Syracuse University

SURFACE at Syracuse University

School of Information Studies - Faculty Scholarship

School of Information Studies (iSchool)

1997

# Reproduced and Emergent Genres of Communication on the World-Wide Web

Kevin Crowston
*Syracuse University, School of Information Studies*

Marie Williams
*Web Architects, Syracuse NY*

Follow this and additional works at: https://surface.syr.edu/istpub

 Part of the Library and Information Science Commons

## Recommended Citation

Crowston, K. & Williams, M. Reproduced and emergent genres of communication on the World-Wide Web. The Information Society, 16(3), 201–215, doi: 10.1080/01972240050133652.

This Article is brought to you for free and open access by the School of Information Studies (iSchool) at SURFACE at Syracuse University. It has been accepted for inclusion in School of Information Studies - Faculty Scholarship by an authorized administrator of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.

# Reproduced and Emergent Genres of Communication on the World-Wide Web


## Description/Abstract

The World-Wide Web is growing quickly and being applied to many new types of communications. As a basis for studying organizational communications, Yates and Orlikowski [1, 2] proposed using genres. They defined genres as, “typified communicative actions characterized by similar substance and form and taken in response to recurrent situations” [1, p. 299]. They further suggested that communications in a new media will show both reproduction or adaptation of existing communicative genres as well as the emergence of new genres. We studied this phenomena on the World-Wide Web by examining randomly selected Web pages (100 in one sample and 1000 in a second) and categorizing the type of genre represented. Perhaps most interestingly, we saw examples of genres being adapted to take advantage of the linking and interactivity of the new medium, such as solicitations for help and genealogies. We suggest that Web site designers consider the genres that are appropriate for their situation and attempt to reuse familiar genres.



## Keywords

new media, communicative genres, genres, World-Wide Web

## Disciplines

Library and Information Science



## Creative Commons License

This work is licensed under a Creative Commons Attribution 3.0 License.

This article is available at SURFACE at Syracuse University: https://surface.syr.edu/istpub/76

The Information Society, 16:201–215, 2000
Copyright © 2000 Taylor & Francis
0197-2243/00 $12.00 + .00

# Reproduced and Emergent Genres of Communication on the World Wide Web

**Kevin Crowston**

*School of Information Studies, Syracuse University, Syracuse, New York, USA*

**Marie Williams**

*Web Architechs, Syracuse, New York, USA*

**The World Wide Web is growing quickly and being applied to many new types of communications. As a basis for studying organizational communications, Yates and Orlikowski (1992; Orlikowski & Yates, 1994) proposed using genres. They defined genres as "typified communicative actions characterized by similar substance and form and taken in response to recurrent situations" (Yates & Orlikowski, 1992, p. 299). They further suggested that communications in a new media would show both reproduction and adaptation of existing communicative genres as well as the emergence of new genres. We studied these phenomena on the World Wide Web by examining 1000 randomly selected Web pages and categorizing the type of genre represented. Although many pages recreated genres familiar from traditional media, we also saw examples of genres being adapted to take advantage of the linking and interactivity of the new medium and novel genres emerging to fit the unique communicative needs of the audience. We suggest that Web-site designers consider the genres that are appropriate for their situation and attempt to reproduce or adapt familiar genres.**

**Keywords** document genre, World Wide Web, structuration theory, web adoption

Received 17 September 1998; accepted 18 February 1999.

We thank Louis Monier of Digital Equipment Corporation's Palo Alto lab for his assistance in understanding Alta Vista and for providing the sample of URLs, JoAnne Yates for her advice on this article, and Devang Shah for his assistance in coding the sample. Three anonymous reviewers and the editors provided useful suggestions for improving this paper.

An earlier version of this paper was presented at the Hawai'i International Conference on Systems Science, 1997.

Address correspondence to Kevin Crowston, School of Information Studies, Syracuse University, 4-206 Centre for Science and Research, Syracuse, NY 13244-4100, USA. E-mail: crowston@syr.edu

The World Wide Web (or the Web) is an Internet client–server communication system for retrieving and displaying multimedia hypertext documents (Berners-Lee et al., 1994). Documents are identified by an address called a uniform resource locator, or URL. The Web's main advantages over earlier Internet systems are its merger of retrieval and display tools, its capacity for handling formatted text, embedded graphics, and other media, and point-and-click links to other documents (hence the name). Also, many browsers are capable of seamlessly retrieving information using older protocols (e.g., FTP, Gopher, and Usenet news) and automatically launching other applications to display diverse Internet data types (e.g., sound, animation).

Over the past several years, the Web has clearly become the most popular Internet application (at least in terms of traffic). For example, Bray (1996) reported that 223,851 servers at 89,271 sites provided 11,366,121 unique URLs in November 1995; Woodruff et al. (1996) collected 2.6 million Web documents. These numbers continue to grow—current size estimates are in excess of 55 million URLs (Pedram, 1997)—as individuals, businesses, and other organizations rush to establish Web presences.

What is less clear is how the Web can or will be used by these diverse groups. Many organizations see the Web primarily as a cheap means of publishing information and are simply moving existing documents to the new medium (so-called brochureware). However, a few are experimenting with its capabilities to communicate and interact in novel ways, creating new genres of communication (Campbell & Jamieson, 1978; Harrell & Linkugel, 1978; Miller, 1984). For example, Andersen Consulting is experimenting with a Web-based agent, Bargain Finder, to allow users to comparison shop for music CDs (Krulwich, n.d.).

The purpose of our study was to describe the range of genres of communication in use on the Web. Communicative genre is defined as an accepted type of communication sharing common form, content, or purpose, such as an inquiry, letter, memo, or meeting. Note that genre is not simply the medium of communication. A document with a memo genre may be realized on paper or in an electronic mail message (two different media), while the electronic mail medium may be used to deliver memos and inquiries (two different genres of documents). However, medium does influence which genres are accepted. We were particularly interested in studying how the adoption of a new communication medium, the Web, is leading to adaptation of existing genres and the emergence of new ones.

The article starts by discussing the concept of document genre and its applicability to the Web. Succeeding sections describe our methodology and discuss the genres we found. Our specific contributions are to document the range of genres found in the sample of Web pages and to identify possibly new genres in use on the Web. We also hope to suggest how many uses of the Web have evolved from earlier communicative practices rather than being created de novo. We conclude by discussing how the notion of genre can be useful to Web-site designers and, as is required by the genre of this article, by considering future research directions.

## THEORETICAL BACKGROUND

Rhetoricians since Aristotle have attempted to classify communications into categories, or genres, with similar form, topic, or purpose. Numerous definitions of genre have been debated in that community (e.g., Campbell & Jamieson, 1978; Harrell & Linkugel, 1978; Miller, 1984; Swales, 1990). Other groups have also struggled with the notion of document types; for example, information scientists have included rules for document types in Standard Generalized Markup Language (SGML) (Dollar, 1994).

