# KNOWLEDGE ORGANIZATION

## M.LIB.-102

**DIRECTORATE OF DISTANCE EDUCATION**
SWAMI VIVEKANAND
**SUBHARTI UNIVERSITY**
Meerut (National Capital Region Delhi)

**SLM Module Developed by :**
**Dr. Ramkumar Pathak**
Ph.D., MLIS, PGDLAN



**Reviewed by :**
**Dr. Javed Khan**

**Assessed by :**
Study Material Assessment Committee, as per the SVSU Ordinance No. VI(2).



**Printed and Published by :** VIDYA Prakashan Mandir (P) Ltd.
Vidya Industrial Estate, Baghpat Road
Meerut-250 002 (Delhi-NCR)
info@vidyaprakashan.com

(M.LIB.—102)

# Knowledge Organization

## Syllabus

**Unit 1:**

- Development of General Theory of Classification; Classification Principles from the 19th Century to the Present; Mapping of Universe of Knowledge: Categorization of Isolates: Historical Perspectives and Modern Trends

**Unit 2 :**

- Mapping of Universe of Subjects in CC, UDC and DDC: Special Classification Schemes; Development of Cataloguing & Catalogue Codes including Modern Trends

**Unit 3 :**

- Contributions of Cutter, Lubetzky, Ranganathan in the Field of Cataloguing; Layout and Rules for the Union Catalogues of Books. Periodicals; Indexing and Abstracting Journals and National Bibliographies

**Unit 4 :**

- Online Cataloguing: OPAC and Web OPAC; Subject Cataloguing: Definition and General Principles; Choice and Rendering of Subject Headings: LCSH

**Unit 5 :**

- Study of UNIMARC & MARC-2; Comparison of CC & UDC with regard to Common Subdivisions and Indicator Digits; Recent Trends in Classification; CRG, BSO, Automatic Classification, Classification in Online Systems, Web Dewey

# CONTENT

UNIT 1



Note

# Classification of Knowledge

## 1.1 Introduction

In this chapter we will learn about knowledge and study the classification of knowledge and why knowledge is necessary. Apart from this, in this chapter, we will learn about the positioning of the world in terms of various classification systems. In this chapter we will talk about the common isolates method. We will also study and learn why this method is necessary for studying and how it facilitates the law of acquisition of knowledge by man.

## 1.2 Need of Library Classification

Library classification has been derived from the word 'classis' which means 'grouping'. The process of grouping and categorising similar items and objects which is very important in formulating groups is called a classification. This process helps the user to arrange, organize and make a logical sense of any collected material to find or locate them in an easy manner. According to S.R. Ranganathan "Library classification is the translation of the name of the specific subject of a book into a preferred artificial language of ordinal numbers and the indivisualisation of the several books dealing with one and the same specific subject by means of another set of ordinal numbers which represent some features of the books other than their thought content. The first of these ordinal number is called the class number of the book. The second ordinal number is called its Book number. It is usual to separate the book number from the class number by a space or to write the former beneath the latter. The class number and the book number taken together constitutes its call number. The call number of a book fixes its position relatively to the other books in the library." Before we attempt to study the theory of library classification it is necessary for us to know the importance

Note

of developing a theory. It is equally necessary for us to recognize the need for such a theory.

### 1.2.1 Importance of Theory

At the outset, one must know what constitutes a theory and how it is important for development of a subject. A theory refers to an organised set of principles, which provides the basis for further investigations into and the development of a subject. It explains the what and why of the existing phenomena. It qualifies the subject to be accepted as a discipline. It provides a scientific basis for a subject and brings respect ability and status to it. Its importance for the growth and development of a subject hardly needs emphasis.

### 1.2.2 Need of a Theory

It we look into the history of library's classification, we find that during the early stages of its development it handled a small number of subjects constituting the whole of knowledge, and a broad classification met the requirements of that time. The schemes were prepared largely in response to the exigency of the time. These schemes seem to have been guided by the purpose on hand rather than a theory that would stand the test of time. These schemes solved the immediate and short term problems. However, with the passages of time the number of subjects into which knowledge could be divided steading increased, proving the existing schemes inadequate with the growing complexity of subjects enshrined in documents it became necessary to classify knowledge minutely.

### 1.2.3 Development of a Theory

In any sphere of life, practice precedes theory. Life fore stimulates man to improvise, design and develop various aids – both at the physical and mental levels. After a long experience is gained with an improvised aid, a theory is developed in order to understand the aid deeply and to systematise, improve, refine and develop it. So also it has been with classification. Within fifty years : after the design of Decimal classification, Richardson add Sayers made comparative studies of then known schemes for classification : and they also evolved a Theory of Classification. It was static and not dynamic. The emphasis at this stage, according to Parkhi in his book Library Classification, Evolution of a Dynamic Theory, was on the description of the practices followed by the classifications in designing their schemes and were considered as norms for designing schemes.

On the other hand, after 1949, Ranganathan and his associates slowly evolved by a Dynamic Theory of Classification. The first consolidated account of this Dynamic Theory was published in 1957 in the Prolegomena

Note

to Library Classification by Ranganathan. This was further refined after the establishment of DRTC at Banglore 1962, which provided facilities for deepening the Theory of Classification and making it more dynamic and applicable both to book classification and article classification. Consequentially, active work in the design of depth classification schedules for the classification of articles progressed. The need for such a dynamic theory is obivous as it only could provide guidelines for the development of subject classification in the future.

## 1.3 scriptive Theory of Library Classification

In the beginning there was no theory, only practice was followed. Practice gave rise to descriptive theory. Thus, the descriptive theory was the first stage in the development of library classification. This theory was able to meet the requirements of the universe of subjects (the totally of subjects comprising knowledge), as it existed at that time. The desciptive theory was based on the practices in vogue based on different schemes of classification then available. The descriptive theory distilled out of the contemporary - schemes, held its sway until the early 1950s. The schemes designed before the 1950s were based on the flair or natural gift of the designers and not on any objectively worked out theory of library classification.

This methods were empirical. The development of the descriptive theory is attributed to several stalwarts like Brown, Richardson, Hulme, Sayers, Bliss and Ranganathan. The period between 1898 and 1937 witnessed the genesis and development of this theory. These stalwarts, through their schemes and writings, enunciated certain principles of library classification. These principles and contributions of the personalities are briefly outlined in the following sub-sections.

### 1.3.1 J.D. Brown (1862-1914)

J.D. Brown was an English librarian, whose contribution to the General Theory of Library Classification. The first of these three was developed in 1894 jointly with J.H. Quinn and was known as Quinn- Brown scheme. This scheme did not make much impact. Three years later, in 1897. Brown independently brought out another scheme and called it adjustable classification. This scheme also proved inadequate even in those days. In the year 1906, Brown published the first edition of his subject classification, the scheme for which he is mostly known. Its second edition was brought out in 1914 and the third edited by J.D. Stewart in 1939.

Brown's subject classification was founded on the principle that every science and art spring from. Some definite source. In the order of things, there were first two factors, viz, matter and force. These, in turn, gave place to life,

Note

in course of time, led to the mind, which in turn gave birth to records. In addition to the shove principle, Brown also advocated - two other principles. The first of these two was his one place theory. According to this principle, each subject has only one place in the scheme inrespective of its aspects and numerous manifestations. For example, the subject of rose may be viewed from the viewpoints of botany, horticulture, history, geography, decoration, bibliography etc. The subject of rose, according to Brown, is concrete, while the various viewpoints represent its aspects. He was of the opinion that the interest of the scholar in 'rose' is constant, unlike that of the bibliographer whose interest is only occassional. He, therefore, preferred to place rose under one concrete or specific heading. It means that his arrangement of books was not discipline or library of congress ; but by topic. It was an experiment, which failed.

### 1.3.2 E.C. Richardson (1860-1939)

E.C. Richardson was the first librarian of Hartford Theological Seminary, USA, and later took over as librarian of the Princeton University Library. Richardson is regarded as the first classificationists to have made a systematic attempt to set down a theory of library classifications. In 1910, he published his book classification, Theoretical and Practical. It was the first textbook on classification, which later influenced W.C.B. Sayers. In the introduction to this work, he enumerated basic laws and principles meant to guide the work of designing a scheme of classification. These principles, called as criteria of classification, are as follows :

Classification should follow the order of things, classes should be arranged in historical sequence :

1. Division of classes should be minute.
2. Arrange things according to likeness and unlikeness.
3. Books are collected for use, they are administered for use, and hence it is the use, which is the motive behind classification.
4. A scheme of classification should be provided with a notation. The notation should be amenable to indefinite subdivisions preferably using a mixed symbol with decimal base and with mnemonic features.

### 1.3.3 E.W. Hulme (1859-1954)

Hulme was the librarian of the Patent office Library, London. In 1911-12, he published his book principles of Book Classification in the library Association Record. In the words of W.C.B. Sayers, the contribution of Hulme was "A valuable leadup to the more complete and satisfactory

Note

theories today." According to Hulme, all classifications could be arranged into two groups :

- Mechanical
- Philosophical

According to this categorisation, book classification is mechanical Hulme's principles of book classification are as follows :

1. Book classification is the plotting of areas pre-existing in literature and coincidence with a philosophical order is no guarantee of accuracy.
2. Book classification is mechanical assembly of material into classes.
3. The division and coordination of classes in literature is determined mainly upon formal and non. philosophical lines.
4. Classification should be based literary warrant. Hulme states that mechanical classification are left uncoordinated. But in book classification, systematic coordination of classes is introduced. His theory of literary warrant immensely attracted the attention of later classification. Literary warrant simply means that a subject cannot be listed in the scheme unless some literature has already appeared on it. Cr, the existing literature on a subject only justifies the inclusion of that subject in the scheme.

### 1.3.4 W.C.B. Sayers (1881-1960)

William Charles Berwick Sayers, an English librarian and teacher of S.R. Ranganathan made a remarkable contribution to the development of theory of classification. He is referred to as the first grammarian of library classification. He is responsible for interpreting and systematising the ideas of other theoreticians. He never designed any classification scheme, though, through his theory he has shown the way for others in the designing of classification schemes. His theory of book classification first appeared in 1915 under the title "Canons of Classification". He expanded the outline of the theory contained in this book in three other books viz. Grammar of classification, Introduction to Library Classification and Manual of Library Classification. It has now been revised by Rita Marcella and Robert Newton in 1994.

## 1.4 Sayors Canons of Classification

Sayors implified his theory of classification by stating 29 principles. He called them canons, meaning rules, regulation standard tests or criteria of classification. The 29 canons can be grouped under six categories as follows :

Note

| | |
|---|---|
| Canons of definition | 6 |
| Canons of divisions | 7 |
| Canons of terms | 4 |
| Canons of book classification | 4 |
| Canons of notation | 4 |
| Canons of book classificattion schemes | 5 |

**Definition :** Classification is a mental process by which things or ideas are grouped according to their likeness. The likeness which exists in the universe of things and in ideas is called characteristic in classification.

**Division :** Assembling things according to their degree of likeness and separating them according to their degree of unlikeness is the process of division. The chosen likeness or characteristic used to divide the given things may be natural or artificial. A natural characteristic is the inherent quality of a thing and hence, is responsible for its very existence.

**Terms :** A scheme of classification is a statement of knowledge using verbal terms. A term is the name for a class. It may be a word or a phrase. The terms should be unambiguous and unique with the same meaning whenever they are used in a scheme of classification.

**Book Classification :** A book classification is a device for the arrangement of books by subject or form in a logical order. It must be capable of admitting any new subject without dislocating the class of subjects already drawn. Book classification schemes must be equipped with :

- a generalia class
- form classes like poetry, fiction, drama etc.
- forms in which subjects are presented like theory, history, dictionary etc.
- a notation and
- an index.

**Notation :** A notation consists of signs of representing the class names in a scheme of classification. A notation should be brief, simple and flexible and have a mnemonic value.

**Book Classification Schemes :** A scheme of classification should provide column schedules in order of precedence of subjects. It is necessary to explain how to use the scheme. There should be machinery for the revision of the scheme to keep it up-to date accomodating new developments in the knowledge.

Note

### 1.4.1 H.E. Bliss (1870-1955)

He Evelyn Bliss devoted his entire active life to the intensive study of the art and science of classification. In addition to the articles, which he contributed in library journals, his theories and principles of classification were expanded in his first work, titled organization of knowledge and the system of science (1929). In this work, he formulated scientific, philosophical and logical grounds for the study of bibliographic classification. This work is regarded as one of the basic texts on the theory of organization of knowledge. He laid down the foundation for a relatively stable, scientifically acceptable and consistent scheme of classification. He also published another basic work on the theory of library classification titled Organisation of Knowledge in Libraries and the

Subject approach to Books (1933, 1939). His work helped in establishing librarianship as a scholarly discipline.

These two basic works convey to be us the fundamental principles of classification which Bliss later tried to apply in his system of Bibliographic Classification (BC) whose outline was first published in 1935.

The basic concepts of classification as expanded by Bliss may broadly be categorised as :

(1) Consensus (2) Subordination (3) Collocation (4) Alternative locations (5) Notation.

### 1.4.2 S.R. Ranganthan (1892-1972)

Right from 1924, S.R. Ranganthan had been developing his theory of library classification. In the first edition of Prolegomena to library classification (1937), he provided an integrated theory mainly descriptive and comparative of the practices in classification then in vogue. Ranganathan went ahead of those classificationists, mentioned in the preceeding sub-sections, by extending the principles put forward by them. He also provided the largest list of normative principles named by him as fundamental laws, postulates. Principles and Canons and evolved a special terminology, which is evident from the first edition of prolegomena. These rightly belong to stage-2 in the development of the General Theory of Library Classification. His theory is now synonymous with the general theory of library classification.

### 1.4.3 Classification Research Group (London)

After the Royal society scientific information conference in 1948 and on the suggestion of the eminent scientist J.D. Bernel, the classification Research group London was established in 1952. It is an unattached society of volunteers pursuing classification as an additional off the job work. They

Note

meet regularly in London. Upto 1996, they have held 308 meetings. Its founder members D.J. Foskett, Bernard Palmer. B.C. Vickery and A.J. Wells were greatly influenced by Ranganathan's work. They mostly came from special, industrial and academic libraries. Their deep and through study led them to believe that none of the published schemes provided a satisfactory system either in arrangement of depth of details. CRG accepted Ranganathan's method of facet analysis though it did not accept his views on the restriction of the number of categories to be five. They named their categories of Entities, Properties and Activities. Nevertheless, in Ranganathan's Dynamic theory, they found a sound base to be built up. They published their manifesto in the periodical Library Association Record which emphasis on the need for a faeceted classification as the basis of all method of information retrieval. B.C. Vickery wrote a small volume on the methods of constructing a faceted classification. Members of CRG designed many faceted classification schemes for specialised subjects ranging from diamond technology to soil science, music to education. Experience gained in designing such schemes led them to believe that the right approach should be to seek new principles for library classification. Though they never produced any new general classification system, their contributions to the development of classification techniques were many and innovating. A prominent member, Miss Barbara Kyle had a limited success in doing away with the necessity of main classes. Another member J.E.L. Farradane incorporated the idea of relational analysis with operators into the construction of a faceted classification scheme. Later, a group member developed the theory of Integrative levels, which arranged entities in an evolving aggregation of complexities. Their work received publicity and wide discussion in their International conference on classification research held at Darking , England in 1957.

## 1.5 Classification principles from the 19th Century to the Present

A library classification is a system of knowledge organization by which library resources are arranged and ordered systematically. Library classifications use a notational system that represents the order of topics in the classification and allows items to be stored in that order. Library classification systems group related materials together, typically arranged as a hierarchical tree structure. A different kind of classification system is called a faceted classification system, which allows the assignment of multiple classifications to an object, enabling the classifications to be ordered in many ways.

Library classification is an aspect of library and information science. It is distinct from scientific classification in that it has as its goal to provide a useful ordering of documents rather than a theoretical organization of

Note

knowledge. Although it has the practical purpose of creating a physical ordering of documents, it does generally attempt to adhere to accepted scientific knowledge. Library Classification helps to accommodate all the newly published literature in an already created order of arrangement in a filiatory sequence.

Library classifications were preceded by classifications used by bibliographers such as Conrad Gessner. The earliest library classification schemes organized books in broad subject categories. The earliest known library classification scheme is the Pinakes by Callimachus, a scholar at the Library of Alexandria during the third century BC. During the Renaissance and Reformation era, "Libraries were organized according to the whims or knowledge of individuals in charge." This changed the format in which various materials were classified. Some collections were classified by language and others by how they were printed.

After the printing revolution in the 16th century, the increase in available printed materials made such broad classifications unworkable, and more granular classification for library materials had to be developed in the nineteenth century.

In 1627 Gabriel Naude published a book called **Advice on Establishing a Library.** At the time, he was working in the private library of President Henri de Mesmes II. Mesmes had around 8,000 printed books and many more Greek, Latin and French written manuscripts. Although it was a private library, scholars with references could access it. The purpose of Advice on Establishing a Library was to identify rules for private book collectors to organize their collections in a more orderly way to increase the collection's usefulness and beauty. Naude developed a classification system based on seven different classes : theology, medicine, jurisprudence, history, philosophy, mathematics and the humanities. These seven classes would later be increased to twelve. **Advice on Establishing a Library** was about a private library, but within the same book, Naude encouraged the idea of public libraries open to all people regardless of their ability to pay for access to the collection. One of the most famous libraries that Naude helped improve was the Bibliotheque Mazarine in Paris. Naude spent ten years there as a librarian because of Naude's strong belief in free access to libraries to all people, the Bibliotheque Mazarine became the first public library in France around 1644.

Although libraries created order within their collections from as early as the fifth century BC, the Paris Bookseller's classification, developed in 1842 by Jacques Charles Brunet, is generally seen as the first of the modern book classifications. Brunet provided five major classes: theology, jurisprudence,

Note

sciences and arts, belles-lettres, and history. Classification can now be seen as a provider of subject access to information in a networked environment.

## 1.6 Mapping of Universe of Knowledge

The concept of the universe of knowledge is a metaphor of great importance in library classification theory. Different authors in different moments of history have provided examples and models on the this concept in relation to classification tree, maps, atlas and even constellation and multiverses are some of the examples that have been used in literature. A common concern of all classificationists reflecting on the universe of knowledge is the necessity of mapping or translating knowledge "out there" into concrete and logically developed classes that compound the classification system. This is one of the core aspects of knowledge organization system. The methodological approach to this mapping has been also considered as an epistemological question. In other cases, the sole mention of universality in classification has been politically questioned from critical and poststructuralists stances. The position is summarized by Olson as follows "The rejection of universals or absolutes in central to postmodernism and poststructuralism.

Knowledge is ever growing, changing and becoming ever new. New subjects constantly emerge, old subjects change their statu, structure and boundaries. There is no universal pattern of all knowledge that could be all things to all users. Therefore individual subjects change their structure and relationships between subjects can be seen in different ways. The prevailing philosophy, material culture, economic and technological needs, cosmic vision, sense of history and values held by the society influence the boundaries status and structure of the stock of knowledge in its possession. Every age and society has a distinct view of the structure of knowledge and relative value of its different fields. For example, in the middle ages theology was considered the queen of sciences and other subjects were valued according to their capacity to serve her. Natural sciences considered as an idle man's curiosity were not valued much then. Even during the times of Melvil Dewey in the late 19th century, philosophy and theology occupied a very respectable position. It is evident from the fact that 1/5th of space in the Dewey's universe of knowledge was occupied jointly by these two classes. Today the scales are tilted towards the study of natural sciences and their economic and technical implications. Sciences rule the roost and have been given the status of a national religion in some secular countries. Empirical and experimental modes of investigation are considered reliable methods to discover new. Knowledge and solve epistemological problems. These days authority, faith, speculation and intuition as sources of knowledge are looked upon with suspicion by the scientific community. thus the status a subject commands

Note

in a society is never constant. Some subjects once important and at the centre stage of knowledge are now relegated to a peripheral position. Once it was industrial production which was important. Today the environmental studies management, biotechnology and research on non-conventional sources of energy are gaining, importance in the information society, subjects like resources management, human/animal rights, information technology, biotechnology environmental protection are pervasive.

### 1.6.1 Classifications are Impermanent

Inter and intra-subject equations are always in a flux. For example, many subjects such as public health, international law, geopolitics, demography which had status of compound or complex subjects in the 6th edition of the Colon Classification gained the status of basic subject in the 7th edition of the scheme. Many similar examples can be given from the DDC. Thus knowledge structure is always perceived to be changing. an example of the scientific relevance of this matter is the recent special issue on "subject ontogeny and knowledge organization system change", published by the journal knowledge organization in 2016.

Classification essentially represents knowledge and it its map. It is a tool to analyze, organize and represent knowledge - thus a total for knowledge management. Therefore, as the knowledge advances by filling gaps we need new classification or adjust and modify the earlier ones. We need new classifications from time to time as the new knowledge develops. We have not only to revise classifications, but have to invent new classificatory techniques to organize new knowledge. S.R. Ranganathan commended the DDC as the best classification for the 19th century literature. At the same time he thought it quite unsuitable to classify 20th century knowledge especially of the post - world wars period. Thus 20th century needed new classification system and techniques and the 21st century may well need new classification particularly for organizing the Internet. Regarding the structure of knowledge we need to limit ourselves to one epoch within one culture to find some firm basis for a unified knowledge.

### 1.6.2 Principles for Mapping the Universe of Knowledge

D.W. Langside, a well-known English librarian, identifies four principles for mapping of the universe of knowledge. These, however are not mutually exclusive.

There are based on some schools of though, or some ideologically held principles. Earlier examples are Christian schemes of the middle ages. Latest examples is the Russian classification system BBK which had made Marxism - Leninism as the centre of the universe of knowledge. To some extent every scheme is based on some ideology. No classification scheme

Note

can be value free or independent of the time and culture of its origin. Every scheme by default is biased towards, the values and culture of the society of its origin. That is why the Dewey decimal classification has to be modified and adapted to classify African and Asian subjects.

