Development of Indexing and Indexes

CHARLES L. BERNIER

San Diego, California 92109

Received December 10, 1984

The importance and benefits of indexes and indexing are so startling that discussion of their past, present, and future development seems useful for indexers as well as searchers. Better understanding of the construction, qualities, specifications, standards, structure, and use of indexes is valuable for it stimulates innovation and may help to eliminate defects, such as blocks of undifferentiated references, external guidance, omission of specifications and standards, and unjustified cross-references and notes. As a prediction, perhaps cooperative intelligence can be developed to decrease loss of relevant material, reading overload, number of literature searches, and need for the digestion of some literature discovered. It may come to provide tailor-made reviews and recommendations for back-up reading so that research can, far more completely than at present, be based on the entire record of the past and bring unimagined benefits to mankind and civilizations.

Construction, qualities, and use of past, present, and future indexes after millenia can still be made more obvious to indexers and searchers with great value to mankind and civilizations. For this reason, a discussion of index importance, construction, editing, quality, specifications, standards, structure, use, and benefits may be worthwhile. A few predictions also seem appropriate in this paper in the silver anniversary issue of the *Journal of Chemical Information and Computer Sciences*. Twenty-five years ago, my earlier paper,[1] "New Kinds of Indexes", was published in the first issue of the parent of this same journal. In that paper, I predicted the development of certain indexes, some of which are now among the many indexes today.

IMPORTANCE

Disappearance of mental and physical indexes would likely disintegrate our civilization in a few hours. Without mental indexes, we would be unable to keep track of things. Without physical indexes, there would be no street indexes to maps and no street signs that physically index streets. Telephone directories and bar and zip codes would vanish. Author, citation, classified, concept, correlative, HAIC, KWIC, KWOC, molecular-formula and -structure, name, notation, number, reaction, ring, subject, substance, substructure, title, topic, and word indexes would all disappear. Indexes for time, such as clocks and watches, would be gone. Card catalogs and concordances would certainly be missed. The absence of indexes that give access to abstracts, databases, extracts, translations, and other surrogates would instantly convert bound volumes, databases, documents, files, journals, magazines, offprints, patent specifications, preprints, published articles and papers, records, and reprints into waste paper and unused tape because these datum sources would be too slow to search page by page, item by item, even if the searcher could keep track of what he sought without mental indexes. Fortunately, outside of nuclear holocaust and science fiction, there seems little likelihood of an indexless world.

QUALITY

Indexes, like diamonds, have different qualities that are usually subtle.[2] Completeness and ease of use are major qualities of indexes. Accurate, specific titling of the kind of index is important. Access to indexes is favored by low price, wide distribution, and knowledge about inquiry, all aided by an acute perception of the magnitude of reading overload.[3] Correct entries preserve guidance. Promptness is a quality. Absence of errors and mistakes discovered during search improves the perception of the quality and reliability of the index and increases future use that saves time and necessary data.

Index quality can be measured and specifications[4] established. This generally has not been done for different reasons,

Charles L. Bernier, Ph.D., has had a long, productive career in information science in professional society, government, industry, and academia. His positions include the following: Editor, *Chemical Abstracts*; Director, Defense Documentation Center; Science Communication Specialist, National Library of Medicine; Senior Research Information Associate, Squibb Institute for Medical Research; full Professor with Tenure, Graduate School, State University of New York. He has consulted as information scientist for over 20 organizations. He is the author of about 75 publications, including books. Some of his travel on six continents has been in a professional capacity. He is listed in *American Men and Women of Science* and other biographical compilations, was twice president of the American Society of Indexers, and has been a member of numerous honorary societies, including The Cosmos Club, Washington, DC. Now emeritus, and active, he lives in beautiful San Diego, CA. He celebrated his golden wedding this June and asked if this picture including Mrs. Bernier could be used with this paper.

0095-2338/85/1625-0164$01.50/0 © 1985 American Chemical Society

one possibly because an index may have no competitors and so the searcher has but two choices: take it or leave it. Favorable specifications compared with established quality standards should enhance searcher confidence and decrease distracting doubts during search. Index specifications can include the following:

(1) Completeness of entries, expressed as a percentage, can be measured by checking, by duplicate (preferably replicate) indexing of a random sample of indexed material, and by complaints of searchers.

(2) Completeness of search, usually more important than is completeness of entries, and specified as a percentage, can be determined by comparison of search results among indexes covering about the same body of knowledge.

(3) Promptness can be measured by the delay in the appearance of the index after the availability of the material indexed and expressed as a function of the reciprocal of the delay.

(4) Ease of use can be expressed as a function of the reciprocal of search time as compared with search times through other indexes to somewhat related material.

(5) Freedom from errors and mistakes, expressed as percentages, can be measured by checking at least a sample of entries and from the complaints of searchers.

(6) Quality of internal guidance, measured crudely by the number of cross-references and notes and determined more accurately by examination and evaluation of the syndetic system, can reassure searchers of syndetic quality.

