CHAPTER 5

ISSUES AND PATTERNS FOR CLASSIFICATION AND INDEXING

Information explosion, has paved the way for tremendous acceleration in the collection, reproduction and storage of data and documents in all fields of human knowledge. Rapidly increasing investments in scholarship and research have produced enormous quantity of reports, articles, papers and other information carriers and created unprecedented problems for librarians and documentalists. Scientists, scholars and researchers are experiencing this considerably in the field of Natural Sciences a great variety of efforts have been made in recent years to cope with this overflow of information. Mechanical and electronic devices for storing and retrieving coded data are under development in a number of research centres and libraries and increasing attention is given to the logic and efficiency of the classification schemes. Where as in the field of Social Sciences very little has been done to sort out the problems raised through the rapidly increasing production of data and documents throughout the world (1). UNESCO's International Committee on Social Sciences Documentation has made valiant efforts to control the increasing world literature in the social sciences through its annual bibliographies. But these only
constitute a first step in the direction. The next step is, to ensure extensive depth classification and indexing. This means essentially to move inside the particular 'information carrier' and to classify its 'hard contents', be it a description of an experimental technique, an account of a particular case, a set of tabulated results, a sequence of calculations or some discussion of analysis procedures. To develop a logically coherent yet efficient classification scheme for such purposes is a gigantic task. The difficulties are obvious. The imprecision of social science vocabularies, the proliferation of specialised terminologies, the marked cultural and ideological contrasts in basic approaches to research. But the challenge is there and all efforts to tackle these problems should be encouraged.

It is only recently that a great deal of serious work has been done by librarians on the symbiosis of the two, that is classification and indexing. But it has been established beyond doubt that this combination may make it possible for information to be organised for the use of advanced research, as well as the comparatively simpler but much more widely known level of arranging books in a classified order on library shelves.
Library classification is one of the techniques of the librarian designed to achieve efficient library service through the organisation of the universe of knowledge embodied in documents, in a helpful sequence(2). To be useful, classification must keep up with the changes in the social pressure as may be sensed from the quality of the readers and the kinds of documents to be organised for service to them. It involves continuous work at several levels at the seminal and research levels (i.e., intuitive and intellectual level) and at the pragmatic level of classifying.

Indexing may perhaps be described briefly as the putting down, from day to day, of layers of information as it is published, each layer being divided into a series of separate (through usually related) subject sections. The accumulation of successive layers, each divided according to the same pattern, will in time build up a vein or lode of information for each subject in the pattern. The characteristic means of putting down the layers is the index card or slip; each day produces a pack of cards, for documents covering a range of subjects. By filing the cards according to a pre-arranged systems, one produces in time a series of packs each relating to documents and the same subject.
5.1 Terminology in Social Science

Terminology (3) is the aggregate of terms representing a system of concepts of a particular subject field. Terminology (4) is the theory of terminology, which is an interdisciplinary field comprising Linguistics, Logic, Information Sciences and individual subject fields.

Terminology is fundamental for information: on the one hand it serves toward the information, on the other hand terms are used for indexing, storing and retrieving the information. Thus specialised vocabularies, called thesauri, are prepared as documentation languages. Terminological work can be divided into three areas, namely:

1. Establishing terminological and lexicographical principles;
2. Preparation of terminologies in particular subject fields
3. Terminological documentation.

One should distinguish between lexicographical and terminological work. Lexicographical work mostly consists of collecting terms of a particular subject field and to present them as lists of terms, translation vocabularies, etc. Terminological work, however, comprises investigation into relationship between concepts (genus/species, part-whole and
so on), construction of the systems of concepts, and the definition of concepts. This work is preferably to be the responsibility of commissions of experts working in the subject field concerned. These systems of concepts are also the basis for the preparation of structured documentation languages, that is thesauri.

The objective of standardization of terminology is to enable unambiguous communication by standardizing concepts, systems of concepts, terms and definitions by competent terminological commissions of experts. Terminological Commission of National Standards Institutions examine terminologies existing so far and prepare the standardised terminology for a certain subject field on the international level. In October 1971 the International Information Centre for Terminology (Infoterm) was established at the Austrian Standards Institute in Vienna with the assistance of UNESCO within the frame work of UNISIST. Terminological data are stored in a computer in order to retrieve quickly not only any information on a given concept but also to produce quickly upon inquiry different types of lists of terms which may be needed for special purposes. Such terminological data banks exist in Canada, France, Germany, Sweden, USSR. Some
of them collect standardized terminology (for instance ISO, France, Germany, USSR).

5.1 Problems of terminology in Social Sciences

There have surprisingly been a few attempts at systematic organization of the varying usage of concepts in the vocabularies of subject fields in the Social Sciences, and in particular Politics, International Relations and Law. The major attempt in Political Science has been that of Kaplan and Lasswell in their 'power and Society' (5), the result of collaboration between a philosopher and a political scientist—an appropriate example of fruitful cooperation between a subject specialist and a student of the philosophy of language and of scientific methodology. It is an ambitious effort at formulating the basic concepts and hypothesis of Political Science— in terms of the empirical propositions of Political Science but excluding the value judgements of Political doctrine—as the study of the shaping and sharing of "power". The authors are thus attempting to work out neutral definitions corresponding with practices in the natural sciences.
The UNESCO has been much exercised about the question of social science terminology and has attempted to look the perennial problem in its face. The second meeting of experts on social science terminology called under its auspices in 1956 recommended "the preparation of interdisciplinary dictionaries of general social science terminology of the Lalande type", that is, something midway between a general dictionary and a specialized encyclopedia emphasis being placed on the conceptual aspect of words rather than on their semantic evolution or historical content; those dictionaries should give definition of the most key notions currently employed in the various branches of social science, regarded as empirical science of social fact (6). The number of concepts to be defined was estimated at 1,000 to 1,200 words. Etymology, semantic evolution, command and scientific usage (with special reference to related or similar terms, synonymus and homonymas) or concepts were to be covered for the purpose. The model for this project is provided in definitions of concepts such as "imperialism", "nationalism" and "power" worked out in International Social Science Bulletin (7).

Dictionnaire de la terminologie du droit international (8) (Union Academic International) does for international law
what Stroud’s Judicial dictionary does for English law. The book examines seven hundred concepts and phrases used in international law. It gives a definition or explanation of each word or phrase and also of any alternative or secondary meaning which it may have, with supporting passages from cases, conference proceedings, and the works of distinguished authorities. But the most ambitious attempt of all is that by centre international de synthese to produce a vocabulaire historique designed to define the terms used by historians and containing the following information:

a) History: Etymology. The history of the word (different meanings, successive or simultaneous). The history of the idea (noting the synonymy and words of similar meanings).

b) Proposed usage: A critical study of the different meanings, differentiation between essential and accidental meanings, and recommendation of the acceptance of the 'most constant and most logical meaning' (9). The project initiated in 1926 would take many more years for its completion.

