ACIG Report July 1996

Subject Authorities in the Nineties
Joint Program of LITA/ALCTS-CCS Authority Control in the Online Environment Interest Group (ACIG) and ALCTS-CCS Subject Analysis Committee
ALA 1996 Annual Conference
Sunday, July 7, 1996, Empire Room, Waldorf-Astoria Hotel

Martha Yee (UCLA Film and Television Archive)
Display of Subject Subdivisions in Online Public Access Catalogs

Yee cited statistics to highlight the importance of subject access for searchers. Studies have found that up to 90% of the subject headings searched in catalogs will also be present with subdivisions, and a study of long subject searches in MELVYL revealed that 72% were subject searches. Large retrieval sets on subject searches cause great difficulties in OPACs, one of library critic Nicholson Baker's complaints. Many software systems limit the size of retrieval sets. Yee disagrees with this approach, and outlines some steps to make displays that include subdivisions more meaningful and more logically organized. Among them are recommendations to include the syndetic structure in the display (e.g. showing broader/narrower term relationships); to practice a structured approach in displays of headings with multiple subdivisions, by using the subfield delimiters to sort subdivisions; to display subdivisions directly below the core heading without interposed elements (e.g. WaterAbstracts should not be followed by Water Beds); to practice compression to summarize long displays; and, to display chronological subdivisions in chronological order.

Jane Greenberg (School of Information Sciences, University of Pittsburgh)
Reference Structures: Stagnation, Progress, and Future Challenges

Greenberg's paper reflected three months of gophering to library catalogs and making queries on the electronic discussion list Autocat to explore the availability or lack of reference structures in online public catalogs. Greenberg identified three principal types of relationships (equivalent, hierarchical, and associative), and noted widespread inconsistencies in how OPACs made use of these relationships, or even displayed them. Greenberg attributed this to the greater interest that OPAC developers have shown in refining search features, at the expense of drawing the relationships among the items retrieved; this was evidence of stagnation. She saw evidence of improvement through systems content and design (many libraries load LC subject authority file, though not making full use of its relational features yet; some prototype catalogs that are working in that direction; better interface design) and research (including work with linguistics, object modeling, and ontologies). She identified future challenges in rationalizing reference structure relationships, integrating multiple subject headings lists and thesauri, developing new encoding formats for subject authority records, and garnering the cooperation and support of vendors and institutions in pursuing these improvements.

Mary Micco (Computer Science Department, Indiana University of Pennsylvania)
Improving Subject Access on the Web

This is a rapidly-evolving area, with new generations of tools appearing all the time (e.g. the Digital Library Initiative). Attempts to exercise some authority control over the Web must begin from a broader view of the concept of "authority:" the Web represents a wide range of information packages (e.g. homepages, gopher, OPACs) with little quality control (the "democracy of the Net" and the result of author-supplied cataloging, classification, indexing, and abstraction) which needs to be accessible to users who may not be sophisticated searchers. Micco claimed that existing access tools (LCSH, thesauri, indexes, class schedules, keyword searching) are inadequate, and was critical that LC classification schedules were not yet online (eliciting a response from Bruce Johnson of LC re: Classification Plus). Existing tools were designed to support catalogers and librarians; they often reflect a parochial, collections-based view, and are geared to books and periodicals. The new tools that Micco identified (probabilistic search engines, intelligent search engines, automatic indexing, filters for sorting and ranking search results) sort the chaos, rather than attempt to impose any order. In a fast-paced, dense presentation, five summary points for improving subject access on the Web...
were offered:

- Take advantage of classification systems
- Build interfaces to assist with filtering search results
- Develop expert systems to assist users to index
- Professionals become involved with design of search engines
- Develop minimum standards for content of Web resources

Karen M. Drabenstott (University of Michigan, School of Information)
Authority Control in Digital Libraries

The UMDL (University of Michigan Digital Library Project) seeks to develop a virtual library of information resources in all formats, principally in the earth and planetary sciences. Drabenstott defined the digital library as a dynamic, federated structure that provides an end user with physical and intellectual access to large amounts of information. Among its salient features: distributed access to many collections/databases, diverse in form and function, with some interactivity. Users must be able to search across the breadth of the digital library, and will use it for functions other than simple information retrieval (one example given was to follow the lifecycle of research on a particular topic). One of the innovations of the project for retrieval in the digital library is the "collection registry," where publishers (which could in future be anyone on the Internet) characterize their publications in general terms to help filter cross-database searches. Authority control in the digital library includes developing language for users to describe what they want, and developing language for describing collections (including non-textual ones), with the goals of targeting collections with the most potential for use in a given search, conserving system resources, and saving users money (assuming that there will be no free lunch). This involves borrowing from existing lists of subject headings, thesauri, genre/form lists, and the like; identifying subject areas that are thus well-described; effecting the links between multiple vocabularies in terms of broad subject areas, "transitioning" across multiple lists; and borrowing from existing vocabularies to extend control to the "have-nots."

Following the program, the ACIG business meeting convened with approximately 15 persons in attendance. James Maccaferri presided. He received a report from Mark Scharff, liaison to ACIG from the Bibliographic Control Committee of the Music Library Association. Scharff reported on the activities of the BCC's Authorities Subcommittee, including the sponsorship of a live demonstration of the Cataloger's Toolkit at the Seattle meeting of MLA, preparations to mount a copy of the Final Report of the Working Group on Types of Composition as a Web document, lobbying LC to allow NACO participants to use the 680 field for public notes in authority records, and preparing a checklist of authority-related specifications to use in evaluating vendor OPAC software. He reported on the series authority workshop held at MOUG, an event whose impetus arose from the Authorities Subcommittee. An area of future interest will be proposals to institute 7xx fields in authority records to allow them to be "internationalized." Little other business was transacted. The new chair of the committee will be Sherry Kelley, from the Smithsonian Institution Libraries.

Respectfully submitted,
Mark Scharff, Chair, Authorities Subcommittee

Last updated February 4, 2000