Likeness Between Ranganathan's Postulations Based Approach to Knowledge Classification and Entity Relationship Data Modelling Approach

Madan Mohan Kashyap

201, Kadambari Niwas, Kadambari Marg, Sector-9, Rohini, Delhi-110085
Email: mmkashyap@bol.net.in

Madan Mohan Kashyap (b.1928) is a distinguished retired teacher from the Department of Library and Information Science, University of Delhi. He was actively involved in teaching and research in the areas of knowledge classification, systems analysis and design, and computer application. Certain Indian universities have invited him many times as a Visiting Professor and Visiting Fellow. He was awarded the Fulbright Fellowship in 1971 and Ranganathan Medal by IASLIC in 1994 for best article presented at the XVI National Seminar of IASLIC.

Kasahyap, Madan Mohan. (2003). **Likeness Between Ranganathan's Postulations Approach to Knowledge Classification and Entity Relationship Data Modelling Approach.** *Knowledge Organization, 30(1).* 1-19. 31 refs.

ABSTRACT: This paper describes the *Postulations Based Approach to Facet Classification* as articulated by **S. R. Ranganathan** for knowledge classification and for the design of a facet scheme of library classification, and the *Entity-Relationship Data Modelling and Analysis Approach* set by **Peter Pin-Sen Chen**; both further modified by other experts. Efforts have been made to show the parallelism between the two approaches. It points out that, both the theoretical approaches are concerned with the organisation of knowledge or information, and apply almost similar theoretical principles, concepts, and techniques for the design and development of a framework for the organisation of knowledge, information, or data, in their respective domains. It states that both the approaches are complementary and supplementary to each other. The paper also argues that Ranganathan's postulations based approach or analytico-synthetic approach to knowledge classification can be applied for developing efficient data retrieval systems in addition to the data analysis and modelling domain.

Introduction

A library is a unique type of human organisation. It is primarily a knowledge-based system or information organisation, which acquires, organises, safeguards, preserves, and gives access to carriers of information and knowledge in all forms (i.e., bibliographic documents) to the library members free of charge.

The organisation of knowledge sources and bibliographic information about them has been the concern of librarians since the inception of libraries. They provide information about the knowledge or the information stored in various types of knowledge/information sources and systematise the bibliographic information about them.

Librarians and library science scholars were possibly the first information specialists who developed theoretical approaches, practical tools and techniques for organising, storing, and retrieving bibliographic documents and the bibliographic data about them, more essentially information about their subject contents.

Among them, Ranganathan stands apart, because he provided a theoretical base in his classical work *Prolegomena to Library Classification* in the form of postulate, principle and canons for classification of the universe of subjects/concepts/entities, for the design of a scheme for library classification, and for classifying documents according to subject order by applying his Analytico-synthetic approach.

This paper has been written to show the parallelism between the *Postulations Approach to Facet Classification* as articulated by **S. R. Ranganathan** (1937; 1957a; 1957b; 1958; 1967) and the *Entity-Relationship Data Modelling and Analysis Approach* set by **Peter Pin-Sen Chen** (1976, 1983); further modified by other experts. Both theoretical approaches are concerned with the organisation of knowledge, information or data, and apply almost similar theoretical principles, concepts, and techniques for the design and development of a framework for the organisation of knowledge, information, or data, in their respective domains.

Ranganathan's *postulations-based approach to subject classification*, also known as *analytico-synthetic approach to knowledge classification*, is based on a set of postulates, principles, canons, and the principle of facet analysis and synthesis, as enunciated in his classic work ***Prolegomena to Library Classification*** (**Ranganathan,** 1937; 1957b; 1967) and further elaborated, extended and interpreted by him in his latter writings. This approach can be applied for

- Designing a faceted scheme for library classification
- Classing and organising bibliographic documents possessed by libraries in a helpful classified order or in APUPA PATTERN, as Ranganathan named the order.
- Developing verbal classification systems like POPSI.
- Formulating specific subject headings and developing thesauri, and so forth
- Organising records of the bibliographic items as library catalogues, which enable the library users to search and identify the bibliographic items of their interest through various access points (Kashyap, 1983).

According to Ranganathan, "library classification is the classification of subjects or knowledge classification" (Ranganathan. 1967, pp.94). According to him, a *subject* is an organised or systematised body of ideas, facts, data, information or knowledge concerning *some thing/things* or an *exposition of theoretical construct(s)* embodied in a bibliographic item. The content of a bibliographic item is the product of intellectual, transintellectual, creative activity or the imagination of man recorded in an intelligible, coherent and communicable form. The focus of study and description in a work or bibliographic document can be about a ***conceptual*** (*abstract*) or ***concrete*** (*physical*) ***entity/entities*** and ***their properties***. The postulate, principle and canons for classification are concerned with developing a framework for a scheme for subject classification and subject indexing system for organisation of documents, information about documents and their contents (i.e., the description of abstract or physical entity/entities or their attributes in the bibliographic documents).

Data modelling is a technique for establishing, organising and documentation of information system data. Data modelling first gained recognition in **Dr. Peter Chen's** article *Entity-Relationship Model: Towards A Unified View of Data* (1976). Since then, data modelling or database modelling has become the standard approach in information systems development or in the design of database systems. The Entity-Relationship (E-R) modelling technique is used to specify a conceptual or universal data model, sometimes called conceptual or logical design of an enterprise database. The conceptual or logical model is converted into a physical model (also called *implementation model* or *technical model*).

The conceptual or logical models depict the data requirement of end users of information systems. They are implementation independent; whereas physical models are implementation dependent and reflect technology requirements. Different database management systems have different database structures for representing data. For example, in relational databases the fundamental structures for representing data are what we call *relational tables*. Thus, a physical model or implementation model of a relational database system will reflect the preferences and limitations of the chosen database management system, such as Microsoft Access or Oracle.

The entity-relationship approach (E-R approach) is an analytical technique, which is used to identify and describe the entities associated with an enterprise/ organisation/problem area or about which an enterprise, organisation or person wishes to store information. This database modelling approach is based on the following fundamental concepts for presenting the conceptual model of an enterprise database:

- **Entities**, which represent the object set (mini world) with which one is concerned. The things about which one wants to collect and keep *facts, data* or *information*.
- **Entity type or class**, which represent a set of distinguishable real-world objects or things (mini world) with common properties and about which an organisation or a person wishes to store *facts, data* or *information*

- **Attributes**, which represent the properties of these objects. An attribute is the data element/item that describes the property of an entity or a relationship.
- **Relationships**, which represent the associations or connections among entities. Examples of typical relationships are "**works-in**", "**member-of**", "**places-orders**" or any verb that connects entities. The entities acquire certain descriptive elements or properties due to their association, connections or relationships with other entities. Relationships between entities are either optional or mandatory, may be one-to-one, one-to-many, or many-to-many.

It may be pointed out here that beside the ***entity-relationship analysis and modelling approach,*** another data modelling approach, named the ***object-oriented analysis and modelling approach,*** is also popular these days. There is some similarity between the concepts or terms used in both of these data modelling approaches. *Object modelling* is a technique for identifying objects within the systems environment and the relationships between them. According to the object modelling technique:

- "An ***object*** is something that is or is capable of being seen, touched, or otherwise sensed, and about which users store data and behavior.
- ***Attributes*** are the data that represent characteristics of interest about an object.
- ***Behavior*** refers to those things that the object can do and the correspondence functions that can act on the object's data (or attributes). In object-oriented circles, an object's behaviour is commonly referred to as a method, operation or service." (**Whitten** et al., 2001, pp. 646-672)

In this paper, the description of the **Entity-Relationship Approach** *of data modelling and analysis,* and the **Postulations-based Approach of Ranganathan** *for designing a scheme for subject or knowledge classification* is given with a view to pointing out or showing the similarity between their theoretical assumptions and practical applications. In fact, these theoretical approaches and their applications are complimentary and supplementary to each other, as both are concerned with the organisation of information or knowledge.

1 Entity-Relationship (ER) Approach

1.1 Introduction

The Entity-Relationship (ER) approach, as a data analysis and modelling approach, was first introduced by **Peter Pin-Sen Chen** in 1976 . Since that time, many experts in the field of information technology have suggested several modifications to the ER model (*see review* by **Teorey** et al.,1986; *also* **Chen**, 1983; **Codd**, 1979, 1982, 1990; **Mair**, 1983). It has presently evolved into one of the most important data analysis and design techniques. Designing a database structure is fundamentally a task in data modelling and a data model is architecture for data. It describes the general structure of how data is organised. The purpose of data analysis is to represent data, as it is perceived by its users, and it is also referred to as the *semantic modelling technique* (**Date**, 1995)

This data or database modelling approach is an analytical technique, which is based on three fundamental concepts : i. **Entities***, the objects or things about which one wants to keep data or fact in a database*, ii**. Attributes,** *which represent the descriptive properties of the entities, and* iii. **Relationships**, *which represent the association between entities and the attributes or properties entities acquire or possess due to their relationships with other entities*. These concepts are the foundation of the ER model, with the help of which we can outline a conceptual model (also called the logical view) of a database. The conceptual level model or design of a database relates to the representation of that part of the real world that the database is about.