(7) Search control is increased in indexes with modifiers and justified cross-references and notes, as well as in indexes that select by correlation through the Boolean operators AND and NOT by the use of modifiers, links, and roles. Lack of control gives a sense of flying blind during search. The searcher is never sure of how many valuable, relevant references he may be missing. He is constantly aware of the possibility of invisible, catastrophic loss that might cost him his job. Search control can be measured by the percentages of adequate modifiers and by justified cross-references and notes, as well as by completeness of search.

(8) Justified guidance means the use of cross-references and notes that actually apply to the index being searched. It can be measured by the percentage of justified cross-references and notes. Unjustified (dangling, frustrated) guidance wastes search time and specifies lower quality of the index.

(9) External guidance requires search in at least two volumes instead of one and thus takes more time. If the searcher forgets, or opts not, to look in several volumes, he almost certainly will miss relevant material. The existence of external guidance can be specified to warn the searcher. "An index should stand on its own feet," E. J. Crane, when Editor of *Chemical Abstracts*, used to tell his staff.

(10) Scattering among identical or closely related entries can be measured by checking and expressed as a percentage. It is a major defect of poor indexes and is eliminated by editing.

Just as cutting costs in health care may, in some cases, lower its quality and not be in the best interests of patients, so cutting index-publication costs may not always be in the best interest of index searchers. Reducing costs, by decreasing checking, editing, internal guidance, justification of guidance, and proofreading, by decreasing type size, by delay in publication, by the use of subject indexers uneducated in the field indexed, by omitting specifications and standards, by resorting to word indexing, etc., can save money but certainly does burden searchers and drastically lowers the quality of the index.

It is more reliable to evaluate indexes by comparisons of measured specifications with quality standards[5] than by accepting the cognitive authority of publishers, just as it is more reliable to use standards and measured specifications to evaluate cameras, stereo equipment, motor vehicles, etc. than to rely on opinions of ad writers, manufacturers, owners, and sales personnel.

DEFINITION

Indexes are guides to, and not surrogates for, what is indexed. Just as nobody would use a street index in place of a map or a street, so nobody would use subject-index entries, for example, to an abstract or article as a substitute for the material indexed. Indexes are not condensed abstracts.[6]

STRUCTURE

Indexes are composed of entries, syndetic system,[7] and title. Entries consist of heading, modifier (modification, modifying phrase, expression, or term), and reference. The heading controls, to a major extent, the position of the entry in the index. A heading in a subject index, for example, represents one general subject area reported in the material indexed. In word and title indexes, the headings consist of words taken *verbatim* from the material indexed—often from title and subtitles. In author indexes, the name of the first (primary) author, is used as heading. Cross-references may be made in author indexes from the names of other authors to the name of the first author. Names are usually inverted, except in Islandic telephone directories. In name indexes, names of people, buildings, places, etc. in the text are used as headings. In number indexes, such as patent number indexes, the headings are the names of countries issuing patents and the kinds of patents issued. Patent numbers are the modifiers. In topic indexes, entries are made to topics used by authors in presenting their works. In classified indexes, the headings are the names of categories, classes, kinds, species, and types. All indexes have headings and references; some indexes lack modifiers.

The modifier modifies the heading much as an adjective modifies a noun. It delimits the scope of the heading. In subject indexes, for example, modifiers tailor the headings to match, as precisely as desired, the specific subject(s) reported by the author of the work indexed. Effective modifiers enable searchers to say "Yes, I want to look up this entry" or "No, I don't want to look up this entry" rather than to say "Maybe I want to look up this entry." Good modifiers eliminate most of the "maybes". Some searchers use modifiers to reject irrelevant entries; then they put the relevant entries and the maybes into ranked order before consulting the references. Even though modifiers are alphabetized under headings, it is nearly always impossible to predict their starting words. Because of this, it is necessary to read all modifiers under all headings searched in subject indexes to find all relevant entries. Indexes that omit modifiers and provide only blocks of undifferentiated reference numbers under headings put the responsibility for differentiating among undifferentiated reference numbers onto the shoulders of searchers. This practice seriously degrades the quality of the index by making it more difficult, discouraging, and slow to use.

References guide searchers to the material indexed. Incorrect reference numbers make it extremely difficult or impossible to locate the work indexed. The percentage of correct reference numbers, based on the total number of entries, can be calculated and constitutes another important index specification. References are usually numbers, such as page or column numbers. They may include the fraction of the page or column and also a code indicating "book", "patent", "review", etc. Fractions and codes improve the quality of the index by making it easier and faster to use.

The syndetic system of the index consists of cross-references, notes, instructions for index use, introduction, and perhaps a thesaurus.

