

Pinpointing preservation possibilities

Washington University uses collection analysis to determine digitization targets **BY BOB MURPHY**

WASHINGTON UNIVERSITY IN ST.

Louis, Missouri, has a new tool in its preservation efforts. And it's one you might not think of immediately.

The OCLC WorldCat Collection Analysis service.

With this new OCLC service, the university is comparing its library collection to two of those participating in Google's massive book digitization project. The goal is to identify unique items the university should digitally scan and avoid duplicating Google's work.

Among the findings:

- 700,000 public domain works not in Washington University's collection will be freely available through Google.
- 150,000 public domain works in Washington University's collection will be freely available through Google and could perhaps be removed from library shelves.
- 400,000 Washington University works are not held by the two libraries planning full digitization with Google.

The collection analysis of Washington University Libraries was fast and easy, according to Shirley Baker, Vice Chancellor, Information Technology and Dean, University Libraries. She got the idea earlier this year.

At the Coalition of Networked Information meeting April 4, Baker heard representatives of the libraries participating in the Google project in a "provocative and delightful panel discussion." She also heard a presentation on the characteristics of WorldCat by Brian Lavoie, OCLC Senior Research Scientist, and Roger Schonfeld, Coordinator of

Research for Ithaka, a nonprofit organization that helps accelerate the adoption of productive uses of information technology for higher education.

"I began to see that with Brian's findings and some additional data, I could provide truly substantive and gripping information to my council," Baker said. So she placed her order for OCLC WorldCat Collection Analysis, and had the data she needed within days.

In July, the university will use collection analysis to identify unique holdings by title. "We will know which of those 400,000 titles no other library holds—probably between 7,000 and 30,000. Those will become our preservation and digitization targets," said Baker.

Baker said numbers from the OCLC WorldCat Collection Analysis service helped make a solid case for continued acquisitions.

"I presented these findings to deans and vice chancellors," said Baker. "Usually they sit politely through library presentations and have a few questions, but this time they were riveted and the chancellor had to cut off discussion so we could go on to other agenda items. The presentation certainly gave them useful cocktail party conversation.

"The most tactically useful outcome from the libraries' point of view was showing that about 82 percent of all library holdings would not be freely available through Google," she said, "and that, since copyright lasts more than 75 years, we would have to continue to acquire current publications."