The ER approach pre-empts identifying and describing of the real world objects or things associated with an enterprise or problem area; that is, the universe of entities, whose database is to be built. This approach is based on the assumption that examining only processes, transactions, outputs, or data flows of a system, or examining all factors gives partial information about the environment of the system. Without a proper understanding of entities, objects or things themselves and their environment, one cannot make reasoned decisions as to what data is needed about those objects or things. The ER approach helps us to get a true or complete picture of the real world or universe whose database is to be built and involves the identification and definition of entities of the concerned real world, entity grouping and description, keeping in view the enterprise or problem area context.

i. ***Entity identification and definition:*** This step involves recognising various entities whose database is to be built, determining why they are important to the organisation, researcher or users of the database, and naming them.
ii. ***Entity grouping and description.*** This step involves grouping of entities of interest (i.e., the given population) into broad and narrowly defined classes, and identifying and describing the relevant attributes of the entity classes and relationships among them, for creating the records of entities in the given context.

Each group or class, broad or narrow, constitutes a number of entities that have a set of attributes in common. Thus, the entity description consists of identifying which attributes of an entity-type (or entity class) are relevant for creating the records of entities in a given context.

Broadly, we can group entities around us into persons (including human organisations), objects, places, events, concepts, and so forth. However, these groupings are too general to handle in a meaningful manner. Therefore, we must form subgroups from each of these classes into subtypes, that is, into different kinds of persons, objects, places, events, concepts. The categorising of entities may be based on criteria, such as, what the entities are, what purposes they serve, what they do, what they look like, how they are used, and so forth.

The entities sharing common attributes are grouped together and form a class. In other words, an entity class could be considered as an aggregation of attributes. Similarly, an entity could be considered as an aggregation of its properties. Selecting appropriate differentiating attribute(s) or characteristic(s) shows the further division of entity classes into subgroups.

One has to specify the recognised entities and entity groups and subgroups at the exact level of precision that ensures that it not so general as to be meaningless, and yet not too specific. For example, to consider entity class PERSON to be of primary importance for creating a database of a library system would be too general or broad, as it can be fragmented into too many subsets. Whereas, to consider entity classes LIBRARY USER, LIBRARY EMPLOYEE and so on, as the entities of interest is a more precise and relevant selection. On the other hand, to consider TEMPORARY MEMBER and PERMANENT MEMBER as entities of interest would be too specific, particularly if member of each group has equal status or privileges for the use of the library materials, and so forth . In this case, '**temporary**' and '**permanent**' are two different standings or states of a library member, and is to be treated as an *attribute* of a library member. A member can alternate between these two states or positions and still is a library member having the same status.

Sometimes it is necessary to divide an entity class into subgroups. For example the entity classes PROFESSIONAL STAFF, SEMI-PROFESSIONAL, NON-PROFESSIONAL STAFF might all be declared as subgroups of the entity class LIBRARY EMPLOYEE. Likewise LIBRARIAN, DEPUTY LIBRARIAN, and ASSISTANT LIBRARIAN would be declared as subgroups of the entity class PROFESSIONAL STAFF. While forming groups of entities, it is necessary that we attempt to use the most general, yet most meaningful, grouping possible. The important consequence of forming a generalised hierarchy of entities or entity groups or classes is that entities lower down the hierarchy inherit the attributes and relationships of entities higher up in the hierarchy. Hence, an Assistant Librarian would inherit attributes of a Deputy Librarian in general, and of library employees in general. Likewise, an Assistant Librarian would inherit the relationship of a library employee to a Library Unit/Department. However, each entity class (super-class or subclass) possesses a unique property or characteristic that differentiates it from other groups.

Grouping entities and subdividing them into subgroups at many levels involves choosing appropriate characteristics (differentiating attributes). Entities belonging to a class (i.e., members of a group) share the same attributes. Further, each entity class has its own set of attributes, which may be numerous. Out of these, only those are picked which are relevant for the purpose of creating a database or are selected depending upon the requirements of the system or enterprise whose database is to be created.

The broad group of an entity set or the first level group is called *family or super-class*. Each super-class or family is constituted of subgroups, and the groups of lower levels can be further subdivided in subgroups (the terms, supertype/super-class and subtype are also used in place of group and subgroup, respectively). Each subgroup within the family or superclass has a characteristic (a *characteristic is some property or attribute that distinguishes one thing from an other*) that differentiates it from other groups of the family, and shares all the attributes of the family as completely as possible.

The differentiating characteristics or attributes from which we form subclasses of a super-class can be very effectively used while designing an effective database model. For example, while creating a database of the entity type EMPLOYEE, many data models treat the class or category to which an employee belongs as an *attribute,* namely, **assigned position/ category/cadre** of entity type EMPLOYEE. This attribute is also the characteristic that can be used to form subclasses of the entity type EMPLOYEE. It can also be seen that the **position assigned** or **held by** an employee of a library system represents a job slot that can be occupied by many persons over time or can be vacant at a given point in time. For example, if there are two positions vacant or open for a deputy librarian's post in a library that would indicate that library could appoint two people. Each of these persons may perform the same type of duties, but they would be filling two separate slots of a specific type of position. It can be seen that **assigned position** also constitutes the attribute of the relationship between entity types EMPLOYEE and POSITION. It may be observed that position forms an attribute of entity EMPLOYEE (or forms value of the attribute position assigned or held by an employee) in one context, and in other contextPOSITION constitutes or is treated as an important entity, which has its own information as well as a many-to-many relationship with the EMPLOYEE entity.

An entity, entity family, entity-type or group is defined within the context of the organisation or the library system whose database is supposed to be built. Thus, an entity belonging to a broad group in one context may belong to a subclass/subgroup/subtype of entities in another context.

Entity groups, at the family level and below, are primarily formed based on the role, which each member of each group plays in the organisation. Further, the entity grouping may be created in the form of a *mutually exclusive* (disjointed) type of subclasses or *mutually inclusive* (partially disjointed or overlapping) type of subclasses.

In some instances however, an entity can play a multiple role. For instance, a member of a university library system can be a teacher, a student, and a library staff member. In the real world, a teacher, a student, and a library staff member belong to mutually exclusive subclasses. In case the library rules do not provide equal loan facilities or other privileges to the members of each group, then in this case, each group must be treated as a mutually exclusive group. However, if each member of each group is given equal status or privileges according to the rules of a library system, then though the roles of each entity-type or group are distinct in the context of university setting , all of these entity groups may be merged into one family type or supertype - Library Member. *Their roles (as teacher, as student, or staff member) may however be treated as their attributes*. The subgroups of entity class Library Member, namely, Teacher Member, Student Member, and Employee Member are to be treated as overlapping or mutually inclusive subclasses.

While classifying entities into families (supertypes - level 1), groups (subtypes - level 2) and subgroups (subtypes of subtypes of level 2), the following factors need to be kept in mind:

- If the roles are mutually exclusive, that is, if the entities can play one role in the organisation and not another, define separate entity families, groups, or subgroups.
- If the roles are distinct, but *not* mutually exclusive, merge the entities into a single entity family or a broader group. (e.g., a library employee who is assistant librarian, can be promoted to a post of deputy librarian, and at next stage to a post of librarian, as such we must merge the entity subclasses LIBRARIAN, DEPUTY LIBRARIAN, and ASSISTANT LIBRARIAN into a single entity EMPLOYEE or need not consider these subclasses for designing a data model for library system. An employee is assistant librarian, deputy librarian or librarian by virtue of the position he holds, so EMPLOYEE and POSITION best be treated as entity types having a relationship for designing a data model)

Grouping of entities and subdividing them into subgroups of many levels involves choosing appropriate characteristics (differentiating attributes). Entities belonging to a class (i.e., members of a group) share the same attributes. Further, each entity class has its own set of attributes, which may be numerous. Out of these, only those that are relevant need to be selected, depending upon the context. The process of classification or grouping entities into supertypes and subtypes is not as simple, as it may appear. It may be pointed out that the Canons of Classification enunciated by Ranganathan can be of much help. The following specific Canons, as enunciated by Ranganathan in his work *Prolegomena to Library Classification* to provide the framework for classification of sub-

ject/knowledge can be used for classification of entities or grouping entities into subtypes and super types, in the domain of data modelling or database design and development:

- The **Canons for Characteristics,** which provide directions for the choice of appropriate attributes or characteristics in terms of securing differentiation among classes, must be *relevan*t for the purpose of classification, should be *definite* and *ascertainable,* and must be *permanent* (i.e., should continue to be unchanged so long as there is no change in the purpose of classification).
- The **Canons of Succession of Characteristics**, which are guidelines for choosing the right characteristics for division, one after another. The guiding principles are: (i) No two characteristics should be concomitant or concurrent (that is, they should not give rise to the same types of subclasses), (ii) they should be relevant to the purpose of classification, and (iii) they should be consistently adhered to in all cases.
- The **Canon of Exclusiveness**, which implies that a set or an array of subclasses formed due to grouping, should be mutually exclusive.
- The **Canon of Exhaustiveness**, which implies that a set or an array of subclasses formed due to grouping, should be totally exhaustive of their respective common immediate universe.

iii. ***Enterprise context*** involves identifying and defining the relationships that exist between the identified and defined entities, their relative importance to the enterprise as a whole, and each specific part/sub-unit/subsystem of the enterprise. The enterprise or problem context also involves identifying the role or function of each of the entities or entity-types within the enterprise or organisation. For example, the entity type or class BIBLIOGRAPHIC DOCUMENT and its subgroups are viewed in different perspectives in the context of a library system compared to the way a business enterprise (publisher or bookseller) may view them.

