DESIDOC Bulletin of Information Technology, Vol. 21, No. 1, January 2001, pp. 3-19

Classified Catalogue Code of Ranganathan: A Proposal to Make it Compatible for Developing Computer-Based Library Information Systems

Madan Mohan Kashyap

Abstract

This paper deals with the differences between the environments of card catalogue and online catalogue, and emphasises on the need of developing computer-based library information systems and services. It describes database technology, kinds of databases, database management system, computerised library information system, and management information system. It covers in detail the database design and compatibility of cataloguing codes for developing databases of computer-based library information systems.

1. INTRODUCTION

The introduction of information technology in libraries, development of online catalogues and bibliographic databases, availability of documents in vast array of new media and formats, and huge amount of information sources made available by the World Wide Web (WWW) (the multimedia part of the Internet) necessitate necessary changes in the rules of Ranganathan's Classified Catalogue Code (CCC)[1] to make it an effective tool for organising and retrieving information in the changed conditions. If Ranganathan were alive at this time, he would have brought radical changes in the rules of CCC to address the changing needs.

The *Canon of Context*, one of the principles of cataloguing, propounded by Ranganathan, mandates that the rules of a catalogue code need to be amended, updated and formulated to keep in step with changes in the features of book, the mode of book production, the nature of organisation of libraries, particularly in the mode and quality of library services, in changed conditions[1]. Thus, in the context of accelerating environment changes faced by the libraries, we must consider to bring radical changes in the rules of Ranganathan's CCC, as decreed by the *Canon of Context*. To meet the need of changed scenario, we must address the need of preparing effective guidelines or comprehensive code for (i) building online library catalogues of vast array of old and new bibliographic materials, (ii) library networking, and (iii) searching and downloading information from online databases of libraries, publishers' catalogues and Internet.

Today, online library catalogues hold a large number of records that were earlier constructed for card catalogues and retain more or less the same underlying structure. The proposed changes in the CCC rules must support conversion and organisation of old bibliographic data available in card catalogues into electronic form in a suitable manner, creating bibliographic records of documents now available in manifold physical forms in various kinds of formats (text files, sounds images, video clips), and storing, organising, and retrieving bibliographic information from online databases of library networks, in a way

that enhanced benefits or better results are achieved by using traditional library techniques along with electronic technology and its applications.

To bring the required changes in cataloguing rules, it is necessary for us to understand the differences between the environments of card catalogue and online catalogue, and the need of developing computer-based library information systems and services.

2. CONTRAST BETWEEN CARD CATALOGUE & ONLINE CATALOGUE

2.1 Card Catalogue

The design of the card catalogue is a model based upon the following objectives, which are restated in line of Cutter's approach[2]:

- To enable a person to find bibliographic items whose one or more attributes, such as author, collaborators, title, subject, etc. are known to him.
- To show what works or bibliographic items the library has:
  - (a) by a given author
  - (b) on a given subject
  - (c) belonging to a series.
- To assist him in the choice of the item, as to its:
  - (a) edition
  - (b) date of publication
  - (c) language and script
  - (d) physical format.

To achieve these objectives; the traditional library codes:

- Lay down the laws, principles or canons and rules of cataloguing.
- Identify and describe the attributes of various types of bibliographic entities or items. (Each bibliographic entity has particular properties or attributes which describe it and whose structured description ultimately helps a person to identify a specific bibliographic item or set of items having similar attributes).
- Specify the source(s) from where the information for the description of each type of bibliographic material or item must be collected.
- Fix the order according to which the entire bibliographic information (or the elements of description) must be organised and various levels of detail in the description.
- Describe the various access or search elements, which are placed at the top position or leading section of catalogue entries to provide access or search points. Access elements constitute names, codes, words or phrases, etc., under which bibliographic records are searched and identified.
- Specify the rules for choice and renderings of each type of access element in a catalogue entry.
- State the rules for the choice and rendering of headings, the leading elements of main entries of a dictionary catalogue (referred to as main entry headings) or the constituents of the heading sections of main entries of a classified catalogue.
- Delineate the order and the style of presenting the value of the attributes or data elements of a bibliographic item on a card (i.e., specify order and style of elements of description or output formats of various types of bibliographic items).

The rules of traditional library codes have been framed for well structured, but fixed display of information or data pertaining to the bibliographic items on cards The order of representing the values of the distinct units of bibliographic information is fixed, starting from the leading element or heading, generally consisting of a name of an author(s), or title of the work, followed by the title of the work, and statement of responsibility, etc.

The organisation of elements of description of various types of bibliographic items, namely, monographs, multi-volume monographs, composite monographs, periodical publications or serials, individual contributions in

monographs or a serials (i.e., independent works or articles forming part of composite monographs or serials and treated as an bibliographic entity) is divided into certain defined sections or areas by library codes. Each of which includes specific attribute or a set of attributes (elements of descriptions or data elements) of the information entities that describe them.

The information about bibliographic items included in their records or catalogue entries, and the formats of displaying the information used by the various codes do not differ very much. Whatever trivial differences in the order or in the form of presentation of information about a bibliographic item exist, the overall information produced or displayed is always similar or the same and is significantly meaningful to the end users.

2.2 Online Catalogue

Online catalogue is essentially a computer-based replacement of the old card catalogue and its purpose and functions almost remain the same. However, the online catalogue is a value-added system, which supersedes the functions of traditional catalogue. It is beyond the capabilities of card catalogue to provide keyword search, Boolean searching and to provide interactive instructions or search facilities to users with possibility of getting outputs in the various formats and information contents they require.

As pointed out by Potter: "The power and flexibility of the online catalogue have caused it to surpass the card catalogue in searching capabilities and have also caused its configuration to change to include three functions that were formerly distinct and unconnected. These three functions are the library catalogue, circulation control, and authority control."[3]

Today, an online catalogue reveals to the reader, on the basis of his/her query, whether a book or set of books of his/her interest are, on order, under processing, on the shelf, on loan, sent for binding, misplaced or lost, etc. By developing single or integrated database for all or near all library functions, or by way of linking the catalogue, the acquisition system and the circulation system, a library provides bibliographic and status information to the user and staff at one spot, the terminal.

