Interface Design

Stan Ruecker *(sruecker@ualberta.ca)*
University of Alberta
Stéfan Sinclair *(sgsinclair@gmail.com)*
McMaster University
Stephen Ramsay *(sramsay@uga.edu)*
University of Georgia
Milena Radzikowska *(mradzikowska@mtroyal.ca)*
Mount Royal College
Alan Galey *(agaley@uwo.ca)*
University of Western Ontario

Unsworth suggests that humanities computing has progressed during the last thirty years through three phases, beginning with the design of text analysis tools, moving to the design of online digital collections, and now returning to the development of a new generation of tools for working with online materials. While many of the first generation of text analysis tools focused on the analysis of a single text, the emphasis in many of the new tools is on data mining and visualization of results across sets of documents or even entire collections.

There are several distinct research agendas involved, including interest in making digital materials more accessible (Lyman), in providing new affordances for doing synthetic research at the level of the interface (Ruecker), and in using visualizations of digital information not only for analysis but also for further access to subsequent information (Flanders).

This session brings together three papers on visualization and interface design. First is Ruecker, Sinclair, and Radzikowska's "The Aesthetic Function: The role of visual communication design in interface research", which examines the role of graphic design in interfaces to digital collections and visualization interface research. Their conclusion is that

> Careful attention to the details of graphic presentation can have a significant impact on the perceived value of a digital collection, the function of a visualization system, the research results available from analysis of visualizations, and the dissemination of findings both within the academic community and for the larger public audience.

Second is Alan Galey's "'Alms for Oblivion': Bringing an Electronic New Variorum Shakespeare to the Screen". Galey emphasizes the importance of W3C advanced standards compliance for the delivery of academic collections online. Interfaces to collections like the eNVS call for digital adaptations that can make traditional scholarly apparatus more manageable:

> The value of the eNVS interface lies in reinventing such fundamental scholarly mechanisms as the textual collation line, the commentary footnote, and the annotated page — three structures from which the variorum derives both its archival power and, for many print users, its aura of cognitive overload.

The third paper is Ramsay's "Mining Shakespeare", which discusses the results and the implications of using the *StageGraph* and *D2K* software systems for semi-automatically mapping scene changes in Shakespeare's plays. Ramsay comments not only on the results of this study, but also on the implications of data mining in humanities scholarship:

> Though it may be used to support or refute hypotheses, data mining is far more useful in the service of the broad humanistic mandate to find new and insightful ways of looking at textual artifacts.

These presentations provide a cross-section of current research on visualization in the humanities, moving, as does the research itself, between intricate detail and broad theoretical principles.

Bibliography

Flanders, Julia. "Text analysis and the problem of pedantry." Paper delivered at CaSTA 2004: The Face of Text. 3rd conference of the Canadian Symposium on Text Analysis, McMaster University, Hamilton, Ontario. November 19-21 2004. 2004.

Lyman, Eugene W. "In pursuit of radiance: Report on an interface developed for the Piers Plowman Electronic Archive." Paper delivered at CaSTA 2004: The Face of Text. 3rd conference of the Canadian Symposium on Text Analysis, McMaster University, Hamilton, Ontario. November 19-21 2004. 2004.

Ruecker, Stan. *Affordances of Prospect for Academic Users of Interpretively-tagged Text Collections.* PhD. Dissertation, Edmonton: University of Alberta, 2003.

Unsworth, John. "Forms of Attention: Digital Humanities Beyond Representation." Paper delivered at CaSTA 2004: The Face of Text. 3rd conference of the Canadian Symposium on Text Analysis, McMaster University, Hamilton, Ontario. November 19-21 2004. 2004.

