

IX CONCLUSION

IX . 1 Overview

The main body of the review discussed experimental or operational approaches to the different activities involved in automatic indexing, namely document analysis, document description, searching and index language generation. It is evident that the approaches to both syntax and semantics are, linguistically speaking, not very sophisticated, particularly in operational systems. The methods being tried in question answering research are more ambitious, but their application to large scale document retrieval systems is still some way off.

The most striking feature of research in automatic indexing is how little there is of it. It is difficult to see it having any impact on the massive operational systems and services which are becoming established. It is true that there are not many basic ideas which have not been tested in some form or other, but the experiments performed in the last decade are simply an arbitrary choice from the very much larger set which could, and presumably should, be carried out. It is therefore particularly depressing that the experiments which have been done should be, in the main, so unrelated to one another, and that they should typically involve such small test collections and/or such defective evaluation.

A further difficulty is presented by the fact that so few direct comparisons between manual and automatic indexing have been made. Admittedly those that have agree in showing that automatic indexing is not markedly worse than manual, as appears in Section VIII. But it would be dangerous to build too much on such slight foundations.

Very broadly, findings for automatic indexing seem to parallel those for manual experiments: for example on the value of syntax. But automatic indexing in some cases offers opportunities which are not always readily available in manual systems: for instance in connection with weighting.

Specific conclusions are as follows:

1. Syntax

In general the syntactic techniques used either to identify important content bearing words or to provide relational information have not been particularly advanced ones. There is no evidence that they are particularly effective for either purpose.

2. Semantics

The statistical approaches used for extraction do not seem very useful; and those tested for classification generation are not obviously advantageous. But statistically based weighting does seem to benefit performance.

Some approaches to the treatment of other system components chiefly developed in automatic indexing research appear to be advantageous: for example, normalising matching coefficients which can be expected to generate a quite discriminating ranking.

When the results of all the experiments are taken together with operational practice, the specification for an automated system, simpler than some current retrieval systems, but possibly as effective as they are, or at least not conspicuously inferior, emerges. Some parts are provided manually for reasons of convenience rather than principle. As retrieval systems nowadays have names, this one may be called MODESTA.

1. input : abstracts (+titles).
2. analysis : selection of all words not stigmatised by a manual stop list (this may include some 'fluff' words as well as syntactic operators).
3. description : attachment of within-document frequency weights.
4. index language : vocabulary control through application of suffix stripping.
5. searching : exploitation of further vocabulary control through use of collection frequency based weights.

The system would usually operate with a normalising matching coefficient. Its performance might be improved by relevance feedback, and its operation made more convenient by document clustering. It would not be very expensive.

IX . 2 Recommendations

The specification of MODESTA just outlined is in some ways rather dank: she does not attract. But the all-singing, all-dancing systems so enthusiastically predicted in the early days of mechanised information retrieval are not readily achieved. It is more useful to think of genuine research problems than pie in the sky. In particular, it is evident that we are very far from understanding why many results are obtained. For example, why do statistical association techniques perform so variably? In my view, the real research need is for experiments which will provide us with explanations for performance results, and not merely results.

A sensible medium range programme would therefore be to carry out rather bigger, and better designed, experiments over a proper range of collections and test variables. Hopefully, these should give us some information about such factors as exhaustivity, for example, or term cooccurrence patterns, which would help us to understand not merely why automatic techniques behave in observed ways, but (ultimately) retrieval systems in general. A specific point is that vocabulary control in MODESTA is rather rudimentary, and it is doubtful whether these simple methods would be adequate for really large collections. A very worthwhile line would therefore be to investigate simple automatic control techniques for large vocabularies, given that more recognition of verbal relations would seem to be required.

B I B L I O G R A P H Y

The references are in alphabetical order of sole or lead author (see definition in the Preface), and in date order for each author.

Abbreviations:

