September 2006

Is Access to Government Information Getting Better or Worse?

Katherine Holvoet

University of Utah, kate.holvoet@utah.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Part of the [Library and Information Science Commons](#)

Recommended Citation

Holvoet, Katherine (2006) "Is Access to Government Information Getting Better or Worse?," *Against the Grain*: Vol. 18: Iss. 4, Article 9.

DOI: <https://doi.org/10.7771/2380-176X.4916>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Is Access to Government Information Getting Better or Worse?

by Katherine Holvoet (Head, Government Documents, University of Utah) <kate.holvoet@utah.edu>

The simple answer is yes, access to government information is getting better and worse. Electronic indexing is improving access to material, and obscuring it. The **Government Printing Office** is attempting some heroic measures to improve access to government information, but doesn't have the budget to control the entire process. Commercial vendors are creating fabulous value-added government information products that are unattainable for many libraries. The upcoming wave of retirements among government documents librarians may create a significant loss of subject expertise, but will open opportunities for newer librarians. In short, there is cause for optimism and concern.

As government information moves online, it is sometimes easier to find material, such as tax forms, and sometimes harder. More agencies are creating databases for information such as technical reports, patents, and research publications. Search engines and Web crawlers will index the search page for these databases, but there are settings that must be present to allow a Web crawler to index the database itself, and some government databases are not set to be indexed by Web crawlers. This effectively places that information in a black hole. The average searcher would have to know exactly which database contains the information for which they are looking.

Currently, government documents librarians are the best finding aids for government information. Comprehensive computer indexing is one of the greatest advantages the digital format has to offer, and is not available

for U.S. documents. Mending this problem is technologically simple. Search engines could index the databases, but setting policies is politically difficult. The **Government Printing Office** does not have the authority to force executive branch offices such as the **Department of Homeland Security** or the **Department of Energy** to allow Web crawlers to index their databases.

For example, if a search is run in **Google** for "patent number 5,003,456," the result is one return, "Circuit for providing fast output current control in a current mode switching power supply." The resulting hit is from *freepatents online.com*, not from the **Patent and Trademark Office (PTO)**. The commercial site includes advertising, and may leave the researcher wondering if the result is in fact an official patent. There are over six million patents available online through the PTO's Web search, and none of them will show up in a basic **Google** search unless they are put online by someone other than the **Patent and Trademark Office**.

Perhaps the most talked about issue today revolves around preserving long-term access to born digital documents, and digitizing and preserving the historical government record. Who is responsible for archiving government information? Should that trust go to a government agency such as the **GPO**, or to library consortia such as the **Association for Research Libraries (ARL)**? What role should commercial publishers who offer terrific but expensive products play?

The **Government Printing Office** serves the pleasure of **Congress**, and executive branch departments serve the President, leaving open the possibility that inconvenient information could be purged from the public record. The significant change in access to online government information in the wake of 9/11 should alert librarians and citizens to the negative consequences of single institution provision of government information.¹ For example, the **EPA** has removed **Risk Management Plans** for facilities such as water treatment plants and dams, from their Website even though the necessary details a terrorist could use for planning purposes were never put online.² Now citizens no longer have ready access to pertinent safety information about the public works installations in their communities.

Cooperatives such as **ARL** may not have the finances or the organizational structure to effectively coordinate and implement such a massive preservation task. Attempting to copy and preserve born digital documents in the absence of a central database of government information would be maddening. Agencies no longer rely only on the **GPO** to publish their materials — now they can simply put them on the Web. Any born digital preservation project would need to continuously crawl all government Websites for new documents, and then save copies to servers. The task of identifying government information online, capturing a digital copy, storing it, and indexing it for easy retrieval may be beyond any organization's capabilities. The closest example to such a project is the **Internet Archive**,³ and the indexing for that is not very user friendly; it works best when one already knows the URL.

In regards to digitizing the historical record, consider the **US Congressional Serial Set** (the official record of **Congress**) which comprises over 12 million pages. Without centralized control of digitization and metadata cataloging, it would be very difficult to ensure every page met project standards. The best way to ensure centralized control would be for **ARL** libraries to all contribute money to pay to have the works digitized by a commercial vendor, and to have metadata cataloging created. Not all **ARL** libraries are willing to pay to create a digital **Serial Set**, and not all libraries would be able to pay even if they were interested. Another option would be to distribute the burden for digitizing the **Serial Set** among **ARL** members, along with standards and guidelines. Not all libraries have the expertise to digitize documents, or access to the necessary material.