The specific description of bibliographic entities and the relationships among them and with other entities within a library system are relevant only within the context of the library system. For example, for creating a database of a library system as whole, we may consider the entity types WebPages, SERIES/PSEUDO SERIES, MONOGRAPH, SERIAL and INDIVIDUAL CONTRIBUTION in a monograph, serial or WebPages (including INDIVIDUAL LINK of WebPages), as mutually exclusive (disjointed) entity subtype of the superentity type BIBLIOGRAPHIC DOCUMENT/ITEM. However, it may be observed that though a SERIES is a serial publication that is published by a publisher, ordered and purchased by a library, it is never treated as a bibliographic item *per se* by a LIBRARY MEMBER and does not form part of an exclusive bibliographic database of a library system. From a reader's point of view a series is simply an attribute of a bibliographic item monograph or a serial and is not a physical entity as any other book, monograph or article published in a serial - a source of knowledge. A library also treats a series as an attribute of a monograph or a serial while creating a bibliographic record for the library catalogue, but treats it as an entity while placing the order for a SERIES with a publisher.

1.2 Entity Relationship Diagrams (ERD)

Entity models are usually mapped out as **entity relationship diagrams** (**ERDs**). The product of entity relationship diagramming is a graphical model of entities (objects of interest), the relationships between entities, and attributes associated with entities associated with an enterprise, or a domain of investigation/ discourse. An entity relationship diagram illustrates the data to be stored in an information system, which constitutes a database of the information system. The data to be stored in a database pertain to an entity or a set of entities in the **real world,** sometimes called the **mini-world**. The data about an entity are essentially the ***values*** of its ***attributes*** (or ***attribute set***) including the attributes of its ***relationships*** with other entities, that is, the attributes the entity acquires or attains due to its relationships with other entities. ***Relationships*** represent connections between entities and meaningful dependencies between them.

An entity or an entity class (i.e., *the object set about which one wants to keep data or fact in a database*) is represented in a diagram by a rectangular box in which is written the meaningful name for entity. A relationship between entities is represented, by drawing a line (sometimes labelled) between relevant entity boxes, in the diagram. An attribute is represented by an oval attached by a line to the appropriate entity. The entity identifier (attribute) is underlined.

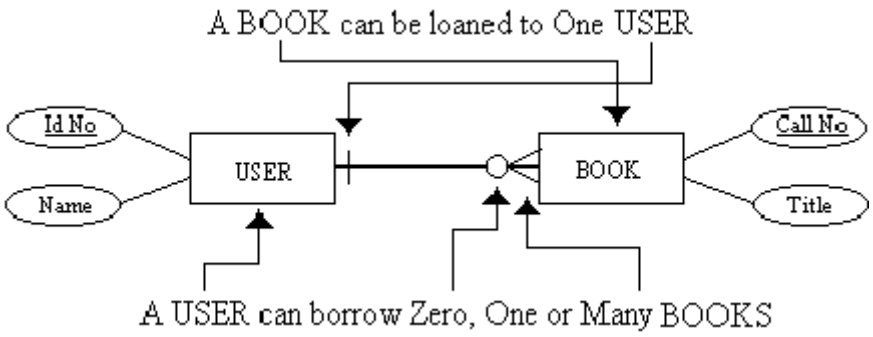

Figure 1: *An illustration of an Entity Relationship Diagram*

Entity-relationship diagrams show *data at rest*; that is, data being stored or to be stored in the proposed system, independently of how that data will eventually be *processed or used*. E-R diagrams complement data flow diagrams or flow charts, which show data in motion - as data flows.

Systems analysts use this tool to develop implementation-independent data models, which allow them to communicate with end users in non-technical language. The implementation-independent data models are usually called ***conceptual, logical*** or ***essential*** *models.* The conceptual model (also called conceptual or logical view of a database) is a concise description of the data requirement of the end user or users of computer based information systems. It describes the various entity-types, their attributes and the relationship between them.

The ***conceptual schema*** might be regarded as an overall logical database description or a global model representation of a database as conceived by the library database administrator or systems analyst and is entirely independent of the physical storage organisation of a database. A conceptual schema can be defined using Entity-relationship (E-R) model concepts and can be displayed by means of the graphical notation known as Entity Relationship Diagrams (Kashyap, 1999, Chap. 7).

An Entity-Relationship diagram (E-R diagram) is conventionally referred to as a conceptual model of a database system. From this conceptual model, we derive an implementation-dependent logical model (also called a physical model, e.g., expressed in the form of a series of tables structured in a third normal form for a relational database system), which may be mapped on to or converted into the database structure of a Database Management System (DBMS). The final step in database development is to produce a physical model. That is, a series of record structures expressed in the syntax of some programming language or DBMS (**Howe**, 1986).

A database conceptual schema is specified during the database design of an information system and is not expected to change frequently.

1.3 Entity Relationship Model Concepts

Entity

In database terms, an ***entity*** is defined as anything (physical or abstract) in the real world about which we store facts, information or data in a database record, and which is capable of independent existence and can be distinctly identified. An entity may be a tangible object with physical existence, such as a particular person, an employee, a library member, a book, a serial, a chair, or a table. It may be a non-tangible object with conceptual existence or an abstract concept, such as an event, a transaction, a job, a procedure, a subject of study, or a university course. In other words, every thing that exits in reality or is perceived as being in existence is an entity.

All entities are distinct from one another in the sense that each possesses a particular set of **properties, attributes or values** that distinguish it from the others. Thus, a bibliographic item has properties such as *title, size, price, cost*, and so forth; a library user possesses attributes such as *name, age, educational qualification, place of residence*, and so forth ; and a library employee has attributes such as *name, address, basic salary*, and so forth. We contrast or differentiate between entities, for some purpose, based on some unique name, identifier, label, symbol or property, even if they belong to the same class of objects. For example, the book entitled "Computer Based Library Information Systems Designing Techniques", belonging to a library and assigned *accession number 245600*

is an entity. Similarly, "Prof. S. R. Ranganathan" *born in 1892* is a particular person (entity) in the universe.

The entities exist in large homogeneous groups where all members are capable of being described in the same manner having identical attributes, or they may be fragmented or considered part of many subtypes, each with a description which is either slightly different, or in some cases radically different, from the other members of the same group. We contrast or differentiate between entities, for some purpose, based on some ***characteristic*** or ***characteristics***, for example:

- **Library users** and **library employees** are both *people*, but a **library user** is a *person* who makes use of a library collection while a **library employee** is a *person* who organises and maintains the library collection for the use of the library members.
- **Monographs** and **Serials** are both *bibliographic items*, but a **monograph** is a *non-serial bibliographic item* (i.e., a bibliographic item complete in one part or finite number of parts, or intended to be completed, in a finite number of parts), but a **serial** is a *periodical publication* or *bibliographic item* published / issued in successive parts bearing numeric and / or chronological designations and intended to be continued indefinitely. Each part is generally made up of distinct and independent contributions, not forming a continuous exposition.

The entities LIBRARY USER, MONOGRAPH, and SERIALwhich we referred to above as entities, can be called entity classes (or entity types) because each denotes a set of objects (individual entities), each of which exhibits the properties/attributes described for the class. Thus, the entity-type LIBRARY USER is made up of individual entities, each of which has attributes, **name, age, sex, educational qualification, occupation, place of residence,** and so forth . Every individual entity belonging to this class possesses all the attributes of the class but differs from the others in the values of their attributes.

In a data-modelling context, the specific definitions of the identified entities have meaning within the context of the organisation or the problem area. The identified entities are grouped into types and subtypes, or defined into broader groups and narrowly defined groups. Each group, broad or narrow, large or small, consists of a number of entities that have a set of attributes in common. Each entity group share the same set of attributes, and share characteristics of the higher group, class or family.

Entities for data modelling purpose must be grouped into reasonably broad classes. For example, we can group entities into the following very broad classes: *persons*, *human organisations, things, places, concepts, and events*. These groups are, however, too general to work with and for purpose of optimal efficiency and usefulness of the database system, each of these six classes must be represented by two or more restrictive classes (subtypes or subclasses) through **generalisation** and **aggregation** processes.