Cross-references guide searchers among headings. They are rarely used to guide within headings because of the requirement to read all modifiers under every subject heading searched. "See" cross-references usually guide from synonyms and to combined headings, such as Plates and Trays. Rarely do they guide to more general terms because indexing is to the optimum specificity. Indexer generalization pushes the author beyond what he actually reports. Scattering of entries among synonymous headings is a serious index error. "See also" cross-references guide to terms related as antonyms, class terms, component names, species labels, and to miscellaneous kinds of relations. During search, cross-references are ignored at the risk of losing relevant data. Unjustified (dangling, frustrated) and circular guidance by cross-references are especially distracting. All cross-references should be justified. Unjustified cross-references external to the index may or may not be dangling. Cross-references and notes interfiled among the index entries help to make the index searchable without consultation of guides, instructions, introductions, and thesauri and thus improve the quality of the index by decreasing loss of entries, by speeding search, and by avoiding the frustration of incomplete guidance. Searchers may be remarkably insensitive to low index-quality factors, such as this.

Notes interfiled among the entries, or less desirably kept separately, provide highly specific instructions as to how a heading has been indexed and give search suggestions. Searchers, as well as indexers, should never remain in doubt as to how any relevant material has been indexed. Notes at appropriate locations in the index have been found to be a very convenient way to inform indexers as well as searchers of indexing policy. Earlier indexes (also separate instructions) guide indexers and decrease unnecessary variability from one index to the next. The time required to justify cross-references for an index introduces negligible percentage delay in the appearance of the index and greatly improves the quality of the index.

Guides and instructions carry general notes and directions affecting the use of the entire index, or large parts of it, by removing doubts and by indicating better ways to search.

Thesauri provide useful guidance among headings and related terms, such as antonyms and synonyms. A thesaurus used instead of justified cross-references degrades the quality of the index by introducing unjustified guidance and makes a second place to search. Vocabulary control is ultimately, of course, by authors and nomenclature authorities and not by indexers, index editors, and thesaurus builders. For example, when the words Penicillin, Plutonium, and Pseudomonas came into our language, the indexer could only use them or a synonym; he had no other options besides skipping them. He could not invent substitutes for these words because searchers could not find the substitutes. Authors and searchers would properly object if he ran "see" cross-references to his substitutes. Controlled vocabularies of preceding indexes or, less desirably, of thesauri are continually slowly obsolescing but are still useful in avoiding many distracting fluctuations in indexing from year to year for an abstract journal, for example.

Introduction can usefully give general data about the index and adjunct indexes.

KINDS OF INDEXES

The kind of index accurately specified by its title can save searchers the time required to look in irrelevant indexes. The most important kind of index, the subject index, guides searchers to specific subjects reported by authors. It is not to be confused with concordance, KWIC index, KWOC index, title index, topic index, concept index, or word index. Author indexes provide the most rapid and precise search of all if one knows the name of the author of the work sought. The names of some authors working in a field can be obtained readily from a citation index to the broad field.[8] Citation indexes were developed to guide searchers to works that cite a given author's publications. In effect, citation indexes use the references that authors have discovered as guides to specific subjects in a field. They can take a considerable load off searchers and, in some cases, provide more rapid access than through subject indexes. Correlative indexes select by the correlation of two or more terms from a thesaurus by Boolean operators, AND, OR, and NOT.[9-17,2,18] Names of indexes, if properly used, are self-explanatory.

Knowledge of the structure and kinds of indexes facilitates search. Acquaintance with indexes of a field precedes, or is concurrent with, search.

SEARCH

Inquiry[19] and *The Conduct of Inquiry*[20] can introduce "search". Introductions to indexes, as well as the literatures of index publishers, also provide useful data. Because scholarship, science, and progress generally build upon the past, knowledge of the past, which is nearly always provided by search, is essential if inadvertent waste by duplication of past work and by the failure to use it is to be avoided. Such knowledge is obtained by inquiry—search. Search precedes and is often concurrent with research to the benefit of both. Search strategy and tactics, which can include cut-off dates and interaction among several kinds of indexes, are improved by using every advantage; rigidity of search is disadvantageous and unnecessary. A careful, dated record of search (attempted and results) saves time and missed entries.

Solutions to the problem of reading overload[3] have yet to be much more widely implemented. It is impossible to avoid, except by chance, mistakes of the past if it is not possible to read as well as to remember the past. Indexes can now discover far more relevant material, even in fairly narrow fields, than can be read completely in the limited reading time available. One solution to overload is through condensed surrogates.[21-24] Another solution may be through artificial intelligence.[25]

INDEXING

The process of building an index is not obvious from the product. Myriad decisions made by indexers and index editors are not revealed by the index.[26] Indexing can be as simple as is using a computer to prepare a title concordance or as complex as is building a quality subject index. Because of its complexity, only subject indexing[27] will be discussed in the following paragraphs.

Subject indexers, to avoid mistakes and errors, need to know how to index as well as to know the basic subject matter of the field indexed. For example, "Steels for Automatics", might inadvertently be indexed under "Guns" whereas it should have been indexed under "Lathes" because automatic lathes require special steels for their tools that cut at high temperatures. This type of mistake can be caught through 100% checking of the index entries by another subject indexer. The requirement of subject knowledge makes subject indexing not a clerical operation. The subject indexer mentally paraphrases the subjects that he selects from the material, selects subject headings that guide to the paraphrased subjects, and provides modifiers that limit the headings to specific subjects indexed. References may be added to index entries by a transcriber if the indexing is dictated. One transcriber can handle the dictation of three or four indexers.[28] Dictation of indexing is more efficient than having the indexer himself write or type index cards. Cards (best 3″ × 5″ slips of twenty-pound bond paper to save space) can be used to facilitate alphabetizing, checking for errors and

mistakes, cumulating, printing, proof checking, proofreading, and typing. Indexing organizations that have computer facilities have eliminated typing and handling of index cards by transcribing the dictated index entries onto the computer discs or inputting directly such index entries on-line.

