KNOWLEDGE REPRESENTATION, CLASSIFICATION AND LIBRARY DATABASE MANAGEMENT

M.A. Gopinath,

ABSTRACT

Analyse the role of knowledge representation, surrogate generation in the context of creation of a database for library collection. Identifies, the interdisciplinary nature of classification studies, surrogation processes for content description, the qualities stipulated to represent subjects and ideas and the basic research done in the area of knowledge representation. The need for formalisation and scientific approaches to representation process is emphasised. Collaborative developments in exchanges of knowledge-bases between general schemes for classification has been indicated. Direction for future research in representation systems for knowledge in library systems is indicated.

1. INTRODUCTION

Classification has been a fascinating field of research in library and information science. For more than a hundred years classification research has attempted to find ways and means to present knowledge contained in documents in a concise and indicative way. It has provided tools and techniques to lighten up the path for scanning, browsing, selecting and retrieving the documents by seekers' of knowledge. Ranganathan defined the function "classification as one of the unavoidable incidents in broadening education into a great highway where upon all can travel all through the life. It is a technique that is inextricably involved in education - education which is not merely the progressive unfoldment of the personality of each individual to the fullest extent and at his own speed but also a composite social process whereby organised community itself develops its won personality and efficiently. The invention of classificatory language is, therefore, not a mere professional indulgence but a necessity brought on the library professional by its new primary mission of education!"(4).

2. TRENDS IN CLASSIFICATION DEVELOPMENT

Thus classification has been a field which sets each specialty in a cognitive manner. The first century of classification was recognised by Ranganathan as movement from enumerative to analytico synthetic one at the Elsionore Conference (3) Recently Elaine Svenonious (6) identified eight ideas which continually fetched progress of classification. Briefly stated they are:

1. The idea of operational definition-that is to define operational aspects of classification and attempt to quantify them. This would lead to establish relationships between quantified concepts and develop a scientific language that could be used to evaluate and compare classifications.

2. To base classification to referent theory. It relates to design of data structures for representing knowledge.

3. The idea of classification as a language. This opens up a panoply of new questions about the semantics, syntax and programes of classification languages structures.

4. The idea of classes being formed by the method of family resemblances. The coincidence, possibly random, of the idea of family resemblances in a philosophical system and the availability of a technology to construct classes based on family resemblances gives the idea credence and survival value.

5. The idea of meaning as use, with related ideas of fluid categories and fuzzy sets. According to Elaine Svenonious, "A potential benefit of a fuzzy-set theory type of approach it that is introduces into classification some shades of grey, there by giving a more subtle and possibly more realistic map of the world of knowledge.

6. The idea of meaning as use, with the related ideas of context dependency and multiplicity of view points. Though traditional classification are able by means of devices like the relative index to show the treatment of subjects from a variety of perspectives, the linearity of there basic structure limits the relationships that can be expressed. The computer with the versatility in the moves it enables and the displays it permits, offers possibilities in relational knowledge representation.

7. The idea of an analytico-synthetic approach to classification, probably, the most fundamental paradigm shift in classification in the twentieth century, has been the shift from largely enumerative to largely analytico-synthetic.

8. The idea that there are general laws and principles underlying classificatory languages, leading towards formalisation and, generalisations.

In the world of information technology, we find

emergence of number and variety databases. Modern databases contain abundance of data classifications. Typically, the searches may have to consult these classifications in separate manuals or files, thus making their effective use difficult. But use of relational algebra operations and integrated with transitive computation help development of representation hospitable to deductive information retrieval.

3. QUALITIES IN KNOWLEDGE REPRESENTATION :

Classification systems are being rested not only for unique representation of subjects, and/or ideas but also to provide for multipurpose connections. Criteria for these can be stated in the context of networking are:

1 Relative simplicity (the wider the range of population served, the simpler the system should be).

2 Multipurpose, multi-fundamental. (multiple associations)

3 Flexibility to absorb and adopt to changes in user requirements in the information world ("Hospitality" in SRR terms)

4 Capable of providing appropriate rules and positive assistance to users by guiding him along paths likely to lead towards relevant information.

5 Requiring minimum of inputting efforts recognition aids and checking for common types of input error;

6 Capable of providing appropriate rules and positive assistance to users by guiding him along paths likely to lead towards relevant information;

7 Representation should be capable of being helpful   to search for unconventional ways (non-formal ways) but not inhibit the conventional searchers; and

8 Alternative solutions for representations in different types of systems such as discipline-oriented, mission-oriented, stretching and management systems, and methods for integrating   variety of demands for multi-functional representations to meet the needs of trade, professions and missions.