The Aesthetic Function: The role of visual communication design in interface research

Stan Ruecker, Stéfan Sinclair, Milena Radzikowska

Since the professionalization of various humanities disciplines in the latter part of the nineteenth century, humanities scholars have been primarily occupied with the interpretation and analysis of existing cultural artifacts, such as texts. The expertise and artistry required to produce the material form of the objects are generally outside the scope of a humanities education; print-making, for instance, is a separate craft. However, since the rise of personal computing and graphical interfaces, many humanities scholars have been empowered to create the interface through which their materials can be studied: the proliferation of digital collections such as *Perseus* and *Rossetti* bears witness to this phenomenon. The distinctions between author, critic, editor and publisher have blurred. Significantly though, the knowledge and perspective of artists and designers have either largely been ignored by the digital humanities scholar, or else have contributed in a manner that has not been subject to direct analysis. Similarly, the growing interest in visualization systems for the humanities is another research area where design issues are relevant (Bradley and Rockwell). The significance of the visual is sufficiently evident in all of these cases that aesthetic factors become intrinsically woven with issues of functionality. Research interests in graphic design and presentation find a new relevance and weight, not only as a contributing factor in the design of computer interfaces and visualization systems, but also as an area of study in their own right.

We address the issue of graphic design contributions to visualization research by making reference to several recent interface design research projects at the University of Alberta and McMaster University. The emphasis is on the functional differences between early and later prototypes, including an analysis of what Frascara calls "the aesthetic function of design" . We argue that aesthetic function is a composite that includes attracting viewers, holding their attention, and compelling their trust and respect. Design, in other words, is of utmost importance to the value and legitimacy of scholarly digital content.

Our first example is from the *TouchGraph* representation of XML data, where initial designs included boxes around each of the individual text items. Although this form of display is commonly used with topic maps, it is also unreasonable for several reasons: it draws the reader's eye to locations that are not particularly meaningful; it introduces unnecessary clutter; and it misappropriates a grouping affordance for a single element, which does not require grouping.

A second example is a prototype system for blocking and reading plays, called *Watching the Script* (Ruecker et al.). The interface design has gone through three distinct stages, beginning with a white parallelogram, moving to a four-colour square, and ending with a very attractive full-colour combination of stage, playback controls, and large coloured dots that clearly indicate character positions. Attention to the details of the graphic design is in this case intrinsically related to the details of the functions of the system, such as the location and movement of characters and text. However, the list of additional qualities would not be complete without acknowledging that part of the attraction of the most recent iteration is the aesthetic function. In its extreme form, this value can result in forms of interface that are in some senses autotelic — they can become an end in themselves for some users, who find their attractions sufficient to make the system worth further attention, outside the context of any particular research task.

The connection between graphic design and academic research also has implications for the ongoing need for improved communication between the academic world and everyone else. Several strategies are required at different levels, including public information campaigns, academic contributions to popular media, and a more significant presence of the academic in the community. One potential role that design has to play is in visually rewarding the reader of research results. However, there are barriers to be overcome, not least of all within the academy. It might even be argued that there is an anti-aesthetic subtext in certain research areas, since effort to engage readers through visual appeal (and its related functionality) might be understood as devaluing more essential research outcomes. However, Pujol points out that the visual qualities of professional design are one of the key signifiers by which we distinguish the individual voice from the institutional. If someone hand letters a sign to advertise a garage sale, we understand the sale as an amateur activity. If that same person employs graphic design skills and produces a glossy poster, we may interpret the same event, at least until we arrive at the site, as the establishment of a new retailer.

Karvonen takes this line of reasoning even further in her study of the relationship between trust and design. The cohort for her project was Scandinavian, with participants from both Finland and Sweden. The long-standing cultural awareness of design quality in those countries is probably a factor in her findings that people tended to find that web sites with a clearly professional design quality were rated as being more trustworthy than more vernacular sites. It would be indefensible to suggest that a professional standard of visual communication design could contribute to the perceived reliability of research results, since there are other, more important indicators that are applicable. However, it is not outside the domain of the possible that graphic quality is potentially a contributing factor not only in the evaluation of research results, but, particularly in the areas of visualization and information design research, also in the results obtained from user study. Careful attention to the details of graphic presentation can have a significant impact on the perceived value of a digital collection, the function of a

visualization system, the research results available from analysis of visualizations, and the dissemination of findings both within the academic community and for the larger public audience.

