Chapter: 1

Introduction

- Introduction
- Library Processes
- Document Selection - a key process
- AI application in LIS
- Literature Review
- Need for present work
- Scope of the present work
- Objectives
- Hypotheses
- Chapterization
- References
Chapter 1 Introduction

1.1 Introduction

We are in the midst of the information revolution. Computer technology has made a major breakthrough from pure numerical calculations to intelligent problem solving known as Artificial Intelligence (AI).

AI research is concerned with developing intelligent systems; that is, systems that exhibit the characteristics we associate with intelligence in human behaviour, solving problems, learning, understanding language, interpreting visual scenes, and so on.

One major sub-branch of AI is Knowledge Based Systems (KBS). These are systems that imitate or substitute for the reasoning processes and knowledge of experts to solve specific types of problems. One unifying factor in these systems is the expertise of people. Systems have been developed in highly domain-specific areas incorporating the expert's knowledge into computer systems and forcing them to take decisions based on the given knowledge. As in many other fields, a few experiments have been carried out in Library and information field also. There are expert systems designed in areas of retrieval information and representation. But it is necessary to take a look at the other processes which could be considered for further modelling work.

1.2 Library Processes

Library or an information centre is a dynamic system. Various sub-systems work within the whole system. There are many parameters which keep varying and thus the system is a dynamic one. The most important is the user needs. In order to cater to the user needs and achieve user satisfaction.
the library identifies locates collects and maintains information sources. Thus the activities involved encompass the areas like identification, location, acquisition, processing, retrieval, and dissemination. While all the above-mentioned are important processes, the most important is the document selection. The information services rendered from a library can only be as good as the collection. Relevant information can only be served by a focused and exhaustive collection. Therefore considerable thought and decision making go into putting together a collection that relates to the user needs and thus forms an interesting area of study dealt hereafter as Document Selection.

1.2.1 DOCUMENT SELECTION A KEY PROCESS

Selection of reading materials to the University Libraries in India rests on the faculty members; whereas in USA, UK, and some other countries, the subject bibliographers working in the libraries are responsible for selecting relevant reading materials for their libraries.

India has one of the largest Higher Education Systems in the World. It has 267 Universities over 10600 Colleges spread in all the States. University Grants Commission is responsible for coordination, determination, and maintenance of standards. Release of grants, Professional Councils, are responsible for recognition of courses.

Central Government is responsible for taking major policy decisions relating to higher education in the country. It provides grants to the UGC and establishes Central Universities in the country. Central Government is also responsible for declaration of Educational Institutions as Deemed to be Universities on the recommendations of the UGC.
Presently there are sixteen (16) Central Universities in the country. They are

Banaras Hindu University (1915)
Aligarh Muslim University (1920)
Delhi University (1922)
Viswa Bharati University (1951)
Jawaharlal Nehru University (1969)
North Eastern Hill University (1973)
University of Hyderabad (1974)
Pondicherry University (1985)
Indira Gandhi National Open University (1985)
Jamia Millia Islamia (1988)
Assam University (1994)
Nagaland University (1994)
Tezpur University (1994)
Maulana Azad National Urdu University (1998)
Dr. B R Ambedkar University (1998)
Mahatma Gandhi Anter Rashtriya Hindi Vishwavidyalaya (1998)

In pursuance of the Mizoram Accord, another Central University in the State of Mizoram is planned. There are 37 Institutions that have been declared as Deemed to be Universities by the Government of India as per Section of the UGC Act 1956.

State Governments are responsible for establishment of State Universities and Colleges and provide plan grants for their development and non-plan grants for their maintenance. The coordination and cooperation between the Union and the States is brought about in the field of education through the Central Advisory Board of Education (CABE).
Education is on the Concurrent list subject to Entry 66 in the Union List of the Constitution. This gives exclusive Legislative Power to the Central Government for coordination and determination of standards in Institutions of higher education or research and scientific and technical institutions.

