



Classification schemes for collection mediation: Cognitive work analysis and work centered design

Albrechtsen, Hanne

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Albrechtsen, H. (2004). *Classification schemes for collection mediation: Cognitive work analysis and work centered design*. Roskilde, Denmark: Risø National Laboratory.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Classification Schemes
for Collection Mediation:
Work Centered Design
and Cognitive Work Analysis**

Hanne Albrechtsen

Ph.D. dissertation

A dissertation submitted to the Faculty of Engineering and Science at Aalborg University, Denmark, in partial fulfilment of the requirements for the Ph.D. degree in computer science.

**Classification Schemes for Collection Mediation:
Cognitive Work Analysis and Work Centered Design**

Hanne Albrechtsen

Cognitive Systems Engineering Centre
Systems Analysis Department
Risø National Laboratory

Frederiksborgvej 399, DK-4000 Roskilde, Denmark
Tel: +45 46775153 – Fax: +45 46775199
E-mail: hanne.Albrechtsen@risoe.dk

Copyright © 2003 Hanne Albrechtsen

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the author.

Abstract

Work domain analysis and work centered design is a new approach to design of classification schemes. Classification schemes are symbolic artefacts that inscribe and mediate concepts and structures of domain semantics. In collection mediation, classification schemes are used to search and keep order in paper-based, electronic and digital collections. Research indicates that work centered design of classification schemes is difficult, and that solutions are often based on generic conventions for detailed design, which do not build on needs analysis of the work domains.

This Ph.D. thesis deals with cognitive work analysis and work centered design of classification schemes to support mediation of collections. Cognitive work analysis is a methodology, which supports empirical analysis of work domains and work centered design of information systems. As foundation for overall design of schemes, cognitive work analysis focuses on understanding structures and concepts in work domains and actors' work activities. Research in classification schemes for work domains is also addressed by other perspectives, which either emphasise actors' concept formations in work domains or structures and concept formations within communities like professional disciplines or work collectives.

The thesis addresses two themes within work centered design of classification schemes for collection mediation: classification schemes as sociological phenomena and cognitive work analysis, with an emphasis on analysis of collaborative work. The themes define two research questions that shape the present research. The two research questions are answered by five individual paper contributions. There exists no generally acknowledged conceptual framework for work centered design of classification schemes. The thesis applies the framework of cognitive work analysis as a methodology for analysing collection mediation work and as a starting point for characterising the use, contents and structures of classification schemes in the work domain. The five publications are placed within the means-ends model of cognitive work analysis in order to illustrate similarities and differences between the contributions.

Four empirical sources contribute to the results. Field studies of three film archives and a public library address various collaborative activities in collection mediation. Primary methods applied in field studies are participant observation and interviews. In addition, participatory design of the empirical studies and empirical evaluation of design results have been used. Results from the field studies are discussed in the light of cognitive work analysis, as well as in the light of concepts developed within research in computer-supported collaborative work and sociological theory of communicative action. The studies of collection mediation in film archives focus on cognitive work analysis of collaboration. Their aim has been

to an empirical basis for design of a web-based multimedia film research collaboratory, where classification schemes may contribute to supporting collection mediation and production of new film knowledge. The study of collection mediation at a public library focuses on collaborative design of classification schemes as boundary objects in a web-based information system that mediates multimedia materials to children and adults.

The primary results of this thesis are:

1) The use potentials of collections define the form and contents of classification schemes for collection mediation. Use potentials is not a static feature of the domain, but is constructed in collaboration amongst users and staff of collections. Constructions of use potentials happen in collaborative task situations, which involve information searching, indexing and classification of collections. The thesis discusses different theoretical perspectives on classification schemes in collaborative work and identifies advantages and disadvantages of the perspectives for analysis and design of classification schemes for collection mediation.

2) An important challenge and opportunity in work centered design of classification schemes for collection mediation is that classification schemes are used in different task processes, and that their affordances are dependent on use situation. Cognitive work analysis addresses this challenge through heuristic models for studying and analysing prototypical task situations and the means and ends of work domains. The thesis proposes a general model of a common workspace for collaborative collection mediation. The model is based on the framework for cognitive work analysis. Preliminary design considerations are stated for how classification schemes can be perceived and used in context by the users.

3) An additional important challenge in work centered design of classification schemes is that the contents and structures must be rooted in work domain semantics. Likewise, a translation of analysis to design is required. Analysis of the domain semantics of collection mediation is a comprehensive task, because the semantics is dynamic and diverse. The thesis proposes a preliminary method for translating analysis of semantics in collection mediation to a classificatory structure.

Preface and Acknowledgements

This thesis deals with work centered design of classification schemes with a focus on cognitive work analysis. The thesis consists of this summary and five individual contributions. The publication record for these five contributions is as follows:

- [1] Albrechtsen, H. and E. Jacob (1998). The Dynamics of Classification Schemes as Boundary Objects in Electronic Libraries. *Classification in the Electronic Environment*, edited by S.L. Star and G. Bowker. *Library Trends* 47 (2), pp. 293-312
- [2] Pejtersen, A.M. and H. Albrechtsen (2000). Ecological work based classification schemes. *Dynamism and Stability in Knowledge Organization. Proceedings of the 6th International ISKO Conference*, edited by C. Beghtol and N. Williamsson. Würzburg: Ergon Verlag, pp. 97-110
- [3] Pejtersen, A.M., Albrechtsen, H, Cleal, B, Hansen, C.B. and M. Hertzum (2001). *A Web-Based Multimedia Collaboratory. Empirical Work Studies in Film Archives*. Roskilde: Risø National Laboratory, Centre for Human-Machine Interaction (Risø-R-1284 (EN), CHMI-02-01)
<http://www.risoe.dk/rispubl/SYS/ris-r-1284.htm>
- [4] Albrechtsen, H., Pejtersen, A.M. and B. Cleal (2002). Empirical Work Analysis of Collaborative Film Indexing. *Proceedings of the 4th International COLIS conference*, edited by H. Bruce, R. Fidel, P. Ingwersen and P. Vakkari. Greenwood Village (CO): Libraries Unlimited, pp. 85-107
- [5] Albrechtsen, H., Andersen, H.K.H, Bødker, S. and A.M. Pejtersen (2001). *Affordances in Activity Theory and Cognitive Systems Engineering*. Roskilde: Risø National Laboratory, Centre for Human-Machine Interaction (Risø-R-1287)
<http://www.risoe.dk/rispubl/SYS/ris-r-1287.htm>

The contributions 3 and 5, which were submitted as part of this thesis, are published separately.

I wish to thank a number of people and organisations that have supported this Ph.D. project. First and foremost, I wish to express my sincere appreciation and gratitude to my supervisors Peter Bøgh Andersen, Aalborg University, and

Annelise Mark Pejtersen, Risø National Laboratory, for their support and encouragement of this thesis.

Special thanks go to my fellow co-authors. Without their help and support, this thesis would never have been produced. Special thanks also go to my good colleagues at the Centre for Human-Machine Interaction, and especially to Hans Andersen, Morten Hertzum and Bryan Cleal, for discussions about the research and for commenting on drafts of this thesis. I would also like to thank Ole Harbo, former dean of Royal School of Library and Information Science, Denmark, for his encouragement and support of this research, and John Law, Department of Sociology, Lancaster University, UK, for supervising the first phase of the research.

A number of organisations, schools and foundations have supported this research. I am grateful to all of them: Department of Systems Analysis, Risø National Laboratory, Denmark, Centre for Human-Machine Interaction, Denmark, Royal School of Library and Information Science, Copenhagen Denmark, the Danish National Research Foundation through its funding of the Centre for Human-Machine Interaction, the European Commission's Information Society Technologies Programme through its funding of the Collate project on collaboratories for annotation, indexing, and retrieval of digitised historical archive material (IST-1999-20882) and the Danish National Library Authority through its funding of the Database 2001 project.

I would like to extend my gratitude to Kjeld Schmidt, IT University of Copenhagen, for letting me become part of the international research workshops on classification schemes and cooperative work that he has organised in Denmark and the United States, during the last three years. I also want to thank my good friends and colleagues Raya Fidel and Jens-Erik Mai, Information School, University of Washington, for their support and encouragement of this research. In addition, I want to thank the informants involved in my field studies at the national film archives FAA, NFA and DIF and Ballerup public library.

Finally, I would like to thank my family and friends for supporting me during these past four years, and especially during the last months of working on and completing this summary.

Contents

| | |
|--|------------|
| Abstract | i |
| Preface and Acknowledgements | iii |
| Contents | v |
| 1. Introduction | 1 |
| 1.1 Collection Mediation through Classification Schemes | 2 |
| 1.1.1 Example: Iconic Classification Scheme | 3 |
| 1.1.2 Classification Schemes and Collection Mediation..... | 5 |
| 1.2 Work Centered Design of Classification Schemes | 6 |
| 1.2.1 Classification Schemes for Collection Mediation | 8 |
| 1.2.2 Cognitive Work Analysis and Classification Schemes | 9 |
| 1.3 Research Questions Summary and Thesis Structure..... | 10 |
| 2. Classification Schemes and Collection Mediation | 13 |
| 2.1 Cognitive Work Analysis, Classification Schemes and Collection mediation | 14 |
| 2.1.1 Collection Mediation..... | 15 |
| 2.1.2 Cognitive Work Analysis and Collection Mediation | 16 |
| 2.1.3 The Means-Ends Abstraction Hierarchy | 16 |
| 2.1.4 Means-ends Analysis of Collection Mediation | 18 |
| 2.1.5 Means and Ends and Decision-Making Activities in Collection Mediation | 20 |
| 2.2 Classification Schemes and Cognitive Work Analysis | 21 |
| 2.2.1 Functional Perspectives of Classification Schemes | 22 |
| 2.2.2 Structural Properties of Classification Schemes | 23 |
| 3. Research Contributions | 27 |
| 3.1 The Dynamics of Classification Schemes as Boundary Objects in Electronic Libraries | 28 |
| 3.2 Ecological Work Based Classification Schemes..... | 30 |
| 3.3 A Web-Based Multimedia Collaboratory. Empirical Work Studies in Film Archives | 32 |
| 3.4 Empirical Work Analysis of Collaborative Film Indexing | 33 |
| 3.5 Affordances in Activity Theory and Cognitive Systems Engineering..... | 35 |
| 3.6 Summary | 36 |
| 4. Research Discussion | 39 |
| 4.1 Classification Schemes | 40 |

| | |
|---|-----------|
| 4.1.1 Traditional Sociological Definitions of Classification Schemes | 40 |
| 4.1.2 Modern Definitions of Classification Schemes in Work Centered Approaches | 41 |
| 4.1.3 A Modern Information Science Definition of Classification Schemes | 42 |
| 4.1.4 Social Practice and Social Worlds in Work Domains | 43 |
| 4.1.5 Experience of Social Practice in Collection Mediation Work..... | 44 |
| 4.1.6 Work Domain Semantics..... | 44 |
| 4.1.7 Classification Schemes and Conceptualisations..... | 45 |
| 4.1.8 Experience: Translation of Interests through a Classification Scheme | 46 |
| 4.1.9 Collaboration and Symbolic Artefacts | 48 |
| 4.1.10 Experience: Classification Schemes are not Mechanisms of Interaction | 50 |
| 4.1.11 Implications for Work Centered Design of Schemes | 51 |
| 4.2 Cognitive Work Analysis and Work Centered Design of Classification Schemes..... | 52 |
| 4.2.1 Analysing the Deep Structure of Collection Mediation Domains | 53 |
| 4.2.2 Means-ends Modelling of Collaborative Work | 54 |
| 4.2.3 Meanings of Physical Resources | 56 |
| 4.2.4 Experience of the 'Affordance' Concept in Work Domains | 57 |
| 4.2.5 Affordances of Symbolic Resources in Collection Mediation | 58 |
| 4.2.6 Linking Representation Tools to Means and Ends of a Common Workspace | 60 |
| 4.2.7 A Preliminary Model of a Common Workspace for Collaborative Film Indexing..... | 62 |
| 4.2.8 When is a Classification Scheme a Conceptual Tool? | 64 |
| 4.2.9 Analysing Implicit Structures in Interactive Communication and Scheme Design | 66 |
| 4.2.10 Transforming Analysis to Overall Design of a Scheme | 68 |
| 4.2.11 Summary | 71 |
| 5. Research Methodologies..... | 73 |
| 5.1 Theoretical Foundations | 74 |
| 5.2 Research Approaches and Methods | 76 |
| 5.2.1 Action-Oriented Dialog Research: Application | 77 |
| 5.2.2 Work Domain Analysis: Application. | 78 |
| 6. Conclusion | 81 |
| 6.1 Work Centered Understandings and Models of Classification Schemes in Collaborative Collection Mediation | 81 |
| 6.2 Limitations | 83 |
| 6.3 Future Research..... | 84 |
| Bibliography..... | 85 |

Appendices 1-3 :

| | |
|---|------------|
| 1. The Dynamics of Classification Systems as Boundary Objects for Cooperation in the Electronic Library..... | 95 |
| 2. Ecological Work Based Classification Schemes..... | 115 |
| 3. Empirical Work Analysis of Collaborative Film Indexing | 137 |
| Summary in Danish (Sammenfatning på dansk)..... | 161 |

Chapter 1

Introduction

Classification schemes contribute to orderings in our environment. When picture books are arranged in piles on the floor of a children's library, this constitutes an ordering that can meet the curiosity and interests of children that are too small to reach for books on shelves and who cannot read or understand the labels about book content on the shelves in the library. Yet a borough mayor, guiding a group of international politicians during an official visit to the same library, can judge that same arrangement as disorder. When books are ordered in piles on the floor, the arrangement can serve the purpose of the children's information needs. When books are sorted in groups and placed on the shelves of a library by use of a standardised international classification scheme, the arrangement can make it possible for the visiting politicians to recognise orderings that are similar to those that are used by libraries in their home countries. When the mayor saw the piles of picture books on the floor in the library, he took the cue of rebellion against the directives for cultural institutions within his area of responsibility. The result was an immediate instruction to the librarians to clear up the disorder and revert to mediating the collection according to the library's standard classification scheme and shelving procedure. This story is a real-life incident that took place in the early 1970s at a Danish public library.

As media for telling about orderings in collections, classification schemes meet users who may have different perceptions of ordering. When classification schemes are used as work tools by staff to arrange materials of a collection like a library or an archive, they meet with changes in the collection's stock of materials as new materials are acquired and some materials are removed. The political goals and directives for maintaining a collection may mean that such changes are not radical. But goals and directives are temporary constructs, and so, policies for acquisition may shift entirely. Change in the collection users' interests is an additional dynamic factor. The collection staff's priorities of keeping order in collections may alter in response to the users' changing interests. The staff and the users can jointly experiment with acquiring new kinds materials for the collection. The staff can produce materials that provide answers to users' requests. The staff can mediate users' sharing of materials and communication about materials. New types of technologies, such as networked access to collections from the home or the workplace, allow users to easily access known items in the collections. In turn, the

demands on the collection staff to resolve complex requests from the users, asking for that which they do not know, will increase. Within the networked environments of electronic or digital libraries, the demands on classification schemes for support of complex mediation tasks challenge previous approaches to scheme design. Yet such new developments may also provide for new opportunities for collection mediation through classification schemes.

This thesis deals with work centered design of classification schemes for collection mediation. This is an emerging design approach for classification schemes, and there is presently no general agree-upon methodology for the task. Design of classification schemes for collection mediation implies understanding how the schemes can support people's ordering and search activities in collections. In addition, design of the schemes requires design models. General principles and codes of practice that support ease of construction and maintenance of classification schemes have been developed within, for instance, library and information science (Svenonius, 2000; Miksa, 1992; Ranganathan, 1967). Design of classification schemes is, however, often conducted in an ad hoc manner, where the choice of models and construction techniques for a scheme is made according to general prescriptions for detailed design, prior to an understanding of the task complexity of work domains. Work centered design of classification schemes for collection mediation implies that the designer's perspective shifts from general construction principles and detailed design to analysing and understanding the contingent practices when classification schemes meet users in collection mediation. Such a shift entails an analysis of how classification schemes are used in work domains and a bridging from analysis to choice of models for overall design. This increases the complexity of the design task.

1.1 Collection Mediation through Classification Schemes

Due to the diversity of definitions of classification schemes and the lack of a generally accepted understanding of collection mediation, I will start by drawing out some main properties of classification schemes for collection mediation from an example. The example illustrates mediation of a digitised collection of multimedia materials - texts, musical recordings and images - through a classification scheme. The scheme was created for Database 2001, a web-based information retrieval system, which has been developed and evaluated in a public library (Albrechtsen and Jacob, 1998; Albrechtsen, 1997).

1.1.1 Example: Iconic Classification Scheme

The web page from Database 2001, shown in figure 1.1, was created for children users of the library. The icons on the web page cover the following eight categories:

1. Computers
2. Astronomy, nature, animals, environment
3. First love, star signs, being young today
4. Records (who's best etc)
5. Horses
6. Thrillers, humour
7. Fantasy, science fiction
8. Books that are easy to read

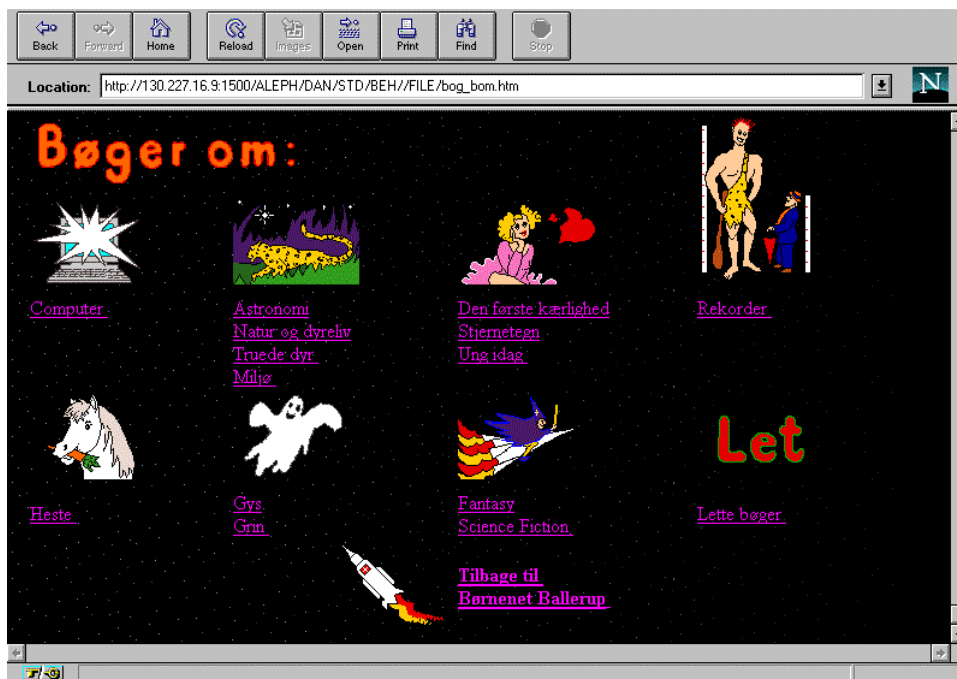


Figure 1.1: Web page with an iconic classification scheme, guiding children's browsing in a multimedia information retrieval system. The classification scheme has eight icons, visualizing the main categories of subjects that the children often ask for when they come to the library where the scheme was developed.

The categories 1-8 and their associated subjects shape a classification scheme. For example, category 2 covers the topics or subjects of astronomy, nature, animals and environments. These subjects correspond to indexing terms. The indexing terms are linguistic expressions used to represent the content of multimedia documents. Database 2001 is composed of a set of web pages with classification schemes for children and adult users, a web-based search client, a database with descriptions of multimedia materials and a repository of images and sound tracks. Figure 1 illustrates some key characteristics of a classification scheme for children on one of Database 2001's web pages. The icons tell the children of particular orderings of multimedia materials. The words that are displayed together with the icons are indexing terms, which tell the children about the contents of a single material or several materials in the database.

When classification schemes meet their users, a number of possibilities and challenges arise. From the web page with icons and indexing terms, the children can start out navigation in a multimedia database. The icons and the indexing terms are stable or invariant elements on the web page. During interaction with the web page, the children can click on the indexing terms and navigate via a search client to entries in the database where the multimedia materials are indexed and described in words, together with images of front pages of books and soundtracks from musical recordings. The navigation between different modalities of operation sometimes makes it difficult for the children to maintain an attention focus on subjects that they have selected on the classification web page. When the children select a subject to explore, the system automatically switches to the search client. From there, the children can click on descriptions and excerpts from materials on the chosen subject.

The children enjoy exploring sound tracks of musical recordings and book descriptions. From time to time, they return to the web page with the iconic classification scheme to find more subjects in the multimedia database. The children then click on the icons of classification scheme on the web page, expecting to find more topics to explore. But nothing happens. They then select indexing terms and start a new navigation in the database. When the children return to the web page with the stable structure of the iconic classification scheme and select on an icon in the scheme, they expect that the scheme will unfold to reveal more about details of the multimedia collections. Their assumption is to have a shortcut through the classification scheme to directly get to the collection materials and circumvent the navigation through search client and database listings. They also assume that the scheme itself will unfold more details of subjects. For instance, that the subject of astronomy category 2 would open up for more detailed inspection of subjects like stars, planets and space travelling.

It appears as if classification schemes may become fairly quickly naturalised in human interaction with information systems. It also appears that users of information systems, including children, have a capacity to deal with fairly complex classificatory structures. This may have to do with children's relative

capacity to grasp the semantic structures of language (cf., eg., Cooper, 2002). But classification schemes are not natural languages. They are symbolic artefacts, constructed to mediate orderings. As artefacts, classification schemes are materially available, i.e. not solely residing 'in the head' of human actors. As symbolic, they mediate meanings that are culturally/socially constructed (cf. eg., Schmidt, 1994, pp. 64-67).

1.1.2 Classification Schemes and Collection Mediation

The above example illustrates how users' exploration of collections can be supported by a classification scheme. A classification scheme can guide users' searching and navigation by giving them a stable overview of collection contents. It can be assumed that the stable features of a classification scheme may counteract users' navigational confusion in searching. Ordering, on the other hand, can be presumed to be related to the way the collection staff chooses to present a collection to the users. Perhaps the task concepts of searching and keeping order are necessary to understand the functions of classification schemes collection mediation. But likely, these concepts are not sufficient to guide design of schemes. The above example illustrates that *stability* is an important property of classification schemes. Exposure of details about the contents of a collection could be an additional important property for searching and keeping order in a collection. When classification schemes unfold details about orderings of collection content, they support people's distinctions between those topics that are within their interest and those that are not.

Context is another important quality of a classification scheme for collection mediation. The iconic scheme in the example is placed within the context of the more low-level modalities of interaction with web pages of an information system. It is also contained within the more high-level modalities of semantic cues of icons. When classification schemes meet their users, they are framed within particular technologies and cultural understandings or conventions of ordering, distinctions and naming. As Star and Bowker (1998) have pointed out, classification involves an understanding of how the informal is embedded in the formal and vice versa. That is, how certain cultural cues are made salient or silenced within formal structures like classification schemes, and how the schemes in turn impact our perceptions of our environment. In Star's words, "we then take cultural cues from the systems so created" (ibid., 1998, p. 185). In the terminology of Habermas (1987, 1984), the life world becomes inexorably linked to the system world through classification schemes.

Searching, browsing, distinguishing, ordering and indexing are important interpretive activities in collection mediation. Stability, structures and unfolding of topical content are important properties of classification schemes. The above example has illustrated such aspects of classification schemes for collection mediation. A first understanding could be that the interpretive activities relate to

‘collection mediation’ as work activities, whereas ‘classification scheme’ is related to structural properties of collections, but how the concepts are linked mandates further explorations.

1.2 Work Centered Design of Classification Schemes

The example has illustrated some key characteristics and challenges of collection mediation through classification schemes. This thesis deals with work centered design of classification schemes for collection mediation. Work centered design of classification schemes is an emerging approach. It involves an understanding of the activities in collection mediation and how classification schemes may support these activities. Early empirical research in the field was mostly concerned with counting the effects of collection mediation. Quantitative research methods were adopted to study the outcomes of mediation, like the number of visits by users to a collection, the number of requests made by users to collection staff and the number of items acquired over time in a collection. As Wilson (1999) has pointed out, very little of this counting revealed insights from where a theory of collection mediation could be developed.

Recent research in library and information science has addressed collection mediation as human information behaviour. Based on empirical studies of communication between collection staff and collection users, and users’ individual search behaviour, a number of models for human information behaviour have been proposed (eg., Kuhlthau, 1993; Ingwersen, 1992; Ellis, 1989). The shift from quantitative to qualitative studies of information behaviour has resulted in research within library and information science that is inscribed in a wider body of knowledge on human behaviour, notably behavioural or cognitive psychology. The positivist tradition of counting effects of mediation tasks has been supplanted with broader cognitive or phenomenological perspectives.

Yet very few models developed within the research on information behaviour have addressed design of classification schemes for collection mediation, but rather the modes of information behaviour to be supported by classification schemes. From a work centered point of view, classification schemes are not only media that support particular modes of information behaviour. They are also *artefacts* embedded in the materiality of collections (Schmidt and Wagner, 2003) and the infrastructures of our environments at large (Bowker and Star, 1999). The body of knowledge on information behaviour has to a large extent ignored the materiality of collections and the social factors impacting scheme design and use. Emerging theories of classification schemes in collections at workplaces, produced within the field of computer supported cooperative work (CSCW), address the mediating role of classification schemes in workplaces (Bowker and Star, 1999; Schmidt and Wagner, 2002, Schmidt and Bannon, 1992).

These contributions are concerned with classification schemes as symbolic artefacts that mediate particular coordinative practices. The contributions are

inscribed in a broad body of knowledge, notably sociology of work, social history of the sciences and library and information science. Contrary to the research in information behaviour, the research on classification schemes as symbolic artefacts is concerned with the socio-material and coordinative practices of collection mediation, rather than with individual actors' information behaviour. Nonetheless, very few contributions within the research in information behaviour and CSCW have come up with methodologies for work centered *design* of classification schemes.

The design of classification schemes, starting out from empirical studies of human information behaviour, has often been approached through imposing general prescriptions in codes of practice for construction of classification schemes in work domains. The emphasis has been on fitting empirical data into pre-established classificatory structures, like thesaurus models and ontology models. It appears as if workplace-oriented design approaches for classification schemes rarely build on methodological approaches for work analysis on the one hand, and methodologies for system development on the other hand. The challenges introduced above for work centered design of classification schemes for collection mediation imply a need for gaining clarity about concepts and possible approaches. The body of work on work analysis and practice-related studies for design of information systems is substantial (cf., eg., Mathiassen, 1998; Schmidt and Carstensen, 1990), and introducing this knowledge and experience in work centered design of classification schemes seems useful.

The framework for cognitive work analysis (Rasmussen, Pejtersen and Schmidt, 1990; Rasmussen, Pejtersen and Goodstein, 1994, Vicente, 1999) is a general approach to empirical analysis of work domains with the purpose of work centered design and evaluation of information systems. The approach is intended to cover empirical analysis and design for several kinds of work domains, ranging from engineering domains to humanistic domains. Amongst the humanistic domains addressed is collection mediation in film archives and libraries, where the units of study comprise mediation tasks as well as contents and environments of collections, and where important results are the design and implementation of classification schemes. The framework for cognitive work analysis suggests a number of general models and methods that can be used for work centered design of information systems and classification schemes.

This thesis explores classification schemes as social phenomena in work domains, and how the framework for cognitive work analysis can guide work domain analysis for work centered design of classification schemes. I choose to address the following themes for bringing support to work centered design of classification schemes for collection mediation:

1. Classification schemes for collection mediation and
2. Cognitive work analysis for work centered design of classification schemes.

1.2.1 Classification Schemes for Collection Mediation

Classification schemes can on the one hand be viewed as abstractions of intrinsic or already existing structures in a work domain. Such structures can be viewed as being an effect of orderings that are broader or more general than the locality of a collection. This could be orderings that have already been constructed to classify knowledge within speech communities like scientific disciplines (Hjørland, 1997) or orderings to sustain coordination of professional practices (Schmidt and Wagner, 2002). As such, intrinsic orderings of a collection can be regarded as being implicit, in the sense, unknown, till wider disciplinary orderings involved in collection mediation are revealed. Conversely, orderings can be understood as intrinsic structures in data produced during collection mediation, for instance data that are created by staff or users during indexing or searching of databases. Such understanding underlies the use of computational techniques of automatic document clustering and linguistic engineering to build up classificatory abstractions of mediation data (eg., Salton and McGill, 1987).

On the other hand, classification schemes can be understood as *constructs* with local as well as global potentials for ordering collections. As such, classification schemes can be perceived and used differently depending on their locality in use. It would then be the properties of multiplicity, plasticity and mobility that constitute the robustness of classification schemes for mediation, rather than their possible inherent match with a 'world out there' (Star and Griesemer, 1989; Bannon and Bødker, 1997) or 'world in here', eg, compliance with individual cognitive structures (eg., Ingwersen, 1992).

The diversity in understandings of classification schemes for collection mediation may be one background for the apparent methodological fragmentation, which currently characterises the body of knowledge on empirical analysis of classification schemes for collection mediation. Nielsen (2001) and Ørnager (1997) claim, that collection users' work tasks and their choice of terms in classification schemes like thesauri are linked through the linguistic expressions that they use to search collections. They therefore suggest that design of a classification scheme to support collection mediation should be based on analysing users' associations to words in a thesaurus that is used to index materials in collections. Other approaches focus on ethnographic studies of users' labelling of documents at work places and how such labelling may reveal task-related orderings of workplaces (notably, Kwasnik, 1989). Iterative prototyping and empirical evaluations of users' interaction with classification schemes in information systems are mostly concerned with design of classification schemes that can optimise users' navigation processes in interfaces (notably, Albrechtsen, 1997).

From an initial distinction between classification schemes as representing immanent objective structures in collections or speech communities versus as representing ongoing construction of orderings created in localities of work domains, it is not straightforward how results from empirical work domain analysis can contribute to understanding classification schemes as sociological phenomena

and to developing approaches for work centered design. The understandings of classification schemes in collection mediation are inscribed in many different research paradigms and design approaches. It is therefore necessary to further explore the concepts of classification schemes as sociological phenomena and their emergence and use in collection mediation work.

This leads to the first research question: *What characterises classification schemes as sociological phenomena, and what does that mean for understanding their roles to support information seeking and sharing of knowledge about collection content?*

1.2.2 Cognitive Work Analysis and Classification Schemes

Cognitive work analysis (Vicente, 1999; Rasmussen, Pejtersen and Goodstein, 1994; Rasmussen, Pejtersen and Schmidt, 1990) has been developed as a conceptual framework to guide empirical analysis of work domains, which considers stable properties of work domains and work tasks as very important units of work analysis. The aim is to bring support to work analysis, which will then guide work centered design and evaluation of information systems. The framework, as a methodology, comprises a number of methods for analysis of data gathered during workplace studies that can inform design and evaluation of information systems. The framework covers a means-ends model for analysing stable properties of a work domain and methods for analysing its recurrent task situations. Cognitive work analysis has previously been applied to analyse and bring support to collection mediation through work centered design of classification schemes in information systems.

Two classification schemes have been developed for mediation of fiction literature in public libraries and are integrated in the icon-based information retrieval system for fiction, the Book House (Pejtersen, 1994; 1989; Rasmussen, Pejtersen and Goodstein, 1994). One classification scheme displays stable attributes of fiction novels. This scheme is used to keep order in fiction materials and formulate search requests. A second scheme displays associative relations between fiction concepts. This scheme is intended for users' navigation in fiction concepts. Mediation of fiction collections is addressed through means-ends analysis of the library domain, the user domain and the document domain. The work centered design of the two classification schemes is based on analysis of means-ends properties in these domains and analysis of task situations of information searching and indexing. Rasmussen, Pejtersen and Goodstein (1994) advocate that when an analyst encounters a mismatch between the practices for keeping order in collections and the practices that users have for searching collections, this may indicate a need to develop new classification schemes.

The suggested distinction between the practices of keeping order on the one hand and the practices of searching on the other hand appears fruitful for understanding the roles of classification schemes as information systems, telling

about orderings in collections, versus as symbolic artefacts of interaction on ordering of collections. However, how the actual coupling between people's work activities and the classification schemes that they use for these activities is to be identified is not a straightforward issue. The coupling is not an effect of a static division of conceptual labour between collection staff that keeps order in collections and users that search collections. Collection staff searches collections, too. Users may keep order in their own collections. Neither is the coupling an effect of classification schemes as metadata control structures for mechanical manipulation of collection substance matter.

Within the body of knowledge on cognitive work analysis it has been argued that there are static as well as dynamic couplings between the units of analysis of people, activities and artefacts like classification schemes. That the coupling between units of work is of a socio-technical, or more broadly, socio-material nature, rather than mechanical linkages or linkages between individual cognitive/mental structures (eg. Vicente, 1999). This latter notion of static versus dynamic couplings between people, activities and artefacts appears to be related to recent sociological theories of the role of artefacts in human work, for instance, to actor-network theory (Latour, 1987; Law, 1994) and symbolic interactionism (notably, Star, 1989). Thus, the notion by cognitive work analysis of dynamic versus static couplings in work domains would seem to be an important issue, demanding further exploration, both for understanding classification schemes as sociological phenomena and for approaches to bringing support to work activities of collection mediation through classification schemes.

This leads to the second research question: *What are the possibilities and limitations of cognitive work analysis in the development of classification schemes for collaborative collection mediation?*

1.3 Research Questions Summary and Thesis Structure

| Theme: | Research question: |
|------------------------|--|
| Classification schemes | What characterises classification schemes as sociological phenomena, and what does that mean for understanding their roles to support information seeking and sharing of knowledge about collection content? |
| Work analysis | What are the possibilities and limitations of cognitive work analysis in the development of classification schemes for collaborative collection mediation? |

The thesis consists of this summary and five individual paper contributions as listed in the preface. The papers deal with different aspects of work centered design of classification schemes for collection mediation and origin from a number of empirical studies.

Chapter 2 of this summary gives some initial definitions of classification schemes and collection mediation and identifies challenges of scheme construction and use. It introduces the main concepts of cognitive work analysis for analysing and modelling collection mediation work, together with some examples of collection mediation from the empirical studies.

Chapter 3 presents the research contributions based on results from the five paper contributions. The five contributions are mapped into the means-ends representation of collection mediation through classification schemes.

Chapter 4 discusses the two above listed research questions by addressing the research results of the five paper contributions and by additional literature.

Chapter 5 deals with the research approaches taken in the thesis, with particular emphasis on discussing and juxtaposing the social theories and conceptual frameworks for work analyses that have guided the present research.

Chapter 6 concludes the work and outlines limitations of the results and draws implications for future research.

Chapter 2

Classification Schemes and Collection Mediation

There is great diversity in understandings of classificatory structures in work domains. This diversity is due to the specific functions and roles that classificatory structures have in different research and practice communities. The variety is also due to different philosophical traditions, which go across research and practice communities. Within the research and practice communities of computing, for example, construction and use of classificatory structures is involved in modelling of interactive information systems. Such modelling activities can rely on different philosophical foundations. Foundations can range from positivist traditions of viewing classificatory structures as constituting true or objective representations of the deep structure of the domains that the information systems serve (eg., Wand and Weber, 1995), to more interpretive traditions that view classificatory structures as transmitting meanings in the communities that are involved in design of an information system (eg., Hirschheim, Klein and Lyytinen, 1995).

Within the research and practice community of information science, the construction and use of classificatory structures address the activities of ordering and searching in collections (eg., Olson, 2002; Svenonius, 2000). Scheme construction can rely on different philosophical foundations. In the positivist tradition, classification schemes for collection mediation are designed inductively through automatic clustering of collection content (eg., Salton and McGill, 1987), through statistical analysis of collection users' queries or terminology (eg., Nielsen, 2001; Ørnager, 1997), or through deductions from assumed universal deep structures, transcending collections and domains (eg., Ranganathan, 1967; Dahlberg, 1974). Conversely, development of classification schemes for collection mediation can be based on more interpretive traditions where the understanding of people's interpretive and communicative activities in collection mediation is central and where classification schemes are viewed as sociological phenomena and as units of analysis in their own right, rather than as more or less direct abstractions of collections or users' discourse. This interpretive tradition has been termed 'the socio-cognitive perspective' (Jacob and Shaw, 1998) and 'domain analysis' (Hjørland and Albrechtsen, 1999, 1995; Pejtersen, 1994, Albrechtsen, 1992) and is an emerging approach within information science.

In this chapter, I will characterise classification schemes for collection mediation. I start out by introducing the main concepts and methods of cognitive work analysis for addressing the work domain of collection mediation and design of classification schemes. The chapter introduces the means-ends abstraction hierarchy of the framework of cognitive work analysis and illustrates how it can be used to represent the properties of a collection mediation domain. The chapter also presents the main understandings and concepts developed in cognitive work analysis for design of classification schemes in collection mediation.

In chapter 3, I will illustrate how the five contributions of the thesis relate to a means-ends model of the properties of collection mediation. In chapter 4, I will discuss and evaluate the lessons learnt from contributions in chapter 3, in the light of selected literature.

2.1 Cognitive Work Analysis, Classification Schemes and Collection mediation

Cognitive work analysis (Rasmussen, Pejtersen and Goodstein, 1994; Rasmussen, Pejtersen and Schmidt, 1990; Sanderson, in press; Vicente, 1999) is a conceptual framework to guide empirical analysis of work domains. Cognitive work analysis has been developed from field studies of different work domains, from engineering design to patient care in hospitals and collection mediation in libraries and archives. As a conceptual framework, cognitive work analysis explains to the analyst what key factors to look for during empirical studies of work domains. The framework covers a number of heuristic models for workplace studies and data analysis. Central models comprise a means-ends abstraction hierarchy for analysing stable properties of a work domain and a method of task situation analysis for understanding people's activities in the domain.

A basic tenet of cognitive work analysis is that humans have different modes of controlling their interaction with the environment. It is suggested that people have specific cognitive resources and subjective preferences, which depend on a proper match between stable features of a work domain and the requirements of various cognitive control modes (Rasmussen, 1986). The features of a work domain are regarded as constituting a work domain structure or a territory of work. The means-ends abstraction hierarchy is used to represent the territory of work. According to cognitive work analysis, people interact with the work environment from three distinct control modes: skill-based, rule-based and knowledge-based (Rasmussen, 1983). The coupling between the means and ends of a work domain and the different control modes of interaction (skills-rules-knowledge) happens through a dynamic world model, constructed 'in the head' of the actors. This world model is assumed to mediate the shift between the three types of control modes. People's construction of world models has been observed through empirical analysis of human decision-making activities in work domains (eg., Rasmussen, 1986). For the

collection mediation domain, Pejtersen (1988) has identified recurrent control modes, which are specific to decision-making activities in this domain. These modes are formulated as recurrent strategies, which people use during their activities in collection mediation. The following introduces the means-ends abstraction hierarchy and the coupling between means and ends and decision-making activities in collection mediation. The aim is twofold: to introduce the main characteristics of collection mediation, as found through empirical investigations of the domains of archives and libraries carried out within the present research, and to introduce the main concepts of cognitive work analysis for understanding classification schemes in collection mediation.

2.1.1 Collection Mediation

A collection is an assemblage of gathered materials. The materials can be artefacts like written documents or images, or more natural objects like samples or specimens. A collection can be held in one physical location, for instance within a particular organisation like a public library, a film archive or an art museum. A collection can also be physically distributed and brought together as one virtual collection through joint administration or collection management by one or several organisations or individuals. Digital libraries or networked digital museums are examples of such distributed collections (eg., Bishop and Star, 2000). Thus, collections may be created and distributed by one or more organisational units. The purposes for the organisations' mediation of collections vary. Collections of books in a public library may, according to official policy statements in alignment with national legislation for libraries, serve a general purpose of public enlightenment and education. Collections may also serve special needs within a community. For instance, a public library may create or use special information services for mediation of specific materials, like fiction literature (Pejtersen, 1994) or create specific collections to serve the information needs of target groups like cultural minorities, visually impaired users, private companies, kindergartens and schools (Thorhauge et al, 1997). Likewise, collections held at national film archives may serve the overall purpose of preservation of cultural heritage, following official national policies for cultural mediation. A national film archive can implement such a policy in its mediation of collections to serve local as well as global needs for research and education (Pejtersen et al, 2001).

Collection mediation is not only about preservation and physical distribution of materials from a collection site to users or to other collections. Collection staff monitors and develops the state of the materials held in collections. The staff also monitors the development of interests in user communities and relates this development to the way they build and manage the collections. For example, a public library may create special databases or web sites to inform the public about topics that are of current interest. A film archive may arrange a retrospective for the public with films created by a particular film director. The collection staff

communicates with users of the collections in order to understand and serve the users' information needs. They conduct searches in electronic repositories of information, for instance databases and the World Wide Web. The staff keeps order in the collections. They index the materials held in the collections and arrange them on shelves in such a way that the collections can be browsed by users or by staff. When collection staff keeps order in collections through indexing, they consider the semantic content and use potential of each material that they index (Soergel, 1985; Albrechtsen, 1993). The staff tries to bring each item to the home of the collection and to the home of the users through interpreting its semantic content and use potential during decision-making activities in indexing.

2.1.2 Cognitive Work Analysis and Collection Mediation

Cognitive work analysis addresses the following recurrent interpretive activities in collection mediation:

- *Keeping order* in collections, that is, indexing of collection items and arrangement of collection items on shelves, in databases or digital libraries and archives. Design and update of classification schemes for indexing and arrangement of collection items.
- *Searching* in collections, that is, browsing collection items on shelves or in databases, communication between staff and users on information needs and formulation of search statements in databases.

Each interpretive activity is addressed empirically through interviewing and observing staff and users in collection mediation in order to determine *what they do* (tasks or functions) and *how they do it* (processes). In particular, cognitive work analysis considers recurrent *task situations*: i.e. *what happens when staff and/or users work on solving a particular problem* (task situations) and *what actors* are involved in recurrent task situations (cf. eg., Pejtersen, et al, 2001; Vicente, 1999). The *environment* in which the activities take place is captured through *means-ends analysis* of the work domain or the overall territory of work in which these activities take place. The method applied for means-ends analysis of the territory of work is the means-ends abstraction hierarchy.

2.1.3 The Means-Ends Abstraction Hierarchy

Means-ends analysis addresses the overall territory of work in terms of (i) domain structures and actors' work strategies on the one hand, and (ii) actors' resources, background and preferences on the other hand. Means-ends analysis is based on two analytical principles: (a) empirical analysis of work domains and (b) mapping of identified domain substance in a means-ends hierarchy (Rasmussen, 1986;

Rasmussen, Pejtersen & Goodstein, 1994). The generic means-ends abstraction hierarchy is displayed in figure 2.1.

| MEANS-ENDS RELATIONS | PROPERTIES REPRESENTED |
|---------------------------------|---|
| Goals and Constraints | Properties necessary and sufficient to establish relations between the performance of the system and the reasons for its design, i.e., the purposes and constraints of its coupling to the environment. <i>Categories are in terms referring to properties of environment.</i> |
| Priority measures | Properties necessary and sufficient to establish priorities according to the intention behind design and operation: Topology of flow and accumulation of mass, energy, information, people, monetary value. <i>Categories in abstract terms, referring neither to system nor environment.</i> |
| General Functions | Properties necessary and sufficient to identify the 'functions' which are to be coordinated irrespective of their underlying physical processes. <i>Categories according to recurrent, familiar input-output relationships.</i> |
| Processes and Activities | Properties necessary and sufficient for control of physical work activities and use of equipment: To adjust operation to match specifications or limits; to predict response to control actions; to maintain and repair equipment. <i>Categories according to underlying physical processes and equipment.</i> |
| Physical resources | Properties necessary and sufficient for classification, identification and recognition of particular material objects and their configuration; for navigation in the system. <i>Categories in terms of objects, their appearance and location.</i> |

Figure 2.1: The means-ends abstraction hierarchy (reprinted from Rasmussen, Pejtersen and Goodstein (1994), with the authors' permission)

The means-ends abstraction hierarchy (figure 2.1) has five levels, from goals and constraints to physical resources. The highest level of *goals* (i) addresses representation of the purpose of the work domain in relation to its functions in the environment. This level applies to the domain's anchoring in cultural, political and economical systems. This can for instance be found in formulations of policy statements for the domain. *Constraints* imposed by outside regulations like legislation or codes of practice are represented in this abstraction level as well. The

second level of abstraction concerns representation of *priority measures* (ii). This level refers to how resources like staff, material and finances are allocated within the domain. The feature addresses how the distribution of activities and materials are managed within a domain, through organisational structure and division of labour and resources. The third level of abstraction is general *functions* (iii). This feature concerns representation of the recurrent work tasks carried out in a domain, irrespective of the physical resources like staff or work tools involved in carrying out these tasks. The fourth level of abstraction focuses on *physical processes* (iv) involved in work activities, which are necessary to establish and maintain the general functions of the work domain. The fifth and lowest level of abstraction comprises representation of an inventory of *physical resources* (v), which are created, used and maintained within the domain. 'Physical resources' also cover the actors involved in activities in the domain, for instance staff and users (Rasmussen, Pejtersen and Goodstein, 1994, pp. 35-55).

2.1.4 Means-ends Analysis of Collection Mediation

The original means-ends abstraction hierarchy, promoted by Rasmussen (1986), has been the object of discussion and further development within the research community of cognitive engineering, concerned with human-machine systems design (eg., Lind, 1999). The original means-ends hierarchy was introduced as a framework for representing engineering control systems like power plants. The means-ends abstraction hierarchy viewed the relationships between substance matter in the five levels of abstraction as governed by laws of nature, the structuring of the control system and by human operators' interpretations and decision-making activities, within these environmental constraints.

The use of the abstraction hierarchy for modelling work domains was later extended beyond the research community of cognitive engineering to research in design and evaluation of information systems to serve a broad range of application domains, for instance within case-handling, patient care in hospitals, aviation and collection mediation. The adaptations of the abstraction hierarchy for such domains are discussed in Vicente (1999), Rasmussen, Pejtersen and Goodstein (1994) and Rasmussen, Pejtersen and Schmidt (1990). Adaptations address the environmental constraints as sources of regularity that *inform* the actors' decisions and freedom of choices, rather than as conditions which causally *determine* the actors' activities and understandings. The result is a means-ends representation, which reflects the territory of work in which the actors will navigate. This approach has been used in the means-ends analysis of the three European film research archives, carried out as part of the present research (Pejtersen et al, 2001). The means-ends representation of a national film archive in figure 2.2 illustrates an application of the abstraction hierarchy for means-ends analysis of collection mediation in film archives (adapted from Albrechtsen, Pejtersen and Cleal, 2002).

A national film archive can have the overall *goal* (i) of contributing to film research and mediating national cultural heritage. The *constraints* (i) for mediation can be codes of practice for mediation, in the shape of rules for indexing and cataloguing. The national archive can have the *priorities* (ii) to preserve national cultural heritage and bringing it home into a collection. The *functions* (iii) in the work domain to serve this priority cover for instance 'keeping order' in collection materials. *Processes* (iv) of keeping order may embrace screening and discussions of films that are going to be part of the archive's collection, creation of subject statements about films and indexing of films. The *physical resources* (v) that support the processes can comprise meeting rooms for film screening, classification schemes for indexing, cataloguing rules for description of films and databases for storage and retrieval of film descriptions.

| MEANS-ENDS RELATIONS | PROPERTIES REPRESENTED |
|--|--|
| Goals | Collect, archive, preserve and restore all film-related materials relevant for the Czech Republic. Ensure commercial viability, via exploitation of materials owned by NFA. National and international collaboration with archives and other professional institutions in and out of film domain. |
| Constraints | Develop international awareness of Czech film culture. Obligated to hold all films released in Czech Republic. International code of practice for cataloging. All income beyond 28% provided by state must be generated by NFA itself. Ownership of rights restricts commercial exploitation of archive information content. Annual audit by Ministry of Culture. Acquisition criteria. |
| Priorities | Dedication to collection building, documentation, preservation and publication of research results. Preservation and restoration of disappearing knowledge. Broaden appreciation for singularity of NFA collection. |
| Functions | Index and describe all materials of collection. Acquire new materials based on archive's own agenda. Publishing activities: production of annual Czech filmography, <i>Iluminace</i> (quarterly) and <i>Film Review</i> (monthly), film books. Collect material from amateur filmmakers, organise retrospectives nationally and internationally, and participate in international film festivals. Participate in international standardisation activities in film indexing, restoration and mediation. Develop and implement new IT resources. |
| Processes | Two groups meet weekly to watch and analyse films. Input information in filmography database. Communication with three sites (buildings) via face-to-face contact, telephone, fax, and email. Cut, edit, and restore film material, restore and scan posters, identify and label new material. Searching films in database. |
| Resources and physical configurations | Full length feature and short films, masters, print and negatives, videocassettes, books, journals, magazines, film scripts, digital laboratory, scanner and colour printer, restoration equipment, workbenches and special chemicals, editing equipment. Cinema. Screening room. Computers, databases, local IT-network environment. Conceptual tools: database format, cataloging rules, list of indexing terms. Human resources: staff, colleagues outside the institution |

Figure 2.2: Means-ends representation of collection mediation at a national film archive

2.1.5 Means and Ends and Decision-Making Activities in Collection Mediation

The relations between the content or substance matter of the five levels in the means-ends representation of the national film archive (figure 2.2) are not static, nor are they logically or causally given. The relationships are dynamically *performed* during actors' activities, for instance during decision making in searching and indexing. It is a major tenet by cognitive work analysis that actors' decision-making activities are coupled to their interpretations of the territory of work and that their action possibilities are given within means-ends abstraction levels, from overall goals to physical resources (cf., eg., Pejtersen and Rasmussen, 1997). From a design perspective, *means ends representations* of the territory of work in information systems make action possibilities visible and ready-to-hand for the actors. Analysis of actors' recurrent task situations, that is recurrent events of individual and collaborative decision-making in the domain, concerns couplings between substance matter in means-ends levels and actors' activities in a domain. In collection mediation, recurrent task situations occur during keeping order and searching in collections.

An example task situation in a film archive: a university student of women's studies is working on an essay on socio-political conditions for women's lives in the Czech Republic during the 1940s, through studies of how national feelings and women's values are represented in Czech movies from that period. The student comes to the national film archive and is introduced to the collection by the staff. The student explores the shelves of the collection together with staff members. Together, they explore databases of the collection, looking for topics that are related to the student's problem. They discuss the results of the searches. Results are books, journal articles and films. The student screens some of the films and discusses the films with the staff. Such exploration of the collection comprises a number of decision-making activities. The initial decision-making activity is to explore as many sources of knowledge as possible. The student's interaction in proximity with collection staff opens up to many kinds of information systems and resources: films, film-related materials, databases, special collections held, staff members who can negotiate with the student about the problem and have experience in solving such requests from previous interactions with users around similar problems. The interaction between user and staff leads to new decision-making situations, like browsing shelves, formulating search statements in databases, screening of films and reading texts, and deciding on the relevance of materials.

The resources that are involved in the task situation are represented as physical resources (level 5) in the means-ends representation of the archive domain (figure 2.2). The goals, constraints and priorities of the archive (levels 1-2) constitute the overall possibilities and limitations for the student to collect the necessary information for her essay. If the archive is only obliged to hold materials produced

in the Czech republic, then the student cannot retrieve relevant materials produced in, for instance, Austria, within this territory of work. If the archive does not index the films (function, level 3), the search (processes, level 4) for film topics cannot be carried out in databases or card catalogues, but could instead involve interaction (processes, level 4) with knowledgeable staff as information resources. If there is no classification scheme available for browsing topics of the collection (physical resources, level 5), the staff and users have no materially available representation of possible orderings of the collection to refer to in their communication.

2.2 Classification Schemes and Cognitive Work Analysis

Cognitive work analysis defines classification schemes for collection mediation along two axes or perspectives:

1. A structural perspective
2. A functional perspective

This corresponds to the general distinction between form versus function of artefacts (cf. eg., Rasmussen, Pejtersen and Goodstein, 1994, pp. 161-202; Hirshheim, Klein and Lyytinen, 1995). Pejtersen and Albrechtsen (2000) state that classification schemes designed on the basis on cognitive work analysis make visible the invariant semantic structures of the territory of work during knowledge exploration of collections. Thus, the functional perspective of schemes (knowledge exploration) appears to be primary. The functions comprise:

- Searching of collection items
- Indexing of collection items

For these functional perspectives, two distinct classificatory structures have been introduced by cognitive work analysis (cf., eg., Pejtersen and Rasmussen, 1997; Pejtersen, 1994):

- An attribute structure based on recurrent perspectives of a collection. This structure is intended to support the actors' knowledge-based behaviour (analytical search strategy)
- A clustered structure based on concepts, which represent subsets of collection content, and linkages between such concepts. This structure is intended to support the actors' rule-based behaviour (browsing strategy)

2.2.1 Functional Perspectives of Classification Schemes

Pejtersen (1988) has formulated the strategies that actors apply to search for information in collections, on the basis of empirical studies of user-intermediary negotiations in public libraries. The strategies are implemented in the Book House system (eg., Rasmussen, Pejtersen and Goodstein, 1994, pp. 265-282). The implemented strategies comprise:

Analytical search strategy. This strategy is used to address information needs in a systematic way. With reference to the example above for mediation of films (2.1.5), the student and the staff may decide to negotiate or construct search dimensions for the student's information needs. Dimensions can for example be subject matter (women's lives, national feelings), place (the Czech Republic) and time (1940s). Likewise, during decision-making in indexing of films, similar recurrent dimensions may be considered. Recurrent dimensions for searching and indexing can be represented as attributes in a classification scheme.

Browsing strategy. This strategy is used to learn more about the content of a collection or to navigate a collection. It is a means for the user to gain new knowledge and get ideas for formulating information needs. For instance, the student in the example (2.1.5) can select some films and other materials from the shelves in the archive, read or screen the materials and discuss the findings with the staff or colleagues. Pejtersen et al (2001, pp. 49-53) give a related example of a search problem in a film archive. In this case, the user is a historian, investigating the cross-disciplinary research problem of the cultural meaning of Austrian films. The staff and the user cannot initially formulate attributes of the search. To begin with, they decide to browse the collection in collaboration. A classification scheme with concepts for subsets of the collection can support a browsing strategy and inspire the interacting actors in sharing knowledge and joint formulation of new ideas for a search.

Further identified strategies are *search by analogy*, *bibliographic search strategy* and *empirical search strategy*. *Search for analogy* relies on comparison between features of collection items, which may be equivalent to attributes of a classification scheme. In the Book House system, the classification scheme with the attribute structure supports this strategy. *The bibliographic search strategy* involves communication between users and the staff about collection items, in terms of attributes of materials as unique works, for instance authors, titles and publication year. The strategy is supported by a bibliographic format for description of materials. Searches for originators of work may also be regarded as subject searches (cf., eg., Hjørland, 1997). For instance, a user can express the need to find materials by and/or about a film director or manuscript author.

The empirical search strategy involves staff's tacit knowledge on the repertoire of information needs amongst users of a collection. This strategy was involved in collaborative design of the icon-based classification scheme, presented as an example classification scheme in chapter 1 of this thesis. The design of this

classification scheme made use of librarians' background with mediation to children, their experience of the kinds of materials that particular groups of children like and their design of visual representations of children's preferred topics in posters exhibited in the children's library. The design of the scheme also integrated new knowledge and ideas for mediation of materials to children. The innovation in the library's mediation occurred as a result of the development of a project-oriented culture in the library, the librarians' introduction to electronic mediation of information and the new opportunity for collaboration with users in the design process (Albrechtsen, 1997).

Overall, *subject searching*, i.e. search for particular topics, can involve a number of search strategies, from browsing and analytical search to empirical search. For analysis and design of classification schemes, the experiences of this thesis primarily considers design of schemes that can support analytical search, browsing and empirical search.

2.2.2 Structural Properties of Classification Schemes

The structural properties of a classification scheme comprise the elements in the scheme and the relationships between these elements. The elements of a classification scheme are concepts. The relationships between the elements constitute an ordering of the concepts. Figure 2.3 illustrates some common types of classificatory structures and their applications. The structures used for design of classification schemes by cognitive work analysis are: *modular structures* and *cluster structures* (emphasised in *italics* in figure 2.3).

Prevalent classificatory structures for collection mediation are: (i) enumerated classification schemes for physical arrangement of collection materials and (ii) thesauri for arrangement of search terms in text-based information retrieval systems. Such schemes unfold collection or database contents (attributes and concepts) in one singular representation. Design of classification schemes by use of cognitive work analysis separates the attribute description of a collection's content and the concept description of this content. Thereby, the actors' entry to collections can be coupled to their preferred search strategy.

Integration of attribute representation and concept representation can be realised ergonomically, that is, in the presentation of the representations in the interface of an information system. For instance, the interface of the Book House fiction mediation system makes visible the two different classificatory structures for fiction as affordances for actors' perception-action in the Book House fiction mediation system (eg., Pejtersen and Rasmussen, 1997; Gibson, 1979). The attribute representation of knowledge about collection content is based on means-ends analysis of recurrent interpretive activities in collection mediation, where searches are formulated in dimensions, as in the analytical search strategy. To the degree such dimensions are identified as recurrent, they are assumed to correspond to 'an externalised deep structure of the work domain' (Pejtersen, Rasmussen and

Goodstein, pp. 123-134, 265-285). Further units of analysis can be actors' understandings of indexing terms used to index materials in a collection. These understandings can be added as concepts linked with the indexing terms in order to build a classification scheme to support the browsing strategy (cf. eg., Pejtersen, 1991).

| Types of structure | Exemplars of classificatory structures | Applications in work activity |
|---------------------------|---|--|
| Class inclusion | Scientific taxonomies Object-oriented models | Theory-building within science Analysis and design of computer systems |
| Tree structures | Enumerated classification schemes | Collection mediation: indexing, shelf Arrangement, searching, browsing |
| Modular structures | Entity-relationship models Faceted classification schemes <i>Dimensions of collection content</i> | Analysis and design of computer systems Collection mediation: indexing, shelf, arrangement, analytical search strategy (1) |
| Network structures | Thesauri Semantic networks Ontologies | Collection mediation: indexing, browsing, Searching Language learning Translation Analysis and design of knowledge-based systems Knowledge management in work organisations |
| Cluster structures | <i>Concept clusters</i> Document clusters | Collection mediation: browsing strategy (2) Browsing in set of documents |

Figure 2.3 Classificatory structures and some of their applications in work activities. The examples shown in **bold** are classificatory structures for collection mediation, applied by cognitive work analysis to support analytical search strategy (1) and browsing strategy (2)

Rasmussen, Pejtersen and Goodstein (1994) claim that classification schemes represent common semantic work territories. They argue that heuristic models like

the abstraction hierarchy and methods for inductive clustering of actors' concepts can be used to design classification schemes that are representative of the deep structures of interpretive work in collection mediation. From a social constructivist perspective on current research on design of information systems for collection mediation, Talja et al (2000) argue that endeavours to abstract and represent deep structures of interpretive work correspond to an objectification of interpretive or contextual knowledge and that such endeavours are based on the notion that context refers to objective reality, which can be described in an exhaustive way. In collection mediation, core activities are interpretations of cultural artefacts like texts, images or films. Some of these activities can be carried out individually. Yet, actors in collection mediation are interdependent, that is dependent on one another through interaction and sharing of knowledge in order to get the work done. No individual actor in collection mediation is omniscient. Interaction amongst actors can happen in proximity, but can also be distributed. Common practices could, from the perspective of communicative action, involve ongoing construction of *intersubjective spheres or common conceptual grounds* amongst actors in collection mediation (eg., Habermas, 1993, 1987, 1984). Common practices could involve the construction of artefacts like *ordering systems* (Schmidt and Wagner, 2002), *boundary objects* (Bowker and Star, 1999) and *mechanisms of interaction* (Schmidt, 1994). It appears that it is not a straightforward issue to analyse work domain semantics in collection mediation and make suggestions for design of classification schemes to support interpretive activities in a work domain.

Chapter 3

Research Contributions

This chapter presents the research of the five individual contributions listed in the preface and below. The contributions illustrate different aspects of collection mediation, cognitive work analysis and design of common representations like classification schemes. The papers 1-2, 4 are enclosed in appendix A, whereas the publications 3, 5 are available as independent parts of this thesis. The publication details for each of the five contributions are as follows:

- [1] Albrechtsen, H. and E. Jacob (1998). The Dynamics of Classification Schemes as Boundary Objects in Electronic Libraries. *Classification in the Electronic Environment*, edited by S.L. Star and G. Bowker. *Library Trends* 47 (2), pp. 293-312
- [2] Pejtersen, A.M. and H. Albrechtsen (2000). Ecological work based classification schemes. *Dynamism and Stability in Knowledge Organization. Proceedings of the 6th International ISKO Conference*, edited by C. Beghtol and N. Williamsson. Würzburg: Ergon Verlag, pp. 97-110
- [3] Pejtersen, A.M., Albrechtsen, H, Cleal, B, Hansen, C.B. and M. Hertzum (2001). *A Web-Based Multimedia Collaboratory. Empirical Work Studies in Film Archives*. Roskilde: Risø National Laboratory, Centre for Human-Machine Interaction (Risø-R-1284 (EN), CHMI-02-01)
- [4] Albrechtsen, H., Pejtersen, A.M. and B. Cleal (2002). Empirical Work Analysis of Collaborative Film Indexing. *Proceedings of the 4th International COLIS conference*, edited by H. Bruce, R. Fidel, P. Ingwersen and P. Vakkari. Greenwood Village (CO): Libraries Unlimited, pp. 85-107
- [5] Albrechtsen, H., Andersen, H.K.H, Bødker, S. and A.M. Pejtersen (2001). *Affordances in Activity Theory and Cognitive Systems Engineering*. Roskilde: Risø National Laboratory, Centre for Human-Machine Interaction (Risø-R-1287)

The following sections present the contributions individually. Each section describes the background of the contribution and the kind of research that has been conducted. Likewise, the main research focus of each publication is introduced and its contributions to aspects of the main research questions are outlined. Finally, the primary results of each publication are listed. The last section gives a means-ends representation for the main properties of collection mediation, addressed in the contributions.

3.1 The Dynamics of Classification Schemes as Boundary Objects in Electronic Libraries

Albrechtsen, H. and E. Jacob (1998). The Dynamics of Classification Schemes as Boundary Objects in Electronic Libraries. *Classification in the Electronic Environment*, edited by S.L. Star and G. Bowker. *Library Trends* 47 (2), pp. 293-312

The paper builds on results from a participatory evaluation study of a collaborative design process, conducted at a public library from 1996-1997 (Albrechtsen, 1997). A project group of staff members in the public library and an evaluation researcher designed a web-based information system for interactive mediation of multimedia materials for users in a community borough. The project involved users in prototyping and empirical evaluation of the system. The multimedia materials were indexed and described in a text database. All entries to the database were enriched with pictures of front pages (books and CD recordings) and sounds (CD recordings and books on tape). The target groups of the new mediation service were children and adults. Amongst the results were two classification schemes for web-based navigation in subjects of the multimedia database. The schemes addressed children and adult users, respectively, and were developed through iterative and collaborative prototyping by the project staff and ongoing empirical evaluation by library users. The schemes were displayed as pictorial icons, from where the users could start navigation in systematic lists of subject indexing terms for the materials in the database. The final version of the application was publicly launched by the library as an experimental new mediation service, addressing the concept of a 'library without walls'.

The focus of the paper is classification schemes, collection mediation and collaborative design of schemes, involving actors with heterogeneous backgrounds in design and use of schemes. The aim is to contribute to opening up the black box of implicit assumptions in information science about the functions of classification schemes for collection mediation. It looks at the gap between theories of classification schemes versus the practices of mediation through classification

schemes. The paper discusses whether Star's concept of '*boundary objects*' may explain evolving local practices of mediating collection materials through classification schemes that target particular user needs (Star and Griesemer, 1989). It also discusses the possibility to bridge the identified gap between theory and practice of classification in collection mediation. The practice of classification for library collections is found to build on implicit epistemological foundations, such as rationalism and empiricism. This means that knowledge is seen as objectively given, as something 'out there', which can be readily captured and stored in the libraries' collections and databases. Similarly, users' search behaviour is seen by some approaches in information science as expressing knowledge 'out there'. The paper argues for a more constructivist orientation to understanding and designing classification schemes for mediation. It is argued that libraries are more than collections and users, but rather constitutes an information ecology. Within a library, viewed as an information ecology, the librarians are not only intermediaries who deliver information to users. They also facilitate knowledge production and construct public arenas for discussions within, for example, cultural mediation. This view challenges the common distinctions between librarians as mediators on the one hand and library users as consumers of information on the other hand. It also challenges the common approaches to design of classification schemes as structures imposed on a more or less static body of knowledge.

The empirical research carried out in the public library was intervention-driven, during my participation in the project. Interventions were negotiated between the library staff and the evaluator and involved collaborative prototyping of classification schemes and interfaces for the web-based information system. The theoretical results of the paper are reflections on how use potentials of libraries may shape the design of classification schemes. In addition, the results comprise a first introduction of a possible bridging between constructivist theory of knowledge representation (notably, Star) and theories of information science of knowledge representation. The paper contributes to the research question of understanding classification schemes as sociological phenomena, from the point of view of the history of classification schemes in libraries, and from the perspective of the constructivist approach of symbolic interaction (Star, 1989). From a means-ends perspective, the emphasis of the paper is on the goals of mediation for local communities, coupled with a priority of collaborative design and classification.

3.2 Ecological Work Based Classification Schemes

Pejtersen, A.M. and H. Albrechtsen (2000). Ecological work based classification schemes. *Dynamism and Stability in Knowledge Organization. Proceedings of the 6th International ISKO Conference*, edited by C. Beghtol and N. Williamsson. Würzburg: Ergon Verlag, pp. 97-110

The paper introduces a new approach to work centered design of classification schemes. The approach builds on the framework of cognitive work analysis and the approach of ecological information systems. The paper coins the notion of 'ecological classification schemes'. Two empirical cases illustrate the possibilities and limitations of using cognitive work analysis for designing ecological classification schemes.

The paper introduces some main concepts for design of ecological classification schemes. The main concepts are the means-ends abstraction hierarchy (Rasmussen, 1986) and Gibson's notion of affordances and invariants (Gibson, 1979). The challenge of designing ecological classification schemes is dealt with, as a fundamental problem of structuring knowledge base contents to meet people's information needs. Knowledge base content is found through means-end analysis of work domains. Thus, design of ecological classification schemes is viewed as the creation of a special facility for knowledge exploration within ecological information systems. The focus of ecological information systems is to situate people's choices and action possibilities within a virtual workspace that makes visible the territory of work. This workspace assembles and arranges the diverse properties of domains into one affordance space that the users can navigate. Stable properties or invariants of the domains are identified through field studies and means-ends analysis of real-life workplaces. Task situations are studied as well in order to address the dynamics. Decision-making activities during task situations and actors' need formulations and search questions are argued to be important backgrounds for designing ecological classification schemes as facilities for knowledge exploration within the affordance space in the ecological information system.

The systems in focus are classification schemes as well as ecological information systems. This dual focus is chosen, because of the assumption that classification schemes must be made visible and available within the context of a familiar territory of work in order to be used and understood by the actors. The basic architecture of an ecological classification scheme is the invariant properties of the workspace, the users' domains and the knowledge contents of the workspace. The first design example of a fiction mediation system initially defines the workspace as collections of books and multimedia materials in libraries. The

invariant properties of fiction mediation are defined as the means and ends of three domains: i) collection contents; ii) users' information needs, and iii) the institutions (libraries) that host the mediation services. The classification scheme constructed to support fiction mediation (Pejtersen, 1989) integrates the initially incompatible properties of the three domains into one structure of five means-ends levels.

The second example illustrates cognitive work analysis, which involved extensive empirical studies of collaboration at a Danish engineering company. The aim of these studies was to explore how cognitive work analysis could guide the design of an ecological information system for concurrent collaboration on product development. Within this workspace, product specification, design, evaluation and marketing are central work activities taking place in multidisciplinary teams. Collection mediation is not a primary activity in this domain, but information searching and knowledge sharing among team members are central work activities to explore possible design ideas and solutions. The paper draws on Carstensen's special study of information exploration in project documentation (Carstensen, 1997) together with means-ends modelling of the different domains involved in the workspace of concurrent engineering.

It is concluded that it may not be possible to generalise results from the fiction mediation case to the concurrent engineering case for creating ecological classification schemes. Concurrent engineering involves interaction between many different domains coming together one workspace, when a development project is carried out in multidisciplinary teams. Configurations of such teams will vary, according to what kind of expertise is needed and according to the phases in development of the product. Information searching, for documents and people's expertise, and knowledge sharing constitute recurrent activities during concurrent engineering projects. The domains engaged in these activities can be subject to means-ends analysis. It is however questioned, if the properties and substance matter of the means and ends of these domains can be translated into one classification scheme. The paper contributes to answering the research question of understanding classification schemes in collection mediation, from the perspective of cognitive work analysis. The main concepts of 'invariants' and 'invariant structures', inspired by Gibson's theory of visual perception, are also introduced and discussed. These concepts are recommended to guide cognitive work analysis for scheme design. The paper addresses an important issue of the main research, namely whether the concept of 'invariant structures' in cognitive analysis is equivalent to immanent or constructed structures in work domains. From a means-ends perspective of the research in this thesis, the emphasis of the paper is on the goal and priority of methods for work centered design of classification schemes.

3.3 A Web-Based Multimedia Collaboratory. Empirical Work Studies in Film Archives

Pejtersen, A.M., Albrechtsen, H, Cleal, B, Hansen, C.B. and M. Hertzum (2001). *A Web-Based Multimedia Collaboratory. Empirical Work Studies in Film Archives*. Roskilde: Risø National Laboratory, Centre for Human-Machine Interaction (Risø-R-1284 (EN), CHMI-02-01)

The publication reports from empirical field studies of three film research archives, carried out during 2000-2001 in connection with the European Collate project (*Collate, 2000*). The three film archives are Deutsches Filminstitut (DIF), Frankfurt, Germany, Filmarchiv Austria (FAA), Vienna, Austria and Národní Filmový Archiv (NFA), Prague, Czech Republic. The vision of the Collate project is to create a web-based and interactive film research collaboratory for professionals working with film media. The problem addressed is possibilities and challenges in the current work practices in the film archives for creating a film collaboratory.

The focus of the report is on collaboration forms and collaborative task situations in the archives' mediation practice. Case examples concern collaborative information searching and collaborative indexing and cataloguing. All authors of the report contributed to means-ends analysis of the archives. I chose to focus on the task situations of collaborative film indexing at NFA, collaborative information retrieval at FAA and collaborative classification and cataloguing at FAA. My analyses address role allocations among staff and users during collaboration and their joint decision tasks, processes and strategies. Furthermore, I studied the work tools and common representations used by the staff. The report states that the archives mostly apply locally developed classification schemes for indexing and searching, which are built in an intuitive and ad hoc manner. It was furthermore a finding that collaboration is at the core at most tasks in the film research archives. Collaboration takes the shape of either teamwork among staff, group work among staff and users and ongoing consultation between staff and users. Collaboration can happen amongst actors in proximity or in a distributed manner.

The report contributes with a comprehensive empirical work analysis of the three archives, based on cognitive work analysis. Current constraints for creating a web-based collaboratory are identified. Amongst the apparent constraints are diverse priorities, competences, practices and work tools in the archives. The study and its results focus on an early stage of the development of a collaboratory, where the concept of a film research collaboratory is still open to discussions and definitions. The study contributes with a work domain analysis of collaborative collection mediation, which deals with mediation activities within and among film

archives in detail. The results indicate varying professional strategies for building, ordering and mediating collections amongst the archives.

While the archives are based on different cultural and historical traditions of collection building and mediation, a shared theme is that their mediation practices are focused on making visible the use potentials of their collections. Use potentials are dealt with in various ways by the archives, but the recurrent pattern found is to either enrol users in description and acquisition of materials of an archive or to enrol users in continual proximity in an archive's search activities. The purpose of user enrolment in the archives' activities is not only to facilitate circulation of materials, but also to participate in production of new film knowledge in a collaborative and situated learning environment. Film research is a cross-disciplinary field of knowledge, characterised by high degree of strategic interdependencies among researchers. This means that people (staff and users), and not documents only, are crucial resources to mediation of film knowledge. It also means that there are diverse perspectives, paradigms and understandings amongst staff and users as well as across cultures about possible orderings of the field. This challenges design of classification schemes for a cross-cultural film research collaboratory. The report contributes to several aspects of the research questions, in particular to the question of understanding collaboration, in the face of diverse mediation traditions and diverse user needs in collection mediation. In a means-ends perspective of the main research of this thesis, the report contributes with a focus on collaboration between institutions and people in collection mediation. In addition, the focus is on application of cognitive work analysis for collaborative work in collection mediation.

3.4 Empirical Work Analysis of Collaborative Film Indexing

Albrechtsen, H., Pejtersen, A.M. and B. Cleal (2002). Empirical Work Analysis of Collaborative Film Indexing. *Proceedings of the 4th International COLIS conference*, edited by H. Bruce, R. Fidel, P. Ingwersen and P. Vakkari. Greenwood Village (CO): Libraries Unlimited, pp. 85-107

The paper explores the case of collaborative film indexing at a national film archive. It draws on a study of empirical analysis of three film archives (Pejtersen, Albrechtsen, Cleal, Hansen and Hertzum, 2001). A means-ends representation of the subset of means-ends properties in the archive is given. The task situation of collaborative film indexing is analysed as work arrangements and decision-making activities. An attribute representation of collaborative indexing is suggested as a conceptual tool, to support discussions and interpretations during the collaborative indexing task. This representation is derived from analysis of the interpretive

activities in the collaborative task situation. The model for the structure of the conceptual tools draws on a previous model for indexing of fiction (Pejtersen, 1994). The limitations of the archive's present conceptual tools for indexing and possible applications of the suggested conceptual tool for collaborative indexing are illustrated and explained in terms of analysis of the collaborative task situation and the information resources and tools applied by the archive staff.

The systems in focus in the paper are conceptual tools for the collaborative processes of subject analysis, indexing and description of films. The primary purpose of collaborative film indexing is to create a national filmography for mediation of the archive's documentary research on national films. A further purpose is to facilitate mediation services of the films held in the archives. The conceptual tools that support the staff's processes are an alphabetical list of indexing terms, a database format, a list of film genres and an international code of practice for film cataloguing. The collaborative task situations are screening meetings where films are analysed and discussed, and collaborative film description. Screening meetings involve staff and representatives of the target groups for the filmography and mediation services. Film description involves the staff only. The archive's conceptual tools for film indexing are applied in the task situation of film description. These tools devise a high granularity in description of films, as unique works, types of material and provenance. However, their support of collaborative film interpretation is inadequate.

The conceptual tool for collaborative film indexing introduced in the paper is intended to provide a higher degree of support for collaborative film interpretation. The attributes of the applied model range from overall goals and values of a film to a film's archive attributes. This sequence of attributes is argued to follow the sequence of attributes in the generic means-ends abstraction hierarchy. The model attempts to cover intellectual access to films, as well as physical access in circulation and storage of collection items. The paper illustrates application of cognitive work analysis for work centered design of classification schemes. The paper contributes to the research question on cognitive work analysis through its application of the approach for detailed analysis of the functions, processes and resources involved in collaborative mediation. This motivates a means-ends placement of the paper in the main research, within the focus of collaborative film indexing, analysis and subject indexing and resources like archives, staff, users and ordering systems.

3.5 Affordances in Activity Theory and Cognitive Systems Engineering

Albrechtsen, H., Andersen, H.K.H, Bødker, S. and A.M. Pejtersen (2001). *Affordances in Activity Theory and Cognitive Systems Engineering*. Roskilde: Risø National Laboratory, Centre for Human-Machine Interaction (Risø-R-1287)

The report covers the research on affordances and human-computer interaction, which has been carried out within the Centre for Human-Machine Interaction (CHMI). This research was initiated to explore the application of Gibson's notion of affordances within human-machine interaction (HMI) and to understand why the concept of 'affordance' has so far to a large degree been understood as direct manipulation of interface objects.

Within HMI, discussions of the affordance concept have on the one hand focused on affordances in a strict Gibsonian sense as 'direct perception and action' of objects in interfaces (notably, Gaver, 1991). Affordances have also been understood as elements already inscribed in information systems that are made visible through display in interfaces of the systems (notably, Norman, 1999). Such understandings imply a view of affordances as 'add-on' features to interfaces. Two well-established perspectives on HMI, activity theory and ecological information systems, understand and apply Gibson's concept of affordances in a more comprehensive sense. In these perspectives, human activity is the central unit of analysis for design of interfaces to computational systems. Human work activity is understood as situated and co-evolving with the states of affairs in the environment, including the artificial environment of an information system. Affordances not only guide direct manipulation of interface objects through direct perception-action, but also *situate* the activity of human-machine interaction. Thereby, affordances are understood as localised and domain-specific, rather than as add-on features of interfaces versus as essentially given.

The focus of the report is how Gibson's affordance concept has been applied within HMI research and how the concept can and cannot be extended to cover interface design that is based on analysis of human work activity. In order to contribute to an understanding of how Gibson's concepts have been applied in HMI in general and how Ecological Information Systems (EIS) appears to have extended the original intent of the notion to cover perception-action only, I chose to explore Gibson's ecological theory of perception and action and recent discussions within ecological psychology. In addition, I chose to investigate the historical background for EIS's applications of Gibson's notion of affordances and the related Gibsonian concepts of invariants and invariant structures. This investigation attempts to address EIS's recent notion of classification schemes as

structured affordance spaces in interfaces, representing the deep semantic structures of work (eg. Pejtersen & Albrechtsen, 2000).

It was Gibson's claim that for affordances to be perceived and acted upon as cues, they must be situated in and stand out from invariant structures of the environment. The results of the investigation of Gibson's original concepts are that Gibson did not address higher cognitive processes, such as human beings' interpretation and use of symbolic representations (signs or sign systems) or artefacts. This is a challenge for applying these ecological concepts for human-machine interaction. However, it is also concluded that the interaction between people and information systems is not a static relation between a human agent and an artificial system environment. The interaction is regarded by EIS and as well as by activity theory as a more dynamic mutuality. This understanding of actors' mutuality with the environment is in line with Gibson's overall conception of how activity in natural as well as artificial environments takes place.

The report investigates and discusses the history of applying Gibson's ecological theory, from Rasmussen and Vicente's mapping of Gibson's affordances towards the means-ends hierarchy to the early development of EIS. The conclusion is that Gibson's ecological theory may be applicable for designing interfaces that hypothesise an ecology where human activity can take place. The mutuality in such artificial environments evolves dynamically through the users' already learned abilities to interpret and construct signals, signs, and symbols in their workspace. Contrary to Rasmussen and Vicente, EIS addresses loosely coupled work domains like research in the humanities. The stipulation of classification schemes as structured affordance spaces is addressed through linking EIS's notion of self-organisation of loosely coupled work domains to recent constructivist theory in social informatics, notably Star and Ruhleder's notions of infrastructure. The publication contributes to the research question on possibilities and limitations of cognitive work analysis for scheme design in collection mediation, in particular with respect to whether a classification scheme, which maps the semantic territory of work, can be regarded as an affordance in the interface of an information system. The emphasis on cognitive work analysis motivates a placement of the contribution within the means-ends perspective of goals and priorities of HMI research in the main research of the thesis.

3.6 Summary

The experiences gained in the five above independent contributions place the contributions within the following mean-ends representation (figure 3.1). The placement of the contributions within each level of the means-ends representation is motivated in the summary of each contribution.

| Means-ends | Properties represented in independent papers | Properties in focus in independent papers |
|--|--|---|
| Goals and constraints | Cultural mediation; classification research; film heritage research; HMI research | Mediation for local communities (paper 1); To contribute to workcenteredclassification research (paper 2); Research in film heritage (paper 3 - 4); To contribute to HMI research in affordances (paper 5) |
| Priorities | Team-building; innovation of mediation; heterogeneous information needs and communities; identify and analyse core tasks in collection mediation; collaboration in collection mediation; collaborative film indexing | Collaborative classification (paper 1); Scheme design for heterogeneous information needs (paper 2); Analysis of collaboration in collection mediation (paper 3-4) Affordances and deep structure of domains (paper 5) |
| Functions | Design of classification schemes; interface design; film mediation services; film restoration and preservation; film research; retrospectives and exhibitions, etc. | Collaborative design of classification scheme and interface (paper 1) Cognitive work analysis and scheme design (paper 2) Information searching and indexing of films (paper 3) Collaborative film indexing (paper 4) |
| Processes | Communication amongst staff and users by phone, e-mail and in person; film screening meeting; discussions of films; consultations amongst staff; planning of entries and information input in databases | Communication amongst staff and users in archives (paper 3) Analysis and subject indexing of films (paper 4) |
| Resources and physical configurations | Collection sites, staff, users, collection materials, design tools, IT-systems, ordering systems, means-ends models and classification schemes | Library, staff, users, multimedia materials, web design tools, classification scheme , database (paper 1) Means-ends models and classification schemes (paper 2) Archives, staff, users, IT-systems, ordering systems, films and film-related materials (paper 3-4) |

Figure 3.1: Means-ends representation of the experiences of classification schemes in collection mediation in the five paper contributions of this thesis

Chapter 4

Research Discussion

The preceding chapter 3 has presented my five contributions individually by summarising the research process and the results of each contribution, as they apply to the two overall research questions. A summary model illustrated how each contribution addressed some properties of the means and ends of the research experience of classification schemes and collection mediation.

In this chapter, I will address the two research questions raised in the introduction. The research questions are:

1. What characterises classification schemes as sociological phenomena, and what does that mean for understanding their roles to support collection mediation?
2. What are the possibilities and limitations of cognitive work analysis in the development of classification schemes for collaborative collection mediation?

In the previous chapters 1-3 it was stated that very little research actually exists on work centered design of classification schemes, based on a methodological approach. Cognitive work analysis is a methodological framework for work domain analysis and work centered design of information systems and offers concepts and methods for work domain analysis and scheme design in collection mediation. Examining the nature of classification schemes to support collection mediation appears to require a close look at how orderings of collections are constructed and used, and how classification schemes contribute to making use potentials of collections visible. This leads to the following claim:

The use potentials of collections shape the structure and contents of classification schemes for collection mediation, defined as classification, indexing and searching of collections. In order to explore this claim, I set forth the following two partial theses:

- Existing or static features of the work domain of collection mediation do not solely give the use potentials. The use potentials evolve in ongoing

conceptualisations about the work domain and decision tasks amongst the actors.

- Conceptualisations develop from actors' social-material encounters and intersubjective constructions of knowledge during collection mediation activities.

The discussions in this chapter will therefore also address the above claim and the partial theses. The discussions will relate the research experience to these suggestions and bring in selected literature, which can challenge their ideas and clarify their implications. The chapter will progressively create models of features and activities in the work domain for how to compose common workspaces for collaborative activities in collection mediation.

4.1 Classification Schemes

The first research question concerns classification schemes and collection mediation: *What characterises classification schemes as sociological phenomena, and what does that mean for understanding their roles to support collection mediation?*

The following sections deal with classification schemes as social phenomena. A distinction is made between traditional sociological definitions of classification schemes and modern definitions of classification schemes. Classical sociology has promoted notorious definitions of how social factors are linked with systems of ordering (4.1.1). Modern definitions are concerned with the social structures or activities in work domains, which specifically affect the *design* of classification schemes. Sections 4.1.2-4.1.4 deal with modern perspectives of classification schemes as social phenomena. In sections 4.1.5-4.1.10, I explore the notions of work domain semantics and use potentials in collection mediation through discussions of my experience cases in this domain, in the light of the literature. Section 4.1.11 summarises the results of the discussions and makes the suggestion that work centered design of classification schemes for collection is faced with the challenge of capturing work context on the one hand and deep semantics on the other hand.

4.1.1 Traditional Sociological Definitions of Classification Schemes

Durkheim (eg., Durkheim and Mauss, 1963) and Levi-Strauss (1966) formulated prominent classical sociological definitions of classification schemes. In Durkheim's conception, any classification scheme is a projection of an innate order of social consciousness onto natural phenomena. In Durkheim's terms, social consciousness is a necessary condition for people to project and perceive order in

the world. The implication of such a view is a definition of a classification scheme as a text, which documents social consciousness at given time and place. In contrast, Levi-Strauss advocates that the order of cosmos comes first, and that classification schemes are socially or culturally specific abstractions of universal semantic and structural primitives. The latter view entails a definition of a classification scheme as a particular social or cultural sign system, a product of social signification or naming, whose form and contents can be reduced to generic primitives, which constitute an assumed structural unity amongst domains and culture, independently of actual social factors. The philosophical implication of such early sociological definitions of classification schemes is *dualism* between a certain social or universal world of thought or meaning and a particular socio-material reality. While Durkheim's and Levi-Strauss' conceptions of how social structures emerge have been abandoned in modern sociology (eg., Ritzer, 2000), the dualism between thought and reality, or mind versus body, inherent in such definitions of classification schemes, still permeates more recent sociological understandings of classification schemes, for instance Douglas (1966), Leach (1976) and Berlin (1974). The classification schemes addressed in this thesis are of much more modest magnitude than the ordering systems dealt with in classical sociology. Nonetheless, it is relevant to mention the classical sociological views of classification schemes here, as more recent sociological research on classification schemes acknowledges the tradition, but attempts to go beyond the dualism of the traditional view (eg., Bowker and Star, 1999, pp.60-61, 290-317; Schmidt and Wagner, 2003).

4.1.2 Modern Definitions of Classification Schemes in Work Centered Approaches

Work centered approaches to understanding classification schemes as sociological phenomena address the issue from the perspective of the actors' activities in a work domain. The core object of study is classification schemes as symbolic artefacts in workplaces. The core unit of analysis is human collaborative work activity. Cognitive work analysis defines classification schemes as invariant structures which make visible the deep semantics of a work domain for users of information systems (eg., Rasmussen, Pejtersen and Goodstein, 1994). By deep semantics is not meant an assumed single generic structure, but rather evolving as well as stable domain-specific structures of the work semantics. Schmidt (1994) defines a classification scheme as a symbolic artefact in the category of coordinative artefacts, like 'mechanisms of interaction'. According to Schmidt (1994), a classification scheme mediates articulation work in a work domain and thereby reduces the complexity of articulation work for individual or collaborating actors. This implies that classification schemes support actors' sense making in collaboration and their construction of work-related concepts. For the domain analyst, this suggests an emphasis on identifying and analysing the actors' ongoing

mutual sense-making and concept development when they interact with classification schemes or related symbolic artefacts during work activity. Bowker and Star (1999) state that classification schemes can be enmeshed in work domains in many different ways, as boundary objects, as information systems and as boundary infrastructures. They also argue that classification schemes are often only visible upon breakdown in work. That is, as long as classification schemes function to support the actors in getting their work done, the schemes are naturalised and invisible embodiments of tacit or implicit knowledge. For work domain analysis, this means not only to look for explicit knowledge in the domain in the shape of texts, organisation diagrams, etc., but also for situations of breakdown where classification schemes no longer function to support work.

4.1.3 A Modern Information Science Definition of Classification Schemes

Traditional information science definitions of classification schemes do not regard the social practice of collection mediation work as a core object of study. The emphasis is on finding universal principles for organising collections, which may support any practice in collection mediation, such as indexing of documents and information searching (Miksa, 1999; Svenonius, 2000). Contrary to the traditional approaches in information science, Hjørland (1997, 2002) claims that the contents and structures of classification schemes for collection mediation are effects of broader social or epistemic structures of knowledge. Hjørland (2002) mentions the case of special libraries for medical information as an illustration in point. He argues that the core performance criterion for all staff activities in such a library is to serve people dealing with health problems, in line with an assumed core performance criterion for doctors to cure the patients. His suggestion is to apply a documentary, desk-study based approach to analysing features of collection mediation domains and classification schemes. Hjørland (2002) recommends studies of sources to explicit domain-specific knowledge, ranging from domain-specific subject guides, to special classifications and special purpose languages.

Hjørland's domain-analytic view entails, that features of collection mediation work domains are to be understood as being defined by high-level goals of particular user groups. For mediation of medical collections, user groups would then be doctors, nurses and paramedical staff. In other words, it is implied that special collections primarily serve fairly homogeneous and professional user communities. Likewise, his view implies that work activity in collection mediation is determined by one particular high-level performance criterion, and that due process in work is a matter of consensus about such a criterion. Because of the wide-spread use of special libraries, like medical collections, outside core user groups, for instance by researchers in sociology, psychology and social work, Hjørland suggests to extend documentary studies to different discourse communities whose knowledge interests may go across specific research

specialties (Hjørland, 1997; 2002). One design consequence which can be drawn from Hjørland's sociological conception of classification schemes is that creation of classificatory structures in collection mediation domains is equivalent to identifying and integrating high-level disciplinary or epistemic structures, which exist 'out there', independently of possible conceptualisations emerging in the social practice of workplaces involved in collection mediation. Thereby, Hjørland's sociological understanding of classification schemes for collection mediation is infused by the dualism in classical sociological definitions of classification schemes, notably Durkheim's: that social order or consciousness comes first and is projected onto the materiality of classification schemes and the social practice of collection mediation.

4.1.4 Social Practice and Social Worlds in Work Domains

From a work-analytical perspective, Gerson and Star (1986) argue that social practice in workplaces evolves through actors' mutual articulation of work in coordination of work. A unit of analysis is *due process*, which evolves from ongoing reconciliation of divergent assumptions and procedures, rather than from one single performance criterion. Explicit models for due process may exist in a workplace in the shape of documents with policy-formulations or explicit procedures and in the shape of common representations like classification schemes. Gerson and Star (1986) argue that common representations are always incomplete, because they are frozen in time, drawn from a knowledge base, which is continually changing and rarely embody all viewpoints of work domain semantics.

From a related work-analytical perspective, which is based on constructivist sociology of science, Star and Griesemer (1989) plead that repositories like libraries and museums and common representations like classification schemes are *boundary objects*, which sustain collaboration across divergent social worlds. As boundary objects, they inhabit several intersecting social worlds and satisfy the information needs of each such world. Repositories have the advantage of modularity. That is, the actors do not need to negotiate differences in use purpose amongst social worlds, but can simply use or borrow from the collection of objects or documents for their own purposes. The notion of collections as boundary objects implies, then, that collections may be conceived of as special or general to begin with, by their creators or their sponsors, but are simultaneously interpreted and used differently by actors from different social worlds.

The concept of 'social world' is not equivalent to Hjørland's concepts of disciplinary or epistemic structures. Social worlds cover all kinds of communities from those that have established social technologies and norms for generalising knowledge claims, like sciences, to more loosely coupled communities like amateur collectors, for the museum domain, and patients, for the medical domain. Likewise, Star and Griesemer's (1989) notion of boundary objects means that classification schemes are *dynamic constructs* co-created locally by actors from

different social worlds. Drawing design consequences from Gerson and Star (1986) and Star and Griesemer (1989) for work centered classification schemes implies analysis of the dynamic and interpretive collective practices of work domains. For work analysis, an important implication is that classification schemes are social constructs for ongoing translation or negotiation between actors from divergent social worlds and not the result of projection of invariant social structures. The core object of study and unit of analysis for work centered design of schemes would then be the dynamics of actors' collaborative translations and negotiations, rather than an exclusive focus on stability of social or epistemic structures, as Hjørland has suggested (2002).

4.1.5 Experience of Social Practice in Collection Mediation Work

In the experience of cognitive work analysis of three European film archives (Pejtersen et al, 2001), an important observation is that the three archives studied share the high-level goal of contributing to preservation of national cultural heritage and international film research. Based on this finding of a common high-level performance criterion, in the face of diverse user communities, it would seem that Hjørland's claim (2002) of homogeneity as equivalent to fusion of high-level epistemic horizons is supported. The study of the film archives found differences in priorities and functions in the archives. One archive gives high priority to collection building and to curation and indexing of materials. As a consequence, the archive has developed detailed codes of practice for indexing and description of materials in order to create semantic coherence in staff's interpretations of collection materials (Albrechtsen, Pejtersen and Cleal, 2002). In Gerson and Star's terms (1986), due process is carried out through closing the system of work locally and temporarily through explicit codes of practice, so that the indexing job gets done. Two other archives give priority to high quality services to archive users and to direct face-to-face communication with users. One archive is a non-governmental institution and builds a large part of the collection on donations and voluntary submission of materials from their users. Activities in the two latter archives do not build on codes of practice for indexing and description of materials. Rather, semantic coherence is continually performed and developed through sharing of experience and mutual learning amongst staff, and especially from knowledge gained through interaction with users (cf., eg., Hertzum, et al, 2001). This latter pattern could be argued to comply with Star and Griesemer's notion of collections as boundary objects, which are used and maintained in such a way that they can serve actors from different social worlds (Star and Griesemer, 1989).

4.1.6 Work Domain Semantics

A preliminary observation from the above literature and experience is that classification schemes are related to the semantics of work domains. It appears that

collection mediation has many semantic layers, which not only have to do with explicit structures of knowledge, but also comprise implicit and tacit knowledge. Collection mediation semantics may be permeated by epistemic and disciplinary diversity (Hjørland, 2002) or, more broadly, by the interests of diverse social worlds (Star and Griesemer, 1989). At the same time, it appears that ongoing negotiations and translations of orderings amongst local actors contribute to domain semantics as well (Star and Griesemer, 1989). From the experience in the three film archives (Pejtersen et al, 2001; Hertzum et al, 2002), it could be argued that the actors' activities of negotiation and translation contribute to creating semantic coherence in collection mediation. Domain semantics can be inscribed and coded in artefacts like common representations, which sustain actors' interpretation of work (Star and Griesemer, 1989). Work domain semantics can be found in infrastructures like codes of practice and work organisation (Pejtersen et al, 2001).

In my partial thesis (1) I state that use potentials are not solely given by existing or essential features of the work domain, but evolve through ongoing conceptualisations and interpretations amongst actors in collection mediation domains. If 'use potentials' is equivalent to 'work domain semantics' in collection mediation domains, then this partial thesis is challenged, because work domain semantics can be given by existing features or taxonomies of work, inscribed in codes of practice and explicitly named through work organisation and task allocations. In the following, I shall pursue the idea that 'use potentials' is related to actors' activities of negotiation and translation of work domain semantics, for instance in a) actors' translations of diverse understandings of domain semantics into one common classification scheme and b) actors' negotiations and sense-making of domain semantics in social-material encounters.

4.1.7 Classification Schemes and Conceptualisations.

Bowker and Star (1999) state that classification schemes represent ongoing conceptualisations of work. A classification scheme is defined as a set of boxes (metaphorical or literal) into which things can be put to then do some kind of work – bureaucratic or otherwise (Bowker and Star, 1999, p. 10). This definition should be understood as broader than formalist or essentialist definitions of classification schemes as entities which comply with particular principles for classification, like consistency, exclusivity and completeness (eg., Svenonius, 2000; Ranganathan, 1963). Rather, Bowker and Star's definition is empirical, or pragmatic. That the work which is involved in building, maintaining and using the kinds of entities that people call classification schemes is the central object of study for understanding classification schemes as sociological phenomena, rather than whether classification schemes are good or bad representations in essence. Bowker and Star (1999) define important properties of a classification scheme as its ambiguity or interpretive flexibility (Pinch and Bijker, 1987) and its ability to travel as an

immutable mobile across contexts (Latour, 1988). In Star and Griesemer's terms (1989), classification schemes are boundary objects, which are robust enough to retain a particular identity across contexts and are plastic enough to be adapted to local or situational needs.

As *boundary objects*, classification schemes sustain collaboration between people across divergent social worlds. Classification schemes arise over time as means for translation between collaborating actors from different social worlds, come to live in multiple social worlds and have different identities in each such world. Classification schemes can function as *information systems*, which can reside in multiple contexts. Classification schemes can be *work objects*, created, maintained and updated by people from divergent worlds, coming together to create common understandings, for instance in cross-disciplinary research projects. Eventually, over time, classification schemes can become naturalised artefacts, which sink into the background, once their history and cultural origin have been forgotten. They may then come to function as *boundary infrastructures*, which by and large function to get the work done, much in line with Gerson and Star's (1986) notions of due process through local closure by means of common representations.

Yet Bowker and Star (1999) emphasise that the process of classification or categorical work is not linear. For instance, boundary infrastructures like international disease classification schemes serve multiple communities of practice across cultures. As such, they may give rise not only to different interpretations as information systems about diseases, but also continue to be open-ended work objects. Classification schemes can become subject to translations to fit particular work contexts, like doctors' and nurses' decisions on what disease categories are relevant in patient case records in a particular hospital, versus translation into international categories of diseases, which can support actors' decisions in domains like immigration authorities, medical research etc. It is argued that any classification scheme can shift from the background, as boundary infrastructure, to the foreground of work, as boundary objects or information systems, as an effect of shifts in actors' assumptions of what kinds of categories may support particular access to knowledge.

4.1.8 Experience: Translation of Interests through a Classification Scheme

My experience with participatory evaluation of a multimedia design project in a public library (Albrechtsen and Jacob, 1998) comprised the design of a classification scheme as boundary object for translating among the interest of diverse communities, from librarians to groups of users (children and adults). In the initial phases of project, a project group of librarians coordinated planning and design ideas according to the existing work organisation of the library, as a community of practice with division of labour according to the actors' educational

and experience backgrounds. In the finishing phase of the project, the project team becomes tightly coupled in daily encounters with design tools and materials to meet an upcoming deadline for launch of a prototype. Likewise, the project group extends activities to collaboration with library users. This phase is characterised by conflict and crises, due to the late deliverance of solution platforms, breakdown of previous collaboration patterns and development of new divisions of labour. This creates tension within the group, as design requires not only a shared vision about how to organise a future artefact, but simultaneously requires specialist knowledge to address design details (Bødker and Grønbaek, 1991; Sonnenwald, 1993). In order to explore and support the actors' negotiations, I suggested the intervention of collaborative creation of a design detail: an iconic classification scheme.

The scheme, which is introduced in chapter 1 of this thesis (figure 1.1), was initially built up from the actor's implicit knowledge, gained in their experience with mediation of library materials to users. Part of this implicit knowledge had previously been represented in posters, created by the librarians and exhibited in the library to guide library users to groupings of materials that went across the standard orderings of materials, according to disciplines and types of materials (books, journals, sound recordings, etc.). During the design of the new classification schemes, the librarians grouped indexing terms for the multimedia information in the database in clusters. Each cluster was subsequently represented in the visual style of the posters in the library. In Star and Griesemer's terms (1989), the iconic classification scheme came to function as a boundary object to translate the interests of divergent social worlds into a coherent semantic structure. Previously, the boundary infrastructure, implemented in codes of practice like standard database formats and classification schemes, had provided temporary closure of the open work system in the library, to get the daily work of cataloguing, indexing and information services done. Thereby, the mediating practices in the library had previously operated within the tension of a singular world-view embedded in universal structures of knowledge of the library's standard classification scheme and the librarians' tacit knowledge about the heterogeneous interests among the library's users. Following the experiment with design of the multimedia information retrieval system Database 2001, the classification scheme was singled out as a separate artefact, which became naturalised in the children's library as a tool for the librarians' selection and monitoring of materials for children.

The experience with collaborative design of a classification scheme for collection mediation showed that translation of interests amongst actors is no trivial matter, but challenge conceptualisations amongst collaborating actors and thereby may create conflict (Albrechtsen, 1997; Albrechtsen and Jacob, 1998). Classification schemes are not innocent inscriptions of translations. Their design requires ongoing confrontation of actors' perspectives in mutual critical assessment of sketches and mutual probing of actors' credibility. A classification scheme can

then be the outcome of negotiated order, or the outcome of biased and misrepresented order (cf. eg., Robinson and Bannon, 1990).

Bowker and Star's process-oriented understanding of classification schemes would seem to support my partial thesis (1) that use potentials evolve in ongoing conceptualisations amongst the actors. Such process-oriented understanding of classification schemes implies that use potentials represent a dynamic dimension of work domain semantics, which may not be inscribed in existing features of a work domain. At the same time, the experience from collaborative design of a classification scheme points towards the direction that ongoing conceptualisations do happen in the actors' social-material encounters, where tacit knowledge comes to the foreground. Likewise, following Bowker and Star's 'lifecycle' for boundary objects, the classification scheme becomes naturalised as a tool for selection of children's materials, sinks to the background, but simultaneously comes to reside in another kind of use context. Yet the case may also point towards a different kind of understanding, based on sociological communication theory, namely that conceptualisations in collaborative collection mediation may be the outcome of actors' construction of intersubjective spheres (Habermas, 1984). The following will explore this line of thought through Schmidt's (1994) sociological concept of 'mechanism of interaction' and the experience case of collaborative film indexing (Albrechtsen, Pejtersen and Cleal, 2002).

4.1.9 Collaboration and Symbolic Artefacts

The concept of 'boundary object' is widely applied as a design concept in the literature of social informatics to enhance semantic integration in large and heterogeneous information structures (eg., Monteiro and Hepsø, 2002). From a work-analytical point of view, however, the notion of 'boundary object' may not be a sufficient concept to understand how conceptualisations among actors develop towards the design of artefacts like classification schemes. The extension of the original definition of boundary objects is broad (Star and Griesemer, 1989). Boundary objects can be repositories of things (libraries and archives). They can also be ideal types for things like diagrams and abstract categories like the biological species, coincident boundaries between entities like physical terrains like the state of California and different maps of the same terrain, and standardized forms like protocols for field data gathering. Are all common resources in the work domains of collection mediation then potential boundary objects? And if so, where does that leave classification schemes in the work domain semantics of collection mediation?

The concept of boundary object has gained a broad and long-term interest within social informatics addressing large infrastructures. The informatic concern of collection mediation is, however, more modest (eg., Bishop and Star, 2000; Thorshauge et al, 1997). In the practice of collection mediation, large-scale technology like the Internet is applied as an infrastructural technology for

providing electronic access to collections of documents, not as a technical work object per se. The collection mediation domain is characterised by *communicative interaction* amongst staff within and across collection sites and between staff and users of collections (Pejtersen et al, 2001; Hertzum et al, 2002). Even though some collections, like research archives, are mandated to publish about generalised findings in scholarly contributions like filmographies and articles, a main activity is mediation through indexing, classification, searching and distribution of materials to users. This means that the outcomes of collection mediation are not always documented, except as, for instance through statistical information about numbers of requests from users or numbers of documents circulated. Collection mediation requires continuous concerted action and communication amongst actors, in addition to professional education and training. In order to reduce the complexity of such activities, the actors create and apply mediating artefacts (Schmidt, 1994).

According to Schmidt (1990), work is inherently a social phenomenon. It requires socialisation of actors to enter the work force, i.e. formal socialisation through education and training in professional standards, and more informal socialisation irrespective of professional specialty, through learning how to concert individual knowledge and activity within a particular sphere of production. Collaborative work arises because of the limited capacity in single human actors, that is, in cases where work cannot be accomplished efficiently to meet the goals of a particular sphere of production. In loosely coupled work domains where the outcome of activities cannot be measured according to technical norms and where social norms are adapted to an evolving sphere of production, actors' activities are interdependent and thus comprise concurrent communication. It is Schmidt's original claim (1994) that in such settings, actors rely not only on communication with one another, but also on artefacts which can mediate interpretations of work activity.

The concept introduced for such mediating artefacts is 'mechanism of interaction' (Schmidt, 1994). As artefacts, mechanisms of interaction are available and consistent, that is, they are not solely in the head of the actors involved, but exist independently of task situation. They make the state of work visible to all (distributed) actors at any given moment. In their role as symbolic, mechanisms of interaction can be applied and developed independently of the state of the field of work. As standardised formats, mechanisms of interaction simultaneously constrain and support articulation of work and provide for an ongoing visibility or representation of the state of work articulation.

4.1.10 Experience: Classification Schemes are not Mechanisms of Interaction

In the experience of collaborative indexing at a European film archive (Albrechtsen, Pejtersen and Cleal, 2002), I apply cognitive work analysis to address collaborative task situations where film archivists and representatives of the archive's user communities watch, discuss, analyse and index feature films. Film screening and discussion take place in a collective, synchronous work arrangement. Indexing and description takes place in a distributed collaborative activity, where archivists use a common database format for entering descriptions of each film at hand. In Schmidt's original definition (1994), the database format is a mechanism of interaction. It is consistently available and exists independently of task situation. It can be used and developed independently of the state of work. It constrains and supports articulation of work and provides for an ongoing visibility of the state of work. As a mechanism of interaction, then, the database format and its associated protocol (cataloguing rules) is applied to reduce complexity of the distributed collaborative work of film indexing and description. The study of collaborative film indexing (Albrechtsen, Pejtersen and Cleal, 2002) found that the representations of films, inscribed by use of the mechanism of interaction, primarily capture the archive's prevalent interest in film preservation and curation, i.e. to describe a film in its unique materiality, rather than its interpretive contexts.

A suggested intervention is introduced in (Albrechtsen, Pejtersen and Cleal, 2002), in the shape of a conceptual tool, which is intended to support actors' negotiations, and understandings of interpretive contexts of a film. As Schmidt (1994) points out, the design of symbolic artefacts is a secondary interpretation of work. Design of symbolic artefacts is different from specifying details of work, yet it is directed to support a particular purpose. In the experience case, the purpose of the suggested conceptual tool is to capture the actors' evolving conceptualisations throughout film discussion, analysis and indexing and to support their intersubjective construction of knowledge. It is conventionally assumed that analysis and interpretation of documents during indexing is an individual cognitive activity (eg. Lancaster, 1991). Intersubjectivity in indexing has previously been addressed mechanistically through measuring 'interindexer consistency' in controlled experiments (eg., Leonard, 1977). 'Interindexer consistency' presupposes the notion that high-ranking overlaps between different indexers' terms imply consensus and hence, validity of the high-ranking indexing statements as opposed to singular statements.

As for instance Cooper (1969) and Rasmussen (1992) have argued, the validity of indexing statements is not a matter of mechanistic consensus, but is a consequence of the right of the better argument. This implies that the quality of indexing is a socio-pragmatic issue, which cannot be measured by crude statistics. Taking this argument to the communicative action of collaborative indexing, I would advocate that quality of indexing also presupposes informed intersubjective

consent. According to Habermas' socio-pragmatic theory of communicative action (Habermas 1984, 1987), intersubjectivity in communication is not equivalent to a match between individual actors' conceptions and terminology. Rather, it is the outcome of open-ended discussions where the actors not only contribute with individual perspectives, but also translate perspectives into a shared discourse, which is entered through the construction of intersubjective spheres.

In the experience case of collaborative film indexing, the intersubjective sphere is entered through socialisation of archivists in their daily practice of indexing and cataloguing and their networking with film specialists. It is this intersubjective sphere of shared stable as well as emerging values for quality and validity of film indexing, which functions as a medium for communicative interaction amongst the actors, rather than their mechanism of interaction. The existing 'mechanism of interaction' (the database format and cataloguing rules) constrains the actors' sense making of films and functions as an explicit model for due process to get the work done. This means that the actors are continually mandated to coordinate with one another in translations of intersubjective conceptualisations to the standardised work procedures, embedded in the mechanism of interaction.

The classification scheme, which is suggested as a conceptual tool to support collaborative indexing in the experience case (Albrechtsen, Pejtersen and Cleal, 2002), is not a mechanism of interaction. It complies with the following conditions in Schmidt's definition: it is symbolic, exists independently of task situation and can be developed independently of the state of work. But it does not conform with the additional conditions: it does not make the state of work visible to the actors, and does not mediate particular procedures for actors' decisions, like the database format and its protocol (cataloguing rules).

Andersen (1994) has previously explored the use of classification schemes for the work practice of technical documentation and has found that classification schemes do not comply with all conditions of Schmidt's original definition (Schmidt, 1994). Andersen defines classification schemes as *symbolic resources*, which reflect the dynamics of an assumed conceptual structure of a common workspace, which is a shared field of work involving collaborating actors (Andersen, 1994, p. 136-137). One design consequence that can be drawn from Andersen's observations and the experience with collaborative indexing (Albrechtsen, Pejtersen and Cleal, 2002) is to analyse the emerging conceptual structure of actors' common workspaces.

4.1.11 Implications for Work Centered Design of Schemes

From the above discussions of key contributions to a social theory of classification in work domains, it can be stated that classification schemes are symbolic artefacts, which reflect work domain semantics. It was discovered that work domain semantics is at the core of the contributions in the literature on classification schemes as sociological phenomena, but that work domain semantics is understood

differently. Domain semantics can be regarded as a conglomerate of relatively stable intersecting social or epistemic interests (eg., Hjørland, 2002). Domain semantics can be regarded as emerging through collective agency, that is, local negotiations and translations of interests in domains (Star and Griesemer, 1989; Bowker and Star, 1999). Domain semantics can be understood as evolving in actors' coordination in work through artefacts (Schmidt, 1994; Andersen, 1994). 'Use potentials' appears to concern a meta-level of domain semantics and to be related to the creation of semantic coherence in collection mediation. The creation of semantic coherence is a dynamic process, which is carried out in actors' work activities of classification, indexing and information searching.

A preliminary conclusion at this stage is: *If use potentials of collections shape the structures and contents of classification schemes for collection mediation (main claim of this research), and use potentials are not only given by existing features in work domains, then work centered design of classification schemes must not only build from analysis of invariant features of work domains. Work centered design of classification schemes would then also need to build on analysis of dynamic properties of collection mediation.*

4.2 Cognitive Work Analysis and Work Centered Design of Classification Schemes

The second research question addresses the use of cognitive work analysis for modelling classification schemes: *What are the possibilities and limitations of cognitive work analysis in the development of classification schemes for collaborative collection mediation?*

Work centered sociological studies of classification schemes indicate that their design is a collaborative task (Bowker and Star, 1999; Schmidt, 1994). The framework of cognitive work analysis regards design of classification schemes as the creation of conceptual structures that support decisions and strategies during use of information systems in work (Pejtersen, 1989; 1994). The approach emphasises the *contexts* of work, where classification schemes can be of support for actors' work activities. Within the paradigm of context-aware computing, Dey et al (2001) define context as any information that can be used to characterise the situation (identity and state) of entities (persons, places or objects) that are relevant for interaction. Cognitive work analysis is concerned with such immediate context information in work domains, but in addition focuses on the semantics, or in the terminology of Rasmussen, Pejtersen and Goodstein (1994), the deep structure of work.

In the previous section 4.1, I discussed understandings of classification schemes as sociological phenomena. Despite the differences in recent understandings of classification schemes in the social practice of work, a shared notion is that

classification schemes are enmeshed in work domain semantics and may be related to ongoing collective construction of coherence in domain semantics. Two main aspects of domain semantics challenge work centered design of classification schemes: invariant semantic properties of work domains and dynamic semantic properties of work domains. Cognitive work analysis handles collection mediation semantics as actors' cognitive resources and subjective preferences for navigation of knowledge within the constraints and opportunities of a territory of work (Pejtersen and Rasmussen, 1997; Pejtersen, 1994). Shared preferences amongst teams of actors for navigation of knowledge are a focus in more recent contributions that address collaborative information retrieval in teams of software developers (eg., Pejtersen and Fidel, 2002) and teams involved in industrial design in the domain of concurrent engineering (Pejtersen et al, 1997). These recent contributions apply the methods of mean-ends modelling of the work domains together with analysis of collaborative task situations where actors engage in information searching activities. The experiences in this thesis focus on the collaborative task situations of film indexing and collaborative design of a classification scheme (Albrechtsen, Pejtersen and Cleal, 2002; Albrechtsen and Jacob, 1998).

In the following, I will address my partial thesis (2) that conceptualisations develop from actors' social-material encounters and intersubjective constructions of knowledge during collection mediation activities. This will be attempted through exploring the means and ends of the common workspace of collaborative film indexing and introducing distinctions between the physical resources, as applied by the actors during decision-making activities in the joint task. Sections 4.2.2-4.2.6 deals with the means and ends of the common workspace of collaborative film indexing and focuses on the meanings and affordances of the physical resources in the workspace for the actors' decision-making. Sections 4.2.7-4.2.10 launch a preliminary model of a common workspace for collaborative activities in collection mediation. In this workspace, classification schemes can be configured together with other symbolic resources for collaborating actors. Analysis of implicit structures in interactive communication and their possible transformation to design of a classification scheme is addressed. In the last section 4.2.9, I summarise the main results of chapter 4.

4.2.1 Analysing the Deep Structure of Collection Mediation Domains

Cognitive work analysis suggests that work domain semantics has a deep structure, and that such deep structure can be made explicit in conceptual tools like classification schemes (Rasmussen, Pejtersen and Goodstein, 1994, pp. 123-134; Pejtersen and Rasmussen, 1997). By 'deep structure' is not meant an intrinsic invariable structure, but rather, relatively stable properties of how the actors interpret work domain semantics. Cognitive work analysis has been applied in the creation of a classification scheme for fiction mediation, implemented in the iconic

interface of the Book House (Pejtersen, 1989). The classification scheme consists of (i) an attribute structure (with five attributes) and (ii) a collection of fiction concepts, grouped in a cluster structure. The attribute structure builds on means-ends analysis of the author domain, the library domain and the user domain (Pejtersen, 1994; Rasmussen, Pejtersen and Goodstein, 1994). As mentioned in chapter 2, section 2.2, this attribute structure can be used for analytical search of fiction and for subject analysis of fiction during indexing. The cluster structure builds on controlled experiments involving library users' and librarians' recognition of fiction concepts in iconic representations of indexing terms for fiction (Pejtersen, 1994). The cluster structure can be used for navigating fiction concepts.

Thus, two methods contributed to the design of classification schemes in the Book House: (1) *means-ends representation* of domains, for the attribute structure, and (2) *experimental studies* of the actors' associations to fiction concepts, for the design of the cluster structure. In the experience cases of this thesis, the focus is on collaborative indexing and collaborative design of a classification scheme. In the first case, an attribute structure has been suggested in the shape of a common conceptual tool (Albrechtsen, Pejtersen and Cleal, 2002). In the second case, a concept cluster structure was designed *collaboratively* by a group of librarians and users (Albrechtsen and Jacob, 1998). If structures of classification schemes can be shaped as means-ends attributes, then how can the attributes be derived? How are the properties and domain substance derived?

The following (section 4.2.2-4.2.6) will explore the common workspace of collaborative film indexing in order to take a closer look at the domain substance in means-ends modelling of this work function.

4.2.2 Means-ends Modelling of Collaborative Work

A means-ends representation of collection mediation at a national film archive was presented in chapter 2 (figure 2.2). The means-ends representation was modelled by use of the generic abstraction hierarchy (Rasmussen, 1986). Domain substance was categorised in five levels, from overall goals to physical properties. The means-ends abstraction hierarchy can be used to model the substance of an entire work domain. For analysing specific task situations, a subset of domain substance in an overall means-ends representation can be extracted to illustrate the territory of work in focus for the actors in a task situation. Petersen (2002) and Petersen and Nielsen (2001) have applied the means-ends abstraction hierarchy for modelling the territory of work of the task situation of manoeuvring of a vessel in the maritime domain.

| MEANS ENDS LEVELS | PROPERTIES REPRESENTED |
|--|---|
| Goals | Develop international awareness of Czech film culture Mediate cultural heritage |
| Constraints | Codes of practice for film cataloguing |
| Priorities | To reach broad and diverse target group of users |
| Functions | Production and integration of new film knowledge |
| Processes | Two groups meet weekly to watch and discuss films. Staff write database entries |
| Resources and physical configurations | Full length feature and short films. Computers. Database. Local IT-network environment. Database format. Cataloging rules. List of indexing terms. Staff. Experts outside the archive |

Figure 4.2.1 Means-ends representation of collaborative film indexing at a national film archive

Figure 4.2.1 is a means-ends representation of the common territory of work for the collaborative task situation of film indexing. The domain substance of the common territory of work is derived as a subset from an overall means-ends representation of a national film archive (figure 2.2). Consider the following entities of domain substance in level 5 (Resources and physical configurations) from the means-ends representation of collaborative film indexing (figure 4.2.1):

- Full length feature and short films
- Database format
- List of indexing terms
- Databases
- Computers

The films in the archive’s collection exist as physical entities, but also articulate meanings, ideologies and values of the culture where they are produced. They are work objects in the archive for functions like information searching, indexing, restoration and preservation. During such work functions, new knowledge about the films is produced. The database format and its associated protocol (cataloguing rules) is a mechanism of interaction (in Schmidts’ original definition) that mediates

standard work procedures for the distributed collaborative work processes of film indexing and description. The list of indexing terms exists as a physical document, but also mediates the semantics of a shared vocabulary used to index and search for films. The databases and computers reproduce the outcomes of actors' work activities of film indexing and description, but also make visible the state of work within the constraints of their organisation. It appears that domain substance in the category of physical properties (level 5 of the mean-ends representation in figure 4.2.1) mediates meanings, and that these meanings are divergent. How are the meanings of things captured by cognitive work analysis?

4.2.3 Meanings of Physical Resources

Rasmussen, Pejtersen and Goodstein (1994) state that domain substance within the five abstraction levels is linked through propagation upwards and downwards in the hierarchy. Thus, in the indexing example above, the physical resources of a database and a database format (level 5) would be linked upward through the work process of entering film information (level 4). Yet the physical resource of a film is also linked upward through the work process of entering film information. A film is a work object throughout all work processes involved in the function of film indexing. According to the notion of propagation between levels in the means-ends abstraction hierarchy, physical resources are linked through their use in the same processes and functions. Such linkage does not reveal what the physical resources may symbolise or mean to the domain actors.

In their exploration of the means-ends abstraction hierarchy for the function of manoeuvring of a vessel in the maritime domain, Petersen and Nielsen (2001) have found that the meanings of the things or physical resources, which shape part of the physical environment (level 5 of their means-ends representation), are not invariantly given by their means-ends relation to the processes they serve (propagation upward to level 4). Rather, it is found that the meanings of things can be dependent on the state of the immediate environment. For instance, the meaning of a rudder for control of water flow is not dependent on processes of the upper level (turning, etc.), but on the state of water flow, i.e. dependent on the state of other things (at means-ends level 5). Thus, the semantics of a rudder is not invariant, but dependent on the situation at hand during actors' manoeuvring of a vessel.

The work domain of a film archive is a living system. As a living system, it is open and maintains and reproduces itself within the constraints of its environment. Yet this work domain is not equivalent to a natural organism or environment. It is a symbolically structured life world (Habermas, 1987). It is inhabited and reproduced by human actors, whose activities contribute to the states of the life world. I choose to use the notion of symbolic structuring from the perspective of the actors: the actors' use, construction and interpretation of common symbolic resources, which structure and reproduce the life world of collection mediation. Within the life

world of collection mediation, actors' decision-making is dependent on their socialisation in the sphere of production. That is, actors' understanding of common goals, constraints and possibilities, and how other actors have acted or will act, with respect to such factors. Common resources inhabit the life world of collection mediation. Some of these resources contribute to the actors' reproduction of the life world. The resources, which the actors use to reproduce the life world of collection mediation, are common symbolic resources that can be maintained independently of the state of work. Cognitive work analysis builds on Gibson's notion of affordances (1979) to deal with the use of symbolic resources for activity in work domains (Pejtersen, Rasmussen and Goodstein, 1994; Gibson, 1979).

4.2.4 Experience of the 'Affordance' Concept in Work Domains

The experience case of Gibson's notion of affordances (Albrechtsen, Andersen, Bødker and Pejtersen, 2001) explores applications of the concept in activity theory (eg., Bødker, 1991) and cognitive work analysis (eg., Rasmussen, Pejtersen and Goodstein, 1994). Recently, Bærentsen and Trettvik (2002) have argued, from an activity theory point-of-view, that affordances of artefacts or tools can either be an effect of their possible uses or intended use in particular social practice. As an example in point, they mention a hammer, whose intended use is to beat nails into solid materials like walls, but whose possible use cannot be predicted. Social convention or assumed commonality of use would then define the intended use of the hammer. The experience case of Gibson's notion of affordances explored the concept with the aim of discovering if affordances may be related to deep semantics of work domains (Albrechtsen, Andersen, Bødker and Pejtersen, 1991). That is, not deep semantics by convention only, but also deep semantics in the sense: action possibilities in work environments.

It is a shared notion by activity theory and cognitive work analysis, that actors' activities in work domains are purposeful and goal-directed. According to activity theory, human activity is always mediated, for instance through language and artefacts (Leont'ev, 1978). Cognitive work analysis shares the notion that activity is mediated through such cultural constructs, but adds the conception that actors progressively build up an internal world model as a conceptual structure, which mediates a situational coupling between and actor and the work environment and allows them to perceive action possibilities in the environment (Rasmussen, 1983). Actors' perception of affordances in work environments is then dependent not only on convention or possible uses of particular tools, but on the actors' internal world model, built up during the actors' experience, social interaction and formal and situated learning. The affordances of a hammer, then, would not only be dependent on its use by social convention or unlimited possibilities. The hammer's affordances would be part of the actors' internal world model together with related affordances of other physical objects. The actor's internal world model comprises a conceptual structure of experience and knowledge of physical resources, which

afford hammering. If there is no hammer in the environment, the actor may for instance pick up the affordances of a stone in the environment for the same task. Accordingly, if an internal world model contributes to mediating affordances, then the relations between physical resources are not only given by essential or conventional classifications of 'things', but also by actors' experience of the affordances of things in the environment.

4.2.5 Affordances of Symbolic Resources in Collection Mediation

Petersen and Nielsen (2002) argue that the meaning of physical resources in the maritime domain, like a rudder, is not statically given by their means-ends relation to processes that they serve, through the propagation upward in a means-ends hierarchy. In section 4.2.2, I made similar observations with respect to the physical resources in the means and ends of collaborative film indexing. Each physical resource in level 5 of the means-ends representation is related to level 4 of processes (figure 4.2.1). However, that does not account for the possibility that the meanings of these things may be divergent, not only in an essential sense, but in a pragmatic sense, i.e. from the actors' point of view. Are the relations between the physical resources in collaborative film indexing given by their affordances? How do the actors construct and pick up affordances of physical resources in the social practice of indexing?

In section 4.1.10, I discussed the notion of 'mechanism of interaction' and classification schemes and concluded that classification schemes do not comply completely with the definition of 'mechanism of interaction' (Schmidt, 1994). Based on empirical analysis of cases in technical documentation, Andersen (1994) has refined the definition of mechanism of interaction and introduced distinctions between physical resources in common workspaces. The distinctions are formulated according to objects of articulation work in the domains, and how these objects are referenced symbolically in the domain and through their functions in the domain. The objects of articulation are (a) conceptual structures, (b) technical resources, (c) information resources, (d) material resources and (e) infrastructural resources (Andersen, 1994, pp. 139-163). For instance, Andersen argues that shared conceptual structures are referenced symbolically in classification schemes, and that the associated functions are to categorise, to relate, to specify and name concepts and objects in the common workspace. These distinctions constitute the basis for a model with an ordered ontology of physical resources in the domain, which sustain different work tasks in the technical production domain, from development to documentation of products. However, the ontology does not build on the notion of possible affordances for action that the resources may have. That is, Andersen's ontology does not articulate possible *uses* of the resources.

In the following categorisation of symbolic artefacts in the common workspace of collaborative film indexing, I retain Andersen's notion that physical resources in work domains are symbolic of particular segments of a common workspace.

Nonetheless, I suggest a grouping according to *use purpose* (figure 4.2.2). Figure 4.2.2 is a suggested categorisation of current physical resources for collaborative film indexing at a national film archive (Albrechtsen, Pejtersen and Cleal, 2002). The functions for collaborative film indexing are (1) production of new film knowledge and (2) integration of new film knowledge (cf. means-ends representation in figure 4.2.1, level 3). Production of knowledge involve the joint decision tasks of:

- Film screening and analysis
- Discussion of subjects in films
- Formulating subject expressions
- Formulating descriptions of films
- Assigning indexing terms
- Assigning descriptions of films
- Entering film information in database

Figure 4.2.2 categorises the physical resources which support collaborative decision tasks into (1) information processing tools (computers, database, IT network environment) and (2) representation tools (cataloguing rules, database format, list of indexing terms). This categorisation is done from two perspectives: (a) their affordances for decision-making about production of new film knowledge and (b) their mode of integration of new film knowledge. The affordance perspective addresses ongoing situational coupling between the actors and the tools during decision-making, that is, what the tools mean to the actors when they are faced with a particular decision during work. The mode of integration perspective refers to how the tools mediate the state of work and the work procedures to the actors engaged in the common workspace.

| Segments of physical resources | (1) Main instances of physical resources | (2) Affordances for actors' production of film knowledge | (3) Mode of integration of film knowledge for common work space |
|-------------------------------------|--|--|--|
| Information processing tools | Computers Database IT-network environment | Entering information Storing information Information exchange | Mediation of the state of work |
| Representation tools | Cataloguing rules Database format List of indexing terms | Deciding procedure for film inscription Translating procedure to attributes Maintaining terminological consistency for film content inscriptions Assigning indexing terms | Reduction of conceptual complexity through mediation of standard work procedures and terminology |

Figure 4.2.2. Affordance-oriented categorisation of physical resources, currently supporting actors' joint task of indexing and describing films at a national archive

4.2.6 Linking Representation Tools to Means and Ends of a Common Workspace

The original definition of 'mechanism of interaction' states, as its characteristic 3 of symbolic artefacts: "the symbolic artefact must provide affordances to and impose constraints on work articulation and it must make the state of work articulation at any given moment publicly perceptible" (Schmidt, 1994, p. 65). I agree with Schmidt that artefacts for mediation of collaborative work are symbolic, that is, that they mediate meanings and are made for joint conceptualisations amongst the actors. Within the common workspace of collaborative film indexing, some

symbolic artefacts mediate the state of work (computers, database, IT-network environment). It is however questionable if the symbolic resources of cataloguing rules, database format and list of indexing terms directly mediate the state of work. Cataloguing rules, database format and list of indexing terms do not make the state of work perceptible during collaborative film indexing. Rather, they make it possible for the actors to negotiate and manage the state of work.

My suggested affordance-related categorisation of symbolic resources (figure 4.2.2) implies also another angle, which has to do with the affordance concept itself. Bærentsen and Trettvik (2002) define affordances as an effect of possible or intended uses of natural or constructed objects. The affordances listed for information processing tools and the representation tools in figure 4.2.2 involve two different types of affordances. The information processing tools do have the affordances listed for collaborative indexing, but in other task situations, they may offer different kinds of affordances. For instance, the computer not only affords entering of information. Its set-up can also yield affordances for other work processes at the same workplace, like writing e-mails, constructing web sites, etc. For such tools, affordances go beyond their application in a concrete common workspace.

Conversely, the affordances of representation tools (cataloguing rules, database format, list of indexing terms) are, in Bærentsen's and Trettvik's (2002) terms, intended affordances. The intended affordances of the representation tools are to support the actors in dealing with the conceptual problem of how to interpret a film and how to translate interpretations. The representation tools have been constructed with such particular purpose. The actors, who inhabit the life world of collection mediation, constructed the tools. The representation tools are, in means-ends terms, linked to the codes of practice defined for collaborative film indexing (level 1 of the means and ends of the common workspace: constraints). The codes of practice, then, symbolically constrain this subset of the life world. The representation tools are symbolic artefacts that structure and reproduce the semantics of the common workspace. Their *affordances* are dependent on the codes of practice defined for structuring and reproducing the semantics.

The codes of practice for collaborative film indexing at the national archive are documented in writing in the archive's filmography of Czech films (Opela, 1998). In addition, the codes of practice are maintained and mediated more informally by the director of the archive (Pejtersen et al, 2001). The codes of practice guide the research process of film indexing and filmography production as technical norms in Merton's sense. The intent is to ensure high quality of the historical documentation of Czech films. The experience of identifying codes of practice at the archive in interviews with staff and studies of the written documentation (Pejtersen et al, 2001) found that an important quality criterion is to publish good descriptions, but especially to ensure that these descriptions address a diversity of target groups for Czech films. This latter aspect is also related to the means-ends level of priorities (level 2) for the collaborative indexing task. Due to this priority,

the archive carries out all processes of film indexing and description together with outside representatives who contribute with knowledge and perspectives to address the needs of the diverse target groups for the filmography. The outside representatives participate in screening meetings and discussions in a collective, synchronous work arrangement. They are also consulted during indexing and description of each film by the staff in a distributed work arrangement. Yet the outside representatives do not have access to use the representation tools of the common workspace and are not involved in maintaining the overall codes of practice.

The aim of collaborative film indexing is to arrive at an intersubjective understanding about the content and possible uses of a film for different target groups. This implies that the actors' contributions and sharing of mutual perspectives are the primary means for conceptualisations in the decision task of film analysis. In Habermas' terms, collaborative film indexing operates primarily from a humanistic research-motivating interest, whose primary medium of knowledge production is communication through language, and whose primary type of knowledge produced is interpretation and understanding (cf., eg., Habermas, 1972). This aspect is embedded in the level of Goals (level 1) in the means-ends representation in figure 4.2.1: develop international awareness of Czech film culture and mediate cultural heritage.

At the same time, the collaborative indexing activity operates within the procedural constraint of closure. Closure means that the actors' discussions and negotiations have to result in a final statement in the shape of subject expressions and indexing terms for a film. How do the archive's tools of representation support the humanistic research motivating interest of interpretation and understanding, implied in the high-level goal for collaborative film indexing, together with the requirement for closure? This question is addressed in two ways in the following. In section 4.2.7, I introduce a preliminary model of a common workspace for collaborative film indexing in order to address the notion of context for classification schemes as symbolic artefacts for this collaborative work task. In section 4.2.8, I address the challenge of rooting the structure of a classification scheme in deep semantics of the domain.

