

## Database Review

# New Portals Enrich STM Menu



by MICK O'LEARY

These days, much of the news about search engines covers new content types such as blogs, user-generated videos, wikis, and other Web 2.0 tools. More quietly, progress is still being made in conventional search engines, the kinds that actually provide useful information. Web 2.0 searching is trendier, but which is more important: some blog-mediated blowhard rant, or research on a promising new medical treatment? This is why we have to keep up on advances in information-based content searching, including news about Science.gov, WorldWideScience.org, Scitopia.org, and ScienceResearch.com.

These four search engines have two things in common: They cover a wide swath of STM subjects, and they use the same search technology. Beyond that, each occupies a distinctive niche:

- Science.gov ([www.science.gov](http://www.science.gov)) is the official search engine for federal government STM research.
- WorldWideScience.org (<http://worldwidescience.org>) is new and extends the Science.gov concept to international STM research.
- Scitopia.org ([www.scitopia.org](http://www.scitopia.org)) searches a rich trove of information in the content of STM societies.
- ScienceResearch.com ([www.scienceresearch.com](http://www.scienceresearch.com)) is from Deep Web Technologies ([www.deepwebtech.com](http://www.deepwebtech.com)), the company whose search technology powers the other three. It is also a portal for searching a wide range of STM content.

## Science.gov

Science.gov is a spinoff of the USA.gov portal for federal government information. In 2002, Science.gov was created to provide a specific focus on technically oriented federal STM content (USA.gov still covers science but favors consumer-level material). Science.gov is a joint project of 16 U.S. government science organizations within 12 federal agencies. The Science.gov website is hosted by the Department of Energy (DOE) Office of Scientific and Technical Information (OSTI) in Oak Ridge, Tenn.

Science.gov provides a one-pass search across a wide spectrum of federally produced research. It covers science in the broadest sense, including the physical and life sciences, medical research, engineering, technology, and science education. It also searches 30 separate databases, from major research sources (PubMed, NTIS, and ERIC) to many more specialized databases. It also covers 1,800-plus federal STM websites. Collectively, Science.gov's databases include all types of content (articles, reports, data sets, images), but most records are citations to technical documents.

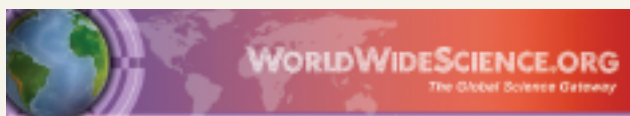


Science.gov uses search technologies from Deep Web Technologies that provide cross-database relevance searching. Searches are run against all Science.gov databases, and the results are merged into a single ranked list. An advanced search option allows database selection, date limiting, and limited field searching. Science.gov version 4.0, released in 2007, added several enhancements.

## WorldWideScience.org

WorldWideScience.org extends the Science.gov model to world science. It includes 28 databases from 18 countries, including research leaders like the U.S. (which contributes Science.gov), the U.K., Japan, India, and Germany. Like Science.gov, WorldWideScience.org was developed and is maintained by the Office of Scientific and Technical Information.

WorldWideScience.org covers a breadth of scientific information comparable to that of Science.gov, and it also has mostly bibliographic databases. WorldWideScience.org also uses Deep Web Technologies search tools.



## Scitopia.org

Professional associations and societies are major producers of STM content, and 15 of them have joined to distribute their publications through Scitopia.org, which was released in October 2007. It follows the general pattern of the others, providing a single access to multiple content sources, and is powered by Deep Web Technologies.



Scitopia.org covers the physical sciences and engineering, but it has no life science, medical, or social science content. Its contributing members include leading science societies such as the American Institute of Physics and the American Physical Society, as well as prominent engineering societies such as the American Society of Mechanical Engineers and the Institute of Electrical and Electronics Engineers. It also has patents from the U.S. Patent and Trademark Office, the European Patent Office, and the Japanese Patent Office. As with the others, most of its content is bibliographic.

## ScienceResearch.com

In addition to powering the other services, Deep Web sponsors its own wide-ranging STM portal called ScienceResearch.com. Most of the content comes from publishers' sites, including journal publishers, technical book publishers, and STM societies. The portal also includes the Library of Congress catalog and Science.gov. Most of the records in ScienceResearch.com are technical journal

## Science.gov, WorldWideScience.org, Scitopia.org, and ScienceResearch.com

### SYNOPSIS

Science.gov, WorldWideScience.org, Scitopia.org, and ScienceResearch.com provide federated searching of scientific, technical, and medical content for U.S. federal science, international STM organizations, STM societies, and general STM resources, respectively. Each is an important resource by itself, but each must be evaluated with awareness of complementary resources, whether technical or consumer-oriented.

### PRODUCERS

Science.gov and WorldWideScience.org, Office of Scientific and Technical Information, 1 Science.gov Way, Oak Ridge, TN 37830, (865) 576-1188; www.science.gov, http://worldwidescience.org.

Scitopia.org, www.scitopia.org.

ScienceResearch.com, Deep Web Technologies, 301 N. Guadalupe St., Suite 201, Santa Fe, NM 87501, (866) 388-1407, (505) 820-0301; www.scienceresearch.com.

citations, but it also has a small amount of consumer-level information from science-oriented websites.

## Where Do They Fit In?

These four science search systems stand roughly in the middle of the broad continuum of STM information. On the technical end of this spectrum are the great comprehensive discipline-oriented databases, such as MEDLINE for medicine, CASearch for chemistry, BIOSIS for biology, COMPENDEX for engineering, and so on. These are far more comprehensive in their respective fields than are Science.gov and its counterparts. On the nontechnical end of the spectrum are "regular language" databases, such as MedlinePlus for medicine, and many other consumer-oriented information sources from government agencies, publishers, and nonprofits. These sites translate technical information into everyday terms, and they occur at the top of the hit lists from the popular search engines. This information is not only comprehensible to the ordinary person, the results are full-text, unlike most of the technical sources. Don't forget one of the



fundamental principles of information retrieval: If you can't get it, it doesn't matter how good it is.

Science.gov and the others fall in between these poles, but they fall closer to the technical end of the range. Most of their content is for experts only. However, they are not stand-alone sources for serious STM research, which still

must include the dominant discipline databases. Consumers have opposite needs; they cannot readily use highly technical information, and they cannot readily obtain the full-text documents that the technical databases cite. Thus, they are actually better served by a Google search that will return reputable medical information from MedlinePlus, WebMD, and many other authoritative sources.

## **Information Literacy Works**

STM research illustrates the practical problems that occur when you don't know what information source works best for your need. In other words, when you are not "information literate." The consumer is generally not well served by Science.gov and the like; the technical researcher cannot rely upon them alone. The serious prob-

lems that arise when either researcher is using the wrong information source are easy to imagine.

This places the burden upon each of us to determine the "right" information source, and given all the possibilities, it's not an easy task. Nevertheless, lots of information is the kind of problem that we like to have. Science.gov, WorldWide-Science.org, Scitopia.org, and Science-Research.com are undoubtedly information overload, but too much is still better than too little.

*Mick O'Leary is the director of the library at Frederick, Md., and a principal in The Data Brokers. His email address is harmonyrd@yahoo.com. Send your comments about this column to itletters@infotoday.com.*