Generalisation is an approach that seeks to discover commonalties among entities or things. It is a process that allows us to group entities having common features or attributes, into a broad class or supertype entity. For instance, a library employs PROFESSIONAL STAFF, SEMI-PROFESSIONAL STAFF and NON-PROFESSIONAL STAFF. There are several attributes that are common to all the entities; for example, **name, gender, marital status, age**, **qualification** and **religion.** We might group them all into an entity supertype called LIBRARY EMPLOYEE. Likewise, entities LIBRARIAN, DEPUTY LIBRARIAN, and ASSISTANT LIBRARIAN are declared as subtypes of supertype entity PROFESSIONAL EMPLOYEE. **Figure 2** shows this generalisation as a hierarchy.

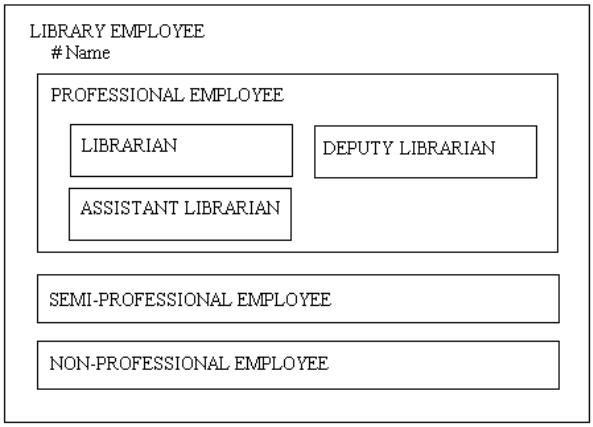

Figure 2: *Subtype and Supertype Entity Classes*

The important consequence of forming generalised hierarchies is that the entities lower down the hierarchy or chain inherit attributes and relationships of entities higher up in the hierarchy or chain. Hence, subtype entities LIBRARIAN, DEPUTY LIBRARIAN, and ASSISTANT LIBRARIAN would inherit the properties of the supertpye entity PROFESSIONAL STAFF in general and indeed of the next higher-class entity LIBRARY EMPLOYEE in gen-

eral. Like wise, entity subtypes PROFESSIONAL LIBRARY EMPLOYEE, SEMI-PROFESSIONAL LIBRARY EMPLOYEE, and NON-PROFESSIONAL LIBRARY EMPLOYEE would inherit the attributes and relationships of LIBRARY EMPLOYEE in general and indeed of the next higher class LIBRARY.

The aggregation process or approach allows declaring certain entities as subtypes of other entities, based on whole and part relationships. The whole (i.e., higher level entity) is considered as an aggregate of parts (i.e., lower level entities). For example, *a country* (whole - super-class) is an aggregate of *states* (parts - subclass); *a state* (whole - subclass) is aggregate of *districts* (parts - subclass a higher subclass of a super class), and so forth.

We know that a *serial* constitutes a set of *volumes*, a *volume* of a serial constitutes a set of *issues*, and an issue constitutes individual *articles* or *contributions*. In other words, a *serial* (whole - super-class) is an aggregate of *volumes* (parts - subclass), a *volume* (subclass of a super-class) is an aggregate of *issues (*subclass of a higher level subclass of a super-class*)*, and an *issue* of a serial (subclass of a higher level subclass of a super-class) is an aggregate of individual *articles* or *contributions* (lower most subclass of a subclass of a higher level subclass of a super-class). The process of decomposing an entity class is opposite of aggregation. That is, it is a process of decomposing entity classes into constituent parts.

We can distinguish between aggregation and generalisation in the following terms. If a hierarchical relationship between two classes is defined in terms of generalisation then both subclasses and super-classes refers to the same thing. That is, if we state A is a Mutivolume Monograph then we also mean A is also a Monograph or Bibliographic Item. On the other hand, if a hierarchical relationship between two classes is defined in terms of aggregation then both subclasses and super-classes refer to different things. For example, a Serial Volume, Issue or a Contribution in a Serial is a distinct thing from a Serial.

The subtypes within a supertype entity class should represent a complete set (meaning one must include all the subtypes possible) and at the same time be mutually exclusive of each other. An exception also occurs when a supertype entity is decomposed into non-mutually exclusive subtypes. For example, EMPLOYEE entity class can comprise of non-mutually exclusive subtypes SUPERVISOR, CONSULTANT, MANAGER, ENGINEER. In this case EMPLOYEE is subtyped in different ways and more than one set of subtype may apply to the same supertype. An employee may be a manager as well as an engineer or a consultant.

It may be pointed out that classification of the various types of entities into their hierarchical subgroups (i.e., into subtypes and supertypes classes) helps us to identify the data records of the entities belonging to an entity class or category, particularly in an integrated database. Because the supertype or subtype classes to which an entity belongs constitute an attribute of an entity, the value of this attribute can help us to identify and retrieve all records of the entities belonging to that class or category. For example, by making provision for indexing the value of the attribute **type** or **class** of the entity bibliographic document, we can search and retrieve records of all the bibliographic items belonging to a subtype, namely MONOGRAPH, SERIAL, or ANALYTIC (i.e., INDEPENDENT CONTRIBUTION in a serial or composite monograph) of the supertype entity BIBLIOGRAPHIC DOCUMENT, in a bibliographic database. It also may be pointed out that Ranganathan's *Canons for work at the idea level* (*cf*., **Ranganathan.** *Prolegomena.* 1977, Part E) can prove very helpful to data modellers in defining subclasses of entity classes or constructing the generalised or aggregated hierarchical entity classes.

Entity Class

An **entity class** *constitutes a set of all entities that have the same attributes* and *relationships with other entities* Essentially, when we use the terms such as Monograph, Serial, Author, Library, Library User, Library Vendor, Order, Invoice, and so forth, to represent an entity, what we actually mean is that there are these entity classes - Monograph, Serial, Author, Library, Library User, Library Vendor, Order, Invoice, and so forth, and that there may be many instances or ***entity-occurrences*** of these types of entities.

A database usually contains records of groups of entities that are similar. For example, a library possessing hundreds of different kinds of monographs stores almost similar types of information or data in bibliographic records concerning each of the monographs. These monographs share almost the same attributes and relationships with other entity-types, but each particular monograph (entity) has its own value(s) for each of its attributes and relationships with other entities.

The concept **entity class** (also named, **object class, object type**, **entity type, object set**, or **entity set**) is

very important for developing logical data models, and to specify real-world items in an application of all things (concrete or abstract) about which we store data in an information system.

The designer of a database has to decide what entities are to be represented in a database and what information ideally should be represented about each entity in a database The identification and description of the entity classes whose data or information is to collected in a database means definition of their relevant attributes and relationships with other entities. Every individual entity belonging to a class possesses all the attributes of the class, but differs from one another in the ***values*** associated with their attributes.

While creating a database, we may consider an entity disjointedly or create a super class from two or more subtypes. For example, we find a library collection constitutes the bibliographic items comprising the subtype entities; namely serials, monographs, and individual contributions in a serial or monograph. All these entity subtypes possess all attributes of the entity supertype BIBLIOGRAPHIC ITEM, in addition to their own distinct attributes. In this case, we store essentially the same information about the entities MONOGRAPH, SERIAL and CONTRIBUTION; we need not consider them entity subtypes and can merely treat these as ***values*** of the **type of bibliographic document/ item** of the entity class BIBLIOGRAPHIC ITEM/ DOCUMENT.

The naming conventions of an entity include using a singular noun in capital letters (e.g., entity type/ class MONOGRAPH). In some books, the name of entity type/class is represented in capitalised plural form (e.g., MONOGRAPHS).

The description of an entity class/type normally includes the following:

a The set of ***attribute-types*** and their ***data elements*** that describe various properties of each entity class, and whose values are actually stored in database records.
b A ***key attribute*** (also referred to as a ***primary key*** or ***ID attribute***) whose value can be used to identify each entity uniquely, and;
c Specification of ***unique Identifiers*** or ***codes*** - names, symbols, or tags of the chosen attributes and their characteristics.

Attribute

Every entity has some attributes that characterise it. **'Name'** of a *person* and **'title'** of a *book* are attributes of the entity-type PERSON and BOOK, respectively. A specific value of an attribute of a particular entity constitutes a fact or datum about that entity. For instance, the value of the attribute '**name**' of particular *person*, for example, 'Swami Dayanad Sraswati' is a particular fact or datum about him, and serves as his identifier.