These attempts in bringing together the varying usage should lead to the standardization of terminology in the social sciences in the long run.
5.2 Classification Systems and Social Sciences

The purpose of a general classification scheme is to enable libraries to arrange their documents and ideas in a sequence that will be as helpful as possible to the user. One of the main criticisms directed against librarians by social scientists has been that the schemes found most widely do not, in fact, produce a helpful arrangement. Very often they scatter material that needs to be used together and no infrequently they abolish subjects completely. For this, the remedy lies not in abandoning the techniques of classification and indexing, but in abandoning those parts of the techniques that have outlined their usefulness, and now constitute a hindrance, to library users. For example, many people are engaged in the study and teaching of Health Education but this subject has no existence in the Decimal Classification, which is used by most education libraries.

An attempt was made to consider the major general classification schemes in the light of the factors we have established as being of importance from the reader’s viewpoint, how social science subjects are collected, the adequacy of the terminology, the extent to which there is compartmentalising into traditional ‘disciplines’ and the
provision or lack of suitable categories of terms to enable the classifier to coordinate concepts as formed in the literature.

5.2.1 The Decimal Classification

The Decimal Classification of Melvil Dewey was the first of the modern general schemes and has become widely known through its use by public and school libraries. Produced in 1876 for the library of Amherst College, the great contribution made by the DC was its principle of 'relative location'- that is, the numbering of subjects, instead of 'fixing location', the numbering of positions on shelves. Relative location allowed subjects to be moved about on the shelves as new books were added. Dewey's scheme also enabled new subjects to be inserted at the appropriate places, which meant that the numbering system had also to allow of such interpolation, and he hit upon the idea of using decimal numbers for his scheme's notation. He divided the universe of knowledge into nine fields which he numbered from 1 to 9 and made a notation for the sub-divisions of each field by simply adding the necessary figures on the right-hand end. Using a minimum of three figures, his nine fields 'main classes' and their sub-divisions were numbered 100-999, and to these he
added a tenth class, numbered 000-999, for general and polygraphic works that could not be fitted into any particular main class. Its early years coincided with the growth of the public library systems of Europe and the USA, and many new libraries adopted the scheme. Its editorial body has now been established in an published in 1971. 19th edition in 1975 and now 20th edition is published. 19th edition was in three volumes whereas the 20th edition is in four volumes, with a greatly expanded introduction. Even though an attempt has been made to bring it up to date, one can still find some confusion, cross-classification and entire absence of important subjects.

The outline of the scheme corresponds to an inversion of the three great classes of knowledge proposed by Francis Bacon: History, Philosophy. A mid-nineteenth century scheme made by W.T. Harris was the DC's immediate parent, but there is, of course, a long tradition of Baconian influence on classifications of knowledge.

Dewey's treatment brings the social sciences into the fourth place in his sequence, at class 300, and as a result of following Harris and Bacon he separates social science from philosophy and psychology on the one hand (in class
100), and from History and Geography on the other (in class 900). Even those scholars who regard History as among the Humanities rather than the social sciences would hardly expect to find, separating the two, the whole of linguistics, purse and Applied Sciences, Fine arts and Literature. The treatment of Psychology in DC has a very curious history and its arrangement will give critical position to the users.

Like this we can find many examples, on feature that appears regularly in DC is the use of sub-divisions beginning at 0 for general matters, such as the Theory and History of a subject. Thus, 300 is used for Social Science-General Works, but since the 15th edition the place 301 has been allotted to sociology. Class 301 itself is subdivided primarily into a series of groups and inter-relations:

301.1 Social Psychology
301.2 Culture and Cultural Processes
301.3 Ecology and Community
301.4 Social Structure.

Public Administration at 350 is separated, by 360 social welfare and 370 education from such public services as Communication and Transport in 380; Demography is at 301.32, but Statistics of Populations (Demographic Statistics) is at
312;390 Customs and Folklore is at the other end of the class from 301, and far from Social Anthropology, which is away in the science class at 572. Slavery at 326 is regarded as a branch of Political Science, and stands next to 327 Foreign Relation; Forced Labour, on the other hand, is an Economic activity, at 331.5 class 330 is for Economics; but Commerce is provided at 381; Marxism at 335.4 appears as a branch of Economics (with other Economic ideologies), with the subdivision 335.411 for Philosophic foundations, and including Dialectical Materialism. Yet at 146.3 Materialism in Philosophy, one can still find the note, 'including dialectical materialism'.

DC is a so-called 'Logical' system; that is to say that its structure is based on the traditional logical method of division by beginning with a Main Class or genus and resolving it into a number of species by the addition, one at a time, of different characteristics. The decimal notation fits this process admirably. But a classification scheme for the literature of a complex subject needs to be able to do more than cater, as it were, for mere descriptions of objects. A classification has to express relationships between entities, as well as the relationships and the entities by themselves.
Taking the subject 'The influence of hereditary factors in the Psychology of delinquent children', one is in a dilemma to classify it by DC, even though all the terms appear to be in the scheme. Since the original number for 'Hereditary factors' in psychology at 136-733 no longer exists, for this purpose one has to rely on the place at 573.2 under Physical Anthropology, where it is directed by the index: Hereditary characteristics see Genetic make-up. Similarly, 'Delinquent Children' have moved from 136. 763 to 155.453 with other types of exceptional children. The scheme thus recognises the existence of the various facets of the subject: Psychology, hereditary factors, delinquent children. What it fails to recognise, however, is the importance of the relationship expressed here as 'influence'. Therefore it is convenient to classify the subject at either 155-453 or 573.2; This leads the scheme to lack of, not only precision, but also consistency of application.

DC does admit the principle of synthesis classifying complexes by assembly of their elements in certain fields. The simplest are the division of subjects by geographical location and chronological period. Dewey himself recognised the need for such synthesis from the beginning, and called it 'number building'. The method has several advantages like
consistency, economies the size of the scheme's schedules, makes ready recognition and fewer mistakes. The geographical numbers are taken from class 900 History, and usually, though not always, follow the symbol 0. Thus as 942 in England and 370 is Education, 370.942 is Education in England.

Provision for synthesis does occur elsewhere particularly in subjects with a quasi-geographical sub-division. There is a section on 'synthesis of notation' in the Introduction, in Volume. The schedule of exceptional types of children from Education at 371.9 may be used in the psychology class at 155.45, and a schedule of industries appears in 620 to 699, and these divisions may be used at suitable places, such as 331.892 strikes:

331.892 Strikes

331.8928 in special industries, divided like 620/699. 622 is the place for Mining; therefore; 331.8928 22 strikes in the mining industry and 331.8929 Strikes, divided by country therefore 331.892971 Strikes in Canada.

So that, in this case, the more specific subject 'strikes in the mining industry in Canada', at 331.8928 22 0971, would precede the more general subject of 'strikes in Canada',
contrary to the supposedly universal Decimal Classification Practice of following the logical principle of 'general' before 'special'.

