Charting the Frontier: The Electronic Literature Directory

Robert Kendall
Electronic Literature Organization
E-mail: kendall@wordcircuits.com

Nick Traenkner
Electronic Literature Organization
E-mail: nick@pixlo.com

ABSTRACT: The Electronic Literature Directory is an extensive descriptive directory with highly flexible searching and browsing features. It was designed to meet the unique bibliographic and typological challenges presented by electronic literature. This paper discusses those challenges and our programming and administrative responses to them.

KEYWORDS: Directories, databases, electronic literature, new media literature, Electronic Literature Organization

INTRODUCTION

In the excitement of the Internet publishing revolution, it is easy to overlook that the very qualities engendering the boom are also hampering the Web's maturation as a publishing vehicle. The ease of entry into Web publishing is paralleled by the ease with which publishers can change hosting services, and hence URLs. The minimal investment required to set up shop, the liberating lack of centralized distribution, and the amenability to experimentation have all contributed to a lack of standards for organizing Web sites, for providing essential bibliographic data, and for handling versioning issues.

To help readers make sense of the Web's intimidatingly vast, variable, and often-disorganized literary riches, a host of online directories have taken shape. Most of these, however, provide minimal detail in their listings, offer limited searching and browsing options, and are not always kept up to date. Few of these resources give comprehensive coverage to literature that exploits electronic elements such as interactivity and multimedia, despite the growing importance of this new media writing.

The Electronic Literature Organization's Electronic Literature Directory (ELD) attempts to fill these gaps with its extensive, detailed, and easily searchable database of new media writing [1]. More than any other resource currently available, it provides centralized access to new media literature without sacrificing the advantages of decentralized Web distribution. It is one of the few literary resources capable of providing all the powerful searching, sorting, and data integrity features of a normalized relational database, and a forthcoming multilingual interface will further set it apart. The ELD meets the typological and terminological challenges posed by electronic literature. It also attempts to resolve the unique bibliographic problems posed by the medium—problems that, if left unresolved, threaten to delay universal acceptance of new media writing, limit the amount of critical and scholarly attention it receives, and hamper the growth of its readership.

The ELD catalogs over 2,100 works, 1,000 authors, and 150 publishers. It focuses exclusively on poetry, fiction, drama, and nonfiction of literary interest. All listed works contain electronic enhancements such as interactivity, multimedia, or text generation. To ensure the comprehensiveness and accuracy of its contents, it combines the work of paid data-entry assistants with contributions from volunteer editors and the listed authors and publishers themselves.

DATABASE STRUCTURE

Most extensive, regularly updated directories of literature online take a decidedly minimalist approach to their listings. For example, the literary listings in the Open Directory Project [2], the Yahoo! Directory [3], and About [4]—as well as those in most smaller directories that specialize in literature—provide nothing more than a URL and a descriptive sentence or two for each author, work, or publisher they catalog. These resources generally store the information for each entry in a single text field, which limits the potential for flexible searching, sorting, and categorization and can create problems with data accuracy and consistency.

The Rhizome ArtBase, a large directory of new media art and literature, includes more description about each entry and separates different types of data into separate fields, but you cannot sort or group on different fields and the directory's search features are limited [5]. John Labovitz's e-zine-list breaks entries into separate fields and includes a keywords field that allows selecting on keywords, but this resource is no longer maintained [6].

As much as possible, the ELD adheres to E.F. Codd's principles of normalized relational database design, which have long been acknowledged to maximize data accuracy and consistency while improving database usability [7]. Essentially this entails providing a separate field for each piece of data, and whenever possible allowing data to be entered by selecting from a list of predefined values. For example, rather than typing "Hypertext poetry for Flash" into a field labeled "Description," a user selects "hypertext" from a list of electronic techniques, "poetry" from a list of genres, and "Flash" from a list of plug-ins. This makes data entry easier and allows the collection of more-specific information; drastically minimizes data entry errors; ensures consistency of terminology across listings; allows much greater flexibility in searching, sorting, and categorization; facilitates localization; and allows data to be displayed in different "views" with varying levels of detail or with focuses on different types of information.

