RISØ-M-2794

THE BOOKHOUSE: MODELLING USERS' NEEDS AND SEARCH STRATEGIES AS A BASIS FOR SYSTEM DESIGN

ANNELISE MARK PEJTERSEN

Risø-M-2794

ABSTRACT

The BOOK HOUSE (and its predecessor, the BOOK AUTOMAT) are computer systems aimed at supporting library users in finding fiction. They are examples of systems using a cognitive work analysis as the basis for design and evaluation - and build on studies of the information retrieval task using actual user-librarian negotiations. These studies lead to the identification of a set of user strategies for retrieving literature as well as a multi-dimensional framework for classifying fiction. Both of these have been necessary prerequisites for the design of the BOOK HOUSE, a flexible on-line data base retrieval system for novice users - both children and adults. The use of icons within a spatial metaphor to facilitate direct manipulation of the displayed objects forms the basis for the interface in supporting the users' navigation in the BOOK HOUSE. The cognitive work analysis which preceeded the design and evaluation of the prototype systems are covered in this report.

In this second edition an appendix is added on page 121 with a summary of the observations made by the librarians of the function of the BOOK HOUSE system in their daily work situation at Hjortespring Library.

LIST OF CONTENTS

SUMMARY

Project description

In the dynamic world of modern information technology and systems, a key issue underlying the work presented in these reports is the support of information retrieval. This task is a crucial element in practically all domains of professional and indeed leisure activity. It is more and more common - e.g. within the integrated work station environment - for users to require access to a large number of different data bases and to combine, process and integrate information from these sources as an important part of their work. However each of these data bases reflects the efforts of specific domain experts and there exists no common frame of reference for indexing and searching and presentation which is based on user needs and ways of problem solving. If this were available, then system solutions could be found which would let the user choose search attributes which covered the specific domain and, at the same time, allow the user to solve the current problem in a natural way.

The project to develop the BOOK HOUSE system represented a step along the way to establish a user-based system design approach - in this instance, within the area of fiction retrieval.

This report describes the empirical analysis of verbal records of fiction retrieval in public libraries. The work analyses of this library task formed the basis for the design of the classification system for fiction, which was used to develop the SPRING database with 3.500 classified novels. This project was supported by the Royal school of Librarianship and Folkebiblotekernes Rådighedssum, both in Copenhagen. The results of this project were used to design the interface of the BOOK HOUSE system.

Project partners were:

The Royal School of Librarianship
Annelise Mark Pejtersen (Project leader)
Inge Cramer
Hans Andersen

Herlev Public Library,
P. Grønholt Pedersen
Lone Gladbo

In cooperation with:

Risø National Laboratory,
Department of Information Technology,
Section for Information Systems and Cognitive Studies,
Annelise Mark Pejtersen (project leader), Len Goodstein,
Finn R. Nielsen.

The Royal Danish Academy of Fine Arts,
School of Architecture,
Institute of Visual Communication,
Steen Agger, Henrik Jensen.

Jutland Telephone's Laboratory
Ove Steenberg

Regnecentralen A/S,
Bent Kryger Sørensen

Project results

The results of the system evaluation were overwhelmingly positive in that users (both children and adults) found the BOOK HOUSE both useful and pleasurable to use. All of the available search facilities were used and the fiction classification system was fully accepted. The utilized iconic/text representation was preferred by a large majority. The designers are convinced that cognitive task analysis is a necessary prerequisite to the design and evaluation of user-work friendly systems, to modify the usually quoted descriptor.

The BOOK HOUSE program has been loaned out to a number of libraries in Denmark. Other groups, companies, institutes - both in Denmark and abroad - have also shown

considerable interest and the system, which can run on a lap-top personal computer, has been widely demonstrated and distributed. The project has been described on two Danish television and four radio programmes as well as in daily newspapers and in both journal articles and international conference papers.

The cooperation between architects, engineers, programmers and information scientists was hectic but productive and hopefully is a forerunner of continued efforts of this kind.

Project reporting

The official reporting of the design and evaluation of the BOOK AUTOMAT and BOOK HOUSE systems is covered in three volumes:

Report nr. 1. THE BOOK HOUSE - Modelling Users' Needs and Search Strategies as a Basis for System Design, by Annelise Mark Pejtersen, Risø National Laboratory, Risø -M-2794, 1989.

