

---

---

# Classification Model for Libraries in the Digital Environment

K S Chudamani

## **Abstract**

*There is lot of research carried out about the models of classification suitable for libraries in digital environment. UDC, CC, DDC, LCC are used in libraries all over the world. People are using existing Classification models and automating them for use in the digital environment. Though this is a right step, one has to look at the operational requirements of the digital environment. Keeping this fact in view, this paper analyses the requirements of the digital environment and points out the adaptations required for the traditional schemes to suit the digital environment .*

**Keywords** : Classification, Classification Models

## **0. Introduction**

Various classification models are used in the organization of knowledge in libraries. The well-known among them are

1. Dewey Decimal Classification (DDC)
2. Universal Decimal Classification(UDC)
3. Colon Classification (CC)

Each of these has certain common characteristics and distinct characteristics. They were adopted by different libraries for serving a variety of purposes like classifying books, articles and technical reports, for providing shelf access, etc. This paper explores the modifications required to these models to augment information retrieval in the digital environment

## **1. Classification Models**

DDC is enumerative and can be easily memorized. The notation is usually brief. Though it has few analytico - synthetic features, they are very limited and as such produces brief notation. This helps user to memorize the number easily. Also, its updated editions are produced regularly and hence it is possible to update the organization of documents without much effort from the classifier. Though its notation is mixed, most part of it is purely numerical. Therefore it is widely used all over the world.

Universal Decimal Classification was conceived as advancement to Dewey in organizing micro documents. Special libraries used it to organize journal articles as well as books. Its notation is also pure for most part of it. It makes use of a colon and other symbols to enable the combination of class numbers. But the notation is not brief when compared with DDC. Its numbers are long and cannot be memorized easily. But the class number is extensive enough to represent micro subjects. Chain indexing can be created for the hierarchies represented by each class. This has an added advantage in retrieval i.e. multiple hierarchies can be accommodated easily

Colon classification was designed by Ranganathan to overcome inherent deficiencies in the above schemes like they are enumerative, no sufficient scope for analysis and synthesis. To design a truly analytico – synthetic classification, he provides a facet structure for each basic subject using PMEST order which may manifest in different rounds and levels. The notation is mixed and makes use of connecting digits like :, ;, ,, ', ,, . His notation is very enriched. It has empty and emptying digits.

Kashyap compares the steps in designing a database with that of designing a scheme of classification. He compares the approach of Ranganathan in designing Colon classification with designing a database using the entity relationship approach. Accordingly, the following table gives the comparative features:

Designing a database	Designing a classification scheme
1. Identifying data entities(objects about which data is to be collected)	Identifying/defining the subject domains for which scheme of classification is to be designed
2. Selecting attributes of data entities of interest to the potential users	Selecting attributes of the entities constituting the subject
3. Selecting data model, a schema to map the entities and their attributes	Selecting classification model ( hierarchic, faceted, freely faceted) for mapping information about the entities(concepts/isolates)
4. Grouping/dividing the data entities by their common attributes and differentiating attributes(Characteristics)	Grouping/dividing the concepts/isolates by their common attributes and differentiating attributes(characteristics)
5. Organising, arranging the groups, subgroups, and units derived at step 4	Organizing, arranging the groups, subgroups and isolates derived at step 4
VERBAL PLANE	
6. Naming fields and data elements	Work on the verbal plane
NOTATIONAL PLANE	
7. Assigning tags to fields, coding ,etc	Work in the notational plane

This shows that Ranganathan's principles are very close to entity - relationship approach of database design. The use of the the freely faceted classification concept suggested by Ranganathan which is used in computerized bibliographic services like INSPEC etc will be a suitable model for libraries. More about this will be discussed in the next section.

Connecting digits and lengthy class numbers makes it rather difficult for the user to remember the class number in both UDC and CC. Also updated versions of the CC scheme are not brought out regularly making the work of the classifier harder. As it is well known, many a times it is difficult to identify the personality. It has extensive analytico synthetic structure both in analysis and notation. It is to be noted here that all these schemes provide a linearly arranged numbers and subjects will be arranged in a linear order. There is lot of closeness between UDC and CC as they are based on facet analysis. However digital environment requires that they should be made mode flexible to suit the new era.