These qualities have to be incorporated in a variety of vocabulary control devices complementing classification systems. The cognitive processes in knowledge domains and concepts which are common but occurring in multiple roles have to be studied. Basic research is needed for this Here there are several directions.

4. BASIC RESEARCH FRONTS

Basic research in classification systems and other complementing devices have to utilise quantitative data, linguistic, semantic, logical, and mathematical models, This approach is manifested in the following studies:

1 Use of databases, classification schemes, thesauri; and investigations of the variety of relations between concepts and in different contexts.

2 Rigorous studies on information search patterns and cognitive aspects of the seekers of knowledge.

3 Use of automatic classification techniques for providing useful quantitative measures.

4 Quantification for bibliometric and scientometric studies of world literature in all subject fields.

5 Studies of the institutional networks (Relationship between research organisations, educational institutions, professional associations, industrial establishments, trade systems, business organisations, management patterns, information modelling centres) on the national level and international level, for recognising the interrelationship between subject fields.

6 Developments and use of general classification systems and micro thesauri to provide for compatible exchange of representations and identify the overlap between subjects in different areas of knowledge.

7 Study of contemporary and emerging new models for structuring of classification and vocabularies of natural languages.

8 Application of citation clustering methods for ascertaining the association, interlinking and structure of concepts and subjects as represented in documents.

9 Study of the modes of formation and development of subject-fields particularly interdisciplinary ones.

10 Comparative study of structures of concept-based systems and word-based systems.

11 Several studies on the theory of concepts, patterns of combination-concepts and their variations in different contexts.

12 Categorical analysis of thesauri and classification elements towards more general approach to subject representation.

13 Studies on universal and special classification schemes and the ways of learning the structures.

14 Research on the fundamental syntactic and semantic structures deduction of appropriate for representation of subjects in different contexts.

15 Research on deductive classificatory language from highly abstract level language. This can represent multiple or hierarchical classifications

with connected basic classification derived from Wall-Picture.

5. TOWARDS SCIENCE OF CLASSIFICATION

Ranganathan strived for developing a science of classification. A theoretical structure continuously guiding practice and any in explicable situation occurring in practice lead back to the theory. Ranganathan called for dynamic theory classification. The same inquiry approach continues to prevail. The application of the classification theory to knowledge classification particularly knowledge representation and also to connectivity sequence of the ideas in artificial neutral networks. Increasingly, formalisation in representation of ideas in the context of new intellectual aspects of computer-communication technologies is taking it towards computer science. Elaine Svenonious on this, she says "The fact-of-the matter is that to date hypertext, cluster-analysis, and knowledge representations supporting reasoning had only a tangential effect on the design of working information retrieval system. The potential of these representation are yet to be realised. I would like to conclude by suggesting those research areas that might contribute to actualisation of potential offered by these knowledge representations. The first is cognitive studies, the second is social epistemology and the third evaluation research" (5). There is still room for basic work to be done in knowledge representation areas.

6. DEVELOPING SCHEMES FOR CLASSIFICATION

Keeping classification schemes upto data have several problems. The first and foremost is to refashion their original structure to meet the current needs. Knowledge organisation these days demand quick restructuring. Knowledge domains such as cybernetics, computer science, representation theory, pattern-recognition, cognitive studies, management studies reveal the shifting paradigms that calls for relocations and re-orientation. While the old is retained, the new approaches have to be represented to give comprehensive knowledge representation. Towards this end, we find a recent phenomena where in Bliss Bibliographic Classification and the Universal Decimal Classification aim to collaborate (7). This investigation is an attempt towards the application of the knowledge-base of one scheme over the other. This exploratory study attempts to test a methodology used to restructure the Universal Decimal Classification class by class. The study uses the facet framework established in Bliss Bibliographic classification, second edition, as the basis' for the restructuring of UDC. The research is being carried out using the discipline of Medical Science. It also derives thesaurus from the restructured scheme (9). The methodologies designed for schemes for classification in India have been attempting to find this universality of import

and export between classification schemes, thesauri and index. The three planes of work demarcated by Ranganathan namely, idea plane, verbal plane, notational plane do have the basic contribution towards this universality. To quote Elaine Svenonious again "Recent literature assessing the adequacy of knowledge representation is skimpy and of a wheel reinventing nature, one writer (Woods) for instance, stipulates two criteria for a knowledge representation for an expert system 1) expressive adequacy, two components of which are the distinctions it can specify and those it can leave unspecified to represent partial knowledge, and 2) notational efficacy, three components of which are computational efficiency, conciseness of representation and ease of modification (Woods). At the Elsinore Conference in 1964 Ranganathan called for a study of abstract classification. Now nearly thirty years later it seems appropriate to echo his call in the context of machine-based knowledge representations; we need basic research, we need to examine the assumption made by such representations, we need to develop a meta-language for comparing them; we need to develop criteria for evaluating them as retrieval structure (8).