This covers the theoretical and empirical work which generated the framework for the design of the BOOK HOUSE, supported by The Royal School of Librarianship and Folkebibliotekernes Rådighedssum.

Report nr. 2. THE BOOK HOUSE - System Functionality and Evaluation, by L.P.Goodstein and Annelise Mark Pejtersen, Risø National Laboratory, Risø -M-2793.

This covers the design and evaluation of the BOOK AUTOMAT and the BOOK HOUSE., supported by Risø National Laboratory, Department of Information Technology, Section for Information Systems and Cognitive Studies, Jutland Telephone's Laboratory, Regnecentralen A/S. and The Danish Technical Council under the 'TUP' programme.

Report nr. 3. THE BOOK HOUSE - VISUAL DESIGN by Steen Agger and Henrik Jensen, Royal Academy of Fine Arts, School of Architecture.

8

This report describes the BOOK HOUSE interface and deals with the reflections and processes connected with the generation and transfer to computers of the visual information provided by the BOOK HOUSE- also supported by Risø National Laboratory, Department of Information Technology, Section for Information Systems and Cognitive Studies, Jutland Telephone's Laboratory and Regnecentralen A/S and The Danish Technical Council under the 'TUP' programme.

Mention should also be made of the following report supported by the Folkebibliotekernes Rådighedssum and The Royal School of Librarianship: Indexing rules for Fiction. Annelise Mark Pejtersen, and Jutta Austin The Royal School of Librarianship, to be published in 1990.

Peter Grønholt Pedersen - Helsingør Public Library

Inge-Lis Sjørslev - Hvidovre Public Library

1. INTRODUCTION

The rapidly increasing use of modern information technology in the interface between humans and their work will lead to a significant change in the content and methods of librarians' and users' search behaviour. Traditionally, tools for professional intermediaries in libraries are specialized and, therefore, often designed to support one separate function. Consequently, the coordination of the individual functions of a task was the concern of the librarians, and was not reflected in the design of tools, only in the selection of tools to use. This has also been the case with the early computerized tools which implemented card catalogues as online search facilities. Since a large number of bibliographical databases are now available with powerful technological capabilities added to the traditional tools, the main concern has therefore been the improvement of the interfaces with new front end systems. However, this situation is now changing, as more integrated tools become available in terms of work stations for particular jobs. Many trivial routines can be automated, and the role of the user will in a rapidly increasing degree be to control the co-function of advanced tools and their adaptation to variations in the work condition. Advanced tools are frequently introduced in order to achieve effective changes of work content in order to match the increasingly dynamic requirements of the environment. In libraries, there is the same need for tools that cope with a dynamic environment: The rapidly and individually changing demands from library users are for example caused by the increase of educational level and the growth of information, the rapid change of user requests and demands, which are also caused by an immediate effect from the mass media and the electronic age. In the future,when work stations for the library work are going to be designed, it is necessary for the library and information science field to consider not only the user-tool interface, but the user-task interaction. User interfaces cannot be added separately as a front end system to a finished functional design of the system in order to achieve a good user-system match. In a dynamic work environment with a computerized work station, the user's job will be knowledge management from a variety of changing information sources. The user should adapt to changing requirements and modify and supplement the conception of the work content and its requirements along with the use and experience of the work station. The user will have to access a number of internal and external databases with different contents and formats. Integration of different databases is not sufficient. Contents and formats of

databases must match users' information needs in a given work situation. Their conception of "aboutness" or contents of different information sources is crucial for the design of databases with access and representation of knowledge that match the users' information need and criteria of choice during their performance of a task. For the design of interfaces to databases, knowledge about the information content of a problem domain, about the task and the users' search strategies and their performance criteria during the search process is necessary.
This analysis of user-task interaction will be the basic problem of the design of integrated work stations in libraries.

This report describes a number of empirical studies of the problem domain of end users and intermediaries during their performance of the complex fiction retrieval task in libraries, including their decision strategies and information needs. It was the aim to gain information from a real life setting through verbal records and use this knowledge as a basis for the design of new information retrieval tools. The results have been implemented in the design of a new information system for fiction retrieval in libraries which is called THE BOOK HOUSE.