An attribute of an entity-type can be ***intrinsic*** or ***natural***, such as *size*, *weight*, *colour*, and so forth, or ***extrinsic*** or ***attained,*** such as *name*, q*ualification***,** and so forth. Entities also attain special types of attributes by having some ***relationship*** or ***connection with other entities***, or ***due to actions performed on them by other entities*** or ***agents*** that bring changes in their state or condition. For example, the attribute **name-of-father** of a person or **name-of-son** of a person is acquired by an entity PERSON due to its relationships with another entity PERSON. The relationships being **father-of** or **son-of** between the person A and B. Similarly the relationships **written-by** and **classified-by** between the entity classes BOOK and PERSON establishes the attribute **name of author** and **name of classifier** of a book entity. The attributes that show relationship or connection between entities and entity classes are called ***relationship-related attributes*** (including one acquired due to actions performed on them by other entities or agents). Some entity attributes do not describe the entities, as such, but describe what it does, how it is used, why it is useful, what is its use, and so forth . The things an entity does are called activities. The attributes that describe the activities of entities are called ***activities-related attributes*****.** We will call ***relationship-related attributes, action-related attributes*** and ***activities-related attributes*** collectively a ***relationship-related attribute*** or attribute ***of a relationship***.

Thus, we can say that ***an attribute is a property, a piece of information or data about an entity,*** or ***about its relationship(s) with another entity or entities***. **Title**, for instance, is an attribute of an entity class MONOGRAPH, and so is **price, size**, **name of the author**, **name of the publisher,** and so forth.

Please note, **name of the author** and **name of the publisher** are the relationship-related attributes of an entity class MONOGRAPH acquired by it due to its relationships **written-by** and **published-by** with entity classes AUTHOR and PUBLISHER, respectively. The relationship, linkage or connection between the entities (or entity classes) MONOGRAPH and AUTHOR, or MONOGRAPH and PUBLISHER are represented by the terms **written-by** or **published-by** and the relationship-related attributes

acquired by entity class MONOGRAPH are represented by the terms **name of the author** and **name of the publisher**.

A set of attributes including relationship-related attributes (or link attributes), that is the attributes acquired by an entity due its association with other entities, enables us to describe, identify or locate an entity. A particular entity possesses specific **value** for each of its attributes. The values of the attributes (its own or those acquired due to association or relationship with other entities) become major parts of the data stored in the database.

An ***attribute*** has one or more ***data elements*** associated with it and are represented by one or more subfields comprising a field, which contains data values about an attribute of an entity in an entity record stored in a storage medium.

Relationship

A relationship is an association or connection between two or more entities. In the real world, entities do not stand alone and acquire certain attributes due to *relationships* or *connections* with other entities, or due to *actions* performed on them by other entities or *agents* that bring changes in their state.

Relationships connect entities and represent meaningful dependencies between them. Relationships specify specific properties or attributes acquired by entities due to some relation or association with other entities. It may be pointed out that a relationship exists between entities, not between the attributes of entities. For instance, the entity PERSON could be anyone. When the values of the attributes **name, age** and **sex** of the entities PERSON are known we can identify men from women, children from adult or adult from old. When attributes of relationship types, **father-of, mother-of, son-of, daughter-of, member-of, employee-of** are added, then we know that we are talking about a set of persons belonging to different groups in a family, or an organisation. The terms **father-of, mother-of, son-of, daughter-of** show associations or relationships between members of a family. The term **member-of** shows relationship between entities (or entity types/classes) INSTITUTION and MEMBER (Person). The term **employee-of** shows relationship between entities INSTITUTION and EMPLOYEE. Similarly, the term **author-of** or **written-by** shows relationship between the entity classes DOCUMENT and AUTHOR.

A **relationship,** thus, can be defined as ***a logical connection or dependency between occurrence of one entity-type and occurrence of another entity-type***. A ***relation occurrence*** is an individual entity's connection with other entity or entities concerning a specific relationship or connection.

Entity relationships are described by:

- Their **Dependencies** on each other; and
- The **Extent** of the relationship.

The entity **Dependencies** are of two types:

i. **Existence Dependency —** that is, one entity is unable to exist in the database unless the other is first present. For example, entity ORDER cannot exist without entity ORDERED ITEM and entity SUPPLIER, to whom the order is placed for the supply of item or items.
ii. **Identification Dependency —** that is:
   a. An entity cannot be uniquely identified by its own attributes.
   b. Identification is possible only through the relationship it has with some other associated entity or entities. For example, in a large group two persons may have the same names. Thus, a person can only be identified uniquely within a large group either by his name and parent's name or by his name and institution, subgroup, or unit with which he may be associated.

The **Extent** of the relationship implies:

i. The ***type of association or degrees of relationship*** that exist between entities.
ii. The direction of the relationship between them.

Both extents can be represented graphically by line and arrow or by a line and a crow's foot, which show mutual dependence.

Properties of a Relationship

There are two properties of the concept relationship, referred to as ***degree*** (or ***cardinality***) and ***participatio*n.**

- ***Degree of relationship*** (also known as cardinality of the relationship) indicates the number of instances involved in a relationship between two entity classes; or the number of occurrences of one entity that can exist for a single occurrence of the related entity, and vice versa. In general, there are three ***types of relationships or degrees of relationship*** that exist between entities:

1. ONE-to-ONE (1:1) relationship.
2. ONE-to-MANY (1:M) relationship,
3. MANY-to-MANY (M:M or N:M) relationships.

For instance, the relationship between *Libraries* and *Librarians (Professional Staff)* can be one-to-one (1:1), if it is defined in the following way:

- A *Library* is **managed-by** at the most one *Librarian (Chief Manager)*
- A *Librarian* (*Chief Manager*) **manages** at the most one *Library*

In contrast, the relationship between *Libraries* and *Librarians (Professional Staff) can* be one-to-many (1:M), if it is defined in the following way:

- A *Library* **employs** many *Librarians* or *Professional Staff Members*
- In a *Library,* only one *Librarian or professional staff member* can hold the post **of** *Chief Librarian*

The relationship between Authors and Books can be many-to-many, and defined in the following way:

- An *Author* can **write** many *Books*
- Many *Authors* can **write** a *Book*

These three degrees of relationships are also called the relationship ***cardinality***. The term **cardinality** refers to how many of one entity is associated with another. The cardinality ratio specifies the numbers of relationship instances that an entity can participate in.

- ***Participation*** (or ***Optionality***) is the involvement of the entities in a relationship. An entity's participation is considered ***optional*** if there is at least one instance when an entity does not participate in the relationship. The participation of an entity is considered ***mandatory*** if in all instances an entity must participate in the relationship. The default participation is mandatory.

Attributes of a Relationship

Attributes can be assigned to relationships as well as to an entity class. For example, the *many-to-many relationship*, namely, **'issued-to'**, **'loaned-to'** or **'borrowed-by'** that exists between the entity-types BIBLIOGRAPHIC ITEM and LIBRARY MEMBER, possesses the attribute, namely, **'date-of-issue'/ 'issue-date'/ 'return-date of item'** of entity class BIBLIOGRAPHIC ITEM to entity class a MEMBER. Alternatively, we can state that the relationship possesses the attribute **'borrowing-date of item'** and **'title-of item'** when considered in the context of the relationship: MEMBER **borrows** BIBLIOGRAPHIC ITEM. The value of the attribute **'date-of-issue of item'/ 'issue-date of item '/ 'borrowing-date of item'** relating to a given instance of borrowing *a particular* **bibliographic item** by *a particular* **library member** shall be single valued. However, the attribute **'date–of-issue of item'/'issue-date of item '/ 'borrowed-on/ borrowing-date'** of the relationship will be multi-valued when characterising or specifying either the entity class LIBRARY MEMBER or BIBLIOGRAPHIC ITEM *alone* since a bibliographic item can be issued many times and a library member can borrow a bibliographic item many times. Performance and storage utilisation would be optimised at database implementation time by assigning this attribute to the relationship rather than to entities. In this case and similar cases, we must define the

a **Date-of-Issue of Item** (*alternatively* **borrowing-date of item** and/or **return-date of item)** and **name-of-borrower** as ***relationship-related attributes*** of the entity class BIBLIOGRAPHIC ITEM having a relationship with the entity class LIBRARY MEMBER who has taken the item on loan, and/or
b **Title-of-borrowed item** and **borrowing date** as ***relationship-related attributes*** of the entity class LIBRARY MEMBER having a relationship with the entity class BIBLIOGRAPHIC ITEM being **issued-to /loaned-to** a borrower.

2 Ranganathan's Postulates for Designing a Scheme for Library Classification

2.1 Introduction

According to Ranganathan, library classification is concerned with "**classification of *subjects*** or ***Universe of Subjects***" (Ranganathan, 1967, P. 94). A **subject** is a theme of a work - an organised or systematised body of ideas, facts, data, or information concerning ***some thing*** or ***things,*** or ***exposition of theoretical construct(s)*** embodied in a bibliographic item. The content of a bibliographic item is the product of intellectual, transintellectual, creative activity or imagination of man recorded in an intelligible, coherent and communicable form. The focus of study and description in a work or document can be about a ***conceptual*** (*abstract*)

or ***concrete*** (*physical*) ***entity*** or ***entities,*** or ***their properties***. A subject of study is itself something created by man that encompasses a body of knowledge about some thing or things and is governed by a set of fundamental laws, hypotheses, postulates, and principles.