I. S. & R. = Information Storage Retrieval

J. Doc. = Journal of Documentation

J. ASIS = Journal of the ASIS.

- 1972 Aitchison, J. and Gilchrist, A. Thesaurus Construction: A Practical Manual, London: Aslib, 1972
- 1970 Aitchison, T.M., Hall, A.M., Lavelle, K.H. and Tracy, J.M. Comparative Evaluation of Indexing Languages, Part II: Results, Project Inspec, Institute of Electrical Engineers, London, 1970
- 1971 Andrews, K. Dissertation for the Diploma in Computer Science, Computer Laboratory, University of Cambridge, 1971
- Armitage
- 1967 Armitage, J.E. and Lynch, M.F. "Articulation in the Generation of Subject Indexes by Computer", Journal of Chemical Documentation, 7, 1967, 170-178
- 1968 Armitage, J.E. and Lynch, M.F. "Some Structural Characteristics of Articulated Subject Indexes", I. S. & R., 4, 1968, 101-111
- 1970 Armitage, J.E., Lynch, M.F., Petrie, J.H. and Belton, M. "Experimental Use of a Program for Computer-Aided Subject-Index Production", I. S. & R., 6, 1970, 79-87
- Artandi
- 1969a Artandi, S. "Computer Indexing of Medical Articles", J. Doc., 25, 1969, 214-223
- 1969b Artandi, S. and Wolf, E.H. "The Effectiveness of Automatically-Generated Weights and Links in Mechanical Indexing", American Documentation, 20, 1969, 198-202
- 1968 Austin, C.J. MEDLARS 1963-1967, U.S.Department of Health Education and Welfare, Public Health Service, Washington D.C., 1968
- 1968 Avram, H.D. The MARC Pilot Project, Final Report, Information Systems Office, Library of Congress, 1968
- 1973 Barber, A.S., Barraclough, E.D. and Gray, W.A. "On-line Information Retrieval as a Scientists Tool", I. S. & R., 9, 1973, 429-440
- Barker
- 1972a Barker, F.H., Veal, D.C. and Wyatt, B.K. "Comparative Efficiency of Searching Titles, Abstracts and Index Terms in a Free-Text Data Base", J. Doc., 28, 1972, 22-36
- 1972b Barker, F.H., Veal, D.C. and Wyatt, B.K. "Towards Automatic Profile Construction", J. Doc., 28, 1972, 44-55

- 1972 Barlow, D.H. "Information Retrieval", The Computer Bulletin, 16, 1972, 250-256
- 1973 Bath University Library, Design of Information Systems in the Social Sciences Project, "Clustering of Journal Titles According to Citation Data", Working Paper No. 11, September 1973
- 1972 Bennett, J.L. "The User Interface in Interactive Systems" in Annual Review of Information Science and Technology (Ed. Cuadra), Vol. 7, Washington D.C.: ASIS, 1972
- 1967 Borko, H. (Ed.) Automated Language Processing, New York: John Wiley, 1967
- Borkowski
- 1970 Borkowski, C. et al. "Structure and Effectiveness of 'The Citation Identifier': An Operational Computer Program for Automatic Identification of Case Citations in Legal Literature", J. ASIS, 21, 1970, 8-15
- 1973 Borkowski, C. and Martin, J.S. "Structure, Costs and Benefits of LEXtractor, An Operational Computer Program for Automatic Extraction of Case Summaries and Dispositions from Court Decisions", Proceedings of the ASIS, Vol. 10, 1973, 21-22
- 1972 Borman, L. and Mittman, B. "Interactive Bibliographic Data Bases in an Academic Environment", J. ASIS, 23, 1972, 164-171
- 1973 Bousselet, P. "Le Système Titus", Nachrichten, Nouvelles, Notizie, 49, 1973, 157-163
- 1969 Brauen, T.L. "Document Vector Modification in On-line Information Retrieval Systems", Information and Storage Retrieval, Report No. ISR-17, Department of Computer Science, Cornell University, 1969.
- 1970 Cagan, C. "A Highly Associative Document Retrieval System", J. ASIS, 21, 1970, 330-337
- 1973 Campey, L. "A Survey of Index Generation Programs", I. S. & R., 9, 1973, 441-448
- Carroll
- 1969 Carroll, J.M. and Roelloffs, R. "Computer Selection of Keywords Using Word-Frequency Analysis", American Documentation, 20, 1969, 227-233
- 1970 Carroll, J.M. and DeBruyn, J.G. "On the Importance of Root-Stem Truncation in Word-Frequency Analysis", J. ASIS, 21, 1970, 368-369
- 1973 Carroll, J.M. and Tague, J.M. "Use of an Automatic Analyser in Preparation of SDI Profiles", J. ASIS, 24, 1973, 277-281
- 1965 Chomsky, N. Aspects of the Theory of Syntax, Cambridge, Mass.: MIT Press, 1965
- 1965 Clarke, D.C. and Wall, R.E. "An Economical Program for Limited Parsing of English", AFIPS Conference Proceedings, Vol.27, Pt.1, 1965, 307-316
- Cleverdon
- 1966 Cleverdon, C.W., Mills, J. and Keen, M. Factors Determining the Performance of Indexing Systems, 2 vols. College of Aeronautics, Cranfield, 1966
- 1970 Cleverdon, C.W. "An Investigation into a Suitable Mechanised Information Retrieval System at the Defence Operational Analysis Establishment", Cranfield Institute of Technology, 1970
- 1971 Clough, C.R. and Bramwell, K.M. "A Single Computer-Based System for Both Current Awareness and Retrospective Search: Operating Experience with Assassin", J. Doc., 27, 1971, 243-253