Bibliography

Bradley, John, and Geoffrey Rockwell. "What scientific visualization teaches us about text analysis." *Consensus ex machina? ALLC-ACH 94 abstracts.* Paris, 1994. 203-204.

Frascara, Jorge. *User-centred Graphic Design.* London: Taylor and Francis, 1997.

Karvonen, Kristiina. "The beauty of simplicity." *Proceedings on the 2000 conference on Universal Usability.* November 2000. 99-99.

Pujol, Monica. "Design as a social practice." Public lecture. Trans. Jorge Frascara. Department of Art and Design, University of Albert, 1 Nov 2001.

Ruecker, Stan, Eric Homich, and Stéfan Sinclair. "Watching the Script of Synge's Playboy of the Western World." Paper delivered at the COCH/COSH. Congress 2004. Winnipeg: University of Manitoba. 2004.

Unsworth, John. "Forms of Attention: Digital Humanities Beyond Representation." Paper delivered at CaSTA 2004: The Face of Text. 3rd conference of the Canadian Symposium on Text Analysis, McMaster University, Hamilton, Ontario. November 19-21 2004. 2004.

"Alms for Oblivion": Bringing an Electronic New Variorum Shakespeare to the Screen

Alan Galey

In this paper, I describe the development of a digital interface for the *Electronic New Variorum Shakespeare (eNVS)*, and explore certain historical and technical issues that bear upon our design strategies. With the considerable burdens of content development and encoding resting on others' shoulders, *NVS* co-general editor Paul Werstine and I have been free to focus exclusively on interface design – an area of humanities computing that, I would suggest, has not kept pace with advancements in web browsers and third-party design standards. Although many computing humanists might see an *eNVS* edition as a document, however complex, we have instead approached it as a data object, where the organizing logic is that of object-oriented programming, not hypertext. The value of the *eNVS* interface lies in reinventing such fundamental scholarly mechanisms as the textual collation line, the commentary footnote, and the annotated page – three structures from which the variorum derives both its archival power and, for many print users, its aura of cognitive overload. With these issues in mind, I will argue that in order to bring electronic editing projects like the *eNVS* to the screen, humanists must also be information architects, who think past documents to embrace the principles of object-oriented and standards-compliant programming and design. Conversely, the historical section of this paper will show that the programmers must also be humanists, who understand the cultural and bibliographical histories of the interface traditions in which they work.

Scholarly opinion differs on the present value and future viability of Shakespeare variorum editions, print and electronic, but tends to agree that, in any form, they rank among the largest information-management projects in Shakespeare scholarship. On one side of the debate, Richard Knowles goes so far as to call variorums "the memories of the profession" (43), though he also stresses that variorums, like all memory, incorporate a principle of selection in their management of heterogeneous masses of data. Maurice Hunt takes issue with the traditional perception of the variorum edition as "the still point in the turning world of texts, a text which would arrest, and even reverse, the processes of textual change and corruption" (62, quoting McGann 93) – a view that consigns these editions to the "tombs/tomes" of an "obsolete modernism" (Hunt 62). Instead, Hunt contends, the variorum structure anticipates postmodernist values in the heterogeneity of its apparatus, which conveys the indeterminacy of the Shakespeare text more than any other kind of edition. Yet for some scholars on the other side, projects like the *NVS* are more about the past than the future, amounting to "admission[s] of failure" and monuments to unachieved textual stability (Rhodes and Sawday 11; Bristol 101). In one instance of pointed criticism, John Lavagnino claims that the uncategorized nature of variorum commentary renders it un-digital in advance, and "not productively open to flexibility of display" (201). He concludes with a call for improvements in display technology (203), which our project echoes and in part hopes to answer. As this range of thought indicates, the challenges facing an electronic variorum are not purely technical, and require a level of interface design that accounts for the historical issues at stake.