A more extensive application, though, is in the seven separate Tables in Volume 1, which are applicable as instructed in the schedules.

1. Standard sub-divisions -01 to -09
2. Areas (geographical) -1 to -9
7. Persons -01 to -8

The editors have now tried hard to come to terms with advances in theory and practice, even claiming that they apply facet analysis, indeed, it would be hard to estimate over the value of the contribution made to the science of classification by Melvil Dewey, and DC can truly claim to have inaugurated a new era, so much so that, for library purposes all previous schemes have been superseded.

5.2.2 Universal Decimal Classification

The originators of the UDC, Paul Otlet' and Henri-La Fontaine, introduced DC to Europe in 1895, urging that it should be used to classify an international bibliography of books and periodical articles covering all subjects on cards.
This vast project was actually begun, and the inadequacy of DC soon became apparent. Permission was obtained for its expansion, which has proceeded ever since, several editions of UDC in various languages having already been produced, and a trilingual abridged version in English, French and German. In this country the British Standards Institution has undertaken the production of an English version as British Standard specification 1000. Intense propaganda for and against the UDC has been carried on ever since its publication, and many specialist libraries and documentation centres have adopted it, because its improvements on DC go far towards overcoming the defects as detailed above. After a detailed examination, however, Kyle has concluded that it cannot be recommended to new users, and that money would be better spent making a new scheme than on revising UDC. UDC is far better than DC, is internationally known and used, and its revisions are prepared by specialists. A.C. Foskett, in the most detailed critical examination of the UDC to date, concludes that its faults, rather than its virtues, have tended to be stressed by previous critics, and that it does in practice serve its purpose well enough to justify continued effort to bring it in line with modern thought both in subject organisation and in classification theory. He has
also made comparisons with American scientific and technical thesauri devised for computerised systems, and again suggested that the performance of UDC is better than might be expected.

The basic structure of UDC is the same as DC, a division of the field of knowledge into nine main classes, with the tenth, class 0, for General works. The three figure minimum has been dropped, but notational symbols, which tend to be long, can be split up by inserting a 'point' after every second or third figure to make small groups that are not so repugnant to the user. The main new feature, from the viewpoint of information retrieval, is the systematic attempt to cater for synthesis of complex subjects by means of a series of Auxiliary Tables, which include tables for geographical and chronological sub-division but also go much further; they are identified by the use of arbitrary signs such as +, /, ( ), and so on as well as .00 and .0. The disadvantage of these, of course, is that they do not have an accepted position in the sequence of numerals or letters, and their sequence has therefore to be laid down, and either memorised or consulted whenever the file is used. But the Auxiliary Tables certainly enable UDC to cope with even
most complex subjects, though in a somewhat cumbersome way. The most used symbol is the colon, which is the sign of relation, and does not actually have a Table as such, since it can be used to join any two (or more) numbers from the scheme. Thus, the example given above from DC, of 'The influence of hereditary factors in the psychology of delinquent children', is dealt with in UDC by the formula A : B, where A is 'delinquent children', B is hereditary factors' and r is the symbol of relation, the colon. The colon device has often been claimed to be the complete answer to the problem of expressing relations by means of synthesis, but it is, of course, no more than a beginning, since the same sign is used for all manner of different relations, influence, comparison, cause and effect, and so on, and there is therefore no means of distinguishing between these. Thus, in the index cards issued by the centre d' Information de Documentation sociales et economique Africaine, in Brussels, one come accross the subjects such as the following;

Mining research in Spanish Sahara

338:622 (661)

1. where 338 is production, 622 Mining and (661) is Sahara (and French West Africa)
2. Economic regulations affecting the cocoa trade on the Ivory Coast

351.824.5: 633.74 (666.8)

where 351.824 is Trade, 633.74 is the cocoa industry and (666.8) is the Ivory Coast.

Clearly, in these examples the colon sign stands for two different meanings, thus introducing homonyms into what is in effect an artificial language. Such a procedure inevitably brings ambiguity, always undesirable, and avoidable if the necessity of distinguishing between different notational symbols. The wide acceptance of the UDC, particularly on the continent of Europe, testifies rather to the power of the new devices for synthesis than to the intrinsic qualities of the scheme itself. The official view of the FID and the British Standards Institution is that the UDC can be brought up to date and equipped with devices able to cope with modern literature in a manner acceptable to specialists, but it is hard to see how this could in fact be done without a complete revision of the basic structure. In reviewing this, Neelameghan also relates the development of UDC as a standard reference code for UNISIST, and emphasises the need to base this on a general theory of classification suitable for documentation of "micro thought".
5.2.3 The Library of Congress Classification

The classification scheme drawn up for the collections of the National Library of the United States Congress was originally based on the defunct expansive classification made in the late nineteenth century by C.A. Cutter, but it was never intended for universal application, and was based entirely on the books in the Library of Congress. The various subject departments of the library compiled systems for their own collections, and each class was therefore developed more or less in isolation, by experts in the bibliography of the subject. The independence of classes goes to the extreme length of listing geographical and chronological subdivisions for each class; little provision for synthesis exists, and the same country may have many different numbers, according to the main class. Its notation also differs from DC and UDC in that it consists of two capital letters followed by numbers used as integers and not as decimal fractions; new subjects are intended to be dealt with by gaps left at what is hoped are suitable points in the sequences of numbers. Each class is revised and new editions are published independently, which leads to the curious result that some classes, such as Medicine, have had frequent revisions, others none.
Expect that it is in a scheme based on the collections of a national library which also acts as the library serving the Government, the social sciences are well treated in LC. Kyle remarked that the arrangement would suit well the 'old-fashioned' sociologist and, as Mills shows, the size of the collections has meant that the scheme contains an immense amount of detail and some very specific subdivisions.

The scheme begins with philosophy at class B, followed by History, extending over C-F, G is for Geography, including Anthropology and Customs, H for Social Sciences, J Political Science, K Law, L Education. Class H published in 1909 and simply reprinted in 1959—hence the 'old-fashioned' appearance—offers a distinct improvement on DC, inspite of odd omissions such as 'Social Change' and 'Stratification' which appear in neither schedules nor index. Class H is almost a circular affair, which begins and ends with theories, and has groups in the middle with activities on either side:

H Social Sciences
HA Statistics
HB Economics (HC History, HE Transport, HF Commerce, HC Finance, HJ Public Finance)
HM Sociology (including social psychology).
These are then followed by social groups:
HQ Family
HS Associations, Clubs
HT Communities, Classes, Races
after which there is a return to activities:
HV social pathology, philanthropy, charities and, finally,
theories of society:
HX Socialism, etc.
An example of failure to modernise the terminology can be seen in HM Social theory in which HM 101 - 121 is for civilization, with three headings among the subdivisions:
HM 106 Evolution Biological Sociology
111 Degeneration
116 Mutation
While a little further it is found
HM 278 Passive Resistance
Crowds, Tumults, Revolutions
HM 281 Theory
HM 283 History.

