

Hotlist or Bibliography? A Case of Genre on the Web

David G. Hendry and Allyson Carlyle
Information School
University of Washington
Box 352840
Seattle, WA 98195-2840
{dhendry,acarlyle}@u.washington.edu

Abstract

Bibliography—the collection, description, and arrangement of information items—is a vibrant activity and a popular form of expression on the internet. Yet, fashionable forms of list making are rarely, if ever, considered bibliography. This is a missed opportunity. In this paper, we examine how the traditions of bibliography and collection development can be used to inform the creation of genre support systems and to inform evaluation of a research into bibliographies. Then, we extend the traditional view of bibliography for the internet and show how this new definition can be used to clarify new systems. This work is a small step towards reinvigorating the practices of bibliography and applying its conceptual power to clarify and undergird information systems for selecting, arranging, and accessing links.

1. Introduction

“The Advanced Hotlist Manager manages your list of links to Internet resources, which is referred to as your *hotlist* or your *bookmarks*. These resources are typically HTML documents, but you can include links to FTP sites, image files, audio and movie files, gopher servers, and search engines.” (Mosaic 2.1 User Documentation, January 1996 [2])

The diversity of information artifacts whose primary purpose is to provide lists of links to electronic material is extraordinary. Consider, for example, a few connotative terms: Hotlist, Blog, Resource Page, Directory, Gateway, Portal, Digital Library, and Social Bookmarks. These are not technical terms; certainly, authoritative, agreed definitions for materials covered by these terms do not exist. Yet, except for efforts to formalize the

underlying infrastructure of such materials [14, 29], this sort of vocabulary is the best we have for discerning differences and identifying similarities amongst popular and evidently important categories of information artifacts.

Prima facie these examples seem to be quite different. To pinpoint the differences, we could invoke the notion of genre and examine the differences in form, communicative purpose, and audience [3]. Erickson [10] defines a genre as: “a patterning of communication created by a combination of the individual, social and technical forces implicit in a recurring communicative situation. A genre structures communication by creating shared expectations about the form and content of the interaction, thus easing the burden of production and interpretation.” By analyzing particular instances of documents, the genre approach allows the mediating powers of particular recurrent textual forms to be understood at multiple levels of scope, from the cognitive to the cultural [30]. It is an empirical approach that has led to incisive analyses [9, 17, 28, 33] and much more empirical work remains to be done because of the size and quickly evolving nature of the genre ecology of the internet [11].

In this paper, we seek to understand the terrain marked by the above vocabulary though traditions that are supported by a rigorous conceptual base. These traditions are *bibliography*, the practice of creating lists of documents, and collection development, the practices and policies surrounding the creation of library *collections*. We shall see that the internet has thrown new adaptive forces upon the traditional, formalized, nature of bibliography. In turn, these forces have radically changed the practice of list making and how people use lists of documents for information discovery. Bibliography, in short, is a lost antecedent for many of the fashionable information genres that we see today on the web.

Three popular activities motivate this analysis. First, list making, in all its varied forms, appears to be very popular on the internet. In January 2004, over 50 million people in the USA published content to the web [24]. It seems quite plausible, therefore, that perhaps millions of people are creating lists of items that are, in turn, used by other people. Much as Vannevar Bush envisioned with his metaphor of trails, people are highlighting and providing access to a wide variety of materials for the benefit of each other.

Second, when lists contain citations to other lists, they combine to form a patchy network of secondary access points. Algorithms such as Page Rank [8], have proven the value of these citation networks for estimating the quality of the cited documents. Thus, in large measure the effectiveness of search engines is due to the individual efforts of list makers.

Third, the creation of lists by individuals and groups supports the development of communities of interest. For example, structural analyses of the web, using graph-algorithmic techniques, have found a fair degree of structure [18]. In fact, using data from 1998, researchers discovered approximately one hundred-thousand highly cohesive communities, such as resources on “oil spills off the coast of Japan” [19]. The availability of these links on such narrow topics supports socialization and discussion.

In sum, bibliography underlies the above three important activities: personal expression, automated information discovery, and discourse around topics of interest. Yet, bibliography, its conceptual base and practices, is not being used deliberately to inform the development and evaluation of systems in support of these activities nor is it being used to clarify the evolving practices of the millions of list makers, working outside of professional settings. This paper is a small step towards reinvigorating this rich tradition and using its historical perspective to analyze this rapidly evolving terrain.

