Indexing:
A Current-awareness Bibliography

Jean Wheeler

Part 2

AUTOMATIC INDEXING
(see also 372, 478, 493, 515, 553)


Describes the bar code data capture simultaneously while microfilming as one of the best new indexing techniques.


Focus is on attempts to improve the FASIT indexing system.


Analyses the indexing of collections of images in electronic form, as well as manual and computer-aided indexing with words; and automatic indexing by content.


Examines the development of research in the area of computer-assisted indexing, types of methods used, economic factors, and the slowness with which the technique has been adopted.


Describes the electronic indexing service, Archie, which is used for finding information on the Internet system.


Includes description of an automated method for categorizing files.


Recommends a new machine-readable file of authority records for topical subdivisions and for enhancements to the current file in view of the limitations of the Library of Congress subject authority records for subject heading assignment and validation.


Explains the principles and development of CAIN, a subsystem of the European Community's system for registering current agricultural research. Keywords and facet codes are being introduced to increase the quality of indexing, as well as to keep indexing costs down.

312 Goodier, John. CAIN, a computer assisted indexing system. Aslib Proceedings 44 no. 6 (June 1992): 251-252.


An overview of the current indexing environment is provided, including simple look-up files, sophisticated machine-assisted indexing environments, and automatic indexing without human intervention. Includes lists of contacts, commercial indexing software; expert system development, neural network, and hypertext environments; graphical user interfaces; and speech/voice recognition.


Advantages of changing to computer from index cards.

INDEXING: A CURRENT AWARENESS BIBLIOGRAPHY


Describes the MedIndEx System, which is being developed to assist users of the MeSH (Medical Subject Headings) thesaurus to index the MEDLINE database in MEDLARS (Medical Literature Analysis and Retrieval System) at the U.S. National Library of Medicine.


326 Savoy, Jacques. Stemming of French words based on grammatical categories. Journal of the American Society for Information Science 44 no. 1 (Jan. 1993): 1–9. Bibliography, Figures, Tables. Shows how inflectional suffixes can be removed from French words, describes a complex morphological analysis, and explains another stemming algorithm based on grammatical categories as well as a dictionary. The results of the study have been implemented, and the performance of this approach tested.


336 Wagner, M.M. An automatic indexing method for medical documents. In: Assessing the Value of Medical Informatics: Proceedings, Fifteenth Annual Symposium on Computer Applications in Medical Care, A
INDEXING: A CURRENT AWARENESS BIBLIOGRAPHY

ECONOMICS OF INDEXING
(see also 306)


Deals with pricing, timing constraints, submission of estimates, as well as problems related to the amount of work in question.

EDUCATION AND TRAINING
(see also 356)


Article followed by Comments by John Simkin.


Article followed by Comments (p. 22).

EVALUATION OF INDEXING
(see also 370, 378, 396, 398, 408, 442, 471, 474, 506, 513, 514)


Discusses evaluation of indexes; length of indexes; alphabetization; presentation of indexes; PRECIS; and user reactions.


Highlights studies that include in-depth examination of the indexes in question.


Discusses the comparative advantages and disadvantages of indexing the EMBASE, JMEDICINE, and MEDLINE databases.


Evaluates the index in *Amos: Victorian Methodist Traveller*, by John Matthews.


Provides examples of poor indexing in scholarly monographs.

INDEX PRODUCTION
(see also 359, 376, 390, 391)


Indexes in revised and repaginatated editions.


Discusses how an index created for Jeremy Wilson’s *Lawrence of Arabia* (1161 pages) was reduced for a concise edition (439 pages) of the same work.


Reprint of article which first appeared in *The Author* (June 1991).


The rationale, method, automation developments, and future prospects for the *Air University Library Index to Military Periodicals*.

INDEXERS
(see also 372)


Includes job descriptions; employment data; education and training; contacts and networks; a future outlook on jobs; as well as how to find a job.


Reviews international aspects of indexing: what indexers do; compilation of dictionaries; speech clichés; subject indexing; and job prospects.


Shows how indexers can communicate with each other through the online system, Prodigy.