Classes are included among communities in HT, with HT 85: for slavery, which is clearly regarded as more than a merely economic phenomenon; Negro slavery appears under the heading
'General' in the schedule, but there is no index entry under the term 'Negro'.

Education, in class L, likewise contains some rather strange groupings and data from 1929. LA History, LB Theory begin well enough, but Theory is followed by Principles, which is joined with practice and leads directly into primary and secondary education. LC special forms has the same sort of progression, beginning with sociology of education and moral education and going off to special classes of persons including women and boys. LC 3951 - 3990 Education of exceptional children has no specific place for the gifted child, who does, however, appear in the index with a direction to the numbers LC 3951 - 90. The rest of class L, LD - LG, is entirely given over to individual institutions, mostly American, but including, for example, the names of Oxford and Cambridge colleges. Psychology, as in DC, is far away with philosophy in class B, and has very little subdivision and, that of a type that certainly could not cope with psychological literature in a way that would appeal to any contemporary school of psychologists. For all its detail, then, the LC scheme cannot claim to offer an adequate system for information retrieval. Certainly, it can show in places an approach to division into categories that appear in the
modern literature, but this division does not give the impression of being systematic and is by no means consistent.

5.2.4 The Bibliographic Classification

Henry Evelyn Bliss, the compiler of the Bibliographic Classification, spent his whole life on this scheme and its theoretical basis, and his work form an important part of modern classification theory. In his books on the organisation of knowledge he considered how subjects develop, and attempted to systematise the relations between subjects and between subject and their parts. He maintained the objective reality of 'natural order', and thought that if a classification scheme accurately reflected this order it would not need radical although many changes continually take place at the frontiers of knowledge, necessitating frequent revision of the details of classification schemes, the great 'main classes' of knowledge are relatively stable, and that a scientifically based sequence of main classes would therefore last for a long time.

BC is based on four principles:

1. Comprehensiveness. The scheme must cover the whole of knowledge as shown by the current consensus of scientific and educational thought.
2. Subordination. The development of subjects from general to specific called by Bliss' gradation by speciality.

3. Co-ordination is a important factor, the collection of related subjects in a manner helpful to readers; alternative collections are provided for some subjects in recognition of the fact that different groups of users may approach the same literature from different viewpoints.

4. Composite specification. This is Bliss' term for the ability to bring together different parts of the scheme in a manner flexible enough to cope with complex subjects.

He extended the use of auxiliary tables, as shown in JDC, very considerably, providing some 45 in all — some general, such as the geographical and chronological tables, some specific to one subject, such as the administration schedules 14 a - d for use in class J Education.

Bliss made a deliberate attempt to find a scientific basis for his sequence of main classes. He has outlined mainly sciences, social sciences, humanities. The sciences of Man begin with class H Anthropology and medicine, and go on to I Psychology, J Education to sociology, L-O History Q Social Welfare, R Political Science, S Law, T Economics.
At the beginning the terminology was more natural but once again it is found that some important terms such as 'Social Change' and 'Stratification' are from both schedules and index. A new edition of BC is under preparation under the direction of Jack Mills, at the School of Librarianship of the Polytechnique of North London. Extensive use has been made of facet analysis in this. In Education, which is largely based on London Education Classification used in the Library of the London Institute of Education perhaps the most important change is the provision of integral schedules for Educational Administration, including Central and Local Government as well as school organisation. The Primary facet, Education and Schools, comes at the end at JL; primary education is JM, and the first direct sub divisions begin at JMM, JMN, etc.

JMAB Aims and objectives of Primary Education, where AB is taken from facet JA principles and philosophy of Education, in which JAB is Aims Objectives.

JMC Administration of Primary Schools, where C is taken from facet JC organisation and Administration of Schools.
JMK Curriculum in Primary Schools, where k is taken from facet JK curriculum.

Bliss classification is in use in a number of libraries already, notably in schools and Institutes of Education and other school libraries; a Bliss Classification Bulletin has appeared regularly.

5.2.5 The Colon Classification

The Colon Classification is the work of S.R. Ranganathan, and was first published in 1933. Colon classification Edition 7 is in practice. It presents the singular paradox that though little used outside India, it has had an outstanding impact on classification theory, and like the DC, marks the beginning of another new era in the science of making schemes of classification.

Among all the other general schemes, CC recognises the absolute necessity and precise specification of subjects for the purpose of information retrieval, that is the length of the resulting notational symbols. By 'depth classification', Ranganathan means the provision of a scheme that is sufficiently detailed and flexible to allow a classifier to deal fully and accurately with any topic no matter how ever
specific it is. Ranganathan himself has opened an attack on the problems of classification in the social sciences with an analysis of their emergence and scope. He has a brief review of their treatment in the other general schemes, and goes on to relate each branch to his 'spiral of scientific method' to show which has a legitimate claim to the description 'science' and which should be included as 'auxiliaries'; these include Geography and History. His sequence includes Social sciences as a class of partial comprehension and similarly, Geography, History Political Science and Economics. Ranganathan has not paid much attention to the problem of aim classes and their sequence. This means that Biology and Medicine are followed by useful Arts, spiritual experience, Fine Arts and so on. Religion and Philosophy lead back to Psychology, Education, Geography, History, Political Science, Economics, Sociology, Law.

In CC the principle of composite specification is applied to the division of the classes themselves, which consist as a result of series of tables from which classification numbers have to be assembled to make up a notational symbol for a subject. Ranganathan has developed a theory that each set of
sub-divisions, or 'facets', can be produced by the use of characteristics related to five fundamental categories, Personality, Matter, Energy, Space and Time. The Personality Facets in Edition 7 include:

Class X Economics

X26 Standard of living

2 Low
3 Medium
4 High
5 Measurement
51 Habitat
52 Furniture
53 Food
54 Transport
55 Ornament
56 Telephone
57 Clothing
5X Purchasing Power

Class Y Sociology

1 By Age
11 Child
111 Newborn (Less than 3 months)
112 Toddler (8 to 12 months)
113 Infant (1 to 6 years)
114 Pre-adolescent (7 to 14 years)
12 Adolescent (14 to 25 years)
13 Post Adolescent (25 to 30 years)
14 Adult
15 Middle Age (40 - 50)
16 Old Age (60 - 80)
17 Very Old Age (80 - 100)

The energy facets include: Matter property is energy in Edition 7

X Economics

1 Quantity

Y Sociology

by Method
<table>
<thead>
<tr>
<th>12</th>
<th>Infra - normal Quantity</th>
<th>1</th>
<th>Defection 11 Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Normal Quantity</td>
<td>2</td>
<td>Prison</td>
</tr>
<tr>
<td>14</td>
<td>Suppo - normal quantity</td>
<td>3</td>
<td>Punishment</td>
</tr>
<tr>
<td>15</td>
<td>By measurement of</td>
<td>4</td>
<td>Extermination</td>
</tr>
<tr>
<td></td>
<td>Quantity</td>
<td>5</td>
<td>Asylum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55</td>
<td>Correction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Work House</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>Relief Work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>Charity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>91</td>
<td>Corporate Punishment</td>
</tr>
</tbody>
</table>