In card catalogue environment, the user moves from physically visible drawer to drawer or from card to card, whereas with online catalogues (more so in network environment), he or she is faced with a sort of invisible and indiscernible universe. The situation as this presupposes that the bibliographic items not only be described and indexed according to rules established for physical description (i.e., cataloguing rules) but also for that of the contents. In this context, the traditional subject indexing technique(s) coupled with online query function and efficiency of search mechanism of an online catalogue becomes paramount element.

The online catalogues, though, add a layer of functionality by providing various computer supported techniques for searching the data, such as use of system dictionaries, Boolean operators for combining semantic categories or terms regardless of any special citation order, stemming, and truncation etc. At the same time, they also add a layer of complexity to this process, because of the hidden nature of bibliographic records. In recent studies, it has been found that the online catalogues are more difficult to use and less serviceable than the card catalogues. Baker in his article published in 1994 in *The New Yorker* put forward this criticism more lucidly, which was widely discussed by the library community.[4]

The limitations of online catalogues can be removed using traditional facet classification and subject-indexing techniques along with computer supported techniques. The use of (i) facet classification and subject indexing techniques for developing post-coordinating retrieval systems for an online catalogue, which use Boolean operators for combining semantic categories or terms regardless of any special citation order, and (ii) truncation of class numbers, particularly provided with the aid of facet classification for retrieving hierarchically different sets of documents has been highlighted by many. Since, Ranganathan demonstrated for the first time the application of his postulates for facet analysis and synthesis,

and the other associated principles in assigning subject headings[5], significant advancement has taken place in the application of his postulates for facet analysis and synthesis in the areas of computerised contextual indexing as well as searching computerised online databases. Besides development of computerised indexing system, such as PREserved Context Indexing System (PRECIS)[6], POstulate Based Permuted Subject Indexing (POPSI)[7,8,9] and Contextual Indexing and Faceted Taxonomic Access System (CIFT)[10], Ingwersen and Wormell[11] have been found to envisage that Ranganathan's Principles for Knowledge Organisation and Facet Analysis Methodology are likely to find application in advanced information retrieval, particularly in the areas of cognitive modelling, neural network techniques and pattern recognition.

Ellis and Vasconcelos[12] point out the continuing relevance of facet analysis as a technique for searching and organising Web-based documents. They argue that facet analysis, which is essentially a concept based approach to indexing is an excellent approach to searching and organising the WWW searches rather than using either search engines or search directories. Finally, they argue that the underlying philosophy of facet analysis is better suited to the disparate nature of WWW resources and searches than the assumption of current information retrieval research.

Neelameghan[13,14] finds his General Theory of Knowledge Classification and Facet Analysis Methodology relevant in designing specialised databases. Godert[15] demonstrates how querying by classification facets or categories allows retrieval of documents that would be more difficult with verbal method.

An empirical study conducted by Iyer, Hemalata[16,17] indicates that structuring of search queries using Ranganathan's Theory of Classification proves very effective in the process of searching bibliographic databases. This approach serves the following purposes:

- Assists in the choice of concepts from the user's narrative statements representing their information needs.
- Assists in formulating search statements by providing a basis for the use of appropriate Boolean operators. Terms representing different facets are combined with the operator 'AND' and those within a facet, or representing different levels of the same facet are combined with the operator 'OR'. The rational of this is based on the degree of bond-strength between facets.
- Provides a method for systematically dropping terms, if the search needs to be broadened. Terms are dropped from the right end of search statement.

Svenonius[18] have developed a model system named DORS (Dewey Online Retrieval System) as an interface to an online catalogue. Its major feature is an automatically constructed chain index, in which a term is shown in as many verbal strings as are needed to show the hierarchical contexts in which a concept occurs. One can develop similar model using Colon Classification. As a matter of fact, Ranganathan's Colon Classification Scheme will prove better than Dewey, because it represents classes or concepts in more strict hierarchical fashion and can provide a much better mean to browse the records of databases (particularly very large databases, including electronic library) arranged in classified order. One can browse and retrieve the relevant information from the online catalogues with the aid of class numbers, specific subject headings, and chain index generated with the aid of Ranganathan's postulate-based subject indexing methodology. The analytico-synthetic approach or facet analysis and synthesis methodology of Ranganathan, based on his theories, principles, and postulates of knowledge organisation, can be effectively used in designing database schema, expert systems,[19,20] thesauri and computer generated subject indexes, and formulating query statements for searching computerised online databases.

2.3 Difference in the Process of Creating a Bibliographic Record for a Card Catalogue & Online Catalogue

To create the record of a document for an online catalogue, the data needs to be entered or structured in one format (called input format), physically stored on an electronic

storage device in another format (called internal or processing format) and displayed in one or more formats, keeping in view the need of different groups of users (called external formats, end users view formats, or output formats). Records in internal or processing format are stored and maintained for retrieval and output. This format is generally predetermined according to the requirements of software (the database management system) used.

In a card catalogue, as one finishes recording the data about a bibliographic item the output is ready. That is, no distinction exists between the input format and the output format of a bibliographic record. However, in digital environment from one input format one can create one or more styles of outputs. One to one correspondence does not exist between input format of an entry and the types of outputs one can get for display of information.

The methods of inputting data in a computer may vary widely from one library to another. Data may be entered directly by the professional staff responsible for creating the bibliographical record and description of information in standard way according to the rules of the catalogue code being followed. Alternatively, a person may fill in the worksheets, which are later checked by the professional staff and the data operators/typists key data in. On-screen (electronic) worksheets and/or general-purpose text-editor may facilitate keying.

Concluding Remarks

The library users make use of library catalogues for three purposes:

- To find whether the known bibliographic items are available in the library collection and if so, what are their locations?
- To find what knowledge resources (bibliographic items) a library possesses on a subject or topics of their interest.
- To seek information about bibliographic items held in a library or in a network of libraries, and make list of those that may be of interest to them.

