

Alejandro Giacometti is Doctoral Research Student in the Departments of Information Studies and Medical Physics and Bioengineering at University College London. Email: alejandro.giacometti.09@ucl.ac.uk.

Susan Brown is Director of the Orlando Project and the Project Leader of the Canadian Writing Research Collaboratory. She is a visiting professor in the Department of English and Film Studies at the University of Alberta, and professor in English and Theatre Studies at the University of Guelph. Email: susan.brown@ualberta.ca .

Brent Nelson is Associate Professor in the Department of English at the University of Saskatchewan. Email: brent .nelson@usask.ca .

Piotr Michura is a lecturer at the Department of Design Fundamentals, Faculty of Industrial Design, Academy of Fine Arts in Krakow, Poland. Email: pmichura@asp.krakow.pl .

For the middle of the book, we have two prototypes in progress. The first, called "dialR for repetition" (Figure 2), uses n-grams to allow the reader to look for patterns of repeated words. N-grams, which are sequences of two or more words (or other text items such as lemmas, stemmed words, or even characters), have recently been popularized through the Google n-gram viewer. The process of using dialR involves typing in a search string beneath one of the radar screens at the bottom of the page, then clicking the go button in the centre of the radar.

What appears as a result is a list of n-grams down the left-hand side of the page. The user can choose how many words are necessary in order for an n-gram to be added to the list — the default is two words. Clicking on one of those words puts the n-gram into the central space of the display, highlighted on what is essentially a volumetric space composed of transparent sheets of text. In the current iteration, the user clicks one of the coloured highlights to call up that passage in the reading panel on the right. In the next iteration, those texts will appear automatically in the reading panel once an n-gram is chosen from the list, with the option to collapse or expand them.

Figure 2. The dialR system provides an exploratory environment for n-grams
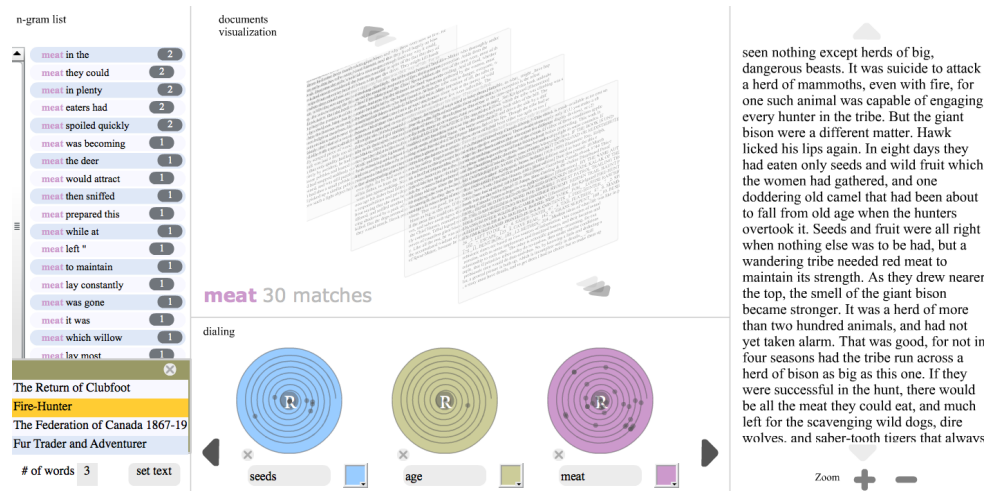
Although we have previously described the dialR prototype (Ruecker, Radzikowska, Michura, Fiorentino, & Clement, 2008) the INKE Year Three iteration is intended to include an improved user interaction model that leverages the idea of the scholarly edition, where multiple kinds of information can interact. On our list of improvements are the following:

CONCEPTUAL
- Embed dialR as one of the available tools to be called up while reading
- Add a microtext location scrollbar on the right of the interface
- The text reading panel should allow the user to expand the text snippets, or else allow the reader to specify how many words to show on the right panel
- The reading panel should not preload, or else if the reading panel does preload, it should scroll as the reader scrolls the plastic sheets

Ruecker, Stan et. al. (2012). The Beginning, the Middle, and the End: New Tools for the Scholarly Edition. *Scholarly and Research Communication, 3*(4): 040153, 7 pp.

- The reader should be able to define stopwords
- We need an easier way to compare two documents
- Can we allow the plastic sheets to emerge for reading?

