

"It's Not Rocket Science": Making Sense of Scientific Evidence

by Paul Barron

The Internet has quite simply become one of the primary tools in a research strategy that aims to pull data from all relevant sources. The Internet simply can't be ignored.

Carole A. Levitt & Mark E. Rosch

The Lawyer's Guide to Fact Finding on the Internet

American Bar Association

In a 2001 *Duke Law Journal* article, Carl F. Cranor wrote, "[L]ittle is known about the universe of approximately 100,000 chemical substances or their derivatives registered for commerce. Surprisingly, for seventy-five percent of the 3,000 top-volume chemicals in commerce, the most basic toxicity results cannot be found in the public record."¹ This might dissuade researchers from searching the Web for information on the reliability of scientific evidence about toxic substances.

Legal professionals have integrated the Web into their research, but they also know that the number of Web documents makes it more difficult to find relevant information. Sources cited by the Congressional Research Service estimate the "Surface Web"—that portion of the Web reached by search engines—contains more than thirteen billion documents. Each day seven million documents are added on the more than nine million Web sites. The "Deep Web"—that portion of the Web reachable only by querying a free database such as PubMed from the National Library of Medicine or fee-based services such as LexisNexis or Westlaw—is much larger. Estimates of its size vary from 150 to 500 times larger than the Surface Web.²

If a law library housed the number of print resources equivalent to the number of resources of the Surface Web, researchers would use more than key words to locate

information. By using the advanced search features of Web search tools, relevant material can be found quickly. For instance, to locate the Cranor piece, a Google site-limited title search returns one result—the link to the full text article. This article will review a search process using advanced search query features in Google, (see figure 1) Yahoo and other search tools to find publicly accessible

Scirus science-specific search engine, OAIster academically oriented digital resources and MEDLINE.⁴ The information gleaned from the two Web sections will supplement more in-depth research conducted in print resources and proprietary databases such as LexisNexis.

The search query syntax is standardized: Search terms of two or more words are

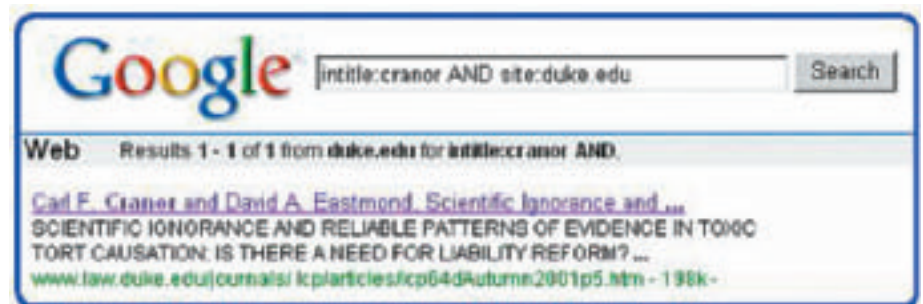


Figure 1

Web-based information on toxic substances and the law and the reliability of scientific evidence about toxic substances. Search tools that perform better with specific topics are searched using queries related to "sick building syndrome." Although Google and Yahoo are used in Surface Web searches, no search engine indexes more than 20 percent of the Web. Additional search engines may be used to search thoroughly the Surface Web's content.³ The Deep Web is searched using the

enclosed in quotes so the search engines search for phrases, not individual words; Boolean operators (OR, AND and NOT) are upper case text; and, multiple Boolean expressions are nested in parentheses to keep like concepts or synonyms together. This method works in the Surface and the Deep Web search tools discussed in this article. Using this standard, the mixed syntax search query format is:

"toxic substances" AND "scientific evidence" AND (reliability OR verification).

Although Google's and Yahoo's advanced search query features are used, queries are run in the basic search templates to avoid the confusion of where to place search query segments in the advanced search template.

For broad topic research we begin with the Scirus science-only search engine recommended by Levitt and Rosch: The search tool locates scientific, scholarly, technical and medical data. Unlike the general search engines, Scirus's advanced search option allows the researcher to find search terms in the article title and to limit the results by content area, date, information type and subject. (see figures 2 and 3)

The date-limited search (2000–2006) Scirus search: **"toxic substances" AND "scientific evidence" AND (reliability OR verification)**

returned relevant resources such as the current edition of the Federal Justice Center's *Reference Manual on Scientific Evidence*, including the full text of the *Reference Guide on Toxicology* coauthored by Mary Sue Henifin, J.D., M.P.H., a partner in a New Jersey law firm and an adjunct professor of public health law at the University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School. The "guide focuses on scientific issues that arise most frequently in toxic tort cases . . . and provides an overview of the basic principles and methodologies of toxicology and offers a scientific context for proffered expert opinion based on toxicological data."

