
Reports and Communications

Papers of 10th Annual Conference, Mathematical-Statistical Area

At the recent 10th Annual Conference of the German Society for Classification, June 18-21, 1986 in Münster/W (see also report in Section FID/CR News of this issue for the papers from the conceptual-contents-related area) the following papers were presented and discussed in the area of (1) *Numerical Classification*: OPITZ, O.: Cluster analysis for qualitative data. - HERDEN, G.: The problem of defining distance measures with ordinal data. - SPATH, H.: Anticlustering: Maximizing partitioning cluster criteria for quantitative data. - STEINHAUSEN, D.: CLUSTER-PC: A menu-guided cluster analysis package for micro-computers. - KROLAK-SCHWERDT, S.: A stopping criterion for single-linkage analysis derived from graph theory. - GAUL, W.: The classification of marketing research institutes according to their use of data analysis methods. - BOCK, H.H.: Multidimensional scaling in the framework of cluster analysis. - ECKES, T.: The sorting procedure for obtaining proximity data in multivariate psychological research. - (2) *Multivariate Analysis*: WERMUTH, N.: Statistical structure analysis based on graphical models. - BAASKE, W.: Principle planes! Visual structure analysis by cluster biplot projection pursuit on multinational survey. - SCHUBERT, L.: Multidimensional scaling with mixed data. - TUSHAUS, U.: A graphical aid for qualitative data analysis. - MATHAR, R.: Metric scaling with restricted configurations. - (3) *Discrimination and Mixture Analysis*: KRAUTH, J.: Modification of the nonparametric discrimination method of Kendal. - STANK, S.: The decomposition of time-varying mixtures. (4) *Biological Taxonomy*: VACH, W., DEGENS, P.O.: A proposal for more robust estimation of ultrametrics. - SCHMITT, H.P., OBERWITTLER, Ch.: Numerical taxonomy of brain tumours: A challenge for modern mathematical classification techniques. - SCHERER, S., BINDER, H.: A comparison of classical and molecular taxonomy in the case of amino acid sequences. - HAESELER, A.v.: Experiences with tree constructions of t-RNA sequences. - POVEL, G.D.E.: Fitness of input demands of a classificatory process to obtain natural results. - (5) *Phylogenetic Trees*: KRÜGER, M.: Reconstruction of phylogenetic trees by Monte-Carlo methods. - LAUSEN, B., DEGENS, P.O.: Variance estimates in reconstructing phylogenetic trees. - PATZLAFF, M.: Reconstruction of phylogenetic relations by graph-theoretical methods: Results and problems in the chemical systematics of cytochrome c. - HOFMANN, H., KÜHNE, K.: A relation between the exon/intron organization of genes and the formation of fibrils in case of collagen. - (6) *Applied Data Analysis*: BÄUMER, H.P., ZAUKE, G.P.: Time series analysis as a tool for classification and exploratory data analysis in the context of biological monitoring. - MIEBACH, B.: Comparing subgroups in LISREL models on the basis of clustering methods. - BALDERJAHN, I.: Cross-validation

of simple and multiple moment structures in the case of consumer behaviour data. - (7) *Applications of Classification Methods*: KOHLSCHKE, A.J.: Cluster analysis methods in the empirical analysis of voting behaviour. - BÜHLER, K.-E.: Proposals for classification methods in the analysis of biographies. - BUSSE, M.: The misuse of classifications! some non-mathematical considerations. - GERNERT, D.: Classification as a proving technique and as a heuristic means in graph theory.

A plenary lecture was given by H.FEGER, Hamburg, on "Estimating relational structures from ordinal data". - A tutorial by P.O.DEGENS on "Estimating procedures for reconstructing of evolutionary trees by dendrograms and additive trees" and one by H.H.BOCK on "Correspondence analysis" were held before the opening of the conference. - The conference proceedings will be available through INDEKS Verlag in the fall of 1986.

Meetings of the IFLA Division on Bibliographic Control. Section on Classification and Subject Cataloguing, Chicago.

At the first business meeting of the Section on Sunday August 18, 1985 nine members were present as well as three observers. *Russell Sweeny*, the United Kingdom representative, had resigned and was replaced by Ms. *J.E. Butcher* of the British Library. Mr. *E. Fritjofsson* of the Uppsala Universitetsbibliotek has been appointed as a new member from Sweden.