In the next section, we establish the background for this analysis by reviewing the notions of bibliography as it applies to print. We also argue that linked bibliographies on the internet are collections, and look to the practice of building library collections for its theoretical contributions. Then, in section 3, we examine how the communication facilities of the internet have transformed the traditional nature of bibliography and introduce the notion of collaborative bibliography and bibliography as information system. Next, in section 4, we use the notion of collaborative bibliography to analyze two examples, showing how this framework can be used to clarify various kinds of collecting. Finally, we conclude with some

observations about the continuing, long-term importance of bibliography.

2. The tradition: Print background and definitions

2.1 Bibliography

Bibliography is an ancient genre. Lists of references to documents have probably existed for almost as long as documents themselves have existed. Records of early bibliographies date back to at least the third century B.C.[6]. In addition, the forms or subgenres that exist are as diverse as their history is long. Print-based bibliographies vary in length and scope from *Bedtime Stories for Children* [25], a brief, contemporary list of bedtime stories, to the massive *Sikuquanshu*, an 18th century Chinese project that involved the creation of references and annotations for every important book written in China available at the time [32]. The creation of bibliographies is so widespread a practice that even in the print environment defining what exactly a bibliography *is* is a somewhat daunting and even controversial task. The extension of the genre to the internet makes this even more challenging.

The types of resources listed in the first paragraph of this paper are ones that we would consider to be bibliographies. As stated above, they vary widely in form, communicative purpose, and audience. We would argue, however, that the widely varying forms are not unlike print forms that we would also identify as bibliographies.

One of the difficulties in laying a claim to the term “bibliography” is a large body of scholarly practice and a formal discipline called *Bibliography*. This discipline is supported by a number of highly regarded manuals and textbooks [7, 31]. The definitions supported within the discipline restrict what may be identified as a bibliography, although total agreement on a definition does not exist. A list such as a hotlist, for example, would not be considered by some within this discipline to be a bibliography because of inadequate information being included in the references.

For a variety of reasons explained below, we prefer to use the most simple definition, “a list of documents or resources”, and refer the reader to Stokes [31] for a

Table 1. Specifications for a bibliography [5].

-
1. *Domain*: the territory from which items are drawn (e.g., the domain of many internet bibliographies is the set of resources searched by Google on such and such a date at such and such a time)
 2. *Scope*: the subject or conceptual area of the bibliography
 3. *Selection rules*: the rules specifying exactly which items will be selected for the bibliography
 4. *Bibliographic units*: the document levels included (e.g., single web page, website, subdivision of a website, or collection of similar websites, or all)
 5. *Information fields*: the types of information included in the resource references (e.g., title, date published, URL, etc.)
 6. *Organization*: the manner in which the references in the bibliography will be organized (e.g., alphabetical, chronological, topical, etc.)
-

discussion of the definitional issues in scholarly practice.

One reason we prefer a simple definition of bibliography is that it gives us the ability to link the print genre to the internet genre. Doing this offers a conceptual base for research and evaluation. In addition, this orientation can be used to frame the development of *genre support systems*, a term we use to refer to systems that allow people to create and manage specific classes of electronic documents (or genre). Blogger.com, for example, is a system that allows a person to create a blog, to specify a visual style, to post items, and to maintain it over time. It is a special case of what we have in mind. In the future, we see the development of more general systems that will allow people to design blogs but also a range of similar systems, including bibliographies. As such, genre support systems are a meta design tool—they support the evolution of genres [13].

The principles at work in the discipline of bibliography have been laid out by Bates [5] as "specifications" for a bibliography. In this article, Bates pulls together traditional principles of

Table 2. Collection development guidelines [12].

-
1. Identify the community that the collection is being developed for and the uses to which the collection will be put
 2. Deliberately consider the subject areas (similar to the scope specification in a bibliography) and formats (similar to, but not exactly the same as bibliographic units) to be included in the collection
 3. Describe roles and managerial responsibilities for making selection decisions for items to be included in the collection
 4. Describe the selection constraints, for example, books receiving negative reviews will not be selected (similar to selection rules)
 5. Perform regular evaluations for quality
-

bibliography creation and adds a new one. She argues that the list of specifications, shown in Table 1, should be recorded with the bibliography to guide the user.

Agre argues that "[d]esign for new media ... requires some rational understanding of who are using the materials, what they are doing with them, and how they fit into an overall way of life." [3]. We would argue further that design for digital genres that have a history in print also requires an understanding of the history and context of those print genres. For instance, in this case the historical frameworks supporting the making of bibliographies, the specifications, can be extremely useful in designing genre-support software, particularly for collaborative environments [16]. In addition, they can be used to frame research methodologies studying bibliographies as a digital genre, and to evaluate the composition or clarity of those bibliographies. Finally, they can be used to help novice bibliography creators understand how to make digital documents with potentially greater impact and reach.