A **work** constitutes the thought contents, information, or knowledge contained in a document or package of information, plus the language in which the thought contents are expressed or communicated. A work is equal to thought contents, which is the ***soul*** of a document and language or any other means of communication of thoughts, which is the ***subtle body*** of a document.

Ranganathan defines a **document** as a "record of a work on paper or other material, fit for easy physical handling, transport across space and preservation through time." The term 'Document' includes any bibliographic item or the record of any kind of work -- macro or micro -- and the physical embodiment is exclusively of one work or is shared by more than one work. It constitutes *subject* (i.e., thought contents, knowledge, information, or the message it holds), plus *language* (communication medium or thought channel), plus *recording material* (paper, electronic media - tape, disk, and so forth; *cf.* **Ranganathan's Prolegomena**. Chapter: TC. 24). *The subject is the soul of a document; language (subtle body), and recording material represent gross body of a document* (**Kashyap**, 1983; 1986, p 200-215).

He defines **entity** as "any *existent*, concrete or conceptual, that is, thing or idea" (**Ranganathan**, 1967, pp. 53). Any *entity* has a distinct *personality* comprised of a set of characteristics or attributes; an **attribute** being "any property or quality or quantitative measure of an entity" (**Ranganathan**, 1967, pp. 53).

If we perceive or know about the attribute or attributes possessed by an entity or entity set we know about it or them. We can also distinguish one entity or entity set from another because of certain characteristic(s) or attribute(s) possessed by one and not by other. For example, we know that the entity ***tree*** is distinct from the entity ***timber***. Because we know that for a living tree it is essential that its root be in soil and for its branches to have access to water, light and air. The 'tree-ness' (the personality of tree) of the living tree is distinct from the dead tree as 'timber' or 'wood'. If soil, water, light and air are removed, there can be no tree left if it turns into timber or wood. The 'tree-ness' of tree is energy (water, light and air) dependent and tied to matter (soil). Further, a tree or specific type of tree can exist in certain conditions and places, or live up to a certain time. Thus, 'tree-ness' of tree is related to matter, energy, space and time in a specific way and none of these elements can be excluded from it. The description of the properties of the tree, its relationships with or dependence on water, soil, air, light and specific environmental conditions, and that it exists in a particular place, and lives up to a certain time, provides us knowledge or information about the entity tree, or portrays its total personality.

Soil, water, air, light, place and time are also entities in their own right, having their *own personalities* that are distinct from the entity tree. Tree is a living thing; where as soil, water, air, light are non-living physical things. Space and time, *per se*, are conceptual entities.

Ranganathan divided the universe of knowledge into traditional **Basic Subjects** or **Main classes** *of subjects* followed by a sub-division of these basic subjects or classes through the application of "***trains of Characteristics***" or "***facets.***" The term '*facet*' refers to a manifestation of any one of the five fundamental categories - Personality (P), Matter or Property (M), Energy or Action (E), Space (S) and Time (T).

According to Ranganathan:

- If a work or document contains an overall description of an entity, an entity set or entity class, which is/are the central theme of a work, or contains an exposition of a theoretical concept or concepts about some entity/entities, phenomenon/phenomena or thing/things then the subject of the work is deemed to be of the type ***Simple Subject.***
- If in a work or a document, one describes only part or portion of the personality of an entity or an entity set, or give description of one or some attributes possessed by an entity or entities, and/or actions on it by or through other entities, in a particular space and time context, then the subject of the work is deemed to be of the type ***Compound Subject***.

  Ranganathan postulated that the make up of a **Compound Subject** constitutes one or more of five mutually exclusive ***fundamental categories***: **Personality** (**P**), **Matter** (**M**) or **Property**, **Energy** (**E**) or **Action**, **Space** (**S**) and **Time** (**T**). This set of fundamental categories for brevity is denoted by the acronym **PMEST**. He also introduced an extended version of these in the form of levels and rounds of their manifestations.
- If a document or a work deals with or contains descriptions of interrelationship, comparison, and so

forth, among two or more basic subjects or compound subjects then such a subject of the work or document is deemed to be of the type **complex subject**.

2.2 *The Postulates for Designing A Scheme for Library Classification*

2.2.1 *Introduction*

The set of postulates formulated by Ranganathan for designing a scheme for library classification, classing the bibliographic documents according to their specific subject, and logically specifying their specific subject statements put the theory of classification, and the work of classifying and indexing the bibliographic documents on a firm scientific basis. These postulates or principles of facet analysis and synthesis were for the first time formally stated by Ranganathan in his paper presented to the *International Study Conference On Classification for Information Retrieval*, held at Dorking in 1957 (Ranganathan**,** 1957a). Though the explicit statement of the postulates or the principle of facet analysis and synthesis was for the first time made by Ranganathan in his paper presented to the Dorking Conference, these were actually applied by him, implicitly, in designing his *Colon Classification Scheme* published in 1933 (Ranganathan, 1933, pp 8). The *fourth edition of Ranganathan's Colon Classification* is the first version of the Colon Classification, whose design is explicitly based on the Postulations approach to facet classification which were for the first time stated by him in a paper presented at the *International Study Conference on Classification for Information Retrieval*, held at Dorking in 1957**.** Ranganathan later incorporated these postulates, in a somewhat elaborated form, in his *Prolegomena of Library Classification. Edition 2* (Ranganathan, 1957b). The enunciation and application of the postulates became more sharpened, clear and simple in his later works, particularly *Prolegomena of Library Classification. Edition 3* (Ranganathan, 1967; Also refer to Ranganathan, 1958).

One of the many major contributions of Ranganathan in the field of Library and Information Science is ***chain procedure***, the first version of which was published in 1938 (**Ranganathan,** 1938). It is a subject indexing technique to derive subject headings from a class number mechanically. The method is based upon Ranganathan's theory of the symbiotic nature of classification and cataloguing.

Ranganathan was of the view that once the class number of a bibliographic item is determined on the basis of a classification system, the *Specific Subject Statement (or Heading)* of a document can be derived by verbal interpretation of the class number and by consequently deriving other related subject headings for a subject catalogue, with the help of the Chain Indexing Technique. However, later he suggested that the Specific Subject Statement or Heading of each bibliographic item could be derived independently of a class number based on the *'Postulates of Facet Analysis'*. In Ranganathan's own words: "The Postulates for Facet Analysis and Principle for Facet Sequence are as much help in finding out the names of the Subject Heading, as they are in establishing its Class Number. The tasks of cataloguing and of classifying are equal beneficiaries of these postulates and principles. The use of one and the same procedure in cataloguing and classifying does not warrant the assumption of subject heading being derived from class number or class number being derived from subject heading, of two different branches A and B of one and same tree, we do not say either A stems from B or that B stems from A " (**Ranganathan**, 1957a).

2.2.1 *The Postulates Enunciated by Ranganathan*

Postulate Of Different Kinds Of Subjects

Ranganathan postulated that a subject of a work or a document can be a Basic Subject, a Compound Subject or a Complex Subject, that is, three types of subjects exist namely: Basic, Compound, and Complex Subject.

– *Basic or Simple Subject*
A subject without any ***isolate idea*** or ***concept*** as a component is a ***Basic*** or ***Simple Subject***. A Basic Subject represents a field of study, a discipline or sub-discipline (branch of learning, e.g., Algebra) or any aggregate of fields of studies (e.g., Mathematics), mutually exclusive and totally exhaustive first order array of subjects of a scheme for classification. *A **basic subject** forms the **Basic Facet** of **a compound subject***. It encompasses a body of knowledge that is concisely governed by a set of fundamental laws, hypotheses, postulates, and principles. Mathematics, Arithmetic, Algebra, Biology, Botany, Cytology, Social Sciences, Political Science, Literature, Sanskrit Literature, Physics, Space Physics, Logic, and so forth, are examples of Basic Subjects.

– *Compound Subject*
A compound subject is a subject having a ***Basic Subject*** (**Basic Facet)** and one or more ***Isolate Ideas*** or ***Concepts (Isolate Facet***) as its components. For example, each of the following subject statement or title of a monograph indicates that the subject matter of the monograph is a compound subject. Examples of compound subjects: **Mining** of ***gold,*** **Chemistry** of ***gold,*** **Biological study** of ***animals, Botanical studies*** of ***flowers,*** or ***Treatment*** of ***cancer*** in the **Ayurvedic System of Medicine**, represents Compound Subjects. (The terms in bold letters represent the basic subject or facet and a term in bold and italics represents an isolate idea or facet of the compound subject). According to Ranganathan, **Isolate Idea** (or **concept**) is any idea or concept which itself cannot be deemed to be a subject but fit to be a component of a compound or complex subject. Further, the make up of a **Compound Subject** constitutes of one or more five mutually exclusive ***fundamental categories***: **Personality** (**P**), **Matter** or **Property** (**M**), **Energy** or **Action** (**E**), **Space** (**S**) and **Time (T)**.