"See" cross-references from indexed synonyms to the subject heading used in the index are mandatory. Replicate indexing under all synonyms discovered during indexing has been found to be futile and also very poor indexing. "See also" cross-references help searchers to avoid missing index entries related as other than synonyms. Notes at subject headings that explain what is and what is not entered can be helpful for large headings, especially if interfiled among the entries. Notes can also explain the structure and arrangement of modifiers under appropriate large headings. Notes that apply to the entire index and to large parts of it are better placed in guides, instructions, and introduction. Multiple entries under the same heading for a given abstract are generally avoided to save searchers from being referred repeatedly to the same material; the modifier is coined broad enough to cover all aspects of the subject. This is usually not true for book indexing.

INDEX EDITING

Because indexers cannot keep track of all the modifiers they coin, synonymous and identical and closely related entries may be discovered scattered under a single heading after the index entries have been assembled. This is especially true if indexers in a group work together, as for an abstract journal. Scattering among headings occurs mainly because of synonyms, class and component relationships, and mistakes such as misspellings. Singleton entries are always suspect in a large index. Scattering can be decreased by editing the index entries before publication. The searcher should assume that all closely related entries have been assembled and all identical entries have been combined in a good index. Index editors who also index part of the time edit best. The editor eliminates scattering, combines identical entries, justifies the use of cross-references and notes (perhaps by means of an inverted cross-reference system),[29] crosses off unwanted subject headings or indicates to nontechnical personnel where this is to be done, makes indentures (indentations) of modifiers to avoid repeating the same starting word or term, adds modifiers to blocks of undifferentiated references, converts "see also" cross-references into "see" cross-references where appropriate, especially scrutinizes singletons, eliminates errors and mistakes, and, in general, readies the index entries for the printer (human or computer). He is also the best proofreader to pick up errors and mistakes before final printing. Computer-aided editing has not yet replaced the editing of index entries in large indexing organizations. While the computer programs check and correct, whenever feasible, index entries for correctness and consistency, the final editing is done by the editor with a computer terminal on-line.[30] This reference also provides an excellent introduction to current indexes for chemical literature. Some of these are illustrated on Figure 1. Additional historical material can be found in references 31–33.

PREDICTIONS

As artificial intelligence (AI) develops, it could provide efficient, elegant, and comprehensive solutions to many problems of indexing, index searching, and reading overload. If correlations among condensed surrogates (abstracts, extracts, terse literatures,[24] translations, etc.) and entries in indexes (author, citation, molecular formula, ring, subject, etc.), for example, can be provided automatically by AI, areas in science and humanities that require research and scholarship should be revealed. Perhaps solutions to some problems thought suitable for research may even be discovered in the literature by the AI system. Automatic correlations among substructures, therapeutic data, and toxicological data might turn out to be an illustration of an elementary way in which AI could benefit research. Comprehensive AI files would help to eliminate most the following: reading overload, literature search, and digestion of literature discovered. The researcher or scholar would simply read the highly condensed literatures (perhaps in the form of a tailor-made review) created by the AI system upon request and then read only the AI-recommended back-up material to enable him to understand, believe, remember, and use what he had read. He would then plan research projects and scholarly work knowing that he would not inadvertently be duplicating anything. The AI system, through its comprehensive files, would forget nothing and miss nothing relevant. Research and scholarship would then be based firmly on the *entire* record of the past. Mistakes of the past no longer need to be repeated.

Quality subject indexing by computer will become possible after computers have been programmed with the general subject knowledge of subject-expert indexers. This is so because the human indexer knows what are old and well-known data and messages and does not reindex them. He indexes only reported subjects that are new, researched, and studied. By way of contrast, concordances and word indexes to all nouns in an abstract, article, book, document, paper, patent, report, review, subtitle, title, etc. have most entries that are unrelated to the specific, novel subjects studied. Searchers may be unable to select only those entries that lead to novel subjects reported. As a result, they may be overwhelmed by a mass of entries that are not useful because they lead to material they already knew. Correlative indexes also may provide too many irrelevant entries for searchers to look up before discarding them. This is so if the precise specificity of modifiers, or less effectively of links and roles, is lacking.