7. CLASSIFICATION THEORY AND ITS PURPOSE

In an excellent review on the role of Ranganathan's classification theory on knowledge-base system which includes library systems, information system and expert systems besides many others, A Neelameghan 92) identified the following features:

1 Organizing concepts in and structures knowledge bases,specialized databases , etc;

2 Designing schemes for classification;

3 Designing and developing vocabulary control tools, such thesaurus, classurus, etc.;

4 Generating various types of structured indexes;

5 Preparing field definition table/data dictionary for databases;

6 Recognizing inter-relationship among concepts and linking them accordingly;

7 Assisting user to browse / negative more conveniently in the database to zero-in on the specific areas of interest to him/her at the moment;

8 Facet analyzing user's query and structuring search expressions for better results in on-line retrieval;

9 Presenting the retrieved records/ information in a sequence helpful to users;

10 Object-oriented analysis and design; and

11 Possibly in neural networks and cognitive modelling.

These aspects are the base for knowledge organization.

8. CATEGORISATION AT GENERIC LEVELS

Further, we see that in knowledge seeking there is a flow from unfamiliarity to familiarity, uncertainty to certainty, ignorance to understanding or from haziness to clarity. In other words, the knowledge seeking phenomenon exhibits a kind of abstract to concrete situation, a kind of no information to a kind of all information. This presents a kind of taxonomy of stage of knowledge consumption. Knowledge organization has to incorporate this taxonomy for convenience of assimilation. Thus every communicator of knowledge aims to reflect this approach. While taxonomic organization of knowledge is a necessity, the number of object, events and their complexities in this world appear, at phenomenal level, as different and calls for unique organization of these object, events etc. If one tackles this issue directly he may not be able to solve it easily and the very factors which he wants to organize may lead to chaos than order. It is here that the scientific method begins to operate. The numerous complex details are reduced to a few fundamental parameters devoid of details. The indicate a kind of simplicity in their relation and help understanding of the patterns which may secure differently in different environments. The crucible of human knowledge processing reduces numerous objects and events to a few fundamental categories. The ramification of the fundamental categories can occur in different combinations and in different context. This is a kind of analytico-synthetic process. By representing the fundamental categories and identifying a typology of the relations that these categories instantiate, we can provide a representation which is coextensive with the idea complex that we are representing. The situations in Information Science and those in knowledge base in the computer communication world adopt a variety of representation systems which exhibit all the features of the components of the knowledge they represent. The constraints particularly hardware ones, which were inhibiting knowledge representation in the computer communication records are being overcome by continuous and productive research in the field of Information Technology. It is this versatility of hardware systems that has given scope for looking at the nuances that are needed to transfer information to cognitive minds for assimilation. The cognitive paradigms in knowledge organization are thus having intensive attention and intellection.

9. TEACHING AND LEARNING IN CLASSIFICATION

In recent years, in the context of information retrieval, classification's role has been well recognized. Teaching and learning calls for strong theoretical foundation coupled with continuous practice and application to specific domain of knowledge warranted by disciplines, missions, problem-solving and decision-making aspects. In this new environment has emerged anew mode of thinking the developments, teaching and learning in classification. Here, I quote G Bhattacharyya (1).

1 Classification in the context of information retrieval is a device for vocabulary control. It is possible to implement this control in the notational plane as well as in the verbal plane. What are the specific purposes for which are plane is more efficient and effective than the other.

2 It is necessary that the classification for organizing the resources in its store should be the same as that for information retrieval.

3 The classifier or the indexer, the learner and teachers of the same have all long been depending upon the designer of the subject indexing language and classification for the supply of a ready-made scheme and for its updating. For the purpose of macro-classification and indexing there has always been some ready-made schemes available. But for micro-classification and indexing there has been hardly any. If there is any, it is rarely found to be suitable for the specific purpose in hand. If there is any suitable scheme, it is found to have been developed locally to meet specific requirement at the local level. In the circumstances should the old tradition of depending upon the designer be continued? Or, should the classifier and indexer be simultaneously be the designer specially when the purpose is depth indexing from the point of view of micro-documents in a specific area? If the later is desirable what would be its implications on education and training for the professional work and service?

4 It is possible to develop a general theory which can form the basis for the derivation of specific purpose oriented subject indexing language especially to ensure compatibility of various SILS are basic or found level? If so, what would be the procedure? What would be its structure? What should be the basic components of any SIL? What should the procedure to derive specific purpose oriented SIL?

The answers to these questions are in the process of taking precise shapes (1).