TYPOLOGY AND USABILITY

The heart of any directory is its categorization system. First the directory editors must decide what to include and what not to include. Then they must decide upon the categories within which to organize the listings in order to give the directory a meaningful structure. Of course the works listed in the ELD can be sorted by author or publisher, but additional methods of organization are necessary to maximize its usefulness.

As FitzPatrick points out, when it comes to categorization, "directories are a trade off between precision and effectiveness" [8]. Categories must be generalized and given easily understandable labels in order to be useful, but this comes at the expense of some precision. There's always the danger that some entries will have to be shoehorned into the category that provides the closest but not an exact fit.

In many typologies there can be a gray area rather than a clearly demarcated line defining the boundaries of each category. This is particularly true in an often-experimental field like electronic literature where categories are often defined by newly emergent and rapidly changing technologies or by unconventional new uses of existing technologies. To further complicate things, authors of e-literature may deliberately attempt to subvert conventional categorization in their work.

The terminology associated with e-literature is still in flux. The term "hypertext" has been defined in a variety of different ways, so that a particular work will be "hypertextual" by one definition but not by another. A case in point is Morrissey's *The Jew's Daughter,* in which clicking on the "links" changes parts of the text in the current page instead of loading a new page [9]. Aarseth has pointed out the problems of ambiguity associated with the term "interactive," which is widely applied to many types of e-literature [10]. Unfortunately, alternative and perhaps more accurate terms (such as Aarseth's "ergodic") are not at all widely understood and are therefore likely to heighten confusion rather than alleviate it. Even the terms "electronic," "new media," or "computer mediated" mean different things to different people when applied to literature. The ELD therefore attempts to use the most commonly accepted terminology and provides introductory pages that carefully define all the terms we use for categorization.

We chose to define electronic literature as work in computer-readable format that could not be published in print without sacrificing or altering significant elements that depend upon the electronic medium. A work must have one or more of the following to qualify for inclusion in the ELD: interactive elements, multimedia elements (audio, video, or animation), algorithmically generated text, or prominent visual elements.

It could be argued that prominent visuals don't legitimately constitute an electronic element, since often these can be reproduced in print. In practice, though, many of the visual elements in Web work would be prohibitively difficult or expensive to print, and many of the odd window proportions and sizes found on the Web would simply defy the printed page. In a similar spirit of wanting to err on the side of inclusiveness, we have also chosen to include digitized audio and video recordings of performances. Though these are often based on old analog recordings that couldn't be considered "new media," the publication of written text and audio together was rare prior to the Web.

For the purposes of typology, the most salient characteristics of a work of electronic literature will usually be its genre and the electronic elements it uses. We therefore decided to use these two characteristics as our primary basis for categorization. Our main page lets users click on links to browse by genre and then length or by electronic technique and then genre (see Figure 1).

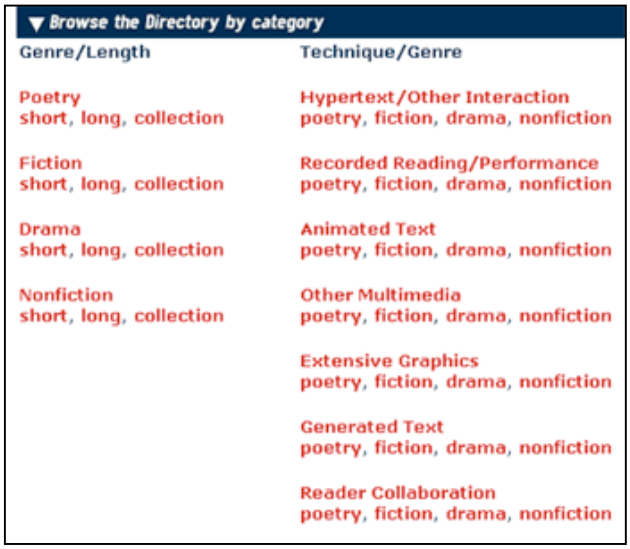

Figure 1: Primary ELD categories.