More recently, Yates and Orlikowski (1992; Orlikowski & Yates, 1994) proposed using genre as a basis for studying communications in organizations. They defined genres as "a distinctive type of communicative action, characterized by a socially recognized communicative purpose and common aspects of form" (Orlikowski & Yates, 1994, p. 543). In other words, given a socially recognized need to communicate (i.e., a purpose; Bitzer, 1968), individuals will typically express similar social motives, themes, and topics in a communication with similar physical and linguistic characteristics (i.e., form); that is, they will communicate in a recognized genre. Some genres are defined primarily in terms of purpose or function, such as a proposal or inquiry; others are defined in terms of the physical form, such as a booklet or brochure; still others are defined in terms of the document form, such as lists or directories. However, most genres imply a combination of purpose and form, such as a newsletter, which communicates "the news of the day," including multiple short articles and is distributed periodically to subscribers or members of an organization.

This article, for example, is an example of the social science paper genre, commonly used when communicating scientific results in a community of social scientists to advance the state of knowledge in a field (as well as the authors' careers). It has a form familiar to a social scientist: a title, authors and affiliations, sections for introduction, theory, method, data and discussion, directions for future work, citations, a bibliography, and so on. Other common genres include letters and memos, project team meetings, and TV sitcoms, all immediately recognizable by their typical purpose and characteristic form.

*Genre Hierarchies.* The notion of genre has been extended in several directions. First, as the examples illustrate, genres can be arranged in a hierarchy, as in the *Art and Architecture Thesaurus* (Petersen, 1994), for example. The social science paper is a special case of a more general research paper genre, which in turn is a type of paper. Other types of research papers are computer science implementation papers, biology research papers, and so on. These genres share some similarities, such as a title, abstract and bibliography, but differ in other particulars, such as the expected section headings, types of arguments, and so forth. Similarly, TV sitcoms are a special case of TV shows in general, project team meetings of meetings, and so on. Rather than argue about the proper level of analysis for a genre, we believe it is most useful to follow Yates and Orlikowski and consider genres at any of these different levels.

*Embedded Genres and Genre Systems.* Second, multiple genres may be linked or embedded to form a more complex pattern of communication. As Orlikowski and Yates (1994) pointed out, some communications use multiple genres simultaneously, such as a proposal embedded in a memo. Multiple communications may also be performed in a recognizable pattern, what Bazerman (1995) called a genre system. Examples include the sequence of examination and cross-examination in a trial, or the cycle of article submission to a journal or conference, reviews, final acceptance or rejection letters, and publication. Features of a genre may enable their use in a genre system. For example, page numbers in a technical paper make it possible to cite concepts or quotations from the paper, thus binding the paper into the literature.

*Genre Repertoires.* Finally, Orlikowski and Yates (1994, pp. 546–547) introduced the notion of a genre repertoire, that is, the set of genres in use within a community. They noted that different communities use different genres

in their communication, and use common genres with different frequencies. These differences provide one source of insight into the communicative (and other) practices of the community. For example, a community of social scientists and computer scientists can be distinguished by the frequency of use of different paper genres as well as by the paucity of computer programs and program documentation created in the former, reflecting different modes of research.

Genres are useful because they make communications more easily recognizable and understandable by recipients. Because we drew on the social science paper genre, for example, another social scientist can more quickly determine the purpose and content of our communication and begin to evaluate its contribution. On the other hand, a genre may be unfamiliar or hard to understand for someone outside of the community. In fact, recognition of a particular genre is one sign of membership in a particular community. Freedman and Medway (1994, p. 14) suggest that incomprehensible genres may even be used deliberately to defend positions of privilege.

## Genre Change

Drawing on Giddens's (1984) structuration theory, Orlikowski and Yates (1994, p. 545) argued that "People produce, reproduce and change genres through a process of structuring." As members of the community draw on their knowledge of a genre repertoire to communicate, they reinforce the use of these genres, making them more appropriate or legitimate for use in the given situation. For example, by creating an order entry Web page that draws on the genre of an order form, a designer reinforces the appropriateness of the order form genre for this type of communication, making its use in future situations more likely. In other words, the set of genres in use (i.e., the genre repertoire) is both a product of and a shaper of the communicative practices of a community.

*Reproduced Genres.* Orlikowski and Yates (1994, p. 547) suggested that in a new situation individuals will typically draw on their existing genre repertoires, reproducing genres they have experienced as members of other communities. For example, traditional genres such as the book or academic article have moved intact to the Web. These reproductions may be immediately accepted, or there may be a transition period during which the limits of the genre are renegotiated. For example, the electronically distributed journal article is still in transition (Kling & Covi, 1995; Harter, 1998). It is being used, but this adapted genre is not yet completely accepted or considered legitimate for all purposes (e.g., as evidence for a tenure case) by the academic community as a whole.

*Adapted Genres.* However, people are also free to modify a genre and communicate in a way that invokes only some of the expected aspects of a form. If these changes become repeatedly used, they too may become accepted and used together with or instead of existing genres, thus extending or altering the genre repertoire. For example, the journal article will likely change as it moves onto the Web to take advantage of the possibilities of linking or embedding information; the eventual form may bear only passing resemblance to the self-contained 20- to 25-page articles of today. Also, modifications of genres that are parts of genre systems may require corresponding changes to the rest of the system. For example, changes in citation habits will be necessary before page numbers can be dropped from the technical paper genre. Such interdependencies between genres will tend to slow the adoption of a new genre.

Because the definition of genre relies on social acceptance, it is impossible to define the exact point at which a new genre emerges from the old one. Acceptance may take many years. However, after some period of coexistence, the new combination of form and purpose may become generally recognized and named as a separate genre. For example, the FAQ (frequently asked questions) has emerged as a distinct genre on the Usenet and Web. An Alta Vista search indicates approximately 170,000 Web pages with FAQ or "Frequently asked questions" in their title. Also, genres may be accepted in different communities at different rates. The emergence of distinctive new genres would be one sign of the formation of a new community with new communicative practices.

## Why Study Genres on the Web?

The Web provides a particularly interesting setting in which to study the use and development of genres and genre repertoires. First, the capabilities of the new medium seem likely to result in the development of new genres of communication. Furthermore, the rapid development of this medium suggests a high level of experimentation with potential genres. Bearman (1994, pp. 160–161), for example, notes the rapid evolution in what he refers to as "forms of material" in electronic media in general.

Second, because the majority of Web sites are public, many examples of Web communication are easily available for study. Furthermore, because there is no central management of the Internet or the Web, there is no explicit management or enforcement of genres of communication, as might happen in the introduction of a communication system in a corporate environment (Orlikowski et al., 1995). Instead, individual Web site developers individually choose how to present their information, drawing on their understanding as members of a community, what Orlikowski et al. (1995) called implicit structuring (in this

case, from the point of view of the Web-page developer rather than the recipient of the communication).

Finally, there are many communities meeting on the Web, bringing experiences with different genres and using the Web for many different purposes. The Web is sometimes used for direct communication where someone with a Web server "delivers" a document to members of a known community by giving them a URL. For example, some academics use the Web to communicate with colleagues by publishing their own papers, and with students by publishing syllabi and assignments. Another example of communication within a predictable community is computer companies' announcing new products, publishing catalogs, or providing troubleshooting tips online for their customers. Since computer users by definition have the computer necessary for Web access, computer companies have been early and heavily into Web-site development in expectation of directly reaching their customers.

