Subject Indexing and Classification, 1998–2001

The major research event of the past three years, at least with regard to presence in the literature, was the study of librarian and end-user understanding of Library of Congress Subject Heading (LCSH) strings with subdivisions conducted by Drabenstott and colleagues. The study showed that librarians misunderstood a set of complicated Library of Congress (LC) headings at a rate of about 50 percent, while for end users it was nearer to 75%. Headings arranged according to the new standard subdivision order were slightly—but only slightly—more likely to be misunderstood. The study description and major findings were reported, with some variance in emphasis and detail, in articles in *Library Resources & Technical Services*, *Cataloging & Classification Quarterly*, *Reference & User Services Quarterly*, and *Technicalities*. *Technicalities* also published brief discussions by Drabenstott and Warner of what the study might mean for LCSH. The full report is available from the University of Michigan School of Information.

Hearn (2000), in his review of the possibilities for machine-assisted validation of LC headings, critiqued Drabenstott’s assumptions about the role of subdivided headings in information retrieval, maintaining that they exist to bring order to large files and facilitate browsing. Browsing also received welcome attention from information scientists such as Allen (1998), who drew a distinction between “conceptual and spatial representations of information” and argued that their usefulness depended on a user’s particular cognitive abilities, and Johnson, who presented a prototype system that draws on syndetic structures to illustrate subject relationships.

At the same time, the prospect of improving LCSH in some way inspired a variety of projects. Cochrane reviewed issues raised in her 1986 book about LCSH for signs of progress. Olson criticized LCSH (2000) and library classification schemes (2001) as culturally biased and reflective of prevailing prejudices rather than truly objective. An alternative subject vocabulary approach, facet analysis, sparked a flurry of interest as Ellis and Vasconcelos (1999) considered it as a Web organization tool while Spiteri (1999) found in an analysis of 14 faceted thesauri little consensus as to what constitutes a facet. Increasingly considered a prime candidate for facet treatment, form/genre access in LCSH was investigated by Miller (2000) and by O’Neill, Chan, and Childress (2001). Wilson, Spillane and Cook (2000), meanwhile studied the impact on circulation of subject headings for fiction.

Issues involving the multiplicity of subject languages continue to get attention. MacEwan (2000) discussed the need and potential for linking LCSH to vocabularies used in non-English-speaking countries. Chan, Lin, and Zeng (2000) developed a pilot project for multilingual subject access to the Web. Heiner-Freiling (2000) reported on a survey of subject-heading languages used in various national libraries (most use LCSH), and Martínez Arellano Yanez Garrido (2000) reported on a survey of classification schemes used in Latin American libraries. Hoerman and Furniss offered a comparative analysis of principles governing LCSH and the IFLA Principles Underlying Subject Heading Languages.

Voorbij (1998) compared the subject retrieval potential of title keywords and subject descriptors, finding that the latter enhanced retrieval for about half the records studied. Martínez Arellano (1999) analyzed records in a catalog with large amounts of both Spanish and English material and found a controlled vocabulary particularly advantageous in that setting. Sclafani (1999) also assessed the relative value of keyword and subject heading searching. The advantages of using classification markers to organize retrieval sets in automated information systems were investigated by Kwasnik (1999), Jörgenson (1999), and Vizine-Goetz (1998). Gordon (2001) showed how everyday activities could be used to identify links between terms in a thesaurus to support subject browsing for a collection of photographs. Gottlieb and Dilevko (2001) studied decision making in individuals’ classification of their Web bookmarks.

The principal focus of research activity now is automated subject indexing. This must be said even though the people carrying out this research are for the most part not librarians. Hardly an issue of the *Journal of
the American Society for Information Science and Technology, the Journal of Documentation, or Information Processing and Management is published now without at least one or two reports of new research projects (e.g., Moens and Dumortier 2000; Mostafa and Lam 2000; Mostafa, Quiroga, and Palakal 1998; Roberts and Souter 2000; Wu, Fuller, and Wilkinson 2001). Theoretical works abound as well, including a major assessment by Anderson and Pérez-Carballo (2001) of the appropriate roles of human and machine indexing as technology continues to advance. Mai’s analysis of the intellectual process of subject indexing asserts that the indexer herself creates the subject matter of the document being indexed and that much depends on the indexer’s social and cultural context (2001). Finally, Fugmann’s essay questions the assumptions behind the quest for totally automated subject access, including the fallacy of user-friendliness based on convenience alone and the inverse relationship between precision and recall (2000). Insisting that “interpretation is a requisite for any text understanding and, hence, for any sensible text processing” (39-40), he issues a call for “research and development in what may be called an information philosophy” (40) as distinct from information technology.

Works Cited


Fugmann, Robert. 2000. Obstacles to progress in mechanized subject access and the necessity of a paradigm change. In Saving the time of the library user through subject access innovation: papers in
honor of Pauline Atherton Cochrane, edited by William J. Wheeler. Champaign, Ill.: University of Illinois Graduate School of Library and Information Science.


Martinez Arellano, Filiberto Felipe. 1999. Subject searching in online catalogs including Spanish and English material. *Cataloging & Classification Quarterly* 28, no. 2: 45-56.


Prepared by Gregory J. Wool, Monographs Science and Technology Cataloger, Iowa State University, gwool@iastate.edu