

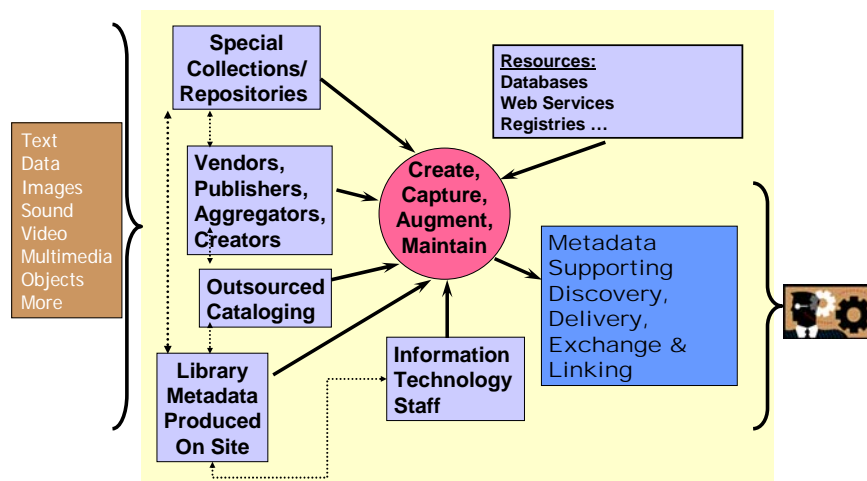
**Response to Members Council on the “Innovative Ideas” Session**  
 Karen Calhoun, Vice President, WorldCat and Metadata Services  
 May 13, 2008

At the February Members Council meeting, Bruce Willms and Tony Ferguson offered creative, even provocative ideas related to library metadata creation and exchange.<sup>1</sup> Bruce suggested automatic generation of bibliographic descriptions using publisher metadata. Tony suggested setting up metadata production centers in Asia’s developing countries that have well educated workers but lower wage costs.

On the surface Bruce and Tony advocated different solutions for today’s resource intensive cataloging practices. Yet both proposals rest upon an emerging widely distributed model for metadata creation and exchange, both within the library world and beyond it. Interestingly, in their proposals both Bruce and Tony moved beyond the local solutions—re-engineering or workflow redesign at the local library level—that have for the last decade dominated best practices to streamline and reduce the costs of library back office operations. Instead, underlying Bruce and Tony’s proposals is the complex, loosely coupled network of firms and organizations (including libraries) that deploy metadata in a global information supply chain, whose purpose is to connect information seekers to content they want, when and where they want it.

In the global information supply chain, metadata plays critical discovery, delivery, and linking roles. Librarians increasingly recognize (at least intuitively) that useful metadata can and does come from many sources. The trend toward technology-based methods for reusing metadata produced outside libraries has been gaining strength since the 1990s. Creating metadata through traditional cataloging operations, staffed on site, is one choice among many for connecting information seekers and content. This realization has emboldened a few library decision makers to embrace outsourcing on a scale that permitted recapturing significant resources for new strategic initiatives. In effect, a growing number of librarians are setting aside prior assumptions about the proper division of labor for producing metadata.

**The Information Supply Chain Network:  
 A Widely Distributed Metadata Creation and Exchange Process**



<sup>1</sup> *Automated Cataloging*, by Bruce Willms, *Open a Cataloguing Factory in Asia*, by Tony Ferguson.

This figure is an attempt to provide a bird's-eye view of the inputs, flows and outputs in a widely distributed metadata creation and exchange process, with players both inside and outside the library. Solid lines in the figure are data flows. Dotted lines represent inquiries and responses among the various players. Some dotted lines have been eliminated to make the figure a little easier to read; but the idea is that all the players are communicating. The output of the processes in the center is metadata. The figure illustrates a supply chain for making end to end connections between content (the box on the left) and those who seek it (the end user on the right), with many organizations playing roles across the supply chain network.

OCLC is responding to the wider distribution of responsibility for metadata creation and exchange among participants in the global information supply chain in a number of ways:

1. In the Next Generation Cataloging pilot, we are automatically generating MARC records using ONIX (**ON**line **I**nformation **eX**change) metadata from publishers and vendors. As suggested by Bruce, OCLC is building the infrastructure to support automated cataloging and developing the necessary alliances with publishers and vendors. At present, OCLC has nearly completed this year's effort to establish proof of concept through the pilot, which is scheduled to conclude in June 2008, with results presented at ALA Annual.
2. We are re-engineering OCLC Contract Cataloging services to offer a technology-based, significantly more affordable option for libraries' and publishers' current, backlog, and "hidden collections" cataloging. OCLC's operations are powered by automation and data mining of WorldCat. For many libraries in the Americas and EMEA, the new service is likely to be less costly than maintaining on site operations, especially for currently received library materials. Feedback thus far from libraries that have chosen fully outsourced OCLC cataloging indicates sufficient savings to permit redeployment of staff to public services, collection management, and other strategically important library projects.
3. We're significantly enhancing WorldCat data loading processes to permit more metadata, and more types of metadata (for special collections, archives, images, digital repositories, etc.), to be loaded faster. The choice of many OCLC libraries to contribute descriptive and holdings metadata as files—rather than one record at a time in the online system—has resulted in an order of magnitude increase in the volume of records requiring processing over the past ten years. A rapid rate of growth in data loading is expected to continue in FY09 due to changes in the metadata landscape described above.

Members Council delegates have been particularly helpful in keeping OCLC apprised of how the library environment is changing and in encouraging OCLC to develop creative solutions to new challenges. In the area of data loading, delegates have for several years been providing helpful advice on the need to shift priorities from almost exclusively adding holdings to equal priority to adding both holdings and new records. OCLC appreciates delegates' valuable advice and their patience as batchloading has transformed from a little-used option for adding holdings to WorldCat to the primary source of new records and holdings.