2.2 Collections and Collection Development

One of the most exciting developments in the transition of bibliographies from print to digital is their metamorphosis from lists of references leading to laborious searches for documents, to lists of references that place the reader within instant reach of the documents they refer to. In this way, the hyperlinked bibliography broadens the scope of the genre to include the library notion of "collection."

Library collections are physical groupings of items at a location or group of locations to support the educational, recreational, and information needs of a specific community. In a sense, in the internet environment, the genre of bibliography has merged with the practice of building collections to become a hybrid genre. Thus, a hyperlinked list of references acts simultaneously as a bibliography and a collection – the list of references itself, through links, makes physical documents available to a group of users or, again, a specific community.

Library collections have also been developed with the support of specific conceptual frameworks and scholarly practice, formally called *Collection Development*. The incorporation of some of the principles and practice of collection development also have the potential to lend powerful support to the development of software supporting bibliography creation as well as informing research and evaluation. Like defining bibliography, defining collection is somewhat controversial [21-23]. For example, is a Hotlist or a Gateway a collection? Some, including Lee [23], would say no. But here we will again adopt a very loose definition in order to take advantage of what the historical antecedents have to offer.

The social and political contexts surrounding the building of library collections have led to precise guidelines that result in collection development policies. Guidelines typically considered in developing collections are shown in Table 2.

The overlap between the conceptual territory covered by principles used in bibliography and in collection development is obvious—it somehow seems unsurprising that the two, which have developed in almost completely separate realms in the same academic community, may now be seen to merge in the digital environment.

One of the important contributions of considering digital bibliographies to be collections is to incorporate the consideration of user community. For whom is this list of resources being compiled? What purposes will be served by this list? Who is allowed to make decisions about what is to be included and what is not? Who will evaluate the effectiveness of the list, and how will be evaluated? Again, these questions may inform genre support systems, and aid in research and evaluation surrounding bibliographies.

2.3 The social context of bibliography

As Agre said, social contexts are critical in understanding and developing genres. Another change in the transition of the bibliography from print to digital involves the social context in which bibliographies are created. Print bibliographies, particularly book-length bibliographies that fit the disciplinary definitions of bibliography, were frequently created by a single scholar or researcher who had total control over the specifications and means by which the bibliography was created. In the digital environment, bibliography has become a genre that can be and is produced collaboratively. The extent to which that expands our conception of the genre is described below.

3. The transformation: Collaborative bibliography and bibliography as information system

In the print environment, bibliographies are clearly artifacts. In the transition of the bibliography to the internet, it can be argued that bibliographies are, in addition, information systems. Indeed, the last three of Bates' specifications concern presentation and lead to functional requirements for how items are to be stored and accessed. The first three of the specifications are more subtle. When defined and deliberately used by a selector or group of selectors, they can help improve the rigor of individual section decisions and, in turn, the coherence of the bibliography. And when made visible, they provide an explicit rationale to the users of the bibliography. As we shall see, on the internet these first three specifications assume great importance for supporting various forms of collaboration; yet, they are typically not articulated.

As described above, bibliography has traditionally been practiced by a single compiler who rigorously selects items. In aggregate, these items establish the bounds for a topic and elaborate it. A bibliography is an instrument. As such, it enables people to discover useful resources efficiently but it also frames a conceptual space, the domain, of a topic. This framing produces instrumental power by *collocating* items in useful presentations. Therefore, the quality of a bibliography should be judged, first, by the degree that its items are on topic but, more importantly, also, by the degree to which the collocation of items is innovative. It is one thing to establish boundaries and elaborate a conceptual space as an individual; it is quite a different matter to engage others

collaboratively in this work. Yet, the internet begs for this to occur.

We saw that in stable environments, such as libraries, collection development policies are used to engage multiple stakeholders in selection decisions. The policies enable stakeholders to participate in rational decision-making. When items are costly and budgets limited, prioritizing acquisitions are critical. But even when resources are free, time is always a limiting factor for both selectors and users; thus, prioritization remains essential.

Nevertheless, the internet, through its diverse modes of communication technology, makes virtual collaboration in the process of genre creation a real possibility, and this has the power to change how people collaborate in decision making. This potential for collaborating in the selecting of items is a major transformative challenge confronting bibliography. We shall see that this potential can be envisioned as a bricolage of interactions around familiar tools or it can be envisioned as a common pattern on a specific system. We examine this challenge in the following sections, seeking to illustrate how the choices of the internet potentially empower the process of bibliography creation but also complicate it.

3.1 Framework of bibliography-like systems

Figure 1 charts the terrain of bibliography-like information systems. In related work [16], we describe some twenty different examples from this simple design space. Here, we wish to delineate four different regions in this space and illustrate the diverse range of approaches that enable groups of people to participate collaboratively in selection decisions.

