tion system by a number of meetings in which questions of methodology were discussed. Solutions are to be found which permit both, systematic shelving of books and other library material as well as establishment of systematically arranged cardfiles. Also the system should allow computer retrieval of particular items in the thematic descriptions of documents. The findings of the present meetings will be discussed in a hearing with library directors before taking any action towards the actual elaboration of such a classification system.

Nachuniversitäre Ausbildung der ZMD

Zum Abschluß ihrer zweijährigen Ausbildung im Fach „Indexing und Retrieval" an der Zentralstelle für maschinelle Dokumentation, Frankfurt, haben unter Leitung von Dr. G. Lustig erstmalig im Dezember 1972 und nunmehr im Februar 1974 insgesamt 9 Hochschulabsolventen in Kolloquien und öffentlichen Vortragsveranstaltungen die Ergebnisse ihrer Abschlußarbeiten vorgetragen und erläutert. Folgende Arbeiten wurden angefertigt, (1972): Dipl. Math. H. Friese: Untersuchungen am Sachregister von Food Science and Technology Abstracts (FSTA) und einige Vorschläge zu seiner Verbesserung. – Dipl. Math. R. Henzler: Quantitative Beziehungen zwischen Textlänge und Wortschatz. – Dipl. Math. H. Hüther: Mathematische Beschreibung von Relationen in der Dokumentation. – H. Jaene, M. A.: Ein Verfahren zur wörterbuchunabhängigen Gewinnung von fachspezifischen Komposita. – (1974): Dr. R. Alsheimer: Probleme der Textaufbereitung im automatischen Indexing. – R. Kuhlen: Flexive und Derivative in der maschinellen Verarbeitung englischer Texte. – Dipl. Phys. H. Hoggenmäller: Das automatische Erkennen von Abkürzungen und Satzbegrenzungen in der maschinellen Textverarbeitung. – Dr. R. Kragenings: Statistische Relationen zwischen Textwörtern und Deskriptoren. – Dipl. Biol. C. Schaab: Zur maschinellen Fragemodifizierung bei der Free Text Search.

BOOK REVIEWS BUCHBESPRECHUNGEN

MALTBY, Arthur (Ed.): *Classification in the 1970's; A Discussion of Developments and Prospects for the Major Schemes.* London, Hamden: Linnet Books & Clive Bingley, 1972. 269 pp. $ 12.00 or £ 4.00, ISBN 0-208-01170-6

A useful general overview has been drawn together here, with at least one common opinion emerging from the eleven contributions: that the two major purposes of documentary classification, namely for shelf-organization and for mechanized retrieval, are *not* well served by a single system unless consciously modified to cater to the two purposes (Maltby, Lloyd, Austin, and Freeman all argue thus; Mills argues for the other, classically Ranganathanian, position; the rest do not express themselves). Another, less well documentable, point can be felt: the growing concern for a "roof classification" or "switching language" to allow for access to several indexed corpora (each indexed with a different classification or thesaurus) – without the user having to have his query reformulated in each relevant indexing language.

The outline is as follows (not all the essays will be thoroughly discussed later, since some either present no points of controversy or say nothing really new): Maltby takes up the point about various purposes thematically; J. Mills argues, in a good general outline of how to revise a classification (namely Bliss' Bibliographic Classification), the need for detail within a valid collocation, thus providing simultaneously for both purposes; M. A. Gopinath gives a dogmatic (*not* a pejorative term, in my usage!) survey of Ranganathan's theories as they have issued into the current expansion of the Colon Classification; S. K. Vann gives a history of the use of (and especially of periods of dissatisfaction with) the Dewey Decimal Classification; J. P. Immroth attempts to justify the Library of Congress Classification theoretically, and carries on his campaign to supplant LC Subject Headings and to create a general index to LC by chain indexing the schedules; G. A. Lloyd concentrates on the use of UDC as a switching system for mechanized retrieval, and points to projects for its improvement that may make that use possible; B. C. Vickery argues that alphabetical indexing systems are not enough: classification is implicit in them and is in fact necessary for superior results, E. M. Keen argues not only against classification but against controlled alphabetical indexing systems, basing himself on a survey of evaluation tests; D. Austin presents a helpfully explicit account of the development of both PRECIS and the integrative-level classification of CRG, both in terms of how they arose and of their theoretical justification; and R. R. Freeman discusses what classification can contribute to the burgeoning development of information networks, especially in the role of switching languages.

Mills' advocacy of a single system as adaptable for the two mentioned purposes relies on his view of librarianship/documentation as being a single activity: "our own subject ... (is) organizing stores for retrieval" (p. 45) – a dictum that, I at least agree, applies to libraries as well as to information centers.