2. FICTION RETRIEVAL IN LIBRARIES

Today, the retrieval of fiction in libraries is based on an alphabetical classification by authors for shelving arrangements supported by bibliographical databases with author/title access including a number of additional bibliographical data- and sometimes supplemented by an annotation about the story of the book. This system offers the advantage of uniqeness in the arrangement, identification and retrieval of an individual document, but it is of little help for the users with needs related to the contents and other aspects of documents. Typically, there are no exact or unique answers to such needs, nor are there within the present computer-based retrieval systems any aids, which the intermediary or end user can utilize to establish a basis for their search strategies. There are two main difficulties: first the problem of identifying the user's often subconscious and intuitively formulated need, secondly, the problem of formulating a relevant search strategy among documents that are not indexed and represented according to needs, but according to formal, bibliographical data. Difficulties may arise here, since the main part of the request/book correlation is usually rather intuitive and it is most likely that many of the habits of the users are based on ill-defined attitudes and associations, which define the fundamental frame of reference for comprehension and interpretation of most users' requests and criteria of satisfaction. This problem can be solved by the design of a data base with access to document contents that is compatible with the users' request/need criteria, and by the design of an interface to the database that supports their choice of search strategy. The design of a database with a representation of document contents that reflect end users' formulations of their information needs requires a new notion of the information content of the problem domain and a new kind of indexing scheme for analysis of document contents. Traditionally classification and indexing schemes have been developed to reflect the contents of a document in terms of its relationships with the knowledge structure of the subject field to which it belongs. The analysis and representation of what a document is about does usually not take the users' requests as a focus. In the following, a new approach to the design of a classification system derived from empirical analyses of users' inquiries in libraries is presented. Before that a short review of previous, traditional approaches will be presented.

Previous classification schemes

Traditionally, imaginative literature is organized by formal criteria such as author's name, language of text or literary form rather than by content. Universal classification systems such as the DEWEY Decimal Classification (DC) or the Universal Decimal Classification (UDC) etc., do not usually contain classes for fiction as specimens. If fiction has to be classified, it is either grouped with the literature about it or, if necessary for the production of catalogues, special subdivisions are sometimes added for that purpose. In the British National Bibliography, for instance, extra symbols are added to the Dewey literature class marks to identify novels as opposed to works about them. In a few classification schemes, separate classes for fiction as specimens were provided from the outset, but they were, even then, most commonly defined and subdivided by formal aspects only. For example, in the Library of Congress Classification, fiction is arranged by period first and then by Cutter number, and in the Danish Decimal Classification (DK5, a Dewey derivative), the first division is by language and the second by literary form.

On shelves, novels have traditionally been arranged alphabetically by author/title. Only one alternative approach has been utilized for some time, both on shelves and in catalogues: Classification by genre. Although this represents an attempt to group fiction by content rather than by formal aspects, the individual genres are usually identified by such general or selective criteria that they hardly constitute a sounder basis for effective retrieval than that gained from grouping fiction by formal criteria.

Examples of classification schemes

Over the past ninety years or so, a limited number of people have experimented with the content classification of fiction. Three typical approaches can be identified:

1. Universal non-fiction classification schemes were adapted for fiction.
2. Special classification schemes were designed for fiction.
3. Fiction was organized into genre-systems, based on the reading appeal of different features of novels.

Universal Systems

One example of a universal system adapted to fiction is Frank Haigh´s 1930´s version of the Dewey decimal Classification, which he used to classify 5000 novels (Haigh 1933). He kept all main classes intact, but the original subdivisions were replaced by an incoherent mixture of familiar genre headings and newly invented subject categories, interspersed with some of the original subclasses. In each main class, subclasses were arranged in numerical, non-hierarchical order. The exact contents of the classes were not easy to establish or predict; for instance, ghost stories were classed under "philosophy", whereas mystery stories and thrillers were found under "sociology". Some novels about love and marriage belonged to "philosophy", while others were classed as "useful arts-homelife". Haigh realized that it would be difficult to mark clear boundaries between his classes. When problems of overlap occur, he suggested that the classifier should "make decisions".