## 2. Classification in the digital environment

One of the oldest abstracting service known is the Science Abstracts started around 1898. This service initially used a list of about 30-40 subjects for arranging the entries. Later on, it started using the UDC for arranging its entries. But with the furthering of the growth of knowledge, the inadequacy to represent micro documents according to UDC was felt. Also, the Science Abstract branched out into Physics and Electrical Engineering Abstracts and a new classification scheme was developed for both the subjects.

The computerized generations of the Abstracts were started in 1967. The Physics Abstract in particular assigns more than one class number to each record i.e. in a sense is very close to UDC model, but does not connect the classes using connecting digits. This can give the flexibility needed to automatically

---

separate the different classes and use the set theoretic approach to retrieval using the classification scheme For ex. 74.20C Union 81.20L . This is particularly helpful in the digital environment. Even online databases can be searched using this approach. What is needed is the classification scheme. It should be noted here that the concept space gets reduced with the use of classification scheme. CD ROM databases can also be searched using this logic. But system designers of the databases, though facilitate such a search, did not open up their classification schemes to users until recently. Now with the availability of such a facility, it might help the database searchers to employ this strategy for searching the database. Further this paper explores the possibility of using similar technique by the libraries to enhance their retrieval capability in the digital world.

### **3. Adoption of modified UDC classification model in the digital environment**

The libraries using DDC, UDC, or LC and even CC can all directly employ the strategy described above, i.e assigning multiple disjunctive class numbers to a book. Others can also adopt a suitable classification scheme of their choice and use the same model for organizing their knowledge. But the maximum number of classes that can be assigned may be 4-5 as against the four employed by the physics abstracts. This may be further increased if necessary, after some experimentation. But the location of the book should be in the primary subject which is to be given first. Provision of more than one class number may be further enhanced by providing more keywords, which may even be taken from special thesaurus with the use of proper modifiers. When this solution is thought of, the problems that may be encountered are:

1. Problems relating to use of standard subdivision
2. Problems relating to use of provisions for synthesising numbers
3. Increase in the number of books with the same class number
4. Thesaurus term different from the subject heading being used

The standard subdivisions are usually combined with the basic class to enhance subject enumeration. Here also the first number assigned can be the usual number assigned to the document along with the actual subject represented by the number. However when numbers are synthesised, they can be limited to a fewer digits to properly express the subject. Additional numbers may be given which represent other subjects covered by it. This includes the subject added by synthesis. Standard subdivisions pertaining to space isolates can be separately assigned as a class number. Cross-references to the records displayed will enhance retrieval.

In the same manner synthesizing of the numbers can be handled. If there is an increase in the number of books with the same class number, then individualizing by year can be thought of. This will be a useful solution for the digital environment also as it has been used as an individualizing element in library classification. Whenever thesaurus and subject headings are used together there may be a mismatch. Then authority files with proper cross references need to be created which is being already done by libraries.

### **4. Impact on retrieval**

The adoption of the above described model helps in the creation of a concept space with a manageable dimension. Chudamani and Asundi use this to create a knowledge base of subject relationships as described in a paper submitted to the OCLC-NISSAT-DRTC workshop on information resource management. This knowledge base when augmented with thesaurus and keyword terms can become a system that enhances retrieval effectiveness using set theoretic rules. This can be achieved by creating

---

a list of all subjects related to a subject as described in the paper. Also navigation using class number hierarchies is possible.

## 5. References

1. Chudamani, K S and Asundi, A Y, use of knowledge base of subject relationship for enhancing retrieval effectiveness in the digital environment, OCLC-NISSAT-DRTC seminar on information resource management, March 10-12, 2002, Bangalore
2. Ranganathan, S R, Prolegomena to library classification, Bombay, Asia, 1967
3. INSPEC, Physics abstract,2002(CD-ROM Version)
4. Kashyap, M M,, Similarity between Ranganathan's postulates for designing a scheme for library classification and Peter Pin-Sen Chen's entity relationship approach to data modeling and analysis, Desidoc Bulletin of information technology, 21(3)2001

## About Author



**Dr. K.S. Chudamani** is Deputy Librarian at Indian Institute of Science, Bangalore – 12, India and holds Masters in Physics, LIS and Information Studies (University of Sheffield, UK) and Ph D. Earlier she worked with Indian Institute of Management, Bangalore and Indian Statistical Institute, Bangalore.

**E-mail : [ksc@library.iisc.ernet.in](mailto:ksc@library.iisc.ernet.in)**