The Indexer Vol. 18 No. 4 October 1993
INDEXES
(see also 348, 351, 394, 398, 408, 417)


The ARTFL (American and French Research on the Treasury of the French Language) is used as an example to illustrate the problems which arise when compressing a large textual database. New algorithms for compressing indices are described.


371 McKenzie, Elizabeth M.; McDermott, Margaret H. CD-ROM journal indexes: Wilsodisc compared to LegalTrac. Law Library Journal 84 no. 3 (Summer 1992): 439–459. The two leading journal indexes for legal research on CD-ROM are Wilsodisc and LegalTrac. This article compares the two along with LegalTracEF, a new version of LegalTrac. The authors conclude that the Wilsodisc software is the most sophisticated.

372 Milstead, Jessica L. No, you can’t be replaced by a computer. Keywords 1 no. 2 (Sept. 1992): 1–2.

373 Newbrough, C. The user-friendly index. Keywords 1 no. 1 (July 1992): 6–9. Examines the accessibility and marketability of a good index, as well as the theory of indexing.


381 Tomaiuolo, Nicholas G.; Trollely, Jacqueline. Science Citation Index/Science Citation Index with Abstracts on CD-ROM. CD-ROM Librarian 7 no. 9 (Oct. 1992): 34, 36–42. Figures.

382 Weaver, Alice O. The Columbia Granger’s World of Poetry. CD-ROM Librarian 7 no. 5 (May 1992): 36, 38, 40, 42, 44. Figures. Illustrates and discusses the indexing for the CD-ROM database.

INDEXING MANUALS


Provides rules and guidelines for ERIC subject indexing. Special attention is paid to coordinate indexing.


390 Indexes: a chapter from The Chicago Manual of Style, 14th edition. Chicago:

All aspects of preparing an index are covered, including general principles, typing the index, making page proofs, and computer usage.


Expands The Chicago Manual of Style chapter on indexing, and discusses the place of indexing in the publishing process, deciding on whether to hire a professional or compile an index on one's own, coping with the indexing guidelines of publishers, and choosing indexing software.

INDEXING NEWSPAPERS AND SERIAL PUBLICATIONS
(see also 347, 348, 355, 368, 369, 371, 400, 530, 540, 550)


Procedures and policies for indexing "The Forum," a daily newspaper, which serves eastern North Dakota and western Minnesota.


Deals with problems which arise when accessing gay and lesbian materials in connection with subject headings and the indexing of periodicals.


INDEXING SERVICES AND DATABASES
(see also 307, 347, 371, 394, 476)


Attempts to respond to such questions as how should the style and content of a subject index be designed; how should it be evaluated for retrieval purposes; how long should it be; and should it be compiled in-house or by a professional indexer. Twelve criteria for evaluating the quality of an index are listed.


Reviews indexing and content analysis.

INDEXING SPECIAL MATERIALS AND FORMATS
(see also 305, 322, 351, 352, 353, 367, 393, 403, 438, 455, 480, 505, 557)


Evaluates and provides an overview of indexing services and indexes that offer access to European Community information sources.


Aspects of indexing policy and database selection are discussed by representatives of indexing and abstracting companies.

400 Dodd, J.A. F & S Index plus Text on Compact Disc: added value for a familiar resource. CD-ROM Professional 5 no. 6 (Nov. 1992): 84-86, 88.

Describes the extensive indexing done for the Predicasts F & S Index series, which covers approximately 1,000 titles in the area of international newspapers, industry trade journals, and general business journals.


The distributed file system, Prospero, provides a framework that ties various indexing services together.


Describes two scientific literature indexing services, one of which was approved by the Germans, the other published illegally.


International Archives Institute, Inc. publishes its new Monographs in Education database, which compiles back-of-the-book indexes of selected monographs.


Deals with copyright, multi-institutional networked access to databases; and licensing local databases.

405 Tenopir, Carol. A day in the life of a database producer. Database 15 no. 3 (June 1992): 17-20, Illus.

Describes the indexing process based on procedures at the H.W. Wilson, Information Access, and UMI/Data Courier companies.

INDEXING SOFTWARE
(see also 328, 371, 379, 509, 511, 519, 529, 535, 552)


Concludes that ZYIndex is the best-selling microcomputer text retrieval package on the market.

INDEXING: A CURRENT AWARENESS BIBLIOGRAPHY


Discusses the problem of author-indexed works.