Special systems

Half a dozen attempts have been made during the last hundred years to develop special classification schemes. The first classification scheme for fiction was designed in 1898. These classification systems were usually based on their designers´personal ideas about fiction and its value for the promulgation of education and culture. A typical example is the first of two different schemes for shelving fiction designed by C.A. Burgess (Burgess 1936). This system was a hierachical decimal classification. It is not altogether clear how Burgess defined his subject classes, but they seem to represent a combination of genres and literary styles: Examples of a main class and its subdivisions:

30 Love stories and erotic romances
31 Simple, non-tendentious, love stories
32 "Cinderella" formula
33 Straightforward domestic formula; including stories of marriages
34 "Triangle" and "polygonal" formulae
35 Erotic: domestic and ruritanian
36 Erotic: exotic, topographical and primitive ("cave-man" etc.)

37 Erotic: picaresque ("love 'em and leave 'em" etc.)

38 Erotic: abnormal

39 Short stories of love and erotic romance.

This classification was intended to be a critical guide and a standard of taste in fiction. His appending notes as to characters and manners were of a considerably moral quality, and contained recommendations as well as warning to children and women against certain "indecent" books. Masterpieces were marked with an asterix. Another example classified fiction according to topics and the author's intention to influence the reader with social values. His ranking of social influence was from low to highest:

1. Entertainment and escape, 2. Challenge of the reader by criticism of life or standard values, 3. Education for citizenship, 4. Socially purposive to the point of tendentiousness 5. Literary or aesthetic significance.

Four categories or literary value for the staff use were added:

A = Literary or "Highbrow"

B = "Middlebrow" competent and having literary ambitions

C = "Middlebrow" without literary ambitions, but competent "best-seller".

D = Frankly "Lowbrow", commercial and mechanized

Symbols such as + or - were used to indicate occasional moments in the book which were above or below an average standard. + - marked erratic or uneven work. Librarians were expected to use these categories when guiding users, either towards a category of higher quality, or towards new subjects in a preferred value category. Burgess stressed the point that "... these symbols are intended solely for staff use - to make accessible to every member of the staff the mature judgement of the librarian or his book selectors. They should never be used as arranging symbols or appear in public records."

The philosophy of application was the conscientious librarian's attempts to raise the readers' standard systematically by recommending a better book within the same subject class or a book with identical quality but of a different subject class. These are two examples of schemes that attempt to match the professionals' and society's perception of the intermediary's task: to perform subject retrieval that matches the user's need and in addition combine it with promotion of literary values.

However, they are based on subjective speculation and arbitrarily chosen cultural values.

Genre systems

Genre systems consist of pragmatic groupings of novels according to a number of selective features popular with readers. When a genre-classification is used for shelving purposes, it invariably leads to the division of a library's fiction stock into two parts, because genre systems are generally so crude that the majority of novels cannot be classified by them but have to remain in alphabetical arrangement by author. Alternatively, alphabetical shelving of all novels may be retained, and an additional genre-grouping is shown on spines and in card indexes by means of symbols such as letter notations, graphic signs or colour codes (Corbett, 1938).

Genre systems consist of co-ordinate classes which are not in the least mutually exclusive. Nevertheless, exclusivity is frequently assumed when novels are organized by genre: each novel is only classified by one of its characteristics and cannot be retrieved via other, equally dominant, features. For instance, a book would be classed either as a historical novel or as a love story, even though it could well be both. In this way, many relevant features are lost for retrieval. (Figure 1).