To help the user to find known bibliographic items, author, title and other access points are provided in card catalogues as well as in online catalogues. To locate the unknown bibliographic items that may be of interest to a user, subject access and other approaches have been developed and used by libraries. Author and title search is easy as compared to subject search, though these are developed various methods of organising bibliographic records according to subjects (e.g., various classifications and indexing techniques). The reasons being that very often users are not able to formulate exact subject queries. Formulating precise subject queries is difficult because the searcher must find ways to articulate his or her intended meaning using terms that match those in the catalogue, whether assigned by author, classifier, or subject cataloguer.

According to Borgman, "query matching is effective only when the search is specific, the searcher knows precisely what he or she wants, and the request can be expressed adequately in the language of the system (e.g., author, title, subject heading, descriptors, dates). Even with graphical user interfaces, the searcher must enter terms and specify relationships that match those in the database. Many, online catalogues allow users to browse authority file indexes, but usually only within the constraints of specific query."[21] Classification access is rarely offered in online catalogues. Perhaps, in an online system, if one can provide browsing facility as provided in a classified Card Catalogue, it could be of immense help to the information seekers, particularly who are not experts in subject knowledge, semantic knowledge (that is, the knowledge of how to implement a query in a given system, e.g., knowledge of Boolean logic). [The difficulties with subject query formation have been analysed by Bates[22]; Crawford, et al[23]; Hildreth[24]; Lancaster, et al[25].

While, revising the CCC, one must also seriously address the issues related to the creation of online catalogues of bibliographical materials now available in vast array of new media, efficient retrieval of bibliographic information, and provision of facility for browsing the records in classified order in addition to author, title, subject search. One

must also address the problems related to retrieval of information using natural language descriptors, faceted classification class numbers combined with chain indexing, and/or PRECIS strings, Boolean search technique and browsing of catalogue or bibliographic records in classified order.

The difference between the environments and approaches in the construction of traditional card catalogues and digital catalogues forces re-examination of the rules of traditional catalogue codes, making them more suitable for the creation of online catalogue or electronic bibliographic databases. These rules need to be revised and updated keeping in view the ease and better data processing and retrieval facilities provided by the information technology and developing computerised library information systems and services using database technology.

3. DATABASE TECHNOLOGY

3.1 Introduction

The current practice of collecting, organising, processing and retrieving information by means of computers is called Database Technology. Database technology began to replace the file system in the mid-1960's and was characterised by the introduction of distinction between logical data and physical data. In a file system, there is rigid or unchanging association of certain data files or data sets with individual programs that operate upon these. In a database system, a collection of logically related data or data records is assembled in a set of linked or integrated files to facilitate the retrieval of information, which may be processed by one or more application programs.

The data becomes the central resource in database environment. Information systems are built around this central resource to give the end users flexible access to data. The end users of information systems or applications use the database management system (DBMS) to input, process, maintain, and access the data and to generate reports or outputs.

The data in a database is organised in variety of ways, and the particulars for structuring the data or data records within a database is called data model. A data model is a statement of data. Logical or conceptual data model specifies the rules according to which database records are logically structured and helps us to collect and analyse information to describe the logical database structure. Entity-Relationship (ER) model is widely used to describe logical database structure or conceptual schema of a database. Low level or physical data models, also called implementation-oriented models, provide concepts that describe the details of how data is stored in the computer storage device. The three most widely used implementation-oriented models are relational, network, and hierarchical models.[26]

3.2 Database

A database is an organised collection of structured data, units of information or data records, independent of any application in a computer memory that serves the needs of multiple users. According to Oxborrow[27], "a database is an organised collection of related sets of data, managed in such a way as to enable the users or application program to view the complete collection, or a logical subset of the collection, as a single unit". An organised collection of related sets of data or units of information (data elements) in a database is called record. A collection of database records constitutes a database file. The collection of sets of data or data records in a database are integrated to reduce data replication, and is managed in such a way that it can fulfil differing needs of its users.

A stored or physical record is a named collection of associated fields. A field contains a particular type of data or information within record that can be separately addressed. Thus, different fields in a bibliographic record can be 'author field', 'title field', 'imprint field', and so on. A field may be divided into smaller units called subfields comprising data elements. For example, 'title field' containing the name(s) of author(s) can be subdivided into still smaller units of information (data elements) or subfields, namely, entry element, other part of name, date of birth, etc. Each field and subfield is given a unique identifier, called tag and

subfield identifier, respectively. A physical record is thus composed of fields and subfields. A collection of records of single type constitutes a file. As such, a database can also be defined as a set of related data files containing facts or information about something, say about an enterprise, people, objects, entities, events or a certain problem area. Identifying what fields and subfields are to be included in an entity or object record and providing them unique field tags and subfields identifiers is an important task of database design.

A database can also be seen as a collection of inter-related, largely similar data or data records in a set of linked or integrated files designed to facilitate the retrieval of information, which may be processed by one or more application programs. Further, the files of the database are organised and administered in such a flexible way that these can be adapted to new, unforeseen tasks. In a database, the data records or related data files are stored with minimum redundancy and organised or structured in such a flexible manner that multiple applications or the functional systems share the entire data, and the end users get the information out of it very quickly. Primarily, a database has the following attributes:

- A database represents some aspect of a real world, sometimes called mini world, that is an enterprise or organisation, or the universe of discourse, study, research or investigation, whose database is to be constructed. The mini world or the universe, whose database is created, constitutes of specific entities and relationships that exist between them. By specific entities, we mean those entities that must be considered or are important in the context of an enterprise or universe whose database is being created.
- A database is designed and built for a specific purpose for the use of intended groups of users and some applications in which these users are interested.
- A database is designed to avoid duplication of data as well as to permit retrieval of information to satisfy information needs of wide range of users.
- A database permits shared access, enhances data independence by permitting application programmes to be insensitive to changes in the database.
- A database can be of any size and varying complexity.
- A database is created and maintained by a database management system (DBMS), such as CDS/ISIS.[28]

3.3 Kinds of Databases

Broadly, databases can be divided into two major categories, namely:

(a) Databases of primary sources of information (also called source databases), and

(b) Databases of secondary sources of information (also called reference databases).