However, in many other cases the audience is unpredictable. Unlike the Usenet or electronic mail groups, there is no clear separation of communities into different channels of communication (as is the case for journals or talks given at conferences, for which the audience is likely to have shared interests). Indeed, it is unlikely that there is a single Web community at all. Therefore, the resulting genre repertoire of a collection of Web pages will be the result of interactions among communities. In some cases, a genre may act as a type of boundary object (Star & Griesemer, 1989), providing a common point of contact between different groups (Freedman & Medway, 1994). In others, this mixing may lead to genre confusion, meaning that there is a practical need to understand the way genres enable communication. For example, organizations have used the Web to publish information such as product brochures, annual reports, country, state, and city home pages, government agency press releases, and so on. These organizations tend to use existing genres when putting information on the Web. However, a person happening to reach a document on one of those Web sites has a good chance of being outside the community in which that genre evolved. As a result, the document may be confusing and the communicative purpose lost.

## METHOD

To document the range of genres currently in use on the Web, we sampled and classified randomly selected Web pages. We chose the individual Web page as the unit of analysis for several reasons. First, we had no way to create a random sample of Web sites, since the available databases of Web sites (e.g., Yahoo!) are typically manually created and reflect the editorial biases of their creators. There are also significant difficulties in drawing boundaries around a Web site, as we discuss in the conclusion.

Instead, we sampled pages without regard to where they appeared in a site. As a result, our sample included parts of documents as well as whole documents, allowing us to see adaptations of existing genres in different parts of an electronic document. However, this procedure means that we were more likely to choose pages from sites with many pages. Our sample contains mostly pages from the interior of sites, since there are many interior pages and typically only a few top pages. (Bray [1996] found that the majority of pages are pointed to only by other pages at the same site).

### Sample

Our sample of Web pages was created by selecting 1000 URLs from the pages indexed by the Alta Vista search engine. The developers of Alta Vista provided us with a random sample of about 8000 URLs drawn from their database, from which we randomly chose 1000 for this study. The sample of URLs was taken in February 1996. At that time, Alta Vista attempted to record essentially all URLs (or at least, all of the publicly accessible parts of the Web, a point we return to in the conclusion). Because of Alta Vista's comprehensiveness at that time, we did not feel it was desirable to add pages chosen from other sources, since that would have resulted in a sample that drew more heavily from pages found in the smaller databases. Nevertheless, in light of the limitations of our sample, we make no statistical claims about the composition of the entire Web. Also, since no search engine still attempts to provide universal coverage, this procedure would not be effective today.

The Web pages and graphics were captured in May 1996 (the delay between sampling and capture was caused by the need to develop and debug the spider used to capture the Web pages and their graphics). By that time, the sample included 128 obsolete URLs (error 404) and 35 URLs to which the server did not respond, leaving a total of 837 pages to be studied (of these, a further 11 were "Custom 404s," as discussed later). We have continued to snapshot this sample every 6 months, although the analysis presented here concerns only the initial snapshot.

### Coding

In their study of genres in electronic mail, Orlikowski and Yates (1994) coded the purpose of each message as well as specific features such as the presence of embedded messages, subheadings, or lists. They then defined genres in terms of combinations of these features. Such an approach was necessary to inductively identify specific genres. Their study required such precision because their messages were mostly quite similar, because the differences in the forms of the genres were small (e.g., all were e-mail messages),

and because they wanted to reliably classify hundreds of messages. However, in our case we had only a few examples each of many different genres and the differences in the forms of most were pronounced, obviating the need for such precision. More importantly, we were primarily interested in the appearance of novel genres and therefore were willing to tolerate slight imprecision in categorizing documents with well-accepted genres. For these reasons, we defined genres based on our experience with the Web and with other forms of communication.

We started with a list of genres and their definitions developed in a pilot study and refined it during the course of this study. Definitions of accepted genres were drawn in many cases from the *Oxford English Dictionary*. Using the definitions and a database of the captured pages, a research assistant did the actual determination of the genre of the majority of the Web pages. After looking at each page, the coder chose the appropriate genres from a pop-up list in the database program. If none of the already defined genres were appropriate, the coder could add new genres to the list.

In a few cases, the coder used Alta Vista's translation feature to examine pages in foreign languages. Interestingly, we found many cases where it was possible to tentatively assign a genre to a page written in a language that could not be translated and that the coder did not read, underlining the importance of form in defining certain genres. Nevertheless, in 47 cases, we found pages that simply could not be classified because they were not in English and had ambiguous forms (24 pages), because we did not know the name for the genre (2 pages), or because the pages did not have a recognizable genre (21 pages). The latter included binary or other nontext documents.

Approximately 40% of the sample was also coded by one of the authors to determine reliability of the coding. The two raters agreed completely on the coding for 68% or 193 of the 285 double-coded pages. For a further 28 pages (10%), the genres selected were similar. In many cases, the two genres selected were specializations of a common type (e.g., article vs. column or sports card vs. employee profile) or parts as opposed to a whole (e.g., an article vs. a newsletter).

In other cases, the assigned genres differed in assumed audience, author, or relationship to other documents—that is, in features of the context of the document. For example, we found a page describing an electronic data interchange transaction set; this page is either a specification, if it describes how a system works, or a standard, if it dictates how it must work. Numerous pages might be either press releases or articles, depending on who published them, a company, or an online magazine. A page with the name and address of a company that stands alone might be a business card; however, if it were part of a collection of similar pages, it might be a directory entry or a classified ad. These coding difficulties reflect problems in the use of the Web, an issue we discuss further in the conclusion. In 22% of the cases the two coders simply disagreed.

The Web pages for which the raters disagreed were reexamined by the coder and one of the authors and discussed to determine the final genre. Our primary purpose in reexamining these pages was to refine the definition of genres rather than to perfect the codes assigned. For example, we paid particular attention to disagreements on coding hotlists, home pages, and bookmarks in order to refine our definitions of these emerging genres. On the other hand, we worried less about disagreements on coding pages as columns or articles (for example). Although these genres are also quite similar and easily confused, both are well-accepted genres and therefore of less interest for our research focus on novel genres. Similarly, we did not try very hard to distinguish the examples mentioned earlier (specification vs. standard, press releases vs. article, business card vs. directory entry or classified ad) since, again, these are all familiar genres. Once we had refined the definitions, the coder then reexamined other pages to ensure they had been coded properly.

As mentioned earlier, some genres are special cases of other, more general genres. The genres assigned were as precise as possible. A sample of genres briefly defined can be found in the Appendix. The structure of this hierarchy was adapted from the *Art and Architecture Thesaurus* (AAT, Petersen, 1994). Terms from the AAT are listed in regular type; terms added by the authors are in italics. Note that this thesaurus groups genres by form and by function (i.e., purpose) as well as by condition of production.

In addition to coding for genre, we used a Perl script to parse the HTML files and code objective features of the pages, such as the number of hypertext links, forms, or images included.

## DATA

In this section we briefly describe some characteristics of the pages we studied to demonstrate the diversity of sources and their use of the features of the Web. The sample includes pages from at least 40 countries, as shown in Table 1. About equal numbers were from educational and commercial sites, while a smaller number were from governmental sites. The majority of pages were in English, but 17 languages were found as were few multilingual pages and pages with only computer codes, as shown in Table 2. In terms of origin, our sample seems representative for the time. *Business Week* ("A World Wide Web," 1996) reported on the location of Internet host computers. Their figures and ours are generally similar, as shown inTable 3.