Tor Henriksen, Chairman of the Section, first reported on his work in preparing for the Conference and getting speakers for the Section's open meetings. *Suzanne Jouguelet* gave the financial report. For the first time, the group requested and received project money to help Nancy Williamson in her study on the use of classification in online searching. Tor Henriksen next summarized the meeting of the Coordinating Board of the Division on Bibliographic Control. The Division will publish two reports on C.I.P. (Cataloging in Publication.) It was also announced that *IFLA Journal* has priority for three months to publish any IFLA papers. The status of project work came next. Though he could not attend, Mr. *Sukiasian* of the USSR had forwarded his report on "Classification systems in national bibliographies: surveys and recommendations." The Section decided to write him to encourage further work in this area with concentration upon "investigation of the extent to which details on edition, expansions variations and options of the various schemes are incorporated" since this is one of the goals of the Medium Term Program. Tor Henriksen, as representative to the International Standards Organization, reported that a standard on thesaurus construction may be forthcoming. He has not yet been able to provide a summary in English of his survey of standards relevant to classification and subject cataloguing.

Nancy Williamson next reported on her study of classification in online searching. She had mailed a questionnaire to 35 libraries with online catalogs and hopes to analyze the results by the end of the year. At present, she has concentrated on online catalogs rather than reference data bases and has considered both the use of

classification within bibliographic records and the role of online classification schedules in formulating searches. Furthermore, she has also looked at the use of PRECIS for subject searching. Once she can obtain an experimental data base, she would like to replicate the Dewey Decimal Classification studies on Library of Congress Classification data base. Finally, she is interested in the best way to display classification information to online catalog users.

The meeting concluded with the election of *Robert Holley* of the United States as the new chairman and *Barbara Kelm* of the Federal Republic of Germany as the new secretary.

At the first open meeting on Tuesday, August 20, *Robert Holley* gave a paper entitled "Classification in the USA" while *Karen Markey* and *Anh Demeyer* reported on "Dewey Decimal Classification Online Project: Integration of a Library Schedule and Index into the Subject Searching Capabilities of an Online Catalog." The second open meeting on Wednesday began with *Suzanne Jouguelet's* presentation on "Subject Indexing in France: Tools and Projects," in which she described the two competing systems for subject access in France. *Nancy J. Williamson* followed with "Classification in Online Systems: Research and the North American Perspective," an excellent summary of the current status of classification research in the United States and Canada.

The second business meeting on Thursday August 22 began with a discussion of *Nancy Williamson's* paper as a starting point for determining the Section's priorities for the project. For the moment, the group agreed to emphasize the online catalog while postponing a study of citation and full-text data bases to some time in the future. In addition, the group agreed to focus on traditional classification, especially DDC because of its international use, before studying thesauri and syndetic subject structures.

The Section then agreed upon its *new terms of reference*:

To be a forum for users and producers of classification and subject indexing tools.

To promote standardization and uniform application of such tools by institutions generating or utilizing bibliographic records.

To initiate and give advice on research in the subject approach to information and to disseminate the results through open meetings and publications.

In addition, the Section voted to recommend a change of its name to the *Section of Classification and Indexing* since indexing is less restrictive and less linked to card catalogs than subject cataloging. The Professional Board, however, must approve the change before it becomes official.

The following topics were proposed for the open meetings in Tokyo:

Classification and indexing in Japan (one or two contributions)

A further progress report on the Classification in Online Searching Project by *Nancy Williamson*

The publication in electronic form of the British ver-

sion of the Universal Decimal Classification by *Tor Henriksen*

As a possible future project, the group was asked to think about efforts to develop an international standard for subject authority file formats. After thanks to *Tor Henriksen* and *Suzanne Jouguelet* for their four-year service as officers, the Section adjourned for another year.

R.P. Holley

Letter to the Editor

Dear Editor!

Would I like to review M.P. Satija's booklet "On the Brevity of Notation in Colon Classification"? (1), this was your question some time ago. I do not feel inclined to review it in a professional way, however, let me share with you some reflections which came to me when looking through its pages. May I continue in German?

Nach Überfliegen des Inhalts dieses Büchleins kamen schöne Erinnerungen aus den Jahren 1972 und 1975 in mir auf: Vier Wochen Bergsteigen in Nepal, dann einige Wochen Zeit, um Indien in seiner ganzen Vielfalt in mich aufzunehmen. Natürlich besuchte ich auch einige Bibliotheken, kaufte mehrere von Ranganathans Werken und sprach mit Anhängern und Kritikern der Colon Classification, die ich 1959 in Cleveland in Jesse Sheras geistreich-spritziger Vorlesung "Theory of Classification" kennengelernt hatte. Mit Ranganathan selbst bin ich 1964 einmal auf der FID/CR Konferenz in Elsinore zusammengekommen. Meine Vorstellungen, daß sicher ein Großteil der indischen Bibliotheken die Colon Classification verwenden, und daß die indische Nationalbibliographie nach der CC angeordnet sei, mußte schon in Kathmandu revidiert werden. Vom Bibliothekar des indischen Kulturzentrums erfuhr ich, daß indische Bibliothekare zwar in der Ausbildung mit der CC vertraut gemacht werden, daß aber mehr als 80% der indischen Bibliotheken und auch die Nationalbibliographie nicht die CC sondern die Dewey Decimal Classification verwenden. Allerdings seien den Eintragungen in der Nationalbibliographie auch CC-Notationen hinzugefügt.