### 1.6.3 Principle of Social Purpose

Vedic system of the division of knowledge into categories of Dharma, Arth, Kam and Moksh is an example of this principle. This is a broad classification which arranges knowledge in an order of decreasing current social utility and in the increasing potential for future use. Dharma is formation of the society, Artha for its governing Karma for its material progress and pleasure pursuits, while Moksha is to prepare for the life after death. This is a theoretical classification which has never been the basis of a library classification or any detailed knowledge classification Ranganathan was a bit influenced by it but he never used it as the basis of his Colon Classification.

### 1.6.3 Principles of Arrangement by Disciplines

A discipline is a major and cohesive chunk of knowledge formed by a single mode, or having the similar objects of study. Major contribution of Mevil Dewey was to adopt the division of knowledge by discipline. The DDC defines a discipline as "an organized field of study or branch of learning dealing with specific kinds of subjects and/or subjects considered from specific points in view."

Disciplines differentiate knowledge into number of logically distinct main classes characterized by the possession of cohesive types of concepts, structure and method of creation and verification of new knowledge. The division by discipline offers hope for a comparatively better, though by no means perfect, solution to the problems of information retrieval and to meet the needs of library users.

## 1.7 Mapping of Universe of Subjects in Different Classification Schemes in Different Classification Schemes

Just as in a geographical map, the surface of the earth or any part of it is drawn and various places, oceans, rivers, mountains are presented in it in relative position along with their boundaries and in it infinite units of past, present and future are presented. They have the ability to provide a place in their relative position, in the same way in a classification system, it is necessary to have the same characteristics as a geographical map to mark or display the position of all the subjects, sub-disciplines, divisions and subdivisions of the subject world. In various classification systems, through the marking process, the demarcation of all the subjects of the subject world or their position is ensured. The relative position of different subjects as

Note

well as different subject groups, their departments and sub-divisions are clearly shown. In a well-organized classification system, it is necessary to make provision for the newly created subjects to be arranged according to the principle of auxiliary less. The position of the subjects to be created in future should also be relative so that the librarian or the readers are not inconvenienced.

According to Dr. Ranganathan, in the classification system, the relative position is determined by analyzing the various units of the subject world first at the conceptual level, after which the designations are given to the different units at the verbal level, and finally the Notation Plane. The subject is presented by providing marks. Thus, the classification method is the presentation of the whole subject world or a part of it at the notation level, in which the broad areas of different subjects, sub-disciplines etc. have to be arranged in their relative positions.

The success of any classification method depends on the extent to which various subjects are given, the relative order in which the different subjects are arranged and the degree to which specificity and precision have been given to the systematic subjects and innovations. How much capacity is there to accommodate the created subjects and multifaceted subjects in-place.

**Different Classification Methods :** Classification is essential for arranging the documents of different subject literature in a suitable order in libraries. With a view to providing a permanent auxiliary sequence to the documents, librarians from time to time have created various classification systems. Some of the major and widely used classification methods are as follows

1. Dewey Decimal Classification Scheme- Melvil Dewey (1876)
2. Expansive Classification Scheme (1893)- C A Cutter
3. Universal Decimal Classification (UDC) (1905)
4. Library of Congress Classification (1904).
5. Subject Classification Method (1906) J.D. Brown
6. Colon Classification Method (1933) S. R. Ranganathan
7. Bibliographic Classification (1935) - H.E. Bliss
8. Riders International Classification (1961) -Fremont Rider)

Out of the above various methods, decimal classification system, colon classification system and universal decimal system are the most prevalent in India. The description of the mapping of the subject world in these classification systems is as follows.

Note

### 1.7.1 Mapping of Universe of Subject in Dowey Classification

1. **Historical description :** The pioneer of the Decimal Classification method, which is popularly known as DDC was Melvil Luis Kossuth Dewey.

   Dewey was born on December 10, 1851, at the Adams Center in the City of New York State. In 1872, he started working as an assistant at Amherst College, USA. He graduated from this college in 1874 and was appointed assistant librarian there. In the year 1874 itself, he made his own classification system to classify the books of his college library and started the work of classification of books.

   The first edition of Melville Dewey's Classification System was published in 1876 as "A Classification and Subject Index Rules for Cataloging and Arranging the books and pamphlets of a Library". This first edition had a total of 44 pages, of which 14 pages contained the introduction, 12 pages the table and 18 pages the index. The second edition of the Dewey Decimal System was published in 1855 under the title 'Decimal Classification and Relative Index,' containing 314 pages. It had 12 volumes published during Melville Dewey's lifetime. After his death in 1931, its 13th edition was published in 1932. The 16th edition of this classification system was published in two volumes for the first time in 1858. In the second volume of this 16th edition, the index was kept. The 17th edition was published in two volumes in 1865. It has a total of 2153 pages. Were. The first volume was of 1251 pages and schedules were given in it. The second volume was 899 pages and contained auxiliary schedules and relative indexes. The 18th edition, published in 1871, had three volumes and a total of 2684 pages. The first volume had 460 pages. In which tables were given. The second volume had 1168 pages in which schedules were given. The third volume was of 1056 pages in which the index was given. Its 19th edition was published again in 1879 in three volumes. It had 3000 pages. Introduction and Tables were given in the first section, Schedules in the second section and Relative Index in the third section. Its 20th edition was published in 1989 in four volumes. It has 3,309 pages. Tables in the first section, Schedules in the second and third sections and the relevant index and manual are given in the fourth section.

In the first section of D.D.C. 7 tables have been given for the expansion of class numbers.

Table 1 : Standard Subdivision

Table 2 : Geographical Area, Historical Periods, Persons

Note

Table 3A : Individual Authors

Table 3B : More than one Author

Table 4 : Subdivision of Individual Languages

Table 5 : Racial, Ethnic, National Group

Table 6 : Languages

Table 7 : Group of Persons

2. **Outline of Main Sections:** Schedules for classifying the Main Classes are given in Sections two and three of the 20th edition of the D.C.C. Dewey has divided the entire knowledge world into ten main sections.

| Class | Number Main Classes |
|---|---|
| 000 | Generalities |
| 100 | Philosophy and Psychology |
| 200 | Religion |
| 300 | Social Sciences |
| 400 | Language |
| 500 | Natural Sciences and Mathematics of Pure Science |
| 600 | Technology / Applied Science Sciences |
| 700 | The Arts |
| 800 | Literature |
| 900 | Geography & History |

Each of the above main classes is further divided into ten divisions (ten parts of each class). Each department is further divided into ten sections or sub-divisions. It provides that each section can be further divided into ten sub-sections and each subsection into ten sub-sections.

In decimal classification, the subject world is classified in a regular way. It is a systematic method based on ten main sections, one hundred divisions and one thousand sub-sections. In this method the classes move from general to very specific. Ten Schedules are given in detail for the ten main sections. In each schedule, the subjects related to the main category have been explained in detail in their natural language and in artificial language.

Note

The different types of tables given in it are associated with each schedule and give the same meaning to the marks, although D.D.C. An effort has been made to keep the relevant subjects side by side. For example, keeping religion with philosophy, history with geography, requiring an auxiliary order is according to. But separating the arts (700) from the social sciences (300) and placing science (500) and applied science (600) between them is contrary to the auxiliary order. Similarly, not keeping literature (800) close to language (100) is contrary to the auxiliary order. Similarly, separating Law (Law-340) and History (900) from Political Science (320) is contrary to the auxiliary order.

Arabic numerals have been used on the basis of the principle of decimal for the arrangement of notation in the decimal classification. It uses pure notation. The decimal point is used after the first three digits.

The relative index used in this method is arranged in chronological order. Synonyms of the subject and many other articles have been given in the posts mentioned in the schedules, which makes it easy and convenient for the class to find its related subject.

The Dewey Decimal Classification System is the oldest and most widely used classification system in the world. It is built on an enumerative basis. There are ready-made classes related to all subjects, simple and accurate notation is used and it has been made more expansive by the use of decimals. At present, new topics are being created as a result of new research, especially in the field of science and technology, and due to the use of pure and limited marks, this method is unable to incorporate new subjects. However, as a result of the increase in the subject world, this system has been continuously revised and its revised large editions are being published from time to time.

### 1.7.2 Colon Classification Scheme

CC was designed by the Father of library science in India Dr. S.R.Ranganathan, in 1933. Dr. Ranganathan was a teacher of Mathematics before coming to library science. After joining Library Science he got so much interested in it that he started thinking towards a new classification scheme. He was very disappointed with the classification schemes known as Colon Classification. CC is a faceted scheme, that is, it does not have schedules of classes or subjects as in other schemes, but each main class has one or more schedules of facets that can be added to the Main Class Number. The universe of the knowledge is divided into some subjects which are known as main classes.

During the life time of Dr. Ranganathan, six editions of CC were published, but after his death his work could not get the attention it deserved . The

Note

seventh edition of CC was published in 1987 without index, so it could not be used by libraries. The libraries which are still using CC are either using its sixth edition or incorporating the revisions, which were brought out in 1963.

The Scheme of Main classes given in CC 6th revised edition is reproduced here :

| Z- Generalia | LX Pharmacognosy |
|---|---|
| 1 Universe of Knowledge | M Useful Arts |
| 2 Library Science | A Spiritual Experience and Mysticism |
| 3 Book Science | MZ Hunanities and Social Sciences |
| 4 Journalism | MZA Humanities |
| A Natural Science | N Fine Arts |
| AZ Mathematical Sciences | NX Literature and Language |
| B Mathematics | O Literature |
| BZ Physical Sciences | P Linguistics |
| C Physics | Q Religion |
| D Engineering | R Philosophy |
| E Chemistry | S Psychology |
| F Technology | O Social Sciences |
| G Biology | T Education |
| H Geology | U Geography |
| HX Mining | V History |
| I Botany | W Political Science |
| J Agriculture | X Economics |
| K Zoology | Y Sociology |
| KX Animal Husbandry | YX Social Work |
| L Medicine | Z Law |

### 1.7.3 Other Classification Schemes

#### 1.7.3.1 Library of Congress Classification (1901)

The Library of Congress (LC) Classification scheme, as its name shows, was designed for the Library of Congress of USA. The scheme serves the requirements of the world's largest library LC and is updated regularly. Although, it is only a practical scheme without much theoretical base, but is quite popular particularly in libraries of USA because of easy availability

Note

of class numbers for most of the books. There is some influence of Cutter's Expansive Classification on main outline of classes of LC Classification.

### 1.7.3.2 Universal Decimal Classification (UDC) (1905)

UDC is the only classification scheme which was not designed for classification of books. It was designed for arranging the entries of universal bibliography at International Institute of Bibliography (Brussels). Two Belgians Paul Otlet and Henri La Fontaine revised the DDC (5th edition) to such an extent that this new scheme UDC was born. So UDC has a base of DDC or we can say that broadly the main classes and their divisions are similar in both the schemes. But, UDC has many more features for building a class number than DDC. Due to these features it is quite popular among special libraries all over the world.

There are some other classification schemes which are now mainly of historical importance. CA Cutter designed Expansive Classification (1891) in seven separate classifications of increasing details. First very broad, second more detailed and so on. James Duff Brown's Subject Classification (1906)was different in the way that it brought together all the material related to a particular topic, instead of using discipline oriented approach used by other schemes. Henry Evelyn Bliss's Bibliographic Classification (1935) is theoretically significant because it attempted to find an order of main classes that was based on scientific, and educational consensus. International Federation for Documentation (FID) published Broad System of Ordering(BSO) in 1978 for dealing with blocks of information. It was not meant forbook classification.

## 1.8 Categorization of Isolates

All the isolates of the world of knowledge can be divided into two parts:

1. Special Isolates, 2. Common Isolates

1. **Special Isolates :** Each major class of knowledge has its own distinct characteristics and isolates representing these characteristics are called special isolates.

   *Example :* isolates number 2 experiment in different forms in library science, botany

   - Library Science-2 (Public Library).
   - Botany-2 (Thallophyta)

2. **Common Isolates :** Common isolates are those isolates whether used with any main class are always in the same form. Dr. Ranganathan has defined Common isolates as :

Note

"an isolate idea denoted by the same isolate term and represented by the same isolate number, quite irrespective of the compound subject in which it occurs, or the basic subject with which the compound subject goes"

*Example :* a (Bibliography) whether used with any main class is always in the same form

1. Bibliography of Mathematics

   Ba

2. Bibliography of Library Science

   2a

### 1.8.1 Common Isolates in Colon Classification

Dr. Ranganathan, after studying the premade classification methods, developed this method in 1933 and he analyzed the common isolates and used them in his system. In the first edition of the two-point classification method, these were named as common sub-division and in the fourth edition, the preceding and subsequent common isolates were used. The seventh edition of this method provides more detail towards common isolates.

In the sixth edition of the two-point classification method, three types of common isolates have been used which are like :

1. Common isolates related to Class Numbers of the text.
2. Common isolates related to Book Number.
3. Common isolates related to Collection Number.

Here are the common isolates used for the taxonomy of the book are divided into two parts, which can be explained with the help of the following chart.

**Common Isolates**

| Anteriorising Common Isolates | Posteriorising Common Isolates |
|---|---|
| Before Space facet [S] | Personality |
| After Space facet [S] | Matter |
| After Time facet [T] | Energy |
| | Space |
| | Time |

### 1.8.2 Anteriorising Common Isolates

Common isolates of this type are added to the main square without a junction sign, and small roman letters are used for them.

Note

There are three types of these common isolates which are shown in the chart above.

**(i) Applicable before space facet only [S]**

Such previous common isolates are used before the (S) side i.e. such common isolates are used before the main class (P) side, The (M) side and (E) side are followed without any connecting sign and the following common isolates come in the series.

| Number | Term | Facet Formula |
|---|---|---|
| a | Bibliography | a [T] |
| c | Concordance | |
| d | Table | |
| e | Formula | |
| f | Atlas | f [T] |
| k | Cyclopaedia | k [P], [P2] |
| m | Periodicals | m [P], [P2] |
| n | Serial | n [p], [p2] |
| p | Conference Proceeding | p [P], [P2] |
| v | History | v [S], [T] |
| w | Biography | w [S], [T] |
| x | Collection | x [S], [T] |
| y1 | Programme of Instruction | |
| y2 | Syllabus | |
| y3 | Synopsis | |
| y4 | Scope | |
| y7 | Case Study | y7 [S], [T] |
| y8 | Digest | |

The above general equations can be explained with the help of the following examples.

1. Bibliography of Library Science 1967.

   2aN7

2. Concordance of Shakespeare

   011, 2J640

Note

3. Table of Economics

XD

4. Formula of Sodium

E11e

5. Geographical Atlas in 1967

UIN67

Normal isolates k, m. n. p. In v (P) side is obtained by Geographic Device (GD) and (P2) side is obtained by Chronological Device (CD) and these common isolates can be explained with the following examples:

1. International Encyclopaedia of Library Science 2k1
2. Botanica (Journal of Botany, Bombay 1950) Im44, N50
3. National Conference of Library Classification (Bombay 1985) 2:51p 44, N 85
4. History of Ayurveda in india upto 1950 LBV 44, N5

In the common isolates w (Biography), different aspects have been given for the biography of any person. (S) and ('T') sides will be used if the biographical taxonomy is to be made collectively and (P) the date of birth of that person will be used if the biographical taxonomy of an individual is to be prepared.

Ex. 1. Biography of Dr. S.R. Ranganathan (1892)

2wM92

Like the common isolates biography (w), the side formula for the common isolates collection (x = ) is given.

Ex. 1. Works of German Scientists (Youngest born in 1950)

Ax55, NS

The formula for I common isolates y7 (Case Study) is also the same as that of w and under y7, the biographies of political persons are made.

Ex. 1. Biography of Kennedy

V73y7M17

2. V73y7M17

2. Common isolates to be used only after the countryside (Applicable after space facet [S] :

Common isolates of this type are used after the countryside without any additive. That is, such common isolates are used only after the main square (FP) facet (M) facet and (E) facet and (S) side.

Note

There are two common isolates in this category which are as follows :

r =Administration report

s = Statistics (If periodical) 8 (T)

Ex. : 1. Administrative report of Secondary Education in India

T244r

2. Annual Statistics of Wheat Production in India

J382: 744s

It is to be noted here that this type of common isolates is used only for sequential data. It is done for this purpose only.

**(ii) Applicable only after Time facet [T] :** This type of common isolates are used only after Kal Paksha and the following common isolates have been placed in this category

S Statistics (if stray)

- t = Commission Report
- t4 Survey Report
- t5 = Plan
- t6 = Ideal
- V = Source Material
- v5 Literature
- v6 = Tradition
- v7 = Archaeology
- v8 = Archive

None of the common isolates in this category have been assigned a perinatal.

Ex. 1. Murders in Rajasthan (Statistics of 1992)

Y: 45112.4431N92s

2. Kothari Commission report, 1964

T4.44N64t

3. Survey report of Secondary Education in Rajasthan, 1962

T2,4437, N2t4

4. Plan of Higher Education In India, 1997.

T4,44, N97ts

Note

### 1.8.3 Posteriorising Common Isolates

Common isolates of this type are used with the additive sign and are usually used only after the (S) side. Common isolates of this type are of three types.

(i) Posteriorising personality common isolates.

(ii) Posteriorising matter common isolates,

(iii) Posteriorising energy common isolates.

**(i) PCI-Personality Common Isolates**

Common isolates of this type are usually used after the (S) facet and thus. The comma () is used for common isolates. The following common isolates fall into this category.

1. **ad af (PCI-Personality Common Isolates) :** Common isolates of this type are usually used after the (S) side and thus. The comma (.) is used for common isolates of r. The following common isolates fall into this category.

   CC-6

   - b = Profession
   - d = Institution
   - e = Educational
   - e2 = Lower
   - e4 = Higher
   - f = Investigation
   - f2 = Observational
   - f3 = Experimenting
   - f4 = Discussional
   - f7 = Yogic
   - g = Learned Society
   - h = Industrial Body
   - k = Commercial Body
   - w = Administrative Department of Government.

   The formula for such common isolates is also given. In this (P) facet is brought from Alphabetical device AD or Chronological device (CD) and (P2) and (E) side is brought from main class History.

   Ex.

   1. Library Profession in India
   2. 44, b

Note

3. Scientific Institution of India
   A. 44d
   Roorkee Engineering University
   D, 44, t4, R
4. Function of Vice-Chancellor of Sukhadia University, 1998
   T4, 44 d. S, 13 33N9s
5. Astronomical Observation at Rajasthan
   B9.4437, 2

2. **PCI-Matter Common Isolates :** In the sixth edition of the double-point classification, there is no purpose for such common isolates, but in its seventh edition, the Matter Common isolates is used in greater detail and for these the additive symbol semicolon (;) ) is used. Some Matter Common Isolates are:
   - a01 = Property
   - a02 = Existence
   - a03 = Function
   - a05 = Relation
   - a06 = Identity
   - a11 = Accuracy
   - a12 = Reliability
3. **PCI-Energy Common Isolates :** Common isolates of this category are used in conjunction with the main class or any side of the main class connecting sign is followed by a colon (;) and the following common isolates fall into this category:
   - b1 = Calculating
   - b2 = Designing
   - 20 Kku laxBu
   - b6 = Measuring
   - c1 = Weighing
   - f = Investigation
   - f2 = Observation
   - f3 = Experiment
   - f4 = Discussion
   - g = Criticism
   - p = Drafting

Note

- r = Reporting
- v = Surveying

These types of common isolates can be explained with the help of following examples:

1. Research on Secondary Education
   T2 : f
2. Observations on Education
   T : f2
3. Discussion of Higher Education
   T4 : f4
4. Critical evaluation of kothari Commission Report, 1964
   T4,44'M64t : g
5. Geographical Surveys
   U : v
6. Weighing of Cotton Yarn
   M71; 1 ; C1
7. Designing of Dresses
   M8 : b2

Energy common isolates according to CC-7

- a02 = By Action
- a012 = Decreasing
- a014 = Increasing

Kku laxBu 21

- a3 = Removing
- aS = Separating
- aM = Ranking
- aN = Classifying
- aP = Naming

### 1.8.4 Common Isolates in Universal Classification

The following two types of tables are described in the Universal Classification System:

1. Common Auxiliaries
2. Special Auxiliaries

Note

Here Common Auxiliaries can also be divided into the following parts

1. Common Auxiliaries of language
2. Common Auxiliaries of Form
3. Common Auxiliaries of Place
4. Common Auxiliaries of Species and Nationality
5. Common Auxiliaries of Time
6. Common Auxiliaries of Specific Approach

Of the 6 types of common Auxiliaries mentioned above, three can be included in common isolates.

1. Common Auxiliaries of Form
2. Common Auxiliaries of Place
3. Common Auxiliaries of Time

1. **Common Auxiliaries of Form :** These types of general helpers give specifics to the physical form of the documents and they can be used with any main class. It is used in parentheses ().

   Ex. : Dictionary of Economics 33 (03)

2. **Common Auxiliaries of Place :** Geographical aspects are represented by such common auxiliaries and they are shown in brackets ( ). Apart from the political departments, the provision of departments according to the direction of the tropical regions of the world, natural departments and polar regions etc.

   Ex. : Public Library Service in India

   (027.5 (540)

3. **Common Auxiliaries of Time :** Such general auxiliaries are used in double inverted commas (" ") and by these auxiliaries century, decade, year, months, hours and minutes etc. are displayed. So it can be said that common isolates are used in all major classification systems of the world.

### 1.8.5 Common Isolates in Dewey Decimal Classification

In DDC, common isolates have undergone several changes in both nomenclature and presentation. They were spelled out by different naives in different editions of DDC. The different names used so far are form divisions, common, subdivisons, view point numbers and standard subdivisions. From the seventeenth edition onwards they have been called standard subdivisons. A complete list of standard subdivision appears as Table 1 in volume 1 of the nineteenth edition of DDC. Types of Common Isolates.