Most pages used some of the capabilities of the Web, as shown in Table 4. About two-thirds had images, although only 18% had a background image. More than 80% had some kind of hypertext links, compared to just under 75%

**TABLE 1**
Top-level domains of sample sites

| Domain | Count | Domain | Count |
|---|---|---|---|
| com (U.S. commercial) | 223 | hu (Hungary) | 4 |
| edu (U.S. educational) | 205 | ie (Ireland) | 4 |
| net (Network) | 49 | br (Brazil) | 3 |
| ca (Canada) | 39 | cz (Czech Republic) | 3 |
| jp (Japan) | 39 | kr (Korea [South]) | 3 |
| de (Germany) | 31 | cl (Chile) | 2 |
| org (Nonprofit) | 29 | my (Malaysia) | 2 |
| uk (United Kingdom) | 28 | sg (Singapore) | 2 |
| se (Sweden) | 20 | tw (Taiwan) | 2 |
| gov (U.S. government) | 20 | za (South Africa) | 2 |
| us (United States) | 16 | co (Colombia) | 1 |
| au (Australia) | 15 | gr (Greece) | 1 |
| fr (France) | 11 | hr (Croatia) | 1 |
| fi (Finland) | 8 | il (Israel) | 1 |
| nl (Netherlands) | 8 | is (Iceland) | 1 |
| be (Belgium) | 7 | mx (Mexico) | 1 |
| ch (Switzerland) | 7 | nz (New Zealand) | 1 |
| it (Italy) | 6 | pt (Portugal) | 1 |
| at (Austria) | 5 | ru (Russian Federation) | 1 |
| dk (Denmark) | 5 | si (Slovenia) | 1 |
| no (Norway) | 5 | sk (Slovak Republic) | 1 |
| hk (Hong Kong) | 5 | tr (Turkey) | 1 |
| mil (U.S. military) | 4 | IP address | 9 |
| es (Spain) | 4 | Grand total | 837 |

**TABLE 2**
Languages used on the sampled Web pages

| Language | Count | % |
|---|---|---|
| English | 708 | 84.6 |
| Japanese | 32 | 3.8 |
| German | 26 | 3.1 |
| French | 14 | 1.7 |
| Spanish | 10 | 1.2 |
| Swedish | 7 | 0.8 |
| Dutch | 5 | 0.6 |
| Finish | 5 | 0.6 |
| Chinese | 3 | 0.4 |
| Danish | 3 | 0.4 |
| Italian | 2 | 0.2 |
| Korean | 2 | 0.2 |
| Russian | 2 | 0.2 |
| Hebrew | 1 | 0.1 |
| Hungarian | 1 | 0.1 |
| Indonesian | 1 | 0.1 |
| Malay | 1 | 0.1 |
| Norse | 1 | 0.1 |
| Slovak | 1 | 0.1 |
| Multilingual | 4 | 0.5 |
| Computer language | 3 | 0.4 |
| No words | 5 | 0.6 |

*Note*: $n = 837$.

of the pages studied by Bray (1996). The most common kind of link was to another page in the same directory on the same server; links to the same page or to pages in a subdirectory of the server were relative rare. To simplify the presentation of these data, therefore, we grouped these different kinds of links into two categories, as shown in Table 5:

- Links outward (to other sites or to pages in higher directories on the same server).
- Links inward (to the same page or to other pages in the same directory or in a subdirectory on the same server).

About half the pages included links outward; two-thirds included links inward. Less than 6% of pages included a form (an HTML construct for submitting information from the browser back to the server).

## DISCUSSION

In our survey we found many examples of reproduced genres, as well as pieces of documents and components of genre systems (see Table 6). Many of these pages showed signs of adaptation, making use of the features of the Web to change their form or to serve a different purpose. We also identified numerous pages with typified communicative purposes and forms unique to the Web, but we felt they were well accepted by Web users and thus constituted potentially novel Web genres. Many pages could not be classified either because we did not understand the language

**TABLE 3**
Comparison of page origins to those reported by *Business Week* ("A World Wide Web," 1996)

| Origin | Our sample | *Business Week* |
|---|---|---|
| United States (.com, .edu, etc.) | 546 (65.9%) | 64% |
| Other English-speaking | 87 (10.5%) | 12.7% |
| Western Europe (excluding Britain and Ireland) | 120 (14.5%) | 16.9% |
| Asia | 53 (6.4%) | 4.0% |
| Eastern Europe | 11 (1.3%) | 0.9% |
| Central and South America | 7 (0.8%) | 0.6% |
| Middle East and Africa | 4 (0.5%) | 0.9% |
| IP address only (no host name) | 9 | |
| Total | 837 | |

*Note*: Percentages are of pages for which the origin could be determined; $n = 828$.

**TABLE 4**
Use of linking and other Web features in sampled document

| Type of links | Count | % |
|---|---|---|
| Links to pages on other hosts | 207 | 24.7 |
| Links to pages in parent directories | 323 | 38.6 |
| Links to pages in the same directory | 511 | 61.1 |
| Links to pages in subdirectories | 87 | 10.4 |
| Links to the same file | 86 | 10.3 |
| No hypertext links | 157 | 18.8 |
| E-mail links | 230 | 27.5 |
| Images | 544 | 65.0 |
| Background images | 151 | 18.0 |
| Forms | 49 | 5.9 |

*Note*: $n = 837$.

or because the page did not have a clearly identifiable form or purpose. Some of these unclassified pages may in fact be emerging genres.

## Reproduced Genres

Most of the pages we studied reproduced more or less faithfully genres or combinations of genres familiar in traditional media, such as the article (36 pages), frequently asked questions (FAQ, 12 pages), meeting minutes (2 pages), or course descriptions (25 pages). In two cases, we recognized the purpose and form of the page, although we were at a loss for a convenient name for the genre. For example, one page described a franchise opportunity and included a form to request additional information. Such pages may represent genres that are common in a community of which we are not members (e.g., franchising). Other pages represented types of communication that are stereotyped but not usually named, such as someone displaying and describing photographs of family members, friends, or a trip (21 pages).

**TABLE 5**
Frequency of hypertext links inward and outward

| | No links outward | Links outward | Row totals |
|---|---|---|---|
| No links inward | 157 (18.8%) | 124 (14.8%) | 281 (33.6%) |
| Links inward | 252 (30.1%) | 304 (36.3%) | 556 (66.4%) |
| Column totals | 409 (48.9%) | 428 (51.1%) | 837 (100.0%) |

*Note*: Links inward include links to the same page and to pages in the same directory or in subdirectories. Links outward include links to pages in parent directories or on different hosts; $n = 837$.

**TABLE 6**
Count of types of genres found

| Type of genre | Count | % |
|---|---|---|
| Familiar genres | 507 | 60.6 |
| New but accepted genres | 239 | 28.6 |
| Apparently new genres | 44 | 5.3 |
| Unclassifiable | 47 | 5.6 |

*Note*: $n = 837$.

Many pages were parts of longer documents, such as part of an index (3 pages) or a catalog (14 pages). We also found many pages that described an organization (5 pages) or product (30 pages) but that appeared to be part of a larger directory of organizations or products. All of these pages were still recognizable as a distinctive genre, although the purpose was sometimes hard to determine.