- 1973 Coates, E.J. "Some Properties of Relationships in the Structure of Indexing Languages", J. Doc., 29, 1973, 390-404
- 1970 Cooper, W.S. "The Potential Usefulness of Catalog Access Points other than Author, Title and Subject", J. ASIS, 21, 1970, 112-127
- Coyaud
- 1966 Coyaud, M. Introduction à l'Étude des Langues Documentaires, Paris: Klincksieck, 1966
- 1972 Coyaud, M. Linguistique et Documentation, Paris: Larousse, 1972
- Cuadra
- 1966- Cuadra, C.A. (Ed.) Annual Review of Information Science and Technology, Vols 1-, 1966-
- 1971 Cuadra, C.A. "On-line Systems: Promise and Pitfalls", J. ASIS, 22, 1971, 107-114
- 1965 Damerau, F.J. "An Experiment in Automatic Indexing", American Documentation, 16, 1965, 283-289
- 1969 Dattola, R.T. "Experiments with a Fast Algorithm for Automatic Classification", in Information Storage and Retrieval, Report No. ISR-16, Department of Computer Science, Cornell University, 1969; reprinted in Salton 1971a
- Dennis
- 1965 Dennis, S.F. "The Construction of a Thesaurus Automatically from a Sample of Text" in Statistical Association Methods for Mechanised Documentation (Ed. Stevens et al.), National Bureau of Standards, Washington D.C., 1965
- 1967 Dennis, S.F. "The Design and Testing of a Fully Automated Indexing-Searching System for Documents Consisting of Expository Text" in Information Retrieval - A Critical View (Ed. Schechter), Washington D.C.: Thompson, 1967
- 1966 Doyle, L.B. and Blankenship, D.A. "Technical Advances in Automatic Classification", Proceedings of the ADI, Vol.3, 1966, 63-71
- Earl
- 1970 Earl, L.L. "Experiments in Automatic Indexing and Extracting", I. S. & R., 6, 1970, 313-334
- 1972 Earl, L.L. "The Resolution of Syntactic Ambiguity in Automatic Language Processing", I. S. & R., 8, 1972, 277-308
- 1973 Earl, L.L. "Use of Word Government in Resolving Syntactic and Semantic Ambiguities", I. S. & R., 9, 1973, 639-664
- Edmundson
- 1961 Edmundson, H.P. and Wyllys, R.E. "Automatic Abstracting and Indexing - Survey and Recommendations", Communications of the ACM, 4, 1961, 226-234
- 1969 Edmundson, H.P. "New Methods in Automatic Extracting", Journal of the ACM, 16, 1969, 264-285
- Fangmeyer
- 1969 Fangmeyer, H. and Lustig, G. "The Euratom Indexing Project" in Information Processing 68 (Ed. Morrell), Amsterdam: North-Holland, 1969
- 1970 Fangmeyer, H. and Lustig, G. "Experiments with the CETIS Automatic Indexing System" in Handling of Nuclear Information, International Atomic Energy Agency, Vienna, 1970

Farradane

- 1967 Farradane, J. "Concept Organisation for Information Retrieval", I. S. & R., 3, 1967, 297-311
- 1973 Farradane, J., Russell, J.M. and Yates-Mercer, P.A. "Problems in Information Retrieval: Logical Jumps in the Expression of Information", I. S. & R., 9, 1973, 65-77
- 1973 Gardin, J.-C. "Document Analysis and Linguistic Theory", J. Doc., 29, 1973, 137-168
- 1972 Gerson, G.M. "A Computer-Aided Classification and Retrieval System for United States Patents", I. S. & R., 8, 1972, 137-150
- 1968 Gotlieb, C.C. and Kumar, S. "Semantic Clustering of Index Terms", Journal of the ACM, 15, 1968, 493-513
- 1973 Hansen, I.B. "CA Condensates as a Retrospective Search Tool. A Commentary", I. S. & R., 9, 1973, 201-205

