

## Keynote Address

### WorldWideScience.org Bringing Light to Grey

Brian A. Hitson and Lorrie A. Johnson  
Office of Scientific and Technical Information  
U.S. Department of Energy

#### Abstract

WorldWideScience.org<sup>1</sup> and its governance structure, the WorldWideScience Alliance<sup>2</sup>, are putting a brighter spotlight on grey literature. Through this new tool, grey literature is getting broader exposure to audiences all over the world. Improved access to and sharing of research information is the key to accelerating progress and breakthroughs in any field, especially science.

WorldWideScience.org has revolutionized access to "deep" web scientific databases. These nationally- and internationally-sponsored databases are comprised of both grey and conventional literature. Consequently, because grey literature is naturally less familiar (and, hence, less accessible) than conventional literature, it receives a disproportionate benefit in terms of usage through its exposure in WorldWideScience.org.

Before expanding on the mechanics and contents of WorldWideScience vis-à-vis grey literature, it is helpful to characterize what is meant by "grey literature." The term "Grey Literature" can be defined in several ways. Wikipedia<sup>3</sup>, for example, describes grey literature as "...a body of materials that cannot be found easily through conventional channels such as publishers..." The National Library of Australia<sup>4</sup> provides a slight variation: "...information that is not searchable or accessible through conventional search engines or subject directories and is not generally produced by commercial publishing organizations." This description goes further to describe electronic grey literature as constituting the "hidden" or "deep" web. Most laypeople, those outside the professional information community, would think of the color "grey" and may be puzzled as to why a color is used to describe literature. To them, the word "grey" likely brings to mind the Webster<sup>5</sup> definition, "an achromatic color between the extremes of black and white."

Traditionally, "white" has been equated with conventional, published literature, but perhaps to better illustrate the point, it could be useful to reverse the "achromatic" color spectrum in this case. The extreme of "black," for example, could be thought of as traditional black ink printed on paper. It consists of words that are very clear and easily accessible to everyone, and makes up the conventional literature such as journals, books, and published proceedings. "White," on the other hand, conveys just the meaning of a blank sheet with no words – simply unrecorded ideas, concepts, and thought. So, then, "grey" is between these two extremes. It includes the kinds of literature that information professionals typically associate with "grey," such as preprints, technical reports, theses and dissertations. More recently, grey literature also includes emerging forms of information such as numeric data, multimedia, recorded academic lectures, and Web 2.0-generated information.

Looking back at the National Library of Australia's definition for a moment, though, it also implies that grey literature comprises the "hidden" or "deep" web. "Grey" is synonymous with "deep" when it comes to the Internet; grey literature, more than any other type, is a body of information that resides in the "deep web" and is not easily found.

To put this concept in context, there is a distinction between the "surface web" and the "deep web." Generally, major search engines such as Google<sup>6</sup> and Yahoo!<sup>7</sup> are searching web pages on the surface web. These are static web pages that are crawled by Google's automatic crawler, where every word on a page is stored in Google's massive index, and the power and sophistication of Google's systems allows it to return millions of hits in milliseconds.