
Reports and Communications

Unified Medical Language System: Progress Report

by Betsy L. Humphreys

In 1988 the Unified Medical Language System (UMLS) project moved from an emphasis on definition of requirements, identification of necessary system components, and exploration of alternative approaches and formats to the building of initial versions of key UMLS components. This change in focus coincided with the June 1988 award of five new three-year research and development contracts for support of the UMLS project. The current UMLS contractors are:

- Lexical Technology, Inc. Principal Investigator: Mark S. Tuttle
- Massachusetts General Hospital. Principal Investigator: G. Octo Barnett, M.D.; Subcontractor: Brigham and Women's Hospital. Principal Investigator: Robert Greenes, M.D., Ph.D.
- University of Pittsburgh. Principal Investigator: Randolph Miller, M.D., Associated Senior Investigator: Bruce Buchanan, Ph.D.
- University of Utah. Principal Investigator: Homer Warner, M.D., Ph.D.
- Yale University School of Medicine. Principal Investigator: Perry Miller, M.D., Ph.D.

Dr. David Evans, director, Laboratory for Computational Linguistics and Associate Professor, Linguistics and Computer Science, Carnegie-Mellon University, is serving as a special consultant to NLM on the UMLS project.

As stated in previous news articles (Oct. 1987, May 1987, Nov. 1986), the UMLS project is a major NLM initiative designed to address the fundamental information access problem caused by the variety of independently constructed vocabularies and classifications used in different sources of machine-readable biomedical information. The UMLS approach will be to compensate for differences in the terminologies or coding schemes used in different systems, as well as for differences in the language employed by system users, rather than to impose a single standard vocabulary on the biomedical community.

For the past six months, NLM's UMLS development team (Donald A. B. Lindberg, M.D., Harold M. Schoolman, M.D., Daniel R. Masys, M.D., Betsy L. Humphreys, Lawrence C. Kingsland, III, Ph.D., Alexa T. McCray, Ph.D., and Peri L. Schuyler) and the UMLS

contractors have been refining and testing the format of the Metathesaurus, a principal UMLS component that will contain information about biomedical concepts, their representation in several different vocabularies and numeric coding schemes, and their frequency of occurrence in selected databases.

On December 14-15, 1988, the UMLS contractors met at NLM to review the results of a test of the creation and editing of 1,700 sample Metathesaurus records and to discuss plans for the construction of the first distributable version of this UMLS component. The first version of the Metathesaurus, to be known as Meta-I, will include all concepts in NLM's Medical Subject Headings (MeSH®) as well as some terms from the Systematized Nomenclature of Medicine (SNOMED) created by the College of American Pathologists, CPT (Current Procedural Terminology) produced by the American Medical Association, and the International Classification of Diseases (ICD). Preliminary versions of Meta-I records will be generated from machine-readable versions of these sources by Lexical Technology, Inc., one of the UMLS support contractors. The records will then be reviewed and edited by persons with appropriate subject expertise. Meta-I is likely to have about 30,000 records. NLM has not yet established a firm schedule for its availability.

Although the development of Meta-I is an immediate priority, work continues on the design of other UMLS components, including a network of permissible relationships among the basic categories of concepts to be represented in Meta-I. There are also ongoing projects on the description of user information needs, frame-based representation of medical concepts, the development of tools for evaluating the usefulness of the results of UMLS research, and the creation of user-friendly interfaces to multiple sources of medical information.

An important objective of the UMLS research effort is to generate immediate improvements to current information services. To date, the UMLS project has resulted in enhancements to Versions 2, 3, and 4 of Grateful Med®, expanded MeSH coverage in several subject fields, the addition of new information to MeSH vocabulary records, the development of an enhanced automated system for creating and maintaining MeSH, and the creation of MeSHLINK, a file for storing MeSH equivalents to terms in other vocabularies. The Grateful Med search engine, a UMLS-inspired enhancement, has in turn been used to imbed access to MEDLINE® in the AI/RHEUM expert system in rheumatology and the QMR (Quick Medical Reference) knowledge-based system developed at the University of Pittsburgh. MicroMeSH, a micro-computer package developed by Massachusetts General Hospital for searching and graphic display of MeSH as well as constructing and executing MEDLINE searches, is also a result of UMLS research. In addition to these products, the UMLS project has generated a number of useful research tools, prototype systems, and publications. A bibliography of articles describing UMLS-related research is available by writing to NLM's Office of Inquiries (Attn: UMLS-BIB).