What about commercial publishers? Could vendors fill the preservation need by digitizing government information for resale? The online **US Congressional Serial Set** by **Readex** is a thing of beauty, with tremendous added value in the extensive indexing. Quotes for this prod-

continued on page 36

Collection Development ... from page 30

ferent from our paper holdings. How curious that we so willingly gave away that control. Of course, we weren't trained to read contracts in the early days of digital resources, and we were doubtless blinded by the simple need to restrain the outrageously escalating costs. But now it is time to step back and ask ourselves whether we're truly willing to give

up the extraordinarily democratic infrastructure we've so recently built for access to our paper collections and to return to a time when a researcher's institutional affiliation defined his access to library resources.

Digital technology has made it possible for researchers to access our collections without regard to *where* they are — in the library, at home or in the office — we should not have to add the caveat that it now depends on *who* they are. 🍀

Endnotes

1. Many of the issues treated here grow out of a study I did recently of collection development patterns in research libraries during the last decade, particularly concerns about the divide between large and small, rich and poor. See my recent article: "A Decade of ARL Collection Development: A Look at the Data," *Collection Building*, 2006 no. 2, pp. 45-51.
2. Jerry D. Campbell, "Changing a Cultural Icon: The Academic Library as a Virtual Destination," *Educause Review*, January-February 2006, p. 20.

Missing Information ... from page 34

insight that when information is collected and distributed by a centralized system such as a cable TV network, a specific type of information is lost: that is, information with little commercial mass appeal no matter how strong local interest may be — a heated discussion at city hall may get a minute or so on the local news, but that's it. The implication of this observation helps define the role of individual libraries within a global library network. When a shared system of databases and interlibrary loans provides easy, reliable access to core items, individual libraries no longer have to spend time and money on providing better and better access to "blockbusters." Librarians can spend their efforts archiving and cataloging what *Wired* Editor-in-Chief **Chris Anderson** terms "the long tail" — that is the virtual infinity of niche items with a limited but still nonzero audience.¹

"Long tail" businesses such as **e-Bay**, **Amazon.com** and **Netflix**, succeed by selling a huge number of low-demand, low sales volume items. Academic libraries are long tail enterprises. The collective activity of many libraries has resulted in an extremely large and diverse knowledge base that offers information with mass appeal as well as information on narrowly focused and highly specialized niche interests. Relatively inexpensive access to enormously large and highly diverse (if infrequently circulating) collections is a major advantage libraries have over other information sources.

The future of libraries may turn out to depend just as much on obsessive bibliographers indulging their own collection fetishes as on grand projects to aggregate back runs of the most vital scholarly literature. The question is, from a virtually infinite selection, how can librarians decide what to add to the collection? The geopolitical and cultural place where a library is located is the most obvious place to start. Local librarians have local knowledge and besides, nobody outside the area may know or care about regional, issues. Individual libraries may also declare responsibility for highly specialized subject collections. These unique materials help create an identity for the library, but they also add value to all libraries that can tap into the resources.

McKibben's essay was published just one year before **Mosaic** was introduced as the first popular graphical Internet browser, and within a few years the World Wide Web changed everything. The Web is the best tool ever invented to distribute niche information, and as a result many types of information that used to be considered second tier — labeled as "ephemera" or "grey literature" — are now readily available full-text and open access. It is practically a librarian mantra that open-access information gets cited more often.

Case Study: Utah's Legacy Parkway

I'll use a case study of the Legacy Parkway freeway construction in Utah, which neatly shows the interaction of government planning, journalism and citizen advocacy in generating

not only a very large mass of documentation (see chart, sidebar, whatever) but also real-life results. This is a highly focused regional issue of the type that generates valuable long tail collections. The publication trail identifies where relevant information resides and whether or not it is likely to be safe for the future. Briefly, the final Environmental Impact Statement which the **U.S. Dept. of Transportation** prepared for a highway project in Utah was contested in court by Salt Lake City mayor **Rocky Anderson** and two citizen advocacy groups: **Utahns for Better Transportation (UBeT)** and the **Sierra Club Utah Chapter**. The plaintiffs won their case and hired their own consultant to write an alternative "Smart Growth" plan for the highway. The Utah Legislature then called a special session to discuss the highway, and all parties agreed to a compromise.

