

**Statement
of
Saul Schniderman
representing
The Library of Congress Professional Guild
AFSCME Local 2910
before the Committee on House Administration
concerning the
World Digital Library
July 27, 2006**

Chairman Ehlers and Ranking Minority Member Millender-McDonald,

My name is Saul Schniderman and I am the president of the Library of Congress Professional Guild, AFSCME Local 2910 representing over 1600 professional employees working at the Library of Congress. These employees are dedicated to providing the best possible service to Congress and to the American people.

On November 22, 2005, James H. Billington, the Librarian of Congress, proposed to create a "World Digital Library" for the purpose of bringing people together "by celebrating the depth and uniqueness of different cultures in a single global undertaking." On the same day Sergey Brin, the multi-billionaire cofounder of Google, Inc., announced that Google was contributing three million dollars to the project. "Google supports the World Digital Library," he said, "because we share a common mission of making the world's information universally accessible and useful."

In fact, the Library's mission is not the same as Google's mission. Mr. Brin says that Google's mission is to "organize all of the world's information and make it universally accessible and useful." Google has taken on a monumental task for one corporation, especially considering the copyright restrictions inherent in such a project. The Library of Congress' mission, on the other hand, seems to be more realistic and focused: "to make its resources available and useful to the Congress and the American people and to sustain and preserve a universal collection of knowledge and creativity for future generations." The Library's mission is based on public service and reflects traditional democratic values of privacy rights, intellectual freedom, and open access.

In the past year, Library staff members have been hearing all sorts of pronouncements from top management that drastic changes are afoot and that the Library is in the process of making a "transition" to digital formats. "Providing access" to our collections, we are told, means digitizing them so that they are not "confined" within our Capitol Hill walls. The World Digital Library project has been announced and the Library will increasingly digitize its special collections for worldwide access on the open Internet. It's all part of a new vision for the Library of Congress.

But it's also a fact that the Library cannot digitize the vast bulk of its holdings while the U.S. copyright law remains in effect. Any digital project is faced with this obstacle: you can't reproduce

content on the Internet without first seeking permission from its author. The popular LC American Memory digitization project is boxed in by this copyright restriction and so its digital materials are either in the public domain or are made available with special permission.

We would therefore caution Congress not to regard the digitization of collections as the Library's central mission or core function. For example, our highly regarded American Memory project has digitized about 10 million items; but the way "items" are being counted needs qualification. Sometimes a scanned item is a photograph, a map, or a manuscript page. On the print side, the Library receives 1,200 new books every working day. The Library's vast collection of 20 million copyright-protected books cannot be digitized for free distribution on the Internet; it remains freely available only to onsite researchers and to other library users throughout the country via inter-library loan.

The point is this: while digitization projects are useful and prestigious, they provide access to only a microscopic portion of the Library's collections, and for that reason should not be regarded as core functions that are more important than existing operations, such as our cataloging and classification work. In its rush to "transition" to the digital age, the Library should not lose sight of its core mission nor abandon its traditional base.

Our union was established at the Library in 1976 and we represent both catalogers (print) and conversion specialists (digital). Many of us have spent our entire careers at the Library and have witnessed significant changes over the years. But when the Library hires consultants for a "Digital Competencies Development Project," and a "Workforce Transformation Initiative," we wonder out loud: what are they really talking about? Where are they leading us? And why?

Part of Library management's vision of a "digital library" is an assumption that collections that are digitized and "freed" from their physical locations on Capitol Hill can be made accessible to the world through the Google search box. This gets a bit complicated so I urge the Committee to visit our website at www.guild2910.org for a more in-depth discussion of whether Google's keyword searching will eliminate the need for LC cataloging and classification.

The problem with the Google keyword search box is this: it is limited in that it displays the search results in "relevance ranked" order yet cannot adequately separate the wheat from the chaff. Google may be great for the quick information seeker, great for business, great for online consumers, great for the undergraduate student sitting in his or her door room cramming for a term paper due tomorrow. But it's not great for the serious researcher who wants an overview of sources relevant to a particular topic. It is not great for the scholar who needs to do extensive and in-depth research.

Typing key words into a Google search box provides access to digitized information through "relevance ranking." When you search on Google your "hits" are "ranked" through an automatic computer algorithm, determined primarily by counting the number of other Internet sites that link to the sites retrieved by the keywords you typed into the box. When I type the words "capital punishment" in the Google box I get over 10 million "hits" ranked according to relevance. Even if I refine my search under Google's Advanced Search technique, I still retrieve all sorts of digital materials of no interest to me. In short, I can't discern the wheat from the chaff and, after a while, I

give up on Google and try something else.

This is why scholars utilize the Library of Congress online catalog or come to the Library to conduct research. Catalogers at the Library create standardized category terms that "round up" all works on the same subject – no matter which keywords their authors have used, and no matter which languages the authors have written in. Because Library catalogers utilize conceptual categories known as "subject headings" researchers are able to systematically retrieve all of the works relevant to their topic, not just those that have the exact words the searcher could think of. Google searching, essentially, is confined to retrieving only the words that get typed in.

For example: if you searched the subject "capital punishment" in a Library of Congress catalog, the catalog would round up such titles as "The Ultimate Coercive Sanction," "To Kill and Be Killed," "Habeas Corpus Issues," "Fatal Error," "A Life for a Life," and many others in both English and a variety of other languages. This is because trained LC catalogers have applied a conceptual cataloging and classification scheme, not a computer algorithm that just looks at keystrokes. It is no wonder that thousands of libraries throughout the nation, large and small, depend upon LC bibliographic records for their quality and accuracy, as the best alternative to Google searching. And it is no wonder that, quite recently, the Executive Board of the American Library Association praised LC cataloging but also raised serious questions about management's apparent willingness to radically change, or even give up entirely, the practice of supplying LC subject headings .

Because of its design, Google searching gives you a super speedy retrieval of "hits." Unfortunately, the search results are often incomplete, haphazard, indiscriminate and largely confined to English language sources. AFSCME member Thomas Mann, a veteran reference librarian in the Main Reading Room and the author of *The Oxford Guide to Library Research* wonders if the national library of the United States is "giving away the birthright of American scholars in exchange for a mess of Internet pottage." His paper can be viewed on our Web page.

The Library of Congress is the nation's oldest cultural institution. It is loved by the American people, by the Congress, and by the staff who work here. The World Digital Library may be very prestigious, but despite the grandeur of its name, only a very small percentage of our collection can be made available over the Internet because of copyright restrictions. Such projects should not be allowed to displace core library functions.

If we go forward with Google, Inc., let's respect each other's competencies and unique roles in society. But let us always maintain support for the Library of Congress and its mission.

##

The Library of Congress Professional Guild
AFSCME Local 2910
Room no. LM G-41
Library of Congress
Washington, DC 20540-9994

www.guild2910.org
guild@loc.gov
phone: (202) 707-6493 or (202) 707-1304
fax: (202) 707-1873