In the Central University libraries the policy of Document Selection is a fairly standardised process. The responsibility of the selection and approval is assigned to a committee which mainly comprises the senior faculty and researchers. Demonstration of the system developed for the present work is for the Indira Gandhi Memorial Library, University of Hyderabad. In this university library, the document selection process follows the standard procedure as in the other Indian Central University libraries. The selection committee comprises several senior faculty by the different departments and programs. In spite of the selection of books by the concerned faculty members, the users express that the library is not having the books they want. This is because of the fact that the selection is not based on scientific judgment, utility, and other parameters. The faculty members who select books are supposed to be subject experts in their areas. Though the user and his needs are the central considerations in the selection process, it is rather difficult to assess the needs in definite terms. Even if the broad subject scope can be easily understood, the intent of the subject is often elusive. Generally, the factors to be considered are:

- Subject scope
- Depth of coverage
- Level of exposition
- Language
- Form of document
- Easy access to information (indexed document)
There are other considerations like the cost and availability against each given subject and decision making regarding which is the best suited among many publications with claim of same subject coverage. The entire process involves intuitive inferencing and learning and re adjustment of the system. If the process could be demarcated into well defined modules and the related processes could be put down without ambiguity then it would facilitate modeling a system using AI techniques.

1.2.2 AI APPLICATIONS IN LIBRARY AND INFORMATION SCIENCE

Till recently applications of AI have proved useful in fields of science and medicine where knowledge can be reduced to series of rules or other well defined relationships. Systems have been developed to provide medical advice (1) (with regard to diagnosis, treatment and the monitoring of the patients) hospital information systems and clinical information systems. Other research covers game playing, machine translation and manipulation of symbols (2). Experiments and applications currently being undertaken in business and industry indicate that the applicability of AI is expanding. For example, systems such as vision controlled assembly by multiple manipulator robot, an image analyzer in quality control or the vision system of an automatic inserter for printed circuit board assembly.

Recently people started applying AI in library operations. Ongoing research experiments and applications indicate its inevitable future use in libraries. Some of the areas where Expert Systems (ES) are likely to make their greatest impact include Reference Service (REFSEARCH) (3).
Chapter 1 Introduction

Indexing (MEDINDEX) (4) Cataloguing (ESSCAPE) and front end systems for database searching (5)

The very first application of computers in libraries were for relieving the burden of routines jobs. But the real challenge of deploying computerised solutions are in the areas which need human intellect. The challenging jobs involve a high degree of decision making and implementation of an action plan. Librarians and Information Specialists have displayed a growing interest in AI KBSs. Particularly for Online search and Reference Assistance.

1.2.3 LITERATURE REVIEW OF AI APPLICATIONS

Knowledge Based Systems have a variety of applications in Library and Information Science fields. Knowledge Based Systems are rapidly becoming part of the Librarians vocabulary. Knowledge Based Systems offer a level of technology sophistication that has never been experienced in the library world. Already work is in progress in areas such as reference work cataloguing classification indexing and intelligent information retrieval. The following sections will briefly describe about their applications as derived from the literature available.
There was a time when library was considered as a storehouse and books were meant for preservation. The Librarian was merely considered as a custodian. A modern library is regarded as a service institution. Its aim is to convert potential users into habitual users. To achieve this, it is essential to establish right kind of contact between the user and the document. Right contact means contact between the right reader and the right book at the right time and in the right personal way. This method is named as reference service. The reference librarian is expected to use the total resources of the library and also the resources available outside the library including national and international.
Reference Librarian involves himself in the following

a) Answering queries received from the users
b) Assisting the users in using Library tools i.e. catalogues and other reference tools
c) Preparing reading lists/bibliographies on request or in anticipation
d) Preparing guides for the users
e) Providing selective dissemination of information (SDI) to users
f) Providing interlibrary (ILL) loan facilities
g) Providing abstracting and indexing services
h) Providing translation/photocopying services etc

All the above activities can be grouped under the following three categories

1. Minimal or conservative approach
2. Maximum or liberal approach
3. Middling or moderate approach

In the first category the role of the Librarian is considered as a guide to the users. The emphasis is on helping the user to locate the information sources answering his information need or rather a kind of self help. The second category is opposite to the first approach where the emphasis is on delivery of information rather than instruction. The third approach is a via media between the two extremes. In this approach, the reference librarian provides the reference service to the extent considered suitable to the users and the occasion

Answering directional questions blends into reference and ready reference and the complex queries require in-depth exploration and extensive search leading to the use of online databases.
The systems developed before early 80s to assist the reference librarian in his reference work have closely resembled database system than knowledge Based system. To answer the simple straightforward queries a database system is sufficient to meet the needs. When complex problems are to be solved like creating information products suitable to the user needs, Knowledge Based model are necessary to meet the user needs. The complex queries require in-depth exploration and extensive search and even use of online databases.

