Overview of Cataloger's Desktop Modernization Project

Background

Cataloger's Desktop is a quarterly web-based cataloging documentation service that provides access to a large number of electronic manuals and procedural documentation that describe standards for performing cataloging and classification tasks in a library. *Desktop* was originally released on CD-ROM in June of 1994 and included 10 LC-authored standards. By November of 1998, *Desktop* had grown to 34 cataloging standards, and included AACR2, the first resource not published by LC. In 2000 *Desktop* began including hypertext links to cataloging resources on the web.

By early 2003 it became clear that the web would provide a more flexible, more capable means for delivering *Desktop*. Work on migrating *Desktop* to the web began in earnest in late fall of 2003, and after extensive subscriber consultation and testing, *Cataloger's Desktop* debuted as a web application in June of 2004. The CD-ROM version of the service was discontinued at the end of 2005.

Moving Desktop to the web provided a number of benefits. Chief among them was that cataloging standards developed by other institutions (such as Library and Archives Canada and the National Library of Medicine) that were distributed via the web could now be fully incorporated into Desktop without the need for LC staff to manually adjust the content. In subsequent years the number of resources that have been incorporated into *Cataloger's Desktop* has risen to more than 270, with 240 of these being fully indexed and searchable.

Changes in the Search Marketplace

Searching in *Cataloger's Desktop* is fundamentally based on keyword indexing. This type of searching is well understood and is particularly useful to users who have some idea of what they are looking for, where to look, and what vocabulary is probably employed. Keyword searching is the most widely used form of searching, and forms the basis for the best-known search portals like *Google* and *Yahoo!*

Although keyword searching will continue to be a very important approach to information discovery, it does not reflect how most people prefer to uncover information. Most do not know what they are looking for, or where to find it at least some of the time. Peter Morville, in his 2005 publication, *Ambient Findability*,[1] makes the case that most searching is not for known items, and is not linear. Indeed, discovery that requires linear, keyword searching is frequently seen as user unfriendly. This reality will become progressively more prevalent as cataloging activities are performed by library staff with less training and experience.

---

[1]Morville, Peter, *Ambient Findability*, O'Reilly, 2005.

Discovery techniques that have seen considerable development in recent years include:

- Fuzzy matching
  - alternative spellings (e.g., "catalog" and "catalogue")
  - approximate spellings (e.g., "cort" and "court")
  - synonyms (e.g., "car" and "automobile")
- Finding or excluding similar resources
  - This approach is extensively used in some e-commerce sites like amazon.com, where prospective search results are adjusted to reflect past search behavior.
- Dynamic drill-downs, based on facets of the content
- Contextual analysis, bringing together resources based on how search results are used in context
- Search relevancy, measuring how well search results match the intent of the user's query, bringing the most relevant result first in the result set
- Search history, remembering the subscriber's searches and refine his/her future searches based on their past searches
- Query federation, enabling the user to run searches against multiple search engines.
- Search through navigation, allowing the subscriber to search by clicking while navigating search results, because the system knows who they are and what they are doing, thereby searching by knowing context and intention.
- Personal searching, enabling the subscriber to search across *Cataloger's Desktop* 3.0's content as well as that of their own PC
- The search engine must mimic subscriber search behavior, bringing together related concepts that are not dependent on predefined hypertext links, enhanceable through subscriber-defined social tagging

The Cataloger's Desktop Modernization Project

Work has begun to modernize *Cataloger's Desktop's* capabilities. In addition to incorporating the search and navigation capabilities in the previous section, the following work environment and interface enhancements will be made:

- Subscriber customizable interface – this will allow the subscriber to display information in a way that reflects how they conceptualize the information. Ideally this will allow them to incorporate other information into the *Cataloger's Desktop* 3.0 environment, such as through "mash-ups."
- Resources organized in a more intuitive way, highlighting the most important resources and tying everything to the MARC 21 formats.
- Provide visual clues and icons that will help subscribers conceptualize the nature of the resource they are consulting and how it interrelates with other resources.
- Subscriber controls how much or how little s/he sees when consulting a resource.
- Incorporate RSS feeds from the Library of Congress and other appropriate sources.
- Facilitate linking into *Cataloger's Desktop* 3.0, and/or incorporation of *Cataloger's Desktop* 3.0 into subscribers' cataloging applications.

- Subscribers will be able to drag and drop shortcuts ("icons") to their computers' desktops for specific, frequently consulted *Cataloger's Desktop* 3.0 resources.
- The interface will build pages on the fly based on result sets and the characteristics of type of results.
- The interface will alert subscribers to changes in *Cataloger's Desktop* 3.0 resources based on the subscriber's search history and their customizations (e.g., notes and bookmarks).

There are some specific data enhancements:

- **Federate searching** – *Desktop* 3.0 will federate searching of the approximately 30 database resources (e.g., LC's online catalog) currently linked to in *Cataloger's Desktop*. This feature will enable the end-user to execute a single search that will query all content resident in the product, while simultaneously executing the same search in all of the database resources utilizing their native search utilities.
- **RSS feeds** – *Desktop* 3.0 will incorporate RSS feeds for both the resources incorporated into *Desktop* 3.0 and resources cited in *Desktop* 3.0.

Timeline

The Library of Congress is collaborating with InfoSolutions of Crestview Hills, Kentucky to accomplish this project. It is anticipated that all or most of the discovery techniques, work environment and interface enhancements discussed in this overview will be incorporated into *Desktop* 3.0. The Desktop Team is currently assessing what is needed to accomplish each enhancement and what their relative priorities are. Given the breadth of change, some may be deferred to after the "day one" product release, but should be incorporated later as development resources allow.

A beta release will be ready in time for the American Library Association annual conference in early July. The target completion date is September, 2009. Any questions can be directed to Bruce Johnson at (202) 707-1652 or bjoh@loc.gov.

February 19, 2009; rev. April 9, 2009