

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

June 2007

Google 'til They Goggle: Trawling Electronic Databases to Build Your Collection and Better Serve Your Client Base

Carol Ottolenghi

Office of the Ohio Attorney General, cottolenghi@ag.state.oh.us

Follow this and additional works at: <http://digitalcommons.unl.edu/libphilprac>



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

Ottolenghi, Carol, "Google 'til They Goggle: Trawling Electronic Databases to Build Your Collection and Better Serve Your Client Base" (2007). *Library Philosophy and Practice (e-journal)*. 126.

<http://digitalcommons.unl.edu/libphilprac/126>

Google 'til They Goggle: Trawling Electronic Databases to Build Your Collection and Better Serve Your Client Base

Carol Ottolenghi

Research Librarian

Office of the Ohio Attorney General

Columbus, OH 43215

Introduction

Like librarians throughout the world, we librarians at the Office of the Ohio Attorney General (AG) are all too familiar with the following statements from information seekers:

"I'll just Google it;"

"It's all online;"

"Nobody uses books anymore."

It is no secret that this attitude has led to much angst within our profession, including the existential question: Are librarians necessary? So it is mildly ironic that the AG librarians use Google, electronic databases, and a variety of online "alert" systems to highlight and increase our value to AG staff.

One of the simplest ways of doing this is also the most visible: the development of a PDF Library that supports the information needs of AG staff throughout Ohio. To collect relevant articles, studies, patents, maps, and consumer information for the PDF Library, we scour the Internet and public and purchased databases. Also, the "alert" options available on Google and individual databases are invaluable in keeping our information hunts timely and thorough. We then save appropriate materials in Permanent Digital Format (PDF).

Our PDF Library is librarian-reliant, as opposed to the Digital Library as defined by Reach et al., which requires little, if any, interaction between librarian and user (369, 371-372). In our PDF Library, the AG librarians conduct reference interviews on an ongoing basis to determine current AG information needs and then locate and disseminate needed materials electronically—solid and traditional librarian duties served up in a different format. The PDF Library showcases our ability to sift quickly through the superfluity of *electronic and print* materials to find needed information. It allows us to assist off-site staff members who previously did not know that the AG had a library; to educate staff about the wealth of high-quality legal, scientific, and social information available; and to add significantly to our collection during a time of major budget cuts.

Under some circumstances, it might be advantageous if the PDF Library were not so librarian-mediated, if, for example, AG staff could peruse the files online, without requiring a librarian to “pull” the file. However, three considerations prevent this from occurring right now. First, many materials we add to the PDF Library are new or unknown to AG staff, so staff would not know to look for them in the PDF Library. In contrast, Reach et al.'s Digital Library was built around well-known legal titles that most lawyers expect to find and use. Second, the time and ability that non-library staff can dedicate to browsing the PDF Library is extremely limited. Third, current AG policies and firewalls make opening the PDF Library to all AG staff a “not-gonna-happen” event.

Background

AG staff requires access to a wide variety of legal and nonlegal information. (Legal information is defined by this author as materials such as statutes, regulations, legislative histories, legal treatises, statute interpretations, and case law. Non-legal materials originate in other branches of knowledge, such as medicine, the arts, and the sciences.) In some respects, the information-seeking patterns at the AG resemble those found in large law firms by Leckie et al. But there is one important difference based on a finding of Leckie et al. that “the organizational context in which the lawyer practices” and the particular “role” a lawyer fills influence information-seeking behavior (173-174). Most of the firms studied by Leckie et al. specialized in only a few areas of law—such as corporate, domestic, tax, or criminal—and in filling certain roles. As a result, their information-seeking patterns remained somewhat static. In contrast, the AG does not have the luxury of limiting its practice or roles, and information-seeking behavior at the AG can morph overnight.