Cogency for executives of accurate, relevant, reliable, and valid data and messages of engineering, scholarship, and science could be improved by developing condensed, semipopularized review literatures consisting only of data and messages judged by researchers and scholars themselves in their own fields to have recognized likelihood of accuracy and reliability. Semipopularization and condensation with minimum generalization would be goals. The fuzziness of generalization and the inaccuracy of popularization and condensation would be accepted by researchers, scholars, and executives as the price of increased cogency leading to useful decisions. Indexes to these semipopularized, condensed literatures would have subject headings and modifiers understandable by educated administrators who are in a position to translate reliable data and messages into appropriate action. Barriers of technical language and reading overload would be lowered by such literatures together with their indexes. These literatures would also help to guide administrative authorities to cognitive authorities who could consult and supply additional, valid data for fruitful decisions.

Indexes to cognitive authorities (personnel locators) have been developed and used in organizations, such as the Navy. A National index or better, a worldwide index, to cognitive authorities would facilitate their contributions to civilizations. The appropriate use of education, experience, and training of cognitive authorities should be enhanced greatly by a computer-based index to scientific and technical personnel for search by administrative authorities of business, governmental, industrial, and other organizations to assist in arriving at valid decisions. Civilizations are based on educated, experienced, trained individuals.

Personal files, in addition to being held in the forms of memory, 3″ × 5″ cards in boxes, steel filing drawers, note-

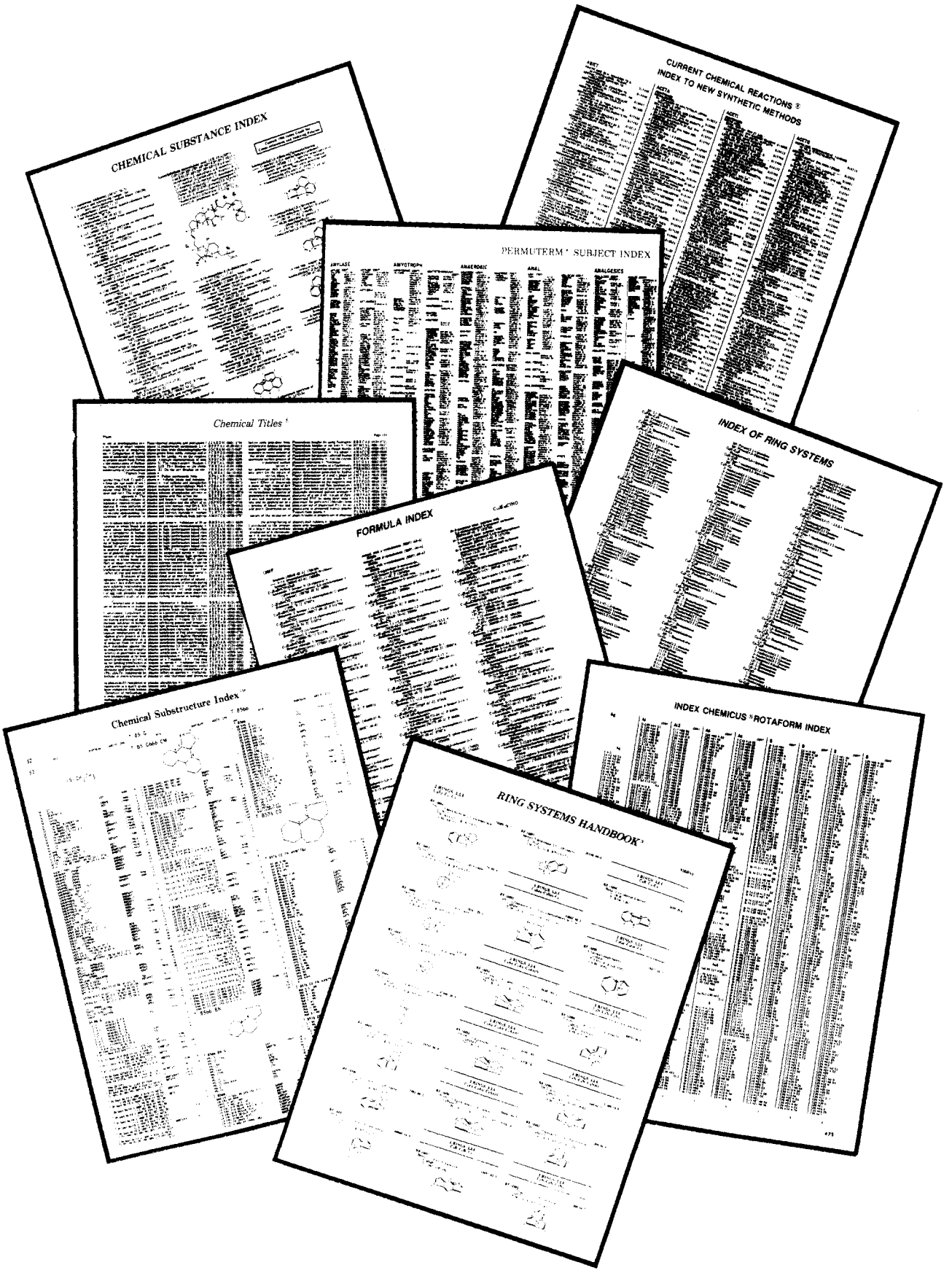

Figure 1. Current indexes for chemical literature. Pages from the *Chemical Substance Index*, *Chemical Titles*, *Formula Index*, *Index of Ring Systems*, and *Ring Systems Handbook* are reproduced with permission. Copyright owned by the American Chemical Society, Washington, DC. Pages from the *Chemical Substructure Index*, *Current Chemical Reactions*, *Rotaform Index*, and *Science Citation Index* (*Permuterm Subject Index*) are reproduced with permission. Copyright owned by the Institute for Scientific Information, Philadelphia, PA.