Figure 1

18
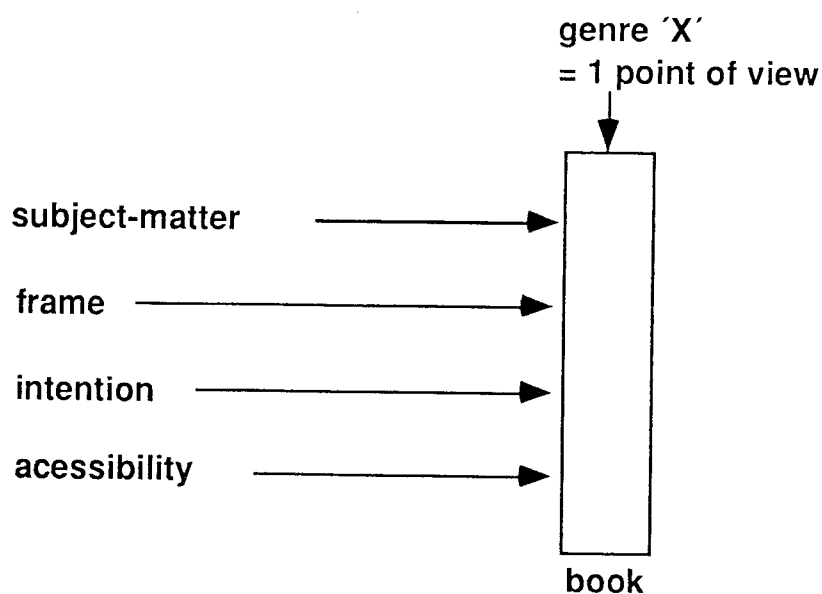
Nevertheless, in the librarian's search strategy and his/her communication with the user, the genre-classification generally serves as a frame of reference, and consequently some libraries have set up a genre-catalogue. The traditional genre classification characterizes only one aspect of the book (for example, a love story) and is insufficient in view of the multidimensional needs of the user. It is inadequate because it ignores relevant combinations of matching books to the user's needs. The employment of heterogeneous criteria in the genre classification, as seen from designations like love story (a subject-matter criterion) and social criticism (a criterion of author's intention) already indicates several dimensions relevant to the ordinary reader. (Figure 2).
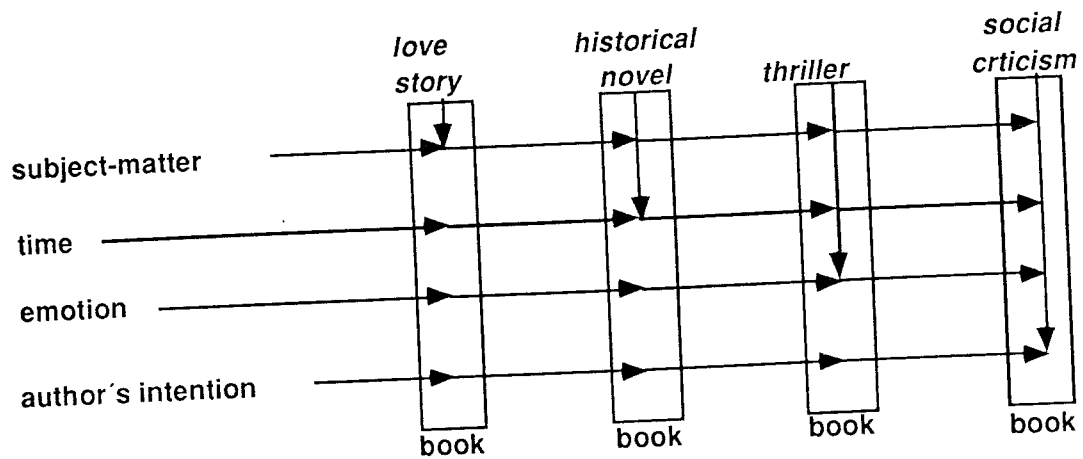
Figure 2

The exclusive and heterogeneous genre-classification has functional implications, which make it ineffective as a model for search strategy and communication.

Conclusion

Early fiction schemes were difficult to use in practice, because:

1. Designers of new systems as well as adaptors of existing ones drew heavily upon their personal sets of values and perception of fiction when they constructed their systems;

2. Genre-classes were regarded as true and adequate reflections of users' needs;

3. These systems did not reflect the typically complex nature of fiction requests, where the subject is only one among several possible and equally relevant criteria (such as period, theme, accessibility);

4. The fundamental incompatibility between the multidimensional character of users' needs and one-dimensional shelving of books was not understood, and systems were therefore intended for shelving as well as for retrieval purposes.

Basically, the system design must focus on retrieval systems that reflect the user's value criteria in terms of subject needs and literary quality and at the same time be flexible enough also to allow the search to be controlled by the professionals' or society's value criteria which might coincide or be in conflict with the end user's need.

3. CLASSIFICATION BASED ON USERS' NEEDS

The present contribution represents an approach which integrates two principal assumptions of the design of an indexing system:

1. The concept of what a document is "about" should be defined on the basis of an empirical analysis of the formulations and needs of groups of users within a specific information retrieval environment.

2. Users' needs are multi-dimensional and the definition of the information content of the problem domain should reflect the different aspects of a general reader's needs. An indexing system should be closely related to the users' formulations of their needs and to their criteria for choosing documents.