Source databases provide information in electronic form or are databases of electronic documents, which contain primary information in electronic form. These databases can be grouped according to their contents. For example:

- **Full-text databases**—Contain the full text of documents.
- **Numeric databases**—Contain numerical data of various kinds.
- **Reference databases**—Lead the users to the sources of information: a document, a person or an organisation. They can be divided in three categories:
  - ***Commercial bibliographical databases***—These are the databases of indexing or abstracting services, booksellers and publishers, which include citations or bibliographic reference of monographs and articles published in serials, and sometime as abstracts of literature.
  - **Library bibliographic resources database**—Provide bibliographic information about the bibliographic documents available in a given library or about bibliographic holdings of a group of libraries.
  - **Referral databases**—Databases of electronic reference sources that contain

general information, such as name, address, specialisation, products, services etc., of persons, institutions, business enterprises, geographical entities, information systems, etc.[29]

3.4 Database Management System

A database management system (DBMS) is a generalised software system, which is composed of a set of programs that create, modify, store, manage, protect, and provide access to the database and enable the end users to retrieve information in a variety of ways. A DBMS supports three levels of data abstraction: conceptual schema, external schemas and physical view (organisation of records in electronic form that are normally inter-related by specific mechanisms e.g., indices, pointers, chains; means of locating or accessing physical records; and techniques used for inserting new records); data definition language; data manipulation language; and important utilities, such as data integrity, transaction management and concurrency control, crash recovery, and security.[30]

The genesis of DBMS is based on the idea of separating the database structure or meta-structure of a database from its actual contents. The certain benefits of database management system software are:

- Different users can share the data stored in the database files and all the requirements or information needs of the different users are fulfilled.
- A common and controlled approach is used for adding or inserting new data and deleting, correcting, modifying and retrieving existing data records or portion of it within a database, and
- Users and applications that access data need not be aware of the detailed storage structure of the data on a computer storage device. Thus, the data are stored in a way that they are independent of one or more applications that use the data.

3.5 Database Design: Approaches & Tools

Database design and its implementation is an essential process that is carried out while developing a computer-based information system for an organisation or person. The design of a database involves an abstract and general description of entity types and their attributes or data elements, whose values need to be stored in the database, and the description of data structures, that is, establishing the logical and physical relationships between entity records that comprise the database. This is achieved, firstly, by separation of the logical definition of data in the database from the physical implementation of data (i.e., storage of data on a storage device), and secondly, by providing application data independence, i.e., separation of overall or global logical data definition of the data into subsets, as viewed by different users in connection with their use and applications.

The logical data definition of data is the expression of the user or programmer's view of data or the way they think about data. It is posed in terms of data model. The data model may reveal little about exactly where each item of data is stored. A DBMS and an operating system work together to convert the logical view of data into physical view of data and vice-versa. The physical view is stated in terms of specific locations in storage device plus internal techniques used to find the data, which is, exactly what the machine does to find and retrieve the data from a storage device.

Thus, the design of an electronic database requires that we need to work at three distinct planes or levels.

(a) the conceptual or logical level of data definition,

(b) external level, to furnish user's views of data, and

(c) Internal level, to provide physical layout of data.

The designs of each level are called the conceptual scheme or model, the external schemes or models, and the internal schemes or models, respectively.

The conceptual schema or model, is application-independent and computer storage independent overall logical description of the database. The conceptual schema (also called global schema or model) is developed using

conceptual data modelling techniques, such as entity-relationship approach, which consists of both analytical and design methods and a data-modelling tool.

The external schemas or models, also called subschemas, are different user-oriented logical pictures of the database. A subschema describes only part of database or a data structure needed for a specific application in the form most convenient to that application. The main purpose of a subschema is to give a particular group of users or application an access only to that portion of the database, which is of their interest. The external schemas are specific parts or portions of the database that may be retrieved, re-organised and displayed in distinctive patterns or configurations, keeping in view the needs of various end users.

The internal schema or model is a layout for physical representation or storing of data on a storage device.

3.5.1 Conceptual Level Schema or Model

The conceptual level schema is an overall logical database description of a database of an organisation. The database schema is representation of that part of the real world that the database is about. For example, the part of real world that an online library catalogue is about all types of bibliographic items commonly collected by libraries, namely, monographs—simple and composite, serials, and independent contributions or work forming part of monographs and serials. Using a high-level conceptual data modelling approach, one can create a conceptual schema or model of a database. The conceptual schema (also called the conceptual or logical view of a database) is a precise description of the data requirements of users. A conceptual schema or model (also called metadata) can be defined using entity-relationship approach to data modelling and analysis and displayed by means of the graphical notation known as Entity-Relationship Diagrams (E-R diagrams).[31,32] One can also apply the postulates, principles of Ranganathan's Theory of Knowledge Classification in designing a conceptual schema or model for specialised databases.[13,14] There is also, similarity between the Entity-Relationship Data Modelling Concepts stated by Peter Chen, and others and Ranganathan's postulate for designing a scheme for library classification, the detail of which will be given in another paper to be published later in this journal.

3.5.2 Internal Level Schema or Model

The internal level schema relates to how the data are actually stored on a storage device. It involves defining the physical organisation of the data file on storage device; that is, the way the data are to be written onto and read from a storage device. It also entails decisions about the arrangement or establishing logical connection of the related physical data or data records on a storage device for the purpose of retrieval, processing, and display.

The layout or physical representation of data on a storage device is the concern of computer programmers or the DBMS. This is the lowest level representation of data. The physical organisation of the database, at this level, is often determined and rendered by the software or database system used, and internal implementation or organisation is unknown to the end users. Here, database is seen as a collection of internal records in electronic form, which are normally inter-related by specific mechanisms, e.g., indices, pointers, chains, and other means of locating or accessing physical records and techniques used for inserting new data.