INTERACTION

- Link the text reader on the right to the items on the left n-gram list
- Make the radar dots interactive to link to the plastic sheets and reader
- Autoscroll the plastic sheets to the currently selected item on the dial
- Add a progress bar
- Could the n-gram list preload with a concordance once the user chooses a text? If so, then the user would need to be able to return to that list
- We need an easier way for users to add a file

BUGS

- Allow the interface to rescale to fit the monitor size
- Stronger colour for the highlights on the plastic sheets
- The thumb in the scrollbar should be scaled to match the amount of text

The second prototype for the body of the document (Figure 3) is a series of small embedded text analysis tools that the reader can activate while reading, in order to be able to examine more closely any claims being made that are supported by text analysis, or else to look at variations of the data (Rockwell & Sinclair, 2008). Although dialR is not currently included on this list, we intend that it will be in future iterations.

Figure 3. This design for the extension of the book reader in its mode of facing pages includes a floating palette of Voyant and other tools
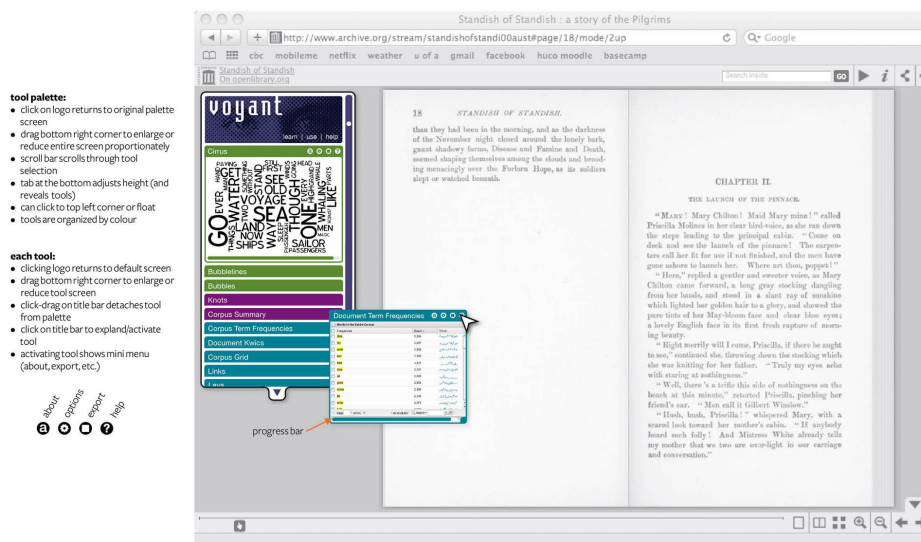
In the current design, the form of the tool palette is in essence a badge (Rockwell, Ruecker, Organisciak, & Sinclair, 2009), with the tools colour-coded into three categories according to the primary affordance they provide: visual, statistical, and fun. In subsequent iterations, we expect to produce additional ways in which the tools can be more tightly associated with the text, so that it is possible, for instance, to apply a tool to selected portions of the document.

Ruecker, Stan et. al. (2012). The Beginning, the Middle, and the End: New Tools for the Scholarly Edition. *Scholarly and Research Communication, 3*(4): 040153, 7 pp.

5

For the end of the book, we are producing a number of interactive visualizations of citations, based on the metadata they contain about time and place (Figure 4), and building on an XML encoding of citation use within the document. The citation visualization system contains a variety of ways of visually expressing citation patterns, whether geographically by publisher, chronologically by date of publication, or by type of use.