It is obvious that from a library standpoint only part of the information stream is well represented, and this is not even a comprehensive list. The Legacy Parkway was a regional controversy for nearly ten years and there were associated Webpages, blogs, emails, discussions, speeches and so on.

Even in the new information environment

Pictured above is **Amy Brunvand**.

there are core information streams and there are ephemera. There is material which will naturally be saved in library collections and there is material which will be lost. It's just that the core has broadened.

If each library can truly identify and represent its own information niche digitally then the community of libraries will build a long tail collection with which no other information source can compete. 🐘

Endnotes

1. "Long tail" also refers to a type of statistical distribution with a few high-frequency values and many low-frequency values. A graph of the data looks like a steep ski slope with a long, flat run-out at the end.

Is Access to Government ... from page 32

uct vary from institution to institution, but the price easily exceeds \$100,000. The *Advisory Board Annual Report* for the Year 2005 posted on the vendor site lists current customers, the vast majority of which are academic libraries. **Boston Public Library** and several state libraries are the only public libraries that have purchased this set. While the products produced by commercial vendors are of excellent quality, they are unattainably priced for most public libraries.

On the upside, as more material goes online, new collection development opportunities are becoming apparent. Historically, it's been a full-time task for government documents departments to process incoming print collections and staff a reference desk. Many departments have not been able to collect extensively about topics and events with high local impact. With the tangible collection diminishing, government information specialists can begin focusing on creating collections in areas that uniquely serve local populations. Through online collections, specialists can gather federal documents, relating state, city and local documents, as well as reactions and policy information from local think tanks, political activist groups, etc. The **Greater Western Library Alliance** has created the **Western Waters Digital Library (WWDL)**, a collaborative project between twelve university libraries in eight western

states. The **WWDL** contains government documents, literature, legal transcripts and more about the Columbia, Colorado, Platte, and Rio Grande river basins. Efforts like this ensure that patrons have access to highly desirable materials without the tremendous effort it would have taken an individual to identify and locate the items.

The **National Institute of Health (NIH)** open access initiative requests and encourages researchers to make peer-reviewed articles arising from taxpayer-funded research available full-text, online, free through **PubMed Central**. While not making free access mandatory, the **NIH** initiative opened up a national dialog about the value of government-funded research, and the current high cost for public access to results. Mega-portal sites such as **FirstGov.gov** and **Google U.S. Government Search** are taking steps in the right direction to move away from using individual agency/department/bureau pages as the primary finding tool for government information.

Exciting collection development projects will come to a screeching halt if the library profession loses too many government documents experts in the near future due to retirements, and doesn't replace them. The attrition of expertise could seriously undermine access to government information. The only way to forestall a permanent loss of expertise is for libraries to choose to have government information specialists on staff, and support training for newly assigned government documents librarians.

continued on page 40

Are New Technology ... from page 38

sold as a traditional library product. Libraries can either subscribe to it or purchase perpetual rights and then provide access to their patrons. Non-subscribers may not access this collection.

The "Second Wave" and Beyond is free. The site is led by three editors, **Judith Ezekiel**, **Stephanie Gilmore**, and **Kimberly Springer**, who are themselves active scholars in the field. By monitoring and encouraging collaboration, the editors instill a spirit of peer review to the site, giving scholars a meaningful online environment in which to flush out their research and writing. Anyone can read and browse the materials posted there, but to participate in the discussions and contribute material, you first need approval from the editors. This form of open access with peer-reviewed participation allows us to keep the site fresh, but at the same time it ensures that the quality of posting remains high.

How's it Working Out?

While it's still in its early days, the initial signs are encouraging. In the past couple of months, we've had some 150 academics, graduate students, and activists sign up and get accepted as participants on the site.

A number of well-known academics have gotten on board. One lively discussion centered on an article published by **Carol Hanisch** in 1970 entitled, "The Personal is Political." This new form of publication offered **Ms. Hanisch** an opportunity both to publish a new introduction to the piece and to candidly discuss the original article and the reaction to it with other scholars.