The teaching and learning process should impart these knowledge and skills to operate in any information environment where knowledge organization plays a vital role.

Teaching/learning the classification process in an in knowledge organisation environment should provide for the following capabilities to information professional and knowledge organisers:

1 Recognise the existence of the context of the whole, its varieties, and component's identity; adds to the process of Analysing.

2 Identify the goal seeking principles for combining components of information; add to the process combination.

3 Provide a mode of executing these components of information in an orderly fashion; adds to the process ordering.

4 Orient the information available to the relevancy of user context; add to the process of Discrimination

5 Provide a mechanism for assimilation by an information seeker, the relevant knowledge, filtered from numerous sources; adds to the process of Assimilation.

6 Organize a database of modulated cumulative organization of knowledge, adds to the process of Conservation.

7 Present a semantic map of existing knowledge and indicates gaps in knowledge; adds to the process of Conservation.

8 Provide a guide to the process of novel ideas; add to the process of Innovativeness.

9 Provide a holistic picture of coherent knowledge at any moment of time; adds to the process Intuition.

These capabilities have to be gained through intensive experience in practical world of design and development of information system and service. Any formal course can provide an in linking to forward into the prospects or gaining these skills.

10. FUTURE DIRECTIONS

Classification research tends towards development of formal theories to guide a variety of applications. It should help development of schemes appropriate to the needs of knowledge organisation. Further, research in the field should work towards the following aspects :

1 Classification should tend towards expressiveness of concepts, subjects, and universe of subjects; there should be movement towards unique representation of concepts and their contexts.

2 Classification system should be made compatible for efficient communication or various interferences needed in situations. That is, it should be able to provide generic structure and in-depth structures.

3 Classification system should have intellectual and surrogational preciseness. Scheme should tend towards compactness, and clarity; appropriate level of abstraction.

4 Different types of knowledge should be capable of being expressed with the same general representation scheme. Universal system. providing for atomised special systems of surrogations.

5. Classificatory language should reflect logical consistency. The neutral exclusiveness should be contextually separated and represented.

6. Classificatory language should be able to provide schemes for multiple representation of knowledge.

7 Development classifications schemes for surrogation that are progressively comparable with searcher's needs-user friendliness.

The areas of research would be towards integrated development of the conceptual structure of subjects, classification structure and surrogational economy. This would help the searchers in moving towards faster retrieval, comprehensive browsing and focusing and interfacing. The role of classification is envisaged by Ranganathan in the following quotation: "Classification can illumine the field of knowledge; it can be prophetic. To make it fit for this new role, it must be placed on a new type of foundation and built in a new way. This calls for considerable search in communication of thought. It has to depend on helpful arrangement and selection. Whatever be the psychology of arranging and selecting process, this is facilitated by the use of artificial counters or symbolics ... Imply as much purposive abstraction as mathematical symbols of formal logic. Research towards absolute syntax which approximates thought process of majority of specialists in a field of knowledge is essential (4).

BIBLIOGRAPHIC REFERENCE

1 BHATTACHARYA (G)
On Classification and Indexing in Science and Technology. *DRTC Annual Seminar*, 20; 1983 x-xii

2 NEELAMEGHAN (A).
Application of Ranganathan general theory of Knowledge Classifications in Designing Specialised Databases (*Libri*. 42; 1992; p 202-206).

3 RANGANATHAN (S R).
Classification and Communication. 1950 reprint 1989, p 260

4 RANGANATHAN (S R).
Library Classification Fundamentals and procedure. 1994, Madras Library Association p 392.

5 RANGANATHAN (S R).
Library Classification through a century. (In Atherton (P). Classification research : *Proceedings of the second International study conference* held at Hotel Prince Hamlet, Sept. 2964, p 23-35.

6 SVENONIOUS (E).
Classification; prospects, problem and possibilities. (In Williamson (N J) and Hundon (N), Editors Classification research and knowledge representation :Proceedings of the 5th National study conference on classification

Research (Toronto) (1991). 1992 Elsevier
Science Publisher B.V p 18, 19.

7   SVENONIOUS (E).
The impact of computer technology on
knowledge representation. (In Neelameghan (A)
et al. eds. Cognitive paradigms in knowledge
organisation; *Proceedings of the 2nd ISKO
Conference* held at (Madras) (Aug. 1992) p 10).

8   ibid - p.12

9   WILLIAMSON (N J) AND McILWAINE (I).
A feasibility study of the Restructuring of the
Universal Decimal Classification into a fully
faceted system. (Albrechtsen (Hnne) and
Oernager (Susanrie) eds.   Knowledge
organisation and quality management :
Proceedings of the 3rd ISKO Conference
Copenhagen, June 1994 Frankfirst Indeks
Verlag p 406-412)