Hillman

- 1968 Hillman, D.J. "Negotiation of Inquiries in an On-line Retrieval System", I. S. & R., 4, 1968, 219-238
- 1969 Hillman, D.J. and Kasarda, A.J. "The LEADER Retrieval System", AFIPS Conference Proceedings, Vol.34, 1969, 447-455
- 1973 Hillman, D.J. "Customised User Services via Interactions with LEADERMART", I. S. & R., 9, 1973, 587-596
- 1971 Hines, T.C. and Harris, J.L. "Columbia University School of Library Service System for Thesaurus Development and Maintenance", I. S. & R., 7, 1971, 39-50
- 1971 Hisinger, B. "A Multidisciplinary and Computerised SDI Service for Industry and Research - Practical Experience and Costs", I. S. & R., 7, 1971, 175-187
- 1973 Hoyle, W.C. "Automatic Indexing and Generation of Classification Systems by Algorithm", I. S. & R., 9, 1973, 233-242
- 1969 Ide, E. "Relevance Feedback in an Automatic Document Retrieval System", Information Storage and Retrieval, Report No.ISR-15, Department of Computer Science, Cornell University, 1969
- IJCAI = Proceedings of the International Joint Conference on Artificial Intelligence
- 1969 First, The Mitre Corporation, Bedford, Mass., 1969
- 1971 Second, The British Computer Society, London, 1971
- 1973 Third, Stanford University, Stanford, Calif., 1973
- 1966 Ivie, E.L. Search Procedures Based on Measures of Relatedness Between Documents, Thesis, Massachusetts Institute of Technology, 1966
- 1973 Jacquesson, A. and Schieber, W.D. "Term Association Analysis on a Large File of Bibliographic Data, Using a Highly-Controlled Indexing Vocabulary", I. S. & R., 9, 1973, 85-94
- 1971 Jardine, N. and Sibson, R. Mathematical Taxonomy, New York: John Wiley, 1971

Jones

- 1967a Curtice, R.M. and Jones, P.E. "Distributional Constraints and the Automatic Selection of an Indexing Vocabulary", Proceedings of the ADI., Vol.4, 1967, 152-156
- 1967b Jones, P.E., and Curtice, R.M. "A Framework for Comparing Term Association Measures", American Documentation, 18, 1967, 153-161

- 1968 Jones, P.E., Curtice, R.M., Giuliano, V.E. and Sherry, M.E. "Application of Statistical Association Techniques to the NASA Document Collection", A.D.Little Inc., Cambridge, Mass., 1968
- 1969 Curtice, R.M. and Jones, P.E. "An Operational Interactive Retrieval System", A.D.Little Inc., Cambridge, Mass., 1969
- 1963 Katz, J.J. and Fodor, J.A. "The Structure of a Semantic Theory", Language, 39, 1963, 170-210
- Keen
- 1967 Keen, E.M. Several chapters in Information and Storage Retrieval, Report No. ISR-13, Department of Computer Science, Cornell University, 1967
- 1972 Keen, E.M. and Digger, J.A. Report of an Information Science Index Languages Test, 2 vols. College of Librarianship Wales, Aberystwyth, 1972
- 1973 Keen, E.M. "The Aberystwyth Index Languages Test", J. Doc., 29, 1973, 1-35
- 1971 Kerchner, M.D. "Dynamic Document Processing in Clustered Collections", Information Storage and Retrieval, Report No. ISR-19, Department of Computer Science, Cornell University, 1971
- 1965 Kessler, M.M. "Comparison of the Results of Bibliographic Coupling and Analytic Subject Indexing", American Documentation, 16, 1965, 223-233
- 1973 Kimball, J.P. The Formal Theory of Grammar, Englewood Cliffs, N.J.: Prentice-Hall, 1973
- Klingbiel
- 1973a Klingbiel, P.H. "Machine-Aided Indexing of Technical Literature", I. S. & R., 9, 1973, 79-84
- 1973b Klingbiel, P.H. "A Technique for Machine-Aided Indexing", I. S. & R., 9, 1973, 477-494
- 1972 Krevett, B. and Griffith, B.C. "A Comparison of Several Zipf-Type Distributions in their Goodness of Fit to Language Data", J. ASIS, 23, 1972, 220-221
- Kuno
- 1965 Kuno, S. "The Predictive Analyser and a Path Elimination Technique", Communications of the ACM, 8, 1965, 453-462
- 1966 Kuno, S. "Automatic Syntactic Analysis" in Seminar on Computational Linguistics (Ed. Pratt et al.) U.S. Department of Health Education and Welfare, Washington D.C., 1966
- Lancaster
- 1968a Lancaster, F.W. Evaluation of the MEDLARS Demand Search Service, National Library of Medicine, Bethesda, Md., 1968
- 1968b Lancaster, F.W. Information Retrieval Systems: Characteristics, Testing and Evaluation, New York: John Wiley, 1968
- 1969 Lancaster, F.W. "MEDLARS: A Report on the Evaluation of its Operating Efficiency", American Documentation, 20, 1969, 119-142
- 1972a Lancaster, F.W. Vocabulary Control for Information Retrieval, Washington D.C.: Information Resources Press, 1972
- 1972b Lancaster, F.W., Rapport, R.L. and Kiffin Penry, J. "Evaluating the Effectiveness of an On-line, Natural Language Retrieval System", I. S. & R., 8, 1972, 223-245