In region A, a very large number of people participate in the selection of relatively small number of items. The Lycos Fifty⁽⁰¹⁾ (50.lycos.com), Google Zeitgeist⁽⁰²⁾ (www.google.com/press/zeitgeist.html), and Technorati⁽⁰³⁾ (www.technorati.com) are examples from this region. By analyzing the behaviors of millions of people (e.g., the queries submitted to search engines or items discussed in blogs), systems in this region are particularly suited to spotting trends. Certainly, these systems stretch the meaning of bibliography, but charting them here helps to establish the bounds for systems that have a strong bibliography orientation.

In region B, we see systems that generally enable incremental growth. As more people join the systems, the sizes of the collections tend to grow. Live Journal⁽⁰⁴⁾ (www.livejournal.com) and Yahoo! Groups⁽⁰⁵⁾ (groups.yahoo.com) are examples of systems where millions of people create localized lists

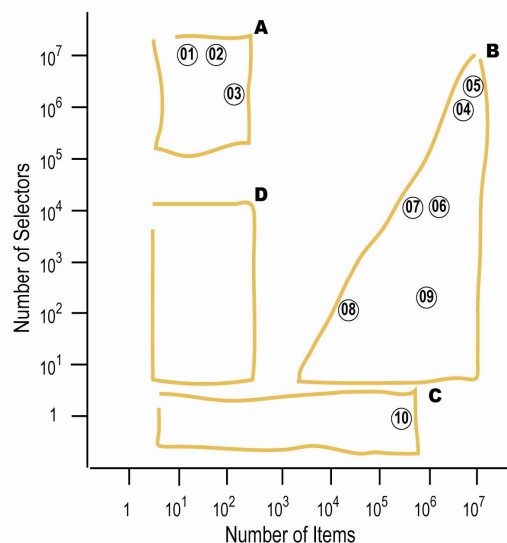


Figure 1. Design space for bibliography-like information systems

of links that, in aggregate, create structured collections of impressive overall size. The systems in this region reveal a tremendous diversity of approaches for enlisting the efforts of others. Some systems organize contributors into a hierarchy to divide labor and to impose a system of top-down review (e.g., Open Directory Project⁽⁰⁶⁾ ~ www.dmoz.org). Other systems are reactive in that they are open to all contributors but through a combination of open-review and revision history interested writers or editors are able to coordinate efforts and maintain quality (e.g., the Wikipedia⁽⁰⁷⁾ ~ www.wikipedia.org is a good example of this approach that is applicable to bibliographic materials). Other systems, informed by collection development policies from libraries, have borrowed highly structured processes to coordinate contributors and enable centralized review (e.g., Librarian's Index to the Internet⁽⁰⁸⁾ ~ www.lii.org). About.com⁽⁰⁹⁾ (www.about.com), while a commercial enterprise, follows a similarly structured approach where contributions are carefully controlled. Many other examples could be plotted in this region.

In Region C, we see the case where a single person, or small number of loosely coordinated people, creates collections. At the extreme end, Cyndi's List⁽¹⁰⁾ of links to genealogy (www.cyndislist.com) is truly impressive: One person created and manages a directory of approximately 250,000 links using a word processor and no other special tools. More typically, we can find many examples of collections in the range of size 10^2 to 10^3 that are created by individuals. While it is difficult to

estimate an absolute number, it seems likely that Region C has the greatest density of examples in this chart.

Finally, in Region D, we see the case where a moderate number of people create collections of relatively small size. Systems for peer review of papers at conferences are perhaps the classic example. Yahoo Groups, when operated in private mode, is the best popular example that we know of.

3.2 Discussion of Framework of bibliography-like systems

If a bibliography is a container, to use a conceptual metaphor, then a selector can decide to either place an item in a container or to dismiss the item as irrelevant and possibly consider other containers. In traditional bibliography, this decision-making process is most often dominated by a single person's judgments [5], ideally guided by rigorous criteria. On the other hand, the systems shown in Figure 1 show great diversity for how selection decisions are mediated amongst people.

This diversity can be approached by examining the different sources of authority that can be conferred to people. First, at the container level, various coordinative activities can take place along a path that begins with a request, such as 'Include item, I, in container, C', and ends with this request being fulfilled or declined. The length and intricacy of this path can vary enormously.

A selector can elicit items from an audience and then make unilateral decisions, possibly informed by a sample of ratings or reviews. Or, the item might be selected through an extended period of joint deliberation and consensus making. Once an item is selected, it will often be necessary to perform additional work, such as adding metadata or rechecking it periodically into the future. The stages and decision-points that make up the selection process, obviously, have a great influence on the quality and character of the items selected; indeed, we believe that differences amongst systems can be best pinpointed by revealing the differences in this selection process that each follow.