Ohio statute mandates that “the Ohio Attorney General is the chief law officer for the state and all its departments ... The Attorney General shall appear for the state in the trial and argument of all civil and criminal causes in the Supreme Court in which the state is directly or indirectly interested” (Ohio Revised Code §109.02). Restated in terms of information-seeking, this means that if an issue could affect how Ohioans are born, die, and spend their time in between, then AG staff will eventually need legal and nonlegal information on it.

To cope with this mandate, the AG is divided into 21 Sections that employ over 400 lawyers, 106 paralegals and 200 investigators spread across Ohio. (The AG's legal Sections include: Appeals; Bureau of Criminal Identification and Investigation; Capital Crimes; Charitable Organizations; Child & Elder Protection; Civil Rights; Consumer Protection; Corrections Litigation; Court of Claims; Crime Victims Services; Education; Employment Law; Environmental Enforcement; Health Care Fraud; Health and Human Services ; Ohio Law Enforcement Gateway; Ohio Organized Crime Investigations Commission; Opinions; Peace Officer Training Academy; Public Utilities; Taxation; Transportation; and Workers' Compensation. AG staff is served by two librarians and two MLIS-seeking interns.) Different types of staff request different types of research assistance, not only in topic, but in discipline. Whisner (2003; 2005) reports that most attorneys do most of their own legal research. (See also the American Bar Association's Legal Technology Resource Center 's 2002 technology survey, cited by Reach et al. [373]). This has been our experience as well. However, very few of the more than 200 AG investigators are attorneys, and they require legal research assistance. In addition, both attorneys and non-attorneys request library help in finding nonlegal information. Nonlegal resources can play a surprisingly large role in the practice of law; Hasko's 2002 study of the United States Supreme Court indicates that nonlegal research was involved in over 40 percent of cases before the Court.

In the past, the library's response to these requests was reactive: a question was asked, library staff found the information, end of interaction. Then staff changes, budget cuts, and a serendipitous screw-up combined in February 2006 to inspire the PDF Library. It began as a low cost and low storage way to funnel information to AG staff across Ohio. Soon, though, we realized that we could adopt a more proactive stance and really market the library by expanding the PDF Library via Google and similar search engines.

Our statistics suggest the results of our shift in stance: research questions have increased more than 20 percent over 2005, despite staff and budget cuts. However, the quantity of questions is not the full story. Every week, we provide information to Sections that had not contacted the library previously. Ever. In addition, the depth and range of research requested has changed dramatically. The vast majority of previous years' research questions revolved around bibliographic instruction and simple retrieval of individual documents requested by name. We still provide these services, but the PDF Library has caused a shift in expectations. Many AG staff members now request in-depth, exploratory, and ongoing research by topic area. Staff in Capital Crimes, Environmental Enforcement, Bureau of Criminal Identification and Investigation (BCI), Education, Health and Human Services (HHS), and Taxation not only request in-depth research, they often ask us to analyze the materials we find and schedule time with us to discuss our findings. It should be noted that the two fulltime Librarians have additional graduate degrees and specialties that facilitate the interpretation of certain materials.