- 1972 Lefever, M., Freedman, B. and Schultz, L. "Managing an Uncontrolled Vocabulary Ex Post Facto", J. ASIS, 23, 1972, 339-342
- Lesk
- 1968 Lesk, M.E. "Performance of Automatic Information Systems", I. S. & R., 4, 1968, 201-218
- 1969 Lesk, M.E. "Word-Word Associations in Document Retrieval Systems", American Documentation, 20, 1969, 27-38
- 1969 Litofsky, B. Utility of Automatic Classification Systems for Information Storage and Retrieval, Thesis, University of Pennsylvania, 1969
- Lovins
- 1968 Lovins, J.B. "Development of a Stemming Algorithm", Mechanical Translation, 11, 1968, 22-31
- 1971 Lovins, J.B. "Error Evaluation for Stemming Algorithms as Clustering Algorithms", J. ASIS, 22, 1971, 28-40
- 1968 Lyons, J. An Introduction to Theoretical Linguistics, Cambridge: Cambridge University Press, 1968
- 1967 Mathews, W.D. "The TIP Retrieval System at MIT" in Information Retrieval - A Critical View (Ed. Schechter), Washington, D.C.: Thompson, 1967
- 1970 Matthews, F.W. "Weighted Term Search and Strategy for Searching a Large File for Document Retrieval by Computer", Proceedings of the ASIS, Vol.7, 1970, 315-317
- Miller
- 1971a Miller, W.L. "The Efficiency of MEDLARS Titles for Retrieval", J. ASIS, 22, 1971, 318-321
- 1971b Miller, W.L. "A Probabilistic Search Strategy for MEDLARS", J. Doc., 27, 1971, 254-266
- Minker
- 1970a Augustson, J.G. and Minker, J. "An Analysis of Some Graph Theoretical Cluster Techniques", Journal of the ACM, 17, 1970, 571-588
- 1970b Augustson, J.G. and Minker, J. "Deriving Term Relations for a Corpus by Graph Theoretical Techniques", J. ASIS, 21, 1970, 101-111
- 1972 Minker, J., Wilson, G.A. and Zimmerman, B.H. "An Evaluation of Query Expansion by the Addition of Clustered Terms for a Document Retrieval System", I. S. & R., 8, 1972, 329-348
- 1973 Minker, J., Peltola, E. and Wilson, G.A. "Document Retrieval Experiments Using Cluster Analysis", J. ASIS, 24, 1973, 246-257
- 1968 Minsky, M. Semantic Information Processing, Cambridge, Mass.: MIT Press, 1968
- 1972 Montgomery, C.A. "Linguistics and Information Science", J. ASIS, 23, 1972, 195-219
- 1972 Moody, D.W. and Kays, O. "Development of the US Geological Survey Bibliographic System Using GIPSY", J. ASIS, 23, 1972, 39-47
- 1972 Murray, D.M. "Document Retrieval Based on Clustered Files", Information Storage and Retrieval, Report No. ISR-20, Department of Computer Science, Cornell University, 1973
- 1973 Myers, J.H. "Computers and the Searching of Law Texts in England and North America: A Review of the State of the Art", J. Doc., 29, 1973, 212-228
- 1971 Negus, A.E. and Hall, J.L. "Towards an Effective On-line Reference Retrieval System", I. S. & R., 7, 1971, 249-270