Auf der Suche nach praktischen Anwendungen der CC fand ich 1972 in Delhi nur eine Gandhi-Bibliographie und den systematischen Katalog sowie den Lesesaal der Universitätsbibliothek der Delhi University. Die gewonnene Anordnung oder Reihung ist zweifelsohne nützlich, aber für den Normalverbraucher sollten die Notationen unbedingt durch ausgeschriebene Klassenbenennungen verdeutlicht werden. Noch nie habe ich einen Lesesaal mit so vielen verstellten Büchern gesehen, wie diesen in Delhi. Ob das auf die komplizierten Sortier-Regeln der CC zurückzuführen ist? Mein damaliges wie jetziges Urteil über die CC: In der Theorie sehr interessant, in der Praxis, zumindest für die Buchaufstellung, nicht zweckmäßig ("a pain in the neck", wie ein indischer Bibliothekar mit USA-Erfahrung sich ausdrückte).

Wie kann man die vielen guten Ideen, die Ranganathan und seine Mitarbeiter im Zuge der Entwicklung der CC konzipiert haben, nutzen, ohne die für das Computerzeit-

alter unnötig komplizierten Notationen und Notationsregeln mitzuschleppen? Einen Weg haben uns die Engländer, die viel von Ranganathan gelernt haben, aufgezeigt (vgl. die Spezialklassifikation für Library and Information Science Abstracts (2)). Man verwendet eine durchgehende Sequenz von Notationen über alle Facetten, z.B. ABO1 bis ZY99. Dann entfallen alle Probleme mit Dokumentationseinheiten, für die einzelne Facetten nicht benötigt werden, wie z.B. "Twentieth Century Economics" oder "Architecture, Windows" (p.28 und 29 in Satijas Buch). Insbesondere entfällt auch die Umcodierung von unüblichen Zeichen, wie "single inverted comma", einem "facet indicator digit". Wenn die Facetten aus der Notation hervorgehen, sind keine verschiedenen Facetten-Indikatoren nötig.

Eine der bekannteren Notations-Erfindungen von Ranganathan ist der "octave device", der später zum "sector device" (vgl.3) verallgemeinert wurde. Nachdem ich in den Prolegomena gelesen hatte, daß der Octave Device von der FID auch für die UDC übernommen worden sei, mir dies in der UDC aber nicht aufgefallen war, fragte ich Herrn Schmidt vom UDC Central Classification Committee danach. Er sagte, das sei wohl mal beschlossen und in einer Aktennotiz festgehalten, aber in die UDC-Praxis nicht eingeführt worden. Jedenfalls hatte Ranganathan für die 7.Auflage der CC zahlreiche Einfügungen geplant, die mit dem Sector Device an die geeignete Systemstelle gebracht werden sollten. Dieses Einfügen an beliebiger Stelle ist in einfacherer Weise möglich, wenn man zum voll flexiblen Prinzip linksbündiger, durchgehender Notationen übergeht, bei dem die hierarchische Ebene nicht durch die Stellenzahl, sondern durch Einrückungen in den Tafeln angezeigt wird. Zweckmäßigerweise hängt man an die Notation auch noch eine Ebenenangabe an, zumindest für die Computerverarbeitung (vgl.4). Während der Octave Device leicht verständlich ist, sind die Regeln des Sector Devices (Satija p.11) mit T, V.X und Z als "Emptying digits" und U, W und Y als Zeichen, die sowohl "leer" als auch "leerend" sind, eher verwirrend. Zum Ergebnis der Einfügungen auf der Hauptebene fragt I.Dahlberg, ob es in Ranganathans Sinn ist, solche "Häßlichkeit" in ein anfangs schönes, harmonisches System einzubringen (5). Wenn man Ranganathans "Preview" zur 7.Auflage liest (6), sieht man, daß es durchaus seine eigenen Vorstellungen waren, die nun zu solchen Resultaten führen.