The data in a database are usually logically and physically organised according to some data model. A data model is a collection of conceptual tools for describing data, data relationships, data semantics, and data constraints. Various experts have proposed a number of data models for the development of database systems. They provide a conceptual basis for logical organisation of related data pertaining to an enterprise or a problem area and physically structuring the database in a database system. In other words, they are formal basis for defining unambiguously the components of an enterprise—the concerned entities, their attributes (data elements), their inter-relationships, operations and a framework for logical and physical organisation of the enterprise data in a database system.

The databases are generally structured on one of the three data models namely, the relational data model, the network data model, and the hierarchical data model, or combination of these three, or some subsets of these three models.

3.5.3 External Level Schema or Model

The external level design, sometimes called subschema, relates to those particular views or formats of data, which are generated as outputs for specific purposes or to serve the needs of different groups of users.

The development of conceptual and external designs (user's views) of a database is the concern of library systems analysts or library database administrators. The physical design of database is the concern of computer programmers.

4. COMPUTERISED LIBRARY INFORMATION SYSTEM

Nowadays, computerised library information systems are developed not only to serve the purpose of online catalogues, but also to serve as comprehensive, unified or integrated systems covering near total library activities or functions, such as, acquisition of bibliographic documents and other materials required for running a library system, creation and upkeep of inventories of the purchased materials, membership registration and allied work, appointment of staff and maintenance of their service records, bibliographic items circulation control, upkeep of financial records, and so forth. Such a library information system serves not only as a bibliographical information system for library members but also as an information system for library management. It supports the operational, managerial, and decision making functions of different levels of library managers, besides providing information about the library bibliographic holdings and their status, to its end users.

In any organisation, a computer-based information system (also called computerised management system) is generally designed to support the operational and managerial decision making functions of the organisation and provides:

- Information necessary for performing day-to-day operations or for taking routine decision by the staff.
- Information to aid tactical planning and decision making (tactical planning is concerned with short-run planning: staff recruitment, re-deployment, promotion, the annual budgeting process etc.), and
- Information for the support of strategic or long-term library plans.

In contrast, a computer-based library information system (also called library management system) serves as an information system for management and as a bibliographic information system for library members. It provides:

- Bibliographic information about the holdings of a library and their status to the library members.
- Information required by library staff and management to carry out their day-to-day library activities, and
- Information which supports managerial and decision making functions of different levels of library managers, particularly that is useful to them for tactical (or short-run) planning and for a strategic (or a long-term) planning.

A computerised library information system integrates or combines near total library operations, and its subsystems are designed to carry out the essential library functions or tasks. These functional or task subsystems are named monograph acquisition system, serial acquisition and control system, bibliographic items circulation system, financial system, personnel management system, and so forth. These subsystems are very often considered distinct from a bibliographic information system designed to provide bibliographic information to library users. This actually is not the correct view.

The library system as a whole is a bibliographic documents handling system that acquires, processes and makes available bibliographic documents, and provides information about these documents and other bibliographic items that may not be available in the library to the library users. It also collects

bibliographic information from commercial or non-commercial bibliographic information systems or services and reprocesses it for internal use. Besides processing and generating bibliographic information, it also creates another distinct type of information found useful for strategic and tactical planning, and taking decisions about day-to-day operations for effective control of library functions and services, by the operational level and managerial level staff of the library system.

One must distinguish and clearly perceive the relationship between library information system as a whole and a bibliographic information system that generates bibliographic information for the use of library members. It may also be noted that a computerised library information system represents a much larger system within which information system for management is also a part, including bibliographic information systems. Information system for management is a subsystem of library information system that supports work practices and information need of library staff for effective management and control. It is often called management information system (MIS).

The end users of a bibliographic information system are library members, whereas the end users of an information system for management are the staff members working in the system, and both the systems constitute subsystems of a (integrated, federal or comprehensive) computerised library information system, also called library management system. A computerised library bibliographic information system in turn constitutes a subsystem of a library system as a whole.

A library information system may be defined as a set of library transaction processing (or functional) systems designed to provide information to library members and to support the operational, managerial, and decision making information needs of library staff. The kinds of concerned library information system includes a computer (hardware and software) as one of its components. The extent of computerised library information systems varies from library to library, but the most effective kinds are those which are integrated. An integrated library information system incorporates near all library functions and provides information to library members and staff according to their needs. An integrated computerised library information system is designed keeping in view that each library functional area is not only concerned with its own area information requirements and support for decision making, but they also share common information needs.

Thus, a computerised library information system or library management system can be viewed as a set of library functional systems encompassing:

(a) Library work practices or procedures

(b) Information technologies—computer hardware and software, having database management system, and a corporate database (it is not simply a bibliographic database but much more), and

(c) Work forces (i.e. library staff).

All geared together for accomplishing the system's tasks and to provide:

- Knowledge sources (i.e. bibliographic documents of all types) to library users for their learning, self studies and researches,
- Bibliographic and any other type of information needed by the library members, and
- Information necessary for the library staff for effective control and management of library affairs.

5. EXPERTISE REQUIRED TO DESIGN & DEVELOP COMPUTER- BASED LIBRARY INFORMATION SYSTEMS & SERVICES

Based on what has been discussed already, one can summarise that to design, develop and manage computerised library systems and services, one needs to have:

- Comprehensive or total view of a library system including its objectives, functions, processes or activities (task systems to be computerised) and the various entities associated with the system, whose database is needed to be created.