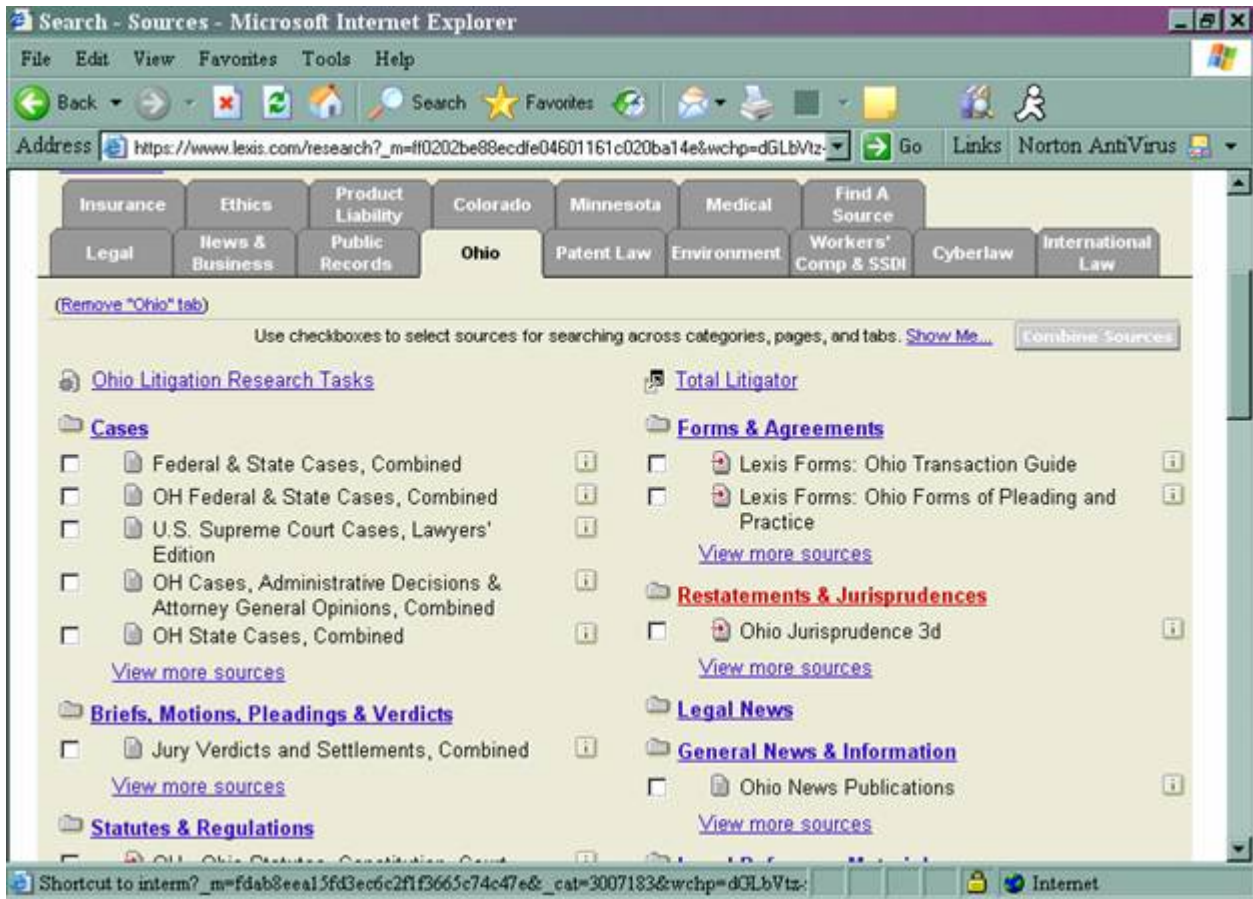
Creating and Using the PDF Library

We opened our PDF Library in three easy steps: 1) We created a file folder to which all library staff could contribute; 2) We called it the PDF Library; 3) We instructed library staff to save electronic materials by exact article title, without quotation marks, into this folder.

Building the user base for the PDF Library was a little more involved. It required a daring combination of traditional librarian research skills and a willingness to trap unsuspecting AG staff members in elevators, at lunch, and in parking garages, grilling them until they cracked and divulged the types of information that they could use. (The author confesses that she is not completely kidding about this.) After the February mix-up alerted us that offsite AG staff members were eager for information, we began using Google and other search engines to scour online databases for studies, articles, patents, and other materials that reflected previous information requests. We then conducted the usual librarian selection tests of checking authority, quality, last update, etc.