The success of an indexing system depends on the degree of agreement between the system's and the users' concept of "aboutness" in documents. Due to psychological factors there is often a "gap" between users' requests and their real needs. If we are to create an adequate superordinate system, which is compatible with users' requests and needs, the ill-defined reading experience and document characterization of users from different viewpoints based on their needs and interests must be made explicit and integrated in the definition of "about" employed by the indexing scheme. As a consequence, the indexing scheme should define the boundaries and attributes of a reasonable and handy number of aspects within several of the dimensions which characterize the documents in relation to users' needs.

Therefore, empirical studies of users' needs based on their requests regarding document contents have been conducted. In the 1970's 500 negotiations between end users (adults and children) and intermediaries were recorded in public libraries (Pejtersen, 1974, 1976). These libraries varied with respect to geographical location (city and country libraries) and with respect to type (central libraries, branches and hospital libraries). After the search, the user and the intermediary were interviewed: If the user was a child, he/she was asked about age and school class and the circumstances which led to his/her inquiry to the intermediary. The intermediary was asked about her knowledge about the documents included in the search and negotiation with the user.

EXAMPLES OF CHILDREN USERS FORMULATIONS.

DIMENSIONS	EXAMPLES
1. Subject-matter	
a. action/course of events	with murder
	about housebreaking
	and boys
	about dirty tricks
	mystery novel, book with action
b.psychological development	love story, book with psychology
	children who beat each other
	about birthdays
c. social relations	about divorce and parents
	family chronicle, not politics
2. Frame	
a. time	historical novel,
	books from 16th and 17th century
	girls in old days
	about the future
	in modern times
b.place	travelogue, books from the countryside
	about castles and mansions
	from the jungle about piracy
c.social/professional setting	about soldiers
	about working class
3. Author's Intention	
a. emotional experience	romantic books,
	relaxing,exciting
	horrifying, humorous
	suspense

3. Author's Intention

b. cognition/information	about narcotics
	about animal life
	their life and food
	how to ride a horse
	about dangers and young people
	philosophical, political
	not too many problems
	love story
	preferably philosophical

4. Acessibility

a. readability	easy, not complicated
b. Physical characteristics	
	for second grade
d. main characters	to read aloud for kids
	of 6 to 7 years
	easy to read with big typography
	with pictures
	serials
	thick/small books
	not heavy books
	not with complicated names
	books in English
	fairy tales
	a book about penny
	the front is black with the picture of a pirate and a castle

Figure 3

EXAMPLES OF ADULT USERS FORMULATIONS.

1. subject-matter

  a. action and course of events

expeditions	crime novel
crime detection	historical novel
fight	war novel
invention, discovery	
sex	family chronicle
technical progress	love story
	travelogue

  b. psychological development

childhood and adolescence	social novel
development	philosophical love story
Bildung	mystery novel
religious conflicts	sentimental book
	suspense novel

  c. social relations

groups	science-fiction novel
family	religious novel
society	criticism of society
	family chronicle

2. frame

  a. time

past	psychological novel
present	historical novel
future	thriller
	biographies

  b. geographical/social environment

place	western
profession	historical novel
organization	sea novel
	medical novel

3. Author's intention

a. emotional experience

horror/thrill	love story
amusement	thriller
relaxation	suspense
	travelogue
romance	family chronicle
entertainment	suspense novel
	biography

b. cognition and information

criticism	social realistic novel
agitation, authenticity	philisophical novels

4. accessibility

a. readability

easy	easy to read
medium	short stories in weeklies
difficult	pulp magazines
composition	entertaining books
language	not-too-heavy books

b. physical characteristics

typography	short stories
length/volume	big letters
novel	pulp magazines
short story	new literature
type of edition/time/place	series
first edition	
nationality	
editor	
anthology, series	

Figure 4

Also, the intermediary's perception of the user's need and the reasons for book suggestions were probed. After the last set of data collection, an individual meeting was arranged with the intermediaries who participated in the searches. Thus they had the opportunity to comment on each search, and the investigators were able to discuss the circumstances of each search. The study showed how users tend to characterize book contents from a number of different angles and led to the conclusion that novels need to be classified in a multi-dimensional way. Four main dimensions were identified in users' requests, (see figure 3 and 4). A user's request may concern only one of these dimensions, but fiction enquiries often contain a combination of book features from different dimensions. Needs formulated within the four dimensions of *subject-matter, frame, author's intention and accessibility* could be specified further by users into several categories within each dimension. Each user would formulate his/her need in terms of one or more of these dimensions and categories. Only a few users would express their needs within all the dimensions and categories identified in the data collection.