We also encountered pages where content in one genre was embedded in a document with another genre. For example, we found:

- An instruction sheet on how to apply for a loan that included the eligibility rules for the program.
- An announcement that included the Unix main page for the software being announced.
- A newsletter that included an events calendar.
- A press release that included a policy statement.

Each embedding created a document of a new genre without completely losing the characteristics of the previous instantiation. This phenomenon is not new, as Orlikowski and Yates (1994) point out, but we believe that it is much more relevant to the Web because of the ease of reusing text in electronic form. For example, e-mail on a particular topic can be collected and made available on the Web (indeed, we found 25 pages containing an e-mail message from an archive). If these messages were selected on a particular topic (e.g., discussions or reviews of a type of product) they create a review or FAQ document, even though the surface genre is a simple listing of e-mail messages.

## Adapted Genres

Many pages showed signs of adaptation to the capabilities or needs of the new media. In the next two subsections, we first discuss pages that exhibit new forms enabled by the use of Web features that are thus potentially new genres. In the following section, we consider pages that serve new communicative purposes and thus are examples of potentially novel genres.

In all, approximately 80% of pages included some kind of hypertext link. Use of links and other Web features does not necessarily change the genre of the page. For example, simply adding images to a page would usually not affect the genre of the document. Many of the document pieces we found used links simply to provide navigation to other document pieces, thus forming a multipage document. These multipage documents had the same purpose and overall form of traditional genres, even though they took advantage of the linking capabilities of the Web to create more manageable pieces.

On the other hand, some pages showed signs of adaptation, with a new form enabled by the linking. For example, links allow information to be accessible from a page without the kind of direct embedding discussed earlier. We saw examples such as genealogies using linking to display and navigate the usually unwieldy amount of data in a family tree (7 pages).

Further, linking can enable a single page to serve multiple purposes and thus be an example of multiple genres. For example, a list of items can include pointers to more information on those items, thus creating both a list and an index. We saw examples such as:

- A paper abstract that included the table of contents of the paper with pointers to pages containing the rest of the document.
- A film review that included links to an order form for the film and was thus part of a film catalog.

These pages were examples of traditional genres that used linking to go beyond a single purpose. At this point, these pages show a mix of genres; if such mixes become more common, they may start to be seen as genres in their own right.

## New Communicative Needs on the Web

A few pages appeared to be novel genres, yet ones that are already well accepted by many Web users, as indicated by references to them in Web design guides, introductions to the Web, and so on. These pages have new genres because they serve communicative purposes unique to the Web. Yates and Orlikowski (1992) suggested that these new genres are most likely derived from earlier genres that might have seemed appropriate to the situation. We briefly describe three sets of genres in particular—home pages, bookmarks/hotlists, and numerous genres related to Web servers—and speculate on their origins.

*Home Page.* An easily identified and commonly accepted genre is the home page, either personal or organizational. We defined a home page as personal or organizational information plus links to other pages reflecting the subject's interests that are intended to introduce the person or organization to the world and to facilitate further contact. Our sample included 36 personal home pages and 24 organizational home pages. One page introduced a city in the same manner.

The antecedents of the home-page genre are unclear and are a promising topic for future research. One hypothesis suggested by JoAnne Yates (1996, personal communication) is that they are adaptations of the .plan files maintained on some Unix machines. (While a minority of current users of the Web use Unix, many of the initial users did.) A .plan file is created by the owner of an account and is printed whenever another user "fingers" or requests the status of the account. It typically includes contact information and a brief description of the owner's interests or job (e.g., "hacking perceptrons for Minsky," Raymond, 1993).

There were also organizationally created home pages, which seem to be adaptations of entries, for example, from a university's faculty profiles book. Further, many organizations maintain home pages as overviews to the organization, their site, or some collection of information.

*Hotlists and Topical Home Pages.* We defined a hotlist as a series of links to material not controlled by the page developer, on a related set of topics. In other words, a hotlist has many outward links and few inward links, as opposed to the table of contents of a document or the directory of a site's content, which it otherwise resembles in form. In our sample, we found 26 hotlists, on topics such as music, HTML, nanotechnology, films, environmental organizations, computer stores, and presidential candidates. A bookmark file (22 pages) is similar to a hotlist, except it includes links to pages on unrelated topics and need not be organized at all.

A related genre we called a topical home page (10 pages). Although hotlists and topical home pages have the same general purpose (to facilitate access to information about a particular topic), we distinguished between them by the amount of information provided on the page. Hotlists provide just the links; topical home pages also provide an overview or introduction to the topic. For example, one topical home page provided an introduction to and categorization of data analysis tools, along with links to other sites and to pages with more detailed descriptions of the tools.

The hotlist form seems to have drawn on earlier forms of posted lists of useful FTP sites or bulletin boards and on the bookmarks file maintained by most Web browsers. While superficially similar to a bibliography, hotlists typically have only the link itself as a reference; rarely are complete author, title, and date of publication information given. Hotlists therefore seems unlikely to have developed from this genre. Hotlists were especially useful before the development of good resource discovery tools, such as Yahoo!, Alta Vista, or Google. They appear now to be

developed as a way to express an interest or to add value to a Web site (such as a commercial site) as much as for personal navigation. As a result, we might expect hotlists to gradually mutate into topical home pages, as authors add more information about the topic of interest.

*Pages About Web Servers.* We found many pages that conveyed information regarding the functioning of Web servers and that can be considered to be novel genres in that they have recognized communicative purposes and distinctive forms. These genres are novel to the Web because their purposes are tied to the functioning of the Web infrastructure. In particular, our sample included pages such as:

- Under construction—place holders for content, often with a "man at work" symbol (16 pages).
- Custom 404—customized pages reporting that the requested page could not be found, typically with information about the service provider and its logo (11 pages).
- Web site has moved—pages reporting a new URL for the requested page (8 pages).
- File directory lists—lists of files available in a particular directory of the Web server, in the form of an FTP directory listing but with clickable links (7 pages).
- Web server statistics—records of the number of hits, bytes downloaded, and so on (9 pages).

The last group, server statistics, provides an example of the potentials for confusion due to nonshared genres. The statistics are reports of interest primarily to the managers and sponsors of the Web site; however, it turns out to be technically easy to provide them via the Web, which makes them available to everyone, even though most are unlikely to find them useful or even understandable.

*Interactive Pages.* Other pages took advantage of the interactivity of the Web to create documents that were more like interfaces to computer programs. These pages have recognized forms and communicative purposes, although in some cases the forms seem to be drawn more from computer programs rather than from documents.

Many of these pages had to do with management of the Web itself. For example, we found pages such as:

- A form to request notification of changes in a page.
- A URL submission form to add a URL to a search engine.
- A search engine interface to search for pages with particular information.
- A page of database search results.

Other pages took advantage of the interactivity to accomplish communications that might otherwise have been done via e-mail or some other medium. For example, we found pages such as:

- A form to submit comments.
- Web-based discussion groups with an archive of old messages and provisions for posting follow-up messages.

We also found examples of emerging genres for online commerce, such as online order forms or shopping carts. These pages resemble traditional order forms but allow the order information to be directly entered into an order system.

Finally, and perhaps most interestingly, we found a few pages that provided access to applications. For example, we found pages such as:

- A Web-based trip planner.
- An instant Web page creator.
- A page for selecting video clips.