The subject 'Growth of Nationalism in India' receives the notation V, 44; 42'N29 - M 87 and the subject 'British Politics, 1964-70' receives the notation W, 6.56'N7 and the subject 'Agricultural Economics receives the notation 4X, 8(J). This illustrates how each facet can be divided without interfering with the others. The same facility exists for synthesis of subjects from different main classes, and in Co, unlike UDC, the different relations are represented by different signs, in the same way as the different facets. Thus, the subject illustrated before, the influence of hereditary factors into the psychology of delinquent children; and numbered as S2 - 65 Og G; 61 In which S is Psychology, S2 is Adolescent Children, - 65 is Criminal, Og is the Phase.
relation 'Influenced', G is Biology and ; 61 is Heredity as Matter Property. This is a complex notation, but it expresses accurately a complex subject.

Similarly, where the same term may appear in several classes, CC makes every possible effort to use the same sequence and notational symbols. Whatever the difficulties are there in CC, there can be none about the power of the method of facet analysis.

5.3 Specialised or Depth Classification

The general classification scheme is one which is designed to cover all the subjects in the universe of subjects. The special classification scheme is the one designed to cover the subjects going with a specific basic class or any other host class. A scheme may be designed to cover any one such field of specialization (10). There are many reasons claimed for making a special classification. These are:

a) Lack of coextensiveness in general schemes
b) Lengthly class numbers.
c) Special requirements or special point of view:
general schemes are designed to take into consideration the majority point of view. Therefore, these are not able to fully meet the special requirements of a particular special library or information centre.

  d) Lack of flexibility in general schemes
  e) Lack of helpful sequence.

There are three different approaches to special classification schemes. These are:

  1) A fully autonomous special classification. Such a scheme may be independent of any general scheme. According to Dobrowalski, "Experience shows that, among documentation centres, those which function best have a classification specially devised for their use, independent of encyclopaedic classifications, which we will call an "autonomous classification". A special autonomous classification is better adopted to the documentation collected by the centre, and to the functions which classification must provide in documentary tasks. (11)

  11) Not fully autonomous special scheme:

A special classification scheme may be prepared so that it is autonomous for special subject(s) and dependent upon general classification for allied subjects. Such a scheme is not a fully autonomous special scheme. This approach would lead to
some saving in the efforts of a classificationist, as well as a classifier.

111) The special classification scheme may be designed in such a way that it is dependent upon a general classification. It may be an extension of the general scheme such an approach may be called the "Do-all classification approach". Ranganathan was of the view that we should have a do-all classification, which can meet all the requirements of the special classification in different subjects irrespective of their extension as well as depth of intension (12). It should be based on a freely faceted scheme for classification worked out by Ranganathan and his school of thought. Colon classification is being developed as a do-all classification.

5.3.1 Advantages of do-all Classification

A do-all classification scheme/depth classification scheme can have its advantages:

1) It should be possible to provide co-extensive class numbers.

11) The length of the class number will also become short.
According to Ranganathan, the class number of a special subject of interest can be replaced by the hyphen (−) or the zero (0). The ordinal value of such a digit can be fixed so that the first position can be secured for documents in the special subject and its sub-division. As a result the length of the class number will also become short.

111) A do-all classification of the freely faceted type should be able to serve special requirements in a satisfactory manner.

11v) By its very nature, a freely faceted scheme is likely to have a good provision for incorporating new subjects in its proper places without disturbing the helpful sequence.

1v) It will be able to offer optimum helpfulness in the arrangement of documents or entries.

1vi) In a do-all classification much effort can be saved because the same common isolates, normative principles, notation, notational techniques and so on can be used. Here once the guidelines have been developed, these can be used for different special schemes.

1vii) Use of the do-all classification can take us further towards standardization and cooperation.
viii) In this classification, it is easy enough for a user to shift from general classification to depth classification, because of the same scheme.

ix) Coding for machine retrieval would be less costly for the do-all classification.

From the above, it is clear that the "do-all" classification approach is the best one. If such an approach is achieved with the help of a freely faceted scheme for classification, guided by a dynamic theory of library classification, then the result would be highly satisfactory.

5.3.2 Design of Depth Classification

There are several steps involved in the design of depth classification or do-all classification.

1) Collection of abstracts on the subject.
2) Analysis of abstracts.
3) The isolates picked up from the abstracts are then posted on to the kernel term record.

The kernel term record gives the following information:

1) Kernel term.
2) Reference to the abstract in which it occurs.
3) Manifestation—If speciator the isolate which it speciates.
4) Characteristic used.
4) These terms will at first be sorted out according to their manifestation, thus bringing all the (P)'s together, (E)'s together and so on. Characteristics are standardized and defined. Thereafter they will be grouped according to the characteristics used.

5) These characteristics or Quasi Isolates (QI) are then arranged in a sequence which it is felt to be helpful to a majority of the users.

6) Later the Quasi Isolates within the group are arranged in a helpful sequence.

7) Then the isolates were listed under each of the QI's.

8) The next step is the allocation of sectors and the assigning of notation, mixed notation comprising of a Roman alphabet (lower or upper case) and Indo-Arabic numbers to be used. The sequence of allocation of sectors has to be drawn from the least value to the highest value.

9) Then the schedule is to be tested by classifying some more abstracts and connections and adjustments to be made will noted.

10) The schedule was then finalized after carrying out the necessary corrections and index to the schedule was then prepared, the Kernel terms serving as the base.
5.4. Indexing and the Determination of Hierarchical Relations/Non hierarchical Relations

The need for new kinds of tools to help in searching information is increasing due to the rapid generation of new ideas. Instead of mere reference to documents, it would be helpful to provide aids in selecting the appropriate documents if those tools are capable of identifying the thought contents of the documents in a condensed and precise form. Thus, research is going on throughout the world to develop a suitable indexing system which will be capable of providing both the location and the essence of the documents.

A good index should relieve the user of the necessity for making a mental image of all the network of possible relations for the term by which he identified the subject of his study.

Indexing is regarded as the process of assigning tags of items of information for the purpose of retrieval. The tags are the index terms (13). Items of information may be tagged for retrieval by subject headings, classification numbers, descriptors or search keys; these are then arranged alphabetically or in some systematic manner in an index. From this point of view, a classified catalogue is as much an
index as the index at the back of a book or a periodical index or a set of search keys.