The selection process is one of the stages of bibliography creation that can be most assisted by the clarity that the traditional frameworks of bibliography and collection development has to offer, for example, by making explicit the rules that govern the inclusion of an item, or highlighting the importance of satisfying the needs of a specific user group. In fact, the specifications could themselves be incorporated

into what we are calling genre support systems, to facilitate the distribution of selection responsibilities.

Second, at the structural level, if we assume that the containers are fixed and organized into some kind of classification system, then a person can be given authority over one or more containers in the classification. If containers are nested, then authority can optionally be inherited from above. Decisions about how containers are structured and how authority is distributed over this organization thus place bounds on how a system will scale. Further, the manner in which selectors inform each other about the scope of their containers—and negotiate and modify their scopes—will significantly influence the overall coherence of the system and its ability to adapt over time. Again, systems to support this process are highly desirable.

And finally, at a policy level, a selector might be given rights to delegate authority for making selections to others. Even more, the selector might be able to change the selection policy and, for example, change a prescriptive, staged process (e.g., Librarian's Index to the Internet) into an open, reactive process (e.g., Wikipedia) or vice versa. This third level of authority, the most powerful form, allows selectors to set policy and indeed to engage in meta-design [13]. We are unaware of any systems in this space that enable this degree of change; however, the ability to make significant modifications appears important for long term sustainability. Certainly, sustainability is a vital concern—fossilized lists are quite common.

3.3 Workflow and coordination policies

We now return to the container level and examine how the selection process might be abstracted into a relatively general framework from which bibliography-specific selection processes might be derived. To begin, consider the following motivating problem scenario:

Top-Ten Problem Scenario. A researcher in Ubiquitous Computing wishes to create a bibliography of the ten central papers in the field. Together, the papers should map the field and enable researchers in neighboring fields to recognize the applicability of their specialties.

Many approaches for solving this problem exist. The researcher, for example, might select a sample of 30 sources and ask a small group of trusted colleagues to rate the items. Based on the ratings and perhaps additional unsolicited items he or she would make a final selection. This approach represents *one* specific kind of participatory policy.

By changing who participates in the process, how feedback is elicited and aggregated, and so on various

kinds of bibliographies can be envisioned. First, an *open bibliography* would allow anyone to submit and rate items. Items can be tagged with keywords from a controlled, or open-ended, vocabulary for access by keyword and aggregate ratings.

Second, a *moderated bibliography* would enable a single reviewer, according to a schedule of rotating reviewers, to judge items. Items might pass through a sequence of states, *submitted*, *accepted* or *declined* and *published*. Items are placed in one of several categories, serving to trigger the reviewer to take the next step.

And finally, a *peer-reviewed bibliography* would enable a group of reviewers, selected randomly from a pool for qualified people, to blind-review submitted items. Here, items would pass through a different set of states, *submitted*, *queued-for-blind-review*, *being-discussed*, *accepted* or *declined*, and so on. To divide labor, items are arranged in a directory of topics. These examples show that by amending a bibliography with a specific selection process, it is possible to readily mutate the traditional definitions of the genre.

Each of these solutions to the scenario make a three-part move: 1) Associate a status indicator with an item; 2) Specify rules for how an item is allowed to pass from one status to another; and 3) Describe the work to be completed when an item takes on a particular status. This approach falls neatly within the Workflow Management Systems framework [1]. This evolving, newly rejuvenated, technical approach offers various models, methods, and systems for modeling business processes as a network of tasks, which are each triggered when well-defined conditions hold. Workflow systems are notorious for their inflexibility which can lead to inefficiencies but recent research has sought remedies through techniques that enable workflow to be modeled in a flexible and adaptive fashion [26, 27]. In digital libraries, for example, creating workflow models for collection development is seen as a crucial feature, and formal approaches have been proposed to create general solutions [4, 14].

So, too, with bibliographies—workflow is critical. Differences in workflow models can be seen to underlie the open, moderated, and peer-reviewed subgenres sketched above. The central feature of each of these systems is a list of citations to documents (i.e., web pages) and it is this feature, which is relatively stable across media, that identifies these systems as bibliographies. Similarly, in conceptual terms, the collaborative selection policies are media independent. However, enacting these policies only became feasible with the advent of the internet, where the convenience of communication makes it possible

to experiment with various forms of joint decision making. In sum, it is through the specification and enactment of workflow policies that new subgenres of bibliography are spawned.