This paper has four sections:

1. History: the variorum interface in print
2. Possible futures: the web browser as design platform
3. Examples of the *eNVS* interface
   (a) Textual apparatus
   (b) Page/screen layout(s)
   (c) Annotations
4. Conclusion: alms for oblivion

1. History: the variorum interface in print

This section will provide a brief outline of the design challenges we have inherited from eighteenth-century editors. Interface issues have dominated the variorum's historical role in Shakespeare studies since Samuel Johnson first applied the format to Shakespeare in 1765. Part of our research mandate is to reinvent the complex layout of the Shakespeare variorum, which has remained largely unchanged – and unloved, many Shakespeareans would say – for over two centuries. As recent scholarship on editing's cultural history shows, the reservations expressed by Bristol and Lavagnino are as old as the Shakespeare variorum itself (see DeGrazia 209-14, and Gondris). I will confine my focus here to the historical problem of too much (Shakespearean) information, which casts a long shadow over any interface design.

2. Possible futures: the web browser as design platform

The *eNVS* is an interface with a 200-year history, and with an eye to the present moment of standards-friendly design in the wake of the so-called browser wars between Microsoft and Netscape. If Bristol is correct that the *NVS's* goals exceed the limits of print (101), and if Lavagnino is correct that complex digital commentary is insufficiently served by hypertext alone (198-200), we might conclude that an adequate *eNVS* interface demands advancement beyond traditional, HTML-era design. I will briefly summarize the case for advanced browser-based interfaces, with particular reference to the W3C's standardization of key web technologies such as CSS, XML, and the DOM – and, most importantly, the implementation of these and other standards in 'postwar' open-source browsers such as Mozilla and Firefox.

3. Examples from the eNVS interface:

3.a Textual apparatus

Known by such tongue-in-cheek epithets as the "band of terror" or "barbed wire" that runs beneath the text (Thomas Berger and Edmund Wilson, quoted in Rasmussen 211), the traditional collation of variants offers the most obvious candidate for a digital reconception. Where many electronic editions at best display variants by way of linked parallel texts (swapping one print interface for another), and at worst simply recode the collation line as a text string, the *eNVS* interface instead generates a properly machine-readable apparatus by means of object-oriented scripting. This allows us to expand the collation line, textually and graphically, into the textual history it compresses and encodes.

3.b Page/screen layout(s)

As Gondris has shown, the Shakespeare variorum page inherited from the eighteenth century constitutes a critical structure that promotes some habits of thought and suppresses others. The consequences of rearranging it will therefore reach beyond readability and convenience – important enough issues in themselves – to impact the production of meaning. This is one of the most challenging aspects of the *eNVS*, not least because of the computer screen's orientation toward vertical scrolling. Again, an object-oriented interface enables multiple layout options without generating redundant files.

3.c Annotations

The question of how best to display electronic annotation remains a central debate in electronic editing. It is also a central concern in our project, since the variorum's primary content is not its playtext, but its notes. But as Lavagnino has pointed out, it is difficult for digital interfaces to improve upon – or even match – the cognitive elegance of the print reader's glance from text to footnote (198-9). This section will demonstrate how our note design works with the page layouts, and with the complex archive formed by the network of *NVS* annotations.

4. Conclusion: alms for oblivion

The paper will conclude with a restatement of the argument for closer integration of textual studies and web programming in the practice of electronic editing, especially in projects like the *eNVS*. Much of the energy that might advance interface design in humanities computing is presently devoted to digitization and tagging, in response to the archival impulse still strong in the humanities. As an invitation to discussion, I will close by reflecting on Knowles's quotation of Shakespeare's *Troilus and Cressida*, which he uses to make the point that all scholarship risks becoming "alms for oblivion, ... good deeds past, which are devoured / As fast as they are made, forgot as soon as done" (3.3.141-4, quoted in Knowles 39). Shakespeare variorums are akin to all electronic preservation formats in that they attempt, in Knowles's words, "to guard against oblivion," even as they are subject to it. The *eNVS* seeks to preserve scholarship into the future by increasing its accessibility and relevance in the present.