(From: NLM News 44(1989) Nos. 1-2, p. 6-7)

Analyse des Données. Apprentissage de Connaissances Symboliques et numériques

An international symposium under the heading "Data Analyses" and "Learning Symbolic and Numeric Knowledge" has been announced to take place in Antibes-Juan-Les-Pins, Sept. 11-14, 1989. The program is available from INRIA (Institut National de Recherche en Informatique et en Automatique, Domaine de Voluceau, Rocquencourt, F-78153 Les Chesnay Cedex). It contains some 63 papers in the two areas having different program committees. The symposium is sponsored by several national and international statistical associations, the International Federation of Classification Societies and the Société Francophone de Classification.

British Classification Society

An invitation has been circulated for a conference on "Ordination in Classification" to take place at the Conference Hall, Rothamsted Experimental Station, Harpenden, Herts., 19. Sept. 1989. It is organized by the British Ecological Society, the British Region of the Biometric Society and the British Classification Society. The program lists papers of the following speakers: Allen D. GORDON (University of St. Andrews), Peter J.A. HOWARD (ITE, Merlewood), Mark O. HILL (ITE, Monks Wood), James PERRINS, Mark WILLIAMSON, Alastair FITTER (University of New York), Jaqueline MEULMANN (University of Leiden), Pete G.N. DIGBY (Rothamsted), Jane L.P. LEFKOVITCH (Agriculture Canada). The Annual General Meeting of the British Classification Society had taken place on Friday, 16 June 1989 at the Open University.

IFLA Section on Classification and Indexing

The Newsletter of June 1989 reports on the following: "The Section on Classification and Indexing faces a major transition over the next year. The founding members of the Section will leave after eight years of service on the Standing Committee. Recruitment efforts were successful. I welcome the seven new Standing Committee members: Svetlana Artamova (USSR), Julianne Beall (USA), Donna Duncan (Canada), Chantal Freschard (France), Ines Lopes (Portugal), Werner Stephan (FRG), and Elaine Svenonius (USA). With two continuing members, Dorothy McGarry (USA) and M. Pino (Portugal), and one reappointment, Joyce Butcher (UK), the new Standing Committee will take over direction of the Section after the Paris Conference.

The Section will also have a new Chairperson and a new Secretary/Treasurer since both Barbara Kelm and I are not eligible for reelection. I will strongly encourage the new officers to continue publishing the annual Newsletter, but they will decide. It is possible that the Newsletter will not appear next year. With increased interest from individuals and libraries, the Section mailing list now has over 100 entries.

For the 1989 Paris IFLA Conference, the Section has a varied program, described below in greater detail. I am especially pleased with the all day workshop on the 20th edition of the Dewey Decimal Classification. Nancy Williamson's completed Section Project on the potential use of the Library of Congress Classification in online catalogues has led to a draft USMARC format for classification. In the long term, the research begun by her Project may eventually result in a UNIMARC format which can accommodate the major internationally used classification schemes.

At the end of my term of office, I wish to thank the supporters of the Section, members of the Standing Committee, and especially Barbara Kelm, Secretary/Treasurer. Best wishes to the new Standing Committee for continued success in achieving improved access to subject information.

Robert P. Holley, Chairperson
IFLA Section on Classification and Indexing"

The Newsletter contains also the announcements of the Section meetings at the Paris IFLA Conference. A report of the Chairman on the Section is scheduled for Monday, Aug. 21, 89; three papers will be presented on Aug. 22, viz.: Marcia BATES: "Designing online catalog subject access to meet user needs", Suzanne JOUGUELET: "L'accès par sujets et le marche de l'information bibliographique en France", and Joyce E. BUTCHER and Robert R. TROTTER: "Building on PRECIS - strategies for online subject access in the British Library". The all-day workshop "On the 20th Edition of the Dewey Decimal Classification" will be held on Thursday, Aug. 24, 1989. Registration is limited to 100 participants. Further information is available from Mr. Peter Paulson, Forest Press, 85 Watervliet Ave. Albany, NY 12206-2082, USA.