We believe that this reshaping of the bibliography through workflows that represent collaborative selection policies will be a long-term phenomena. Indeed, the workflows associated with the systems such as those identified in Figure 1 enable stakeholders to take on various roles and to coordinate work in very intricate ways. The difficulty, however, is that the workflow policies are very tightly bound to the system and it can be difficult to disentangle the workflow that enables coordination from the system's structural, navigational, and presentational features. If we better understood the recurrent forms of workflow and if we were able to associate specific workflows with collections then we predict an even more rapid expansion of genres on the internet, allowing systems to be more responsive to localized needs for information collection and sharing.

The barrier, which prevents people without technical resources from exploring collaborative bibliography, lies with end-user programming [13]. While we can readily conceptualize various workflow policies, such as those sketched above, it is extremely difficult to implement support systems that allow for the flexible specification and enactment of coordination policies by non-specialists [16]. Under the current state of the art, implementing even these comparatively simple policies requires a web development project, putting the exploration of these options out of the reach of most people. In short, the conceptual vocabulary and implementation tools for creating collaborative bibliographies lag behind our abilities to conceive solution scenarios. The technical challenge is to narrow the gulf between policy specification, on the one hand, and system implementation on the other.

4. Examples

4.1 Social bookmarking

Social bookmarking is an interesting approach for information management and sharing, allowing people to assign keywords, which act as references in the traditional sense of bibliography, to web pages and, in turn, allowing people to inspect a list of page citations for a tag. Further, when the tags and page citations are made public, a community of users can benefit, or not, from everyone's tagging behavior. Thus, tagging web pages creates a secondary space of keywords for mediating knowledge representation and

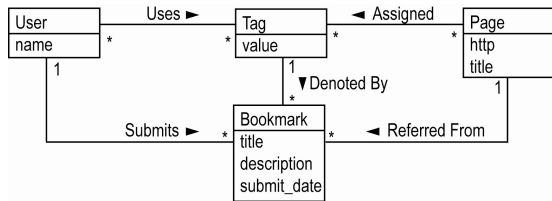


Figure 2. Simplified data model for social bookmarking systems.

information discovery. A number of systems for social bookmarking, some popular and some academically oriented, have emerged in the past two years [15].

Figure 2 presents a simplified data model for a social bookmarking support system. Inspection of the figure reveals that the following information presentations are readily supported when the model is deployed in an open fashion: 1) Show all BOOKMARKS that a USER has submitted; 2) Show all BOOKMARKS that a USER has submitted under a given TAG; 3) Show all TAGS that a USER uses; 4) Show all PAGES that have been assigned a given TAG; and 5) Show all TAGS that have been assigned to a PAGE. These entities provide a set of interconnected views of internet resources.

To illustrate this first consider two examples of *information discovery*. First, if a user discovers a relevant page and notices that an unexpected tag, T, has been assigned to it, she can test the conjecture “Perhaps other pages that have been tagged with T are also relevant.” Or, if a user discovers a relevant page and notices that other users, U1 and U2, have also tagged that page, she can test the conjecture “Perhaps U1 and U2 have tagged pages that are also relevant to me.”

Next, consider two examples of emergent and *dynamic knowledge representation*. First, if a user notices that several other people have tagged a page with T1 but she has used a different tag T2, she can assert “I will rename my T2 to the community’s T1 to better reflect conventional use” Or in the reverse, if a user notices that several people are using tag T in a sense significantly different from her usage, she can assert “I will rename T to a more obscure term” to create a degree of privacy. By navigating through views and by renaming tags, we can see that various interesting strategies for adaptive knowledge representation and information discovery are available through social bookmarks.

Tags are an expressive form that can be used for positive and mischievous purposes. For example, to make a point or to be abusive, a user, U, could tag particular kinds of pages with commonly known tags,

leading to a spamming effect. To counter this effect, a support system could suppress the presentation of all bookmarks submitted by U. Or, more powerfully, it might show bookmarks only for those users that the owners have placed a particular GROUP. This grouping abstraction would allow construction of a ‘name space’ of limited size, lying in region D of Figure 1.

Even within a defined community, tags have a range of expressive properties. A vocabulary of tags could be defined and, by prior agreement, used in a disciplined fashion to make selection decisions. Or, tags might be used haphazardly. Tags might be used to denote specific, carefully scoped, containers for storing bookmarks, to denote provisional containers whose scope will emerge over time, or to simply enrich the term space for pointing to web pages. Quite clearly, Bates’ specifications underlie each of these possibilities but with the twist that selection decisions of many people establish the scope. Current social bookmarking services, however, do not enable people to stipulate selection rules and scope. Just as Bates pointed out some 30 years ago with print bibliographies, users are forced to unearth the scope and selection rules from the items themselves. Nevertheless, when tags are used by many people to identify containers of provisional or emerging scopes, new forms of joint decision making become possible.