Categorizing by genre is fairly straightforward, since the vast majority of entries falls clearly into one category or another. Categorizing by the electronic techniques listed in Figure 1 is more complex, since a single work may fall into several of these categories at once. For example, there are many works that include hypertext, animated text, and extensive graphics. To accommodate such situations, the ELD allows any work to be listed under multiple categories.

Assignment to multiple categories can also resolve problematic situations in which one electronic technique may be hard to distinguish from another. Such ambiguity can arise, for example, with "generated text," which usually refers to text strings that are processed by a program and then output as a page of text that differs each time the program is run. This work can involve complex syntactical algorithms and large vocabulary lists or simply a small number of words strung together in random orders. Should a work such as Komninos Zervos's "A Kidz Story" [11]—in which sequences of graphic images cycle repeatedly to create random combinations of text lines—be listed as generated text or simply animated text? Though strictly speaking no text is being processed, the results are very similar to those of simple text generation programs.

The flexible structure of our database will allow us in the

future to implement filters that let users exclude certain categories of work from their searches. So if someone doesn't consider recorded readings or visual poetry to be true electronic literature, they will be able to ignore this type of work. In this way, the ELD will accommodate competing ideas of what constitutes e-literature.

Most directories that must handle large numbers of categories do so through hierarchical tree structures in which each main category includes several subcategories, which may in turn contain further subcategories, and so on. (See, for example, [2], [3], and [4].) Hierarchical categories present several problems, however. As FitzPatrick observes, every time a user must click to burrow down into a new level of subcategories, the risk increases that she will simply give up [8]. This is especially true if she must go down and back up through several layers of the hierarchy in search of something that could be in any one of several places. Furthermore, many of the categories that are useful for electronic literature don't easily lend themselves to hierarchical organization.

A widely accepted category of e-literature is "Flash poetry." Most of this work focuses on animation and audio, but since Flash pieces can incorporate any of the electronic elements listed in Figure 1, subcategorizing this work under "animated text" or "other multimedia" is problematic. Other useful categories are "game" and "Interactive Fiction" (a term for text-based role-playing games). One might expect to find "game" as a subcategory of "interaction" and "Interactive Fiction" as a subcategory of "game," yet many games rely heavily on other elements besides interaction and some authors of Interactive Fiction may not consider their works games.

FitzPatrick advocates natural language queries as a less confusing alternative to hierarchical menus based on multiple tree structures, and this is the direction in which the ELD is heading [8]. Our database has a variety of fields that can hold predefined values for plug-ins, authoring systems, and other technical information. It also allows such keywords as "game," "Interactive Fiction (role playing)," and "visual poetry" to be attached to records. Any of these values can become a de facto category, regardless of how it relates to other values or our main categories. In the future, our search mechanism will be expanded so that the user can search for any of these values singly or in combination. Thus the user will be able to browse such ad hoc categories as "animated DHTML poetry" or "Storyspace hypertext fiction" simply by entering these terms into the search box. This feature could also be supplemented by a query form that will let users select values from lists of plug-in names, keywords, and so on. In this way the ELD will be able to accommodate the numerous overlapping categories and alternative categorization systems applicable to e-literature.

BIBLIOGRAPHIC CHALLENGES

Obtaining accurate author attributions, publisher names, and publication dates for Web publications can be surprisingly difficult, because this information is often ambiguous, incomplete, or difficult to find on the Web sites where the works themselves are published. We hope to ameliorate this problem by providing this information in the ELD for each listed work, even when it isn't available on the publishing site hosting the work.

Making bibliographic data easily accessible is important for a number of reasons. Incomplete attribution information may hinder critics and scholars in properly citing works of e-literature. The enjoyment of general readers may be diminished if they can't put a work into its larger context by knowing who wrote it, who published it, and when it was launched. An apparent laxity concerning bibliographic details also undoubtedly contributes to the perception that still lingers in many quarters that Web publication is not as legitimate or serious as print publication.