The following are the different types of common isolates in DDC :

| | |
|---|---|
| –01 | Philosophy and Theory |
| –016 | Indexes |
| –02 | Miscellany |
| –22 | Illustrations and models |
| –028 | Techniques, procedures, appratus, equipment, material |
| –0285 | Data processing |
| –0288 | Maintenance and repair |
| –03 | Dictionaries, encyclopedias, concordances |
| –05 | Serial publications |
| –06 | Organisations and management |
| –07 | Study and teaching |
| –072 | Research |
| –08 | History and description of the subject among groups of persons. |
| –09 | History and geographical treatment. |

A note along with Table 1 states that the notation are never be used alone but may be used in required with any number from the schedules. It is thus, clear that the above numbers are not used independent of the core numbers from the subject schedules. Every number in the above table is preceded by a dash which merely shows that the number never stands alone. The dash is to be omitted when it is added to a core number taken from a subject schedule.

### 1.8.6 Space and Time Isolates in DDC

In DDC, the part of volume 1 is devoted to area numbers. Broadly, the division of geographical space is represented as under :

| | |
|---|---|
| –1 | Area, regions, places in |
| –2 | Persons regardless of area, region and place |
| –3 | The ancient world |
| –4 | Europe |
| –5 | Asia |
| –6 | Africa |

Note

–7 North America
–8 South America
–9 Other parts of the world

Having seen the space isolates in DDC, let us now examine the time isolates in it. Provision of time isolates is not as extensive in DDC as it is in UDC and CC. The time isolates in DDC are in the form of historical periods and given in Table 1 as part of the standard subdivisions. Their use is limited. The treatment of time in subject schedules is on the basis of a few very broad historical periods as under :

0901 — to 499 A.D.
0902 — 500-1499
0903 — Modern period, 1500
0904 — 20th century, 1900-1999
0905 — 21st century, 2000-2099

## Conclusion

In this chapter we have learned about knowledge and studied the classification of knowledge. Apart from this, we also learned about the positioning of the subject world in various classification systems. We have also studied the general method in this chapter so that we can easily acquire knowledge in any subject.

## Important Terms

1. Knowledge is the sum total of civilization preserved information.
2. As a result of curiosity, curiosity and new discoveries, there is a continuous development of man's ideas and along with it the knowledge is expanded.
3. Classification is necessary for arranging the documents of different subject literature in a suitable order in the library.
4. All the isolates ideas of the world of knowledge can be divided into two parts-
   (i) specific isolates
   (ii) Normal isolates.
5. Each class of knowledge has its own distinct characteristics and the units representing these characteristics are called specific units.
6. Common isolates are isolates that are not used with any of the main classes.

Note

## Exercises

### Very Short Answer Type Question

1. Define knowledge?
2. Why is classification of knowledge necessary?
3. How many types of isolates are there?

### Short Answer Type Questions

1. Explaining the origin and development of knowledge, give a brief description of the classification of knowledge?
2. Briefly describe the common isolates in a two point classification system?
3. What are indicator digits? Explain in brief.

### Long Answer Type Questions

1. Describing the representations of knowledge in classification systems, explain their differences.
2. Describe in detail classification of isolates.

Note

UNIT 2

# Mapping of Universe of Subjects

## 2.1 Introduction

In this chapter, we will study the colon classification system and also learn about why Ranganathan wanted to create this method. Along with this we will also study the determination of the position of the subject world in the Universal Decimal Classification System. Along with this some special classification scheme will also be known.

## 2.2 Mapping of Universe of Subjects in Colon Classification

**Historical Description:** Colon Classification System is the Analytico Synthetic Scheme. The originator of this system, Dr. S.R. Ranganathan, was originally a mathematician and he was a professor of mathematics before being appointed librarian at Madras University. Madras University sent him to London for training in library science, where he joined the School of Librarianship at the University of London for training. During his training, he visited several libraries in London and studied and analyzed the classification systems prevalent there, mainly the Library of Congress Classification System (LC) and the Decimal Classification System (DC). He came to know of many flaws and shortcomings in these systems. These systems are calculative and at times almost biased, in which, on the basis of past, present and inference, the categorization of future subjects has been done and this is contrary to the nature of development of the subject world, so Ranganathan was not satisfied with these prevailing methods. While thinking, the idea of creating a one-sided classification system came in his mind, in which only the main classes, some general units and some special units should be enumerated and according to the subject of the book, on the basis of some pre-formed rules and principles, by adding them together.

Note

Required class numbers can be prepared. Ranganathan constructed the colon classification method on the basis of this hypothesis. He named his method the colon or double point system because the subject matter he first made the provision of colon (:) to connect different sides. The first edition of this system was published in 1933. Its second edition was published in 1939. The idea of five fundamental categories was introduced in this edition. Its third edition was published in 1950. It presented the Facet Formula for each base class in the form of basic fundamental categories. In 1952, the fourth edition of the Colon Method was published. This version featured 4 different additive symbols for the basic categories - the coma (,) for the Matter side (;) for the Personality side, the semicolon for the Energy side, the colon (: ), dot (.) for the space side and (-) were used for the time side. The idea of Rounds and Levels was also introduced in this edition. The fifth edition was published in 1957, in which a new chapter of the schedules related to the phase relation was added. The sixth edition of this method was published in 1960. In 1963, a revised version of the sixth edition was published in which the Inverted comma was used for the Time side. In 1987, the seventh edition of the Colon Classification System was published, in which the number of main classes has increased to 105.

Three versions of the colon classification method published so far have been accepted. The first, second and third editions deal with the first version, in which the pattern was rigidly faceted. The fourth, fifth and sixth editions are referred to as the second version, in which the co-system is the free faceted method. The revised reprint of the sixth edition and the seventh edition are known to be its third versions. In these, this method is called a completely-freely-faced classification method. Its seventh edition has not yet been fully published. The plan is to publish the seventh edition in eight parts, out of which only four parts have been published as volume one.

**Outline of Main Sections:** Dr. Ranganathan wanted to create a classification system that would be able to place all possible subjects of past, present and future without causing any discrepancy between the classes located earlier in the schedule. Keeping this idea in mind, Dr. Ranganathan created the analytical-synthetic classification method. The number of major classes in the colon classification method exceeds all other methods. Dr. Ranganathan has divided the entire subject world into four zones, the details of which are as follows:

**The first zone (Zone-11):** Gonoralin approach has been kept under this zone. Small Roman alphabets have been used to represent these main classes.

- Generalia Bibliography - a
- General Encyclopaedia - k

Note

- General Periodical - m
- General Biography - w
- Indology - z44
- Gandhiana - zG
- Jainology - z(Q3)

**Second Zone (Zone -2):** Subjects are included which cannot be organized under the traditional main categories. Arabic Numerals have been used for these. Like :

- Universe of Knowledge - 1
- Library Science - 2
- Book Science - 3
- Journalism

**Third Zone (Zone-3):** Traditional main sections have been kept under this zone. To represent them, Roman Capital Alphabets have been used as notation. All capital letters A to Z are used. As:

- Natural Sciences - A
- Mathematics - B
- Physics - C
- Engineering - D etc.
- Law - Z

**Fourth Zone (Zone- 4):** Newly emerging Methodologies and Techniques are organized under this area which are usually from Common Isolates or Main Traditional Classes.

- Evaluation Technique (g)
- Conference Technique (p)
- Administration Report Technique (r)

Dr. S.R.Ranganathan, under his method, has divided the subject world into 26 main categories, out of which the first thirteen sciences and the last thirteen humanities and social scientists.

The Greek description delta 4 is kept for division between them. It has been designated as Mysticism and Spiritual Experience. The main classes are arranged in a very logical manner in the colon classification system. It is a sequential method. Applied Sciences are placed after Fundamental Sciences. Naturality is placed on the basis of increasing artificiality of social sciences moving towards materiality.

The notation of the colon classification system belongs to the category of

Note

mixed notation. In this method, along with the Schedules, many devices have also been used. For example Alphabetical Device, Chronological Device, Geo. Like Method (Geographical Device), Gap Device, Superimposition Device, Subject Method etc. Apart from these, there is special importance of System and Special, Phase Relations, Common Isolates etc.

Thus we can clearly say that it is a suitable and completely efficient classification method to cope with the multidimensional and multi-heterogeneous growth of the subject world. It is a completely independent one-sided domestic-analytical method based on unbiased regulatory principles and scientific principles, which is fully capable of incorporating the new developments of the world.

## 2.3 Mapping of Universe of Subject in Universal Decimal Classification Scheme

1. **Historical description:** Universal Decimal Classification (UDC) is an international extension and use of Decimal Classification (DC) which appeared in 1805 as a result of the completion of the first International Conference on Bibliography in the city of Brussels. happened. Special experts involved in this conference discussed ways to prepare a classified index of all published literature. Malville Dewey's decimal classification method was the most popular and successful method at that time. The Institut International de Bibliography modified the decimal classification method with the consent of Melville Dewey, led by two of his zealots and civil servants—Henri La Fontaine and Paul Otlet. After the adaptation, it and the entire international edition were published in French in 1905 under the name "Manual de Répertoire Bibliographique Universelle ". Schedules of classes in 4 parts in the name of Classification Decimale Universelle under the year 1927-33 in the language published with UDC The third edition of the German language was published in ten parts in the name of Decimal Classification in 1953, starting from 1934. Its fourth edition was started in the English language by the name of Universal Decimal Classification in 1936 by the British Standards Institute which was completed in 1968. Its abridged version was published in English in 1948, French and German versions have also been published. Its fifth edition in the French language was published in 1936, the sixth edition of the Japanese language in 1950 and the seventh edition in the Spanish language in 1955.

   In 1931 the name was changed to International Institute of Documentation (IID). In 1937 it was again renamed as "International Federation for Documentation-FID". Its headquarter is located in

Note

'The Hague' in Holland and it has a Secretary General appointed. To date, abbreviated editions of UDC have been published in 13 languages.

2. **Outline of the main classes:** The basis of this method is the decimal classification method, so the order convention and basic outline of UDC is according to DC. In this also, the arrangement of the main classes is like the decimal classification system and the entire subject world is divided into ten main classes, like:

   0. Generalities
   1. Philosophy, Metaphysics, Psychology, Justice, Ethics (Philosophy, Metaphysics, Psychology, Logic, Ethics)
   2. Religion
   3. Social Science, Economics, Law, Administration, Education (Social Science, Economics, Law, Administration, Education)
   4. Philosophy, Linguistic Languages
   5. Mathematics and Natural Science
   6. Applied Science, Medicine Fine Arts and Recreation
   7. Fine Arts and Recreation
   8. Literature
   9. Geography, Biography, History

   By adopting this basic structure, the principle of moving from general to specific has been used in the subdivision of these main classes. In 1963, class 4 was left vacant and language was arranged with class 8. In this system, a provision has been made to connect two sides or two different subjects with connecting marks.

   This method has the following connecting symbols

   - Connection of Non-consecutive numbers
   - Connection of Consecutive numbers
   - Relation
   - Language
   - Geographical division
   - Form sub-division
   - Place subdivision
   - Race and Nationality
   - Point of view
   - Special Analytical Sub-division
   - Alphabetical device


By using the above indicator signs, the marking base of DDC. is mixed. In this, synthesis has been used in detail. In this system, mixed classes can be formed by analyzing different aspects together, in this method there is also a system of relative indexing, so that classes can be found easily.

The Universal Decimal Classification method is being used by many libraries around the world. This system is especially used in Indexing and Abstracting services. It is very useful from the point of view of micro classification for special libraries.

## Conclusion

After studying the determination of the position of the subject world in the above mentioned three major classification methods, we come to the conclusion that:

1. Dewey Decimal Classification System (DDC) has adopted the computational method. In this, not only the basic subjects but also the status of the combined subjects has been determined on the basis of the computational system. In this, the infinite process of division goes on and every effort is made to determine the position of the definite subject (Mapping). In this, the principle of hierarchy has been adopted while dividing various subjects and determining their position. Although as a result of its computational information, its ability to give suitable place to new subjects has been limited, that is, it lacks flexibility, but its new editions are published.
2. In determining the subject-matter position in the colon classification method, there is a significant difference from D.D.C. and U.D.C. In its structure, similarity is given to the Faceted method. Due to its open-sided nature, it also mentions related isolated ideas under different aspects. By analyzing and synthesizing, there is a facility to provide proper place to the combined and mixed subjects according to their mutual relations, thus this method is able to give proper place to the continuous flow of the subject world.
3. In the administration of Universal Decimal Classification (UDC), the structure of DDC itself has mainly been adopted. In this also, by adopting the resultant method, the status of basic and combined subjects has been determined. In this, an attempt has been made to bring the co-related subjects together with the help of marking tips. The use of common and specific auxiliaries and special notation marks has given it a kind of partisan structure, in which the process of synthesis is visible. Therefore, it has got the flexibility to determine the status of new subjects to some extent.

Note

## 2.4 Dewey Decimal Classification

The first and most popular system in the field of library classification was created by Melville Dewey (1651–1931) in 1876. In this system, after dividing the entire knowledge world into ten main classes, Indo-Arabic numerals are represented by a decimal system of pure and simple notation. The inclusion of new subjects in the hierarchy becomes easy through the decimal notation system and the position of the documents in this is relative and not predetermined. Among other important features of science is its relative index and division of knowledge according to the specific discipline. It is available in three major versions: the full version of DDC-22 (2003), the abbreviated DDC-14 (2004) and the electronic version called Web Dewey. It is available in both full and abridged versions.

As stated earlier, DDC is the first general library classification scheme. It was designed bye Melvil Dewey in 1876. The first edition of DDC had only 44 pages including a 12 page schedule or scheme of subjects. Dewey used Indo-Arabic numerals as notation for the scheme. It is the most popular scheme in the world. Since its first edition which was published in 1876. DDC has been regularly revise to keep it upto date. The 23 edition of DDC has been published in May 2011. Its electronic version Web Dewey 2.0 has also been released simultaneously.

Every class number in DDC consists of three digits. For an aid to Eye after third digit, a dot inserted and then fourth, fifth and more digits are written. In the beginning of schedule three summaries of class numbers are given. First summary gives 10 main classes, second summary gives 100 divisions where each main class is divided into 10 branches and third summary gives 1000 sections where each divisions is further divided into 10 branches. The first summary of DDC 22 edition is reproduced here below :

| | |
|---|---|
| 000 | Computer science, Information and general works |
| 100 | Philosophy & Psychology |
| 200 | Religion |
| 300 | Social Sciences |
| 400 | Language |
| 500 | Science |
| 600 | Technology |
| 700 | Arts and recreation |
| 800 | Literature |
| 900 | History & geography |

Note

In the second summary each main class in further divided into 10 branches called divisions. The 10 divisions of main class 000 are given below.

| | |
|---|---|
| 000 | Computer Science, Knowledge & Systems |
| 010 | Bibliographies |
| 020 | Library & Information Sciences |
| 030 | Encyclopedias & book of facts |
| 040 | (Unassigned) |
| 050 | Magazines, journals & Serials |
| 060 | Associations, Organizations & museums |
| 070 | News media, journalism & publishing |
| 080 | Quotations |
| 090 | Manuscripts & rare books |

In the third summary each division in further divided into 10 sections. The sections of 020 Library & Information sciences are given below :

| | |
|---|---|
| 020 | Library & Information Sciences |
| 021 | Library relationships |
| 022 | Administration of physical plant |
| 023 | Personnel management |
| 024 | (unassigned) |
| 025 | Library operation |
| 026 | Libraries for specific subjects |
| 027 | General Libraries |
| 028 | Reading and use of other information media |
| 029 | (Unassigned) |

### 2.4.1 Use of Special Classification System

In some subject areas, more specific classification systems and subject-title lists are used than general ones. Example Ariadne, a useful guide to computer science resources on the web, 'ACM' computerizing classification system. Omni (Organized Medical Network Information) which is a catalog of Internet websites related to health and medical science subjects, visit NLM for viewing resources. (National Library of Medicine) Classification System NLM Provides access to subject headings and medical subject headings (MeSH).

- Edinburgh Engineering Virtual Library (EEVL) (www.eevl.ac.uk/paper1.html/)

Note

- Uses a specialized classification system developed by Engineering Information Inc.

### 2.4.2 Library of Congress Classification

The source of the Library of Congress Classification (LCC) is C.A. Cutter's "Expansive Classification" (1893) can be considered. The system consists of 21 classes divided into 29 parts, which are described in about 11,000 pages of 48 volumes. This system forms a series of coordinated schedules. Each main class has its own The distinct form is almost independent with geographic divisions and indexes. It is a purely enumerative classification, that is, the entire taxonomy is available for use and has little provision for synthesis. Its other prominent feature is its being based on the principle of literary warrants. All classes are based on Library of Congress Classification and developed for the same. There is no theoretical system of organization of the world of knowledge. This system is also available in CD, ROM version titled "Classification Plus". In this edition the subject titles of the Library of Congress (LCSH) are also given along with the schedules along with several additional entry accesses. It is very easy to use and update. Following are the salient features of the latest published version of this system.

Library of Congress, Subject Cataloguing Division Processing Services, L.C. Classification Outline Servo Version, Washington D.C. LC 1900. ISBN 0-8441-0881-8 Its main classes are as follows:-

- A - Miscellaneous
- B - philosophy
- C/F,G - History and Geography
- H/L - Social science
- M/P - Humanities: P Language and Literature
- Q/T - Science, technology and medicine
- U/V - Military science
- Z - Library and Information Science

The letters 1, O, W. X and Y are still empty. Each main category is redistributed in alphabetical order. for example:

- QA - Maths QD Chemistry
- QB - Astronomy QE Geology
- QC - Physics QH Biology

Each class is further divided into arithmetic structures. For example:

- QA - 1-937 - Mathematics
- 9-10-.3 - Mathematical logic

Note

- 75.5-76.95 - Computer science
- 101-141.8 - Basic Mathematics
- 150-272 - Algebra

The use of decimal expansion has recently been introduced to accommodate new subjects. LC The final form of square numbers has become quite complex due to the use of very mixed notation. an LC The class number may be extended beyond the author number and publication year of the cutter after subject classification. Due to which the class number becomes very long. For example:

- HV 1481.S82G34 - Old Age Pensions in Sweden, author P. Gahrten
- PS 1331. H4 1974 - Mark Twain, author A. Henderson (1974)
- PR 3629, 216 - The Rape of the Lock, editors Logie and McHenry
- QC 16. E5 - Albert Einstein, a physicist

### 2.4.3 Experiment, Modifications and Future

Although this method cannot be called scientific from the point of view of basic nature and other devices, it is counted among the most popular three systems in the library sector. This system is completely enumerative, so much so that form and location assistants have been extended in it. This system is very popular in the United States. Some libraries in other continents also use it. According to an estimate, 60 percent of US research libraries are L.C.C. make use of. Its class numbers are given in mark records and CIP data as well as in some cooperative and commercial language lists.

This system is continuously modified. Its most important feature is that the system is built and its main use is done in the same library. Ale. The changes made to the CC are published in its quarterly journal 'LC Classification: Additions and Changes'. The CD ROM version of this system is titled "Classification Plus". which is available at annual fee and includes L. C. Subject headings are also included. CD version is very easy and useful as it has a separate index created by the series indexing method of each section for its specific topic with the help of Key word, class number, Proximity Options and Boolean operator. Cumulative index will be published after the complete system is completed.BC2 is constantly being revised by J Mills with the help of members of CRG. This is a very effective system and it is expected to get its proper place in future.

### 2.4.4 Board System of Ordering

Apart from this, the Broad System of Ordering (BSO) is the only system that has been sponsored by an international organization since its inception. BSO has been approved by UNESCO as F.I.D./B.S. Developed as a

Note

switching language for its Unisyst program in 1971 in collaboration with O Panel. Switching language refers to a secondary language that performs the transformation from one indexing language to another. Therefore, it is not a classification system in the usual sense but acts as a tool due to the exchange of information between information retrieval systems.

### 2.4.5 Structure

B.S.O. the panel mainly completed the first draft in 1975 under the supervision of Eric J. Coates and G.a.Lloyd. Now the following version is available

Coates Eric J. and other Broad System of Ordering (BSO) three 8.5 inch discs one. In S. DOS Format: Hort UK B.S. O. Panel.

The main subject areas of BSO are as follows :

| | |
|---|---|
| 100 - Knowledge in general | 500 - Humanities and Social Sciences |
| 200 - Science & Technology (Integrated) | 600/890 - Technology |
| 300 - Life Science | 910 - Language and literature |
| 460 - Pedagogy | 940 - Art |
| 480 - Sports and games | 970 - Religion |

In this system, in addition to a total of about 6800 broad subsections, schedules of two general aspects, place and time are also given. It uses pure notation of Sahasranshu (divisible by 1000) and Shatabhisha (divisible by 100) numbers in the pattern of 3, 2, 3. In this only two joining marks hyphen (-) and comma () are used. This is a one-sided system based on modern principles of aspect analysis, in which there is a proper provision of synthesis. In this, Institutional Warrant has been used in place of Literary Warrant. It is very useful for traditional and multidisciplinary areas as well as for mission based subjects. For example

- Agricultural Biology - 360,20
- environmental pollution - 395,60
- Chemical Pollution (Where Chem-230) - 395,60-230
- Agricultural Economics - 360-580

Thus the hyphen (-) is used as a sign of a conditional relation. It also ensures to be used in places of A/Z marking.

### 2.4.6 Experiment

This system is available only in 3x5 floppy and, even more ironically, it is not used today as a switching language for the purpose for which it was

Note

created. However, in some libraries it is being used to organize books on the shelf. No further research is being done on this. However, due to the order in which the main sections are arranged and subject synthesis and easy marking, the usefulness of this system will remain.

## 2.5 Development of Cataloguing

The history of cataloguing is indeed a fascinating area of study. When looking at the early catalogues, we can find that their compilers were not always sure how to do deal with problems related to collections, anonymous and pseudonymous works, translations, and so on. There were however, some attempts at systematization of cataloguing methods, such as the cataloguing rules of Conrad Gesner (1548), Florianus Treflerus (1560), Andrew Maunsell (1595) and John Durie (1650). The first national cataloguing code was the French code of 1791 which made author entry compulsory, specified the form of catalogue card and included rules for accessioning and guiding. However, it is the events of the nineteenth century onwards, when modern cataloguing practice 'as shaped and patterned, that are of most interest to present-day -students. Developments during this period have taken place due to the contributions of several groups -of people. For example, there are the librarians who have contributed towards improved methods by their examples and published works; there is the influence wielded by the great institutions such as the Library of Congress and the British Museum Library (now the British Library); there are the organizations which have banded librarians together (such as IFLA) and the conferences which have enabled cataloguers to meet, discuss problems and exchange views; there are the codes produced in an attempt to standardize cataloguing practice; and there are the examples of many cataloguing services and networks. Some of the recent trends are discussed in this Unit.

