

# NAL & WorldW

**One of the major advantages of the Internet is that search engines cast a wide net when it comes to gathering up websites.**

But one of the disadvantages of the Internet is that these search engines cast that wide net without any weighting for the credibility of individual sites. How many sites must a researcher or policymaker sort through to reach only those with science-based content? And how much more daunting is the task if the information is from another country and in another language?

One way to deal with this issue is to do the search through WorldWideScience.org (WWS), a global Internet portal that simultaneously and only searches nationally sponsored science websites in more than 71 countries, many of which are not readily available through any other search engine.

Search  
Engines  
Google  
Yahoo  
WWW

## A science-based inter

“WorldWideScience makes the most efficient use of search time for those who need to find data and program information from verifiably scientific sources or for those who need to know what work is being done worldwide without having to wade through a mass of irrelevant or unscientific web pages,” explains National Agricultural Library (NAL) deputy director Eleanor Frierson, who serves as the principal representative for the United States to WWS.

Why worry about ensuring an agricultural search is international? “Because the problems that scientists are trying to solve don’t stop at one country’s borders,” says Frierson. “For example, if American scientists are looking for biocontrols for the brown marmorated stink bug, which came from Asia, access to literature in China is going to be helpful. With WWS, what other countries are actually doing about the stink bugs becomes easy to find without having to go through hundreds of folk sites.

“The key is that searchers can trust that every hit they get will be science based, so it makes the most of a searcher’s time,” she says.

But the international search access of WWS goes much further than that, adds Frierson. “It opened a vast reservoir of previously under-accessed scientific knowledge, which promotes international scientific collaboration and interaction and can reduce duplication of research efforts.”

For websites and databases to be part of WWS, they must be nominated by an authoritative agent. NAL is the designee for sponsoring most U.S. Department of Agriculture websites.

NAL sponsors its USDA nominations first to Science.gov, the Internet gateway designed to unify and simplify access and information retrieval from U.S. federal science websites. Currently, Science.gov provides cross searches of more than 45 U.S. government scientific databases and 200 million pages of science information with just one query. Science.gov, as the

# WorldWideScience.org

## national search engine

U.S. national science portal, then feeds directly into WWS.

"Before NAL proposes adding a new website, we look at its potential value to the target audiences of WorldWideScience and Science.gov, which include everyone from policymakers and planners to students, from bench scientists to public interest groups," says Frierson, who also serves as co-chair of Science.gov's governing alliance.

Many of the databases searchable through WWS simply do not show up with conventional search engines such as Yahoo, Google, or even Google Scholar. These search engines essentially work by regularly sending out "crawlers" that construct an index of websites. When a user conducts a search, the search engine consults its index. But the crawlers usually cannot conduct searches of content stored within databases. This means many topics in these information repositories never make it to the index because the content remains invisible to that search engine.

Such unsearchable content resides in what is termed the "deep web." By some estimates, the deep web is more than 500 times the size of the surface web, and perhaps 99 percent of all the web-accessible scientific documents are in deep-web databases.

To access sites of the deep web, WWS automatically converts a search request into the query style needed by the affiliated websites, and then WWS produces a single list of hits, ranked by relevance.

A test of 33 typical scientist search queries chosen across a wide range of scientific disciplines was conducted in early 2011. WWS search results were uniquely different from Google and Google Scholar results 92.7 percent of the time. Within only the top 50 results from each, WWS results were 97.6 percent unique (overlap being only 2.4 percent among these top 50 results).



In the last year, WWS enhanced its international accessibility by adding multilingual capabilities that let users query in English, Chinese, French, German, Japanese, Korean, Portuguese, Spanish, Russian, and Arabic and automatically search websites in all of the languages. The search results are translated into the language of the searcher's choice.

"This is an important benefit to the English-speaking science community because non-English sources are growing exponentially. And of course, it benefits speakers of the other languages available in WorldWideScience who need assistance with English," Frierson says. "NAL is proud to be part of the way WorldWideScience is advancing international access to the true wealth of scientific literature and data."—By J. Kim Kaplan, ARS.

*Eleanor Frierson is at the USDA-ARS National Agricultural Library, 10301 Baltimore Ave., Beltsville, MD 20705-2351, (301) 504-5248, [eleanor.frierson@ars.usda.gov](mailto:eleanor.frierson@ars.usda.gov).*

Search  
Engines

ARS