books, reference works, textbooks, offprints, journals, bound volumes, etc., could also be stored in the files of personal computers. These files, at present in the forms of discs, ROM and RAM memories, and tape, could include conclusions to which authors and owners have come. Many conclusions, gleaned from experience, literatures, and research, could temporarily be stored in memory (human) and on scraps of paper, backs of envelopes, 3″ × 5″ cards, notebooks, tapes of cassette recorders, etc. Most of these could be entered into personal computers. Conclusions, the pay-off of research, scholarship, and thinking, can be expressed tersely as sentences averaging 20 words and ranging from 6 to 46 words each. Such terse conclusions[24] are easily subject indexed, usually by the words they contain. Because the owner is both indexer and searcher, problems of indexing, syndetics, and vocabulary will be minimal, if they exist at all. The index entries could be for subject and author indexes, and possibly for molecular formula, substructure, classified, patent number, and other indexes as well. Index entries could be stored best in the owner's computer files.[23] Search could produce sets of terse conclusions or other terse literature related as the owner wishes. The printout could be arranged in any order desired. The product of this organization could closely resemble a tailor-made, specific minireview. Correlation and ideas obvious from reading the review could suggest further action. Such a computer file is indefinitely expandable and would forget nothing. Reviews could be had as often as requested and on subjects of current interest. The latest review would always be up to date.

A last prediction: There will be rapidly accelerating development of cooperative information systems through the use of cooperative intelligence. Such systems will include both producers and users of data and messages. In extremely complex civilizations such as ours, we realize better all the time that we most often survive by cooperation rather than by conflict with others. Many different specialists are now absolutely necessary to provide myriad products and services for our survival in peace with dignity. Cooperative intelligence means a significant extension of the current cooperation between human intelligences and artificial intelligences (for example, of decision-making machines, such as programmed computers). Artificial intelligence and human intelligence each will continue to contribute during their cooperation what each does best in order to achieve goals unattainable or too slowly attainable without such cooperation. However, with this extension of cooperation, more goals will become feasible and achieved than at present. Intelligence is revealed, measured, and defined by appropriate action. Accordingly, successful computer programs that achieve desired goals show intelligence by its definition. At present, human intelligences can supply databases vastly greater than can be supplied by artificial-intelligence systems, have the ability to handle flexibility and variety several orders of magnitude greater than can be supplied by artificial intelligence, and can supply superior creativity, indexing, ingenuity, inventiveness, and surrogation than can be supplied through artificial intelligence. The database used by these cooperative-intelligence systems will be highly condensed surrogates such as terse conclusions and datum signatures prepared in the same language by cooperating human intelligences of both the producers and users of research results. Thus, producers and users working in highly specialized fields that have been defined in scope by cooperative intelligence will support cooperative information systems, one for each field. These systems will track the individual interests of cooperating members in order to supply them with highly specialized, individual reviews and handbooks of data that omit nothing relevant and contain minimum material irrelevant to the interests of those receiving them. Their supply will be accomplished without subsequent request or search by users and will be updated automatically. As the interests of members change, the cooperative information systems will track these changes through the use of reading material requested and used by members, index entries read, citing material, cited material, specific requests by groups, etc. Thus, members of groups will be kept up to date and completely informed in the highly specific fields desired—and without search or further request. Such a cooperative information system of producers, users, and artificial intelligence will constitute a functioning ecosystem in the fields of research and development. This ecosystem will use cooperative intelligence and will be made far more productive than current systems by members being completely informed, by their being relieved of search, by the elimination of most digestion and recording of search results, and by the elimination of updating searches. In this last prediction, motivators for the prompt implementation of cooperative information systems include the following: (1) acute realization by employers that underinformed employees are too costly and are similar to underprogrammed computers; (2) realization by employers, engineers, researchers, scientists, et al. of the magnitude and the extremely harmful side effects of the massive reading overload that we all now experience; (3) knowledge that most often we survive by cooperation; cooperation among specialists will be greatly enhanced by cooperative information systems—along with other cooperative efforts; (4) cost, frustration, and waste of current information systems that unnecessarily divide producers from users of technical data will be eliminated; current systems require expensive, inefficient, and repeated searches by users, some of whom may be ill-trained for the task; (5) growing conviction that the progress of civilizations, cultures, societies, and their organizations rests on knowledge of current research results and that this knowledge is now incomplete because of reading overload; (6) perception of the urgent need to make complete the knowledge of authorities (administrative and cognitive), experts, professionals, specialists, and the like will increase and will cause beneficial promotion of cooperative information systems that will use cooperative intelligence.