- 1970 Niblett, G.B.F. and Price, N.H. "Mechanised Searching of Acts of Parliament", I. S. & R., 6, 1970, 289-297
- 1973 Olive, G., Terry, J.E. and Datta, S. "Studies to Compare Retrieval Using Titles with that Using Index Terms", J. Doc., 29, 1973, 169-191
- 1968 Parker, E.B. "SPIRES (Stanford Public Information Retrieval System)", 1968 Annual Report, Institute for Communication Science, Stanford University, 1968
- 1973 Petrick, S.R. "Transformational Analysis" in Natural Language Processing (Ed. Rustin), New York: Algorithmics Press, 1973
- 1972 Prywes, N.S. and Smith, D.P. "Organisation of Information" in Annual Review of Information Science and Technology (Ed. Cuadra), Vol. 7, Washington D.C.: ASIS, 1972
- 1969 Reintjes, J.F. "System Characteristics of Intrex", AFIPS Conference Proceedings, Vol. 34, 1969, 457-460; see also adjoining papers.
- 1972 Rettenmayer, J.W. "File Ordering and Retrieval Cost", I. S. & R., 8, 1972, 79-93
- van Rijsbergen
- 1971 Jardine, N. and van Rijsbergen, C.J. "The Use of Hierarchic Clustering in Information Retrieval", I. S. & R., 7, 1971, 217-240
- 1972 van Rijsbergen, C.J. Automatic Information Structuring and Retrieval, Thesis, University of Cambridge, 1972
- 1973 van Rijsbergen, C.J. and Sparck Jones, K. "A Test for the Separation of Relevant and Non-relevant Documents in Experimental Retrieval Collections", J. Doc., 29, 1973, 251-257
- 1966 Rocchio, J.J. "Document Retrieval Systems - Optimisation and Evaluation", Information Storage and Retrieval, Report No. ISR-10, The Computation Laboratory, Harvard University, 1966
- 1970 Rolling, L.N. "Compilation of Thesauri for Use in Computer Systems", I. S. & R., 6, 1970, 341-350
- 1970 Rowlands, D.G. "The Unilever Research SDI System", I. S. & R., 6, 1970, 53-71
- Rush
- 1971 Rush, J.E., Salvador, R. and Zamora, A. "Automatic Abstracting and Indexing. II. Production of Indicative Abstracts by Application of Contextual Inference and Syntactic Coherence Criteria", J. ASIS, 22, 1971, 260-274
- 1973 Mathis, B.A., Rush, J.E. and Young, C.E. "Improvement of Automatic Abstracts by the Use of Structural Analysis", J. ASIS, 24, 1973, 101-109
- Sager
- 1972 Sager, N. "Syntactic Formatting of Science Information", AFIPS Conference Proceedings, Vol. 41, 1972, 791-800
- 1973a Sager, N. "The String Parser for Scientific Literature", in Natural Language Processing (Ed. Rustin), New York: Algorithmics Press, 1973
- 1973b Grishman, R., Sager, N., Raze, C. and Bookchin, B. "The Linguistic String Parser", AFIPS Conference Proceedings, Vol. 42, 1973, 427-434

Salton

- 1966- Salton, G. (Project Director), Information Storage and Retrieval, Reports Nos. ISR-11, 12, 13, 14, 16, 18, 21, Department of Computer Science, Cornell University, 1966-1972
- 1968a Salton, G. Automatic Information Organisation and Retrieval, New York: McGraw-Hill, 1968
- 1968b Salton, G. "Search Strategy and the Optimisation of Retrieval Effectiveness" in Mechanised Information Storage, Retrieval and Dissemination (Ed. Samuelson), Amsterdam: North-Holland, 1968
- 1968c Salton, G. and Lesk, M.E. "Computer Evaluation of Indexing and Text Processing", Journal of the ACM, 15, 1968, 8-36
- 1968d Lesk, M.E. and Salton, G. "Relevance Assessments and Retrieval System Evaluation", I. S. & R., 4, 1968, 343-359
- 1969a Salton, G. "A Comparison between Manual and Automatic Indexing Methods", American Documentation, 20, 1969, 61-71
- 1969b Salton, G. "Search and Retrieval Experiments in Real-Time Information Retrieval", Information Processing 68 (Ed. Morrell), Amsterdam: North-Holland, 1969
- 1969c Lesk, M.E. and Salton, G. "Interactive Search and Retrieval Methods Using Automatic Information Displays", AFIPS Conference Proceedings, Vol.34, 1969, 435-446
- 1969d Ide, E. and Salton, G. "User Controlled File Organisation and Search Strategies", Proceedings of the ASIS, 6, 1969, 183-191
- 1970a Salton, G. "Automatic Text Analysis", Science, 168, 1970, 335-343
- 1970b Salton, G. "Automatic Processing of Foreign Language Documents", J. ASIS, 21, 1970, 187-194
- 1970c Salton, G. "Evaluation Problems in Interactive Information Retrieval", I. S. & R., 6, 1970, 29-44
- 1971a Salton, G. (Ed.) The SMART Retrieval System: Experiments in Automatic Document Processing, Englewood Cliffs, N.J.: Prentice-Hall, 1971
- 1971b Salton, G. "Automatic Indexing Using Bibliographic Citations", J. Doc., 27, 1971, 98-110
- 1972a Salton, G. "Experiments in Automatic Thesaurus Construction for Information Retrieval", in Information Processing 71 (Ed. Freiman), Amsterdam: North-Holland, 1972
- 1972b Salton, G. "The 'Generality' Effect and the Retrieval Evaluation of Large Collections", J. ASIS, 23, 1972, 11-22
- 1972c Salton, G. "A New Comparison between Conventional Indexing (MEDLARS) and Automatic Text Processing (SMART)", J. ASIS, 23, 1972, 75-84
- 1972d Salton, G. "Dynamic Document Processing", Communications of the ACM, 15, 1972, 658-668
- 1973a Salton, G. "Automatic Processing of Current Affairs Queries", I. S. & R., 9, 1973, 165-180
- 1973b Salton, G. "Recent Studies in Automatic Text Analysis and Document Retrieval", Journal of the ACM, 20, 1973, 258-278
- 1973c Salton, G. and Yang, C.S. "On the Specification of Term Values in Automatic Indexing", J. Doc., 29, 1973, 351-372

