Weblogs In Libraries: Opportunities and Challenges

Adam Mathes
Reference Sources and Services - LIS404A
Graduate School of Library and Information Science
University of Illinois Urbana-Champaign

April 2004

Abstract

Weblogs are an increasingly popular form of content on the World Wide Web. While they are not a new concept, having been around in one form or another arguably since the very beginning of the Web, they present a number of issues and opportunities for librarians. In examining how weblogs could be used by libraries, there are two fundamental issues. First, what are the important aspects of weblogs that librarians should evaluate and consider. Second, what are the important aspects of traditional library and information science lacking from weblogs that should be considered when using them in a library context.

1 What is a Weblog?

Although the term is becoming increasingly ambiguous, for purposes of this paper weblogs are defined to be web sites that present a reverse chronological ordering of posts. Posts are defined to be relatively short chunks of content, anywhere between a few words to a few paragraphs. This is an important distinction as most web sites traditionally were created and organized at the page or document level.

The first use of the term "weblog" was by Jorn Barger [1] and was a reflection of its purpose: to log the web sites that he thought were valuable. This paper will examine weblogs that are focused on links with short commentary. This is not necessarily in line with the common use of the term weblog today, which tends to encompass short-form journals that are organized in reverse chronological order, as well as any other number of web sites. Sites holding recent news of an institution or site are often called weblogs now as well. While these sites may also be of interest in a library context for informing patrons of news about the library in the way a newsletter could, I believe it is the weblog as web annotation and commentary system that is of more importance and worthy of consideration.

2 Weblog Positives

There are a number of important elements of weblogs that have made them popular and useful to a large number of people. First, focusing on the foundational elements, the reverse chronological ordering makes it extremely easy to find the most recent content. This encourages and rewards repeat visitors. It is also simply a good user experience and human computer interaction design decision to make some of the most important and often sought material (in this case, the most recent) prominent. The reverse chronological ordering does have a number of problems for longer term use, which will be discussed later.

The prominence of the "post" as the fundamental unit is also worth noting. By allowing updates at a much smaller level of content, it is easier for the maintainer of a weblog to add content that is not a large, monolithic block. Also, it has encouraged web links often to be followed by a short, but meaningful, description. This is far better than simply using bookmarks or lists of URLs as it allows annotation, criticism, commentary, and important context to be available with a link to the web resource.

Although not a feature of weblogs in particular, most of the software that people use currently to maintain weblogs fosters an ease of publishing unavailable to most users previously. Products like Blogger, MovableType, Diaryland, Pitas, and Radio Userland [2] generally reduce the work needed to update a weblog down to writing in a text box and pressing a "publish" button. This is drastically different from manually updating a weblog, which could take dozens of steps involving cutting, pasting, logging in to remote servers, file transfers, and other technological issues. By lowering the barriers to keeping a website frequently updated, these tools can make it more feasible to update with material on a constant and regular basis. This ease of publishing has also made tools to update weblogs popular as general purpose site update tools, and led to the general broadening of "weblog" to refer to frequently updated sites, no matter the content.

A more general aspect of weblogs is that they are showing themselves to be an effective way to manage information overload on the web. Weblogs can serve as filters. (In fact, one of the earliest sites that might be considered a weblog, or forerunner of weblogs was called "Filtered for purity." [3]) That is, while much of what one reads may not be worth commenting on or recommending, the small portion that is can be shared with a larger audience through a weblog. By leveraging the searching, reading, and interpretation of other web users, weblogs can help people to find better content. This is particularly important given that weblogs are often run by a single person with a distinctive voice - by finding weblogs one likes, one has possibly found a constantly replenished source of new material to examine they might not otherwise be exposed to. Domain experts running weblogs offer a powerful resource in understanding and exploring that domain.

3 Weblogs are Not Digital Libraries

There are some important aspects of libraries that weblogs clearly lack. Although weblogs are generally a controlled collection of links with commentary by the collector, they are not what we might think of as digital library collections.

One of the limitations with most weblogs and current weblog tools concerns the lack of metadata concerning the resource described. Most weblogs and weblog tools treat posts as a block of text, possibly with a date. Metadata, such as subject, author, publication date, institutional information, are generally not present in weblog posts, and the tools are not equipped to make adding metadata about the sites referenced in a coherent manner. It would likely be important in a library weblog to standardize on a consistent format for storing more information than simply a URL and descriptive text.