Locating recent research is accomplished by the OAIster Deep Web search tool, a project of the University of Michigan

Digital Library Production Service. OAIster's mission is to provide links to free difficult-to-access, academically-oriented digital resources. A search for **"toxic substances" AND law** returns eleven results. One article, titled "Regulating Toxic Substances Through a Glass Darkly: Using Science Without Distorting the Law," is by Cranor and concludes, "Legal regulation of toxic substances by the tort (or personal injury) or regulatory law can be addressed by sensitively designing scientific and legal burdens of proof for the legal and public health problem in question." (see figure 4)

Another specialized database is the National Library of Medicine's PubMed that contains over fifteen million citations from MEDLINE and other life science journals. New material is added Tuesdays and Saturdays. The database's content is indexed using the controlled vocabulary Medical Subject Headings (MeSH). Since the MeSH terms are precise, the controlled vocabulary should be reviewed prior to searching the database. For example, the MeSH for "toxic substances" is "hazardous substances" and the descriptor for the toxic "black mold" *Stachybotrys atra* is satratoxin H, described as "a toxic metabolite of *Stachybotrys atra*." Using the controlled vocabulary in the search: **"satratoxin H" AND buildings** returns six results from journals and articles studying the adverse health effects to occupants after exposure to satratoxin H in water-damaged buildings. The first result links out to a full text copy of the article. (see figure 5)

Researchers can also search for non-MeSH terms in PubMed; a title search for **"sick building syndrome"** returned 173 results from journals such as the *Archives of Environmental Health* ("Studies on the Role of Fungi in Sick Building Syndrome")

TOXNET is a useful National Library of Medicine database. It is a cluster of databases on toxicology, hazardous chemicals, environmental health and toxic releases.⁵ Running the "sick building syndrome" search returned 500 results from U.S., Belgian, British, Dutch, German and Scandinavian journals.

Figure 2: Scirus Science-specific Search Engine Advanced Search Template

Figure 3: Scirus Results

After exhausting specialized search tools, the next step is to search the Surface Web. An initial Google search for **“scientific evidence” AND “toxic substances”** returns more than 340,000 results from .com, .org, and .gov sites and sites from Canada and the Cocos Islands. While the number of results may reflect the popularity of an issue, no one can wade through 340,000 hits.

However, the thirteenth result is surprisingly relevant; the Web site belongs to a board-certified civil trial lawyer with a bachelor’s degree in chemical engineering. The site’s subject categories include Research Sites for Chemical and Toxic Properties, Scientific Evidence Resources for *Daubert/Frye* Issues, Resources for Specified or Classes of Toxins, and

Litigation Support Resources.⁶ A rule of thumb: When the results are high, review the first fifteen hits.

One useful technique to refine a search returning too many results is to find only Web pages with a specific title since a Web page entitled “Scientific Evidence” probably focuses on that subject. The title search syntax in Google is: **intitle:** and the revised query is:

intitle:“scientific evidence” AND (reliability OR verification) AND “toxic substances”. (see figure 6)

All forty-seven of the results have the phrase “scientific evidence” in the title. The second result, an online version of the peer-reviewed *American Journal of Public Health*, has an abstract of the full text article, “The Weight of Scientific Evidence in Policy and Law,” which can be purchased for ten dollars. The fifth result from Defending Science.org provides a full text copy of the same article for free.⁷

Another effective search technique to reduce the number of results is to limit the results to sites with specific top-level domains such as an .edu, .gov, or .org. Educational sites may provide articles by faculty. Federal and state government sites provide full text of laws, and organizational sites express viewpoints about an issue or law. In the .edu domain-limited search **“scientific evidence” AND “toxic substances” AND (reliability OR verification) AND site: edu** the second result is the 2001 *Duke Law Journal* article by Cranor. The fourth result in the .edu-limited search connects to the Harvard University site Sound Science in the Courtroom. The home page mentions the Atlantic Legal Foundation mission to “ensure that whenever science is used in a courtroom that it shall be sound science.” (see figure 7)

Another search option limits the results to a specific site by running a site-limited search. To search only the Atlantic Legal Foundation site for information about toxic substances and scientific evidence, the search query is: **“toxic substances” AND “scientific evidence” AND site:atlanticlegal.org**.