- Saracevic
- 1968 Comparative Systems Laboratory, An Inquiry into Testing of Information Retrieval Systems, 3 vols. Western Reserve University, 1968
- 1971 Saracevic, T. "Selected Results from an Inquiry into Testing of Information Retrieval Systems", J. ASIS, 22, 1971, 126-139
- 1972 Scheffler, F., March, J. and Bernados, J. "An Experiment to Study the Use of Boolean NOT Logic to Improve the Precision of Selective Dissemination of Information", J. ASIS, 23, 1972, 58-65
- 1971 Schiminovich, S. "Automatic Classification and Retrieval of Documents by Means of a Bibliographic Pattern Discovery Algorithm", I. S. & R., 6, 1971, 417-435
- 1968 Schultz, C.K. (Ed.), H.P.Luhn: Pioneer of Information Science - Selected Works, London: Macmillan, 1968
- 1965 Sharp, J.R. Some Fundamentals of Information Retrieval, London: Andre Deutsch, 1965
- 1970 Simmons, R.F. "Natural Language Question-Answering Systems", Communications of the ACM, 13, 1970, 15-30
- 1963 Sokal, R.R. and Sneath, P.H.A. Principles of Numerical Taxonomy, London: Freeman, 1963
- 1969 Sommar, H.G. and Dennis, D.E. "A New Method of Weighted Term Searching with a Highly Structured Thesaurus", Proceedings of the ASIS, Vol.6, 1969, 193-198
- Sparck Jones
- 1964 Sparck Jones, K. Synonymy and Semantic Classification, Thesis, University of Cambridge, 1964
- 1967 Sparck Jones, K. and Jackson, D.M. "Current Approaches to Classification and Clump-Finding at the Cambridge Language Research Unit, The Computer Journal, 10, 1967, 29-37
- 1970a Sparck Jones, K. "Some Thoughts on Classification for Retrieval", J. Doc., 26, 1970, 89-101
- 1970b Sparck Jones, K. and Jackson, D.M. "The Use of Automatically-Obtained Keyword Classifications for Information Retrieval", I. S. & R., 5, 1970, 175-201
- 1971a Sparck Jones, K. Automatic Keyword Classification for Information Retrieval, London: Butterworths, 1971
- 1971b Sparck Jones, K. and Barber, E.O. "What Makes an Automatic Keyword Classification Effective?", J. ASIS, 22, 1971, 166-175
- 1971c Sparck Jones, K. "The Theory of Clumps" in Encyclopedia of Library and Information Science (Ed. Kent and Lancour), Vol.5, New York: Marcel Dekker, 1971
- 1972 Sparck Jones, K. "A Statistical Interpretation of Term Specificity and its Application in Retrieval", J. Doc., 28, 1972, 11-21
- 1973a Sparck Jones, K. and Kay, M. Linguistics and Information Science, New York: Academic Press, 1973
- 1973b Sparck Jones, K. "Index Term Frequency and Quartile Deletion" (Letter), J. ASIS, 24, 1973, 166-167
- 1973c Sparck Jones, K. "Collection Properties Influencing Automatic Term Classification Performance", I. S. & R., 9, 1973, 499-513