In sum, this brief analysis shows that the issues addressed by traditional bibliography must also be addressed by systems for social bookmarking. By analyzing such fashionable, innovative systems against traditional bibliography one gains a greater appreciation for what is truly new, what is given up, and what ought to be held onto.

4.2 Folk collections

The internet contains many examples of *folk collections*, collections of items, often photographs, on a particular topic created mostly for fun and interest. Often such collections are quirky. One of our favorites is www.drainspotting.com. At this website, you can browse a collection of over a 1,000 manhole covers from around the world. Some one hundred people have uploaded photographs.

By default, the photographs seem to be arranged by date uploaded, which is not a very informative way of browsing the collection. It is also possible to browse all images submitted by a particular photographer. Perhaps more usefully, people who upload photographs can assign keywords to them. Then, one can view all manholes that appear under, for example, the tags *universityofwashington*, *hexagon*, or *fireworks*. Josh Larios, the creator of the

site, explains the intended meaning of some of the tags, saying in part: "What's the difference between [the tag] *grid* and *waffle*? Good question; I'm not entirely sure myself..." [20]

Once a list of manhole covers is shown for a particular tag, users can click on a specific image to view details about it, including other keywords. In turn, these keywords can be clicked to browse further categories. Like blogs, users are also able to leave behind comments about the images. It is not known if these users are able to modify keywords once they are entered into the system.

Why discuss drainspotting.com? First, it illustrates the challenge of genre definition and identification on the internet. Is it a collection, a bibliography, or both? The images of manhole covers are not, themselves physical items; they are references to those items. Viewed as a list of references, drainspotting.com may thus be defined as a bibliography.

Second, while one might be inclined to dismiss it as trivial, we believe that it illustrates some very general points. For example, we consider drainspotting.com to be a special case of a social bookmarking system, where the most important distinguishing feature is that the authority to assign keywords to images is restricted to the submitter of the image. This is clearly a choice concerning workflow. In addition, the site itself—that is, images of manhole covers—establishes the scope of the collection; thus, selection decisions at the collection level are relatively straightforward. The proposed item is either an image of a manhole cover or it is not! Nevertheless, as the collection of images grows, second order decisions about what items to collocate become problematic. Here, specifications on scope and selection rules are crucial even for manhole covers.

In addition, a significant level of programming is required to develop drainspotting.com. This is a barrier that prevents non-programmers from readily participating in the development of such systems. Yet, surely, there is a significant unfulfilled demand to design a subgenre such as this, to seed it, and to let it evolve just as Josh Larios has done.

Finally, scholarship, social activism, and community engagement are often sparked by a collection of interesting materials. In drainspotting.com, we see the potential for enlisting the efforts of many ordinary people in the collection of references that might, in turn, be used by others in social or scholarly pursuits. Giving users the power to specify workflow policies and to seed bibliographies, offers splendid opportunities for community involvement and scholarship. Yet, to take full advantage of these opportunities the gulf between

envisioning possibilities for collaboration and implementing a specific system must be narrowed. We hope that this brief analysis shows the benefits of examining folk collections, even the quirky ones.

5. Conclusion

Bibliography is a genre with a rich history that is thriving in the current electronic environment. Further, it is a genre supported by a substantial body of principles and practice. The internet is transforming the genre and the processes surrounding it in fascinating ways that will become clear only through more empirical and analytical work. Given its ancient history and traditions, we believe that bibliography can help us better understand and nurture its many current permutations. The various forms of collaboration, supported by embedded models of workflow present splendid opportunities for new subgenres, with collaborative features, to evolve.

10. References

- [1] Aalst, W. V. D. and Hee, K. V., *Workflow Management: Models, Methods, and Systems*. Cambridge, MA: MIT Press, 2002.
- [2] Advanced Hotlist Manager, Mosaic 2.1 User Documentation, January, 1996, Retrieved June 1, 2005 from http://archive.ncsa.uiuc.edu/SDG/Software/mosaic-w/releaseinfo/2.1/WBook_42.html#HEADING59
- [3] Agre, P. E., "Designing Genres for New Media: Social, Economic, and Political Contexts," in *Cybersociety 2.0: Revisiting Computer-Mediated Communication and Community*, S. G. Jones, Ed. Thousand Oaks: Sage, 1998, pp. 69-98.
- [4] Anderson, K. M., Anderson, A., Wadhvani, N., and Bartolo, L. M., "Metis: Lightweight, flexible, and web-baes workflow services for digital libraries.," presented at Proceedings of the 3rd Joint Conference on Digital Libraries, Houston, TX, 2003.
- [5] Bates, M., "Rigorous Systematic Bibliography," RQ, vol. 16, pp. 7-26, 1974.
- [6] Blum, R., *Bibliographia: An Inquiry into its Definition and Designations*. Dawson: American Library Association, 1980.
- [7] Bowers, F. T., *Principles of Bibliographical Description*. Princeton: Princeton University Press, 1949.
- [8] Brin, S. and Page, L., "The anatomy of a large-scale hypertextual search engine," presented at Seventh International World Wide Web Conference, 1998.
- [9] Crowston, K. and Williams, M., "Reproduced and emergent genres of communication on the World-Wide Web," presented at Annual Hawaii International Conference on System Sciences, 1997.
- [10] Erickson, T., "Making Sense of Computer-Mediated Communication (CMC): Conversations as Genres, CMC