There are a number of reasons why bibliographic data is often not readily available from Web publications. Sometimes omissions may result from the hastiness of an overworked editor. More often, though, the omissions seem the result of a minimalist design aesthetic that dictates having as little ancillary text as possible on the opening pages of a work. In these cases one can usually get an author attribution from the contents page of the site, but it's not uncommon for all contributor information to be relegated to an out-of-the-way place. Attributions can be particularly elusive in the case of collaborative works (such as those published by *391* [12]), which may deliberately downplay authorial individuality.

Additionally, standardized conventions for presenting bibliographic information have not yet had time to solidify on the Web as they have in print publishing. To accommodate the needs of bookstores and libraries, a work of recently printed literature will usually provide full bibliographic data on its copyright page. Since Web publishers aren't under the constraints of print distribution, they aren't compelled to follow the same bibliographic conventions. Thus the cataloguer of recent Web publications faces problems of missing or ambiguous publication data similar to those associated with printed literature from an earlier era before conventions arose for providing such data. Kirschenbaum demonstrates how many of these challenges presented by electronic texts are similar to those that have long confronted the textual studies community [13]. New challenges have also arisen, though.

Identifying the publisher of a Web work can sometimes be problematic when it isn't named on the title page. Sometimes a site will host several independent publishing areas, and it will be unclear whether the main umbrella site or its smaller constituents should be considered the publisher of a given work. It can also be unclear whether the work is self-published on an author's home page or put out by an independent publisher, because home pages often bear all the trappings of independent publishers. Although self-publishing doesn't carry the same stigma on the Web that it carries in print, it is useful to know whether a work has gone through an editorial selection process and possibly been subjected to editorial revisions. We therefore designate self-published works as such.

In print publishing the physical containment of a work within a set of covers generally makes it easy to distinguish the boundaries of that work. These boundaries are more likely to become blurred on the Web. It may also be

hard to determine whether a site consists of a single multisection work that should constitute a single entry or several discrete works that should be entered individually. To further complicate things, when a piece appears on a Web site, there may be a number of different apparent titles for it. For example, one version of the title might appear in the table of contents, a second version on the work's opening page, and a third in the browser title bar. It seems that the form a title takes on the opening page of a Web work doesn't necessarily have priority over other forms of the title, contrary to the practices of contemporary print publishing.

Publication dates are simply nowhere to be found for a surprisingly large percentage of the e-literature published on the Web. In some cases, the Web seems to challenge the notion of definitive publication dates. For example, if the work was originally put online in draft form before the finished version was released on the same site, or if the work is permanently "in progress," one might wonder about a precise publication date. In such cases, however, the date the work was first put online could easily be given, along with the date the work was completed, if applicable. (The ELD has provisions for recording all such details.) One suspects that many Webmasters don't actually consider "putting it on the Web" to be a true instance of publishing, and therefore don't consider a copyright date necessary. Or they may deliberately shun such conventions of print publishing as outmoded. These are unfortunate attitudes that contribute to the misconception that the Web is not a truly legitimate publishing channel.

E-literature publication dates may not seem important now to many people, since most of this work has been published during the last few years. But in the not-too-distant future this widespread lack of dating will pose problems for critics and historians. Unless the situation changes soon, it may be impossible for future generations to reconstruct the chronology of much early e-literature, which would be a significant loss.

The best-known bibliographic problem plaguing Web literature is the instability of URLs. An up-to-date directory provides the obvious benefit of letting a user easily find a work by name, even after it has moved to an online location different from the one she had last bookmarked. We plan to implement a "persistent URL" feature (discussed in the next section) that will help even further in this regard.

A less obvious URL-related problem is that a work may be available at more than one location on the Web. This may not even seem like a problem, until one considers that duplicate URLs rarely represent mirror sites. Often a work will appear on an author's home page and then subsequently appear on a publisher's site as well. Or it may appear on multiple publishers' sites. Because authors frequently revise Web works to fix bugs or typos, to resolve incompatibilities that arise with new browser releases, or simply to make improvements, these works may evolve after publication. Often such revisions are made only to one online copy of the work, essentially deprecating the other copies.