Figure 4. This visualization shows a pattern of citations on a collapsible timeline
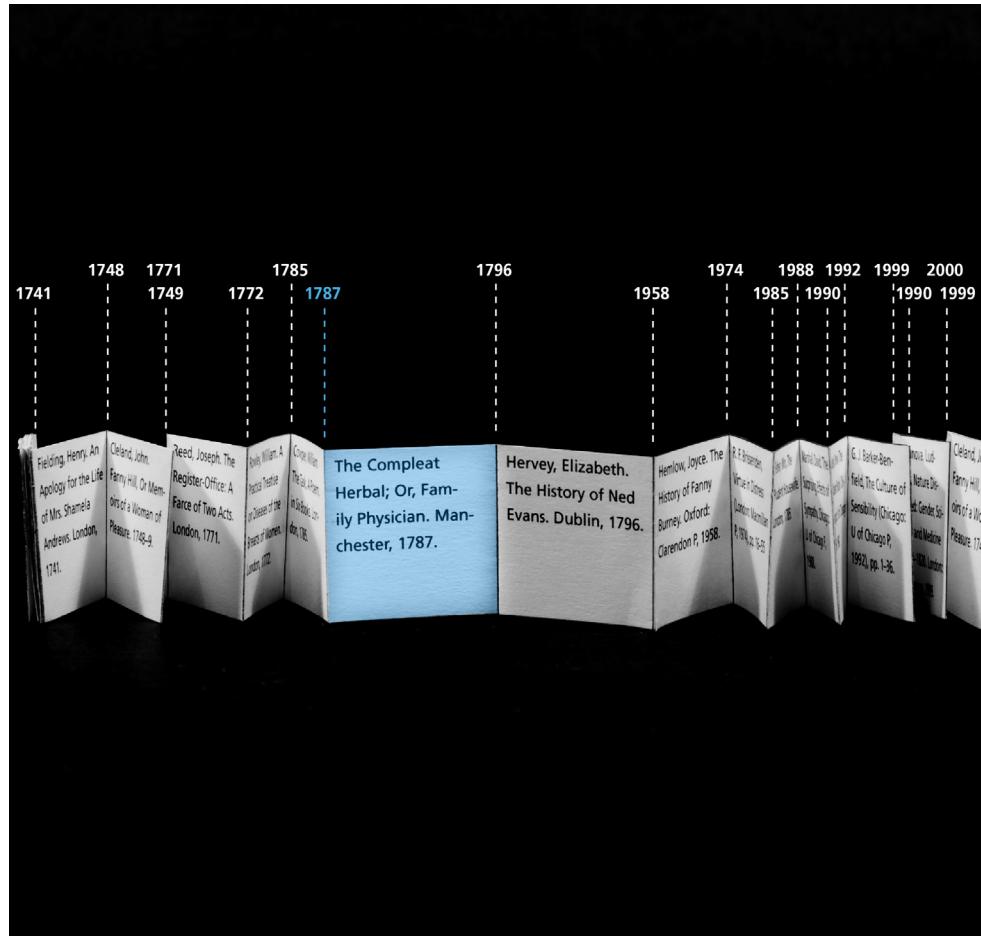
The INKE ID team has begun its activities for Year Three, focusing on experimenting with new knowledge environments for the scholarly edition. However, the process is an iterative one, with many overlapping cycles of design, prototyping, and testing. We anticipate changes, hopefully in the form of improvements, to each of the prototypes described in this article, with the ultimate goal of having the opportunity to provide some of them to an ever-expanding group of users in the digital humanities community and beyond.

References

Nelson, Brent. (2011). *The TOC*. Paper presented at SDH-SEMI 2011 conference. University of Fredericton, NB.

Rockwell, Geoffrey, Ruecker, Stan, Organisciak, Peter, & Sinclair, Stéfan. (2009). *Ubiquitous text analysis*. Paper presented at the Digital Humanities 2009 conference. University of Maryland, MD.

Rockwell, Geoffrey & Sinclair, Stéfan. (2008). *Now Analyze That: Obama and Wright on Race in America*. URL: http://tada.mcmaster.ca/Main/NowAnalyzeThat [June 16, 2012].

Ruecker, Stan, Brown, Susan, Radzikowska, Milena, Sinclair, Stéfan, Nelson, Thomas M., Clements, Patricia, Grundy, Isobel, Balasz, Sharon, & Antoniuk, Jeff. (2009). The Table of Contexts: A Dynamic Browsing Tool for Digitally Encoded Texts. In Lucie Dolezalova (Ed.), *The Charm of a List: From the Sumerians to Computerised Data Processing*. (pp. 177–187). Cambridge, UK: Cambridge Scholars Publishing.

Ruecker, Stan, Radzikowska, Milena, Michura, Piotr, Fiorentino, Carlos, & Clement, Tanya. (2008). Visualizing Repetition in Text. Digital Studies/Le champ numérique. Reprinted from *Reassembling the Disassembled Book: Symposium of the Congress of the Humanities and Social Sciences*, University of Saskatchewan, May 2007. Brent Nelson, (Ed.). Special Issue of CH Working Papers.

Ruecker, Stan et. al. (2012). The Beginning, the Middle, and the End: New Tools for the Scholarly Edition. *Scholarly and Research Communication, 3*(4): 040153, 7 pp.

7