Durchgehende (nicht-hierarchische) Notationen werden ausnahmsweise aber in steigendem Maße auch in der CC verwendet, z.B. beim "telescoping in array" (p.19-23). Ein Nachteil der CC ist jedoch ihre Komplexität durch ihre Ausnahmen von Regeln und Ausnahmen von Ausnahmen (vgl.(7)).

Um die Verwendung der CC für den Benutzer und diejenigen, die Zettelkataloge und Bücherregale in Ordnung halten sollen, zu vereinfachen, könnte man zu einer doppelten Notierung übergehen (vgl. das in der BR Deutschland in den 70er Jahren von Kluth und anderen betriebene Projekt einer doppelt notierten Einheitsklassifikation). Allerdings müßte dann eine zentrale Institution mit entsprechendem Management das Notationssystem ständig auf dem laufenden halten, ähnlich dem Management der Library of Congress Classification oder

der Dewey Decimal Classification. Am DRTC in Bangalore war man 1975 für diesen Vorschlag nicht sehr aufgeschlossen. Auch Satija beanstandet (8, p.136), daß die Benutzer mit ihrer Erfahrung nicht zu Wort kommen.

Satijas Anliegen ist es, zu zeigen, daß Ranganathan in vieler Hinsicht daran gedacht hat, zu kurzen Notationen zu kommen. Bei der Entscheidung eines Anwenders, welches Klassifikationssystem er verwenden soll, spielt die Kürze aber nur eine untergeordnete Rolle. Oft ist der Umstand von sehr hohem Gewicht, daß man für viele der vorhandenen und künftigen Bücher fertige Notationen, z.B. LC-Notationen von der Library of Congress beziehen kann, und dann noch in maschinenlesbarer Form.

Eine zukünftige, wichtige Anwendung von Klassifikationssystemen wird die Gliederung von sehr umfangreichen Datenbeständen auf laser-optischen Platten sein, die mit Mikrocomputern verwendet werden. Facettierte Klassifikationssysteme, z.B. auf einem "Classaurus" beruhend, haben hier durchaus gute Chancen, in Kombination mit anderen Methoden, wie Frei-/ Volltext-Retrieval. Bei geeigneter Gliederung braucht nicht im gesamten Datenbestand gesucht zu werden.

Die Bibliographie mit 34 Hinweisen nach dem Stand von 1979 ist recht nützlich, etwa als Rückwärts-Ergänzung zu Satijas neuerer Bibliographie (8). Leider hat das kleine Werk kein Register, ganz im Gegensatz zu Ranganathan, der immer sehr gute Register zu seinen Büchern machte.

Abschließend noch ein Zitat (p.23), das mich schon beim Lesen der Prolegomena (Sec.JE6) zum Schmunzeln brachte: "To a classificationist saving of one digit gives the same joy as the saving of a phoneme gives to the grammarian. According to Pannini, the ancient grammarian of India, saving of a phoneme gives to a grammarian as much joy as the birth of the first son". Als glücklicher Vater von zwei Töchtern kann ich da natürlich nicht mitreden.

References:

- (1) Satija, M.P.: An epistle to Dean Shera on the Brevity of Notation in Colon Classification. Amritsar, India: Rajat Book Corner 1979. 66 p. = Kaula Series in Library Science No.3
- (2) Daniel, R., Mills, J. et al: A classification of library and information science. London: Libr.Assoc.1975. Reviewed in Int.Classif.2(1975)p.123-4
- (3) Ranganathan, S.R.: Prolegomena to Library Classification. 3rd ed. Bombay, etc.: Asia Publ.House 1967. 640p.
- (4) Körner, H.G.: Notationssysteme für Dokumentationssprachen und Klassifikationen. Eggenstein-Leopoldshafen: Fachinformationszentrum Karlsruhe. 1980. 135p. = BMFT-FB-ID 80-013
- (5) Dahlberg, I.: Classification - East or West? (Editorial). Int. Classif.12(1985)No.3, p.127
- (6) Ranganathan, S.R.: Colon Classification, 7th ed.: a preview. Libr.Sci.with a Slant to Doc. 6(1969)p.192-242
- (7) Bhattacharyya, G.: Towards an integral theory of notational language. Libr.Sci.Slant Doc. 10(1973)p.322-345
- (8) Satija, M.P.: Colon Classification. Literature Analysis 1970-1984. Int.Classif. 12(1985)No.3, p.133-142

Horst G.Körner

Beratung für computeris.Inform.u.Dok.systeme
Zittelstr. 5, D-8000 München 40.