Immroth's justification of LC is almost amusing in its reliance on the putative criticism of Ranganathan; even the answers to the feared charges are phrased as near as possible the way a CRG-man would have written them five or ten years ago. Amusing too are such claims as that "alphabetical order (in array) ... is one of the major advantages of the LC Classification" (p. 126), or that "the collocation of related classes within LC Classification in some instances is done in far better fashion than in the Decimal Classification. This is particularly true in the collocation of language and literature into the single class P". Most will agree that this is a superior collocation (witness UDC's recent similar move); but the example is poorly chosen: no better instance of inconsistency of collocation (of, indeed, *any* desideratum) *within* a class can be found anywhere than class P in LC. He is guilty of not checking his facts when he states that LC's approximately 100,000 entries (plus class K, law) make it

https://doi.org/10.5771/0943-7444-1974-1-47 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

"the largest general vocabulary base of any general classification scheme", (p. 127), since not only is UDC's 150,000-to-200,000 entries larger on the face of it but the fuller provision of synthetic devices in UDC would make the resultant corpus of possible expressions immensely larger than that which LC could generate. His opinion is that "certainly there are some possibilities of cross classification but these can be removed by proper use of phase relations" (p. 126) – without pointing out how phase relations enter into LC practice. There is a puzzling paragraph on p. 141 about classifying index languages; he appears to argue for an "inductive" approach (see below, on Austin); but the next paragraph carries on the point by agreeing that syndesis does in a sense classify LCSH, and that this is not displayed; and the next again argues that "a chain index removes most of these inconsistencies". It appears that he has never seen the arrowgraph technique (Rolling) as I have applied it to LCSH (*Libri* v. 15): subject headings assuredly *can* be classified without it being done "as the list is developed."

Vickery's essay is accompanied by a sample of a comparison between TEST, BC, CC, DC, and UDC in regard to Thermodynamic Properties, as well as a tabular summary of the results of all the comparisons made in a 1969 Aslib test. It would have been more helpful for the number of hierarchies involved in each of the classified systems to have been recorded: this one of the thorniest aspects of concordancing between hierarchical systems.

Keen mounts a full-scale attack on all those who suppose they can discern superior performance from the structure of an indexing system. But his argument is at least in part vitiated by his own methodological presupposition that "evaluation testing has rightly covered only the use of linkage once the search request has been finally formulated." (p. 202). In Freeman's terms this is to make the formulation of the request part not of the retrieval system but of its environment! Freeman quotes C. E. Trotter, in reporting on a 1969 testing of a hierarchical atomic energy retrieval system: "the time of search formulation would be significantly less, with consequent large savings in manpower and costs if the specialized hierarchical vocabulary, is used, and ... the construction of the hierarchical vocabulary, including its arrangement and notation, plays a major role in the retrieval process, especially owing to the ability to retrieve ranges or classes of information which are not specified by a single term." (p. 261). It is this presupposition on Keen's part that allows for part of the total system to escape from his purview, with the result that he can honestly argue that "the gains [from classification] are outweighed by losses. In providing controlled languages their artificiality and complexity introduce new opportunities for misunderstanding and error." (p. 209). But, if this presupposition is granted (and granting that if search as well as storing is uncontrolled, there is no omission of the formulation stage in the analysis (since it never occurs at all as a separate event)), his conclusion is, though not new, still as crushing as ever: "relatively uncontrolled languages used at the indexing stage cannot be improved on by controlled language". (*ibid.*)

Austin's essay is in many ways the most interesting as a piece of theory. As noted above, "it is now believed (by the CRG, at least by its dominant sector in the work here reported) that these different aims (shelf classification / machine retrieval) are not compatible" (p. 215). This thesis is justified by a decomposition of the complex classificatory expression for use in a machine retrieval system – but not without preservation of its context –: whereas the shelf system concerns itself with the subject of the document as a whole, and must decide which partial concept is most important (in order to generate a linear order), the retrieval system "has to operate at the level of the individual term or concept" (p. 222). This does not mean that he ignores "the relationship which exists between citation order and meaning" (*ibid.*), but we do find him saying that "this idea of organizing terms according to their relative significance was not only responsible for the loss of meaning in some of the index entries, but that it is also entirely inappropriate for any machine-held retrieval system" (p. 242). Unless we realize that the crucial phrase "relative significance" means something other than syntactic context (which, in its overwhelming helpfulness, Austin strongly supports), we can seriously misinterpret this passage. The apparently anomalous emphasis of the CRG work on *objects* as aginst *disciplines* is made intelligible by his pointing out that their classificatory work was initially directed wholly to analysis of terms into absolutely general categories – in other works into a thesaurus encompassing not the universe of *subjects* but the universe of *concepts*, "the intention being to organize the basic ideas out of which all subjects are constructed" (p. 226). This is the deductive part of the work; the inductive part is the work on PRECIS at BNB: each term is *ad hoc* taken as a lower term, its higher terms are stored in association with it; and these higher terms are not always genera but sometimes wholes of which the lower are parts (e. g., what is the genus of "veins"? is not "circulatory system" as its including whole what matters for retrieval?). This agrees with Maltby's suggestion that if a logical classification is wholly in terms of genera and species, library classification is illogical in order to be helpful, faceted classification most of all – but one may well quibble with so narrow a definition of 'logical'.

Freeman notes that "There is a growing convergence taking place between the practices for organizing and displaying classifications and thesauri, with more specificity emphasized in classifications and more hierarchy in thesauri" (p. 256) – surely a feeling borne out in both Lloyd's and Austin's essays.

There is one serious misprint: on p. 130 the LC class code JV is given as IV (!).

The collection is certainly uneven: no control over orientation seem to have been exercised by the editor (and of course none over substantive points). Not everything can be agreed with here, sometimes from causes deeper than mere factual accuracy. But this volume should not be ignored by any who would intend to discuss these matters seriously today.

J. M. Perreault
UAH Library, Huntsville, Alabama

https://doi.org/10.5771/0943-7444-1974-1-47 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY