Paper for the 6th ASIS Classification Research Workshop, to be held at the 58th Annual Meeting of ASIS in Chicago, Illinois.

# The Scandinavian Book House:
## Indexing Methods and OPAC Development for Subject Access to Scandinavian Fiction Literature

**Annelise Mark Pejtersen, Hanne Albrechtsen**
**Ringa Sandelin, Lena Lundgren & Riitta Valtonen**

## INTRODUCTION

While potentials for information searching in Scandinavian libraries are improving due to the rapid technological developments, the librarians and library users in Scandinavia are still facing the problem of limited access to fiction literature. Even though the interest in providing subject access to fiction literature has a long history, and various classification schemes for shelf arrangement have been proposed, the mediation of fiction by Scandinavian libraries has been constrained due to i.a. i) the hitherto often underprivileged status of fiction compared with non-fiction as sources to knowledge, and related to this: ii) the very few empirical studies of users' needs and goals for fiction searching, and iii) the special problems related to subject analysis of fiction.

During the 1970s and 1980s, Pejtersen (1979; 1980; 1992b; 1992c; Pejtersen & Austin, 1984, 1985) conducted several field studies of the mediation, retrieval and use of fiction in Danish public libraries, resulting in a) the multi-dimensional AMP scheme for fiction, b) a set of indexing rules for fiction, and c) a supportive retrieval tool, The Book House, whose icon-based interface visualizes both the classification scheme, the work space of the fiction databases, including a choice between different search strategies, and an associative term network for browsing in fiction concepts. Recently, a new version of the Book House has been developed for MacIntosh which in addition to a search tool provides the indexer with a cataloguing and indexing tool, whose interface visualizes the workspace of the indexer. (Book House Search and Book House Write for both children and adults, which is now available as a shareware program; Pejtersen 1989, 1992b, 1993, 1994a, 1994b)

This paper reports on the work done by the Scandinavian Book House project (Pejtersen, Albrechtsen, Lundgren, Sandelin & Valtonen, 1995). The Scandinavian Book House Project, which is supported by grants from the Nordic Council of Ministers, attempts to clarify and test the possibilities of a common approach for subject access to Scandinavian fiction. The project places itself within a firm tradition for cultural collaboration, due to the Scandinavian tradition for the sharing of cultural values, based on i.a. shared views of enlightment and democracy.[1] Further, it places itself within a tradition for joint research and professional networking between library associations, library schools and public libraries across Scandinavia. The project has developed from an informal networking between researchers, librarians and library consultants in Denmark, Sweden and Finland, sharing a common interest in increasing the accessability of Scandinavian fiction by identifying the readers' requests and needs in this area.

The ultimate goal for The Scandinavian Book House Project is to contribute to an increased mediation of fiction in Scandinavian libraries by i) establishing common indexing principles for Scandinavian fiction and ii) further development of The Book House to function as a common Scandinavian OPAC for fiction retrieval. In order to reach this goal, the project has elicited requirements to indexing methods and OPAC development by:

i)      Gathering knowledge of the ongoing approaches to fiction indexing in Scandinavia from indexing experts in national library services and from librarians engaged in local indexing projects for fiction, and disseminating and discussing ideas with Scandinavian libraries experts and thus identify points of convergence/ conflict in approaches, and suggest improvements of the hitherto shallow indexing methods applied for fiction;

ii)     Suggesting a new approach for multi-dimensional subject access to fiction, realized in The Book House OPAC, which can be developed to function as a multi-cultural subject access approach;

iii)    Empirical evaluation of the Book House as a common Scandinavian OPAC by conducting pilot projects on indexing and retrieval of fiction in Scandinavian public and school libraries, using new the Book House version in MacIntosh as a testbed for indexing experiments.

## ONGOING DEVELOPMENTS IN SUBJECT ACCESS TO SCANDINAVIAN FICTION

Investigations of the historical development of classification schemes and systematic shelf arrangement for fiction have shown that the early classification systems were difficult to use, because they reflected normative, and even paternalistic approaches, due to the personal sets of values of their creators rather than the multi-dimensional values of library users and the needs of librarians for mediating information (eg.Pejtersen, 1980, 1986b). As a consequence, subject access to fiction in Scandinavia has until recently merely been provided by access to call numbers, reflecting primarily alphabetical shelf arrangements of fictional works.

The AMP-scheme was developed to i) circumvent the practical problems and ideological bias, identified in the early classification schemes by providing a framework reflecting the users' needs and ii) to act as a subject analysis model for fiction indexing. This indexing method has acted as the major inspiration for recent developments in subject access to fiction in Scandinavia, in particular in Denmark and Finland. On the whole, the national library services in Norway, Sweden, Finland and Denmark are giving verbal subject access in the national databases a high priority in their provision of services to users of the national online catalogues.

### Fiction Indexing

Today, special attention is being devoted by Scandinavian library services to developing better indexing methods, including special term lists and subject analysis models for fiction. Term lists are usually alphabetical indexing terms for fiction, defined by the indexers (cf. eg. Skaarup, 1995; Hjortsæter, 1990); one exception is the Swedish term list of fiction concepts for children, which includes terms compiled from children's requests to librarians (Lundgren, 1994; Jansson, 1994).

In Denmark, a special model for subject analysis of fiction has been developed, which is inspired by the dimensions of the AMP-scheme.

The ongoing national indexing projects for fiction in Scandinavia show great promise for creating better subject access tools for library users. However, there is a great diversity among them with respect to indexing depth, time coverage, as well as varieties in terminology, degrees of pre-coordination of indexing terms and methods for subject analysis. Added to this are linguistic varieties in terminology. In order to overcome the conceptual, cultural and linguistic barriers imposed on fiction searching across national online catalogues, a cooperation on indexing methods and thesaurus compilation for fiction is needed.

In addition, local projects in fiction indexing are conducted by public libraries in Scandinavia, for instance in the public library of Stockholm, the library at the Peter Lykke school in Copenhagen, Denmark, and in public and school libraries in Ekenäs, Finland. Contrary to the indexing approaches taken by the national library services, these projects rely on first-hand information of fictional works. The first-mentioned project in Stockholm is however constrained by focussing on explicit topics in the books, whereas the projects at Ekenäs and Copenhagen apply the indexing rules developed for The Book House, which means that they address dimensions such as intention and readability.

## Thesaurus Development

One very promising approach, supporting users' browsing in fiction concepts as well as bilingual searching, is the special thesaurus for fiction, recently compiled by Kirjastopalvely = Biblioteksservice, Finland. In Finland there exists a Swedish-speaking minority having the same linguistic rights as the Finnish-speaking majority. Cooperation between the two groups among other things manifests itself in the endeavour to make information searching bilingual, i.e. the subject terms are reciprocally connected and users are thus able to use their native language and to achieve titles indexed originally in Swedish or Finnish.

In 1993 a decision was made at Kirjastopalvelu = Biblioteksservice in Finland to compile a special thesaurus for fiction. The thesaurus will be published in both Finnish and Swedish. The Finnish-language version will be published in 1995 (Kaunokki, 1995). The Swedish-language version will appear later in the year, and together the two thesauri will support bi-lingual subject access to fiction in the Finnish national catalogue, FENNICA. The layout of the thesaurus of fiction indexing terms features a systematic section, where the terms are organized into facets such as genre, theme, actor, environment and time, and it has an alphabetical index to the systematic section. The thesaurus uses narrow, broader and related terms. The facets of the thesaurus reflect the dimensions of the AMP-scheme for fiction.

## OPAC Development

A majority of automated library systems in Scandinavian libraries feature special 'add-on' OPACs, or clients, for end-user searching (Larsen, 1994; Hyldegaard, Sinding & Albrechtsen, 1995). These OPACs typically offer query by form, menu-based searching or support for command-driven searching. In addition, some Scandinavian library services, for instance in Denmark, plan

to implement advanced OPAC facilities, such as automatic text processing of search statements and near-hit searching to support end-user searching in their gateway OPAC services, thus transferring the research done at City University London, in a real search environment (Petersen, 1995; Hancock-Beaulieu, 1990). Graphical user interfaces, featuring iconic representations of system functionalities and generic search objects, such as documents, are entering the market.

Even though the OPACs in Scandinavian library systems and gateway services are intended to support end-user searching, their development is driven largely by the professional searchers' requirements. Further, support for browsing in the semantics of the databases does not address any possible diversity of concrete concept structures, such as different vocabularies, provided by the data producers.

## END USERS' PROBLEMS WITH ONLINE CATALOGUES

The history of OPACs reflects a development from a crude implementation of the old manual card catalogue concept for computerized tools with precoordinated searches and very little information for subject access, to more directly developed systems based on a more sophisticated concept, including improvement along primarily two lines: i) Enrichments of bibliographical records both for free text and controlled subject access and ii) a flexible interface allowing both command and menu searches as well as input in natural language. In addition, they often contain spell checking mechanisms and a possibility to evaluate the retrieval results. These so-called third generation systems (Hildreth, 1990) are all heavily text based. Studies reveal the following major problems and inadequacies in current OPAC systems, which need to be investigated and improved (Hildreth, 1989). These are:

1. Navigational confusion and frustration for the user during the search process caused by bad interfaces and unsupportive user-system dialogues.

2. The vocabularies in bibliographic records are incompatible with the language brought to the system by their users. Users fail to match their own need formation with the system's subject vocabulary.

3. Lacking understanding and use of the system facilities. Partial use of the options provided by the system and missing opportunities to complete initiated searches.

There is a general recognition that the design of third generation OPACs should guided by what can be learned from user behaviour and searching strategies employed in shelf browsing as well as in the traditional catalogue. A variety of subject searching capabilities to match various information seeking needs and users' search objectives, including shelf browsing, and automation of complicated search routines, are needed (Markey, 1989). It has been claimed that improved on-line facilities should be incorporated to support interactive discovery and user-system interaction, as well as negotiation of better expressions of information needs (Hancock-Beaulieu, 1990). These, and many similar suggestions, point towards a need for a general conceptual model for OPAC design and development that integrates these suggestions for design improvements.

# A NEW APPROACH TO MULTIDIMENSIONAL SUBJECT ACCESS IN OPACs: THE BOOK HOUSE

To circumvent these problems, design is to be based on the concept of an adequate 'resource space' instead of support of one particular, normative procedure for retrieval interactions. A high degree of compatibility must then exist between the system and user. Compatibility cannot be secured by intelligent interfaces or conversion programs. OPAC systems will only be really useful and widespread, when i) the domain and task characteristics provide the basis for the addition of supplementary information to the existing description of book content in card catalogues, which then matches end users' intentions and needs and ii) when the user interface to the databases is configured as an integrated and uniform set of displays matching user capabilities and limitations.

## A Multidimensional Framework for OPAC Design

The Book House represents a new approach to OPAC design, where the creation of classification schemes and conceptual structures, indexing rules and the icon-based interface constitutes an integrated development, based on analysis of the different dimensions or domains involved: i) the author's domain, ii) the user domain, iii) the library domain and iv) the system domain. Figure 1 shows the framework, which maps the multiple domains involved in the design of a supportive OPAC for fiction and i.a. acts as a metamodel to elicit requirements for multidimensional classification schemes, which thus includes a three dimensional mapping of domain criteria (for users, library and author) with respect to required information about classification of book content.

## A Fiction Classification Scheme

Figure 2 shows a multidimensional classification scheme for fiction, based on empirical analysis of users' (children and adults) requests and behaviour during information retrieval involving a series of user-librarian conversations. The study showed how users tend to characterize book contents from a number of different angles and led to the conclusion that fiction need to be classified in a multidimensional way. The classification scheme has four independent, facet-like main categories called "dimensions" as shown in figure 2. They reflect closely the levels of abstraction of the representation of users' needs and book contents as shown in figure 1. These range from reference to the authors' goals, their value criteria, their choice of book content and their level of communication with the reader to the publisher's choice of physical representation of the document. Note, that the order of the dimensions within the classification scheme reflects the order of importance as identified in laboratory tests to be most effective in support of the indexing task. Consequently, it is different from the order defined by the conceptual level of abstraction (figure 1). Also, the order of presentation in the interface may be different, being chosen from studies of the natural sequence of queries of the average user of a particular category of users as identified in field studies. (In the children's database, the appropriate reading age will be at the top of the presentation).

This classification scheme differs significantly from the exclusive classification schemes identifying a particular document with one and only one location in the formal system. In the present classification system, a document is classified with reference to its location at all five dimensions

and, therefore, a high resolution in the classification can be obtained by only a few categories within each dimension due to the large number of possible combinations. Each dimension characterizes fiction literature according to its own set of criteria and is subdivided into a few broad categories which reflect the structure embedded in the users' needs and made explicit in their queries. 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.

The dimensions and categories are not all mutually exclusive, but supplement each other since all characteristics may be found in the same book but with different dominance. When a book is classified, it is simultaneously placed in every dimension and category in which it would be reasonable to find, depending on the weight placed on its different aspects by its author.
There is no consistent logical order in the established categories according to generic and part-whole principles; e.g., one class is not logically dependent on the other classes in a tree structured hierarchical order and classes may overlap. The dimensions and classes are coordinate and do not comprise uniform subjects, nor do they belong to one well-defined group. This makes it easier to meet user's needs which often are multidimensional and thus are not centered around one single dimension of the book.

## Book House Indexing Language

The Book House does not have a full text database, but offers a short representation of the document content based on the classification scheme and a short annotation of content containing keywords for retrieval within each category of the classification dimension. Development of an indexing language was based on a combination of international, linguistic thesaurus standards and the pragmatic aspects of the domain language employed by the user group - as an iteration among top-down and bottom-up approaches to the establishment of rules to achieve a consistent practice in choice of vocabulary. It was influenced by the recordings of users' actual choice of language, domain concepts and vocabulary, when they were evaluating document content, performing relevance feed back and formulating and revising a need etc. The choice of terminology and indexing vocabulary was to be comprehensible to the different age level of users being both children and adults according to the assumption that users were primarily lay people.

Special attention was paid to the level of exhaustivity of the analysis of books, as user requirements varied within each dimension. For each dimension it was necessary to develop special requirements to exhaustivity of the analysis and description of content. The exhaustivity of the annotation in each record was also a balance between the trade off of providing sufficient information within all dimensions about the content for the user's decision about relevance - without revealing too much about the course of events in a book and, in particular, its ending. The indexing therefore included only the most dominant and characteristic features of a book and not minor aspects.

Another concern was the specificity of the language needed to represent book contents in a language level compatible with users' formulation of their needs and queries. In a number of laboratory experiments users' and librarians found a specific indexing language favorable for recognition of needs, that they had not been able formulate in specific terms. It was then decided

to index with a specificity corresponding to the level chosen by the author of the book. Generally, this resulted in a more specific level than that expressed in users' queries in field studies. This was done to communicate to the reader as precisely as possible the level chosen by the author for writing his/her story for comparison of match with need- and to support recognition of more specific concepts to be searched by users. It was therefore a deliberate policy to employ a general vocabulary meeting most frequent user queries as well as a specific vocabulary meeting authors' levels of writing in order to meet any level of language and concepts chosen by potential users. At the same time, this would enrich the vocabulary of search and access terms. To retain some control of indexer's choice of a consistent vocabulary in keywords and book annotations, a thesaurus was developed by a 'bottom up' method, empirically using terms employed in document descriptions as indexing progressed, and building a controlled vocabulary of hierarchical and associative relationships.

## Subject Indexing of Fiction: Analysis and Rules

Special indexing rules were developed from field studies of users' relevance judgements of documents within the same subject matter content. It was found that three invariant features influenced the users' judgements: i) course of events, ii) psychological aspects of characters and iii) the relationships among characters and course of events. The identification of these features lead to the development of three different viewpoints to be considered by the fiction indexer. These are: (1) The point of view of the main character(s) (2) The main character(s) and the event(s) taking place (3) The relationship between the main character(s) and event(s) taking place. Linguistic theories of discourse analysis and narrative structures in fiction (thematic roles, text composition, eg. Greimas, 1983; Rummelhart, 1977) provided the theoretical framework to develop models for subject analysis of fiction literature (Pejtersen and Austin 1985; Albrechtsen 1993). The result was a set of standard structures informing about the pragmatic function of a fiction work, which are described in detail elsewhere (Pejtersen, 1992a, 1993, 1994; Pejtersena, Albrechtsen, Lundgren, Sandelin & Valtonen, 1995).

## EVALUATION OF THE BOOK HOUSE IN SCANDINAVIA

The Book House was evaluated by users, i.e. librarians and end-users, in public libraries and school libraries in Denmark and Finland, and special emphasis was given to validating the functionality of the multi-dimensional classification scheme for building up queries and the usability of the specific indexing rules for fiction. In order to experiment with the latter aspect, the new MacIntosh version of The Book House was applied, available in two Scandinavian languages (Danish and Swedish).
The users found the multi-dimensional classification scheme very helpful to distinguish between different aspects of information contents in fiction, and it helped them to formulate their needs more precisely than when searching in conventional OPACs. Furthermore, the classification scheme helped the users in the iterative task of building up queries as search sessions developed.

Furthermore, the indexing rules applied for fiction resulted in many subject access points to each fictional work, which made it possible to retrieve books from many different angles and viewpoints during a search session. The users who experimented with the new cataloguing and indexing tool found it highly motivating for the indexing task and easy to use.

On the whole, the empirical evaluation of the Book House in Scandinavia has shown a great satisfaction with the multi-dimensional subject access to fiction. Furthermore, it was found that icon-based searching was preferred to command-driven text-based searching; the associative term network for the subjects of the databases and represented as icons, was very inspiring for gaining a broad view of the contents of the database in addition to finding the right access points during subject searching.

In addition to these empirical evaluations, the new version of the Book House has been demonstrated and discussed with Scandinavian libraries experts, representing both national library services and data producers (reported in: Albrechtsen, 1995). Their comments and advice have elicited a number of new requirements to the Book House, which in particular concern a need for thesaurus delevopments and scenarios for exchanging subject data for fiction. There is a consensus among these experts that the research done by the Scandinavian Book House project in OPAC development and subject indexing of fiction is highly important for a further development of the hitherto provision of subject data. However, the special rules for fiction, developed by Pejtersen (eg. Pejtersen, 1992a, 1993, 1994b) will require additional training of indexers, and their application in national online catalogues will require a harmonized extension of the national cataloguing formats with fields reflecting the different dimensions.

## DISCUSSION

Subject access in Information Retrieval (IR) systems is more than search keys and online user assistance. The problems of subject access to fiction for end users cannot merely be overcome by increasing the number of search keys in online catalogues, nor by offering support in OPACs in the form of menus, browsing facilities, transparent Boolean retrieval etc. The design of classification schemes and indexing methods should be integrated in the design of IR-systems. The multidimensional character of users' perceptions of fictional works, the different cultural/ societal values held by users and mediators of fiction, the difference in their needs and goals, and variety of authors' goals within literary schools and professional paradigms call for an integrated design philosophy for IR-systems, based on analysis of such different collective knowledge interests, or work domains, which is coupled dialogically to analysis of subjective characteristics such as authors' individual writing style, and users' domain knowledge and intuitive search strategies. These requirements for subject access have been met by the development of The Book House system. Furthermore, special considerations have to be made for subject access to fiction across national and linguistic borders, which is a special case presented for the Scandinavian Book House Project.

### Fiction Indexing in Scandinavia

From an organizational perspective, the indexing projects for fiction initiated by the national library services constitute top-down drives in that their primary aim is to provide central, national services for fiction indexing. At the same time, the compilation of special thesauri for fiction in by Kirjastopalvelu=Biblioteksservice, Finland, can be seen as applied research, departing from the basic research done by Pejtersen in the development of The Book House. Conversely, from the perspective of user requirements for better access to and mediation of fiction, the experiments in fiction indexing at the libraries of the Peter Lykke School, Denmark, Ekenäs, Finland, and

Stockholm, Sweden, can be seen as reflecting a bottom-up drive in the development of fiction indexing in each Scandinavian country, initiated to fulfill the needs of their prospected users, librarians and library users.

In other words, the current developments in fiction indexing in Scandinavia reveal a rather complex picture when it comes to transferring knowledge of how to provide a better subject access to fiction, which calls for a more explicit formulation of means and goals in order to reach an aim of a common Scandinavian methodology in this area. In particular, it will be necessary to provide a uniform subject analysis model, such as the model developed for the Book House, in order to ensure a common depth in the Scandinavian indexing approaches for fiction.

## Thesaurus Development in Scandinavia

The current provision of indexing terms for fiction has hitherto primarily adressed the development of models for subject indexing and compilation of term lists, supporting the indexers' task. In order to further support this task, and in order to support the searchers' task of navigating fiction concepts in the databases, it will be necessary to develop thesauri for multi-lingual retrieval. The compilation of the Finnish-language thesaurus for fiction, which will be followed by a Swedish language version, is one important initiative in this direction.

However, further work is needed in this area in order to reach the goal of common access to fiction in Scandinavia. Even though most languages in Scandinavia share common linguistic roots, and a high degree commonality between individual states and regions with respect to cultural values exists, it would be naive to insist on a complete Scandinavian harmonization of fiction concepts and structures. Recent experiences in bi-lingual thesaurus development in Canada and Belgium (cf. eg. Hudon, 1992, and Goossens, 1991) have shown that the conceptual domains covered by each language will vary, i.a. that concepts will not always overlap, and that conceptual structures developed in one monolingual thesaurus cannot be translated directly to a thesaurus in another language, even within regions sharing common cultural and social values. Therefore, it will be necessary to develop thesauri for fiction in Sweden, Norway and Denmark as well, which can later be functionally integrated with the Finnish thesaurus in order to support multi-lingual fiction searching in Scandinavia in the future.

## OPAC Development

OPAC development is currently being emphasized in Scandinavia as important for increasing the library users' access to online catalogues. At present, however, the concrete OPAC facilities provided for library automation systems and gateway facilities in Scandinavia show great diversity, and with the exception of the Book House OPAC, they can be characterized as 'add-on' OPACs, which do not assist the users in access to the semantics of the databases, nor support navigation in concepts and choice of adequate search strategies. A new model for OPAC design is needed in order to ensure a better use and exploitation of fiction resources in Scandinavian libraries. The design philosophy underlying the development of the Book House system, has been evaluated by libraries experts in Scandinavia, and by Scandinavian librarians and users, as a very important contribution to realizing the goal of a common Scandinavian OPAC.

## CONCLUSIONS AND FUTURE WORK

The possibilities to support a better mediation of fiction in Scandinavia are growing with the increased provision of library networking and library automation and with the ongoing projects on fiction indexing. This paper has discussed how some of the problems related to the hitherto barriers for an increased mediation can be overcome. The results of the work done by The Scandinavian Book House projects have shown that it is feasible to use common indexing principles and to further develop The Book House as a common Scandinavian OPAC. In order to reach the goal of a common access to Scandinavian fiction, the project will address the issue of how to support cross-cultural communication in OPACs for fiction by building testbeds for Scandinavian online catalogues for fiction in Denmark, Norway, Sweden and Finland, thereby investigating the feasibility of reusing subject data between Scandinavian providers of bibliographic information, and experiment with bottom-up construction of thesauri in the Scandinavian languages, which can be integrated to function as multi-lingual and multi-cultural searching and indexing tools in a Scandinavian OPAC for fiction.

**Notes**

1.The Nordic Council of Ministers, formed in 1971, is the organization for co-operation between between the five sovereign states Denmark, Finland, Iceland, Norway and Sweden, and the three autonomous territories Greenland, The Faroe Islands and Åland. These countries and regions share a cultural heritage over more than a thousand years. With the exception of Finnish, Sami and Greenlandic, the Scandinavian languages can be traced to common linguistic roots. The Greenlanders and the Faroese speak Danish, and some Finns speak Swedish as well. (cf. eg. Tivéus & Emborg, 1995).

## References

Albrechtsen, Hanne (1995). Det Nordiske Boghus - rapport fra seminar afholdt på Danmarks Biblioteksskole. Rapport udarbejdet til NORDBOG/NORDFOLK. Copenhagen: Danmarks Biblioteksskole. [In Danish: "The Scandinavian Book House - a Report from a Seminar held at the Royal School of Librarianship, Copenhagen"]

Albrechtsen, Hanne (1993). Subject Analysis and Indexing: from Automated Indexing to Domain Analysis. **The Indexer** 18 (4), 219-24.

Goossens, P. (1991). Multilingual Bibliographic Access via Subject In Europe. **Library Automation and Networking - New Tools for a New Identity.** Proceedings of the European Conference, 9-11 May 1990, Brussels. München: K.G.Saur, pp. 206-214.

Greimas, Algirdas (1983). **Structural Semantics: An Attempt at a Method.** Nebraska: University of Nebraska. - lvi, 325p.

Hancock-Beaulieu, Micheline (1990).Evaluating the Impact of an On-line Public Catalogue on Subject Searching Behavior at the Catalogue and the Shelves. **Journal of Documentation** 46 (4), pp. 318-338.

Hildreth, Charles (1990). End Users and Structured Searching of Online Catalogues: Recent Research Findings. In: **Tools for Knowledge Organisation and the Human Interface. Proceedings of the 1st International ISKO Conference, Darmstadt, 14-17 August 1990.** (Part 2) /edited by Robert Fugmann. Frankfurt: Indeks Verlag, pp. 9-24.

Hildreth, Charles, ed. (1989). **The Online Catalogue: Development and Directions.** London: The Library Association.

Hjortsæter, Ellen (1990). **Emneordskatalogisering: Forslag til norsk standard.** Oslo: Bibliotekscentralen. 92 p. [In Norwegian: Verbal Subject Indexing - a Proposal for a Norwegian Standard]

Hudon, Michele (1992). Term Definitions in Subject Thesauri: the Canadian Literacy Thesaurus Experience. **Classification Research for Knowledge Representation and Organization: Proceedings of the 5th International Study Conference on Classification Research, Toronto, Canada, June 24-28 1991** /edited by Nancy J. Williamson & Michele Hudon. Amsterdam: Elsevier, pp. 255-262.

Hyldegaard, Jette; Sinding, Elisabet; Albrechtsen, Hanne (1995). **Beyond the Interface - a Theoretical Framework for User Interface Design and Evaluation in LIS.** Paper presented at the 1st British-Nordic LIS-Conference for Researchers and Educators, at the Royal School of Librarianship, Copenhagen, 22 - 24 May 1995. - 10 p. (mimeo)

Jansson, Eiler (1994). **Ämnesord - kompass i bogskogen.** SAB: the Committee for Children and Young People. - mimeo [In Swedish: Indexing terms - a Compass in the Book Forest].

**Kaunokki - Kaunokirjallisuuden Asiasanasto** (1995) /edited by Jarmo Saarti et al. - Helsinki: Kirjastopalvelu. [In Finnish: Special Thesaurus for Fiction Literature - Finnish Language version]

Larsen, Gitte (1994). **Advances in OPACs in Europe: An Overview.** Publications of Essen University Library 17, pp. 237-252.

Lundgren, Lena (1994). **Barn bibliotek datorer. Rapport från ett projekt.** Stockholm: Länsbiblioteket i Stockholms län. 107 p. [In Swedish: Children, Library, Computers - a Report from a Project]

Markey, Karen (1989). Alphabetical Searching in an Online Catalog. **Journal of Academic Librarianship** (6), pp. 353-360.

Pejtersen, Annelise Mark; Albrechtsen, Hanne; Lundgren, Lena; Sandelin, Ringa & Riitta Valtonen (1995). **Subject Access to Scandinavian Fiction Literature: Indexing Methods and OPAC Development.** Copenhagen: The Royal School of Librarianship. - 80 p. (Report compiled for the Nordic Council of Ministers; available from: Hanne Albrechtsen, The Royal School of Librarianship, Copenhagen)

Pejtersen, Annelise Mark (1994a). A Framework for Indexing and Representation of Information Based on Work Domain Analysis: A Fiction Classification Example. in: **Knowledge Organization and Quality Management** /edited by Hanne Albrechtsen & Susanne Ørnager. Frankfurt: Indeks Verlag, pp. 251-263.

Pejtersen, Annelise Mark (1994b). A New Approach to Design of Document Retrieval and Indexing Systems for OPAC Users. in: **Proceedings of the 17th International Online Information Meeting, London 7-9 December 1993.**

Pejtersen, Annelise Mark (1993). The Role of Domain, Task and Users in Providing Subject Access to Information. Keynote paper in: **Cataloguing Australia** 19 (3/4), pp. 85-116.

Pejtersen, Annelise Mark (1992a). **Emneindeksering af skønlitteratur: Analyse og Regler.** Roskilde: Risø National Laboratory. - vi,108 p. (Available from the author at Risø National Laboratory [In Danish: Subject Indexing of Fiction Literature: Analysis and Rules].

Pejtersen, Annelise Mark (1992b). The Book House. An Icon Based Database System for Fiction Retrieval in Public Libraries. in: **The Marketing of Library and Information Services, 2** / edited by Blaise Cronin. London: ASLIB, pp. 572-591.

Pejtersen, Annelise Mark (1992c). New Model for Multimedia Interfaces to Online Public Access Catalogues. in: **The Electronic Library, The International Journal for Minicomputer, Microcomputer and Software Applications in Libraries** 10(6).

Pejtersen, Annelise Mark (1986a). **Rules for Fiction Indexing.** Copenhagen: The Royal School of Librarianship. - 56 p. (Internal report)

Pejtersen, Annelise Mark (1986b). Implications of Users' Value Perception for the Design of a Bibliographic Retrieval System. in: **Empirical Foundations of Information and Software Science** / edited by J.C. Agrawal and P. Zunde. New York & London: Plenum Press, pp. 23-39.

Pejtersen, Annelise Mark & Jutta M. Austin (1985). **Concept Identification and Subject Control in Fiction Indexing.** Copenhagen: The Royal School of Librarianship. - 66 p. (Internal report)

Pejtersen, Annelise Mark & Jutta M. Austin (1984). Fiction Retrieval: Experimental Design and Evaluation of a Search System based on Users' Value Criteria. Parts I-II. **Journal of Documentation** 39 (4), pp. 230-246, and 40 (1), pp. 25-35.

Pejtersen, Annelise Mark (1980). Design of a Classification Scheme for Fiction Based on an Analysis of Actual User-Librarian Communication, and Use of the Scheme for Control of

Librarians' Search Strategies. in: **Theory and Application of Information Research** / edited by Ole Harbo & Leif Kajberg. London: Mansell, pp. 167-183.

Pejtersen, Annelise Mark (1979). The Meaning of Aboutness in Fiction Indexing and Retrieval. in: **ASLIB Proceedings** 31, pp. 251-257.

Petersen, Per Mogens (1995). DANBIB udviklingsplan 1995. **danbib-bladet** 1, pp. 4-9. [In Danish: Development Plan for DANBIB, 1995]

Rummelhart, David E. (1977). Understanding and Summarizing Brief Stories. in: **Basic Processes in Reading: Perception and Comprehension.** /edited by David Laberge & S. Jay Samuels. Hillsdale (NJ): Lawrence Erlbaum Associates, pp. 265-303

Skaarup, Peter (1995). **Indexing of Fiction Literature in BASIS**. Dansk BiblioteksCenter (personal communication)

Tivéus, Ylva & Kerstin Emborg, eds. (1995). **Nordic Reflections**. The Nordic Council of Ministers. 88 p.