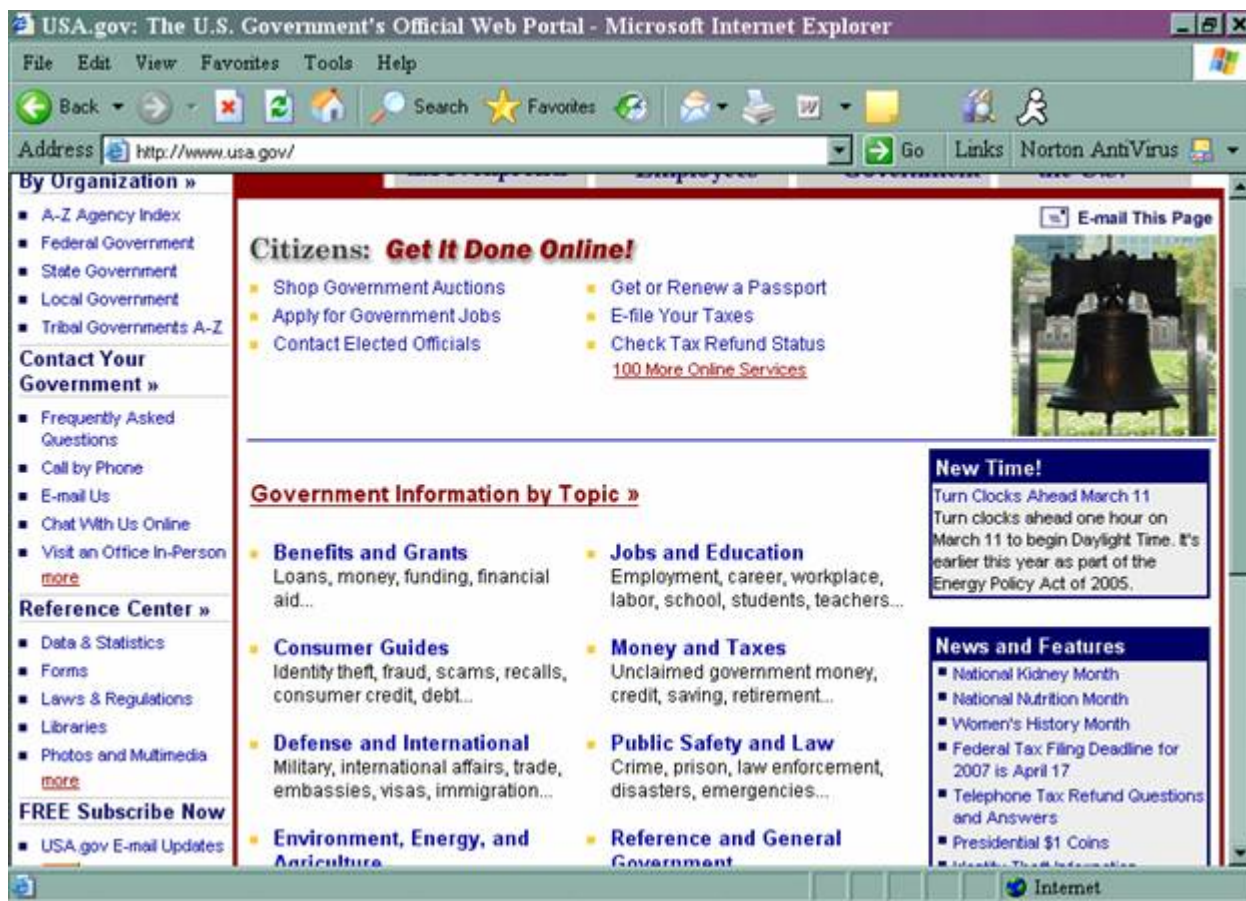
Currently, most of the materials that we harvest are from government databases or from AG-purchased databases from which we may disseminate materials. The databases we search differ for each question, but three stand out as being useful on a daily, sometimes hourly, basis: Lexis, Science Citation Index Expanded, and USA.gov. Lexis and Science Citation Index Expanded are proprietary databases that may be available through state consortiums ; USA.gov is a federal website, open to all.