Several things are apparent from the quotation below:

"This is Carol Hanisch responding. I'm delighted to be participating in this pioneering use of the Web for discussing feminist theory. Greeting to all, and especially to Chude Allen, who I knew as Pam in New York Radical Women (hereafter referred to as NYRW) nearly 40 (YIKES!) years ago.

*Actually, NYRW was my first major experience with feminist organizing and thought. I'd read Betty Friedan's *Feminine Mystique* the summer of 1964 while working as a UPI reporter fresh out of college in the Des Moines Bureau (which was also where I encountered major work discrimination). There had been rumblings of feminism in college (Drake University in Des Moines), especially over dorm hours for women while the guys had none. Although I found it interesting, Friedan's book did not strike a major chord with me. As an "Iowa farm girl," I had no contact with suburban American housewifery except on TV and the movies, and I had no desire to become any kind of housewife, though I assumed I would marry and have children. Obviously I hadn't thought that one through, but this was a gut reaction, not an ideological one!"*

You can see immediately that this type of discourse is very different from traditional forums. It's much less formal, more conversational, and above all much more personal. This last has been a pervasive characteristic of the site so far. Already there are a number of articles that talk about historical events in personal terms. For example, **Stephanie Gilmore** talks about her experiences growing up in Alabama in 1985 (<http://scholar.alexanderstreet.com/x/OAc>) and how her high school teacher reacted to her notions of feminism.

It's informal perhaps, but the editorial control means that the site is largely free from the trivial, the factually inaccurate, and the irrelevant. It's more formal and stable than email, and more interlinked and supported than discussion group conversation.

Users can follow each comment or page independently, via RSS subscription. Discussions or pages can be printed or copied using **MS Word** or in **PDF**. Users can view and compare how the page has been edited and by whom. And finally, scholarly users who have been granted access by the editors can comment on or change what's been posted.

Pictured above is **Stephen Rhind-Tutt**.

Academics have also been using the site to post scholarship and images that otherwise would still be languishing in their file cabinets. Here, for example, is a posting by **Judith Ezekiel** of a Gay Pride demonstration that she attended in New York in 1977.

Ms. Ezekiel is also developing a comprehensive chronology of the women's movement on the site. Other users are now referencing and editing the chronology, adding new events to the sequence, linking to primary materials describing these events, and elaborating on the experience and significance of them. Through this process, the scholars are collectively creating an authoritative sequence of events.

The site allows postings of a wide range of varied and often hard-to-find material including syllabi, lists of links, and reviews. Each of these objects now lives in a formal way on the Web, where it can be referenced and accessed easily later.

Going Forward

As the community begins to use the site more, we will learn much and also expect the site to grow in value and utility. Already we're seeing requests for navigational improvements and for training on how better to use the technology.

We're also planning a second site called **Teaching Women** that will be aimed at the teaching community. Further down the line, we plan to introduce additional scholarly communities to complement our other existing collections. We aim to offer interactive communities for disciplines including American history, music, religion, theatre, and film. Researchers will not only benefit from the material we publish, but they will have an opportunity to add to it and read what colleagues worldwide have to contribute.

The technologies that are so well used by **Wikipedia**, **Flickr**, and others hold great promise for scholars. The "Second Wave" and **Beyond** is a small example of how they can be applied. Our experience so far suggests that we're seeing the emergence not just of a new technology, but also of a new model of scholarship and of literature that will be more accessible, easier to understand, conducted in real time, and conducted with participation from many scholars at the same time.

Is Access to Government ... from page 36

What benefits does the future hold for government information? Ideally, every Internet connection will be a depository library, and every librarian a government information expert. While the distribution mechanism is imperfect, the information is out there, and freely available. To create a brighter future for government information, library organizations must make developing and promoting government

information collections a priority (especially local interest collections), and support permanent, free access to government information whenever possible through local library policies, financial support for access projects, and legislative support for open access initiatives.

Government information is funded by the people, and created for the people, and should be forever accessible. Current and future projects and programs must facilitate access through policy and technology as well as keeping costs within the reach of libraries and citizens.

Endnotes

1. <http://www.ombwatch.org/article/articleview/213/1/1>.
2. "Let the People Know the Facts," pg. 9, *Law Library Journal*, 2006-01.
3. <http://www.archive.org/index.php>.