**Yleinen suomalainen asiasanasto** (1995) / [julk.] Helsingin ylioiston Kirjasto; [toimittajat: Eeva Kärki et al. Helsinki: Kirjastopalvelu. - 499 p. [In Finnish: General thesaurus for Subject Indexing]

## Coupling Users' Domain to Authors' Domain Through Library and IR Systems

| Means/Ends Representation in IR Task Situation | 1.Users' Work Domains → | 3. Library Domain | 4. IR System | 2. Authors' Domain ← |
|---|---|---|---|---|
| **Work Domains** | Goals, constraints. Task situations and functions. Information sources and tools | Goals: Information dissemination, cultural mediation. Constraints: Laws, quality, budget | Boundaries of domains and their means/ends levels and attributes to be co-vered in the database | Goals: Informative, ethical, aesthetic, emotional within pro-fessional paradigm; Social, cultural con-text / subject content, time, place, setting. Form, accessibility. |
| **Task Decisions** | Users' state of know-ledge, information needs, analysis and evaluation of problem space and solutions | Librarians' state of knowledge about the library domain and users' situation and needs | Faceted, multilevel classification scheme. Indexing principles and rules | Author's perception of user's situation and information needs. |
| **Task Language** | Language in users' queries about domain / task situation | Librarians' task lan-guage, knowledge about users' domain language | Language and vocabulary for representation of database information | Document language |
| **Task & Search Strategies** | User strategies: analytical: specific need, well defined means/ends attributes Search by ana-logy/browsing: intuitive need, fuzzy set. | Librarians' empirical shortcuts | Retrieval functiona-lity, search algorithm: Analytical: Boolean best match; analogy / browsing: prob-abilistic, partial match | Author's strategies for communicating goals and intentions |
| **Mental Models** | Semantic networks: Analytical: Complete means/ends representa-tion; Analogy / Browsing: Partial Means/ends Representation | Librarian's empirical models of prototypical document attributes and user stereotypes | Representation of several different semantic networks of classification dimensions and indexing language | Author's perception of subject domain, literary paradigm and preferences |
| **User Characteristics** | Age, education, novices, experts, domain knowledge | Librarians as expert users | Icon & text interface displays for novices, command interface for experts | Author's perception of target reading population |
| **Subjective Preferences** | Domains, language, strategy, interface. | Same | Access to multiple representations of domains, indexing languages, database networks and interfaces | Author's individual style, language, composition. |

**Figure 1** shows the framework and the multiple domains involved in design of an IR system.
Means/ends Domain Representations

| 1. Author's goals and values | |
|---|---|
| Why? | (1) *Author's intention and goals:* Communication of information, education, promotion or stimulation of ideas and emotional experiences or other goals. |
| Why? | (2) *Author's affiliation and attitude:* Subscription to professional paradigms and associations, political and cultural movements, or other value criteria. |
| **2. General and Specific Topical Content in Book** | |
| When? | (3) *General frame/time content:* Time, year, historical period etc. Temporal aspects of the topic. |
| Where? | (4) *General frame/setting content:* Place and setting, geographical, historical, social, class, professional setting etc. as environmental aspects of the topic. |
| What? | (5) *Subject matter content:* Specific topic & plots. Psychological and social phenomena. Ending of book. Events from e.g., natural or social science |
| Who? | (6) *Living beings, institutions and artifacts:* Main characters, persons, animals, plants, institutions etc., that are involved in the topic, and their important attributes etc.. |
| **3. Communication and Presentation of book format** | |
| How? | (7) *Literary form:* Fact, fiction, genre, lexicon etc. |
| How? | (8) *Accessibility level:* Readability of language, content and form etc. |
| How? | (9) *Physical characteristics:* Color and picture of front cover of books, Printing characteristics, book format, art of pictures, maps, cards, photos etc. |
| **4. Publication attributes** | |
| Who? | (10) *bibliographical data and identification numbers:* Author, title, translator, illustrator, publisher, etc. |
| Where? | Place of edition, etc. |
| When? | Year of edition, etc. |
| What? | Type of document, ISBN number etc. |
| How? | Size and format etc. |
| **5. Library attributes** | |
| Why? Where? | (11) *Library goals and policies:* (12) *Local library conditions:* Shelf location |
| When? | Loan and reservation of books |

**Figure 2**.

Empirical analysis shows how a user population tends to characterize book contents at five different levels of abstraction and from a number of different angles (Pejtersen 1980, 1986). The database content must be rich enough to support users' dynamically shifting information needs among these levels: Why a user needs information? What kind of information content and other book attributes will be required? At what levels should the content be communicated and in what form should the content be communicated by the author to the reader ? Some examples are shown.