### 2.5.1 AACR2 : The Developments

Attempts to devise a common cataloguing code for the UK and US began early this century, and resulted in the Anglo-American Cataloguing rules of 1908. Work on the revision of this code began in the 1930s, but the UK members of the Committee had to withdraw on the outbreak of the World War II in 1939. The American Library Association (ALA) continued the work and produced a new code, the ALA Rules, in 1949. Problems inherent to this code led to an International Conference, the International Conference on Cataloguing Principles (ICCP) held in Paris in 1961, which asserted the importance of drawing up a code of rules based on sound theoretical principles rather than on adhoc solutions to practical problems. This led to the production of the Anglo-American Cataloguing rules in 1967 in two different versions : the British text and the American text.

Note

Despite these tremendous developments one major problem still remained. If one wants to have the international exchange of bibliographic information, then it is necessary for international standards to exist that would ensure that the bibliographic records produced by different agencies in different places are compatible. One of the areas that needed urgent attention for standardization was the area of description of items. In 1969 IFLA sponsored an international meeting of cataloguing experts at Copenhagen which led, in 1971, to the publication of the International Standard Bibliographic Description (ISBD). This was further developed for various non-book materials, leading to the publication in 1977 of a generalized version of ISBD, called ISBD(G). ISBD laid down rules on a number of matters including the specific parts of an entry, the various data elements to form-an entry, preferred sources of information, and so on. Much of this was incorporated in the unified edition (the British and the American texts no longer existed) of AACR, called AACR2, published in 1978. Major revisions of AACR took place over the years and a new edition, called AACR2R, appeared in 1988, and further minor amendments were made in 1993.

The above discussion shows that there is now a very large measure of international agreement on the content of catalogue entries, and AACR2 (with all the revisions) has become the de facto international code for cataloguing. In recent times technological innovations and advances in IT have called for major revolution in cataloguing. With technological developments, newer media for storage and publication of information appeared ranging from microforms (microfilms, microfiche, etc.) to magnetic (floppy discs, tapes, etc.) and optical media (CD-ROM families). Consequently, the catalogue code had to incorporate new rules to handle these new media.

One of the most significant developments has been emergence of the Online Public Access Catalogue (OPAC) which is basically the gateway to a library's collection. This is an interface through which users can search a library's catalogue, and can also perform a number of other operations, such as borrower record checking, reservation of items, online renewal, and so on. Although OPACs made their first appearance in the mid-1970s, it was only at the beginning of the next decade that libraries in significant numbers began to switch from card to automated catalogues. Library management software packages usually come with an OPAC module which can be used as it is, though libraries prefer to tailor the standard OPAC module to suit their specific needs. The library can identify and design specific menus for staff and users, dialogues on specific fields, such as author, title, subject,

Note

ISBN, etc., with right truncation facilities, and keyword searching with Boolean search facilities. With the recent advent of graphical user interfaces, it has become possible to include a range of information retrieval facilities menus or as options on buttons or check boxes.

The other developments include various online cataloguing networks and services, sometimes called bibliographic utilities. These networks, or bibliographic utilities, such as OCLC (Online Computer Library Center), RLIN (Research Library Information Network), etc., provide cataloguing facilities to their member libraries_ These are basically central' cataloguing facilities whereby member libraries can download the catalogue entries of specific materials. The largest of these is the OCLC. OCLC's WorldCat (the OCLC Online Union Catalog) offer access to catalogue resources that no other single institutions possesses.

The most recent, and probably the most prominent, development of the era has been the Internet which has drastically changed the various aspects of our lives. Internet, and 'we of its most prominent services, the World Wide Web (WWW), has become the most widely used source of information for all concerned, and due to its exponential growth it has been extremely important to catalogue, and thereby exert some kind of bibliographic control on the information available on the Internet.

### 2.5.2 Online Public Access Catalogue (OPAC)

An online public access catalogue is a library catalogue accessed via a computer terminal for the benefit of library users. OPACs were introduced in the US in the late seventies and in the UK in early eighties. These have also been introduced in a few Indian Libraries. An OPAC provides the users online access to the library's catalogue allowing them to search and retrieve records from the online catalogue, and depending on the underlying library management software it also offers other facilities such as online reservation, borrower status checking, and so on. With the advent of the Internet, and more recently with the World Wide Web (WWW), most library OPACs can now be searched by remote users, anywhere in the world, who have an access to the Internet.

In an OPAC what the users see is an interface containing a menu, say, that works on several databases, the primary one being the library catalogue database. A catalogue database can simply be perceived as the machine-readable version of the card catalogue in a library. An online library catalogue is composed of a number of bibliographic records. A bibliographic record can be defined as a collection of data elements (author, title, publisher, ISBN, etc.) organized in a logical way which represents a bibliographic item. Thus,

Note

for any OPAC, there is a catalogue database working behind the scene that contains a number of bibliographic records, each record representing a bibliographic item. Each bibliographic record.represents a catalogue entry containing all. the bibliographic elements such as author, title, publisher, place of publication, year of publication, ISBN, and so on.

Now, the question is how does an OPAC work? In other words, how does the system tell the user whether the library has a book written by a given author, say by G.G. Chowdhury, and then if it is available, how does the user get the details, *i.e.*, how does the system-retrieve it, and moreover, where does the cataloguing part come in this picture? In order to have an idea about the processes involved, we can simply divide an OPAL into four major components: the user interface, the main master file of bibliographic records, the index, and the display/print format.

(i) The user interface allows a user to interact with the system. Typically it offers various search options, such as author search, title search, keyword search, subject search, and so on. Although this is controlled by the underlying library management software, some cataloguing knowledge has gone into its design. The various search keys, author, title, keyword, subject heading, etc., are the parameters used by the users- for searching a library catalogue and these have also been included as search keys in the online, catalogue.

(ii) The main catalogue database file is something like the card catalogues, each card being converted into a bibliographic record. However, in order for the computer (actually the software) to understand and manipulate the various data elements that form a bibliographic record, some measures are taken at the design stage of the catalogue database. Like in any database structure, a catalogue record is composed of a number of fields, and subfields, each field (or subfield) representing a data element (name of an author, title of a book, and so on).

Now, you may ask, how many fields (and subfields) should we create for a bibliographic database and how can it be standardized so that any two databases in a country, or anywhere in the world, will have the same structure. Here, the skills of a cataloguer, the standards and AACR2 play an important role. While designing the database, an important step is concerned with assigning the appropriate tags and attributes to each field (and subfield). This is extremely important and needs high level of standardization because this does not only ensure standardization (every catalogue database using the same fields and attributes), but this also helps computer (the library

Note

management software) to handle and manipulate the data appropriately. Here, comes the standards such as MARC and its various versions :

(i) **The index** file, called the inverted file, which may be conceived as a back-of-the-book index file containing all the index terms (here, authors name, title words or full titles, keywords, subject headings, and so on) and appropriate pointers pointing to appropriate record(s) in the catalogue database. Here again some cataloguing knowledge is required, such as which fields are to be indexed, how subject headings are to be assigned and indexed, whether any vocabulary control tool will be used at the indexing stage or not, and so on.

(ii) **The display/print format** directly relates to cataloguing and to AACR2. Catalogue records, in a catalogue database file, are not arranged in the same way as we see on the catalogue card cabinets. In fact, each bibliographic record contains raw data associated with various attributes, and when we want to view it on the screen, or want to print it, the underlying software picks up the data and presents it according to an AACR2 entry. This job is performed by the display format (which could also be used as print format) which is designed in accordance with the rules for generating entries according to AACR2.

### 2.5.3 Standards For Cataloguing

The effective exchange of bibliographic data between agencies can be accomplished only if the records conform in respect of : (1) the structure, (2) the content designators and (3) the data element definitions.

The Universal acceptance of the ISO 2709 record structure (International Organization for Standardization. ISO 2709-1981 Documentation: format for bibliographic information interchange on magnetic tape. As a basis for exchange formats, addresses the first issue and thus has enormously benefited the information community. There are equivalent national standards too.

The second component relates to the content designators which are represented in most bibliographic formats by tags, indicators and subfield codes. There are several bibliographic formats; many of them come from the MARC family, viz. MARC, UKMARC, INDIMARC, and' so on.

Since the early 1970s, an extended family of more than 0MARC formats has grown up, whose paths have diverged owing to the different national cataloguing practices and requirements. The differences in these formats mean that editing is required before the records can be exchanged. One solution to this incompatibility was to develop an international MARC format, known as Universal MARC (UNIMARC) which would accept

Note

records in one national format and provide a means whereby records in one national format could be converted first into UNIMARC and then into another national format. This would in effect require each national agency to write two conversion programs : one to convert from the national MARC format into UNIMARC, and another to convert from UNIMARC into the national format, instead of having to write a separate program for each other national format with which the records were to be exchanged.

The first version of UNIMARC was published by the International Federation of Library Associations and Institutions (IFLA) in 1977. This was followed by a second edition in 1980 and by the UNIMARC Handbook in 1983. All of these focused primarily on the cataloguing of monographs and serials and reflected international progress towards the standardization of bibliographic information as represented by ISBDs.

Other agencies have also bibliographic formats to suit specific needs, such as CCF (Common Communication Format) from UNESCO, MIBIS (Microcomputer-based Bibliographic Information Systems) from IDRC (International Development Research Centre, Ottawa, Canada), and so on. The third component in standardization, the form and content of data elements, is controlled by the catalogue code used as well as by the way the different data elements, as prescribed by the rules, are divided up and separately identified by the format. This is taken care of by the software used to develop the OPAC. Several additions and improvements took place in MARC and other standard formats. This has happened to cope with the various new types and formats of information sources, and also to cope with the changing needs of users. For example, CCF came out as two sets of standards, one for bibliographic records (called CCF/B) and the other for factual records (called CCF/F). Recent additions to MARC have also included fields for creating hyperlinks among various documents. All the MARC formats, such as USMARC, UKMARC, UNIMARC, etc., have included a field (Field 656 - Electronic location and access) to hold data on electronic resources; for example, such as Web pages (discussed in the next section). A number of fields have also been added to link records; for example, in the USMARC fields 760, 762, 765, 767, 770, 772, 773, 775, 776, 777, 780, 785, 786, and 787 have been added as various linking entry fields. Similarly field 536 has been created to hold finding information note, field 561 for holding information on ownership and custodial history, and so on.

### 2.5.4 Cataloguing Networks

One of the major objectives of forming cataloguing networks is to perform the cataloguing operations centrally by one library/organization

Note

while other participating members can use the catalogue. They are also called bibliographic utilities for they provide bibliographic utility services such as catalogues of materials to the members. Library of Congress, in Washington, D.C. (USA), since 1901, has made printed cards available containing cataloguing data, classification numbers (Library of Congress as well as Dewey Decimal Classification number) and subject headings. Mechanization was introduced as early as in 1970s to meet the growing demands for this service among libraries. Currently cards are printed from MARC records on demand and the Library of Congress (LC) also operates an `alert service'. A participating library establishes a profile consisting of the subject areas that are of interest. Then, when LC enters a bibliographic record that matches this profile into its MARC database, a catalogue card is automatically produced and is sent to the library.

The United Kingdom was half a century behind the United States in providing a centralized service similar to that of LC. Though the British National Bibliography (BNB) started in 1950, a printed card service, called BNB cards, began in 1956. These BNB cards could be obtained by libraries for producing their own catalogue records. A MARC distribution service began in UK in 1969 and a large number of UK libraries and other institutions use BLMARC records for local cataloguing and housekeeping operations.

The United States led the world in setting up cataloguing networks and bibliographic utilities such as OCLC (Online Computer Library Center), RLIN (Research Libraries Information Network) and WLN (Western Library Network). Among these OCLC is the pioneer and is the most highly used bibliographic utility service.

OCLC is a nonprofit computer service and research organization whose network and services link more than 27,000 libraries in 64 countries and territories. OCLC services help libraries locate, acquire, catalogue, access, and lend library materials. OCLC, the world's largest library information network, offers the following services that help libraries further access to information and reduce information costs :

1. WorldCat (the OCLC Online Union Catalog). The world's largest and most comprehensive bibliographic database. Libraries use the WorldCat® database and OCLC's computerized telecommunications network to process materials and share information. WorldCat offers libraries and their users resources that no single library could provide :
   (i) 38 million records in eight bibliographic formats representing 400 languages with holdings information;

Note

(ii) Growth by approximately 2 million bibliographic records each year; and

(iii) Easy searching from FirstSearch, EPIC, and other services

2. The OCLC system. Extensive online cataloging, interlibrary loan, union list, and selection services. OCLC FirstSearch® service. Online reference service that provides flexible searching and subject access to over 70 databases for end users.
3. OCLC FirstSearch Electronic Collections Online® service. Provides remote access to large collections of journals through the Web. Now also available through the First Search Web interface.
4. OCLC Access Services. Telecommunications and stand alone systems that facilitate online and offline cataloging, resource sharing, reference, and selection services.

Now a question comes how can a library benefit from a networked facility. Let us take the example of the services offered by the largest cataloguing network or bibliographic quality service in the world, viz., OCLC. OCLC offers you a range of cataloging services that help you increase productivity and reduce cost. From batch services to CD-ROM, you can transfer records quickly, efficiently, and cost-effectively to your card catalog or local system. The OCLC Cataloging service is the largest and busiest online cataloging system in the world. It handles over one billion transactions a year and supports a range of services that help libraries increase productivity and reduce costs. And with OCLC Cataloguing, only one library needs to originally catalogue an item. OCLC offers a range of cataloguing options to meet your library's needs. You can :

(i) have bibliographic records automatically delivered at the time your vendors ship books using the OCLC PromptCat service;

(ii) catalog online, where the hit rate is more than 95 percent for English materials;

(iii) input your own original records;

(iv) export records from OCLC and edit them on your. local system;

(v) tapeload your cataloguing that is done outside OCLC's cataloguing systems;

(vi) have OCLC convert your older cataloguing to machine-readable form;

(vii) move records to an offline environment to edit and export and then batch your edits and holdings back with OCLC;

Note

(viii) cataloguing Micro Enhancer software;

(ix) have OCLC staff catalogue current materials, foreign language materials, or backlogs;

(x) seamlessly link selectors, acquisitions department, book vendors, and your local system to an online, ready for filing; and

(xi) order custom-printed cataloguing cards according to your specifications. Cards arrive at your library - ready for filing.

### 2.5.5 Cataloguing with WorldCat

Cataloguing with WorldCat gives you an incomparable knowledge resource. You can :

(i) use high-quality records to improve search results;

(ii) catalogue all materials;

(iii) reduce staff time for copy cataloguing with OCLC PromptCat service; and

(iv) custom order machine-readable tapes or catalogue cards.

The OCLC Catalogue service is an interactive online system that provides access to WorldCat (the OCLC Online Union Catalogue) and facilitates shared cataloguing. OCLC Cataloguing increases productivity and reduce cataloguing costs. Participating libraries retrieve and use bibliographic records contributed by other libraries and enter records for items not already cataloged. OCLC Cataloguing provides :

(i) global record sharing through WorldCat;

(ii) keyword and browsable searching;

(iii) full-screen record editing and cut, copy, and paste capabilities;

(iv) a high hit rate-find over 95 percent of the English records you need; and

(v) access to name and subject authority records.

Electronic export capabilities to local systems in North America join OCLC through their OCLC-affiliated Regional Networks. Libraries outside North America receive OCLC services through OCLC Asia Pacific Services, OCLC Europe, OCLC Latin America and the Caribbean, or via international.

One of the major problem facing libraries in many countries, including those in India, relates to what is known as retrospective conversion or Retrocon in short. This relates to a basic question as to how libraries can convert their existing manual catalogues (card) to machine-readable catalogues. Thus, simply speaking, retrocon is a process that is used in libraries to convert

Note

existing manual catalogues to machine-readable catalogues. There are several options available and each one has its own merits and demerits. The simplest way is to manually convert (i.e., keying-in), each and every catalogue record which is impossible for a large library collection where several hundred thousand catalogue entries are available in card forms, because this will be very time-consuming and expensive process. One alternative way could be to decide a cut-off year for the retrocon process. Indeed, whichever method is used, each library has a cut-off date, beyond which catalogue records are not available in machinereadable forms; users looking for materials beyond certain date have to consult card catalogues. The cut-off year is decided by the library concerned on the basis of a study that determines beyond which year library materials are scarcely used.

Another option to retrocon is to scan the card catalogues (provided they are available in printed/ neatly typed form) and then to use a custom-built program to convert the scanned files to machine-readable catalogues. This method has now found useful in certain situations, and several companies have taken this approach to do retroconversion job for libraries on a contract basis. The major difficulty of this process is that the success rate depends on the nature and quality of the scanning process. If the cards are not neatly printed or typed, then the quality of the scanned image and the corresponding ASCII file will be poor, resulting in number of errors in the catalogue entries. Hence, a significant amount of editorial time will be needed to ensure the quality of the machine-readable catalogue, Yet another, and the most commonly followed method, is to use the services of bibliographic utilities like OCLC. As already mentioned earlier in this section, one can avail the WorldCat service form OCLC for downloading catalogue data of all the materials held in a library. This process is Much easy and does not need any local expertise. However, there are two major problems. First, this may prove to be expensive, particularly for libraries in India, because the price to be paid for such service is calculated in the US Dollar and thus may be quite a big amount in local currency. The second problem may be that local materials and non-English language materials may not be available in OCLC's databases.

Each of the above options has its own merits and demerits. One option, that has been chosen by many, is to set up a centralized cataloguing agency, mostly at the national library level, that produces catalogue entries for all types of materials and makes available to other libraries in the country in machine-readable form. This approach helps off-set the huge cost and other resource requirements of retroconversion.

Note

### 2.5.6 New Challenges : Cataloguing In The Internet and Digital Library Environment

Internet can simply be defined as a network of networks; it is an amalgamation of interrelated computer networks permitting electronic communications on a global scale. Internet has now changed the way we generate and use information. World Wide Web (WWW) is one of the most recent, and most widely used, service of the Internet that mimics a spider web of information, spread all over the world. The volume and variety of information available on the Internet and WWW is increasing at a very rapid rate. While on the one hand the Internet has made it easy for us to have access to information available anywhere in the world from our own computer, it has made our life difficult too. Due to the rapid growth and lack of appropriate mechanism for organization and control of information, the Internet has become, what some people call, an anarchy. It is difficult to know what is available on the Internet at a given instance. Libraries have exercised control over their collection, however huge it may be, through appropriate cataloguing of their resources. Now the question is, can similar measures be taken for Internet resources? In other words, how can the information resources available on the Internet be catalogued? This has become a new challenge for librarians. It can be conceived of an information retrieval tool that is characterized by its use of hypertext links to information sitting on various servers all over the world. Several attempts have begun and a number of research projects are underway to resolve various issues of cataloguing materials available on the Internet and the World Wide Web.

Cataloguing of Internet resources can be done by using a set of tools called ROADS. ROADS stands for Resource Organization And Discovery in Subject-based services. It is a set of software tools designed to help create information gateways on the Internet. At its simplest, an information gateway can be just a simple list of links maintained by hand. The information gateway approach allows a subject based list of links to be created and maintained. ROADS allows Internet resources to be fully described, allowing the end user to judge their relevancy before trying to access them and it is properly organized. The organization and description of resources adds value to an information gateway. ROADS allows you to create a database of resource descriptions (also called templates). This is done via World Wide Web forms that you have to fill-in. This form is a data entry form containing a number of fields, such as title field, keyword field, etc. ROADS contains tools which will automatically create classified listing of the resources and a search mechanism.

Note

Digital library is another new concept that can be defined as a library having all the information available in digital form. The digital library system, unlike traditional library systems, reside in a building or at any particular location. It is distributed virtually all over the world, and user would get information, as needed, on the screen of his computer. In addition to traditional text-based information, data accessible through the digital library system will include :

(i) non-text information (such as photographs, drawings, illustrations, art, etc.);

(ii) streams of numeric data (such as satellite information, chronological data, meteorological data, etc.);

(iii) digitized sound and moving visual images; multidimensional representation of forms (e.g., holograms); and

(iv) the capacity to integrate these data into new representations drawn from many different sources.

Given this wide variety of information sources and their rapidly growing volume, it is not unlikely that a user will easily get drowned in the ocean of digital information unless the information is properly organized. One way to organize the information in the digital library environment is to properly catalogue all the materials, and to provide easy and multiple access points to the user so that he can easily retrieve the required, and only the required, information. Several large-scale digital library research projects have been undertaken all over the world to bring out appropriate solutions to cataloguing and other problems. OCLC, the Library of Congress, the British Library, etc., are again taking a leading role in these research activities.

## Conclusion

In this chapter we have studied the colon classification system and also learned about the formulation of this method by Ranganathan. Along with this, we studied about the Universal Decimal Classification System and also learned about the use of some specific classification systems.

## Important Terms

1. The three volumes published in Colon Classification System are divided into three versions.
2. Universal Decimal Classification is an international extension of the Decimal Classification which originated in 1895 due to the conclusion of the International Conference on Bibliography in the city of Brussels.

Note

3. In the Dewey Decimal Classification, after dividing the entire knowledge world into ten main classes, the Indo-Arabic number is represented by a pure and simple notation system.
4. In some subject areas specific classification systems and subject indexing systems have been used more than the general ones.

## Exercises

### Very Short Answer Type Questions

1. What is the Colon Classification Method?
2. What do you understand about the Universal Decimal Classification System?
3. What do you understand by Dewey Decimal Classification?

### Short Answer Type Questions

1. Explain the use of a specific classification system.
2. Describe the important trends of research in classification systems.
3. Explain the structure of the Brand System of Ordering.

### Long Answer Type Questions

1. Describe in detail the problem and possibility of classification of Internet resources.
2. Describe in detail the determination of the position of the subject world in the Universal Decimal Classification System.