Related to this problem is categorization. Some weblogs categorize their posts, though most do not. This categorization scheme is usually an ad hoc system developed by the individual weblog author. Librarians may want to consider a better controlled vocabulary for their weblog post categories, or consider ways to integrate it into existing schemes such as Dewey and Library of Congress.

The categorization problems suggest another problem with weblogs: while the reverse chronological format is extremely valuable for quick visits, it is not necessarily ideal for browsing a collection of resources at a later date. Browsing by subject, author, institution, or other information is something a library weblog should consider. Web directories, such as Yahoo and Dmoz [4] offer some ideas for managing a large collection of links, but are only a start. Finding ways to integrate weblog content into more sustainable, browsable, and useful websites is an open question that librarians and information architects should examine more thoroughly.

There are substantial issues regarding the archiving and permanence of the resources discussed in a weblog. The Web is ephemeral. Since the resources are generally only linked to by a weblog, there is no control over the later accessibility of this resource. Web resources disappear for any number of reasons. There are efforts to archive the web, the most prominent of which is The Internet Archive [5]. Additionally, for resources that are contained within a page, such as a news or journal article, it is feasible to save or cache the resource for later use. There are numerous technical and legal issues involved in offering caches of pages to users that librarians should consider, but they are beyond the scope of this paper.

4 Recommendations

Weblogs are an excellent way to allow quick access to selected web resources. Libraries should consider the weblog format as part of an overall strategy of using online tools and offering them in a coherent way to their patrons. As part

of this strategy some aspects to consider are:

- *Format*: Consider a consistent format for posts, that may include more metadata than traditional weblogs. The Dublin Core is one place to start. [6]

- *Editorial voice*: While many weblogs employ a distinctive, subjective voice, the tone of a library weblog may perhaps need to be more neutral. Clearly marking objective summaries as distinct from subjective commentary on the resource is another possibility.

- *Bibliographies*: Consider using subject or domain specific weblogs as living, updated, annotated, online bibliographies for patrons. Those subjects may be those that the librarians have an interest or specialty in, or subjects that a large number of patrons have been inquiring about.

- *Offline Resources*: Consider how it might be possible to use the weblog format for traditional, offline resources to keep patrons updated.

- *Pruning*: Most weblogs do not prune or change their posts after their initial publishing. However, just as librarians prune their collections, and replace sources with newer or better materials, there is no inherent reason a weblog could not be treated with the same care.

- *Site Organization*: As discussed earlier, offering a reverse chronological ordering of material may be helpful and is an effective means of allowing frequent visitors to find updated content, but repurposing that content into a more accessible directory may be something to strive for.

Librarians have traditionally been architects of information spaces. With the explosion of web based resources, librarians should use their skills in organizing, preserving, and presenting information in the online world. Applying basic library and information science methodologies to the weblog format could help both libraries and the larger online world. Hopefully, librarians can help drive innovation in the tools and formats being used on the Web, and improve the global information infrastructure.

5   Notes

1. Jorn Barger's "Weblog resources FAQ" written in September 1999 [www.robotwisdom.com/weblogs/index discusses his use of the term in 1997. This is generally credited as the first use of the term in this context, as opposed to web server access logs. His 1997 weblog page is still available: [robotwisdom.com/log1997m12.html]

2. There are a number of software products and web services that can be used to maintain weblogs. Those mentioned here include: Blogger - www.blogger.com; MovableType - www.movabletype.org; Diaryland - www.diaryland.com; Pitas - www.pitas.com; Radio Userland - radio.userland.com.

3. Michael Sippey, writer and publisher of the web site stating the obvious [www.theobvious.com] began a regular feature on the site called "Filtered For Purity" in 1997. The first was published on December 26: [www.theobvious.com/archive/1997/12/26.html]

4. Yahoo [www.yahoo.com] although now a "portal" site still maintains its original hierarchial web site directory. Dmoz [www.dmoz.org] is a similar non-commercial project with volunteers maintaining the directory.

5. The Internet Archive [www.archive.org] has a number of interesting projects, including the "Wayback Machine" that allows you to surf cached versions of pages they have archived over the years.

6. Dublin Core [dublincore.org] - "The Dublin Core Metadata Initiative is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models."