In a sense, these pages embody the promise of the Web to provide ubiquitous distributed access to all kinds of information and processing. They also indicate the difficulty of achieving this promise, since their novel forms and purposes are not always apparent.

## Unclassified Pages

As mentioned earlier, we had difficulty assigning genres to a number of the pages, most often when we agreed there was a genre but simply did not know the name. In other cases, we could not determine the purpose of the communication, making the assignment of a genre problematic.

However, some of these cases may be examples of genres in the process of adaptation to the Web. At this early stage, they are usually seen as variants of an accepted genre, missing some features and possibly adding others. For example, the selected collections of e-mail were not considered proper e-mail archives because they were incomplete and not sorted by time or author, but they were not quite FAQs either, because they were not edited into a coherent document. If their use continues, then they may eventually become independent genres.

Several pages seemed to be part of games while others were part of virtual reality tours. It was difficult to say what the genres of these pages were, because we did not know how to use them or what they were for. In other words, the form and purpose are not yet socially recognized (at least, not in the communities to which we belong). Others, like splash pages on Web sites, may represent dead ends, experiments with linking that do not become commonly accepted usages.

## CONCLUSIONS

We argue that genres provide a useful theoretical tool for analyzing uses of the Web. In general, the concept of genre seemed to be easily applied to the Web pages we studied, with some caveats that are discussed later. We found numerous examples of genres' being reproduced on or adapted to the new medium as well as a few examples of new genres, such as the hotlist and home page, which might be studied. We believe that the size of the genre repertoire is a reflection of the many different communities on the Web and their varied uses of the medium. There are already some puzzles, like the origins and antecedents of the home page. Therefore, we hope to study the Web over a longer period to better document the processes by which genres are being adapted and new genres are emerging. Since many users spend a lot of time revising and improving their home pages, this genre will likely continue to evolve and be a good subject for future research.

### Future Research: Creating a Hierarchy of Document Genres

As mentioned earlier, genres can be arranged in a hierarchy, since some genres are special cases of others. A few thesauri do include genres, although sometimes in restricted domains. For example, the Library of Congress has a thesaurus of genres for graphic materials (e.g., pictures and paintings) (Library of Congress, Prints and Photographs Division, 1995), the ACRL has prepared a thesaurus of terms for rare books (Bibliographic Standards Committee of the Rare Books and Manuscripts Section, ACRL/ALA, 1991), and the *Art and Architecture Thesaurus* (Petersen, 1994) includes a fairly broad set of terms for document genres and information artifacts.

One difference between genres of Web pages and those of traditional media is that genres distinguished primarily by difference in physical characteristics (e.g., a brochure vs. a booklet vs. a flyer) were not very useful for classifying Web pages. Indeed, we used only a few of the terms listed by the *Art and Architecture Thesaurus* (Petersen, 1994) as information artifacts rather than genres. Instead, we focused more on the purpose of the information, such as product description, services, and so on. We believe it would be useful to characterize more precisely the common purposes for which information is distributed and how these are reflected in different genres. Such a guide would be particularly useful for a designer trying to determine how best to present a particular kind of information.

### Genre Systems or Multipage Documents

Many of our pages seemed to be parts of genre systems as opposed to independent pages. Clearly, a limitation of our study was the choice of single pages as the unit of analysis. As people start to build more interactive Web applications composed of multiple pages (e.g., to support internal and external processes rather than simply to broadcast information), defining genre systems will be more important. Conversely, we note the ease with which electronic documents can be reused. Therefore, future research should also address the question of embedded genres, as documents of one genre are reshaped and repurposed.

It might be useful to define the form of a hyperdocument (and thus the overall genre) by the pattern of links it exhibits. For example, a hotlist is a linear list of links, all of which go to other sites; an online book is (usually) a sequential list of chapters, each linked to the next, perhaps with a table of contents pointing to each chapter; and a hyperdocument has a pattern of densely interlinked pages. Similarly, a glossary could be identified by the links throughout a set of documents to regular points on a separate set of pages. Thus, the genre of a hyperdocument might be determined in part by examining how its component parts are linked together. We are currently applying this strategy in a study of FAQs (Crowston & Williams, 1999).

However, increasing the unit of analysis from the single page to the Web site presents significant difficulties. A key advantage of analyzing pages is that single pages are simple to identify. It is much more difficult to identify Web sites, for two reasons. First, since the same untyped HTML links serve to both tie together the pages of a single Web site and to refer to other sites, it is difficult to establish with precision the boundaries of a site. A link might be to another page in the site or to an entirely different site. Second, since links are unidirectional, it is impossible to determine whether a particular page is the "top" of a site or whether there are other pages that link to it and are part of the same site.

### Directed Sampling

Perhaps our biggest surprise was just how mundane our sample was. It included only a few radically new pages such as the video clip page. It is worth noting that our sample is imperfect for identifying innovative genres. A random sample of pages tends to return more leaf nodes simply because there are more leaf nodes than "top-level" pages. More heavily trafficked top-level pages might show more innovation in use of genres, which would not be reflected in our sample. For example, many sites appear to have adopted a "newspaper" genre at the top level but not for lower-level pages.

A future study might address some of the limitations imposed by the construction of our sample. At the time of our study, Alta Vista attempted to be a comprehensive index of Web pages, but given the growth of the Web,

we are unaware of any comparable current source (though Google may be close). As a result, it is no longer feasible to construct a truly random sample of current Web pages (if it ever was). Further, some of the most interesting pages may be hidden from the search engines used to create the database from which we drew our samples. In particular, Web search engines often do not index dynamically created pages. Therefore, we believe that it would be interesting and more practical to conduct a follow-up study using a theoretical sample of "interesting" pages. It would be quite difficult to create a truly representative sample of genres this way, but representativeness need not be a key concern in a study describing innovative uses of the Web and their antecedents or the process of adaptation.

## Implications for Web-Site Design

Finally, we believe that our research has some implications for the practice of Web-site design. First, designers may want to draw on accepted genres when they are appropriate for their purpose. One HTML design book takes this approach, offering samples of home pages for individuals and large and small organizations, brochures, surveys, and hotlists (Arronson, 1994). Most do not, focusing instead on the mechanics of formatting a page to the exclusion of the communicative intent. A surprising number of Web pages in our sample (and pages we have encountered since) are unintelligible. The literature suggests that if the designer of a page draws on accepted genres, it will be easier for readers to determine the communicative intent of the page. For example, a designer might look through the AAT (Petersen, 1994) list of genres by function to see how a particular communicative purpose has been expressed in the past.

Designers should also be aware of users' expectations of a genre. For example, readers are accustomed to tables of contents and indexes that list the entire contents of the book they appear in. It can be very disconcerting, therefore, to encounter site home pages that appear to be tables of contents or searches that appear to be indexes but that are incomplete or are actually hotlists. Of course, the diversity of audiences on the Web compounds the difficulty of meeting a reader's expectations. Even so, on too many pages we found it hard to determine the communicative intent, if any.