Note

UNIT 3

# Contribution of Catalouging

## 3.1 Introduction

In this chapter we will study the subject of catalouging and study the contribution of scholars like cutter, Lubetzky, Ranganathan etc. in cataloging and know how they have expressed their views on the concept of cataloging. Apart from this we will also study abstraction and indexing and bibliography list and also the layout of catalog codes in books.

## 3.2 Contributions of Different Scholars in The Field of Catalouging

1. **Ranganathan's Contribution :** Ranganathan's chief technical contribution to library science were in classification and indexing theory. His Colon Classification introduced a system that is widely used in research libraries around the world and that has affected the evolution of such older systems as the Dewey Decimal classification. Later he devised the technique of 'Chain indexing' for deriving subject-index entries. Other works of his included classified catalogue code (1934), Prohegomena to Library classification (1937), Theory of the Library catalogue (1938), Elements of Library Classification (1945), Classification and International Documentation (1948), Classification and Communication (1951), and Headings and Canons (1955). His five laws of Library Science (1931) was widely accepted as a definitive statement of the ideal of Library Science. He also drafted Plans for a national and several state Library Systems, founded and edited several journals, and was active in numerous professional associations. Dr. Ranganathan has told this through his chain process, it can be used in all forms of classifier and indexer.

Note

The components of the composite subject are automatically arranged in the chain in the order of time, place, action, substance and personality (PMEST). This is the contribution of Ranganathan.

2. **Cutter's Contribution :** Cutter was the first to consider the concept of a specific subject in his 'Rules for dictionary catalouge' in 1876. He entered a work under its specific subject heading, and not under the subject heading of the class that is under that subject heading. Cutter also made rules for the entry of multilingual titles. He said that the entry of the name of a composite subject should be used with its first word. Reverse the order of the word group where certain other words are definitely more important or are used alone in the meaning of the full name. But here the cutters do not explain how to obtain that one component or element of a plural compound subject that is important.

3. **Lubetzky's Contribution :** Lubetzky published three books that influenced the discipline of cataloging and that are still influential in area of information technology. Librarianship is particular and information science in general had not been revolutionized as much since the likes of Antonio Panizzi, Charles Ammi Cutter or Paul Otlet. Cataloging rules and principles and principles of cataloging as well as several periodical articles solidified Lubetzky as one of the most significant influences in his field. He developed a rationalized approach to catalog code design, one that is even more relevant today as current cataloging principles are revisited and revised for a digital environment.

   His unfinished book, Code of cataloging Rules unfinished draft was the basis for modern cataloging adopted by the first international conference on cataloging principles (1961) held in Paris, France called the 'Paris principles'. The code which eventually emerged from the conference became the basis for cataloging practice in the 20 century. In 1967 these concepts were encoded in the Anglo American Cataloging rules, which were used in US, Canadian and U.K. Libraries.

   Lubetzky is credited with renewing an emphasis on the work in library catalogs. This had been a feature of books catalogs of the early 19 century but was not carried through to the card catalog. Whereas Charles Ammi Cutter, a later 19 century influence on cataloging had not distinguished between the idea of a 'book' and the idea of the 'work' in formulating his objectives. Lubetzky contrasts the two ideas, bringing back into play Antony panizzi's original emphasis on

Note

the relationship between a title and all of the different editions of that title might exist. Lubetzky's idea was that the relationships among all the editions and variations of a given work and the author of that work, in all variations of the author's name must be established and brought together so they can be found in one place. All the works of a given author, in all their editions, should be linked together.

## 3.3 Union Catalogue

Union List refers to the combined list of different libraries of a region. This list can be of international, national, regional and local level and has been defined by various scientists as follows:

1. According to Dr. Ranganathan, "The list of all the books of two or more libraries in which the names of all those libraries are given where copies of those books can be obtained and a union list can include all types of books Or it may be limited to any one type of texts among them.
2. According to Larson, "A catalogue which lists all or part of the textual material of two or more libraries in a sequence.

### 3.3.1 Objectives and Use of Catalogue

The following are the purposes and utility of the Union List:

1. Union List is used in Book Selection and Reference Service.
2. The maximum consumption of the Union List is done in the endowment library.
3. With its help, necessary bibliographical information is obtained.
4. This list helps in research and preparation of reference and bibliographic reference list.
5. It tries to prevent duplication of book selection and acquisition in the libraries of an area.

### 3.3.2 Compilation of Union Catalogue

While preparing the Union List, the objectives of the plan should be decided and after that the following things should be discussed -

1. **Participating Libraries :** While preparing the Union List, it should be decided that the level of the Union List should be of International, National, Regional or Local level. The preparation of the national level union list is a very difficult task and the preparation of the local union list is easy because under this, information about the text material stored in a regional library can be obtained immediately.

Note

Therefore, it was considered appropriate that regional union lists should be prepared.

2. **Participating Material :** By the material included here, we mean the type of text material i.e. books, periodical publications and non-book text material etc. Therefore, while preparing the union lists, it should be decided what type of material will be included? In making lists of periodical publications and books, compilation expenditure is less and it is very easy to keep it updated.
3. **Types of Entries :** While preparing the Union List, the type of entries and the list codes should be expounded.
   Their rules have been included in Ranganathan's List Samhita. According to Ranganathan, the taxonomy of the main entry is marked, then it gives complete information related to that document and similarly the creation of auxiliary entries should also be done according to the list code.
4. **Limitation of Details of Entry :** This point is also very important in the preparation of the Union List. In Ranganathan's catalog code, the following information is included in the main entry of the journals:
   - Class Number
   - Title
   - Sponsoring Body
   - Volumes
   - Cumulative Indexes

   Similarly, in the supporting entries also information should be given according to the list code.
5. **Arrangement of Entries :** There are generally three types of arrangement of entries in union lists:
   (i) Purely Alphabetical Arrangement
   (ii) The main entries have been arranged in alphabetical order and chronological index of subjects as in 'Regional Union Catalog of Scientific Serials' compiled by INSDOC.
   (iii) The main entry should have an alphabetical arrangement and an alphabetical index of the subject.
6. **Physical Form :** According to Ranganathan, the physical form should be sheet form. Although book format is also useful for magazines because it can be easily seen and used easily, for books only leaflet format should be adopted because they are kept at a central place.

Note

7. **Methods of Compilation & Revision:** Following are the methods of compilation and revision.
   (i) The information should be collected and compiled personally by visiting the included libraries.
   (ii) The lists of included libraries should be taken and assembled.
   (iii) The largest library among the included libraries should circulate the list of its collections to other included libraries where each library can make use of it.

## 3.4 Layout and Rules for the Union Catalogues of Books

In order to try to bring uniformity in the cataloging practice, it is necessary that in the list codes of different levels, the normative principles at the international, national, regional and local level, the rules related to the writing style and the general opinion should be followed in relation to the layout of the library catalogue code.

This layout is based on the Practice for Layout of Libraries catalog by the Indian Standards Institute. The indexing and abstracting magazine is being laid here:

### 3.4.1 Parts of the Catalogue Code

A catalogue code consists of the following parts, which are as follows:

1. Preliminary material
2. Accessories
3. General Entry
4. Specific Entry
5. Lists other than library lists
6. Documents other than traditional books and periodicals.

1. **Preliminary Material:** Information regarding Guiding Principles, terminology and personal name should be given under the introductory material for indexing and abstracting journals.

   Here the sub-formula rules and principles come under the directive principle, which are as follows:

- Canons of Cataloguing:
  1. Canons of Ascertainability
  2. Canon of Prepotence
  3. Canon of Individualization
  4. Canon of Sought Heading

Note

5. Canon of Context
6. Canon of Currency
7. Canon of Permanence
8. Conon of consistence
9. Canon of Recall Value

- General Normative Principle & Laws :
  1. Five Laws of Library Science
  2. Law of Interpretation
  3. Law of Impartiality
  4. Law of Symmetry
  5. Law of Parsimony
  6. Principle of Local Variation
  7. Principle of Osmosis.

Similarly, under the terminology, the definition of technical words and the artistic side of the words used are given and the structure and functions of simple names and compound names are kept under personal names and their information is given as follows:

Entry elements, Secondary elements and Indivalation elements.

Example: RANGANATHAN (SR) (1892-1972).

2. **Auxiliaries :** It includes the following elements:
   - **(i) Determination of Authorship :** Under this, rules should be kept to clarify the difficulties faced in the problems of determination.
   - **(ii) Rendering :** Under the hypothesis of various names, the rules related to the names of individuals, population rules, books and books, and the rules related to the hypothesis of artificial names should be given.
   - **(ii) Recording :** For the listing of entries under Recording, rules related to the language and script related system, writing style, printing, abbreviation and the aspects of arrangement or arrangement of entries should also be given.
   - **(iii) General Entry :** General entry includes general subject entry and cross reference entry and the list of subject headings or the use of chain process should be given in detail.
   - **(iv) Specific Entry :** Under the special entries, entries of a volume,

Note

general book, multi-section book and indexing and abstracting journals are included and in this the main entry the rules for CIE and BIE entries should also be given.

(v) **List of Documents other than Library Catalogue :** Apart from the library list, rules should be given in these catalogue codes for the creation of other types of lists. In this way, the rules for the preparation, arrangement and layout of joint lists of books and periodicals, national lists should be given.

(vi) **Documents other than traditional books and periodicals :** In addition to traditional books and periodicals, subjects should also be given in the catalogue code for other specific types of text materials. It includes the following study material.

- Handwritten text
- Reprograp
- Tape record
- Cinema reel
- Pictures
- Instrumental recording of accompanying documents and visual objects.

## 3.5 Periodicals

A periodical is defined as "a publication issued in successive points usually at regular intervals and as a rule intended to be continued indefinitely" by A.L.A. Glossary of Library Terms. On the other hand, Ranganathan defines a periodical publication as a Conventional Document of kind I with the following attributes :

1. A volume, or a small group of volumes of it, is intended to be published or completed normally once in a year, though irregularity in interval is not reeded out.
2. Each successive volume or periodical group of volumes, is usually distinguished by the year of publication and/or by a number belonging to system of simple or complex ordinal number. Such a number is usually called the volume number.
3. The intention may be to continue the publication of the volumes for ever, though not actually carried out.
4. They intention may be to continue the same title in all the volumes, though this may not actually be carried out.

Note

### 3.5.1 Indexing

Presently, as a result of the discovery of the latest technologies in paper making and printing art, the publication of text materials has started increasing at a rapid pace. Nowadays, text materials are being published in many languages in the world every day, as a result of which readers are taking interest in the text material related to their specific subject instead of taking interest in any one specific book or writing. People engaged in research have also started turning away from books like articles published in current journals. If we look at the history of the library, in the past, on demand by the readers in the library, overall marks of the text materials were given, but due to this the readers had to waste a lot of time in searching for their useful articles. Therefore, in order to solve this problem at present, the articles published in the periodical magazines are now listed in alphabetical order according to the subject. Due to which it has been facilitated in the selection of useful topics without wasting our time in turning the pages of the magazine. The method of listing published articles in this way is called indexing and list is called Index.

#### 3.5.1.1 Definition

According to the Oxford English Dictionary – *"An alphabetical list, placed (usually) at the end of a book of names, subjects etc. occuring in the with indication of the places in which they occur".* —Oxford English Dictionary

NS. According to C. Vickery – *"An index is a working tool designed to help the user to find way about the mass of documented information in a given subject field"* —B. C.Vickery

#### 3.5.1.2 Functions of Index

The index plays an important role in making the user familiar with its text material due to the influx of literature, some of its functions are as follows:

1. It works to overcome the problem of language barrier due to special headings given in the user's language.
2. Provides information about the progress / development of subjects happening in all fields globally.
3. Through the index, the literature published in any field is collected and kept in one place.
4. An information acts as a source or tool between the source and the user.
5. Text material is easily available in this, which is very important for a searcher.

Note

6. It contains complete details of the text material up to the date of publication of the index.
7. It is a process through which it provides service to the readers.

### 3.5.1.3 Indexing Language

Method of obtaining documents on the basis of specific headings and submission of entries under subject headings has been prevalent among catalogers and indexers during the time of the C.A. Cutter. In presenting specific subject indexes, many difficulties have to be faced due to multifaceted development of subjects, inter-disciplinary research and specialization and abundance of fine documents. Therefore, more importance is being given to the need of coordinating or combining many terms in indexing. Coordination of index terms is considered an important aspect of indexing language. In the language of indexing, coordination actually means coordination of concepts.

Indexing language is the method of naming a subject for indexing. Like any language, it also has two parts. Vocabulary and syntax is a list of vocabulary terms that are included in the method. Syntax index shows the relationship between terms in a phrase.

There are two types of indexing languages. 1. Derived term form and 2. Assigned - term. In the derived term system all the terms are obtained from the document itself, whereas in the assigned -word system, the indexer creates the index term. This method is better by which the correct and proper subject title is assigned after getting the appropriate specific subject of the document. Thus author indexes, text indexes, citation indexes and indexes in natural language are derived term systems, whereas all indexing languages in which terminology is controlled, such as subject title lists, theories and classification systems, are assigned term systems. Derived-word systems are almost clerical and can be mechanized very easily. On the other hand, the assigned term systems introduced are intellectual, so they take more time and money. The structure in indexing generates and makes the index unusable

1. **Natural Language :** Natural language has its own importance. Its terminology can be updated and new concepts can be incorporated. It has its own syntax and rules of grammar. Which are able to show the true meaning of the specific subject. But natural language suffers from problems due to its antonyms and synonyms.
2. **Artificial language :** Themes are representative of title classification methods and theoretic indexing languages. Due to the controlled vocabulary, they are well organized and because there is a provision

Note

of reference entries see and see more in these, there is no problem of synonyms and antonyms in them. Nowadays, several components are needed to represent the specific themes of micro-documents. So these components are sorted according to the syntax rules of the indexing method. For example, in systems such as chain process, precision popsy, etc., the subject of the subject document is formulated on the basis of their own syntax rules.

When a topic is formatted in a selected index language, it provides only one access to the index file. In order to provide access to other constituent terms, the index language provides for mechanization of permutation of constituent terms. In permutation, each constituent term is used one after the other as a leading term. Since these forward terms are related to other terms, so in each condition the forward terms are also added with the referenced terms.

### 3.5.1.4 Features of Indexing Language

1. Such language is developed and used for a specific purpose. The exchange of ideas, along with being a means of service purpose, has to fulfill some specific purposes as well.
2. Generally the word used in indexing language has a very precise and clear meaning.
3. Words having the same meaning and same pronunciation are removed.
4. As far as possible, an effort should be made to establish a unitary relationship between a language concept and terms. When terms have that kind of relationship, those terms can serve as a better directory of information in an index file.
5. Although it is not possible that we should completely eliminate words having the same meaning. The indexing language controls the occurrence of scattering or explosions in the indexing subjects when a term is selected to represent a particular concept from among the many words of similar meaning available in a natural language.

### 3.5.1.5 Recall and Precision Devices in Indexing Languages

Due to the availability of literature in abundance, an ordinary researcher has to face great difficulty in finding information. Many times a situation also arises that the user does not even understand what the related information can be about him and where it can be available.

Note

Indexing language is the only medium that tells the users where the information related to a particular topic is located. In what form? Keeping all these important facts in mind, the importance of indexing language has been understood.

The following are the advantages of creating or using an indexing language:

1. By this it is possible to list the literature scattered in books and magazines on a specified subject at one place and it is very useful for researchers.
2. Indexing language serves as a communication between the source of information and the recipient of information.
3. It is estimated that progress is being made in various fields on the world level.
4. The topic title is in the language of the users, so the language barrier is also removed.
5. Previously published literature is easily surveyed.
6. Saves time.

### 3.5.1.6 Recall and Precision devices in Indexing Language

The effectiveness and quality of any indexing system depends on a few factors. Two important variables known as 'Recall and Precision' play an important role in measuring the effectiveness of information retrieval.

They are also related to each other. If a user wants 'High Recall' then he accepts 'Low Precision' and similarly the user who wants 'High Precision' will have to get 'Low Recall'. In this way they are highly related to each other. That is to say that the recovery effectiveness in the indexing system can be measured by 'Recall Ratio' and 'Precision Ratio'. With respect to 'Query' we can say that not all related documents can be retrieved but only some parts of them are retrieved. Generally not all retrieved documents are related. Similarly, the same situation prevails for unrelated documents. This situation can be understood as follows:

| Documents | Retrieved | Non-Retrieved | Total |
|---|---|---|---|
| Relevant | a | b | a + b |
| Non-relevant | a + c | b + d | a + b + c + d |

In this way, based on the above discussion, the recall ratio and precision ratio can be obtained.

1. **Recall Ratio :** The retrieval of related documents by a system is known as Recall. The number of documents retrieved from all the related documents through the indexing system is called Recall

Note

Ratio. In mathematical interpretation it is described as a/a+b. Let's calculate the Recall Ratio in the following way:

$$\frac{\text{Number of relevant items retrieved}}{\text{Total number of relevant items}} \times 100\%$$

or

$$\frac{a}{a+b} \times 100\%$$

For example, we assume that the total number of related documents in a file is 100 and the index is able to retrieve only 75 documents out of them and the remaining 25 documents are not able to be retrieved, then we can get its Recall like this

$$75/75 + 25 \times 100 = 1\%$$

or

$$75/100 \times 100/175\%$$

2. **Precision Ratio :** The ability to restrict unrelated documents in the system is called Precision Ratio. In mathematical form, finding out the number of documents retrieved out of all the related documents out of all the related documents i.e. out of the total number of all related documents is called Precision Ratio. In mathematical interpretation it is described as a/a + c.

Let us calculate the Precision Ratio in this way.

$$\frac{\text{Number of relevant items retrieved}}{\text{Total number of relevant items}} \times 100\%$$

or

$$\frac{a}{a+c} \times 100\%$$

For example, we assume that the number of documents retrieved in a system is 150, out of which 75 are related, then its Precision Ratio will be as follows.

$$75/100 \times 100\ /1\% \quad \text{or} \quad 50\%$$

But it is very difficult to find the total number of all the documents from the collection. For this reason it is very difficult to calculate the recall value in an indexing system. For this reason Recall and Precision Value are very important in any indexing method. In this way we can say what should be found in the indexing system for the effectiveness of retrieval and what documents should be restricted? Recall and Precision prove to be very helpful in telling this.

Note

Books for any reader who comes to the library to get their information, the material related to it is always available in the collection of library. Precision Ratio plays an important role in making the desired information available to the reader in a broad way.

3. **The Recall :** Precision Curve: In this we can divide the classes of four types of

   (i) A U B (ii) A - B (iii) B - A (iv) (A∩B)'

   All these are in the form of a matrix.

| | Retrieved | Non-retrieved | Total |
|---|---|---|---|
| Relevant | AB | A – B | A |
| Non-relevant | B – A / B | (A ∩ B)'/B | A/L |

On the basis of the above format, we can say that some writers understand this format better than other formats.

When we try to develop Recall by putting this form into practice then we get that the importance of Precision also goes on increasing and so on. If we unfortunately describe group A to a very small extent. It is probably used only under experimental conditions.

In this way we see that the quality and success of any indexing system has an important role. They represent the average performance of any indexing system and not the totality and their results also depend on the suitability of the conditions.

## 3.6 Major Indexing Services

1. **Biological and Agricultural Index :** This is a monthly indexing journal, which has been published since 1964 by the New York publishing body H.W. Wilson Company. This is a proper subject index published in the English language. In which about 190 periodicals are indexed. The compositions of biology, agricultural science and allied sciences are indexed in the same chronological order by author and subject. The index of book reviews is given in separate sections according to the authors. Each entry is complete from the point of view of bibliographic information.
2. **British Technology Index :** Now published as 'Current Technology Index'. It has been organized and published by the Library Association, London since 1962. In a year, 11 issues and an annual volume are published. This magazine is a great tool for exploration and recovery of specialized and innovative information available in the fields of business knowledge, engineering and manufacturing technology.


Published in around 300 British magazines, it helps in establishing contact with technical creations. Its indexing methodology is based on the Word Index and Trade Name Indexing (CATNI) method which is accessible separately.

3. **Index Medics :** Index Medicus has been compiled and published by the National Library of Medicine, Washington since 1960. It indexes 150,000 works annually. It contains more than 2000 medical journals. An index is prepared under author and subject. Monthly scores are published as cumulative volumes under the heading 'Cumulated Index Medicus'. Reports published in foreign languages are translated into English and the respective language is indicated in parentheses. Subject headings are given for each entry as required.

4. **Chemical Titles :** This fortnightly index journal has been published by the American Society since 1960. In this, articles of about 700 periodical journals of chemistry are indexed. Each article, author and keynote is indexed using the Indexing Method (KWIC). Its main objective is to keep chemists, researchers and chemical engineers fully aware of the specific and current information related to their subject area.

5. **Index India :** This is a quarterly documentation index which was started by the University of Rajasthan in 1967. It contains material from about 1000 selected publications. Efforts are being made to include in it the published literature related to India in the field of social sciences and humanities. It does not include legal reports and material related to the firm and the bank.

   In its main categorical part the entries are arranged in chronological-categorical order, while in the second part the three indexes alphabetically, author and review indexes are arranged alphabetically letter by letter. A list of current periodicals included is given at the beginning of each issue.

6. **Library Literature :** This bi-monthly index journal is being published by the H. W. Wilson Company of New York since 1934. Its annual volumes and bi-annual cumulative volumes are also published, about 225 journals published from different parts of the world available in the field of library and information science are included in it. In addition, on a selective basis, journals related to other subjects included in Wilson's other indexes are also included in it. Indexing of booklets, films, film plates, microcards, microfilms, dissertations, research articles, etc., related to library and information

Note

science is also done. Microcard and microfilm related material that has been published in print is omitted. In this, the author, title and subject entries are arranged chronologically in the same sequence.

7. **Citation Indexing :** The basis of this new method lies in the failure of classification systems. who have been unable to cope with the abundant emergence of knowledge. In which subject the original research works related to these indexes are cited from the author and the report. All contributions are grouped under each citation. The cited works of each author are placed in chronological order. The credit for indexing the citation goes to Eugene Garfield of the Philadelphia-based Institute for Scientific Information. This organization has been publishing 'Science Citation Index' since 1963 and 'Social Sciences Citation Index' since 1973'. Now the publication of the Humanities Citation Index has also started.