Obviously no one wants to read or write about a work that contains needless bugs, mistakes, or display problems. In the case of duplicate URLs, it is therefore usually desirable to list only the one that leads to the most current version. However, when a work is listed as the offering of a particular publisher and a newer version is housed on an author's home page, we're constrained to list that publisher's URL as the primary one for the piece. We must then note in the listing that a newer version is available at an alternative URL. If there are identical versions on the sites of two different publishers, we list the publisher's site where the work appeared first.

There are some cases in which two distinctly different versions of a work exist and neither one has absolute priority over the other. An example of this is *Forward Anywhere,* by Judy Malloy and Cathy Marshall, which exists in a self-published Web version [14] and in a later version published on disk by Eastgate Systems [15]. In such cases we create separate listings for each version, using a "version" field to distinguish between entries. Our decision to avoid separately cataloguing versions containing only minor variants or bug fixes is motivated mostly by a desire not to confuse readers. In the future, however, this may change as the ELD becomes a front-end for the Electronic Literature Organization's forthcoming Preservation, Archiving, and Dissemination initiative. It may become important for archivists to record these minor variants in the ELD, even if they aren't reflected in the main listings presented to the browsing public.

URLs create additional problems when it is not possible or desirable to link directly to the title page or opening screen of a work. Often clicking on the link to a work in a publishing site's table of contents will open the work in a new browser window that is set to specific dimensions. If the work is not loaded into this custom-sized window it may not display properly. In such cases, if an ELD entry provides the URL of the work's opening page, the work will display incorrectly in a default-sized browser window. Many publishers now provide a launch page (a separate page containing a link that launches the new window) to allow easy linking to such works. But in the absence of a launch page, the ELD listing must provide the URL for the publisher's table of contents page and may have to include instructions about what to click on to launch the work. We plan to add database fields for browser window height and width so that the URL for a listing will open in a correctly sized window when necessary.

CONTRIBUTOR INVOLVEMENT

Most large-scale literary directories rely on a paid staff or volunteer editors, often soliciting entries from potential listees but generally giving these listees minimal direct control over their listings. ELD employs a paid staff and volunteer editors but also allows authors and publishers editorial access to their own listings.

The bibliographic problems discussed above often cannot be resolved without consulting the author or publisher of a work. To ensure the accuracy of listings, our staff frequently sends out queries to authors and publishers and encourages direct input from them whenever possible. All listed authors and publishers have their own accounts in the ELD, which enables them to create or revise listings

for works they have written or published, which we then check over. When authors or publishers create listings they can also provide information that is unavailable to our editorial staff: details about the work's content or previous publication history, for example.

We are constantly looking for ways to increase contributor involvement. One means toward this end is a "Personal Directory" feature that we hope to implement in the future. This will offer each author or publisher listed in the ELD their own page that can function as a standalone directory of their publications. The works of the contributor will be sortable by date published, by genre, and (in the case of publishers) by author, as well as alphabetically by title. (Currently the listings of one author or publisher can be sorted only by title.) Searches will also be confined to just the contributor's works.

Each Personal Directory (PD) will have a mnemonic URL allowing anyone to link directly to it. For example, the PD URL of the online magazine Beehive would be something like "publishers.eliterature.org/beehive/home." Contributors could then treat their PDs as adjuncts to their own sites. For example, a magazine publisher could link to its PD to give readers a single index of all the contents scattered throughout various back issues. It would be much easier for publishers to maintain such indexes in the ELD than to build them on their own sites.

We also hope to offer a persistent URL feature for all contributors that will, for example, let a link to something like "authors.eliterature.org/deena_larsen" forward the browser request to the current URL of Deena Larsen's home page. When the actual home page URL changes, the author can update it in the ELD. Thus when other sites link to Larsen's home page using this persistent URL, the links will not be broken even should her home page relocate to a new server.

The success of such an approach of course depends upon the long-term survival of the ELD. While the Directory's continued operation in the near future is assured by sponsorship from UCLA, it will ultimately be subject to the economic vicissitudes that plague any nonprofit organization. The Electronic Literature Organization has a volunteer board of directors engaged in ongoing fundraising, and it also has provisions to keep the ELD operational if funds run dry. Because it relies entirely on open-source software, it could be moved at any time to a low-cost commercial hosting service or piggybacked onto another Web site and maintained by a core group of volunteers. This survival strategy has proved successful in the past.