409 Bell, Hazel K. On the indexability of butterflies. LOGOS 3 no. 3 (1992): 149-152.

The advantages and disadvantages of indexing fiction.


Extracting a narrative work for indexing purposes.


Attempts to define a data model for organizing and managing a pictorial information retrieval system's indexing structure.

413 Cawkell, A.E. This is IT. London: Aslib, May 1992.

Includes a chapter on indexing and retrieving graphics and images.
Looks briefly at the Kodak microfilm indexing system at the Bank of Melbourne.


Focus is on the problems of distant definitions, statutory coding, name changes, and the missing name.


Abbreviated version of a paper which appeared in Stamp Lover (April 1991).


Discusses the use of records management indexes.


419 Grant, Rose. Cookbook indexing is not easy as ABC. American Society of Indexers Newsletter no. 98 (May/June 1990): 1, 4.


Considers item indexing; indexing documents in case files; assigning a catalogue or bibliographic reference to books, reports, and journal articles; and the subject indexing of correspondence.


422 Locke, Christopher. The dark side of DIP. Byte 16 no. 4 (Apr. 1991): 193–204. Illus. Argues that document image processing (DIP) tends to underestimate the complexity of indexing documents properly. Instead of an exercise requiring knowledge engineering skills, the indexing is treated as a low-order clerical task.


Discusses indexing as the dark side of document image processing.


Describes the Carnegie Mellon University Libraries’ in-house database, Archipics, which provides references to architectural drawings and photographs.


Basics of document image processing (DIP) with focus on CD-ROM, WORM, Hypertext, and indexing.


Deals with British versus American English: Anglicisms and Canadian versus standard French situations.


The difficulty of trade or brand name indexing in the area of drugs for major bibliographic databases.


Uses electronic schemas, architectural drawings, and topographic maps to illustrate the AMSTERDAM system for image indexing, as well as the storage and retrieval of paper documents based on pictorial or textual information. The architecture, algorithmic background, and main characteristics of the system are presented.


Discusses the coding of Landsat, weather, and medical imagery.


INDEXING SYSTEMS

(see also 303, 305, 330, 332, 345, 353, 369, 378, 407, 435, 465, 474)


SIGNPOSTS is an indexing/recording/cataloguening system for computer databases. It provides a matrix for creating a fully articulated index on any subject.


INDEXING TECHNIQUES

The Indexer Vol. 18 No. 4 October 1993

Discusses document indexing techniques and tools as well as image management systems.


446 Taylor, John. Index of progress: embedded entries. American Society of Indexers Newsletter no. 102 (Jan./Feb. 1991): 1, 20. Examines embedded indexing capabilities of publishing software, which would allow indexes to be developed with the book to which they refer.


INDEXING VOCABULARIES
(see also 477, 487, 550)


452 Gendina, Inna E. The Russian Rubricator: a unified system of classificatory indexing languages. International Classification 19 no. 3 (1992): 126–130. Presents an overview of the idea and purpose of the study, its components, the development of local rubricators, development and maintenance of the system, and its publications.


KEYWORD INDEXING
(see also 311, 474, 530)

461 Ballard, T.; Lifshin, A. Prediction of OPAC spelling
INDEXING: A CURRENT AWARENESS BIBLIOGRAPHY


Two methods for indexing databases are considered: automatic extraction of natural terms from texts, and assigning descriptors from a thesaurus.


Describes creation of an inverted index for the purpose of providing keyword-based access to a large text file.


The Encyclopedia of Computer Science (Iwanami Book Co.) is discussed, including its entry terms and the KWIC (Keyword in Context) index table.


Steps in the production of the KWIC index are discussed, along with KWIC index components, use of the index, the KWIC program, source listing, further improvements, and copyright policy.

LEGAL ASPECTS OF INDEXING
(see also 404)


NAME INDEXES


Suggests every-name indexing rules in the area of genealogical research.


Topics covered include pseudonyms, co-authorship, multiple authors, non-corporeal authors, and citation.

PRECIS
(see also 345, 477)


SUBJECT INDEXING


The involvement of the Library in document indexing, compilation of an index to international treaties and agreements, the Index to Nigeria, and book indexing.