The following entry points have been provided to search for content in the Citation Index

1. Population Index Division
2. Quote index
3. Source index, and
4. Permutation Words, Subject Index -

All those items which become its specialty under the name of an organization and all those publications which are published in the said period are included in the population index.

The citation authors are arranged in chronological order in the citation index and the year cited is arranged chronologically.

The source index is an alphabetically arranged author index of 5,40,000 articles and editorial material, each entry containing the author/s name, language sign, article count,

Name of the magazine, volume number, issue number, number of pages, year of publication, number of references and Full address of the first author. This index is helpful in the use of citation and permutation word subject index.

Permutation word subject index is created by computer. The important words of the story of each article are changed in order and added to each other word of the story with which it is included in the belief of any article. The exploratory technique is used to follow the instructions quoted in the related works.

Note

## 3.7 National Bibliographies

The term 'bibliographic record' is a relatively new term. The use of this term is associated with library computerization. In general terms, a bibliographic record is the sum of the elements and fields used for the description, identification or retrieval of the physical fields (such as publications, documents, etc.) of the subject matter of any information.

In the words of Gradle and Hopkinson, "A collection of data elements that are logically arranged and represent a bibliographic subject is called a bibliographic record." Gradle and Hopkins have defined the textual subject as:

"Any document of human communication, book publication or any other transcript or group of documents or part of any document etc. may be taken as an entry.

Any bibliographic record format that is valid for data exchange must have the following three basic components:

1. **Physical Structure :** Rules for the adjustment of data on the computer storage medium.
2. **Content Designators :** Codes to identify the various data elements of a record (i.e. author, title, date of opening of journals, etc.)
3. **Content :** The content of any record that is bound by the rules for the creation of various data elements, they are associated with the content designator. Data components from any exchange format are identified by a code that is not only in the form of content, but that the record is suitable for use by any other agency.

Exchange of any bibliographic data between different agencies is possible only if their data set includes the appropriate three components. In any information system the records that make up the database are mainly available in various formats. Some of the main formats are

1. A format to be incorporated into the record system.
2. Such a form which is suitable for storage for a long period.
3. A format that helps in easy information retrieval.
4. The form in which the record may be displayed.

Along with this, it is also necessary that if two or more organizations want to exchange data among themselves, then they should use a common standard. Using a uniform standard is a must for any networked world. Every organization should be able to convert its records to a standard that allows for smooth and smooth exchange of data. Although at present there

Note

are many national and international standard exchange formats available and there is a lot of uniformity in them, due to their not being uniform, data exchange from one format to another is not easy. Keeping this fact in mind, ISBD was developed by IFLA.

## 3.8 ISBD-International Standard for Bibliographic Description

IFLA held a meeting of the ICCP in 1967 to standardize bibliographic description and as a result step towards the preparation of a standard for various bibliographic descriptions. In this sequence, different I.S.B.D was developed. In 1974 the first I.S. B.D was published which was related to the monograph. After this other I.S. B D. was formed as follows. These are discussed further.

| Type Publication | Year | Detail |
|---|---|---|
| ISBD (a) | 1980, 1991 (Second Edition) | For publication of old monographs (Anti Quarian) |
| ISBD (cf) | 1990 | |
| ISBD (cm) | 1997 | For details about the map/atlas |
| | 1987 (Revised Edition) (Cartographic) | |
| ISBD (CR) | 2002 | For serial and other serial resources |
| ISBD (ER) | 1997 | For electronic resources |

## 3.9 Abstract: Definition and Concept

When the details given in a document are presented in a concise form, then that brief description is called Abstract. Through this summary, information is provided to the reader about what has been highlighted in the related document. In this way the process of preparing a summary of documents, articles, etc. is called abstracting.

Various definitions of essence have been given by various scientists, some of the major definitions are as follows:

According to Kallison, "A summary is a summary of all the important things that are given in the original document according to the order of the information given in an original primary article. The author presents all those points in a systematic and concise form in his own language.'

According to Dr. Ranganathan, a summary is a summary of the essentials of the subject described in the article of a general journal, which is done by

Note

a professional other than the author and the details of the availability of documents are indicated in it.

According to Mazel, "The abstract presents the purpose, scope, and conclusion of the document. It is mainly produced for users who do not have access to the original document.

According to Ashworth, "A summary is a collection of information." It is an abbreviation with which the context of the search is included. Often nowadays this short summary form is indicative of information contained in an article in a magazine.

In the context of these definitions, it can be said that a brief statement of the characteristics of the essential and important subject matter of an original document, articles of a particular type of magazine is called a summary. It provides information to the user to determine whether the said document can meet their requirement. In the summary, the related document or the bibliographical details of the publication are also mentioned which helps in getting the document. Therefore, in a nutshell, it can be said that the summary is that it is a concise representation of the content of a document in a style similar to that document.

### 3.9.1 Abstract: Scope And Importance

The details of the scope and importance of the abstract can be presented as follows:

1. Literature is published in more than 80 languages of the subject. Normally the user is fluent in two or three languages. Therefore, he cannot use documents that have been published in other languages. All readers can use the document to summarize such documents from the respective language.
2. The reader's time can be saved by the abstraction service.
3. The efficiency of the indexing service can be increased by this service.
4. It plays an important role in literary search.
5. Casual Consciousness Service (CAS) can be expanded.
6. Repetition of research and research work can be prevented.

### 3.9.2 Types Of Abstract

A concise, accurate and succinct presentation of any document is called a Abstract. Abstracts are of many types due to their purposes and forms. These can be described as follows:

1. Indicative Abstract
2. Informative Abstract

Note

3. Informative-indicative Abstract
4. Critical Abstract
5. Slauted Abstract
6. Auto Abstract
7. Telegraphic Abstract
8. Graphic Abstract

1. **Indicative Abstract :** In this type of abstract, the content of the document is presented briefly. It mentions the area of the document. These abstracts do not provide much information about the document but point to them. Such abstract documents are helpful for the purpose of selection, but they cannot be used in place of the original document. These abstracts are considered very useful for providing Current Information Services (CAS) i.e. indicative abstracts are important in guiding readers.
2. **Informative Abstract :** Under this abstract, quantitative and informative information of the document is provided. This abstract can also be used in place of the original document. That is why it is also called Abstract Synopsis. The idea of the original document, principles, methods used, results and techniques are presented in this way in an informative summary. This is done so that the user does not even need the original document. An average of 200 to 300 words are used in this abstract.
3. **Informative :** Indicative Abstract: This type of abstract includes the characteristics of both informative and indicative abstracts.
4. **Critical Abstract :** In the critical abstract, the original document is presented in a critical way. Through this, not only the summary of the subject given in the document is presented, but the evaluation of that work is also done. The depth of the work is clarified by this summary.
5. **Slauted Abstract :** In the present era, some documents are useful and informative for experts and scientists of different fields due to the development and impact of interdisciplinary subjects.

   In this summary, more importance is given to any one part of the document.
6. **Auto Abstract :** This abstract is created with the help of a computer. Under the automatic summary, the important words that come in the document are selected through the computer and when these

Note

words come again and again, the sentences are sequenced and the summary is made.

7. **Telegraphic Abstract :** In this abstract, the key words are selected based on the original document. In this abstract these words are not arranged in any definite order.
8. **Graphic Abstract :** Graphic abstract is used for communication in chemistry. These abstracts are created in the language of Molecular and Structure Formula.

### 3.9.3 Abstracting Procedure

After selecting the documents for preparation of abstract, the following procedure is followed while summarizing them:

1. After this, a final copy of this rough abstract is made. For this, an effort should be made to end the abstract in an isolated paragraph.
2. In this abstract , it is checked to rectify the mistakes, in which the correct use of words like spelling, comma, punctuation mark etc. is done under the rules.
3. A rough abstract is prepared from the selected key words.
4. After selecting the key words of the document, they are noted down.
5. First of all, the subject-matter, area, conclusion etc. mentioned in the abstract document are carefully studied. For this work, the introduction, conclusion, suggestions, ideas etc. given in the document prove to be helpful.

### 3.9.4 Canons of Abstracting

Dr. S. R. Like the principles of classification and cataloging by Ranganathan, the canons of abstracting have also been presented. All these canons give a theoretical form to the abstraction work. These canons are as follows -

1. Positive Canon,
2. Negative Canon

1. **Positive Canon :** What elements should be included while abstracting a document? They are selected by this canon.
   (i) The content described in the original article should be summarized.
   (ii) In the abstract, the basic and elementary progress in the field of knowledge should be clearly mentioned.

Note

(iii) Secondary progress in the field of knowledge should also be mentioned.

(iv) The methods, technical findings, etc., described in the document should be included along with their characteristics.

(v) The important data given in the document should be briefly summarized.

2. **Negative Canon :** Which elements should not be included while abstracting. It is determined by the negative canon.

(i) Abbreviations of subjects should not be used in the abstract.

(ii) Information should not be provided which reveals the report of the document.

(iii) The examples mentioned in the document should not be elaborated.

(iv) The sentences of the document should not be used in the abstract, hence telegraphic language should be used.

### 3.9.5 Principles of Abstracting

To make an abstract answer and useful, the following principles should be kept in mind:

1. **Size :** The size of any summary makes it useful and important. Various scholars have rendered their views regarding its size. According to the summary,

(a) According to Borko and Vernier – 1 to 1,000 words

(b) Informational Abstract (as per DDC) – 150 to 200 words

(c) For Serials – 50 to 100 words

(d) Best Summary – 200 to 250 words

2. **Writing Style :** According to this principle, the abstract should be written in the simplest style. The ideas, concepts, activities etc. presented in the document should be written in the past, but the conclusion should be presented in the present.

3. **Word Selection :** The sequence of the words selected for the summary should be arranged. One should observe frugality in the use of words.

4. **Sentences :** The title of the document should not be repeated under the summary and whatever sentences are used should be completely clear and concise.

Note

5. **Language :** The language used in the abstract should be simple and include abbreviations, punctuation marks. And trade names should not be used. In addition, presented in the original document. The subject matter should be presented completely.

### 3.9.6 Qualities of a Good Abstract

The characteristic of an answer summary is to present maximum information about the subject material contained in the original document in the least amount of words. An abstract should have the following characteristics to save the reader significant time

1. Brevity,
2. Accuracy,
3. Simplicity,
4. Reliability
5. Authenticity
6. Clarity

1. **Brevity :** Conciseness is the main feature of the abstract. The abstract should be concise as compared to the original document but it is necessary that it should have full capacity to express the purpose, method, conclusion etc. of the original document. Adopting brevity in abstracts saves time for the reader and staff as well as space if the summary is concise. So the cost of presentation, printing and publication is reduced. The sizes of the bars vary. But in how many words it is to be presented, it depends on the document. A good summary should be 200 to 250 words.
2. **Accuracy :** While preparing the abstract, the abstractor should try that the essence should be impeccable, that is, the purity of the essence is very necessary so that any kind of confusion cannot arise. To avoid any errors in the abstract, first a rough summary of the original document should be made and this type of re-examination should be done. Most errors in the abstract are related to mathematical base material, proper names and punctuation marks, which must be verified.
3. **Simplicity :** The abstract made from the original document should be simple and the abbreviation and trade name should not be used in the abstract. Therefore, the language and writing style of the abstract should be simple, which can be easily understood by the users.
4. **Reliability :** The credibility of the essence is very important. Therefore, its rules should be followed in summarizing and presenting the summary. For this The abstractor should have thorough knowledge of all the techniques of abstracting.

Note

5. **Authenticity :** Authenticity is the main quality among all the characteristics of essence. The authenticity of the essence depends on the person doing the abstract. The conclusions presented in the abstract should be according to the authentic and original document of principles, techniques and methods, for this work the abstract should be expert and trained.
6. **Clarity :** The language, style, words etc. used in the abstract should be completely clear and understandable. The words selected while preparing the abstract should be taken from the original document itself. Because the words used by the abstractor can be ambiguous.

Therefore, the essence of an original document should be clear, concise, accurate, reliable and authentic so as to provide maximum information to the user.

## Conclusion

In this chapter, we studied the contribution of scholars like Cutter, Kotsa and Ranganathan in cataloging and the concept of cataloging, to know their ideas, abstraction, indexing and bibliographic list. Also the layout of the Cataloguing of the books.

## Important Terms

- Cutter introduced the concept of a specific subject in the 'Rules for Dictionary Catalouge' in 1876.
- Union List refers to the combined list of various libraries of a region.
- An indexing list is a list that is usually at the end of a book with respect to a name, subject, etc., in which the place to find them is indicated.
- Any document of human communication, book publication or any other set of transcripts etc. can be taken as one entry.
- Abstraction is a concise presentation of the content of an original document using the same style that was available in the original document.

Note

## Exercises

### Very Short Answer Type Questions

1. Write the meaning of indexing?
2. State the meaning of abstraction.
3. What do you understand by bibliographic record?

### Short Answer Type Questions

1. Write a note on the layout of the Union Catalogue Code of Books.
2. Describe the contribution of Cutter in the field of Cataloging.
3. Describe the contribution of Ranganathan in the field of cataloguing.

### Long Answer Type Questions

1. Defining indexing, describe its functions and characteristics.
2. Defining abstraction, describe different types of abstracts and briefly explain the process of abstraction.

Note

UNIT 4

# Online Cataloguing

## 4.1 Introduction

In this chapter, we will learn about the online indexing system OPAC and Web OPAC. You will also learn about the subject listing and what its purpose is and will also study in depth on this topic. We will study why LCSH is used in the library and what kind of facilities are available in running the library by using it.

## 4.2 OPAC

The short form of On-line Public Access Catalog is called OPAC. This catalog proves to be more user-friendly than the catalogs traditionally provided in libraries and helps the users in their information discovery process at every step.

Due to the explosion of information in the modern science era, various networks are providing cataloging work in the form of magnetic tapes in the spirit of cooperation. The OCLC system is providing the Library of Congress (LC) cataloging data via the Network. Users of OCLC. ask for bibliographical description of their desired information in magnetic tapes. For this reason, many libraries from the 1970s onwards moved rapidly towards computerized catalogs in place of their card catalogs and made their catalogues available in the form of Micro Fiche. In the modern era, most libraries are making their catalogs available through OPAC.

Most libraries in the era are making their catalogs available through OPAC. OPAC is a computer-generated catalog available in the form of magnetic tapes. This type of catalog can be used by any user through a computer. Thus the desired document is searched through All Line. The search for any document and information can be done by OPAC through the subject approach.


Today more and more libraries and information centers are using OPAC. Because the electronic format of this catalog proves to be more user-friendly for the users than the card catalogue.

### 4.2.1 Merits of OPAC

Following are the advantages of using OPAC in libraries and information centers :

1. When the extension service (SDI) selected by OPAC is over, Smt. Reminders or requests for extension of date can be received through e-mail.
2. Through OPAC, the user can request for the acquisition of his desired book or document. Apart from this, reservation of desired documents can also be done online.
3. OPAC can also conveniently search for information through Boolean operators.
4. OPEC, being an on-line facility, provides the facility to search information through various indexes, such as publisher index, seminar location index, Quick and Quoc (KWIC/KWOC) etc
5. OPAC allows users the freedom to search for content from different approaches. Through this, the user provides the facility of searching documents and information by author, title, subject, publisher etc.
6. Through OPAC, the user himself can search for his desired material on-line and he does not need any reference staff for this work.
7. Most of the users are not aware of the content of the available content. In this regard, it proves to be User Friendly and also facilitates the user to select alternative material in addition to the desired material.
8. With the help of this, the searcher can be searched for his desired material from remote areas online at one place.

## 4.3 Web Opac

The concept of Web OPACs is the recent origin and it is serving as a gateway to the resources not only held by the respective library but also to the holdings of other participating libraries without to local collection but going, beyond further to regional, national, international levels. It allows users to interact with documents stored on computers all over the world and makes easier access to catalogue data in the form of bibliographic records. It becomes another search engine-referred as 'web cat' and as an 'information gateways'. It can support such as Telnet HTTP, FTP and Gopher. According to Online Dictionary for Library and Information Science defined as : "An

Note

online Public Access Public catalogue that uses a World Wide Web, as opposed to a text-based interface accessible via Telnet."

### 4.3.1 Features of Web OPAC

The features of Web OPAC are :

1. It is accessible through the internet.
2. It is possible to search independently by Author, Keyword, Title or Year.
3. Displays complete bibliographic information as appeared on reprints.
4. Features of traditional OPACs such as storing bibliographic and sometimes full-text databases, providing direct access to a library's bibliographic database by means of terminal of PC, search result in readily understandable form, reference help etc.
5. It has the ability to use hypertext links to facilities navigation through bibliographic records.

## 4.4 OCLC

OCLC is the abbreviation for On-line Computer Library Center. It is the library and information network of America and the largest information network in the world. This network is an important center for resource sharing and information acquisition between libraries and information centers. This network was named "On-line Computer Library Centre" by "Ohio College Knowledge Organization" in 1971.

This information network was established in 1967 in Dublin, Ohio. This network was the most popular regional network at that time. On-line services were also started by this center in 1971. As a result its nature was also changed. In this regard, the Marc format proved to be a good and important step. At present, this network is a huge network, which is not only the network of information and libraries in America but has been established as the largest information network in the whole world.

### 4.4.1 Aims of OCLC

Following are the main objectives of the system:

1. To make available the database of documents in the form of magnetic tapes to the users.
2. Mechanized all the activities of the library.
3. To provide inter-library loan service for the purpose of information communication between libraries and information centers.
4. To support and promote research and development (Rand D) activities.

Note

5. To promote resource sharing among all libraries and information centres.
6. Automation of the process of acquisition and cataloging in libraries and information centers.
7. To make available to the users their desired information and documents on-line.
8. Establish an international level network of libraries and information centers.

### 4.4.2 Activities of OCLC

The main activities of OCLC are:

1. **Union Catalog :** The facility of on-line cooperative list is provided through this network; these union lists are made available by member libraries and information centers of newly Local Services -received documents. This information is very important and incomparable in relation to newly received documents in any library. The main feature of the program is to provide union lists . This feature saves the time of the users.
2. **Local Services :** It is not only an international network but also provides local services to on-line users. Many library units have been created through these services, in which centralized cataloging service, on-line and information service is being provided.
3. **MARC Tapes :** Through this center, services of all the library members are made available through marc tapes. This service can be done off-line sub. is being made available. This information can also be purchased. Material published before 1968 is not included in the mark tapes. The tapes contain about 6 million documents which are available in 468 languages. These can be searched online through Marc from anywhere in the world.
4. **Database Services :** The collection of this center is in the form of a database which is a collective record of all the member libraries, such as marc tapes of Library of Congress, National Medical Library, GPO . Monthly list etc. Through this network, databases of documents published up to and before 1968 can be obtained.
5. **Inter-Library Loan Service :** Inter-Library Loan Service (ILL) service is being provided by this network under Reference and Documentation Services. A sub-system of this system is providing Service on-line between member libraries and information centers within the state of New York.

Note

6. **Printing of Catalogue Cards :** In 1972, this center merged the catalogs of all the member libraries and information centers and the Library of Congress (LC) together so that their titles could be identified and catalog cards can be created. For this task, a Catalogue Production System was started through this network. This system is being served through MARC Tapes and is being distributed to the Member Libraries and Information Centres.
7. **Full Cataloguing of Periodicals :** The entire cataloging process of periodicals is done by this center. The facility of inquiry, demand and union catalog is also provided by this network. Through this system, the facility is available in the list on-line.
8. **Acquisition Sub-system :** Through this network the facility of acquisition sub-system is provided to the user. This system is called SUBY / OCLC. Through this system, the concerned librarian can be ordered to acquire the desired documents. After this, this center receives and prints that order and sends the order to the concerned book publisher to send the book.

## 4.5 Subject Catalouging

This list was first published in 1923 by Minnie Earl Sears. This list was compiled by Sears to meet all the subject heading needs of medium-sized public libraries and schools. It is an abbreviation for LCSH. positions in this list.

The arrangement is a dictionary. The purpose of this list is that all the documents available in the library on a particular subject should be kept under one form of subject headings. Apart from this, in this list, after x, xx, for mutual references, those terms have been given from which the subject "See" and 'See also" can be formed.

The arrangement of words in this list is lexical. Words used are printed in thick and dark ink and synonyms/titles which cannot be used as titles have been printed in light ink.

The titles in this list are printed in a variety of letters:

1. Headings where the user is directed to select the appropriate title by the See Also (SA) directive.
2. See references as required from those headings before which x is marked
3. Those headings which are preceded by xx signs, see it as per the need (See also references) instructions are made. From the 15th edition, these symbols have been replaced by BT, RT, NT, SA and UF, which are used as related terms.


### 4.5.1 Subject Headings Format

Subject headings in SLSH are designed in a variety of ways by expanding descriptive phrases containing two or more words into isolates word headings and isolates concepts of compound and compound subjects.

1. Single word heading
   Agriculture
   Religion

Individualizing element is given with headings to identify all words with different meanings. for example

Seals (Animals)

Seals (Law)

2. Compound heading
   Bow and Arrow
   God and Evil
3. Adjective phrase heading
   American Literature
   Civil Engineering
4. Prefix phrase heading
   Electricity in Agriculture
   Freedom of information
5. Complex subject
   Education-History
   Mathematics-Study and teaching
   United States-History-1945-1953

### 4.5.2 Entry Format

Sears List also uses different symbols with headings, as in 'LCSH'. Subject headings are printed in bold letters and referential terms in lowercase letters. From the fifteenth edition the 'see' and 'see also' are interchangeable references and the ' and 'xx' symbols have been replaced by UF/USE, BT, NT, RT and SA.

### 4.5.3 Adding Headings by Lister

In the Sears list, listers are allowed to add headings as needed, as follows:

(i) Names of persons

(ii) Family names

Note

(iii) Names of places

(iv) Name of Nationality

(v) Name of the National Language and Literature

(vi) Names of wars and battles

(vii) Names of macro bodies

(viii) Common names such as names of animals, fruits, sports, diseases, parts of the body, chemicals and minerals.