- Command on library tools and techniques (i.e. library classification, cataloguing and indexing techniques), and
- Expertise in the application database design and development techniques. One must be able to:
  - Identify and describe the entity types (objects about which information is to be collected for use) and their attributes and relationships (i.e. data elements).
  - Recognise, designate and define the essential and optional data elements that may comprise database record structure of an entity type.
  - Provide unique identifiers and names to the data elements, i.e., assign unique codes to identify the different data elements in the record. These unique codes are also called content designators, which may consist of field tags, indicators, and subfields' codes and serve as means of identifying data elements or values of fields in a machine readable record, and
  - Specify:
    - Guidelines or rules for identifying and using authentic sources for collecting data or pieces of information about an entity. That is, the appropriate sources from where the values of the attributes (data elements) of an entity need to be collected for description.
    - The domain of attributes that define the valid values of the attributes and guide decision in case there are exceptions.
    - Input data format or formats in which records of entities or objects will be input to the system (i.e. the form in which the given pieces of information about an entity will be logically recorded).
    - Rules necessary for uniform or consistent description or rendering of the values of the attributes (data elements), or contents of fields and subfields comprising database records.
    - Formats in which entity records or information will be displayed (i.e., the formats of required outputs or reports).
    - The type of access points or elements which could prove very useful for user to locate or search the relevant records of bibliographic items from a database, and the method or techniques, which may be used to construct the elements of search or to formulate queries to search the database.

In other words, collectively or blended, these techniques form special skills to design and develop computer-based library systems and services.

The use of computer technology in libraries forces us to examine and describe afresh almost all traditional library policies, rules (including cataloguing rules), procedures, and techniques as a whole and incorporating certain additional elements from the database design and development technology.

6. COMPATIBILITY OF CATALOGUING CODES FOR DEVELOPING DATABASES OF COMPUTER-BASED LIBRARY INFORMATION SYSTEMS

The traditional cataloguing codes, like AACR-II[33] and CCC of Ranganathan[1], presently only serve as tools to develop manual catalogues and to limited extent electronic bibliographic databases. To develop a database of library information system as a whole, one needs to:

- Expand the scope of these codes.
- Identify and define all possible entities or entity types that are connected with library systems, including the various types of bibliographic documents added in a library collection.
- Make efforts to revise them so that they cover the rules and guidelines to create databases of these entity types.

These codes need to incorporate guidelines to define conceptual schema, and logical data

design or structure of a library database. One must also:

- Absorb the components of Common Communication Format (CCF)[34,35,36,37,38,39] in these codes.
- Include the traditional library techniques for knowledge organisation, subject indexing and information retrieval appropriate for computer environment, such as facet analysis approach to classification and indexing devised by S. R. Ranganathan.
- Include the techniques for query formation and search strategies appropriate for computer environment.

The revised code developed incorporating all the aspects discussed may be called Code for the Design and Development of Computer-based Library Databases or Information Systems and Services.

Critical analysis of the CCC reveals that this code (that has inherent potentiality), with some modification, can be used as a code or guide for creating a comprehensive library database covering all types of entities associated with a library system.[40,41]

The careful study of CCC also reveals that it is a code, which much before the development of database technology, identifies and specifies different bibliographic entity types, namely: serial, monograph, and their subsets. Even the list and definitions of various kinds of conventional and non-conventional bibliographic items provided by Ranganathan in the code are still valid.[1] This code also lists and defines all possible attributes of the entity set—bibliographic document. That is, it specifies all the discrete attributes or data elements of various types of bibliographic items, and has framed the rules for description or rendering of their values and grouping and fixing their relative order or sequence within a printed bibliographic record or a catalogue card. These guidelines can be effectively used for entry and display in a computer.

The rules framed in part H and J of CCC are so general that these can be used to render the names of:

- **Persons:** Personal authors, collaborators, library members, library staff, and so forth.
- **Corporate Bodies:** All types of organisations. Typical examples are: Governments, associations, corporate business enterprises, social, cultural, educational and research institutions. A corporate body can be an author of a work, a publishing agency (or publisher), a sponsor under whose auspices a work is published, or a business enterprise—library vendor, book supplier, a library contractor, a co-operating library, etc.
- **Conferences**: International, national, and regional.
- **Geographic Entities:** Continents, countries, constituent states of a large country, cities, town, sees, and so on.

However, to revise and to convert this code into a code for library database design and development, we need to identify and describe the various types of entities, bibliographic or non-bibliographic in nature that constitute part of the library system or have affiliation with it, and their relevant, appropriate or context dependent attributes **. Some of the entities, other than bibliographic in nature, which are associated or have linked with the library system are: library members and employees (potential, existing, or retired); library suppliers; parent organisation(s): library departments or units; library systems and institutions; equipment; orders of items; invoices of items; inventories of library items, and so forth.[37,38,40-43] Besides incorporating the details of various types of entities that have affiliation with the library systems and specifying their attributes, one also needs to incorporate the rules for

**Entities possess numerous attributes, including attributes attained due to relationships with other entities and play different roles or have different identities in different conditions or contexts. Entities have divergent personifications or their personalities are transformed or look different in changed circumstances, positions, situations, or locale. For example, a person can be a teacher, a writer, a father, an employee, or member of an organisation. A given set of attributes (or their values) of a person, enables one to perceive, specify or identify him in a particular role, personification or context. Name of a person is universal identifier. However, if a person is an employee, then in this case, for his proper identification, one must at least specify or know the name of organisation in which he may be working and the name of his job.

rendering the values of the identified attributes of these entity types.