All spoken languages consists of two basic elements: a vocabulary, and a grammatical structure (syntax). Writers on information retrieval often use the term natural language to denote spoken languages. It is possible, in indexing to employ natural language, that is to use language simply as it is spoken, or used in documents. Such an index is often referred to as a natural language index. Alternative to natural language indexing is to use an artificial language adopted to our needs, that is an index language. Such a language will reflect a controlled vocabulary, that is to say that careful decisions will have been made as to the terms to be used and the meaning to be attached to each. So an index language is an artificial language which is adopted to the requirements of indexing (14). The vocabulary of an index language will be controlled. An index language may also seek to indicate the relationships between the terms in its vocabulary. In this regard the paper of C.J. Van Rijsbergen and others tackle, the problems of how one might select good search terms, using relevance back feed, given to search terms in the query (15). There is a criterion for the selection of terms. Terms are the external form of concepts. Concepts
may be regarded as a trio consisting of (A) a referent, (B) the characteristics of a referent and (C) the external, communicable form of the referent and its characteristics, the term (16). The relationship of these three can be shown graphically in the form of a triangle.

Fig. 6. The Concept of Triangle

Whitelock's article presents proposals for the organization and creation of a Descriptor Bank of terms used in social science. (17) Clement T.Y discussed a methodology to the construction of term classes. Heuristic methods to the construction are presented (18). Harris discussed the 'terminology change: effect on index vocabularies (19). The selection of terms to be used in our indexing system should be based primarily on a survey of the literature.
An indexing language which has a controlled vocabulary, and which attempts to indicate the relationships between terms in the index vocabulary, is said to the structured index language. There are various kinds of structured index language, examples classification schemes, Sears list of subject leading, Engineers Joint Council Thesaurus of engineering and scientific terms. Careful subject analysis is the basic to all structured index languages. The vocabulary of an index language may be verbal or coded. A classification scheme employs a coded vocabulary in the form of its notation. Sears list of subject leadings employs a verbal vocabulary. Dahlberg describes establishment of compatibility between indexing languages (20). The first step in subject analysis is a grouping process by which we seek to group together terms which share an essential relationship. Grouping is a classificatory process, and it can eventually lead to the production of a classification scheme. Main purpose is to establish the relationships which a structured index language should display. The grouping process in index language constructive seeks to assemble out terms into broad categories. In classificatory terms this is referred to as division, since the terms within the subject, or basic class, are divided into a number of categories for
class) that is a characteristic of division is the idea which governs our grouping of terms into a category around 1957. Ranganathan realised that fact analysis is not a classification or the technique is not exclusively designed for classification, it could be used in the precise formulation of the subject requirement of the reader, in the construction of subject leading of a document, in the construction of thesaurus and in the formulation of the query of the user (21 & 22). In 1964 Ranganathan demonstrated the use of fact analysis is based on the set of postulates and principles forming part of the general theory of classification in choice of the name of subject (subject representation and its rendering without using class number (23). After period of time, the deficiencies of chain procedure were realised and led to the development of postulate based permitted subject indexing PQPSI in 1969. now a days it could be constitute as a semi-automatic indexing system (24). PQPSI has multipurpose use.

There are several steps to be followed in formulating the coextensive representation of each of the specific subjects of a document or a query (25). These are.
Step 1 - Specific subject: Let us of the specific be selected example "priority in employment for jobless engineers in India".

Step 2 - Component ideas
Determining by analysis, the specific component ideas in the subject and the broader or superordinate ideas or upper links implied by each of the component ideas.

<table>
<thead>
<tr>
<th>Specific idea</th>
<th>Superordinate Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economics</td>
<td></td>
</tr>
<tr>
<td>Human Resources</td>
<td></td>
</tr>
<tr>
<td>Engineers</td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td></td>
</tr>
<tr>
<td>Employment priority</td>
<td>Manpower plan; priority</td>
</tr>
<tr>
<td></td>
<td>Assignment.</td>
</tr>
<tr>
<td>India</td>
<td></td>
</tr>
</tbody>
</table>

Step 3 - Sub-structure formation
linking up the component ideas recognised in step 2 to form sub-structures.

The use of postulates and principles based on a theory of subject classification would help in formation of sub-structure of the subject.
Facet-type Idea  
Economics  
Human Resources  
Manpower Plan  
Priority Assignment  
Employment Priority

Facet-type Idea  
Engineers  
Unemployed

Associate Speciators - Engineers - Unemployed.

The hierarchic relations among the ideas are shown by linking together the successive super-ordinate ideas for each component idea.

Man Power Plan  
Priority Assignment  
Employment Priority
denotes a hierarchical or inclusion relation.

Step 4-sequence within sub-structure.

The components within each sub-structure are arranged in a helpful sequence.

Human Resources-  
Engineers - Unemployed
Man Power Plan Priority
Assignment Employment
Priority.

Step 5 - sequence among sub-structure
'Economics Human Resources Engineers
Unemployed Manpower Plan Priority
Assignment Employment Priority India.

step 6 - Indication of Sequence and Relationships,
'Economics, Human Resources - Engineers - Unemployed,
Manpower Plan Priority Assignment Employment
Priority India.

step 7 - Standard Terms.
'Jobless' is a term used synonymously for the term 'unemployed'.

Index languages may be divided into two broad groups, distinguished by the way in which they specify compound subjects (that is, subjects consisting of two or more concepts in combination). The first group is known as Pre-coordinate, and post-coordinate index languages. Pre-coordinate means that the elements of a subject are put together by the indexer, before a search, to form a single
heading under which documents and that subject are entered in
the index. A post-coordinate index language is one which
includes only terms representing a single concept with no
provisions for their combination by the indexer to specify
compound subjects, in a post-coordinate index the terms
necessary to specify a compound subject are brought together
(co-ordinate) by the user after (post) the search.

An interesting variation of the ordinary author catalogue
which uses tracing by authors' names as a clue to the
development of a subject is the 'citation index' devised by
Eugene Garfield, the objective of constructing a systematic
subject index, is to create a network of terms which brings
together, by a series of consecutive entries or by cross-
references, those terms which are used together or in a
formal relationship in the literature, so that when one term
is chosen as the key with which the searcher enters the
index, all other relevant terms will be recalled
automatically. The Subject index has to rely for its
efficiency and what is actually put in, which means that the
whole of the linkage network has to be the result of
conscious action (26). A general theory of subject indexing
Language(SIL) has developed through logical abstraction of the
structures of outstanding subject indexing languages such as
those of Cutter, Dewey, Kaiser and Ranganathan (27). POPS! (postulate Based Permitted Subject Indexing) and classaurus, that is a faceted systematic scheme of hierarchical classification having all the necessary attributes of a conventional thesaurus has developed.

There are various types of indexes. Four main types of index are in common use:

(1) dictionary or straightforward alphabetical;
(2) alphabetico - classed, a mixture of dictionary and systematic;
(3) classified based on a scheme of classification;
(4) indexes based, not on manual searching, but on the use of machines of varying degrees of complexity.

