

The Digital Word: Text-Based Computing in the Humanities

George P. Landow and Paul Delany (editors)
(Brown University and Simon Fraser University)

Cambridge, MA: The MIT Press
(Technical Communication and
Information Systems Series, edited by
Edward Barrett), 1993, xii + 362 pp.
Hardbound, ISBN 0-262-12176-X, \$39.95

Reviewed by
Mark Olsen
University of Chicago

The sixteen essays assembled by Landow and Delany in this volume cover a broad array of topics connected by a concern with "text-based computing" that have applications to the "humanities." The unifying theme in the collection is, to a certain extent, an attempt to assess the impact that the rapid proliferation of networked text corpora and digital communications is having (or might have) in text-oriented disciplines. The editors suggest that networked computing facilitates analysis of very large corpora, or a *docuverse* (p. 23), which they contend is a vital component of research in the humanities. This does not mean that researchers can simply scale up their methods to meet the demands and possibilities of new technical capabilities. Rather, they suggest that humanists must

confront a boundless universe of text and . . . draw from it elements needed for a given project. . . . [T]ext-based computing implies openness to massive amounts of text and a recognition that we require new procedures to keep from ineffectual wanderings in the labyrinth. (pp. 23–24)

While the network opens up new resources and problematics to humanities scholarship, the editors are equally convinced that the "digital wor(l)d" is changing the nature of scholarly publishing, communications, and writing. Thus, in addition to an examination of computer-aided research in the humanities, Landow and Delany offer essays looking at the pragmatics of electronic mail and conferences, the impact of computer technology on writing, and how to navigate this new (should I say yet another) "boundless universe." As might be expected, the editors are only partly successful in these ambitions.

The essays reflect a clear tension between describing the potentials (both methodological and theoretical) of this new networked universe of digital text and providing descriptions of the current state of the art in the humanities. For example, the first essay outlines the TGB (Très Grande Bibliothèque) effort of the Bibliothèque de France, which envisages the digitization and electronic consultation of the holdings of this great library. It is followed, however, by two essays describing commercial and public-domain text management tools aimed at personal computers and workstations, which provide little new information on navigating anything much more than your own hard disk. Similarly, Jeremy Clear's discussion of the British National Corpus or Paul Delany's personal *docuverse* bear little conceptual or theoretical relationship to, for example, Ian Lancashire's very solid essay on Atwood's *The Handmaid's Tale* or

David Miall's proposal of computer-aided examinations of reader responses to literary texts. The gap between current practice of textual computing in the humanities and the "boundless universe of text" being deployed on the networks arises, to a certain extent, from the fact that few scholars in the humanities have integrated networked information systems into the practice of computer-aided textual analysis. An important part of the problem, however, arises from the fact that the digital wor(l)d challenges the traditional notions of textuality that have informed the humanities.

In the paper world of the humanities scholar, a text is simply a fixed organization of writing, often finding its limits at the covers of the printed volume. Landow points out that text-based computing in the humanities relies on "electronic simulacrum" of printed editions (p. 239) retaining the notions of textuality inherited from print and paper. Delany, however, argues that "inside the docuverse world texts are dematerialized—which means that they are freed from the constraints that accompany embodiment" (p. 189). He continues by pointing out that the entire corpus of French literature, held in the Trésor de la langue française database may be consulted as a single entity or as dynamically defined subsets. In this context, the traditional categories of text based on a title or an author are simply virtual subsets of the querying software, which are equally as valid as subsets defined as "all the texts published in the 18th century" or all documents containing units identified as correspondence found in novels and collections of letters. The distinction between text and context, in this dematerialized environment, requires serious rethinking, since querying software so readily shatters the traditional divisions. Furthermore, by discarding its paper referent, Delany's dematerialized text can no longer be considered a simple typescript, representing the physical characteristics of a printed volume. Coombs, Renear, and DeRose argue, in "Markup Systems and the Future of Scholarly Text Processing," that digitized text must be conceived of as a logical structure of textual components, which they call "descriptive markup." Schemes to represent the dematerialized text, such as the Text Encoding Initiative Guidelines (based on SGML), make a critical distinction between the logical content of textual structure and a particular instance of a fixed representation on paper. Thus, the docuverse dissolves many of the fixed points that guide humanities research, by suggesting that particular instantiations of text are arbitrary constructs. Alan McKenzie summarizes the statelessness of the docuverse:

Electronic texts, having been rendered independent of ink, paper, agent, and human memory, are always already in a state of excitement. They weave their way along lines that are largely inexplicable and only partly determinate. Texts that arrive over a network have no margins to speak of; they present themselves to every gaze in a hierarchy yet to be established, and even then they are infinitely subject to revision. (p. 201–202)

If the digital wor(l)d has challenged traditional notions of textuality, it also bypasses the established mechanisms of communications in the humanities. The explosion of LISTSERVs and news groups dedicated to themes or issues in the humanities, and the widespread distribution of electronic texts over the network, challenges the notion of publication. Landow's essay on "Electronic Conferences and Samisdat Textuality" is an important effort toward understanding the pragmatics and anthropology of this new communications medium. Landow suggests that electronic conferences retain the informality, rapidity, and spontaneity of oral communication, which results in nonlinearity, democracy, and open-endedness—all good things, which, in Landow's view, come at the cost of overwhelming masses of information and floods of verbiage.

The distinction between personal correspondence and publication, in this environment, break down completely, because postings to the network can have, as Renear and Bilder suggest, extremely wide readerships. With democratization of communications and low-cost network publication come related issues of quality assessment, in terms of the reliability of network resources (such as text files and databases) and the caliber of discussions found on lists.

The main thrust of *The Digital Word* makes it an important collection for anyone involved in text-oriented disciplines, for it is clear that computer-mediated textuality in its many forms is fast becoming an issue for both critical research, by opening up new forms of textuality, and the communication of research results. While not all of the essays live up to this large task, the collection as a whole makes an important contribution to understanding the implications and potentials of the digital wor(l)d.

Mark Olsen is Assistant Director of the ARTFL Project at the University of Chicago and Technical Review editor of *Computers and the Humanities*. A French Revolutionary historian, his research interests include computer-assisted analysis of political and social discourses, reception of texts in the 18th and 19th centuries, and philosophy of history. Olsen's address is: ARTFL Project, Department of Romance Languages, University of Chicago, Chicago, IL 60637; e-mail: mark@gide.uchicago.edu.