Indexing consistency and access term selection are examined.


Study conducted by librarians at the State University System of Florida.


Discusses a study analysing MeSH (Medical Subject Headings) terms for indexing the literature of AIDS.


Keyword posting is found to be more effective in many cases than either vocabulary or thesaurus indexing.


The National Library of Medicine’s (NLM) Medical Subject Headings (MeSH), which pose challenges to users searching MEDLINE.


Discusses indexing for the RINGDOC pharmaceutical literature documentation service produced by Derwent Publications Ltd.


Queries were indexed under Library of Congress Subject Headings (LCSH), Dewey Decimal Classification (DDC), and PRECIS, with PRECIS showing the best performance.


Identifies indexing features of the National Library of Medicine’s Medical Subject Headings (MeSH), which pose challenges to users searching MEDLINE.


Discusses problems which arise when indexing archival and manuscript materials. Suggests categories that should be included in such indexing, and provides an overview of efforts at present to address this need.


The availability of specialized index sources in the area of women’s studies.


Attractive indexing features of the CD450 Music Library database.

483 Hersh, William R.; Hickam, David H. A comparison of two methods for indexing and retrieval from a full-text medical


Indexing and classification schemes are shown to rely on an understanding of “subject.”


Provides rules, guidelines, and examples for indexers and lexicographers working on the Thesaurus of ERIC Descriptors. (For Part 2 see item no. 507.)


Discusses the unrestrained use of the term, “syndrome.” and the huge volume of syndrome literature cluttering databases in the MEDLARS system. Also shows how “syndrome” has deteriorated as a specific and useful MeSH (Medical Subject Headings) term.


Bibliography. Tables.

Survey conducted at Montana State University Libraries in 1989.


Bibliography. Figures. Tables.

Papers were analysed bibliometrically to determine language: existence of abstract; abstract type; depth of indexing; and inclusion of author’s address. Works originally written in English are compared with those in Slavonic languages.


Examines controlled indexing versus natural language in searching for patent information.


Bibliography. Figures.


Bibliography. Figures. Tables.


Bibliography. Figures. Tables.


Bibliography.


Illus.


Figures.

Describes creation of the Leeds University Library’s BCMSV (Brotherton Collection Manuscript Verses) computer database. Indexing problems which arose in the course of the work are discussed with emphasis on subject indexing.


Bibliography. Figures. Tables. (In German)

Attempts to determine whether the Schlagwortnormdatei or SWD (German subject authority file) is a controlled indexing vocabulary that is based on the Regeln für den Schlagwortkatalog or RWSK (Rules for the alphabetical subject catalogues). Also illustrates terminology control in the SWD, and points out the reasons for the file’s shortcomings.


Bibliography. Figures. Tables. Deals with term frequency.


Use of Bayesian networks for structuring index terms.


End user and intermediary searches on the Physician Data Query (PDQ) system were observed.


Procedures for preparing the index.


Includes sections on subject cataloguing and indexing, as well as thesaurus and indexing approaches.


497 p.

Provides a definition of identifiers, their function and guidelines for their creation and use. Compares identifiers which appear in the Identifier Authority List with those in the ERIC Database. The alphabetical display, as well as identifier format, categories, and rules are also examined.

507 Wellar, Carolyn R.; Houston, Jim, eds. Vocabulary
INDEXING: A CURRENT AWARENESS BIBLIOGRAPHY


THESauri
(see also 310, 316, 319, 355, 388, 462, 472, 486, 497, 503)


515 Buchan, Ronald L. Quality indexing with computer-aided lexicography. Information Services and Use 12 no. 1 (1992): 77–84. Bibliography. Examples to illustrate computer-aided indexing activity were taken from projects conducted at the U.S. National Aeronautics and Space Administration (NASA).


524 Dimitrova, Ginka. Thesaurus semiotics and thesaurus lexicography. International Classification 19 no. 2 (1992): 86. Specific features of thesaurus semiotics are discussed, and problems of thesaurus lexicography pointed out.


INDEXING: A CURRENT AWARENESS BIBLIOGRAPHY

Consider the quality of thesauri and indexing as part of a study to determine how professional searchers use search terms.


Provides a definition of “user-friendliness” in library and information science systems, followed by a discussion of facet analysis, concept analysis, and concept organization.