Normally the Reference Librarian goes through the following steps:

1. the person has an information need
2. the person verbalises that need
3. the Reference Librarian interviews the person in order to rephrase, refine or augment the query statement
4. the Reference Librarian considers the type of access points to be searched
5. the Reference Librarian selects specific reference books to search

Following are some of the systems developed to meet the reference needs of the users:

1241 Refsearch

Refsearch is a system that supplies patrons the recommended sources to lookup for certain questions. The system can be used to teach students reference skills or as a computerised aid for practising reference librarians and information specialists. Refsearch incorporates a small indexing language that can simultaneously categorise a question and the desired works that may answer it. Current advances in software make it feasible to
redesign the system so that it can use natural language and run on a microcomputer in libraries (10). Refsearch was used to characterise about 145 works in the Library School's general reference teaching collection at the University of California Berkeley campus. The software tried was Insight 2+ which incorporates knowledge based system building tools a Pascal compiler and dBASE III within the same shell on an IBM microcomputer.

The Reference Librarian makes use of two sorts of expertise in handling general reference questions. Knowledge of procedures to be followed during question negotiation and knowledge of types of information found in reference books. Insight 2+ allows to model the Librarian's procedural knowledge of how to conduct a question negotiation. The rule type knowledge base with backward chaining inference maintains the experts' knowledge as a series of rules in if then form. This is well suited to modelling the reference librarian's procedural knowledge of question negotiation. The Factual Knowledge about reference books is maintained in dBASE III files. The inference engine allows the knowledge based system to draw conclusions from the information in the database. A backward chaining inference engine will test for all routes to the goal retrieval of a particular book and fire all sequences of rules it finds.

1242  POINTER

POINTER was the early successful working application of computer system in the area of reference work. POINTER was created in 1984 for federal documents at the SUNY Buffalo. This system runs on an IBM PC written originally in LISP and later converted into BASIC. The program was developed to meet the shortage of staff at the Lockwood Library (6)
POINTER only directs the users to the reference sources. It does not tell them how to use the sources. POINTER is not a knowledge-based system but a computer-assisted reference program. POINTER is written in a straightforward manner. Dialog with the user is primarily through menu choices. Input from the user is tested against IF THEN condition statements (which is as close as this system gets to 'rules'). In response to the menu questions, the user makes appropriate choices, and the system at the end of the interaction responds with the names of specific reference sources, their call numbers, and brief instructions about using them.

Librarians get a lot more complex questions, but POINTER is not capable of handling them. Furthermore, if the person is looking specifically for one of the commercial reference books by title, POINTER will not help.

### 1243 ORA (Online Reference Assistance)

ORA is a system developed at the University of Waterloo Library by a team of Librarians (7). It is intended to simulate the services of an academic reference librarian for questions of low and medium level by using several technologies: a videotex-like database, computer-assisted instruction modules, and knowledge-based systems. Like other systems, it gives help in performing subject searches, but it also gives assistance in tracking down publications for which the reference is already known. ORA consists of directional transactions like library locations, services, and policies. Holdings transactions like the library holdings and substantive transactions like factual information for help in finding information on a topic and so on.
ANSWERMAN

National Agricultural Library (NAL) Maryland has developed a knowledge based system called ANSWERMAN (8) to help Library users find answers to ready reference questions on agricultural topics. It uses series of menus to narrow down the subject of the questions and the type of tool needed (directory encyclopaedia, atlas etc.) It can function as either a consultation system or as a front end to external databases and CD ROM reference tools.

PLEXUS

PLEXUS is a system developed at the University of London by Alina Vickery and Helen M Brooks (9) as a referral tool for use in Public Libraries. It includes knowledge about the reference process, information retrieval, certain subject areas, reference sources and Library users. The system uses various forms of knowledge representation to handle these different kinds of knowledge: heuristics, rules, frames, and semantic nets, and its sophisticated deployment of natural language processing. Frames and semantic networks for handling the reference interview for subject queries.