What has been discussed so far in this paper, can be summarised that to convert CCC into a code for library database design and development, one needs to:

1. Revise and restate the Canons of Cataloguing of Ranganathan to make them applicable not only to develop a library catalogue but a comprehensive database for library systems, which may include data about bibliographic as well as non-bibliographic entities.
2. Incorporate the Ranganathan's Postulates and Principles for Designing a Scheme for Library Classification and Concept Based Subject Indexing Systems, and the Logical Data Modelling Technique of Peter Pin-Sen Chen and further modified by other experts as methodologies for the design and development of library database systems.
3. Identify and describe the various types of entities (e.g., the various types of bibliographic items, library members, staff, equipment, records, etc.) associated with a library system, their attributes and relationships, or data elements, whose values comprise an entity record in a library database.
4. Designate and define the essential (or mandatory) and optional data elements of entity types that may comprise database record structure.
5. Assign unique codes to the essential and optional data elements, which serve as means of identifying data elements or values of fields in a machine-readable record. These unique identifiers or codes may be composed of field tags, indicators, and subfield codes.
6. Specify a set of rules for uniform or consistent description or rendering of the values of the attributes, i.e., the data elements or contents of fields and subfields comprising the database records.
7. Stipulate the guidelines or rules for identifying and using authentic sources for collecting data or pieces of information about an entity, i.e., the appropriate sources from where the values of attributes (data elements) of an entity need to be collected for description.
8. Delineate the domain of attributes (data elements), which define the valid values of an attribute and guide decision in case there are exceptions.
9. Outline:
   - Input formats in which the data of various entities or objects will form input to the system (i.e. the form in which the given pieces of information about an entity will be logically recorded), and
   - Output or display formats in which entity records or information will be displayed (i.e., the formats of required outputs or reports).
10. Identify and specify the access elements or points—the instruments of searching the database records, which also constitute component of a search mechanism used to retrieve information from the database.
11. Specify the methodology for developing post-coordinating and pre-coordinating information retrieval systems appropriate for computer environment
12. State the techniques of query formation and search strategies appropriate for computer environment.

Concluding Remarks

The development of the proposed code and making it an effective tool for designing and developing computerised library information systems is a Herculean task. It requires teamwork and support of a nation- level organisation. One such national agency that can take up this task is National Information System for Science and Technology (NISSAT), Department of Scientific and Industrial Research (DSIR), Government of India, New Delhi.

In the revised code, one must specify the various entity types associated with a library system, identify and define their attributes and relationships, and subsequently construct a standard list of fields and subfields (data elements) and provide them standard codes, tags or identifiers. Generally, electronic records of different entities comprised fields and subfields associated with an entity type, and

whose values are useful to locate, identify, and get information about an object, an action or a process etc. In addition, an electronic record contains certain fields, which contain data required for the management of database system or for other administrative purposes. For example, one such field that forms a component of electronic record is 'Record Number', representing unique number of an entity record of a database. The data fields and their constituent data elements comprising the records of different types of entities need to be identified and defined.

In the proposed revised code, besides covering the various aspects listed in the previous sections, one may also consider to redefine the Laws and Canons of Cataloguing as enunciated by Ranganathan in the light of new milieu. For example, 'Canon of Ascertainability', the principle prescribing the chief sources of information to collect data for creating bibliographic records of various bibliographic items need to be revised to include the appropriate sources of information for collecting data about other entity types associated with library systems.

With the arrival of information technology in library arena, the mindset of some library professionals about implementation of 'Law of Parsimony' needs to be changed. In the light of 'Law of Parsimony', Ranganathan had suggested to provide only brief but essential description about bibliographic items, while preparing catalogues of libraries other than catalogues of national libraries or preparing national bibliographies. This approach based on the principle of 'trade-off' (that is, to get something you need, you give up something you need less urgently) though good one, is no more valid for creating online catalogues or union catalogues of libraries, which are used not by library users but also by those who are involved in book trade and book publication.

With passage of time, due to the above-mentioned trade-off practice in the CCC, many professional librarians have developed a notion that CCC has no provision to provide detailed bibliographical description, as is the case in AACR-II. In CCC, no doubt, there is direction to provide only essential information and ignore the less essential data, like price of a book, while creating catalogues of small libraries. Nevertheless, it provides a comprehensive list of all the data elements that may comprise bibliographic records of various types of bibliographic items in a catalogue (see CCC Chapter FQ: Section of Entry). However, due to mindset, the most professional librarians ignore this direction and think that CCC rules limit us to provide detailed bibliographic description about a bibliographic item in the catalogue card, unlike AACR-II. This mindset needs to be changed by way of interpreting the 'Law of Parsimony' or the 'Principle of Trade-off' in the present context. The parallel example of application of 'trade-off' in information technology itself can be seen in the form of Y2K bug and old hardware and software designs. Actually, implementation of Y2K concept is not a bug at all; it is a result of application of Law of Parsimony or principle of 'trade-off' by software and hardware designers of 19th century. To get more space on disk and in memory, they gave up the precision of century indicators, which is proving as a flaw in software and hardware designs, in the context of year 2000 and future applications in almost every field.[44]

The present cataloguing rules were laid down for creating card catalogue of bibliographic items manually or semi-automatically. These rules need to be revised, updated keeping in view the ease and better data processing and retrieval facilities provided by the information technology in developing the information systems, and application of new approaches as demanded by the technology.

A brief outline of the coverage, which may form the part of the proposed code for library database design, shall be presented in another paper.

BIBLIOGRAPHIC REFERENCES

1. Ranganathan, S. R. Classified catalogue code: With additional rules for dictionary catalogue code. Ed 5. 1988.
2. Cutter, C. A. Rules for a dictionary catalogue. Ed. 4. Government Printing Press, Washington, D.C., 1904. P 12.

3. Potter, William Gray. Online catalogue in North America: An overview. *Program*, 1986, **20**(2), 120-30.

4. Baker, N. Discards. *The New Yorker*. April, 1994, 64-86.

5. Ranganathan, S. R. Subject heading and facet analysis. *Journal of Documentation*, 1964, **20**(3), 109-19.

6. Austin, Derek. PRECIS: A manual of concept analysis and subject indexing. Ed 2. The British Library, London, 1983.

7. Bhattacharyya, Ganesh A. General theory of subject indexing language. (Karnataka University, Ph.D. Thesis, 1980). (Unpublished)

8. Bhattacharyya, Ganesh A. POPSI: Its fundamentals and procedure based on a general theory of subject indexing language. *Library Science with Slant to Documentation*, 1979, **16**, 14-5.

9. Bhattacharyya, Ganesh A. POPSI: A source language for organising an associative classification. *Library Science with Slant to Documentation*, 1982, **19**, 249-52.