– *Complex Subject*
A complex subject is a two or more phased subject. It is represented by a subject statement, which shows some relation (e.g., bias, comparison, influence, and so forth.) between two or more simple subjects or compound subjects, for example: Physics *compared to* Chemistry or Psychology *for* Doctors.

Postulates of an Isolate Idea or Concept

An **Isolate Idea** (or **concept**) is any idea or concept which itself cannot be deemed to be a subject but fit to be a component of a compound subject or complex subject. Isolate concepts or the terms representing them such as: Gold, Animal, Child, Air, Flower, Cell, Hardness, Goodness, Red, Blue, and so forth, do not convey any meaning for us unless we associate them with some Basic Subject. Further, each isolate idea of a compound subject is deemed to be a manifestation of one and only one of the ***five Fundamental Categories of ideas, defined below.***

Postulate of Five Fundamental Categories

There are five and only five Fundamental Categories of Isolate Concepts or Ideas - *namely*, **Personality (P)**, **Matter (M)**, **Energy (E), Space (S)** and **Time (T).**

An isolate idea or concept belongs to any one of these five ***Fundamental Categories***. These categories are represent by a respective symbol – (P), (M), (E), (S), and (T).

The fundamental category (FC) "Personality," "Matter," "Space" and "Time" can manifest many hierarchical levels or may comprise *facets* of different *levels. An isolate idea* belonging to the fundamental category "Personality," "Matter," "Space" or "Time" may pertain to any one of the *hierarchical levels of the category.*

The fundamental category *Energy* may manifest itself in one and the same subject more than once. The first manifestation of the fundamental category (FC) "Energy" is taken as end of *round one* of the manifestation of the three fundamental categories **"Personality"** (1P), "**Matter**" (1M), and "**Energy**" (1E). The second manifestation of FC "Energy" is taken to end *round two* (represented by symbol [2E]), followed by (2P), (2M), (2E), and so on. The fundamental categories "Space (S)" and "Time (T)" are supposed to manifest or are represented only in the last of Rounds in a subject statement.

– *Personality ([P];* i.e., Entity)

The fundamental category ***Personality*** (P) that represents entity types is the most crucial or essential category among the five categories of Ranganathan. There was a time when many scholars in the field of library science, particularly western scholars, felt that "Ranganathan has not attempted strict definition of personalities" (Vickery, 1958; Roberts, 1969).

It appears Ranganathan chose the term 'Personality' to represent the focal point of ***description or key object or objects of study or description in a work***. A basic subject or domain of study is concerned with the study and description of a conceptual or physical object (entity) or a set of objects (entities). The domain of study Botany is concerned with the study and description of plants (found in nature) as well as Forestry. Physics is concerned with physical phenomena or entities - Heat, Light, Sound, and so forth. These entities can also form attributes of certain other entities. Medicine is concerned with the study of the Human Body (a biological entity), and its organs (subtypes). It may be observed that although study of the entity Plant is concerned with the discipline Botany as well as Forestry - the context or focus of study is different in each case. The Heat, Light or

Sound may constitute properties (attributes) of certain other entities under study.

As **Dahlberg** (1978a, pp. 11-17; 1978b) points out "man is always concerned with two realities: (a) entities or being, such as *inanimate beings* (e.g., gold, earth, water, and so on), *animate beings* (plant life, animal life), *divine beings* and *mental beings*, and (b) *predications* (known/established facts) about the entities or beings, that is, determinations of beings". Correspondingly, Ranganathan divides the subjects dealing with the various kinds of entities, that is, *the living or non-living natural beings (systems), or the material, intellectual, cultural, and spritual products of man and society*, into different **Basic Classes** or **Subjects**. Further, he postulated that the study and description of the entity class or its subtypes, including constituent parts, organs, subunits, and so forth, of a supertype entity or its subclasses falling within the domain or purviewof a Basic Class or Basic Subject (i.e., a field of study or any aggregate of fields of study) are deemed to be the component of the '***Personality***' facet of that Basic Class or Subject. Whereas the *predication about,* or *perceptions of entity-types*, or *the description of discovered, established or known attributes or facts about the entitytypes (supertypes as well as subtypes) falling within the domain or preview of a Basic Classs or Subject* are deemed to be the component of one or more of the remaining ***four facets*** or **Fundamental Catagories - *Matter* (*Property*), *Energy* (*Action*), *Space*** and ***Time***. It appears that Ranganathan picked up the terms for these four Fundamental Categories from the domain of physical sciences.

The category "Personality" or "Personality Facet of Compound Subject" stands for any ***physical*** or ***conceptual entity*** *(*or ***entity set)***, or ***theoretical construct(s)*** about which facts, information, explanation, knowledge, or mental images are formed in the mind of a person, and described or embodied in a work or a document.

It may be observed that when *we collect and outline or explain facts about an entity, we actually give description about its whole personality or part personality - that is we describe its attributes, character, nature, relationships with other entities (i.e., influence, impact, control, etc. on it, of or by some other entity or entities), or its interaction or symbiosis with some other entity or entities, as well as about its being present or existing in certain space-time context.*

According to Ranganathan the physical or conceptual entities or entity types that constitute focus of study or subject of study in a discipline are deemed to belong to the fundamental category "Personality' of that discipline or basic subject.

– *Matter ([M]* Represents Matter Material [MM] and Matter Property [MP] type attributes of the focal entity types)

The isolate ideas or concepts representing the intrinsic matter, material, properties, attributes of entity or entity class; that is, qualities, quantities, functions, activities, processes, growth, change of state, behaviour or characteristics of entities or objects belong to this category. Examples are *Colour, Inflammation, Reliability, Weight, State, Harmony, Authority, Constitution, Structure, Hardness, Softness, and Iron (as content of* Blood or as a *material of* Iron-Table)

Prior to the third edition of his Prolegomena, it was considered by many that the fundamental category "Matter" or "Material Facet" included only physical properties or attributes of an entity. However, in his third edition of Prolegomena, Ranganathan explains that the "manifestations of fundamental category "Matter" are taken to be of two kinds – "*Material* and *Property,*" named as *Matter Material* (MM) and *Matter Property* (MP). He also points out that "it may look strange that the property should be taken along with material. But let us take a table as an example. The table is made of the material timber or steel, as the case may be. The material is intrinsic to the table, but not to the table itself. Moreover, the same material can figure also in several other entities. So also, the table has the property of being 2-1/2 feet in height and the property of having a soft-top or hard-top." *He further explains that each of the isolates ideas,* ***Morphology, Physiology, Disease, and so forth,*** *which were earlier considered as the manifestation of the fundamental category* ***"Energy (E)"*** *are being looked upon as* ***property isolates***. (**Ranganathan**, 1989, pp. 400-401).

– *Energy ([E];* i.e., Action; Represent attributes indicating action on the focal entity types)

The category "Energy" (E) (or Action) covers the isolate ideas or concepts that represent the attributes attained or acquired by the focal entity or entity class due to *action(s)* performed on it by another *entity or entity class or* due to its *interaction* or *relationship* with another entity or entity class that bring changes in its state, properties or charac-

teristics. The concepts or terms representing energetic actions, operation (mental or physical) or impact on the focal or core entity (i.e., the object or thing affected by action on it, by or through another entity) are deemed to belong to the "Energy." In other words, this category demarcates isolate ideas or concepts of the attributes acquired by an entity or entity class relating to the energetic actions or interactions that "may be among and by all kinds of entities – inanimate, animate, conceptual, intellectual and initiative" (Ranganathan, 1967, p. 400).

Isolate Concepts such as *Measurement, Treatment, Evaluation, Diagnosis Calculation, Critical Evaluation, Control, Influence, Impact,* and so forth, belong to this category.

As was pointed out earlier, the fundamental category ***Energy*** may manifest itself in one and the same subject more than once. The first manifestation of the fundamental category (FC) "Energy" is taken as the end of *round one* of the manifestation of the three fundamental categories "Personality" (1P), "Matter" (1M), and "Energy" (1E), which is followed by second round of "Personality" (2P), "Matter" (2M), and "Energy" (2E), and so on. For example, the resulting (value) of measurement (*action* [1E]) is the *property* (1M) of an *entity* (1P) measured with the aid of *yardstick, instrument, or tool (agent) of measurement* (2P), another entity. It may be observed that Ranganathan's concept of *Energy,* referring to action or interaction between the two entities has parallelism in application with the concept of ***relationship*** of ***Entity-Relationship modelling approach.*** In the example, we find that the tool of measurement (entity) has a relationship with the entity being measured and the result of the measurement process (action-measurement is indicative of relationship between entity A and B) is the finding of the specific value of the attribute possessed by the entity measured.

– *Space ([S]* Represents the location of the focal entity types in given time)

The category "Space" (S) denotes concepts, isolate ideas such as geographical location, or place where the entity resides, or where some event or action takes place in relation to an entity or entity set. Location of an entity is also a distinctive property or attribute of an entity.