To moderate these problems, we believe that it is useful to make clear the genre of even a single Web page. With the growing use of indexing systems such as Excite, Lycos, or Alta Vista, it is not uncommon for a user to start reading in the middle of a document as we did, suggesting that the purpose and form of even a single page should be evident. In a physical document, pagination is determined by the physical dimensions of the book and so is not usually meaningful. On the other hand, division of a Web documents into pages should reflect the actual structure of the communication. In other words, a single Web page is more like an article pulled out of a magazine than a page pulled out of a book. However, we noted that numerous sites lacked navigational aids to help a reader figure out where they were in a longer document or the purpose of that communication. To continue the analogy, when photocopiers became common, many journals added citation information (journal title, date, page numbers) to the first page of each article so the source of an article could be determined from a copy. Web pages require similar appellation if they are to be useful by themselves.

On the other hand, the technology of the Web enables novel applications, such as a shift from static documents to "live" data. For example, Yan et al. (1996) describe how patterns of user access can be used to suggest which information should be viewed next. Designers should be free to modify or reject genres when it is necessary to take advantage of the technology. In creating novel applications, however, designers must be aware that new genres are often misunderstood or resisted. Therefore, more attention is needed to clearly define the community in which the communication makes sense and to identify already accepted genres that can serve as a bases for evolution. We believe that such explicit attention to genre will speed the wider acceptance of newly emerging genres of communication unique to the Web.

## REFERENCES

A World Wide Web for tout le monde. 1996, April 1. *Business Week*, p. 36.

Arronson, L. 1994. *HTML manual of style*. Emeryville, CA: ZD Press.

Bazerman, C. 1995. Systems of genres and the enactment of social intentions. In *Genre and the new rhetoric*, eds. A. Freedman and P. Medway, pp. 79–101. London: Taylor & Francis.

Bearman, D. 1994. *Electronic evidence: Strategies for managing records in contemporary organizations*. Pittsburgh: Archives and Museum Informatics.

Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H. F., and Secret, A. 1994. The World-Wide Web. *Communications of the ACM* 37(8):76–82.

Bibliographic Standards Committee of the Rare Books and Manuscripts Section, ACRL/ALA. 1991. *Genre terms: A thesaurus for use in rare book and special collections cataloguing*. Chicago: Association of College and Research Libraries.

Bitzer, L. F. 1968. The rhetorical situation. *Philosophy and Rhetoric* 1:1–14.

Bray, T. 1996. Measuring the Web. *Fifth International World Wide Web Conference*, May 6–10. Paris, France.

Campbell, K. K., and Jamieson, K. H., Eds. 1978. *Form and genre: Shaping rhetorical action*. Falls Church, VA: Speech Communication Association.

Crowston, K., and Williams, M. 1999. The effects of linking on genres of Web documents. *Thirty-second Hawaii Int. Conf. Systems Science (HICSS-32)*. January, Maui, Hawaii.

Dollar, C. 1994. Electronic objects circa 2001: Problems or opportunities? . . . Yes. *Conf. on Electronic Records in the New Millennium*, pp. 25–38. Vancouver, BC.

Freedman, A., and Medway, P. 1994. Locating genre studies: Antecedents and prospects. In *Genre and the new rhetoric*, eds. A. Freedman and P. Medway, pp. 1–22. London: Taylor & Francis.

Giddens, A. 1984. *The constitution of society: Outline of the theory of structuration*. Berkeley: University of California.

Harrell, J., and Linkugel, W. A. 1978. On rhetorical genre: An organizing perspective. *Philosophy and Rhetoric* 11:262–281.

Harter, S. P. 1998. Scholarly communication and electronic journals: An impact study. *Journal of the American Society for Information Science* 49(6):507–516.

Kling, R., and Covi, L. 1995. Electronic journals and legitimate media in the systems of scholarly communication. *The Information Society* 11(4):261–271.

Krulwich. n.d. *An agent of change*. Andersen Consulting Center for Strategic Technology Research, Northbrook, IL, USA. <http://bf.cstar.ac.com/bf/article1.html>

Library of Congress, Prints and Photographs Division. 1995. *Thesaurus for graphic materials*, vol. 2, *Genre and physical characteristic terms*. Washington, DC: Library of Congress, Cataloging Distribution Service.

Miller, C. R. 1984. Genre as social action. *Quarterly Journal of Speech* 70:151–167.

Orlikowski, W. J., and Yates, J. 1994. Genre repertoire: The structuring of communicative practices in organizations. *Administrative Sciences Quarterly* 33:541–574.

Orlikowski, W. J., Yates, J., Okamua, K., and Fujimoto, M. 1995. Shaping electronic communication: The metastructuring of technology in the context of use. *Organization Science* 6(4):423–444.

Pedram, M. 1997. *Lycos*. Kansas City Public Library. <http://www.kcpl.lib.mo.us/search/lycos.htm>

Petersen, T. 1994. *Art and architecture thesaurus*. New York: Oxford University Press.

Raymond, E. 1993. *The new Hacker's dictionary*, 2nd ed. Cambridge, MA: MIT Press.

Star, S. L., and Griesemer, J. R. 1989. Institutional ecology, "translations" and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. In *Social studies of science*, vol. 19, pp. 387–420. Newbury Park, CA: Sage.

Swales, J. M. 1990. *Genre analysis: English in academic and research settings*. New York: Cambridge University Press.

Woodruff, A., Aoki, P. M., Brewer, E., Gauthier, P., and Rowe, L. A. 1996. An investigation of documents from the World Wide Web. *Fifth International World Wide Web Conference*, May 6–10 Paris, France.

Yan, T. W., Jacobsen, M., Garcia-Molina, H., and Dayal, U. 1996. From user access patterns to dynamic hypertext linking. *Fifth International World Wide Web Conference*, May 6–10 Paris, France.

Yates, J., and Orlikowski, W. J. 1992. Genres of organizational communication: A structurational approach to studying communications and media. *Academy of Management Review* 17(2):299–326.

## APPENDIX

Characteristics of some identified genres, organized according to the *Art and Architecture Thesaurus* (Petersen, 1994)

| Genre | Characteristics of form and purpose |
|---|---|
| <information artifacts> | |
| <information artifacts by physical form> | |
| books | Long works (one or more volumes) on any topic, often divided into chapters, with a table of contents and index |
| <books by internal form> | |
| albums | |
| photograph albums | Albums made up of mounted photographs, with or without identifying information GAHLM (AAT) |
| pamphlets | Small treatises occupying fewer pages or sheets than would make a book, issued as separate works unbound, on subjects or questions of current or temporary interest, personal, social, political, ecclesiastical, or controversial, about which the writer desires to appeal to the public (OED) |
| <information artifacts by function> | |
| <identifying artifacts> | |
| <identifying cards> | |
| business cards | Small cards produced since the 19th century bearing the name and address of a business concern and one of its representatives, and intended more for information than advertisement (AAT) |
| <document genres> | |
| <document genres by form> | |
| catalogs | Enumerations of items, usually arranged systematically, with descriptive details; may be in book or pamphlet form, on cards, or online (AAT) |