PLEXUS is tested in the subject area of gardening in agriculture. National Agricultural Library (NAL) began development of an advisory system for ready reference in 1985. The objective was to create a system in a relatively narrow field of agriculture that could help the users in locating either appropriate information sources or specific answers to their questions without professional assistance (10). Long range goal was to create a system which could be used by the Librarians to develop advisory system for different fields of knowledge and would permit those knowledge bases to be linked together in a single growing system covering much broader
subject areas For this purpose an expert system shell "1st CLASS was used which permitted development on an IBM PC. Other reference advisors developed by using 1st CLASS expert system shell include Biographical Reference Advisor at the Goucher College (11) MEDSTAT at the National Library of Medicine Reference Advisor on American Literature at Austin College in Sherman Texas (12) and an advisor on Engineering at the California State College at Pomona Kenneth Quinn. A number of reference advisors were developed later including an advisor on business information sources by Ralph Alberico of James Madison University using the VP Expert Shell.

All the above systems are advisory systems for locating reference source books and factual data. Professional reference often deals with complex concepts in depth like citation retrievals. Full text retrievals and concept of searching etc. A new concept search package TOPIC uses object oriented programming which enables Boolean searching. TOPIC treats a hierarchical outline describing the concept as a search subject.

A number of Knowledge Based Systems building tools now offer hypertext capability as well including 1st CLASS HT and Knowledge Pro. NAL contracted with a Knowledge Engineering Group that used Knowledge Pro (an object oriented tool) to assist NAL in creating an intelligent reference work covering African agriculture.

12.5 CATALOGUING

Cataloguing is one of the oldest library crafts. These were prepared only on the basis or rules of thumb without taking into consideration the functions to be performed by a catalog. The present day codes such as AACR2
Chapter 1 Introduction

contain rules formulated in a systematic manner based on normative principles. A Library catalogue is an essential and important tool in a Library. It is a key to the Library holdings. A cataloger studies the documents for the purpose of recording and interpreting them for its potential users. He prepares the catalogue with the help of rules given in cataloguing code. Cataloguing consists of various processes involved in the preparation of entries and maintenance of a catalogue. Each entry is designed for satisfying a particular approach of a user. The objectives of the Library catalogue are

1. To enable a person to find a book of which
   a) the author or
   b) the title or is known
   c) the subject

2. To show what the Library has
   a) by a given author
   b) on a given subject
   c) in a given kind of literature

3. To assist in the choice of a book
   a) as to its edition (bibliographically)
   b) as to its character (literary or topical)

Cataloguing is concerned with three separate but related activities. They are

1. Describing the item and choosing the name and title access points
2. Assigning classification number
3. Assigning subject headings
The first activity is governed by a complex and detailed set of rules. Activities 2 and 3 involve subject analysis.

Recent attempts to automate cataloguing through knowledge based systems have focussed on descriptive cataloguing because it is considered rule based (AACR2). There are two approaches for applying artificial intelligence techniques to cataloguing (13):

a) A human machine interface where the intellectual effort is divided between the intermediary and the support system.

b) A knowledge based system with full cataloguing capability linked into electronic publishing systems so that as a text is generated on line it can be passed through a knowledge based system cataloguing process without any intellectual input from an intermediary.

The first project to automate cataloguing using KBS techniques was carried out by Davies and James (14) at the University of Exeter, England. This study emphasised a software interface to Prolog rather than creation of knowledge base but the prototype system was never implemented.

An attempt was initiated at Linkoping University, Sweden (5). The project ESSCAPE (Expert System For Simple Choice of Access Points for Entries) involved implementing a set of rules from Chapter 21 of AACR2 (Anglo American Cataloguing Rules 2nd edition) that deals with the choice of access points. EMYCIN, a shell system for building knowledge based system was used to implement the rules. The authors concluded that the AACR code is flawed with reference to the choice of access points.
Chapter 1 Introduction

The reasons mentioned are
1. too much is taken for granted
2. the base entries are not properly defined
3. some concepts overlap and as a result the rules are poorly constructed
4. exceptions are made on an ad hoc basis
5. the relationships between general and specific rules have not been properly thought out

For a KBS to generate the required conditions for a specifically defined goal, the logic in the knowledge base must be consistent. There have been problems in every attempt to convert AACR2 into the highly structured rules necessary to run a KBS.

Literature on KBS is generally optimistic and sometimes even sensational. Current research on automatic cataloguing has had mixed results but researchers continue to suggest further work. The impact of this technology on AACR2 will be great because of the demand for consistency. When one looks at the previous decisions of those responsible for AACR, it is safe to conclude that more changes in the code will be forthcoming.

12.6 CLASSIFICATION

Classification is the fundamental activity in the organisation of knowledge. For this reason, it is prominent in all systems for organising knowledge in libraries and information centres and also in the knowledge bases of knowledge-based systems.
Classification may be defined in two ways

1. Creating categories of like objects, actions, or concepts that are useful for some defined purpose and arranging these classes into a conceptual structure based on the relations among them.

2. Identifying individual entities as instances of classes in an existing classification system.

The objects, actions, and concepts as entities can refer to the conceptual structure of a classification system. The examples include Dewey Decimal Classification (DDC) and the Library of Congress Subject Headings (LCSH).

Application of knowledge-based systems in the area of classification in libraries include the following:

1. CoalSORT
2. EP X and
3. BIOSIS automatic indexing system
4. DRTC Classification research and systems development

**1261 CoalSORT**

CoalSORT is a conceptual browser designed to serve either as a search or an indexing tool. CoalSORT consists primarily of a frame-based semantic network and the software needed to allow users to display portions of it and to move around in the conceptual structure. The domain is very narrow and covers catalyst applications in coal liquefaction and contains 500 concepts that have been used to index about 100 documents. The expert knowledge...
in the system is embodied almost entirely in the semantic network. There is no procedural knowledge in the system.

The frames for the concepts in the system contain dictionary-like definitions of the concepts and relationships of the concepts to other concepts in the network. The inference component of the system is a multi-window screen with a mouse, which allows users to select concepts they wish to add to their search statement or to explore further. The windowing allows users to view the current query, the concepts they have already examined, and the documents found as of that time.

### 1262 EP X

The Environmental Pollution Expert (EP X) (15) has certain things in common with CoalSORT in that both are concentrating on enhancing interfaces using a knowledge-based approach. The knowledge base of EP X consists of a hierarchical frame-based semantic network of concepts and a set of templates that express the patterns called the pragmatic relationships among concepts. These patterns are referred to as contextual information.

### 1263 BIOSIS

The earlier two systems were basically networks of term relations used to help the user browse in the system or to help the system make intelligent suggestions to the users to modify strategy. In BIOSIS we can see fuller implementation of a KBS (16). BIOSIS uses a knowledge base including a significant amount of procedural knowledge to assign documents to categories automatically. Like the MEDLINE project, BIOSIS is designed as an indexer aid. BIOSIS uses the information in the titles of biological
documents to assign as many categories as possible of those that would be assigned by human indexers

It can be concluded that there is a close link between classification and KBSs. The indexing languages are structured and practical representations of information that can be used to very good advantage of AI applications.

1.2.6.4 DRTC Classification research and systems development

According to F.J. Devadason (17), the Deep Structure Indexing System is based on a set of postulated Elementary Categories of the elements fit to form components of names of subjects, a set of syntax rules with reference to the Categories, a vocabulary control tool such as Classaurus, a set of indicator digits to denote the Categories and their sub-divisions, and a set of codes to denote a few of the decisions of the indexer. Devadason has attempted generation of different types of subject index entries based on deep structure of subject indexing languages.

A.R.D. Prasad (18) has developed Prometheus, an automatic subject indexing system using NLP techniques. The work demonstrates that NLP approach has much potential in subject indexing as it attempts semantic analysis. The system adopts the deep structure of SIL as a semantic structure of document abstracts. That is, the deep structure subject indexing language serves as Meaning Representation Language (MRL) for a NLP system. The parser of the system works at the syntactic and semantic levels. At the syntactic level, the grammar formalism identifies the noun phrases and at the semantic level, the system identifies the roles of each noun phrase in order to build a frame based knowledge representation of the thought content of the information processed. Once the meaning of a
A given abstract is represented in a frame based model the system generates the subject strings.

Following the development of Prometheus another project was carried out on the design and development of knowledge representation model for automatic indexing. It resulted in a frame based model built for generating subject indexing strings according to POPS1 (19) (20). There have been other efforts such as application of NLP techniques to convert natural language queries to Boolean operators (21). A model for thesaurus construction has also been developed with the application of frame based models and inference built in PROLOG code.

A system Viswamitra (22) has been developed using frame based representation which automatically assigns Colon Classification numbers to input expressive titles. This system follows all the prescribed steps in practical classification and assigns the class number for the titles. Following the trend in the representation techniques evolved earlier Vyasa an integrated system which attempts the management of analytico-synthetic classification system has been developed (23). It adapts the earlier model in Viswamitra for representation of the Colon Classification schedules as the work starts with the problems in classifying documents. It recognizes the new terminology that need to be accommodated in the schedules and resolves through frame representation and rule based inferencing mechanism and fixes them in the appropriate position in the schedules. It thus enables automatic updation of the schedules.
127 INDEXING

Indexing of periodicals is another area where knowledge based systems are being developed. Indexing a periodical article involves identification of concepts discussed in the articles, translating these concepts into verbal descriptions and selecting and assigning controlled vocabulary terms that are conceptually equivalent to these verbal descriptions. But the intellectual aspects of the indexing process however are not yet automated. The reason for automating the intellectual aspects of indexing is to improve the indexing consistency and quality. A variety of indexing errors occur because understanding indexing rules are relatively complex. Knowledge based systems are very helpful in this regard because they include not only the semantic networks but also procedural rules. Based on the information provided by the indexer, the knowledge based systems can arrive at appropriate preferred terms automatically to assign relevant subdivisions. The system can make inferences and based on the inference it can take appropriate action. A prototype of such a KB indexing system is already developed and tested at the National Library of Medicine (NLM). The knowledge base of the indexing system is encoded in the form of frames.

1271 MedIndex

MedIndex Project at NLM investigated a KB approach to indexing the periodical literature. A prototype MedIndex system (24) has been developed on NLM's VAX 11/780 minicomputer under UNIX operating System. The system is designed to interact with trained MEDLINE indexers. For each input consisting of a MEDLINE citation to a document, there are two forms of output describing the subject content of the document.
a) Indexing frames created by an Indexer using the System

b) Set of Medical Subject Headings (MeSH) indexing terms generated automatically by the application of knowledge base rules to the indexing frames

Very few library users have interacted with knowledge based systems. In general, users have had very little contact with these systems due to the fact that most of them are not perfected well enough to be used by the everyday library patron. Most of the systems are in experimental stage and very few library KBS are being marketed by vendors.

12.8 ACQUISITION

In the words of Sorgenfrei and Lane (25) Reference librarians' experience plays an important role in collection development and it should be applied in systems modelling. This is further endorsed by the views of Ferguson (26) who emphasizes that along with the library staff and selectors, the users have a significant role to play in building electronic collections and that their help and advice should be solicited in the process. Several systems have been built which incorporate the techniques discussed above.

Monograph Selection Advisor (27), a pioneering effort in applying this emerging technology in another area of Library Science, is building library collection. Specifically, the task modelled is the item by item decisions that a subject bibliographer makes in selecting monographs.

The aim of Knowledge Based systems is to make use of the library easier. Therefore, transfer of intelligence from person to the machine and vice
versa must be effective. The knowledge base has to be broad enough and the interfacing aspects must be easy enough for the library to get the desired information from the machine. The degree of user success with KBS will depend on how well the system is designed, the completeness of the knowledge base, the ease in using the system, etc.

We will witness several new applications of expert systems to libraries in the near future as a result of more sophisticated shells and new developments in control strategies for the knowledge base. Advancements will surely be made in the level of intelligent problem solving offered by knowledge-based systems. Such systems are expected to reach a level of performance comparable to that of a human expert in a specialized problem area.

1.3 Need for the Present Research

The aim of the library and information center is achieving user satisfaction through efficient information services. The information services should be backed by strong and relevant document collection. The process which leads to a pertinent and useful collection involves some complicated inferencing. There are various criteria to be considered before the document is pronounced as selected to be part of the collection.