LC Cataloging Distribution Service: Latest Products

The Library of Congress Subject Headings is now available in four formats: 1) casebound 12th edition (the red book), \$150; 2) CDMARC Subjects (a single CD-ROM disc) for \$300 US, \$370 international; 3) microfiche, \$80 US, \$85 international; and 4) Subject Authorities (magnetic tape in USMARC format), \$3,120 for a 1989 subscription and \$857 for the 1986-88 retrospective file. In addition the new third edition of "Subject Cataloging Manual: Subject Headings" is now available. Prepared by LC's Subject Cataloging Division, this work substantially revises the 1985 edition and includes all the update pages in 1987. For the first time, the manual comes as a two-volume binder set. It includes many more usage guidelines for free-floating subdivisions, all in an updated form. Also, 60 new instruction sheets explain how to use specific subdivisions. A useful bound and hole-punched index is included with the basic text. The \$65 subscription price includes the basic text and an annual update. Extra copies of the index are \$15 each.

Another recent publication guides readers through the electronic world of information retrieval. "Intelligent In-

terfaces and Retrieval Methods for Subject Searching in Bibliographic Retrieval Systems" explores software technologies in 'intelligent interfaces'. It also examines 'intelligent' retrieval systems and online public access catalogs, which offer users powerful capabilities, such as suggesting alternative search strategies during search and retrieval operations. This special purpose-software employs logic to help solve many user problems automatically. It is available for \$20 from the Library of Congress Cataloging Distribution Service, Customer Services Section, Washington, DC 20288-5017.

At the IFLA Conference in Paris, CDS will be represented at the exhibition. It sent out invitations and offers to give away – at this opportunity – free demonstration diskettes for their CD-ROM system and sets of their brochures.

Meetings 270 and 271 of the British Classification Research Group, London

On Febr. 16, 1989 the 270th meeting of the CRG took place at University College, London, with ten Persons present. Discussions centered (1) around the so-called Prolegomena Classes of the Bliss Classification, 2nd ed. (BC2), introduced by Mr. MILLS, treating the Universe of Knowledge, Communication Skills, Communication, and Documentation, Library and Information Science. A definition of Prolegomena Classes was proposed as classes consisting of those studies which deal with the production of knowledge and are too general to be subsumed under any one of the disciplinary classes. It was, however, decided to exclude Systemology from the Prolegomena Classes and place it at the beginning of Science and Technology (AY); (2) the classes for general science, which Bliss had provided in AK (Formal Science), AL (Logic and Mathematics), and AY (for Empirical Science and Technology). Mr. Mills concluded the meeting by saying that some tidying up of the Common subdivisions would have to be undertaken as the result of the discussions.

The 271st meeting on April 27, 1989 treated subject analysis with special reference to Prolegomena Classes. Mr. Langridge referred to his recent book on the subject analysis of documents. He stated that the overriding principle in library classification was 'helpfulness to readers'. Subject analysis should be completely separate from the

translation of terms resulting from an analysis into the symbols of a classification scheme. A first step must be to decide, what kind of work one was concerned with. He considers fundamental forms of knowing as necessary complements to fundamental categories of phenomena. The Prolegomena Classes should consist of fundamental prescriptive works on how to think, write and read, above which could be a class on the study of knowledge in books. But he thinks that there are no grounds for a comprehensive communication class. The discussions following in this meeting are very well reported, however, because of their many controversies a short summary would not be adequate. The meeting of July 13 was to be on "How to manipulate a classification schedule using a computer".

DGD-KTF again active

At the Deutscher Dokumentartag (the Annual Conference of the German Documentation Society), 4-6 Oct. 1989 at Bremen, a special session is scheduled for papers from the Committee on Classification and Thesaurus Research (DGD-KTF). The 3 papers to be presented are (in German): SCHMITZ-ESSER, W.: Thesaurus requirements from the point of view of new application fields. – LUKAS, E.: The thesaurus as part of the new information technology. – SCHMIDT, B., ENDRES-NIGGEMEYER, B.: Representation of a faceted classification in a frame system.

Colloquy on the Subject Authority File (of the German Library), 5-6 Oct. 1989

The Deutsche Bibliothek, jointly with the University Library Augsburg are sponsoring a two-day colloquy at Frankfurt on the Schlagwortnormdatei (SWD), (an authority file of subject headings) in order to inform interested colleagues and to discuss problems in application areas. The following four topics have been pinpointed for discussion: Cooperative handling of the SWD, SWD and OPAC applications, a more refined classification and SWD, SWD and authority files. For further information turn to Deutsche Bibliothek, Abt. Sacherschließung, Zeppelinallee 4-8, 6000 Frankfurt 1.