All are techniques of dealing with units that need to be assembled in different ways from time to time to meet the differing requirements of users, and all are based on analysis of the contents of documents so that the concepts contained in them may be co-ordinate with the same and related concepts in other documents. Due to the result of machine index, (MARC) Machine Readable Cataloging - produced first by the LC and then in collaboration with the BNB since its inception in 1950, the BNB has been classified by the DC of Melvil Dewey, and its subject index has not followed the
Relative Index of DC, using the 'chain procedure' derived by Ranganathan. When the BNB changed to DC 16th edition and computerised printing, chain indexing was abandoned and a new system known as PRECIS (preserved context indexing system) has been introduced. This has been devised by Derek Austing, and is the outcome of several years of research, beginning with the classification Research Group's project for a new general classification for information control. (28).

5.4.1 Relationship Among Ideas

The relationships among ideas forming components of subjects may be categorised as (1) hierarchical relationship and (2) non-hierarchical relationship (29). These may be sub-divided broadly as shown in the following chart.

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Hierarchical</th>
<th>Non-Hierarchical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genus-species relationship</td>
<td>Part-whole relationship</td>
<td>Associative equivalence relationship</td>
</tr>
<tr>
<td>Genus-species relationship</td>
<td>Part-whole relationship</td>
<td>Associative equivalence relationship</td>
</tr>
</tbody>
</table>

In a thesaurus the HR is usually represented as Narrower Term (= NT)/ Broader Term (= ST) and an NHR is represented as a Related Term (= RT). In some thesauri, the part-whole
The NHR equivalence relationship indicates terminological relationship between terms denoting the same idea—that is, synonymous.

Example:

'India' and 'Bharat'

Generally, the term commonly used by a majority of users of the system is selected from among the alternative terms for a given idea. This preferred term is linked with the alternative terms (unpreferred terms) by UF (used for) and to unpreferred terms by U (use).
Other type of NHR between any two ideas may arise from the way in which the ideas are associatively treated by an author and also from the way in which the ideas may be associated with each other in the users queries. Theoretically there is a wide variety of ways in which authors and users may associate two ideas. However, an associative relationship should normally be established only if it is likely to be helpful in information retrieval. Example of NHR

\[ CR = \text{Coordinate Relation} \quad PR = \text{Phase Relation} \]
\[ FR = \text{Facet Relation} \quad SR = \text{Speciator Relation}. \]

<table>
<thead>
<tr>
<th>S.No.</th>
<th>NHR</th>
<th>Example of subjects</th>
<th>Faceted representation of Relation (d) (C) and thesaurus entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Process occurring in sequence</td>
<td>Settlement of disputes by arbitration</td>
<td>Disputes: settlement, Arbitration</td>
</tr>
<tr>
<td>2.</td>
<td>Process (1) and its property (2)</td>
<td>Impact of non-violence</td>
<td>War - violence; Impact non-violence RT</td>
</tr>
</tbody>
</table>
3. Process\(^{(2)}\) Voting age and property of object (1) associated with the process

4. Coordinate Types of ideas that is, ideas in Pathology one and the same array derived by division of a superordinate idea on the basis of a single train of characteristics.

---

5.5 **Construction of Vocabulary Control Devices**

The performance of an information retrieval can be improved by the use of controlled vocabularies, such as classification schemes, subject authority lists, and thesauri\(^{(30)}\). Recent trends in the field of design of different types of controlled vocabularies have developed exclusively, principles and rules for each one of these devices. While it is admitted that there are certain essential difference in each of these types of controlled vocabularies, there are also certain features which emphasize basic similarity, interconvertibility and compatibility among them. According to Kent, thesauri is, a compilation of terms of a given
information retrieval system's vocabulary, arranged in some meaningful form and which provides information relating to each term that will enable the user of information file to predict the relevance of responses to questions when this particular control mechanism is used (31). Thus a thesaurus acts as a device, which helps users to formulate their questions precisely. It also acts as an aid to indexer in assigning the preferred descriptor to subject of documents.

The function of a classification scheme is more or less similar to that of a thesaurus. The essential unity of the functions of a thesaurus subject heading list, and classification scheme is successively indicated by John H. Schneider (32) in the following figure:

Uncontrolled Alphabetical Three word Multi levels Vocabularies subject head level trees hierarchical authority Thesauri classification lists.

5.5.1. Guidelines for the Construction of Thesaurus

A fairly good number of documents have been published to aid the construction of a thesaurus. The more important of them include: UNSECO's Guidelines for the establishment and
development of monolingual thesaurus (33), and ISO: 2788 Guidelines for the establishment and development of monolingual thesauri (34). In addition to this several treatises are written on thesaurus. They include Lancaster's Vocabulary Control for Information Retrieval (35), Soerges's Indexing Languages and Thesauri Construction and Maintenance and Aitchison and Gilchrist's Thesaurus Construction, A Practical Manual (36).

5.5.2. General Features

A thesaurus, as a structured subject of natural language, describes the subject content of the documents, objects of collection of data.

A particular thesaurus should accurately reflect the information content of the body of documents or other items in a collection to which the thesaurus applies. Thesaurus is used for post-coordinate representation.

A thesaurus should contain terms and cross-references appropriate to the subject matter taking into consideration both the language of the information and the needs of the users.
A thesaurus essentially functions as a controlling device for the vocabulary used in a subject field or as field of activity. It generally entertains the users by accepting their own terms into its vocabulary (usually known as entry vocabulary). And, if the user happens to come with a non-preferred term it refers him to the preferred term by providing cross-references. It also indicates to the users, the standard vocabulary of the language.

A thesaurus provides words or terms to express meanings that are implied by the term relationships given in the thesaurus (in contrast to a dictionary). In a thesaurus, each preferred term collects together, all its associated terms. This can be treated as atomised presentation of relationship.

Thesaurus may be arranged like an alphabetical index and the terms in a thesaurus may be used to construct and index. However, the thesaurus itself is not an index. An index to a collection must have addresses or locators of an index. A thesaurus classifies terms by arranging them in hierarchical classes. As a "term classification system" a thesaurus has some similarities with subject matter classification system, for example, by the universal decimal classification.
A thesaurus is one kind of authority list, that is, the preferred terms in a particular thesaurus are required, indexing and retrieval terms for an information and documentation system. There are other kinds of natural language based authority lists, such as subject heading lists. However, these do not have the hierarchical structure of the thesaurus.

Cross-references in a thesaurus make explicit the ways in which entries relate to each other in a network of concepts. Criteria for the choice of a lead - in term (sought heading) is explicitly stated in the form of Canon of sought heading in the subject heading rules. In a thesaurus, the descriptor can be characterised as an authorised and formalised term or symbol used to represent unambiguously the concepts of documents and queries.

Descriptors in a thesaurus may be terms denoting concepts or concept combinations, terms denoting individual entities (these are also called proper names or identifiers, like project names, nomenclatures, trade marks, acronyms, etc.). The definition of the term 'descriptor' in thesaurus is explicit.
In thesauri not unsing preferred terms, all terms included in the thesaurus may, in principle, be descriptors. The concept - denoting terms not permitted in indexing must be regarded as unauthorised terms. They are called non-descriptors.