Subject expertise and concept familiarity is the topmost requirement in putting together useful collection. That is the reason in the academic libraries the book selection process often involves a committee of teachers and researchers in the respective areas of study. But more often than not,
the collection does not meet requirements of the users satisfactorily. There can again be many reasons for this. To point out a few of them:

- The process of selection may be treated as routine and recommendations may be made casually rather than seriously.
- There may not be as many documents of a recommended subject in the market by reputed publishers.
- If there are good publications, the library may be unaware of it.
- There may be inconsistencies in the recommendation procedures.
- There may be inconsistencies in the acquisition procedure.
- Cost of the publication may be prohibitive.
- Document is in the electronic form only but recommended only if in print OR vice versa.

For all such factors which come into play, it is necessary to formulate a plan of action to deal with the issues in a decisive manner and unambiguous manner.

Deploying computerized solutions for the document selection process imposes rigor and demands highly defined set of rules for the procedures. This is true of AI-based models developed for library functions. Though the processes like representation and retrieval of information have been given enough attention in modeling the area of document selection has been neglected. It is necessary to justify the budgetary allocation for collection development by giving a document collection which caters to all the needs of the users. But often the aim is missed because there is no rigor in the whole process. Because of the defects in the existing system, it is proposed to design a Knowledge-Based System for document selection using...
computer technology. This KB document selection model based on scientific method will become a meaningful exercise and lead to user satisfaction. This model will help the persons responsible for selecting books rather than replace him.

1.4 Scope of the Thesis

The scope the present work falls in the area of Library and Information science particularly the document selection process. The effort is to test the feasibility to employ the AI techniques in building a model that is capable of making crucial decisions in Document Selection.

Knowledge Based Systems is a branch under AI where many applications have been developed successfully in the areas of information processing and retrieval. The expertise in a given field is built into the system so that the inferencing component has the knowledge of the expert. The present system is supplemented with the PLAN OF ACTION for document selection in the university library. The work involves interpreting the decisions taken at various stages in Document Selection. The activities and rules have to be defined. Later the process is to be incorporated into a Knowledge Based System which would aide in the process of document selection. Therefore the work involves the design of a knowledge based system incorporating the knowledge of the documents selection process so that it could be deployed for selection process.
15 Hypotheses

1 A Knowledge Based Document Selection system will help to build useful collection
2 Document Selection procedure is amenable to AI modelling
3 A Knowledge Based Document Selection System will streamline the process of Document Selection and eliminate inconsistencies and inaccuracies in the process

16 Objectives

1 Study and understanding of AI KBS and its impact in Library applications
2 Study and understanding of Document Selection process
3 Designing a Knowledge Based System model for Document Selection
4 Establish usefulness of this model in the University of Hyderabad Library Environment

17 Chapterization

Chapter I Describes the task of document selection and its significance in the library process. Introduces the concepts of AI technology and its applications in various fields. Further, this chapter also mentions the AI applications in the area of Library Science. A review of literature for various applications in the field of library science is presented. It describes need
for the present work its scope hypotheses and objectives

Chapter II  
Describes in detail the trends in AI. It describes the important branches of AI. Gives an overview of the relevant technologies under AI useful for the present work.

Chapter III  
Describes in detail the trends in Document selection including item by item selection model developed manually.

Chapter IV  
Presents various design issues of KBS model for document selection.

Chapter V  
This chapter presents the rule based inference engine developed for the decision making process in the document selection procedure in the University of Hyderabad library according to the criteria fixed.

Chapter VI  
Discusses the new dimensions to document selection in the wake of Digital Libraries and Internet. Discusses a model for incorporating Internet based document selection especially the subscriptions to online journals.

Chapter VII  
Covers the research results, conclusions and possible future research directions in this area.
18 Conclusion

Information related activities like retrieval and organisation are deemed as highly intellectual processes. But the process of document selection has been neglected. There is ample opportunity for scientific and methodical principles to be implemented for document selection so that finally the collection relates to the needs of the users. The aim of this thesis is to make definite modules for the work involved and incorporate the process in a Knowledge Based System that would streamline the activities related to Document Selection. There are enough advances in artificial intelligence field to facilitate the modelling work. The next chapter gives an overview of the artificial intelligence techniques that are related to the present work.
19 References


2. MARTIN (W A) AND FATEMAN (R J) The MACYSMA system. *In Proceedings of the second symposium on symbolic and algebraic manipulation* 1971 pp 59-75