Systems as Genre Ecologies", Retrieved Sept. 1, 2005 from http://www.pliant.org/personal/Tom_Erickson/genreEcologies.html

[11] Erickson, T., "Social Interaction on the Net: Virtual Community as Participatory Genre," presented at Annual Hawaii International Conference on System Sciences, 1997.

[12] Evans, G. E., *Developing Library and Information Center Collections*, 4th ed. Englewood, CO: Libraries Unlimited, 2000.

[13] Fischer, G., Giaccardi, E., Ye, Y., Sutcliffe, A. G., and Mehandjiev, N., "Meta-Design: A manifesto for End-User Development," *Communications of the ACM*, vol. 47, pp. 33-37, 2004.

[14] Goncalves, M. A., Fox, E. A., Watson, L. T., and Kipp, N. A., "Streams, structures, spaces, scenarios, societies (5S): A formal for digital libraries," *ACM Transactions on Information Systems*, vol. 22, pp. 270-312, 2004.

[15] Hammond, T., Timo, Hannay, Lund, B., and Scott, J., "Social bookmarking tools (I)," *D-Lib Magazine*, vol. 11, 2005.

[16] Hendry, D. G., Jenkins, J. R., and McCarthy, J. F., "Collaborative Bibliography," *Information Processing and Management*, in press.

[17] Herring, S. C., Scheidt, L. A., and Wright, E., "Bridging the gap: A genre analysis of weblogs," presented at Annual Hawaii International Conference on System Sciences, 2004.

[18] Kleinberg, J. and Lawrence, S., "The structure of the web," *Science*, vol. 294, pp. 1849-1850, 2001.

[19] Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A., "Trawling the web for emerging cyber-communities," presented at Eighth International World Wide Web Conference, 1999.

[20] Larios, J., "Drainspotting: About Tags," 2005, Retrieved June 1, 2005 from <http://www.drainspotting.com>

[21] Lee, H.-L., "Collection development as a social process," *The Journal of Academic Librarianship*, vol. 29, pp. 23-31, 2003.

[22] Lee, H.-L., "The concept of collection from the user's perspective," *Library Quarterly*, vol. 75, pp. 67-85, 2005.

[23] Lee, H.-L., "What is a collection?," *Journal of the American Society for Information Science*, vol. 51, pp. 1106-1113, 2000.

[24] Lenhart, A., J. Horrigan, and Fallows, D., "Content Creation Online," Pew Internet & American Life Project, 2004, Retrieved June 1, 2005 from http://www.pewinternet.org/pdfs/PIP_Content_Creation_Report.pdf

[25] Thunder Bay Public Library, "Bedtime Stories for Children," 2005, Retrieved June 1, 2005 from <http://www.tbpl.ca/internal.asp?id=282&cid=1021&doc=3>

[26] Narendra, N. C., "Flexible support and management of adaptive workflow processes," *Information Systems Frontiers*, vol. 6, pp. 2470262, 2004.

[27] Nutt, G. J., "The evolution towards flexible workflow systems," *Distributed Systems Engineering*, vol. 3, pp. 276-294, 1996.

[28] Shepherd, M. and Watters, C., "The evolution of cybergenres," presented at Annual Hawaii International Conference on System Sciences, 1998.

[29] Smith, M. A., "Portals: toward an application framework for interoperability," *Communication of the ACM*, vol. 47, pp. 93-97, 2004.

[30] Spinuzzi, C., *Tracing Genres through Organizations: A sociocultural Approach to Information Design*. Cambridge, MA: MIT Press, 2003.

[31] Stokes, R., *The Function of Bibliography*. London: André Deutsch, 1969.

[32] Tarsala, C. B., "What is an Author in the Sikuquanshu?," UCLA dissertation, 2001.

[33] Yates, J. and Orlikowski, W. J., "Genres of Organizational Communication: A Structural Approach to Studying Communication and Media," *The Academy of Management Review*, vol. 17, pp. 299-326, 1992.