Lexis is an extensive legal database of laws, regulations, caselaw, treatises, reference materials, law review articles, etc. We use Lexis "alerts" (discussed later in this paper) to keep abreast of additions as they are made.



Science Citation Index Expanded provides citations to published works in science. This resource is particularly useful for our Criminal Justice, Environmental, Workers' Compensation, and Health and Human Services Sections.

USA.gov is a free gateway to all things federal, from agency websites to consumer guides, ongoing research, and manipulable census data. The site provides RSS feeds (discussed below) to alert users when new materials are added to selected topics.



We respected and continue to respect copyright during the PDF collection process. Since the AG is a major law enforcement agency within Ohio, we feel an added imperative not to violate copyright laws; not only would a copyright violation be an ethical breach, it would also be very embarrassing to be featured in the press as "The Librarian Lawbreakers of the AG's Office." Dissemination rights to the few materials not found in purchased or governmental databases are obtained in writing from their copyright holders.

After we gathered the first batch of materials and saved them as PDFs, we offered it to AG staff who had previously indicated interest in the topics. We also listed selected PDFs in the AG's in-house newsletter. When AG staff contacted us to get the materials, we emailed the PDFs with a variation on the following blurb:

"BTW - We conduct ongoing research on a number of topics for members of AG staff. If there is anything specific that you'd like us to research for you, please email us at Library Requests."

AG staff responded to this blurb with requests for research on topics that ranged from eminent domain to fingerprinting, gangs, drugs, terrorism, photography, white-supremacist groups, search and seizure, excessive use of force, jury note-taking, cybercrime, crimes against consumers, and health care fraud. We fill ongoing

requests on, roughly, a biweekly basis. Most of the time we email relevant PDFs, but when the best materials are only available in print format, we forward those. This underscores the fact that *not everything is available online*.

We also created ongoing searches for information that was not requested but that seemed a logical fit with AG Section mandates and current events. We continue to add documents located through these searches to the PDF Library and advertise their availability to staff.

As PDF files accumulated in the PDF Library, we needed an easy way to locate files that answer individual AG staff members' information needs. Adding this information to the library catalog would have been perfect, but that was not immediately possible. So, we created a simple finding tool using Microsoft Word's "table" function, which can be searched by keyword. Once we identified the needed materials with the PDF Library Finding Tool, it is the work of a minute to open the PDF Library and email the files to AG staff.

Each entry in our PDF Library Finding Tool has eight columns. Filling these columns is the most onerous part of maintaining the PDF Library, but, like any cataloging or inventory procedure, it is vital. Our column headings have evolved over the past eight months to be: **Citation**; **Keywords**; **EF**; **Intranet**; **Sent to**; **Section**; **Date sent**; and **Asked for**.

Citation: The PDF Library file lists materials alphabetically, so in the finding tool we first list the exact title of the article as it was saved into the PDF Library file. Then we list the first three authors, followed by the journal or sponsor's name (i.e., DOJ for a Department of Justice-sponsored paper) and the volume and issue number. Finally, we add the publication year and, if the materials are also available at a website, give the website address.

Keywords: These are the core of the finding tool; they allow us to find all PDF Library materials on, for example, charter schools or weather patterns. Some articles have keywords listed below the abstract. We cheer when we find those. We also read the materials while considering what words we might use to search for this type of information in the future. We plan to create a PDF Library thesaurus, but time constraints are... time constraints.

EF: EF identifies the Electronic Form of the material. This column gets either a PDF, a W for website, or both.

Intranet: This is the date that the material was advertised in the AG inhouse intranet newsletter, hypothetically letting all staff know that the material—and the library—exists.

Sent to: Here we list AG staff members who have received the material. This prevents duplication and gives us another tool for identifying staff members who

might be interested in something. For example, if persons A and B are involved in similar activities, then if A finds something useful, we will send it to B also.