Bibliography

Bristol, Michael D. *Shakespeare's America, America's Shakespeare.* London: Routledge, 1990.

De Grazia, Margreta. *Shakespeare Verbatim: The Reproduction of Authenticity and the 1790 Apparatus.* Oxford: Clarendon P, 1991.

Gondris, Joanna. "'All this Farrago': The Eighteenth-Century Shakespeare Variorum Page as a Critical Structure." *Reading Readings: Essays on Shakespeare Editing in the Eighteenth Century.* Ed. Joanna Gondris. Madison, WI: Fairleigh Dickinson UP, 1998. 123-39.

Hunt, Maurice. "New Variorum Shakespeares in the Twenty-First Century." *Yearbook of English Studies* 29 (1999): 57-68.

Knowles, Richard. "Variorum Commentary." *TEXT* 6 (1994): 35-47.

Lavagnino, John. "Two Varieties of Digital Commentary." *Textual Performances: The Modern Reproduction of Shakespeare's Drama.* Ed. Lukas Erne and Margaret Jane Kidnie. Cambridge: Cambridge UP, 2004. 194-209.

McGann, Jerome J. *A Critique of Modern Textual Criticism.* Charlottesville, VA: UP of Virginia, 1983.

Rasmussen, Eric. "Richly Noted: A Case for Collation Inflation." *Arden: Editing Shakespeare.* Ed. Ann Thompson and Gordon McMullan. London: Arden Shakespeare-Thompson Learning, 2003. 211-8.

Rhodes, Neil, and Jonathan Sawday. "Paperworlds: Imagining the Renaissance Computer." *The Renaissance Computer: Knowledge Technology in the First Age of Print.* Ed. Neil Rhodes and Jonathan Sawday. New York: Routledge, 2000. 1-17.

Mining Shakespeare

Stephen Ramsay

"[T]he computer," writes Susan Hockey in her 2000 book *Electronic Texts in the Humanities*, "is best at finding features or patterns within a literary work and counting occurrences of those features" (Hockey 66). For many areas of inquiry, such finding and counting is eminently useful. Word-frequency analysis in the context of computational linguistics, concordance generation as a prelude to the study of word usage, and the various search functions that have now become an ordinary part of the task of research in so many disciplines represent clear examples of the utility of computational tools. For scholars engaged in the task of literary critical interpretation, however, such finding and counting can seem beside the point. As Hugh Craig put it:

> The leap from frequencies to meanings must always be a risky one — Lower-level features are easy to count but impossible to interpret in terms of style; counting images or other high-level structures brings problems of excessive intervention at the categorization stage, and thus unreliability and circularity.
>
> (Craig 103)

The risk of which Craig speaks is not merely a matter of interpretive caution. In most cases, low-level features simply don't assert themselves in any obvious way into the broad, complex patterns upon which literary critical interpretation depends.

Data mining provides a suggestive set of methods for bridging the gulf between low and high. It has its roots in a number of statistical techniques with venerable histories in digital humanities (in particular, the use of factor analysis in the study of literary and philosophical texts)[1], but introduces an exploratory dimension far more conformable to the elaborate task of prompting meaningful critical insight. Data mining techniques operate on low-level features, but use a variety of statistical and logic-programming methods to discern broad complex patterns in the data set (such as classifications, categorizations, and prediction models) that are not conceived in advance. In other words, data mining lets us ask what's interesting about an apparently disparate set of low-level features without having to form any concrete expectations in advance.

This paper presents research on the structure of Shakespearean drama and its relation to genre categorization using two programs: *StageGraph* and *D2K*. *StageGraph* generates directed graph visualizations of scene changes and character movements from XML representations of plays (figure 1). It can also use the generated graphs to produce matrices of individual graph properties (e.g. degree number, number of cycles, diameter, chromatic number). While such graphs and matrices provide fertile ground for interpretive reflection, the utility of the graphs is greatly enhanced when the generated properties are themselves 'mined' for broad patterns and features, and the results presented in the form of an open-ended visualization.