A thesaurus which includes preferred and non-preferred descriptors will be taken as equivalent to our alphabetical index to a schedule. A descriptor in a thesaurus may consists of one or more words.

A descriptor in a thesaurus should reflect the terminology of the subject, but it should contain as few words as possible, and preferably only one.

In a thesaurus the most widely accepted spelling of the word should be adopted. In cases, where, due to varying usage, more than one spelling of a word is accepted, both spellings should be included in the thesaurus and reference made from one to the other. A well established dictionary can be chosen to act as arbiter, whenever this problem arises.

In a thesaurus many current technical terms have arisen by translation from other languages but sometimes, a modern foreign language, Latin or Greek term is incorporated into the specialised vocabulary for a particular subject when both
the foreign language term and its putative translation coexist with the same meaning, both should be included in the thesaurus and reference made from one to the other.

The problem is further complicated when the foreign language in question is written in a different alphabet. This is particularly true in the case of identifiers (proper nouns).

Descriptors should be preferably in noun form. The use of the singular or the plural form of descriptors should be decided in accordance with the usage in the language of the thesaurus. Abbreviations should be avoided. Well established acronyms are acceptable as descriptors. Punctuation marks in descriptors should be minimised except for specialised nomenclature, only Parenthesis and the Hyphen are needed as descriptors. The different meanings of homonyms must be marked and distinguished by specifying symbols. Scope note should be brief explanation of a descriptor.

Descriptor interrelationships are a) Equivalence relationship (Perferential relation), b) Hierarchical relation, c) Associative Relation. The authenticity of the descriptors should be verified by consulting dictionaries, other indexing or standardized vocabularies, current usage in the literature and especially the opinion of subject specialists.
Thesaurus assimilates the neologisms and special person that proliferate in expanding fields of basic and applied research. The thesaurus should be able to accept new terms as and when it is used in the technical articles or reports published in the subject of field.

5.6. Searching in Social Science Information Retrieval

The term 'Information Retrieval' has been in current use for years, and is valuable because it describes succinctly an important new stage in the development of Library services (37). It was defined in 1950 by Calvin Mooers as 'Searching aid retrieval of information from storage according to specification by subject'. The 'retrieval' aspect of information retrieval is no new; library catalogues have always been tools for retrieval. In a classified library, the location of the book is determined by its subject. But articles in the same periodical may deal with a variety of subjects; and in most libraries periodicals are shelved in separate sequences, thus the subject approach of the specialist must often be through an index, a highly specialised and specific form of subject catalogue. Information retrieval has developed as an important part of information service, the spectacular progress of the last
years has brought with it new methods and new appraisals of traditional methods. This progress has been most notable in the fields of science and technology than in Social Sciences.

Information retrieval consists of four main stages:

1) Identifying the exact subject of the search;

2) Locating this subject in a guide which refers the searcher to one or more documents;

3) Locating the documents.

4) Locating the required information in the documents.

Looking at the state of research in the social sciences, one can see emerging some alternatives of the circumstances which acted so favourably and the growth of information service in science libraries. a) the literature itself grows in quantity at a rate not far short of that of science; b) need problems are thrown up, and the achievements of modern technology, c) publication and exchange of information proceeds as rapidly as research and field work, d) the effect of improving standards of data collection and analysis, emphasising the trend towards out increasingly scientific and objective approach. Social Science research also tends to resemble natural science research in
becoming more and more cumulative. The fact that social science literature has a longer-lasting validity does not alter the case; it merely aggravates the problem of information retrieval by increasing the quantity of literature needing to be organised. The importance of data cumulation, of course, is that it permits of a division of labour by making tolerable the delegating of literature searching. This leads to the tendency of team work.

In order to carry out information retrieval, the material flowing daily into libraries has to be organised in such a way that it can be assembled by subject over a period of time. If one selects from day to day those items that are relevant for each of a group of subject specialists a continuous record of the selection forms a cumulating list of references to each of the specialist subjects. The tools with which librarians organise the literature for its information content are classification and indexing. A systematic classification and an intelligent bibliographic index of this should influence a great advance.

Some considerable improvements have been made, and progress reports have appeared in the UNESCO chronicle in 1955 and in the International Social Science Journal in 1962. UNESCO
and the ICSU Abstracting Board sponsored a feasibility study on a world information system known as UNISIST between 1967 and 1971 and it has now been decided that it will include the Social Sciences. In the United States several organisations have joined together to form the National Federation of Science Abstracting and Indexing Services, which includes services in the Social Sciences. The new publication will fall into three main groups: indexes and abstracts of current literature, narrative surveys dealing with specific topics, and guides to the literature of specific fields.

Guides to the literature of specific fields appear from time to time. Two general guides that illustrate different approaches are Hostlitz’s A Reader’s Guide to the Social Services, 1970; and The Literature of the Social Sciences, 1960 compiled by P.R. Lewis and the use of Economics Literature edited by John Fletcher, and use of Criminology Literature edited by Martin Wright have already been published.
In science successful dissemination of information has been achieved because information libraries have been established to cope with the flood of literature and its readers. But this is not the case in the Social Sciences. Due to the influence of modern techniques of documentation in Social Sciences this situation is changing, especially in research libraries, where the librarian himself is a scholar in a field of social science. Modern systems of interlibrary lending, union catalogues, such as a British Union - catalogue of periodicals in Social Sciences, make it possible to locate within a moment; techniques of document reproduction, and the photostat and microfilm, use of microcards and xerography make the very quick information retrieval process in the Social Sciences.

5.7 Conclusion

In this chapter an attempt is made to study the development of terminology in Social Sciences. The survey of the treatment of social science subjects by the main general classification schemes shows that none of them offers a wholly satisfactory answer to the needs of research workers.
Steps for Specialised or Depth Classification Scheme enunciated by Ranganathan was studied. Depending upon this methodology the next chapter explains the Depth Classification for 'Labour Problem'. Various indexing languages and their hierarchical and non-hierarchical relations have been studied with the help of this section, the next chapter will discuss the indexing devices for 'Labour Problems'. In the development of an effective and efficient thesaurus for information retrieval purpose, considerable intellectual effort is necessary. In the field of social science observation is made that, there is just as great a need to ensure research. Time is used to the best advantage, and while we certainly do not need to initiate scientific systems without due regard to the, different nature of social science data, much can be created from general knowledge of these systems and the adoption of them to our own purpose.
REFERENCES


3. Felber, H "A outline of international terminological activities". In Intern Classification. 1, 1974, p 89-91.


7. Ibid.


13. Sevnonius, Elaine "good indexing: A question of evidence" Lib sc 12, 1975, Paper D.


22. Ranganathan S.R. "Subject heading and document indexing system". Annual seminar (DRTC) 7, 1969, Paper U B.