- 1973d Sparck Jones, K. "Does Indexing Exhaustivity Matter?", J. ASIS, 24, 1973, 313-316
- 1973e Sparck Jones, K. "Index Term Weighting", I. S. & R., 9, 1973, 619-633
- Stevens
- 1965a Stevens, M.E. Automatic Indexing: A State of the Art Report, Monograph 91, National Bureau of Standards, Washington D.C., 1965
- 1965b Stevens, M.E., Heilprin, L. and Giuliano, V.E. (Eds.) Statistical Association Methods for Mechanised Documentation, National Bureau of Standards, Washington, D.C., 1965
- Stone
- 1967a Stone, D.C. Word Statistics in the Generation of Semantic Tools for Information Retrieval Systems, Thesis, University of Pennsylvania, 1967
- 1967b Rubinoff, M. and Stone, D.C. "Semantic Tools in Information Retrieval", Proceedings of the ADI, Vol.4, 1967, 169-174
- 1968 Stone, D.C. and Rubinoff, M. "Statistical Generation of a Technical Vocabulary", American Documentation, 19, 1968, 411-412
- 1972 Svenonius, E. "An Experiment in Index Term Frequency", J. ASIS, 23, 1972, 109-121
- Syntol
- 1964 Cros, R.C., Gardin, J.C. and Lévy, F. L'Automatisation des Recherches Documentaires. Un Modèle Général: Le SYNTOL, Paris: Gauthier Villars, 1964; 2nd Edition, 1968
- 1965 Gardin, J.C. Syntol, Graduate School of Library Service, Rutgers, The State University, New Brunswick, N.J., 1965
- 1967 Coyaud, M. and Siot-Decauville, N. L'Analyse Automatique des Documents, The Hague: Mouton, 1967
- 1970 Bely, N., Borillo, A., Virbel, J. and Siot-Decauville, N. Procédures d'Analyse Sémantique Appliquées à la Documentation Scientifique, Paris: Gauthier Villars, 1970
- Tell
- 1970 Tell, B.V., Larsson, R. and Lindh, R. "Information Retrieval with the Abacus Program" in Handling of Nuclear Information, International Atomic Energy Agency, Vienna, 1970
- 1971 Tell, B.V. "Retrieval Efficiency from Titles and the Cost of Indexing", I. S. & R., 7, 1971, 241-243
- 1972 Thiel, L.H. and Heaps, H.S. "Program Design for Retrospective Searches on Large Data Bases", I. S. & R., 8, 1972, 1-20
- 1968 Thorne, J.P., Bratley, P. and Dewar, H. "The Syntactic Analysis of English by Machine" in Machine Intelligence 3 (Ed. Michie), New York: American Elsevier, 1968
- 1970 Vaswani, P.K.T. and Cameron, J.B. The National Physical Laboratory Experiments in Statistical Word Associations and their Use in Retrieval, National Physical Laboratory, Teddington, 1970
- 1974(?) Vernimb, C.O. "Abstracts on Microfiche for On-line Retrieval", New Developments in Storage, Retrieval and Dissemination of Aero-space Information, AGARD, Neuilly sur Seine, 1974(?)
- 1971 Vickery, B.C. "Structure and Function in Retrieval Languages", J. Doc., 27, 1971, 69-82

- 1973 Walker, D.E. "Automated Language Processing" in Annual Review of Information Science and Technology (Ed. Cuadra), Vol.8, Washington D.C.: ASIS, 1973
- 1969 Weiss, S.F. "Template Analysis and its Application to Natural Language Processing" in Information Storage and Retrieval, Report No. ISR-16, Department of Computer Science, Cornell University, 1969
- 1972 Wellisch, H. "A Flow Chart for Indexing with a Thesaurus", J. ASIS, 23, 1972, 185-194
- Wilks
- 1971 Wilks, Y.A. Grammar, Meaning, and the Machine Analysis of Language, London: Routledge, 1971
- 1973 Wilks, Y.A. "The Stanford Machine Translation Project" in Natural Language Processing (Ed. Rustin), New York: Algorithmics Press, 1973
- J. Williams
- 1968a Williams, J.H. "Computer Classification of Documents" in Mechanised Information Storage, Retrieval and Dissemination (Ed. Samuelson), Amsterdam: North-Holland, 1968
- 1968b Williams, J.H. and Perriens, M.P. "Automatic Full Text Indexing and Searching Systems", IBM Federal Systems Center, Gaithersburg, Md., 1968
- 1971 Williams, J.H. "Functions of a Man-Machine Interactive Information Retrieval System", J. ASIS, 22, 1971, 311-317
- 1972 Williams, M.E. "Experiences of IIT Research Institute in Operating a Computerised Retrieval System for Searching a variety of Data Bases", I. S. & R., 8, 1972, 57-78
- 1972 Winograd, T. Understanding Natural Language, Edinburgh: Edinburgh University Press, 1972
- 1971 Wolff-Terroine, M. and Rimbert, D. "Computer-Aided Automatic Generation of a Structured Documentary Language: Preliminary Study". J. Doc., 27, 1971, 111-124
- Woods
- 1970 Woods, W.A. "Transition Network Grammars for Natural Language Analysis", Communications of the ACM, 13, 1970, 591-606
- 1972 Woods, W.A., Kaplan, R. and Nash-Webber, B. The Lunar Sciences Natural Language Information System, Final Report, Bolt, Beranek and Newman, Inc., Cambridge, Mass., 1972
- 1973 Woods, W.A. "Progress in Natural Language Understanding - An Application to Lunar Geology", AFIPS Conference Proceedings, Vol.42, 1973, 441-450