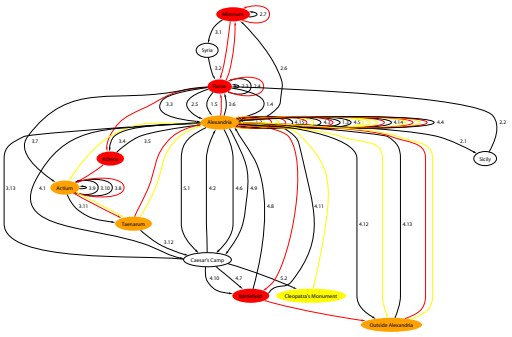

Figure 1

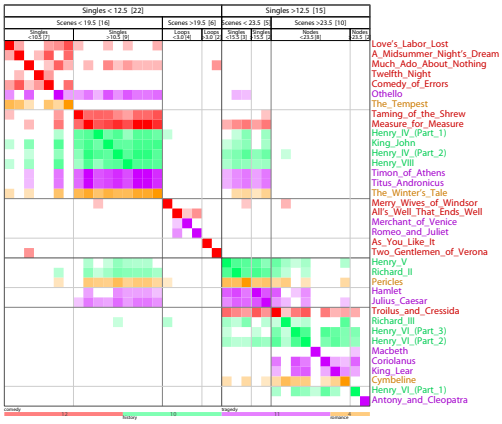

Figure 2

The author and a research assistant at the National Center for Computing Application used the *D2K* software to conduct naive Bayesian analysis and decision tree generation (two standard data mining techniques) on the graph property matrices to try to see if the low-level structural features of Shakespeare's plays assemble themselves into clusters that correspond to the traditional genre categories of comedy, tragedy, history, and romance. We then used a technique pioneered by one of the project participants that combines concept-tree clustering with shaded similarity matrices to generate a visualization of the degrees of similarity among Shakespeare's plays in terms of genre (figure 2). The results are extremely suggestive, in that the visualization not only groups the plays broadly into traditional generic categories, but also contains anomalies that correspond to some of the more influential insights into Shakespearean genre (for example, Susan Snyder's argument that *Othello* possesses the basic structure of comedy appears to be confirmed by the data mining algorithm).

Most interesting of all, however, are the anomalies that represent neither traditional classification nor the product of exegetical insight. For it is here, I believe, that data mining and visualization present the most promise for text analysis practitioners in literary study. Though it may be used to support or refute hypotheses, data mining is far more useful in the service of the broad humanistic mandate to find new and insightful ways of looking at textual artifacts.

Bibliography

Bradley, John, and Geoffrey Rockwell. "Watching Scepticism: Computer Assisted Visualization and Hume's Dialogues." *Research in Humanities Computing* 5 (1996): 32-47.

Burrows, J.F., and D.H. Craig. "Lyrical Drama and the 'Turbid Montebanks': Styles of Dialogue in Romantic and Renaissance Tragedy." *Computers and the Humanities* 28 (1984): 63-86.

Craig, Hugh. "Authorial Attribution and Computational Stylistics: If You Can Tell Authors Apart, Have You Learned Anything About Them?" *Literary and Linguistic Computing* 14 (1999): 103-13.

Hockey, Susan. *Electronic Texts in the Humanities.* Oxford: Oxford University Press, 2000.

Synder, Susan. *The Comic Matrix of Shakespeare's Tragedies: Romeo and Juliet, Hamlet, Othello, and King Lear*. Princeton: Princeton University Press, 1979.

Wang, Jun, Bei Yu, and Les Gasser. "Concept Tree Based Clustering Visualization with Shaded Similarity Matrices." *Proceedings of the 2002 IEEE International Conference on Data Mining (2002).* 2002. 697-700.

---

1. cf. J. F. Burrows and D. H. Craig's studies of Romantic and Renaissance tragedy and John Bradley and Geoffrey Rockwell's use of cluster analysis for the study of Hume's *Dialogues*, op cit.