LOCALIZATION

An important goal of the ELD is to reach beyond the English-speaking electronic literature community. Electronic writers and publishers are particularly active in Latin America and France, and we have many listings for works written in Spanish or French. We also list works in Portuguese, German, Italian, Swedish, Chinese, Korean, and other languages. In order to encourage interaction among writers, readers, and publishers on an international level, we are currently localizing the ELD to serve Spanish-speaking users.

The highly normalized structure of the ELD makes it relatively easy to present most data in both English and Spanish, since data is stored primarily in the form of keyed predefined values (such as "poetry" or "hypertext") that are selected from lists during data entry. We merely need to do a one-time translation of these English values into Spanish. A forthcoming upgrade to the ELD will present all listings in Spanish within an alternative Spanish-language interface. A later upgrade will provide Spanish translations of all the data entry instructions and forms. Once the framework for the Spanish interface is constructed, we will be able to add similar support for other languages merely by adding translated text to an XML file. We hope to create a localized version of the ELD in French sometime in the future, with other languages perhaps to follow.

CRITICAL COMMENTARY

One problem with a directory as extensive as the ELD is that its huge quantity of listings may confuse someone who is simply looking for a few good pieces to read. For some users, a small selection of recommendations, such as those provided at About.com, may be preferable to extensive listings of everything available.

To serve these users, we plan in the future to attach brief recommendations to some ELD listings and to add links to off-site reviews. Readers will then be able to confine their browsing to recommended works, authors, and publishers or view only entries recommended by a particular reviewer. At first these recommendations will be provided by ELO-selected commentators. ELO doesn't want to get into the business of canon-building, though, so we hope eventually to build a flexible recommendation mechanism that will allow a wide range of readers and critics to express their preferences. Thus the Web's diversity of voice will be preserved, but all the voices will be able to speak through a central portal where they have a better chance of being heard.

REFERENCES

1. Electronic Literature Directory. (2000–2003) http://directory.eliterature.org. Access: March 29, 2003.

2. Open Directory Project. (1998–2003) http://dmoz.org. Access: March 29, 2003.

3. Yahoo! Directory. (2003) http://www.yahoo.com. Access: March 29, 2003.

4. About. (2003) http://about.com. Access: March 29, 2003.

5. Rhizome ArtBase. (1998–2000) http://rhizome.org/artbase. Access: March 29, 2003.

6. Labovitz, John. John Labovitz's e-zine-list. (1993–2001) http://www.meer.net/~johnl/e-zine-list. Access: March 29, 2003.

7. Codd, E.F. A Relational Model of Data for Large Shared Data Banks, in Communications of the ACM 13, 6 (1970), pp. 377–387.

8. FitzPatrick, Greg. Universal Directories: Web Services as Human Services, in Proc. of WWW Conference 2002 (Honolulu, Hawaii, May 7–11, 2002), http://www2002.org/CDROM/alternate/729/index.html.

9. Morrissey, Judd. The Jew's Daughter. http://www.thejewsdaughter.com. Access: March 29, 2003.

10. Aarseth, Espen J. Cybertext: Perspectives on Ergodic Literature. Johns Hopkins University Press, Baltimore and London, 1997.

11. Zervos, Komninos. A Kidz Story. http://users.bigpond.net.au/mangolegs/kidzstory.html. Accessed March 29, 2003.

12. 391. http://www.391.org. Accessed March 29, 2003.

13. Kirschenbaum, Matthew G. Editing the Interface: Textual Studies and First Generation Electronic Objects, in Text 14 (2002), pp. 15–51.

14. Malloy, Judy, and Marshall, Cathy. Forward Anywhere. (1995) http://www.csdl.tamu.edu/~malloy/html/beginning.html. Access: March 29, 2003.

15. Malloy, Judy, and Marshall, Cathy. Forward Anywhere. (On disk.) Eastgate Systems, Watertown, MA, 1996.

MelbourneDAC 2003