10. Robins, Matrus & Anderson, James D. Contextual indexing and facet taxonomic literary access system. *Drexel Library Quarterly*, 1983, **19**(3), 91-109.

11. Ingwersen, Peter & Wormell, Irene. Ranganathan in the perspective of advanced information retrieval. *Libri*, 1992, **42**(3), 184-211.

12. Ellis, David & Vasconcelos, Ana. Ranganathan and the net: Using facet analysis to search and organise the world wide web. *Aslib Proceedings*, 1999, **51**(1), 3-10.

13. Neelameghan, A. Application of Ranganathan's general theory of knowledge classification in designing specialised databases. *Libri*, 1992, **4**(3), 202-226.

14. Neelameghan, A. Concept categorisation and knowledge organisation in specialised databases: A case study. *International Classification*, 1991, **18**, 92-97.

15. Godert, Winfried. Faceted classification in online retrieval. *International Classification*, 1991, **18**(2), 98-109.

16. Iyer, Hemalata. Online searching: Use of classificatory structure. *In* Tools for knowledge representation and the human interface, edited by R Fuggman. Proceedings of Ist ISKO Conference (Darmstadt) (1990/91). INDEKS Verl, Frankfurt. pp. 159-67.

17. Iyer, Hemalata. Subject representation entropy. *International Classification, 1992*, **19**(2), 15-8

18. Svenonius, Elaine F. The use of classification in information retrieval. *Library resources and Technical services*, 1983, **27**(1), 76-83.

19. Stiles, W.G. Ranganathan, cognition and expert system. *Canadian Journal of Information Science*, 1895, **10**, 16-24.

20. Vickery, Alina & Brooks, H.M. Plexus—the expert system for referral. *Information Processing and Management*, 1987, **23**, 99-117.

21. Borgman, Christine L. Why are online catalogs still hard to use. *JASIS. 1996*, **47**(7), 493-503.

22. Bates, M. J. Subject access in online catalogue: A design model. *JASIS*, 1987, **37**, 407-24

23. Crawford, J. C., *et al.* A survey of subject access in academic libraries catalogues in Great Britain. *Journal of Librarianship and Information Science*, 1993, **25**, 85-93.

24. Hildreth, C. R. Beyond boolean: Designing the next generation of online catalogs. *Library Trends*, 1987, **36**, 647-67.

25. Lancaster, F. W., *et al.* Identifying barriers to effective subject access in library catalogs. *Library Resources & Technical Services*, 1991, **36**, 377-91.

26. Vossen, G. Data models, database language and database management systems. Addison-Wisley, Wokingham, 1991

27. Oxborrow, E., Databases and database management systems: Concepts and

issues. Ed. 2. Chartwell-Bratt, Chichester, 1981.

28 Unesco, Division of Software Development and Applications, Office of Information Programmes and Services. Mini-Micro CDS/ISIS 2.3: Reference manual & CDS/ISIS pascal manual. UNESCO, Paris, 1989.

29. Chowdhury, C. G. Introduction to modern information retrieval system. Library Association, London. 1999.

30. Dates, C. J. An introduction to database systems. Ed. 3. Addison-Wesley, Reading, MA. 1981.

31. Chen, Peter Pin-Sen. The Entity-Relationship model—towards a unified view of data. *ACM Transaction on database Systems,* 1976, **1**(1), 9-36.

32. Chen, Peter Pin-sen, Ed. Entity-Relationship approach to information modelling and analysis. North-Holland, Amsterdam, 1983.

33. Anglo-American Cataloguing Rules, edited by Michael Gorman & Paul W Winkler. Ed. 2. Revised. American Library Association, Chicago, 1988.

34. CCF/B: The Common Communication Format for bibliographic information. Paris, UNESCO, 1992. (PGI-92/WS/9).

35. CCF/F: The Common Communication Format for factual information. Edited by Alan Hopkinson & Peter Simmons. Paris, UNESCO, 1992. (PGI-92/WS/9).

36. Hopkinson, Alan, Comp. Implementation notes for users of the Common Communication Format for factual information (CCF/F). UNESCO, Paris, 1993. 47p.

37. Kashyap, M. M. Integrated database design for a library system: Employing library techniques developed by Ranganathan and CDS/ISIS DBMS. *Journal of Library and Information Science (India),* 1993, **18,** 82-141.

38. Kashyap, Madan Mohan. Online library catalogues and standards for common application in a network environment: Paper presented at XVI IASLIC Seminar, Bombay, 1994. *In* IASLIC Seminar (Bombay) (1995) (16). Networking of libraries: Problem and prospects. 1995. P. 1-25.

39. CCF: The common communication format. Edited by Peter Simmons & Alan Hopkinson. Ed. 2. (PGI-88/WS/2). UNESCO, Paris, 1988.

40. Kashyap, M. M. Database system: Design and development. Sterling Publishers, New Delhi, 1993. x+382p.

41. Kashyap, M. M. Database technology: Basic concepts and terminology. *Library Herald,* 1991, **30**(1), 1-19.

42. Kashyap, M. M. Common communication format for online ordering, a proposal for book trade and libraries: Paper presented at 15th Annual Convention and Conference at the Society for Information, held at Bangalore during 18-20 January 1996. *In* Digital libraries: Dynamic storehouse of digitized information, edited by N. M. Malwad, *et al.* New Age International, New Delhi, 1996. pp. 75-89. 26-33.

43. Kashyap, Madan Mohan. Computer-based library information system designing techniques. Sterling Publishers, New Delhi, 1999. pp. 259-60.

44. Ullman, Ellen. The myth of order. *Span,* 1999, **XL**(6), 46-50.

Contributor: **Prof Madan Mohan Kashyap** has retired from Dept. of Library & Information Science, University of Delhi. His residential address is 201 Kadambari Niwas, Kadambari marg, Sector - 9, Rohini, New Delhi - 110 085. Tele (R): 786 2699; E-mail: mmkashyap@bol.net.in; kashyap1@indiatimes.com