The realization of this last prediction will usually obviate search and recording of the results of search and updates and will decrease reading overload and its disastrous side effects such as the functional obsolescence of employees and employers. Let us hope that these predictions will actually become underpredictions of the marvels of AI systems and of new indexes and of their elegant ways for organizing and presenting our magnificent treasure, all literatures including the chemical, for vastly improved use.

CONCLUSIONS

By creating and using new indexes, by consulting existing indexes, and by believing in valid research results revealed by indexes, civilized man and his organizations could benefit almost beyond imagination. Greater success than at present in finding, believing, remembering, and using more of the valid record of the past (chemistry included), usually located through indexes, could tremendously benefit our civilization by decreasing war,[34] accidents, aimlessness, boredom, cultural poisons,[35] ignorance, poverty, suicidal lifestyles, the arbitrary and unilateral control of others when it is harmful, bigotry, conquest, epidemics of caries and coronaries and lung cancer, ideologies based on fantasy worlds, illiteracy, infections, unnecessary loss of freedom, malnutrition, obesity, nonsense customs, lack of courtesy, sedentariness, spectatorship, starvation, subjugation, overevaluation of administrative and cognitive authority,[36] repetition of mistakes, and myriad other counterproductives. Opportunities today, perhaps through AI, for improved and far more comprehensive indexing in all fields

including chemistry are certainly remarkable.

REFERENCES AND NOTES

(1) Bernier, C. L. "New Kinds of Indexes". *J. Chem. Doc.* **1961**, *1* (1), 62-67.
(2) Bernier, C. L. "Correlative Indexes. X. Subject Index Qualities", *J. Chem. Doc.* **1964**, *4*, 104-107.
(3) Bernier, C. L.; Yerkey, A. N. "Cogent Communication: Overcoming Reading Overload"; Greenwood Press: Westport, CT, 1979; pp 3-17, 222-234.
(4) Bernier, C. L. "Index Specifications". *Indexer* **1974**, *9*, 9-12.
(5) Bernier, C. L.; Leondar, J. C. "Subject-Index Standards". *Proc. Am. Doc. Inst.* **1967**, *4*, 208-212.
(6) Bernier, C. L. "Surrogates". In "Encyclopedia of Library and Information Science"; Kent, A.; Lancour, H.; Daily, J. E.; Nasri, W. Z., Eds.; Marcel Dekker: New York, 1980; Vol. 29, pp 241-258.
(7) Bernier, C. L. "Syndetic Systems". In "Encyclopedia of Library and Information Science"; Kent, A.; Lancour, H.; Daily, J. E.; Nasri, W. Z., Eds.; Marcel Dekker: New York, 1980; Vol. 29, pp 334-357.
(8) Garfield, E. "Citation Indexes for Science". *Science (Washington, D.C.)* **1955**, *122*, 108-111.
(9) Bernier, C. L. "Correlative Indexes. I. Alphabetical Correlative Indexes". *Am. Doc.* **1956**, *7*, 283-288.
(10) Bernier, C. L. "Correlative Indexes. II. Correlative Trope Indexes". *Am. Doc.* **1957**, *8*, 48-50.
(11) Bernier, C. L.; Heumann, K. F. "Correlative Indexes. III. Semantic Relations Among Semantemes—The Technical Thesaurus". *Am. Doc.* **1957**, *8*, 211-220.
(12) Bernier, C. L. "Correlative Indexes IV: Correlative Chemical-Group Indexes". *Am. Doc.* **1957**, *8*, 306-313.
(13) Bernier, C. L. "Correlative Indexes V. The Blank Sort". *Am. Doc.* **1958**, *9*, 32-41.
(14) Bernier, C. L. "Correlative Indexes VI: Serendipity, Suggestiveness, and Display". *Am. Doc.* **1960**, *11*, 277-287.
(15) Bernier, C. L.; Dyson, G. M.; Friedman, H. J. "Correlative Indexes. VII. Trope Vocabularies and Trope Indexes for Chemistry". *J. Chem. Doc.* **1962**, *2*, 93-102.
(16) Bernier, C. L.; Crane, E. J. "Correlative Indexes. VIII. Subject-Indexing vs. Word-Indexing". *J. Chem. Doc.* **1962**, 2, 117-122.
(17) Bernier, C. L. "Correlative Indexes. IX. Vocabulary Control". *J. Chem. Doc.* **1964**, *4*, 99-103.
(18) Bernier, C. L. "Correlative Indexes". In "Encyclopedia of Library and Information Science"; Kent, A.; Lancour, H.; Nasri, W. Z., Eds.; Marcel Dekker: New York, 1971; Vol. 6, pp 189-205.
(19) Bernier, C. L. "Inquiry". In "Encyclopedia of Library and Information Science"; Kent, A.; Lancour, H.; Daily, J. E.; Nasri, W. Z., Eds.; Marcel Dekker: New York, 1974; Vol. 12, pp 1-6.
(20) Kaplan, A. "The Conduct of Inquiry. Methodology for Behavioral Science"; Chandler: Scranton, PA, 1964.
(21) Bernier, C. L. "Terse Literature: I. Terse Conclusions". *J. Am. Soc. Inf. Sci.* **1970**, *21*, 316-319.
(22) Bernier, C. L. "Terse Literatures: II. Ultraterse Literatures". *J. Chem. Inf. Comput. Sci.* **1975**, *15*, 189-192.
(23) Glazener, T. T.; Bernier, C. L. "Terse Literatures. III. The ARTEC Computer Program". *J. Chem. Inf. Comput. Sci.* **1978**, *18*, 16-17.
(24) *Bernier*, C. L. "Terse Literatures". In "Encyclopedia of Library and Information Science"; Kent, A.; Lancour, H.; Daily, J. E.; Nasri, W. Z., Eds.; Marcel Dekker: New York, 1980; Vol. 30, pp 312-330.
(25) Winston, P. H. "Artificial Intelligence", 2nd ed.; Addison-Wesley: Reading, MA, 1984.
(26) Bernier, C. L.; Crane, E. J. "Indexing Abstracts". *Ind. Eng. Chem.* **1948**, *40*, 725-730.
(27) Bernier, C. L. "Subject Indexes". In "Encyclopedia of Library and Information Science"; Kent, A.; Lancour, H.; Daily, J. E.; Nasri, W. Z., Eds.; Marcel Dekker: New York, 1980; Vol. 29, pp 191-205.
(28) Bernier, C. L.; Langham, C. C. "Dictating Machines as Indexing Aids". *Am. Doc.* **1955**, *6*, 237-238.
(29) Borko, H.; Bernier, C. L. "Indexing Concepts and Methods"; Academic Press: New York, 1978; pp 86-87.
(30) Metanomski, W. V. "Indexes in Chemical Literature—A Brief History". Presented before the Division of Chemical Information, 188th National Meeting of the American Chemical Society, Philadelphia, PA, Aug 30 1984.
(31) Baker, D. B.; Horiszny, J. W.; Metanomski, W. V. "History of Abstracting at Chemical Abstracts Service". *J. Chem. Inf. Comput. Sci.* **1980**, *20*, 193-201.
(32) Rowlett, R. J., Jr. "An Interpretation of Chemical Abstracts Service Indexing Policies". *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 152-154.
(33) Zaye, D. F.; Metanomski, W. V.; Beach, A. J. "A History of General Subject Indexing at Chemical Abstracts Service". *J. Chem. Inf. Comput. Sci.*, in press.
(34) Bernier, C. L. "Changing Behavior". *Chem. Eng. News* **1984**, *62* (29), 4, 38, 51.
(35) Patterson, A. M., former Editor of *Chemical Abstracts*, suggested the appropriate term "cultural poisons" to designate toxic substances, such as caffein, ethanol, street drugs, and tobacco, widely used in a culture.
(36) Bernier, C. L. Book review of P. Wilson's "Second-Hand Knowledge. An Inquiry into Cognitive Authority"; Greenwood Press: Westport, CT, 1983; *J. Am. Soc. Inf. Sci.* **1984**, *35*, 255-256.