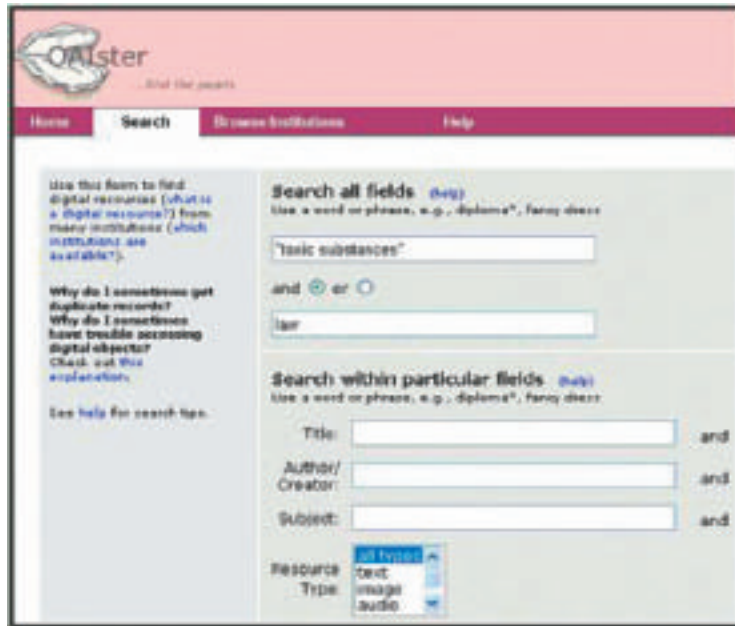


Figure 4: OAIster Digital Library

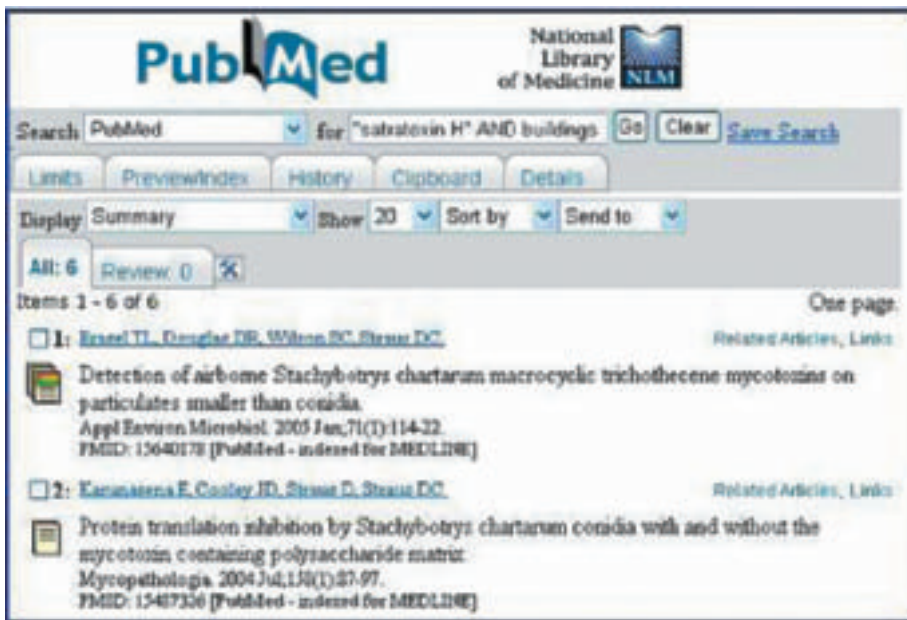


Figure 5: PubMed Results Display

The second result is a 5,300-word article by a lawyer explaining the *Daubert* standard for the admissibility of scientific testimony. The *Daubert* standard only applies in federal courts; some states rely on the earlier *Frye* standard that established a threshold rule for assessing whether scientific testimony had sufficient foundation to be considered by a jury.⁸ (see figure 8)

One of the sites linked to by the Florida attorney specializing in toxic tort is the Agency for Toxic Substances and Disease Registry (ATSDR), a federal public health agency of the U.S. Department of Health and Human Services.⁹ A review of the site indicated that ATSDR's mission is to serve "the public by using the best science ... to prevent harmful exposures and diseases related to toxic substances." A site-limited search with the query "toxic substances" AND "scientific evidence" AND site:atsdr.cdc.gov returned eighty five results from only the ATSDR site. Along with case studies one of the results is an extensive study about the social and psychological effects of exposure to toxic substances.