### 4.5.4 Key Headings

The following main posts have been given as a guide for efficient performance by the lister. According to which alternative provision has been made for other main heads.

(i) Person: Shakesphere, William, 1564-1616

(ii) Names of families; Indian of North America

(iii) Place: United States, Ohio, Chicago -

(iv) Language and Literature: English literature

(v) War: World war 1939-1945

Cross references are also given which are very helpful in finding the most suitable topic title. The website of MeSH is

http://www.nlm.nih.gov/mesh

An updated printed version of this controlled glossary is published in January every year.

## 4.6 Library of Congress Subject Headings (LCSH)

The Library of Congress Content Title Catalogue is a comprehensive English-language lexical control method that is used by library readers to access lexical lists of college, university and major public libraries in the United States, Canada, and many other countries. This subject title list, which originated in the Library of Congress in 1897, has been continuously revised and updated to this day. It is currently in its twentieth edition. This table of contents is now available in print form, on microfilm and CD ROM as well. It can also be obtained on the Internet.

This list was created by the Library of Congress of America. This list is mainly used as a tool for indexing books, periodicals, and information materials acquired in the Library of Congress. Its main objectives are as follows:

(i) To prepare such a map of knowledge in different subjects in which

Note

the structure of the subjects can be understood and how the concepts under these subjects are related to each other, etc.

(ii) To provide a standard controlled dictionary with the help of which the subject headings of speech documents can be determined on the basis of consistency.

(iii) Providing such subject title to both the indexer at the time of indexing and the searcher at the time of search so that the same term can be used in indexing and search and the desired information or document can be accessed.

(iv) To express a concept, to provide a word and a list of its synonyms, from which the appropriate word can be selected.

(v) To act as a guide while accessing a database in which the accessor can get the desired information or document by being guided through counter instructions.

(vi) Showing the hierarchical relationship between the terms so that the searcher can broaden or narrow his search.

(vii) Can express new concepts in the existing concepts by means of words at appropriate places.

### 4.6.1 The Structure and Function

Library of Congress Lists Subject Headings is a controlled repository of simple natural language words. It has a display of words and forms that have been established without any grammar.

**Headings**

The concepts in this list are expressed in two types of headings.

(a) Posted Headings or Preferred Headings

(b) Non-posted Headings or Non-preferred Headings or Lead in Headings

Following are some examples of primary and secondary headings:

(a) Preferred Headings
Economic Policy
Economic Security
Aeronautical Libraries
Aeronautical Laboratories

(b) Non-preferred Headings
Navigation (Aeronautics)
Aeronautical Journalism
Aerodynamic Laboratories

Note

Libraries, Aeronautical
Economic Planning
Security, Economic

### 4.6.2 Creating Topic Headings

Library of Congress subject headings contain predetermined subject headings. For subject access, the indexer or indexer has to select the title corresponding to the content of the document from this subject title list and present it in the catalog or index. LCSH has a variety of methods for constructing topic headings that can represent a single concept with a single verse and descriptive phrases combined with two or more verses.

### 4.6.3 Single Term Heading

There are many documents which are related to only one subject and they can be presented with a single word/word. Titles containing such one word represent an object, person or concept.For example

| | |
|---|---|
| Disease | Railways |
| Dogs | Gandhi |
| Hardness | Schools |

There are generally two problems in the selection of such titles.

1. To differentiate homographs, and
2. Whether the words are presented in the singular or in the plural.

To show the difference in consonantal words, their meaning or sum with their subject title in parentheses, for example

| | |
|---|---|
| Tank (Military) | Pitch (Music) |
| Tank (Water) | Pitch (Cricket) |

Singular and plural form are fixed- like

### 4.6.4 Singular

(i) Abstract ideas, literary forms, such as:

Biography
Density
Essay

(ii) Biosciences class, such as

Monkey
Palm


(iii) Names of fruits, such as

Peach

Pear

### 4.6.5 Plural

(i) Tangible thing or person, such as
Airplanes
Engineers
Students

(ii) Large group like
Monkeys
Palms

(iii) Compilation e.g.

Biographies

Essays

### 4.6.6 Phrase Heading

When a concept cannot be properly expressed by a single term, a phrase (i.e. the title of two or more words) is used. Phrase headings have many patterns, such as

**(a) Adjective Phrase headings:** These are very common types of phrase titles which are made up of noun or cognate phrases in which qualifier or adjective denominator words are attached. One of the two-syllable headings representing a single idea refers to the central idea and the other word represents the merit of this central idea. For example, in the subject title 'Local Taxation', Taxation is the central idea and 'Local' is a qualifying adjective.

**(b) Conjunctive Phrase Headings:** LCSH Conjunctive Phrases uses headings in which two or more nouns contain or are free from noun phrases or both qualifying indicators. Combined words are attached with 'and'. For example:

### 4.6.7 Literature and Science

### Technology and Civilisation

**(c) Prepositional Phrase Headings:** In these types of headings, two or more nouns, noun phrases or both with or without qualifying expressions are joined by a preposition. For example:

- Photography of Animals
- Children in motion Pictures

Note

Mixed titles are those that express party relations. Compound headings are those which express the conditional relations i.e. the interrelationship between two subjects.

- English Literature-20 Century-History and Criticism-Bibliography
- Working Mothers-United States-Attitudes

### 4.6.8 Subdivision

There are four categories of subdivisions in LCSH, which are named episodically, chronologically and geographically. Examples of these are as follows

(A) Topical Subdivisions Automobiles-Motors Carburetors

(B) Form Subdivision United States-History-Periodicals

(C) Chronological subdivisions American Literature 19th Century

(D) Geographical Subdivisions Steel Industry-United States

(E) Free Floating Subdivisions

Free movable subdivisions are either tabled separately (excluding main headings) or are placed under representational main headings called Pattern Headings. Which are appended with appropriate main heading as per the requirement. For example Birth Control-religious aspects-christianity.

### 4.6.9 Entry Format

Subject headings that are being used in the library list in LCSH have been included in this list. Each subject title included in this is for the dictionary list of the library, there is a provision for different types of subject heading and other reference terms. Subject headings are listed in dark ink in bold and cross-references in light letters. A heading may be followed by a May subd geog statement indicating that titles can be divided by locations and by Library of Congress class numbers. In a separate paragraph, a detailed note is given to guide the use or meaning of the heading. The references are then grouped under the codes of UF (Used for) / Use BT (Broader Term), RT (Related Term), SA (See Also) and NT (Narrower Term) by headings. In the list Use and UF can be seen to mean 'See also from' and SA and NT as 'See also' interchangeable references. Different punctuation marks are provided for different types of headings.

### 4.6.10 Scope Notes [SN]

For consistency in the use of a subject, its subject area comment is written after the word indicating it. like :


### 4.6.11 Academic Libraries

SN-Works on Academic Libraries in Special fields, such as law school libraries, are entered under headings for special libraries, e.g., Law Libraries, Business Libraries, etc.

### 4.6.12 Special Libraries

SN-Here are entered works on libraries having specialized subjects, containing materials in a special format, or serving a specialized clientele.

This Subject Area Comment does not give the full meaning of a title, but indicates when to use the title.

Relationship Between Headings

Like thesaurus, three types of relationships are shown in LCSH.

1. Equivalence Relationship
2. Hierarchical Relationship
3. Associative Relationship

### 4.6.13 Layout and Display of Headings

All the primary headings in this list are arranged alphabetically. USE, UIF, BT, NT and RT are given under each primary heading. A sample entry of LCSH is given below :

**Calculators (May Subd Geog)**

- [HF 5688-HF 5689 (bookkeeping)]
- [QA 75 (Mechanical devices)]

SN - Here are entered works on present day calculations as well as on calculators and mechanical computers of pre-1945 vintage. Work on modern electronic computers first developed after 1945 are entered under computers.

UF Adding-machines

- Calculating-machines
- [Former heading]
- Computers, Mechanical
- Desk calculators
- Mathematical machine (Calculators)
- Mechanical computers

BT Mathematical instruments

- Office equipment and supplies

Note

RT Arithmetic Computers
- Tabulating machines

NT Accounting machines
- Comptometers
- Electronike BZ-19M (Calculator)
- Graphic calculators
- HP-12C (Calculator)
- Mark 1 (Calculator)
- Pocket Calculators
- Programmable calculators
- Sharp 5100 (Calculator)
- TI 30 (Calculator)
- Circuits

BT Electronic circuits
- Calculators, Mental
- Use mental calculators
- Calculators, Pocket
- Use Pocket Calculators
- Calculators, Programmable
- Use Programmable calculators

Explanation Title "Calculators" should be used as the primary title "USE" should be changed to "SEE" and "UF" should be changed to "See form". Also BT and RT should convert to "See Also From" like so :

| | | |
|---|---|---|
| Adding machines | See | calculators |
| Calculating machines | See | calculators |
| Computers, Mechanical | See | calculators |
| Desk calculators | See | calculators |
| Mathematical machines (Calculators) | See | calculators |
| Mechanical computers | See | calculators |
| Mathematical instruments | See also | calculators |
| Office equipment and supplies | See also | calculators |
| Arithmetic computers | See also | calculators |
| Tabulating machines | See also | calculators |
| Accounting machines | See also | calculators |


| | | |
|---|---|---|
| Comptometers | See also | calculators |
| Electronics BZ-19M (Calculator) | See also | calculators |
| Graphic calculators | See also | calculators |
| HP-12C (Calculator) | See also | calculators |
| Pocket calculators | See also | calculators |
| Programmable calculators | See also | calculators |
| Sharp 5100 (calculator) | See also | calculators |
| T130 (Calculator) | See also | calculators |
| Calculators-Circuits | See also | calculators |
| Calculators, Mental | See | Mental calculators |
| Calculators, Pocket | See | Pocketcalculators |
| Calculators, Programmable | See | Programmable calculators |

### 4.6.14 General Reference

In LCSH, not a particular title, but the whole group is represented by a simple reference See also. SA is used as a suffix before a simple reference statement. As

**Cavalry drill and tactics**

[UE 157-UE 302]

SA --- Subdivision Cavalry Drill and tactics under names of individual armies, e.g., Great Britain, Army-Cavairy-Drill and tactics, and sub division.

- Drill and tactics under names of individual cavalries, e.g., United States
- Army, Cavalry-Drill and Tactics.

Great Britain Army

- Cavalry
- Drill and Tactics.

United State, Army, Cavalry

- Drill and tactics

Subdivisions of Preferred Headings to use different concepts as a heading. The following four subdivisions are shown in LCSH :

Note

### 4.6.15 Topical Subdivisions

Thematic subdivision of a topic is called thematic subdivision. In LCSH it is represented as :

Automobiles – Fuel Systems
Lubrication
Motors

### 4.6.16 Form Subdivisions

A material is redistributed according to form such as:

Bible – Bibliography

Bible – Criticism, Interpretation, etc., - Bibliography

Bible – Manuscripts-Catalogs

Bible – Chronology-Charts, Diagrams, etc

### 4.6.17 Chronological Subdivisions

Chronological division is used in LCSH to limit the primary heading to a certain point in time, e.g.

- Arabic Poetry
- Arabic Poetry To 622
- Arabic Poetry To 622-250
- Arabic Poetry To 750-1258
- Arabic Poetry - To 1258-1800/
- Arabic Poetry 1801
- Arabic Poetry 20th Century

### 4.6.18 Geographical Subdivisions

In LCSH, two types of denominators are shown after a heading.

(Not subd. Geog) i.e. the title should be followed by the geographical location.

For example:

- Civil Rights (May Subd. Geog)
- Civil Rights and socialism (May Subd. Geog)


- Civil Service (May Subd. Geog)
- Civil Development (May Subd. Geog)
- Child psychology (May Subd. Geog)

(Not Subd. Geog) That is, it is not necessary to indicate the geographical location after the title.

For example:

- Antipodes Islands (NZ) (Not Subd. Geog)
- Chepo river (Panama) (Not Subd. Geog)
- Bego Mountain (France) (Not Subd. Geog)

If a primary heading is not followed by (May Subd. Geog) or (Not Subd. Geog), if so, that title should be used as it is. For example, if in LCSH, the following headings are shown in

Youth (May Subd. Geog)

- Employment (May Subd. Geog)
- Law and Legislation (May Subd. Geog)
- Mental Health and services (May Subd. Geog)
- Utilization (May Subd. Geog)

So they should be written in the subject title like this.

Youth – India

Youth – Employment - India

Youth – Employment Law and Legislation - India

Youth – Mental Health and services - India

Youth – Mental Health and services - Utilization – India

### 4.6.19 Free-floating Subdivisions

This subdivision is used by the indexer to determine the subject title in case a suitable heading is not available to it in LCSH. The indexer should do free-floating subdivision according to the rules and instructions used in its library.

Note

**Some Important Subject Heading Lists**

1. Library of Congress Subject headings (LCSH)
2. Medical Subject Heading (MeSH) (http://www.nlm.nih. gov. mesh)
3. Canadian Subject Heading (CSH)
4. Sears List of Subject Headings (http: www.hwwilson.com/secrist-18th.cfm)

## Conclusion

In this chapter we studied the online indexing system OPAC and Web OPAC. To know about the listing and also study its purposes. Studied in depth about the use of LCSH in the library.

## Important Terms

- Library of Congress Subject Headings Made in America It is a tool used for indexing acquired books, periodicals, and information materials.
- The Sears List was published in 1923 and aims to keep all the documents available in the library on a particular subject under a single form of subject headings.

## Exercises

### Very Short Answer Type Questions

1. Name any three subject heading lists.
2. What is Free Floating Subdivision?
3. State any two objectives of OCL?

### Short Answer Type Questions

1. State the main objectives of the Library of Congress subject headings.
2. Briefly describe the types of headings divided into LCSH concepts.

### Long Answer Type Questions

1. Explain the Sears List of Subject Listing in detail.
2. Describe in detail the Library of Congress subject headings.

UNIT 5



Note

# Recent Trends in Classification

## 5.1. Introduction

In this chapter we will study what is mark format and why it is needed in the library and also get information about Unimark. Along with this, we will also study about UDC that there is a difference between CC and UDC. We will also study this topic.

## 5.2 MARC Format

MARC is an abbreviation for Machine Readable Catalog. The MARC model was first started on a pilot basis in 1966 by the Library of Congress in Washington, and after its successful run, it was adopted as a long-term project in 1968. The full writing of the abbreviation of Mark makes it appear that it is either a Catalogue or refers to the principles of Cataloguing, whereas in reality Mark is neither a Catalog nor a Cataloging Principle. In fact, mark no is just a short and simple word to denote each part of a catalog record so that it is easy to use by computer. The MARK format was developed especially with libraries in mind, so that bibliographic data can be easily exchanged between libraries.

Gradle and Hopkins have defined MARC as "Mark is a set of various formats for use by a bibliographic data computer that provides a set of conversions for their identification and adjustment. MARC Format issued by the Library of Congress After this, many different versions came, some of which are prominent.

(i) US Marc
(ii) UK Marc
(iii) Kenmarc (CAN Marc)
(iv) Uni Marc (UNI Marc)
(v) Marc 21 (MARC-21)

Note

All the mark formats mentioned above have the following similar properties:

(i) All these are based on ISO-2709 or its equivalent national standards.

(ii) Most of these are national formats based on the respective National Library or National Bibliographic Agency.

### 5.2.1 MARC-21

Mark 21 is created by combining MARC and CAN MARK (CAN MARC- Canada). Its first version came in 1999. Since 2001, the British Library has also accepted Marc-21 leaving UK Marc.

Marc 21 format is the set of code and content designators for encoding machine readable records. This format is defined for the following five types of data

(i) Marc-21 Format for Bibliographic Data

(ii) Marc-21 Format for Holding Data

(iii) Marc-21 Format for Authorized Data

(iv) Marc-21 Format for Classification Data

(v) Marc-21 format for community information

Each Marc 21 record consists of the following components:

**(i) Field:** Each field is a tag with three characters that stores two types of information

**(ii)** Information relating to the type of material, its physical properties, specific bibliographic details, which are necessary for the processing of the record.

**(iii)** About the bibliographic data elements (i.e. author, title, publisher etc.) which are necessary for cataloging.

The bibliographic data of Mark-21 consists of the following areas:

**(i) Controlled Area (001-006):** This includes the control number and other control and code related information required to mark bibliographic records. For example date and time document types (books, electronic resources, maps, etc.).

**(ii) Control Zone (007):** Specific information about the physical property in coded form.

**(iii) Control Zone (008):** This is the area of 40 character space (00-39), which gives information about the record and its bibliographic subject.

**(iv) Number and Code Area (01X-04X):** This is the area of control and sex numbers which contain the standard number and other information.

Note

(v) **Classification and class number area (05X 08X):** Area related to class number.

(vi) **Main Entry Area (1xx ):** This is the Heading Area of a name or a form heading, which is used as the main entry in the bibliographic record. for example:

100 Main Entry – Personal Name
110 Main Entry – Corporate Name
111 Main Entry – Conference Name
130 main entries – Uniform Heading

(vii) **Heading and Heading related field (20X24X):** Heading and Heading related information area

For example:

210 short title
222 Main Title
240 Uniform Title
245 title statement

(viii) **Edition and Publishing Area (250-270):** This edition contains the publisher's address and other publication related information.

For example:

250 Version Statement -
256 Computer File Properties
260 Publications Distribution etc.

(ix) **Physical Description Area (3XX):** This page contains information regarding number, size, volume etc.

For example:

300 Physical Description
306 Playing Time
310 Current publication period .

(x) **Sequence Statement Area (4XX):** This is the area of sequence related information which starts from 400 and 490 tags.

(xi) **Comment Area:**

Part-1 (50X53X)
Part-2 (53X58X)

Note

It is a field of different types of comments in which relevant information is included.

(xii) **Subject Learning Area (6XX):** Information related to a particular topic or word of learning is given in this area.

For example:

600 Subjects Additional Entry (For Personal Name)

650 Topical Names

651 Geographic Names

(xiii) **Additional entry area (70X-75X):** It provides information for other accesses to the bibliographic record. This information is related to the main entry (1XX), subject access (6XX), sequence area (4XX), title area (20X-24X).

(xiv) **Link Entry Area (76X-78X):** This area Information for linking various Bibliographic Items

(xv) **Sequence Additional Entry Field (80X-830):** Title, name etc. used in this sequence (Series)

(xvi) **Holding, Location, Alternate Graph etc. Fields (841-88X):** Provides information relating to the details of data elements in a deep bibliographic record.

## 5.3 Universal MARC

Unimark is an abbreviation for Universal MARC, whose creation and development goes to Ifala. This ISBD (ISBD-International Standards for Bibliographic Description).

### 5.3.1 Unimark's Evolution

After the creation of the mark format by the Library of Congress in 1966, the bibliographic agencies / national libraries of different countries were influenced by it and started developing mark formats for their respective countries. Since the different mark formats belonged to different countries, it was natural for them to have local differences. Due to some differences in formats, difficulties began to arise in the exchange of data which is obstructing the fulfillment of the purpose of other international agencies including the Library of Congress. With a view to overcome these difficulties, IFLA (International Federation of Library Associations and Institutions) started work towards the development of a universally accepted mark format. As a result, in 1977, IFLA developed an international mark format known as UNIMARC (Universal Mark). The initial purpose of this mark format was only to enable international exchange of data in machine readable

Note

form between various national bibliographic agencies. The second edition of Unimark appeared in 1980, while the Unimark Handbook was published in 1983. Unimark was the first bibliographical record format that was developed by the ISBD. (ISBD) based. The latest edition of the Unimark manual was published in 1994.

### 5.3.2 Unimark Structure

A field of any Unimark record consists of the following components:

(i) **Tag:** It is of three digits which indicates the areas of bibliographic data.

(ii) **Indicators:** This is a set of two single digits which are written after the tag. It either indicates the area definition or describes how the areas are to be used for cataloging. In this, the no sign is used for the blank number.

(iii) **Subfields:** The data within any field is divided into different subfields which are according to the type or function of the information. Various sub-region designers are indicated by the dealer symbol #.

(iv) **End of field mark:** This is indicated by the @ sign.

### 5.3.3 Blocks of UNIMARC Format

Unimark consists of the following 10 sections:

(i) **OXX Identity Block:** Stores the information needed to identify the record. e.g., 001 0198 122622/- In this example the identification number of the document is information.

(ii) **Coded Information Block:** Area 100 stores general information related to data processing. Area 101 stores language information. For example 1011#$@eng$ c fre@. Language related information is stored in this example.

101 gives language related information.

1- Provides translation information.

$ @eng - English translation

$ c fre- Translation into French

(iii) **2XX Descriptive Information Section:** This stores the subject description information, title and statement of responsibility information.

For example

2001/$@(NSB) Then (NSB) List domain $ f Alasin - Fournier $g translated from the French by Frank Davision $ g after - word by John foweless $ g ilustrated by lan beck @

Note

The example here contains the following information.

200 Heading area

1. The title is Significant.

### 5.3.4 Unimark Field-tabulated Form

| Type of Data | Block | Field |
|---|---|---|
| Identifier | 0 | 0XX |
| Coded Information | 1 | 1XX |
| Descriptive information | 2 | 2XX |
| Comment | 3 | 3XX |
| Entry Link | 4 | 4XX |

There are a total of 143 regions in the Unimarc format, of which only a few regions are ISO. 2709 standard. This format consists of at least one other field in addition to the record label, directory, record indicator. The structure of Unimark is mainly based on four basic facts.

