EB article on Information Processing

"The information revolution comprises the tremendous technological advances made during the past few centuries in human capabilities to encode, record, reproduce, and disseminate information....The electronic computer, together with its peripheral equipment, provides electromechanical capability for modifying and reprocessing stored information to produce vast new stores of information."

"The information revolution produced a knowledge explosion."

"All of these technological advances together have made information a new basic resource that supplements the familiar natural resources of matter and energy. Accumulated world knowledge takes on an entirely new meaning and significance as techniques for mining, storing, sharing and using information in new ways are learned. Knowledge in the form of newly stored information is not degraded or destroyed with use."

distinction between primary and secondary forms of information- primary forms are those forms in which the records are first disseminated; and secondary forms are those which result from their further analysis description, and synthesis

"Primary media can be classified according to the physical forms in which the information appears (that is, as printed, auditory, or visual medai) and again, according to their format (as textbooks, monographs, periodicals or journals, or as report literature or reprints).... Secondary media - as the size of a literature grows, the greater becomes the nece-sity for developing techniques for searching it (refers mainly to abstracts and indexes)"

Indexing

"Thus an index serves to guide its user to the information he seeks and at the same time usually provides some basis for screening or selection before the page or document containing the information is itself sought. Principally an index directs the inquirer to a specific piece of information, but it may also make known to him related items of information brought together by the proximity of their headings. In general, however, items related by subject may appear at points widely scattered in an alphabetic listing. A good index should in that case carry cross references from each one to each of the others and so minimize the disadvantages of such scattering. With the object of bringing together more effectively related material, some indexes are ordered according to a classificatory scheme rather than alphabetically. A classified index is in general more difficult to use than is an alphabetic index. In principle, a well designed index can incorporate the advantages of both types of arrangement."

"The basic hypothesis of Uniterm coordinate indexing is that, for practical purposes of finding specific information, one can dispense with the recording of relationships among words and index a document with a set of individual, relatively brief terms. In the search process, a document, in order to retrieved, must contain some specified co-occurrence of index terms -- whether or not such terms bear the intended relationship to one another.
Thus a document indexed 'hard,' 'soldering,' 'aluminum,' and 'alloys' will be retrieved in response to a request for all documents described by any one, any two, any three, or all four of those terms. To illustrate the type of failure attributable to lack of relationships, that document would be retrieved in a search based on 'hard' and 'alloys,' though it may be quite irrelevant if the inquirer's intent is to find information on 'hard alloys.' The fact that irrelevant retrieval can in principle occur through the accidental coordination of unrelated terms should not obscure the pragmatic question of whether this happens to an excessive extent.
Any system of indexing and retrieval can be presumed to lead to some amount of irrelevant information, and the inquirer himself must finally judge the usefulness of what is retrieved and screen out material which does not serve his purpose."