Armed with a useful Web site, a researcher can use a "link check" application similar to a *Shepard's* process. Quality Web sites link to other quality Web sites and may expand the content of the "linked to" site. Yahoo is recommended for link checks because the search engine will run complex Boolean and top-level domain-limited searches. Note: "http://" must be included in the link check search query in Yahoo or the search will fail. To find sites that mention scientific evidence and toxic substances that are linked to the Atlantic Foundation Web site the query is: link:http://www.atlanticlegal.org AND "scientific evidence" AND "toxic substances". (see figure 9)

Three results are returned for the search; the first is the site of the board-certified civil trial lawyer located in the first search in Google.

The results from these searches of the Surface and Deep Webs remind us that the Web is a vast and ever-changing information source. By using advanced search features in specialized search tools and



Figure 6: Google Title Search Results



Figure 7: .edu Top Level Domain-Limited Search Results



Figure 8: Site-limited Search Results



Figure 9: Yahoo Link Check Results

general search engines, we can locate relevant information to supplement print and proprietary databases.

Summary of Web Search Strategies

- Determine appropriate search engines to recover information in both the Surface and Deep Web.
- Structure the search query with punctuation and groups for the maximum effect.
- Use date restrictions to narrow the results.
- Consider narrowing searches by using in-title, domain or specific site searches.
- Use link check to “Shepardize” the results. ☞

Endnotes:

- 1 www.law.duke.edu/journals/lcp/articles/lcp64dAutumn2001p5.htm
- 2 Congressional Research Service. (2003). *Internet Statistics: Explanation and Sources* (Order Code RL31270). Rita Tehan: Author.
- 3 Lawrence, S. & Giles, C. L. (08 July 1999). Accessibility of Information on the Web. *Nature* 400, 107. To verify the limited overlap in search engine results, run a search query in Thumbshots: (ranking.thumbshots.com). Other recommended search engines are: Ask Jeeves (www.ask.com), MSN (www.search.msn.com), and Yahoo (www.search.yahoo.com).
- 4 OAIster (oaister.umdl.umich.edu/oaister), Scirus (www.scirus.com), and PubMed (www.ncbi.nlm.nih.gov/entrez/query.fcgi)
- 5 <http://toxnet.nlm.nih.gov>
- 6 www.heindllaw.com/Toxic_Tort_Chemical_Florida_Lawyer.htm
- 7 www.defendingscience.org
- 8 See the National District Attorneys Association site for a list of states and whether they follow *Daubert* or *Frye* (www.ndaa-apri.org/pdf/hgn_state_chart_summary_2005.pdf)
- 9 www.atsdr.cdc.gov


**For more information
about the
Virginia
Association of
Law Libraries**
 see
<http://law.richmond.edu/vall>

Yes, Virginia
continued from page 34

manuals, citators, digests, encyclopedia, statutes, ordinances, opinions, regulations, legislative materials and *selected* periodicals. Excluded from the survey are law reviews because subject coverage expands beyond topics specific to Virginia. Excluded too are CLE seminar outlines and a few titles available from Virginia CLE that are not Virginia-specific.

Excluded from the survey's statistics are sites offering opinions from the federal courts sitting in Virginia, although their URLs and dates of coverage are included for researchers' convenience.

- 2 Gary Wilbert, phone interview by author, 17 October 2005.
- 3 Lawrence Lessig, "Let A Thousand Googles Bloom," *Los Angeles Times*, 12 January 2005, p. 11.
- 4 Kevin J. Delaney and Jeffrey A. Trachtenberg, "Publishers Challenge Google's Book-Scanning Plans," *Wall Street Journal*, 20 October 2005, p. A1.
- 5 Thomas Mann, *The Oxford Guide to Library Research*, 2d. (New York: Oxford University Press: 1998), 139.
- 6 Mann, *Oxford Guide*, 134.
- 7 Catherine Sanders Reach, David Whelan and Molly Flood, "Feasibility and Viability of the Digital Library in a Private Law Firm," *Law Library Journal* 95 (2003): 381.
- 8 Mann, *Oxford Guide*, 256.



Paul Barron manages the Stafford Library at the University of Mary Washington's College of Graduate and Professional Studies. After retiring from the U.S. Marine Corps, he earned a library degree from the University of Texas at Austin in 1997. He served in public and special libraries and as the Technology Director for Rockbridge County Schools in Lexington, Virginia. He is a National Teacher Training Institute Master Teacher, a former adjunct faculty member at Virginia Military Institute and Dabney Lancaster Community College and a past president of the Shenandoah Valley Technology Consortium. He has provided over one hundred workshops and presentations on Web searching at regional and national conferences in Virginia, California, Maryland, Texas and Washington, D.C.