3. VICKERY (A) and BROOKS (H M) PLEXUS the expert system for referral. *Information Processing and Management* 23 (2) 1987 pp 99-117

4. HUMPHREY (Susanne M) MedIndEx system medical indexing expert system. *Information Processing and Management* 25 (1) 1989 pp 73-88

5. HJERPPE (Roland) OLANDER (Birgitta) and MARKLUND (Kari) Project ESSCAPE expert systems for simple choice of access points for entries applications of artificial intelligence in cataloging. *In Paper presented at IFLA 51st General Conference 1985 August 18-24 Chicago* IL LIBLAB Linkoping Sweden 1985 p 17

6. BRAJNIK GUIDA and TASSO An expert interface for effective man machine interaction. *In Cooperative interfaces to information systems* Springer Verlag Berlin FRG 1986

7. SMITH (Karen F) and SHAPIRO (Stuart C) Final report on the development of a computer assisted government documents reference capability. First phase St Unrv of New York Buffalo 1984
Chapter 1 Introduction

8 PARROTT (James R) Expert systems for reference work Microcomputers for Information Management 3 (3) 1986 pp 155 171

9 WATERS (Samuel T) ANSWERMAN the expert information specialist an expert system for retrieval of information from library reference books Information Technology and Libraries 5 (3) 1986 pp 204 212

10 VICKERY (A) and BROOKS (H M) PLEXUS the expert system for referral Information Processing and Management 23 (2) 1987 pp 99 117


12 ALBERICO (Ralph) Expert systems and microcomputers Small Computers in Libraries 7 (1) 1987 pp 10 13

13 CLARKE (Ann) and CRONIN (Blaise) Expert systems and library/information work Journal of Librarianship 15 (4) 1983 pp 277 292

14 DAVIES (Roy) and JAMES (Brian) Towards an expert system for cataloguing some experiments based on AACR2 Program 18 (4) 1984 pp 283 297
15 SMITH (Philip J) SHUTE (Steven J) CALDES (Deb) and CHIGNELL (Mark H) Knowledge based search tactics for an intelligent intermediary system *ACM Transactions on Information Systems* 7 (3) July 1989 pp 246-270

16 VLEDUTS STOKOLOV (Natasha) Concept recognition in an automatic text processing system for the life sciences *Journal of the American Society for Information Science* 38 (4) 1987 pp 269-287

17 DEVADASAN (F J) Computer Based Systems for Generating Different Types of Subject Indexes and Alphabetical Classaurus Based on Deep Structure of Subject Indexing Languages Karnataka University Ph.D Thesis 1984 (unpublished)


19 APTAGIRI (Devika V) Design and development of a knowledge representation model for automatic indexing *A Guided Project Submitted for the Partial Fulfillment of the Course Leading to Award the Associateship in Documentation and Information Science* 1992 - 1994 DRTC Bangalore

20 APTAGIRI (Devika V) GOPINATH (M A) and PRASAD (A R D) A Frame Based Knowledge representation paradigm for automating POPS Knowledge Organization *Knowledge Organisation* 22 (3-4) 1995 pp 161-167
Chapter 1 Introduction

21 SRISHAILA (Smitha) Natural Language Interface for Bibliographic Databases *A Guided Project Submitted for the Partial Fulfillment of the Course Leading to Award the Associateship in Documentation and Information Science* 1995 – 1997 DRTC Bangalore


23 MADALLI (Devika P) A Knowledge Representation Model for Management of Analytic Synthetic Scheme A case study in Colon Classification University of Mysore PhD Thesis 2001 (unpublished)

24 HUMPHREY (Susanne M) MedIndEx System medical indexing expert system *Information Processing and Management* 25 (1) 1989 pp 73 88

25 SORGENFREI (R) and HOOPER LANE (C) Book selection responsibilities for the reference librarian professional benefit or burden? Library Collections Acquisitions and Technical Services 25 (2) 2001 pp 171 178

26 FERGUSON (A W) Digital Library selection maximum acess not buying the best titles libraries should become full text Amazon corns *Journal of Administration* 31 (2) 2001 pp 27 38

27 SOWELL (Steven L) Expanding horizons in collection development with expert systems development and testing of a demonstration prototype *Special Libraries* 80 (1) 1989 pp 45 50