(i) Modularity in the form of sections for ease of understanding

(ii) I.S.B.D. Support

(iii) Support for all physical forms of documents stored by the library

(iv) Support at all levels for cataloging of these documents

### 5.3.5 Other mark formats based on Unimark

(i) South Africa (SAMARC South African MARC)

(ii) Taiwan (according to China's mark format which is based on Unimarc)

(iii) Japan (Marc)

(iv) Croatia (Yu/UNIMARC)

(v) France Inter MARC

(vi) Italy

### 5.3.6 Marc Format used by Different Countries

| UK Marc | Unimarc | US MArc |
|---|---|---|
| Australia | India | USA |
| Great Britain | Bulgaria | Malaysia |
| Spain | Italy | Netherland |
| USA | Japan | Singapore |

Note

## 5.4 UDC System

This classification system was first introduced in 1895 by Paul Otlet and Henri La Fontain. Between 1904-07, 33,000 were placed in the first edition in French, and 70,000 in the second edition, published in 1927, in 1933. In the 1951 edition the number of subdivisions was increased to 140,000. Later on, the British Library... The Museum Library had to prepare a classification for the classification of pure and applied sciences. To meet this need, in 1936 the Universal Decimal System... edition was published in English. 6000 for Science and Technology in it. At the request of Eslive, the British Standards Institute took over the publication of the English version. This massive work was finished in 1985. This text was given the standard issue BS1000 M by the British Standards Institute under contract from the International Documentation (FID). You. The impact of D.C. on libraries can be understood from the fact that abridged editions have been published in 23 languages. Its new English version was published in 1961... Every year a list of amendments to the system is published.

### 5.4.1 Characteristics of UDC Systems

The basic outline of DC Decimal Classification by UDC - Main class and principal has been adopted. As a result, UDC like Dewey Decimal, there is also a general classification system encompassing the entire world. The nature of thematic sub-divisions is to move from general to specific.

UDC Synthetic classification system. Actually the U.D.C. is considered to be the first developed synthetic system in the world. It is mainly an enumerative system, in which an attempt has been made to divide the classes from subject to subject to the smallest subject. As a result of this division, the hierarchy has been divided into sections and sub-sections, and notation has been fixed for each subject, although at present the D.D.C. Some Auxiliary Tables have also been included for the subsections. These tables arrange that for the sub-classes of some subjects, the subsections of other subjects that have the same knowledge organization can be used effectively. There is still an enumeration of subjects in UDC, but the method of notation for complex subject headings is also included in this system, by which class numbers are made by adding different parts of notation from other parts of the system or from about a dozen auxiliary tables. This is required for a synthetic system like UDC.

As compared toDDC, in UDC modern classification theories have been given importance and now this system has become a more facted scheme of classification system. This system has been made more robust in terms of synthesis ability or combination of themes and concepts through helpful devices.

UDC Presently available in full medium size and abbreviated versions.

Note

UDC- The system is kept up to date by the International Federation of Documentation, but due to the complex process of revision, it takes considerable time to incorporate the latest branches of knowledge science

### 5.4.2 Functions of Various Symbols used in UDC

| Symbol | Functions |
|---|---|
| + Plus sign | To combine the non-trailing digits |
| + Stroke | To combine the trailing digits |
| : Colon | To combine two or more related concepts of equal importance |
| = Equal Sign | For Subdivisions of 4 Linguistics Section |
| ( ) brackets | Indicates the geographical extent of Place Notation Subject in Brackets |
| : used with indo arabic numerals | To show the physical appearance displayed in the document zero in parentheses |
| (=) equal sign in parentheses | to indicate nationality and race |
| reverse coma | to show time |
| .00 Point zero zero | to show perspective |
| - hyphen | To show analytical representations of concepts to demonstrate perspectives |
| .0 Point zero | Analytical - but for the analysis of concepts |
| ' Apostrophe | Interposition device to abbreviate notation to prevent repetition of two elements contained in a document title that have the same number |
| [ ] Square brackets | Investment strategy |

### 5.4.3 Sequence of symbols used in UDC (Filling Order)

| Logo | Example |
|---|---|
| = | = 30 German language |
| (0.) | (0.035.22) Microfilmed documents |
| (1/9) | 430 Germany |
| (= ...) | (= 1.410) British nationals |
| " =" | "19" Nineteenth Century (Nineteenth Century) |
| + | 622.341.1 + 669.1 Iron-ore mining and ferrous metallurgy. (Mineralization and Ore Metallurgy in Iron Mines) |
| / | 622.341.1/2 Mining of iron ores and manganese ores |

Note

| Simple No. | .341.1 Iron-ore mining |
|---|---|
| : | 622.341.1 : 338.124.4, Economic crisis in metallurgy in iron mines (class number can also be anagrammed as 338.124.4 622.341.1. |
| :: | 622.341.1 338.124.4 Economic crises in iron-ore mining (may not be reversed) |
| = | 622.341.1 = 30 Documents in German about iron ore mining<br>= 622.341.1 (0.035.22) Microfilmed documents about iron-ore mining<br>= 622.341.1 (430) Iron-ore mining in Germany<br>= 622.341.1 (= 1.366) Iron-ore mining among the ancient British<br>= 622.341.1 "18" Iron-ore mining in the nineteenth century |

## 5.5 Web Dewey

Edition 22 is the first edition of the Dewey Decimal Classification to be produced in the context of the web environment. Hence, Web Dewey means the database of Dewey Decimal Classification on WWW (World Wide Web). WWW is the network of computers across the world interconnected together on the internet, and using the concept of hypertext to link internet sites and information ion the internet. Like print versions of DDC, Web Dewey is also found in both full and abridged versions; they are Web Dewey and Abridged Web Dewey.

### 5.5.1 Key features

(i) All content from DDC 22 and Abridged Edition 14, including quarterly updates (new developments, new built numbers and additional electronic index terms).

(ii) An easy-to-use, browser-based interface that allows you to search the DDC (and related terminology) efficiently and navigate intuitively.

(iii) Thousands of Relative Index terms and built numbers not available in the print DDC.

(iv) Library of Congress Subject Headings (LCSH) that have been intellectually mapped to Dewey headings by DDC editors.

(v) Selected LCSH mappings from the Forest Press publication, People, Places & Things.

Note

(vi) LCSH that have been statistically mapped to Dewey numbers from records in WorldCat (the OCLC Online Union Catalog).

(vii) Links from mapped LCSH to the LCSH authority records.

(viii) Quarterly updates, incorporating the latest changes to the classification and new LCSH mappings, index terms and built numbers.

(ix) An annotation capability, which allows you to add your own notes into Web Dewey to reflect local classification practices.

### 5.5.2 Web Dewey Updates

One of the main features of Web Dewey is its updates. OCLC make updates periodically to the data accessible via the Web Dewey or Abridged Web Dewey service, using its best efforts to make these updates available on a quarterly basis. The updates available till date are as follows :

| | | | |
|---|---|---|---|
| 2007 September | 2007 May | 2007 February | 2006 November |
| 2006 August | 2006 May | 2006 February | 2005 November |
| 2005 August | 2005 June | 2005 February | 2004 November |
| 2004 August | 2004 April | 2004 February | 2003 December |
| 2003 November | 2003 October | 2003 August | 2003 July |
| 2003 June | 2003 April | 2002 November | |

### 5.5.3 Latest Enhancements

This release is scheduled for the September 23, 2007 install. This September release consists of updating the DDC 22 and Abridged 14 databases in Web Dewey and Abridged Web Dewey respectively. These database updates contain the latest changes to the DDC and monthly postings (New and changed entries) for May 2007, June 2007 and July 2007.

This September 2007 release includes :

(i) All updates to Dewey Decimal Classification, Edition 22 and Abridged 14 through August 2007 (corrections, new developments, new built numbers, and additional electronic index terms).

(ii) Thousands of Library of Congress Subject Headings (LCSH) that ave been statistically mapped to Dewey numbers from records in WorldCat (the OCLC Online Union Catalog) and intellectually mapped by DDC editors.

(iii) Thousands of Relative Index terms and built numbers not available in print.

Note

(iv) Links from mapped LCSH to the LCSH authority records.

(v) Selected mappings from Medical Subject Headings (MeSH).

(vi) LCSH that have been intellectually mapped to Dewey headings by DDC editors, including mappings to Abridged Edition 13 numbers from the OCLC publication, Subject Headings for Children.

(vii) Mappings between abridged Dewey numbers and the 2005 update to the 18th edition of H.W. Wilson's Sears List of Subject Headings.

### 5.5.4 Subscribe to Web Dewey

A Web Dewey subscription can begin at any time and extends for one year from the starting date. Requires an annual fee only; no per-use charges. It is available on a single-user or site-license basis.

There are two ways to order :

(i) Online using OCLC's secure Web store

(ii) By mail or fax using a printed order form

Web Dewey is available on an annual subscription basis, according to the table below.

| Product | Type of License | Subscription Price |
|---|---|---|
| WebDewey (OCCLC Cataloging Services Edition) | Single user (for use by one staff member) | $250.00/year |
| | Site License (for use by 2-9 staff members) | $250.00/year |
| | Site License (for use by 10 + staff members) | $825.00/year |
| WebDewey | Single User (for use by one staff member) | $300.00/year |
| | Site License (four use by 2-9 staff members) | $665.00/year |
| | Site License (for use by 10 + staff members) | $975.00/year |
| Abridged Web Dewey (OCLC Cataloging Services Edition) | Single user (for use by one staff member) | $72.00/year |
| | Site License (four use by 2-9 staff members) | $165.00/year |
| | Site License (for use by 10 + staff members) | $225.00/year |

Note

| | | |
|---|---|---|
| Abridged WebDewey | Single user (for use by one staff member) | $82.00/year |
| | Site License (four use by 2-9 staff members) | $200.00/year |
| | Site License (for use by 10 + staff members) | $275.00/year |

## 5.6 D.C.C. Problems & Popularity of DD

The main criticism of DDC is being made for its pure and limited notation. In which there is no proper solution to the problems of classifying new subjects in their proper place and long class numbers. Apart from this, its westward orientation has also been a major cause of criticism. But despite these problems and criticisms, D.D.C. is the most popular general classification system used in more than 2 lakh libraries in 135 countries of the world. In the United States, 95 percent of school libraries, 25 percent of other academic libraries and 20 percent of specialized libraries use it. It is translated in Hindi, Arabic, Chinese, French, German, Greek, Hebrew, Italian, Korean, Spanish, and Vietnamese 30 other languages. DDC has been used in 62 national and many merchant music lists. DDC has the largest number of Indian libraries among Asian countries. Class Number of Mark (MARC) is also given in Tapes and in CIP data.

### 5.6.1 Important Terms

Mark bibliographic data is a set of different formats for use by computers that provide a set of conversions for their identification and adjustment.

Unimarc is an abbreviation for Universal Marc, the credit for its creation goes to Ifla. Its aim is exchange of data internationally in machine readable form between various national agencies. Web Dewy is available on the Internet in the form of multiple learning points and presentations.

## 5.7 CRG

Classification Research Group is a group of working librarians and others interested in classification research in London. When Ranganathan's ideas of faceted classification began to make an impression in the Western World, the Royal Society's Scientific Information Conference was held in London in 1948, where classification as a method of subject organization, was discussed as one of the themes. Dissatisfied by the prevailing methods of subject organization, a committee with J.D. Bernal as secretary was set up to examine the existing systems and suggest possible improvements. No progress was however, made until 1951, when B.C. Vickery was invited to form a group to take over the work of the committee. This heralded the formation of CRG in 1952. The group consists of people who are keenly

Note

interested in classification research. A perusal of the reports produced by the CRG indicates that the group has been actively involved in the creation of several classification schemes for such organizations. The theoretical work of the group has involved the study of facet analysis, relational operators and the theory of Integrative levels." The efforts of the group are directed in two directions : Classification and data mining. The classification effort focuses on both methodological research and particularly novel, non-standard application. The work in classification has significant overlap with other areas, including machine learning and pattern recognition, so that the publications appear in a wide literature.

## 5.8 BSO (Broad System of Ordering)

Unesco in 1971 needed a switching language for its UNISIST Programme. Work for such a language started in 1972 as a part of European Communication Project, after the Budepest meeting in 1972, FID set up a new working group ..... FID/SRC whose scope was set out as "For purpose of interconnection and cooperation between information systems, the working group will design and develop a broad subject ordering scheme for all fields of knowledge and usable in manual and mechanised information systems." The first draft was prepared by Eric J. Coates and G.A. Lloyd. Third edition was published in 1978 by the FID.

Its major subject areas are outlined below :

| | | | |
|---|---|---|---|
| 100 | Knowledge in general | 500 | Humanities, Cultural and Social Science |
| 200 | Science and Technology | 600/890 | Technology |
| 300 | Life Sciences | 910 | Language, Linguistics and literature |
| 460 | Education | 940 | Arts |
| 480 | Sports and games | 970 | Religion |

It was designed to serve as a switching language *i.e.* to work as an intermediary or conversion language to transfer information from one indexing language to another. It means that is was designed for classifying information centres rather than documents. But, instead of its intended purpose, it is being used as shelf classification in some libraries.

## 5.9 Automatic Classification

Automated text classification (ATC) is a relatively new research field spun off from traditional information retrieval research. ATC is defined as the task of automatically classifying documents into a pre-defined set of classes.

Note

Rule-based methods and Machine Learning are two dominant approaches to ATC until the late '80s and since the early '90s, respectively. In rule-based approach, knowledge engineers and domain experts define a set of rules to be applied for identifying the class of unclassified documents. In ML approach, the parameters of the models are statistically determined using a corpus of manually pre-classified documents. The categories of unclassified documents are predicted from the parameterized models.

**Automated Classification Projects based on Large Classification Schemes**

In recent decades, classification schemes-based ATC projects have been undertaken in the information community as an effort to arrange heterogeneous information resources under one single organizational structure. This paper focuses on the approaches for classification scheme-based ATC and presents an analysis of the conceptual models used in the ATC projects. A total of 16 research projects and studies published across the period from 1992 to 2011, which are outlined at the table 1. The selection of the projects and studies are not randomly made. Instead, its primary target is to collect all studies based on the three popular library classification schemes, Library of Congress Classification (LCC), Dewey Decimal Classification (DDC), and Universal Decimal Classification (UDC). Other large scale classification systems such as web directories are also considered for the collection. A project and study often utilizes more than one model for a comparative study with the adopted methods; the first column of the table below indicates the key classification models only.

**Table : ATC research projects based on large classification schemes in chronicle order.**

| Key Classification Models | Classification Projects/Studies | Classification Schemes/Best Result |
|---|---|---|
| Clustering algorithm : creating a classification cluster for an LCC class, with inner product for matching: between classification cluster vectors and documents vectors | Automatic selection of Library of Congress Classification numbers based on the titles and subject headings in MARC records in LCC class Z (Bibliography, Library Science and Information Science) : (Larson, 1992) | Library of Congress Classification (LCC) About 47% in accuracy with the test set of 283 MARC records into major ranges of LCC class Z |


| | | |
|---|---|---|
| Term weighted algorithm | Automatic classification of a total of 660 databases in a Wide Area Information Server based on keywords, descriptions, and subject terms in the two top levels of UDC : Nordic WAIS/World Wide Web project (Ardö et al., 1994) | Universal Decimal Classification (UDC) A 'fair' number of misclassifications |
| Vector space model : Latent Semantic Indexing that associates terms with the relevant categories in the LCC Outline | Automatic classification of news articles in Usenet newsgroups based on news article headers in the 4214 classes of the LCC Outline : Alexandria Digital Library Project (Dolin et al., 1997) | LCC No evaluation result has been reported |
| Vector Space model: with a version of SMART (System for Manipulating And Retrieving Text) system | Classifying web resources into DDC: Scorpion project (Thompson, Shafer, &Vizine-Goetz, 1997) | DDC No evaluative outcome is reported |
| Clustering algorithm : a cluster corresponding to a DDC class, with cosine normalization for the distance measure between clusters and input documents | Classifying web pages into DDC : Wolverhampton Web Library (Jenkins, Jackson, Burden, & Wallis, 1998) | DDC Report 40 percent accuracy without further details in the experimental setting. |
| Vector space model: using term weight-based algorithm | Classifying web pages: DESIRE (Koch & Ardö, 2000) European Union-supporting Project | Engineering Information classification system Report about 60% accuracy in a test with about 1,000 web pages |

Note

| | | |
|---|---|---|
| Vector Space model with the cosine angle to measure the similarity between web pages and subject categories | Classifying web pages: (Chung & Noh, 2003) | DDC<br>0.77 in precision over 757 subclasses of the DDC main class of 'Economics' (330) |
| Machine Learning :<br>Support Vector Machine to create a linear classifier combined by pairwise | Automatically assigning a LCC to a work given its set of Library of Congress Subject Headings (LCSH):<br>INFOMINE (Frank & Paynter, 2004) | LCC<br>55% in accuracy in a test with a collection of 50,000 LCC/LCSH pairs for the entire LCC Outline |
| classification for classifying websites | | |
| Probabilistic model: with Naïve Bayesian (NB) | Classifying scientific papers about computer science : (Avancini, Rauber, & Sebastiani, 2004) | ACM classification scheme consisting of 1,474 categories<br>No test evaluation is reported |
| Probabilistic network model : with ExpNet (Yang, 1994) | Classifying web pages in Google and LookStart web directories (Avesani, Girardi, Polettini, & Sona, 2004) | Google web directories & LookSmart web directories<br>A range from 9.8% to 55.1% in F measure with the subdirectories of 9 top-level Google directory categories and a range from 6.8% to 39.5% in F with the subdirectories of 10 top-level LookSmart directory categories, in a test of classifying about 10K web pages |


| | | |
|---|---|---|
| Probabilistic model: with hidden Markov model (HMM) | Classifying dissertation abstracts : (Yi, 2005) | LCC<br>Higher accuracy classification with HMM over with NB in classifying 625 dissertation abstracts into 25 secondlevel subclasses in Q (Science), S (Agriculture), and N (Fine Arts): More than 45% accuracy with HMM vs. less than 5% with NB |
| Machine learning technique with Support Vector Machine | Classifying web pages: (Liu et al., 2005) | Yahoo! web directory<br>As low as 0.24 in F measure for about 246,000 categories |
| Term weight-based method | Classifying web pages into classes of the Ei classification system : (Golub, 2006) | Engineering Information thesaurus and classification scheme<br>No formal evaluation is reported in classifying a sample of 1,000 web pages (preselected and classified by librarians) |
| Clustering algorithm: with k-nearest neighbors (KNN) and the cosine similarity measure | Classifying Reuters articles : (Pong, Kwok, Lau, Hao, & Wong, 2008) | LCC<br>Report 0.8 in F measure as the best performance in a test of classifying 254 documents into 67 LCC classes or subclasses |
| Machine learning technique with Support Vector Machine | Classifying bibliographic data : (Wang, 2009) | DDC<br>Report a range from 0.49 to 0.85 in F measure in various tests with 4,861 DDC classes |

Note

| | | |
|---|---|---|
| Citation network-based term weighting method | Classifying scientific literature into DDC : (Joorabchi & Mahdi, 2011) | DDC A range from 0.61 to 0.89 in F measure in classifying 400 documents into 8-level DDC sub-classes |

**ATC Studies : Approaches and Models**

Below I summarize the key models/algorithms that have been adopted and utilized in the classification projects and studies.

- **Vector Space model (5)*** : A set of categories and test documents are represented as vectors in a multi-dimensional vector space. The degree of similarity between categories and documents are generally evaluated by a cosine of angle between category vectors and document vectors.
  - General vector space model (Chung & Noh, 2003; Godby & Stuler, 2001; Koch & Ardö, 2000; Thompson, Shafer, & Vizine-Goetz, 1997).
  - Latent Semantic Indexing (Dolin et al., 1997)
- **Clustering algorithm (4) :** Clustering algorithm group similar objects into a same subgroup called cluster. In classification, a new document is assigned to the majority class of the cluster closest to the document.
  - K-NN (Pong, Kwok, Lau, Hao, & Wong, 2008)
  - Other clustering algorithm (Jenkins, Jackson, Burden, & Wallis, 1998; Larson, 1992)
- **Machine learning (4) :** Machine learning is a statistical learning process that builds a classifier by acquiring the characteristics of the categories from previously existing examples.
  - Support Vector Machine (Frank & Paynter, 2004; Liu et al., 2005; Wang, 2009)
  - Hidden Markov model (Yi, 2005)
- Term weighting algorithm (3) (Ardö et al., 1994; Golub, 2006; Joorabchi & Mahdi, 2011): A metric to measure the degree of similarity (or score) between subject categories and new document is based on term weighting schemes involving indexing terms and subject terms.

* The number inside the parenthesis next to the classification model/algorithm refers to the number of classification projects and studies in which the classification model/algorithm is used as its key model/algorithm.

Note

- Probabilistic model (2): Given a sample input, probabilistic model-based classifier is able to predict a probability distribution over a set of classes.
  - General networks (Avesani, Girardi, Polettini, & Sona, 2004)
  - Naïve Bayesian (Avancini, Rauber, & Sebastiani, 2004)

Automated classification models can be classified into the following two groups, depending on types of learning :

- Supervised model (Pong, Kwok, Lau, Hao, & Wong, 2008; Wang, 2009; Yi, 2005): The category learning process is "supervised by the knowledge of the categories and of the training instances that belong to them." (Sebastiani, 2001, p. 9) Training data include both the training documents and the correct categories.
- Unsupervised model (Chung & Noh, 2003; Golub, 2006; Jenkins, Jackson, Burden, & Wallis, 1998): The category learning process is made "in a way that reflects the statistical structure of the overall collection of input patterns." (Dayan, 1999) This model is not provided with the correct categories of the training documents.

The ways of dealing with automated classification can be divided into the following two based on whether or not considering structural relationships among categories :

- **Flat classification approach :** In flat classification, categories are considered as independent and the relationships among categories such as a hierarchical structure in a classification scheme are ignored. Thus, the computation cost is proportional to the number categories. In a large scale of categories, its cost increases to an infeasible level. Classification models find the most relevant class, based on the assumption that all classes are in a flat structure. (Larson, 1992; Dolin, 1998; Thompson et al., 1997).
- **Hierarchical classification approach :** In hierarchical classification, a large-scale classification task can be decomposed into smaller tasks. Such a divide-and-conquer strategy greatly reduces the computation cost. Classification models find the most relevant class which a new record is the most likely to belong to while considering the hierarchical structure of a classification scheme.

Note

## Exercises

### Very Short Answer Type Questions:

1. What is Web Dewy
2. What is Mark Bibliographic Data?
3. What do you understand by Online Classification

### Short Answer Type Questions:

1. Write a short note on UNIMARC.
2. Briefly describe the characteristics of the UDC system.

### Long Answer Type Questions

1. Describe in detail Marc's bibliographic data.
2. Describe in detail the Universal Unimarc.