## APPENDIX

Characteristics of some identified genres, organized according to the *Art and Architecture Thesaurus* (Petersen, 1994) *(Continued)*

| Genre | Characteristics of form and purpose |
|---|---|
| forms | |
| *comment forms* | Fill-in forms for comments about a product or service |
| order forms | Fill-in forms for information needed to arrange payment and shipment of a product |
| *URL submissions forms* | Fill-in forms for information about Web sites (e.g., for inclusion in a directory) |
| lists | |
| <lists by form or function> | |
| *bookmark lists* | Lists of Web sites not controlled by the list's author, unorganized |
| *course lists* | Lists of numbers, titles, and description of courses offered in some department or departments, possibly including information such as prerequisites, instructor, meeting time and place |
| discographies | Catalogues raisonnés of gramophone records; lists of the recordings of a single composer or performer (OED) |
| *e-mail lists* | Lists of links to e-mail messages |
| *FAQs* | Edited collections of questions and answers on some topic; often labeled as such |
| *file lists* | Lists of files in a computer directory |
| filmographies | Lists of the films of a particular director, producer, actor, etc., or of those dealing with any particular theme (OED) |
| *hotlists* | Lists of Web sites not controlled by the list's author, organized by topic |
| *projects lists* | Lists of research projects under way in some group |
| *publication lists* | Lists of publications by members of some group |
| software | |
| <*software by form*> | |
| *source codes* | Computer programs as written by the programmer. |
| <versions of documents> | |
| summaries | |
| abstracts | Brief summaries that provide the essential points of written works (AAT) |
| <document genres by form: partial documents> | |
| *indexes* | Alphabetical Lists, placed (usually) at the end of a book, of the names, subjects, etc., occurring in it, with indication of the places in which they occur (OED) |
| tables of contents | Summaries of the matters contained in a book, in the order in which they occur, usually placed at the beginning of the book (OED) |
| <document genres for literary works> | |
| essays | Compositions of moderate length on any particular subject (OED) |
| histories | |
| chronicles | Detailed and continuous registers of events in order of time; historical records, especially those in which the facts are narrated without philosophic treatment or any attempt at literary style (OED) |
| genealogies | Accounts of one's descent from an ancestor or ancestors, by enumeration of the intermediate persons; pedigrees (OED) |
| <document genres for oral or performed works> | |
| scripts (documents) | Typescripts of cinemas or television films; texts of broadcast announcements, talks, plays, or other material (OED) |
| <document genres by function> | |
| <declaratory document genres> | |

*(Continued on next page)*

## APPENDIX

Characteristics of some identified genres, organized according to the *Art and Architecture Thesaurus* (Petersen, 1994) *(Continued)*

| Genre | Characteristics of form and purpose |
|---|---|
| advertisements | |
| *classified advertisements* | Short paid announcements appearing in periodicals sorted according to the goods or services being offered or requested |
| announcements | Printed or published statements or notices that inform the reader of an event or other news (AAT) |
| *custom 404s* | Web pages announcing that requested Web pages cannot be found on the server |
| news bulletins | |
| press releases | Official or authoritative statements giving information for publication in newspapers or periodicals GAHLM (AAT) |
| *under construction* | Web pages announcing that particular Web pages are still in the process of being developed |
| *Web site moved* | Web pages announcing that requested Web pages have been moved to new URLs |
| *home pages* | Personal or organizational information plus links to other pages reflecting the subject's interests; intended to introduce the person or organization to the world and to facilitate further contact |
| manifestoes | |
| *political party platforms* | Public declarations of the principles and policies on which a political party proposes to stand (OED) |
| testimonials | Letters of recommendation of a person or thing (OED) |
| eulogies | Articles in which memories of a person, thing, or event are preserved |
| instructional materials | |
| guidebooks | Books of information on places or objects of interest in a locality, city, building, etc. (OED) |
| instructions | Information in the form of outlines of procedures, as directions or commands GAHLM (AAT) |
| manuals | Books, usually of compact size, containing concise information, often rules or instructions needed to perform tasks or processes (AAT) |
| *computer manuals* | Instructions on the use of a computer program or of a computer programming system |
| *problem sets* | Collections of problems assigned in a course |
| prospectuses | Documents designed to inform prospective investors or patrons, consisting of a formal summary of a proposed commercial, literary, or other venture, or a description of the features or services of an institution or business (AAT) |
| *ratings* | Numeric evaluations of products or services |
| records | |
| <records by form or function> | |
| accession records | Records documenting additions to a collection, whether acquired by transfer under a legally based procedure or by deposit, purchase, gift, or bequest. ICA (AAT) |
| accessions registers | Lists of books acquired, e.g., during a particular time period |
| administrative records | |
| minutes | Records of the proceedings at a meeting of an assembly, corporate body, society, company, committee, or the like (OED) |
| personnel records | |
| résumés | Brief accounts of people's careers (OED) |
| rules (instructions) | Rule prescribed for the management of some matter or for the regulating of conduct; governing precepts or directions; standing rules (OED) |

## APPENDIX

Characteristics of some identified genres, organized according to the *Art and Architecture Thesaurus* (Petersen, 1994) *(Continued)*

| Genre | Characteristics of form and purpose |
|---|---|
| architectural records | |
| contract documents | |
| specifications | Detailed and exact statements of particulars, especially statements prescribing materials, dimensions, and workmanship for something to be built, installed, or manufactured AHD (AAT) |
| census records | |
| *demographic data* | Data on the characteristics or composition of a population |
| <records by provenance> | |
| school records | |
| *course descriptions* | Brief summaries that provide the essential points to be covered in a course |
| *sports records* | |
| *box scores* | Records of players and plays in a game |
| *Web server statistics* | Records of accesses to a Web server |
| <reference sources> | |
| directories | Books containing one or more alphabetical lists of the inhabitants of any locality, with their addresses and occupations; also a similar compilation dealing with the members of a particular profession, trade, or association (OED) |
| <directories by subject> | |
| *faculty directories* | Short descriptions of individual faculty members |
| *topical home pages* | Web pages providing overviews or introductions to a topic along with links to other Web pages about that topic |
| reports | Formal statements of the results of an investigation or of any matter on which definite information is required, made by persons or bodies instructed or required to do so (OED) |
| reviews | |
| *concert reviews* | Retrospective descriptions of concerts |
| *film reviews* | Descriptions and evaluations of films to advise potential viewers |
| *product reviews* | Descriptions and evaluations of products to advise potential purchasers |
| <document genres by conditions of production> | |
| articles | Literary compositions forming part of a journal, magazine, encyclopedia, or other collection but treating specific topics distinctly and independently (OED) |
| newspaper columns | Special features, especially a regular series of articles or reports (OED) |
| *newswire articles* | Articles prepared for news wires, with a dateline and news content |
| publications | |
| newsletters | Medium-length works, including multiple articles, titled, issued periodically (e.g., with date or volume), communicating "the news of the day," distributed periodically to subscribers or members of an organization |

*Note*: Terms are organized in a hierarchy, with more specific terms indented below the more general term of which they are a special case (e.g., "order forms" are a special kind of "forms"). Terms in <angle brackets> are "guide terms," that is, terms used to bring together related genres when no appropriate general term exists. Terms in regular type were found in the *Art and Architecture Thesaurus*; terms in *italics* were added by the authors. Some descriptions were drawn from the *Oxford English Dictionary* (labeled OED) or from the *Art and Architecture Thesaurus* scope notes (labeled AAT). Others were developed by the authors. General terms are not described if more specific terms were used instead. From *Art and Architecture Thesaurus* (Petersen, 1994), *Oxford English Dictionary*, and Crowston and Williams (1999).