A piece of user-intermediary negotiations may illustrate this:

User:	Would you recommend a book for me?
Intermediary:	Yes, what would you like to read about?
User:	I don't know. I've just begun to read books.But I like books with action, exciting, not too heavy or sad, and not to many difficult and strange names. Preferably a novel about a family.
Intermediary	Have you read Linnemann's "The House with the Green Tree"? It is two volumes and tells about a merchant's family in Copenhagen and in the West Indies.

The user's need represents several dimensions: "Action" and "family", refer to the *subject-matter* dimension, exciting and sad refer to the *author's intention* of conveying *emotional experience*, "heavy" to the author's intention of *cognition/information* and the *readability* dimension, "difficult and strange names" to the *accessibility* dimension. The intermediary's document characterization includes the *subject-matter* dimension (family) and the *frame* dimension of *place* (Copenhagen and the West Indies) as well as of *social setting:* profession (merchants). Generally speaking, a user will expect to have his reading need satisfied by an *emotional experience* (by an exciting or sentimental book), or he will expect his reading need to be satisfied by *cognition and information* (about human and social conditions, or

finally a combination of both). The user's expectations of an emotional experience and cognition/information are correlated with the *subject-matter* and the *frame* dimension. A further decisive factor for a successful reading experience is the *accessibility* of the book, with regard to the level of communication employed by the author. The majority of users' formulations about the author's intention belongs to the class "emotional experience", which presumably is due to the most common perception of the function of fictive literature as opposed to fact literature. Secondly, it may be due to difficulty in formulating needs about more abstract qualities like the ideology and philosophy in documents. Instead the user probably utilizes earlier experiences with correlation between emotional experiences/subjects and the author's intention, when selecting a subject or an emotional experience. Since all four aspects are of importance to the user - at different levels - a classification scheme must comprise all the relevant dimensions (see figure 5).

The main difference between the adults' search behaviour and childrens' search behaviour is found in the *accessibility* dimension. Not surprisingly, the accessibility of documents is by far the most frequently used value criterion with 111 formulations per 100 conversations. In addition, this dimension contains criteria that do not often occur in the data about adults' accessibility perception. The readability requests are for instance differentiated with respect to language and content relative to school grades and age groups ("for 1st grade, for big children, for small children, for 6 years old etc."). Readability formulations also include difficulty of language, of content and of literary form relative to the social use of texts: whether they are read by a child himself/herself or read aloud by a grown up. Also, requests for serials, information about serials, and enquiries for something similar play a more important role. Another point of difference is the children's demand for a specific sex, i.e. male or female characters in the book. Also, a request for pictures is a frequent feature of the literary form formulations that does not occur often in adults' formulations of needs.

The users' perception of "bibliographical" data of relevance to the retrieval of "known" items includes not only traditional bibliographical data such as authors' names, titles and year of edition (modern, old). It also includes the physical appearance of documents, such as colour and pictures on cover or front page (this occurs in 13% of the conversations), but also name, age and other characteristics of

main characters in a story were used as bibliographical data of a known item. In the recorded user-intermediary negotiations children used the cover of documents in two different ways:

1. Children described the cover of a specific item known or seen by them to the intermediary as a piece of "bibliographical" information in a search for this particular item. For example: "On the front of the book there is a boy and a rhinoceros, and it is red".

2. Children perceived the book covers of documents as a means of communicating their contents to the reader and therefore evaluated the book contents on the basis of information from the cover (for example: "This looks good" or "it means a lot what the front page looks like").

In our collection of real life records, children communicate a denotative perception of the phenomena on the cover when they search for a known item. Thus their memory and identification of the document is based on references to concrete things and colours like "robber", "castle", "black", etc. No relationship between phenomena is referred to, and no situational perception of actions occurring on the cover is expressed. When children evaluate books by the cover, they are described by the connotative meaning of their message in terms of "summary" value judgements, like the example above, but without any reference to specific features of the cover. Hence, value judgements of book covers do not seem relevant in bibliographic searches for known items.