Examples:
Geographical Space - Continents - Asia, Countries - India, Districts, and so forth.
Population Clusters - Cities, towns and villages, and so forth.
Physiological Formations - Mountains, Deserts, Rain, Forests, Rivers, Lakes, and so forth.
Water formations - Oceans, Seas, and so forth.

– *Time (T)*

The category "Time" (T) includes isolate ideas or concepts, such as *Millennium, Century, Decade, Years, Day, Night, Winter, Rainy day, Dry day, Hour, Second*, and so on. Presence or existence of an entity in specific time is also a distinctive property or attribute of an entity.

Concluding Remarks

It is hoped, after going through the descriptions as presented in the paper one can easily infer that there is likeness between the theoretical suppositions of the ***Entity-Relationship Data Modelling Approach*** and the ***Postulations or Analytico-Synthetic Approach to Knowledge Classification of Ranganathan.*** Both approaches are not only alike in the use and application of a theoretical framework but also are complementary and supplementary. Particularly, some of the concepts of Ranganathan's postulation approach to knowledge classification can be effectively applied in the designing or modelling of computer-based information systems, as pointed out in the paper.

Neelameghan (1991, 1992) also points out that the major steps involved in the three planes of work, namely idea plane, verbal plane and notation plane, as propounded by Ranganathan in his prolegomena (Ranganathan, 1967, Chapter, MA please clarify with-MA is chapter number) are also similar in both domains, as shown below:

Designing a database	*Designing a scheme for subject classification*
Idea Plane	*Idea Plane*
1. Identifying data entities (objects about which data is to be collected)	Identifying/defining the subject domains for which a scheme for classification is to be designed
2. Selecting attributes of data entities of interest to potential users	Selecting attributes of the entities constituting the subject
3. Selecting data model, a schema to map the entities	Selecting a classification model (Hierarchic, faceted,

and their attributes 4. Grouping/dividing the data entities by their common attributes and differentiating attributes (Characteristics) 5. Organising, arranging the groups, subgroups, and units derived at step 4 *Verbal Plane* 6. Naming fields and data elements *Notational Plane* Assigning tags to fields, coding, and so forth,.	freely faceted) for mapping information about the *entities* (concepts /isolates) Grouping/dividing the concepts/isolates by their common attributes and differentiating attributes (characteristics) Organising, arranging the groups, sub groups and isolates derived at step 4 (*Work on the verbal plane)* (*Work in the notational plane*)

Library classification and indexing systems based upon the theory of knowledge classification of Ranganathan or his facet analysis and synthesis approach, are formal devices which have been used very effectively for organising the bibliographical documents in helpful order, and to indicate their subject matter with a purpose of identifying the documents on specific subjects and related subjects available in a library. Facet analysis and the synthesis approach of Ranganathan (also called concept based indexing approach) can also be effectively used to organise and search huge volumes of information being generated on the Internet or on the World Wide Web (or the WWW), the multimedia part of the Internet.

The utility of using facet analysis to organise and search WWW resources is well demonstrated by **Ellis** and **Vaconcetos (**1999). They also point out that ''if Ranganathan were alive today he would be aware of the potential of his ideas for searching and organising WWW materials. Indeed, the genius of Ranganathan is attested to by the very portability of his ideas across time, technology and culture, simply because they address the very foundation of effective information storage and retrieval. Perhaps, in this respect, contemporary WWW developers may find themselves having more in common, or have more time for, Ranganathan, than the current generation of IR researchers appear to have, despite their borrowing IR research results in the form of search algorithms embedded in the various search engines."

The application entity-relationship approach is limited to the domain of data analysis and modelling, a technique for documenting, organising and representing enterprise data. Whereas, Ranganathan's postulations or his analytico-synthetic approach to knowledge classification is applied to create and handle very complex database systems and efficient data retrieval systems relating to the entire knowledge for the use of very diverse users and specialists of diverse fields. It enables us to identify entities, their attributes and relationships with other entities associated with any domain of study. We differentiate and classify these entities into subtypes and supertypes easily, because of his Canons for Characteristics, Canons for Succession of Characteristics, Canon of Exhaustiveness, and Canon of Exclusiveness.

It can be stated that Ranganathan's ideas address the very foundation of knowledge classification, organisation and representation, and that we should devote more time to ponder over the profoundness of his ideas. What was considered decades ago as a "paradigm shift" in library classification theory could surely emerge as a potential "paradigm set" in knowledge representation. Such is the power of Ranganathan's ideas that they have transcended all barriers of space, time, technology, and culture (**Binwal & Lalhmachhuana**, 2001).

Bibliographic References

Binwal, Jagdish Chandra, & Lalhmachhuana. (2001). Knowledge representation: Concept. Techniques and the Analytico-Synthetic Paradigm. *Knowledge Organisation*. *28*(1), 5-16.

Chen, Peter Pin-Sen. (1976). The Entity-Relationship Model: Towards a unified view of data. *ACM Transaction on Database Systems*. *1*(1), 9-36.

Chen, Peter Pin-sen. (Ed.). (1983). *Entity-Relationship Approach to Information Modelling and Analysis*. Amsterdam: North-Holland.

Codd, E. F. (1979). Extending the Database Relational Model to Capture More Meaning. *ACM Transactions on Database*, *4*(4), 397-434.

Codd, E. F. (1982). Relational database: a practical foundation for productivity. *Commun ACM, 25*(2), 109-117.

Codd, E. F. (1990). *The Relational Data Model for Database Management System: Version 2*. Reading, MA: Addison-Wesley.

Dahlberg, Ingeraut. (1978a). *Ontical Structures and Universal Classification*. Bangalore: Sarda Ranganathan Endowment for Library Classification.

Dahlberg, Ingeraut. (1978b). A Referent-oriented Analytical Concepts Theory for INTERCONCEPT. *International Classification*, *5*, 142-51.

Date, C. J. (1995). *An Introduction To Database Systems* (6th ed.). Reading, MA: Addison-Wesley.

Ellis, David, & Vasconcelos, Ana. (1999). Ranganathan and the net: using facet analysis to search and organise the World Wide Web. *Aslib Proceedings, 51*(1), 3-10.

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

Howe, D. R. (1986). *Data Analysis for Database Design* (2nd ed.). London: Edward Arnold.

Kashyap, Madan Mohan. (1998). A proposal for the introduction of an integrative approach in the study and teaching of library procedures, cataloguing and database organisation and designing techniques. In N. M. Malwad (Ed.), *Towards the New Information Society of Tomorrow: Innovation, Challenges and Impact: Paper Presented to the 49th FID Conference and Congress, New Delhi, 11-17 October 1998*, 159-169.

Kashyap, Madan Mohan. (1999). *Computer-Based Library Information System Designing Techniques*. New Delhi: Sterling Publishers.

Kashyap, M. M. (1983). Algorithm for analysis and representation of subject contents of documents in a documentary language. *Library Herald*, *22*, 1-29.

Kashyap, M. M. (1986). Ranganthan's Postulational Approach to Classification, Its Development and Impact. In T. S. Rajagopalan (Ed.), *Ranganathan's philosophy, assessment, impact, and relevance.*

Mair, D. (1983). *The Theory of Relational Database*. Rockville, MD: Computer Science Press.

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

Neelameghan, A. (1991). Concept categorization and knowledge organisation in specialized databases: A case study. *International Classification, 18*, 92-97.

Ranganathan, S. R. (1933). *Colon Classification*. Madras: Madras Library Association.

Ranganathan, S. R. (1957a). Library Classification a Discipline. *Proceedings of the International Conference on Classification for Information Retrieval, Dorking*. 3-14.

Ranganathan, S. R. (1958). Postulations approach to facet classification. *Annals of Library Science*. *5*, 33-51.

Ranganathan, S. R. (1937). *Prolegomena to Library Classification*. Madras: Madras Library Association.

Ranganathan, S. R. (1957b). *Prolegomena to Library Classification* (2nd ed.). London: The Library Association.

Ranganathan, S. R., & assisted by Gopinath, M. A. (1967). *Prolegomena to Library Classification* (3rd ed.). Bangalore: Sarda Ranganathan Endowment for Library Science.

Ranganathan, S. R., & assisted by Gopinath, M. A. (1989). *Prolegomena to Library Classification* (3rd ed.). Bangalore: Sarda Ranganathan Endowment for Library Science.

Ranganathan, S, R. (1938). *Theory of Library Catalogue*. Madras: Madras Library Association.

Roberts, Norman. (1969). An examination of personality concept and its relevance to the Colon Classification Scheme. *Journal of librarianship*, 1 (3), 131-148.

Teorey, T. J., *et al.* (1986). Logical design methodology for relational database: Using the extended Entity-Relationship model. *ACM Computing Survey*, *18*(2), 197-222.

Vickery, B. C. (1958). *Classification and Indexing in Science* (2nd ed.). London:Butterworths.

Whitten, Jeffrey L., Lonnie D. Bentley and Hevin C. Dittman. System Analysis and Design Methods. 5th ed. (2001). New Delhi: Tata McGraw Hill.