History of Citation Indexes for Chemistry: A Brief Review

EUGENE GARFIELD

Institute for Scientific Information, Philadelphia, Pennsylvania 19104

Received January 11, 1985

The *Science Citation Index* (*SCI*) was the first comprehensive citation index for chemistry. But its use in chemistry was not obvious, even though *SCI* covers every important journal of chemistry. However, citation-based searching bypasses dependence on chemical nomenclature. Finding applications of synthetic methods and physical-chemical equations is simplified. These are fields where use of traditional indexing is difficult. An extension of citation indexing, co-citation clustering, is now also used for automatic hierarchical classification and mapping of literature. The value of citation indexes to the historian of chemistry will continue to increase as *SCI* coverage is extended back to include the pre-1955 literature.

This year marks the 25th anniversary of the *Journal of Chemical Information and Computer Sciences*. Appropriately, the journal's name has changed since it was founded in 1961 as the *Journal of Chemical Documentation*. When I contributed a paper to the first issue,[1,2] the field seemed to be moving quite slowly. In retrospect, it was difficult to imagine how rapidly problems would be solved that then seemed insolvable.

Over the past 25 years the scope of the journal has broadened as the field of chemical documentation has advanced. This explains why the journal's title was changed in 1975 to reflect the growth of the information and computer sciences. Of the many advances made in chemical information retrieval over the past 25 years, I have been asked to review the history of citation indexes for chemistry.

It is sometimes difficult to discuss problems that are specific to chemical information because they cannot really be separated from the broader problems of scientific information. Modern chemistry is unavoidably multidisciplinary. *Chemical Abstracts* (*CA*), which calls itself a chemical information service, is used in medicine, engineering, and other disciplines, as well as in chemistry. It can be argued that the first multidisciplinary *Science Citation Index* (*SCI*), which covered the literature of 1961, was the first comprehensive citation index

0095-2338/85/1625-0170$01.50/0 © 1985 American Chemical Society