The possibility of integrating natural language processing and text retrieval are explored.


The VIXIT system facilitates access to a human resource management database in the Quebec civil service. Indexing is conducted by way of the SATO (Système d’analyse de textes par ordinateur) software.


Describes the keyword and thesaurus schemes of the NIKKEI-TEC newspaper article service, as well as the frequency list of assigned keywords.

531 Heinz, G. Data transfer between CD-ROM databases (‘MEDLINE’) to own databases (‘LIDOS’), with the example of a databank on schizophrenia. II. Transfer of descriptors. Software Kuvier fuer Mediziner und Psychologen 4 no. 6 (1992): 198–203. (In German)

Describes the treatment and identification of descriptors during the transfer of large databases from CD-ROM, such as MEDLINE, to local databases (MeSH—Medical Subject Headings). The identification and transfer are based on the INDEX MEDICUS medical thesaurus.


Describes the initiation of a project to create a preliminary music thesaurus, by converting Library of Congress Subject Headings (LCSH) to data format for ARIS (Anderson Rowley Information Systems) thesaurus construction software.


A method of thesaurus construction used to classify ship operations.


Discusses the implementation of a clustering system to generate a thesaurus.


The CD-ROM version of the Robert dictionary was used for research carried out for a course at the Ecole de bibliothéconomie et des sciences de l’information at the Université de Montréal. The SATO (Système d’analyse de textes par ordinateur) was also used for the project.


Covers such areas of interest as the semi-automatic classification and structure of thesauri, as well as thesaurus construction.


Features of the controlling thesaurus are included.


Describes the development of a European Community expert system for retrieving information in the AGRICOLA, AGRIS, and CAB Abstracts databases by using the CAB Thesaurus and AGROVOC.


Discusses the design of a prototype illustrating the link between the U.S. National Library of Medicine’s Integrated Academic Information Management System (IAIMS) and the Unified Medical Language System (UMLS).


Unified Medical Language System Metathesaurus was prepared by the U.S. National Library of Medicine.
INDEXING: A CURRENT AWARENESS BIBLIOGRAPHY


546 Petersen, Toni. Art & Architecture Thesaurus, supplement 1. New York: Oxford University Press, 1992. 221 p. Includes new terms, deletions, and changes. Sources used to support the changes are listed. The supplement is to be used in conjunction with the three-volume Art & Architecture Thesaurus, published in 1990.

547 Rasch, M. THEOBALD—an EDP-based thesaurus. Nachrichten für Dokumentation 43 no. 1 (Jan.-Feb. 1992): 11-18. (In German) The function of THEOBALD is to complete the ARCHIBALD archiving system by controlling the terminology and index term hierarchy.


550 Rodriguez Vela, C. Documental languages in political information and current events databases. Revista Española de Documentación Científica 15 no. 1 (Jan.-Mar. 1992): 13-23. (In Spanish) Questions the appropriateness of applying a thesaurus to such sensitive areas as a daily and weekly current events press. Suggests that the combination of free-text as well as descriptors is most successful.


553 Smeaton, A.F. Progress in the application of natural language processing to information retrieval tasks. Computer Journal 35 no. 3 (June 1992): 268-278. Recent developments in automatic and semi-automatic text indexing, the generation of thesauri and text retrieval, as well as abstracting and summarization of text.


558 Thesaurus of archaeological site types. 2nd ed. Compiled by Roger Leech et al. London: Royal Commission on the Historical Monuments of England and English Heritage, 1992. 213 p. The alphabetical list, with over 300 terms, is preceded by an introductory section which reviews the purpose of the thesaurus, subject coverage, sources, structure and inter-relationships, vocabulary control, the classes, using the Thesaurus for indexing and retrieval, future development, updating and maintenance.


560 Walker, Alvin; Mulholland, Sarah N. The Thesaurus of Psychological Index Terms: a historical review. Behavioral and Social Sciences Librarian 11 no. 2 (1992): 39-57. Bibliography. Tables. Special challenges posed for thesaurus-based indexing are discussed in the context of computerized information retrieval. New features added to the sixth edition of the Thesaurus include a rotated alphabetical terms section, scope and posting notes, and an index term clusters section.