In the post-negotiation interviews with the children, their replies to the question about the cause of their inquiry led to the following observation: the most predominant pattern is a search motivated by a need for emotional experiences in relation to leisure activities. A second predominant pattern is a search motivated by a need for cognition/information either related to a general interest as part of leisure activities - or related to school problems. In both cases the level of books plays an important role either in relation to adults' "read aloud" or children's "read alone" activities. The source for the children's questions are recommendations from friends, from school or the school library and from TV, enumerated according to significance, (Pejtersen, 1986).

The AMP Classification Scheme for Fiction

DIMENSION 1: Subject matter

    a. action and course of events

    b. psychlogical description

    c. social relations

The subject content of the novel: What the story is about.

DIMENSION 2: Frame

    a. time: past, present, future

    b. place: geographical, social environment, profession

The setting in time and place chosen by the author as the scenario of his work.

DIMENSION 3: Author's intention

    a. emotional experience

    b. cognition and information

The author's attitude towards the subject. The set of ideas and emotions which the author wants to communicate to his readers.

DIMENSION 4: Accessibility

    a. readability

    b. physical characterstics

    c. literary form

    e. main characters

    f. age of main characters

The level of communication in terms of those properties which facilitate or inhibit communication, such as difficulty of contents and language, composition, typography etc.

Bibliographical data:

Title, author, illustrator, translator, editor, publisher, pages, year of 1st edition, year of classified edition, cover of book, serial, illustrated, screen version, title country and year of original edition, etc.

Figure 5

Characteristics of the AMP classification scheme.

The classification scheme for Analysis and Mediation of Publications (the AMP classification scheme, figure 5) therefore has four independent, facet-like main classes called "dimensions". Each dimension characterizes fiction according to its own set of criteria, and is subdivided into a few broad categories.

For the sake of convenience and consistency, dimensions and categories are usually listed in the order shown above. The dimensions are not mutually exclusively but supplement each other since all characteristics may be found in the same book but with different dominance when a novel is classified, it it simultaneously placed in every dimension and category in which it would be reasonable to find it, depending on the weight put upon its different aspects by its author. All the sub-classes within the dimensions are not exclusive. This multidimensional approach means that the classification cannot be used for shelf order.

Because fiction collections and their user communities vary a great deal, no attempt has been made to provide a priori subdivisions below the level of categories. Agencies using the system should individually develop subclasses to suit their particular needs. Alternativly, the system can be used without adding any further subdivisions, especially when applied to very small collections.

There is no logical order in the established classes; one class is not logically dependent on the other classes. There is no hierarchy of classes. The classes are co-ordinate and do not comprise uniform subjects, nor do they belong to one definable group. This makes it easier to meet the user's needs which are not centred around one single aspect of the book. The one-sided stressing of the hierarchical relation found in the traditional classification is set aside so as to take advantage of a combination of some of the functions and aspects of fiction. The searcher can enter the system anywhere without first having to place the subject in some super-ordinate class and then go on searching for a subordinate class. On the contrary, any relation between designations and subjects may be expressed. As already mentioned, the classification system aims at being operational in daily routine. Even with a small number of dimensions and sub-divisions in the four dimensions, the resolution of the scheme will be high due to the number of possible combinations. The reader might argue, quite rightly, that it is incorrect to call this system a "classification". With its non-exclusive, multiple-place structure, and the lack of a notation , it certainly differs a good deal from traditional library classifications. It would also be just as incorrect

to use the label "indexing system" in its traditional sense, since book features are not listed alphabetically but organized into meaningful groups. Perhaps the system should be seen as a "hybrid": principles from traditional classification and indexing are combined in an attempt to find new ways of organizing fiction for retrieval by content. An appropriate name for this new scheme does not readily come to mind; until it does, or until someone offers a more useful suggestion, both classification and indexing can be used as an appropiate label.

Conclusion

The AMP classification scheme has been developed further into a number of rules for its application to all kinds of novels. A number of small test databases have been evaluated and experiments with different kinds of retrieval tests and user groups have been conducted. A database with classified poems has been developed as well based on the same principles. These experiments all proved that the AMP classification system is a very useful tool, both for analysis and representation of book contents as well as a tool for control of the user's search process. Based on these preliminary results a library prototype system called the "BOOK HOUSE" was designed