AG Section: Knowing which Sections are interested in what specific topics helps to guide us when hunting for relevant articles, studies, cases, and patents.

Date sent: Used for internal record-keeping.

Asked for: We fill this column with a “Y” when staff asks for material by name; an “N” when we send the material as a “cold call;” a “Y, by topic” when staff asks us to look for materials on a specific topic; a “Y, intranet” when staff requests an article or study that we advertised on the intranet; and a “Y, g-email” when staff responds to a Section-wide email about the material. This information lets us know how we are reaching the rest of AG staff.

Expanding the PDF Library toward Future Information Needs

A variety of publicity techniques, combined with the “Hey, where did you get *that?* I want one, too” effect, continues to grow the list of staffers requesting ongoing research. Anecdotal observation indicates that this effect is particularly evident within the Sections most involved in police work. The author posits that this may be related to information-seeking behaviors in these Sections that are similar to the information-seeking behaviors Leckie et al. noted in engineers, specifically the heavy reliance on one’s immediate colleagues for information.

The following “age of Google” tools are invaluable in filling these requests.

- Real Simple Syndication (RSS) allows online content providers to send part or all of their material automatically to anyone who requests it. More important to us, it allows us to receive emailed Tables of Contents (TOC) and other updates from selected journals as soon as they are posted online. Many journals, blogs, and other websites provide RSS feeds. Once registered with the sites we want to monitor, we don’t have to check the sites weekly; update information arrives in our email.
- Google Alerts, and similar options on organizational, governmental and educational websites “alert” us by email whenever relevant new information is posted on topics we select. To create a Google Alert, go to <http://www.google.com/alerts>.
- Similarly, specialized databases like Lexis have alert options that notify us when court cases and law journals address topics of interest.

Depending on the topic being searched, it can take several hours to craft a sufficiently inclusive and/or exclusive search query for the alert. However, these alerts save us hours each week in search time and help ensure that our searches are complete. Equally important—perhaps even more so—these alerts give us hints about the future information needs of AG staff. For example, Google Alerts may inform us of

news concerning religious issues in Alabama corrections facilities; a Lexis alert may signal changing DNA acceptance in a California court; and a Science Citation Index Expanded alert may provide clues to what AG forensic scientists and litigators need to know.

The AG mandate is so broad that every day relevant topics are in the news or the courts somewhere in the United States. The immediacy of the alerts allows us to make AG staff aware of issues as they arise. The alerts keep us abreast of the changing legal landscape and how rulings in another state may affect our attorneys in Ohio. The alerts provide us with added information to help AG staff frame their questions and give us a heads-up in time to find relevant resources and organize them in the PDF Library. The age-of-Google tools help us do the job librarians are meant to do: figure out what information is needed, find it, organize it, and send it to the people who will use it.

Works Cited

Hasko, John J. "Persuasion in the Court: Nonlegal Materials in U.S. Supreme Court Opinions." *Law Library Journal* 94 (2002): 427-57.

Leckie, Gloria J., Karen E. Pettigrew, and Christian Sylvain. "Modeling the Information-seeking of Professionals: A General Model Derived from Research on Engineers, Healthcare Professionals, and Lawyers." *Library Quarterly* 66.2 (1996): 161-93.

Office of the Ohio Attorney General. *Protecting Ohioans: A guide to the Office of Attorney General*. 2005.
<http://www.ag.state.oh.us/online_publications/OverviewLayout72.pdf>.

Ohio, State of. *Ohio Revised Code*. Current as of September 20, 2006.

Reach, Catherine Sanders, David Whelan, and Molly Flood. "Feasibility and Viability of the Digital Library in a Private Law Firm." *Law Library Journal* 95 (2003): 369-81.

Whisner, Mary. "Researching Outside the Box." *Law Library Journal* 95 (2003): 467-73.

Whisner, Mary. "On Not Doing Research." *Law Library Journal* 97 (2005): 391-97.