4.2.7 A Preliminary Model of a Common Workspace for Collaborative Film Indexing

Based on the observations in sections 4.2.2-4.2.6, it can be assumed that symbolic resources constructed and used by actors in collection mediation are important characteristics of common workspaces for collaborative tasks in the domain.

From a means-ends analysis of the common workspace of collaborative film indexing, I proposed a distinction between physical resources (level 5 of the means-ends representation in figure 2.2):

- Symbolic resources
- Actors
- Work objects

This distinction is retained in the suggested model of a common workspace for collaborative film indexing below (figure 4.2.3). The model of the workspace partitions the means-ends representation (figure 2.2) into the following elements:

- Data processing tools (level 5)
- Representation tools (level 5)
- Actors (level 5)
- Common goals/constraints (level 1) and priorities (level 2)
- Situational activity (level 3 and 4)

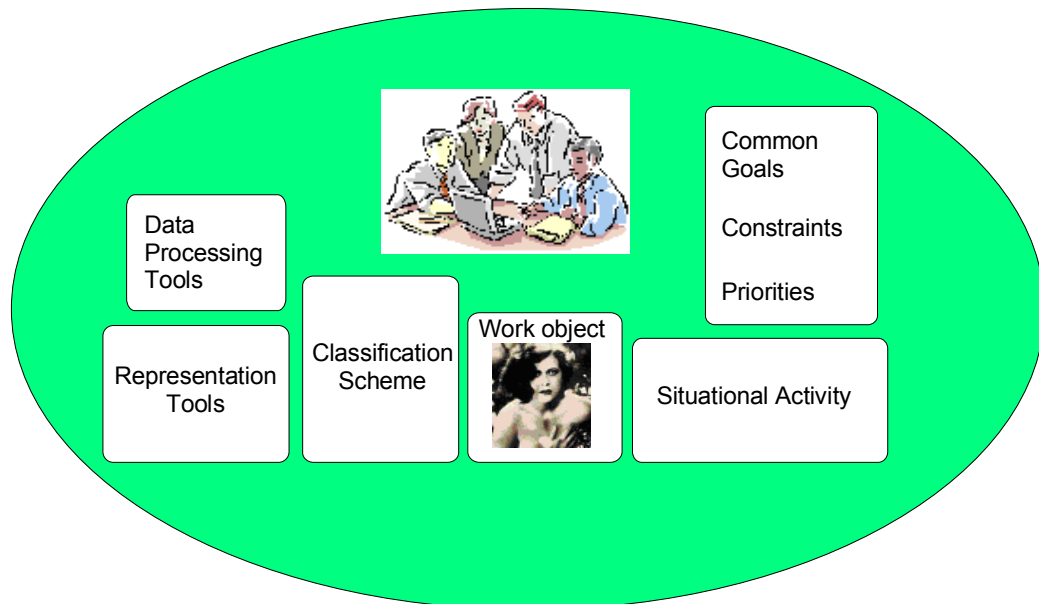


Figure 4.2.3. Common workspace of collaborative film indexing

The model (figure 4.2.3) also covers the symbolic resource of a classification scheme. This element is placed within the common workspace in order to illustrate a possible support by a classification scheme for the collaborative task. In the case of collaborative film indexing (Albrechtsen, Pejtersen and Cleal, 2002), film

interpretations are performed orally in discussions at screening meetings and are subsequently represented as indexing terms. Closure is reached by use of representation and data processing tools. But it is also reached through subsequent oral consultations amongst staff and between staff and users.

Joint situational activities in the common workspace are:

1. Screening and discussing a film
2. Negotiating interpretations of a film
3. Negotiating closure of interpretations
4. Creating subject expressions
5. Translating subject expressions to indexing terms
6. Entering indexing terms into a database

In the experience case (Albrechtsen, Pejtersen and Cleal, 2002), joint situational activities 1-3 took place in a synchronous work arrangement at a meeting, involving staff and outside experts. In contrast, the joint situational activities 4-6 took were distributed and involved ad hoc consultation amongst staff and between staff and outside experts. The symbolic resources of data processing tools and representation tools were only used in the situational activities 4-6. The proposed model of the common workspace shows the overall context of work: symbolic resources, work object, actors and their goals/contraints and priorities. Together, these elements constitute the possibilities and constraints for action in the common workspace.

4.2.8 When is a Classification Scheme a Conceptual Tool?

In my partial thesis 2, I argue that conceptualisations evolve from actors' social-material encounters and intersubjective constructions of knowledge during collection mediation activities. In sections 4.1.10 and 4.2.5, I discussed how joint interpretations of films in collaborative film indexing are currently brought to closure by staff's use of representation tools and data processing tools so that the filmography gets produced. First of all, actors' conceptualisations are brought out orally, in communicative interaction at screening meetings and during ad hoc consultations amongst actors. The collaborative task is performed within an intersubjective sphere, which is shaped by the actors' commitment to contributing to national film research. Likewise, the actors' conceptualisations are brought out during encounters with artefacts.

In the task situation of film screening and discussion, the encounter with a film brings up themes for the oral discussion and draws out the actors' film knowledge. In the task situation of bringing film interpretation to closure in indexing statements, some of the actors (the staff) encounter the materiality of representation tools and data processing tools. The experience case identified a semantic drift in

interpretations and argued that this was due to the lack of affordances in current artefacts to express an overall conceptual view of the semantics of a film. Based on this argument, I drafted a classification scheme as a conceptual tool, intended to make visible a possible shared semantic structure (Albrechtsen, Pejtersen and Cleal, 2002). The vision was to address joint film indexing as concurrent through the classification scheme as a *mediating* symbolic artefact, which could be maintained and used independently of the state of *conceptual* work.

In section 4.1.8, I discussed the notions of classification schemes as 'boundary object' and 'mechanism of interaction'. 'Boundary object' was found to be fruitful for characterising ongoing conceptualisations amongst collaborating actors and inscriptions of actors' conceptualisations in classification schemes. 'Mechanism of interaction' was found to be fruitful for characterising artefacts, which make the state of work continually perceptible to all actors involved in a joint task. If the concept of 'state of work' were to be understood in a restricted sense, to be the referent for actors' articulation of work processes and tasks, then classification schemes would be artefacts for coordination in terms of work allocation. Collection mediation does involve coordination of work, for instance in distribution of collection items. Artefacts like computers and networking technology are, for instance, used for this purpose. However, collection mediation also addresses the semantics or the contents of collection items and users' needs, and especially during information searching, indexing and classification of collection content. This is a different kind of semantics, which has to do with the knowledge content of collections and its possible uses for diverse user communities. I suggested the concept of 'use potentials' to designate the evolutionary and socio-material nature of this semantics.

As shown in the previous sections 4.2.2-4.2.6, the shaping of use potentials during indexing happens in oral communication amongst actors and translations of interpretations through symbolic resources in order to reach closure. Evidently, work centered design of classification schemes for this task is faced with a tension between 'use potentials' versus 'closure'. The shaping of 'use potentials' is not directed from one single authority, but unfolds in joint situational activity. Closure, on the other hand, is shaped by means of representation tools and not least, through practices of *what* to include in the outcome of work, and as importantly: *who* to include in such decisions. When collaborative film indexing approaches closure, a small group of staff is responsible for closure. This would seem to echo the concept of 'heterogeneous engineering', coined within actor-network theory to explain how technological projects reach closure (Law, 1994), and the work-analytic concept of 'due process' (Gerson and Star, 1986).

Yet, now bringing in the communicative notion of 'intersubjective spheres', the experience cases of this research indicate that in various degrees, actors meet one another and the materiality of collections with expectations and pre-understandings, which are sometimes actualised as shared understandings and sometimes demand development of new understandings and knowledge. There is

an element of socialisation in collection mediation, which appears to be different for instances where actors collaborate in long-term work arrangements as in the film indexing case, and instances where actors meet in apparently casual encounters. Encounters between users and staff in an archive where they negotiate about information searching may, however, unfold recurrent themes. In the following section, I will illustrate how cognitive work analysis approaches analysis of recurrent themes in actors' encounters and how the results of such an analysis may be transformed to overall design of a classification scheme.

4.2.9 Analysing Implicit Structures in Interactive Communication and Scheme Design

For analysis and design of classification schemes for collection mediation, cognitive work analysis considers situational activities, where explicit semantic structures like existing classification schemes are used to shape interactive communication or where semantic structures arise. In chapter 2, five search strategies identified by cognitive work analysis were introduced together with two types of classificatory structures: an attribute structure to support actors' negotiation of search dimensions and a clustered structure to support navigation in concepts. Structures, then, may arise in two ways, either through negotiation about search dimensions or about concept clustering. The example task situation in a film archive, introduced in Chapter 2, sections 2.1.5 and 2.2.1 is used here to illustrate identification of search dimensions.

A university student of women's studies comes to a film archive to find materials and seek inspiration for an essay on socio-political conditions for women's lives in the Czech Republic during the 1940s. He is especially interested in how national feelings and women's values are represented in the films from that period. The student's research problem is formulated within a research area that is characterised by high degree of strategic uncertainty (Whitley, 1989). In addition, the student is a newcomer to film studies where the essay is going to be submitted for exams. His education background is in literary studies. In his literary studies, he specialised in socio-cultural perspectives of gender and learned to work with critical discourse theory of literature, including Kristeva's notion of intertextuality. The new research area of socio-cultural aspects of gender is presently gaining interest within the field of film studies in his country. Yet so far, the knowledge production is fragmented, with a high degree of diversity in concepts and terminology and research paradigms. The associated research area is characterised by a high degree of strategic dependence between experts, who are contact in informal professional networking in order to move the field forward; there are no journals or textbooks dedicated to the field. In addition to this interdependence in the research community, students and researchers are dependent on expert intermediaries of collections, whose insight into different kinds of media is used to inspire the research. The student's university supervisor has encouraged him to

visit the film archive and also hinted what staff he might want to talk with in the beginning of his exploration of film knowledge.

The high degree of strategic task uncertainty in the student's project implies that it is difficult to formulate a search request on the spot. The student discusses the problem with a staff member. They browse the collection together during their first encounter. In cognitive work analysis terms, they start out with a browsing strategy. This strategy yields some exemplars of films and film-related materials, which the student analyses. At a later visit to the archive, the student talks with another staff member. The student now has a clearer picture of what he is looking for. The staff member listens to the student's ideas and knowledge. The student and staff member progressively enter a shared discourse about the problem. The student adapts his concepts to what he learns about the archive's concepts through the intermediary staff. The intermediary adapts her concepts to the student's and learns something new. The intermediary's background is in drama studies, and so, she has some background in text analysis, but not from a gender studies perspective nor from a critical hermeneutic perspective. As an intersubjective sphere is gradually entered through their shared discourse, the intermediary attempts to translate some of the student's needs into search strategies. One strategy is to go ahead from exemplars that the student found relevant for the problem. This corresponds to search by analogy. Another strategy is to negotiate dimensions and properties of the student's information need. The properties that the student and staff member formulate together are for example:

1. A film's promotion of particular understandings of gender roles, liberation versus tradition. The film director's explicit or implicit affiliation with particular socio-cultural values. Censorship.
2. The plot of a film. The heroes and villains. The ending. The socio-cultural setting of the plot, including place and time. The theme of the film.
3. A film director's narrative techniques, including inter-textual elements, like allusions to other films, myths and texts or intersecting plots.
4. The public reception of the film; contributions by film critics; the film's national or international impact; the career and life of one or two main characters in a film, i.e. how the 'embodiment' of the plot contributes to the message of the film.
5. A film's version, i.e. exists in full or as a fragment; a film's availability or the accessibility of film-related materials within the constraints of the student's time to write and submit the essay.

The list of properties in the student's information needs above could be regarded as constituting five dimensions of the information need, or, in cognitive work analysis terms, the semantic territory that the student and staff member explore and structure together. The properties of these dimensions are interrelated, and each property is important for decisions about the search. If a film or film-related materials are not available within the constraints of the student's deadline, then the student may decide to ask for a short description or an abstract of the materials. If the career of a main character has reflected leaps in the kinds of characters that she has embodied in films, then that aspect may be relevant, but not crucial for the essay. Findings of intertextuality in a film can ease the student's formulation of a methodology for the essay, due to his background in studies of Kristeva's theory. The dimension of subject content like theme and plot is important for finding as many films as possible from where the student can decide an empirical focus. The aspect of the film director's affiliation with particular views and paradigms about gender in society and culture is vital to the search. This is the highest interpretive value, as seen from the student's point of view and his background in critical hermeneutic literary theory. The staff member contributes with the idea of censorship history to reflect the degree of provocation of a film's overall message or elements, which may mirror the director's affiliation with paradigms or values vis-à-vis the socio-cultural values at a particular time.

4.2.10 Transforming Analysis to Overall Design of a Scheme

Pejtersen has promoted the notion that stable properties arising in formulations of information needs between library users and intermediaries reflect hidden structures of semantics in the domain, which can be expressed as attributes, in the manner of a means-ends perspective (eg., Pejtersen, 1989). This notion was based on extensive field studies of communicative task situations in fiction mediation, where recurrent patterns of formulating search dimensions were found. The recurrent dimensions were transformed to a classification scheme with an attribute structure for analytical search of the semantic territory of fiction mediation. In the case of cognitive work analysis of national film archives, related dimensions and properties of interactive communication between users and staff during mediation were found (Pejtersen et al, 2001).

The present research has attempted to explore the idea of recurrent dimensions through suggesting a classification scheme for film indexing, based on cognitive work analysis of collaborative film indexing (Albrechtsen, Pejtersen and Cleal, 2002). The suggested scheme has six dimensions or attributes, which are argued to constitute a means-ends representation of the semantic territory of film analysis and description. From a modelling perspective, i.e. from the point of view of the means-ends abstraction hierarchy (Rasmussen, 1986), the scheme is however not a means-ends representation. It has six levels instead of five. Its properties express substance from different domains, from the film director domain and the film

content domain to archive attributes. A semantic territory is in itself an abstraction. From a formal perspective on abstraction, abstract properties of one dimension cannot be a means to an end of a property in another dimension. For instance, a film distributor's policies may be related to the archive's goals and policies, but is not related to local archival conditions for shelving (referring to possible linkages between level 5 and 6 of the proposed classification scheme). Yet, the empirical analysis found that the actors *perform* these linkages during the task of collaborative film indexing. Thus, the scheme may be relevant as an explicit structure reflecting the hidden structures in actors' communication and decision-making in indexing.

The contribution about the experience advocates that the classification scheme may be relevant for analytical search of films and film-related materials (Albrechtsen, Pejtersen and Cleal, 2002). Section 4.2.8 gave an example of an encounter between an archive user and a staff member of a film archive, where hidden structures were identified. This kind of encounter is recurrent in all three national film archives (Pejtersen et al, 2001). Yet at one national film archive, the most common end-user search strategy found was 'known item search' (Hertzum, 2003). I agree that 'known item' search is a common search strategy in mediation of films and film-related materials, but I also argue that this strategy does not exclude the importance of negotiations about film content when users visit an archive to resolve a complex information need. A known item can function as an exemplar for further exploration of collection materials, for instance. If classification schemes are not available as symbolic artefacts to express recurrent dimensions of requests, then the analytical search strategy is not supported, and may hence not be considered by the users or staff. In the practice of collaborative film indexing at one of the national archives, the database format was used as a conceptual structure for indexing as well as searching. The mediating practice in this archive is that users rarely encounter the staff, but that interaction happens between intermediaries in institutions. Students will, for example, ask a librarian at their university to contact the staff of the national film archive about a search. Thereby, end-users' formulation of information needs in the archive is constrained not only by the lack of proximity to expert staff, but also by collaboration practices between institutions.

For design of classification schemes in collection mediation, cognitive work analysis considers the semantic dimensions of the search activity as well as the indexing activity. Pejtersen's empirical studies of collection mediation were conducted in public libraries, where users meet staff in proximity (eg., Pejtersen, 1989). The experience cases of national film archives (Pejtersen et al, 2001) identified quite different mediation practices in the archives. In one case, the practice of indexing has high priority, in other cases, the practice of interactive communication between users and staff in proximity is emphasised. During indexing, staff members make decisions for closure of the task. Information searching is more open-ended, but at some point, closure is decided in negotiation

between staff and users. Interdependencies are strategic, as opposed to the technical interdependencies in closure of indexing tasks. Pejtersen has addressed a possible reconciliation of diversity in users', intermediaries' and indexers' practice in fiction mediation through a model of the means and ends of the semantic territory of the different domains involved: the user domain, the library domain and the author domain (Pejtersen, 1994). From this model, a common classification scheme for information searching and fiction has been derived and implemented in the Book House (Pejtersen, 1989).

The empirical analyses, shaping the cases of the present research, focus on the practices of collaborative film indexing and collaborative design of classification schemes (Albrechtsen and Jacob, 1998; Albrechtsen, Pejtersen and Cleal, 2002). The empirical analyses have in addition addressed collaborative information searching, but the main approach has been to explore the practice of keeping order in collections. Very little research exists on this practice in collection mediation. The information science community dealing with information behaviour focuses almost exclusively on collection users' search behaviour and concepts (eg., Ingwersen, 1999). Classification research by this community deals with identification of users' concepts during search, and the suggestion is to add on such concepts as surface 'affordances' for users to access existing collection classifications (eg., Bates, 1986). Classification research emphasising a sociological orientation is conducted within several disciplines. The research focuses on understanding classification schemes as social constructs or epistemic structures and has developed new concepts for analysis (eg., Star and Bowker, 1999; Schmidt and Wagner, 2002; Hjørland, 2002). This latter research does not address methods for modelling of classification schemes. Cognitive work analysis comprises concepts and methods for empirical analysis of collection mediation domains and design of classification schemes. It also covers methods for analysing and modelling common workspaces. I have utilised this opportunity of the framework to make a suggestion for a common workspace for collaborative film indexing and have also made a suggestion for a classification scheme to mediate between actors' interpretations and representation and data processing tools (Albrechtsen, Pejtersen and Cleal, 2002).

At this stage, it can be concluded that design of classification schemes for collection mediation requires not only empirical analysis of information searching practice, but also of indexing and classification practice in collection mediation domains. All three practices can be expected to reveal hidden structures of the deep semantics of the domain. Because the purpose of indexing and classification practices is to support information searching, the information search practice is a necessary, but not sufficient unit of analysis for empirical analysis of collection mediation. Actors involved in indexing and classification have expert knowledge about the domain semantics. Their practice is defined by making visible the use potentials of collections. Yet, when actors meet in collaboration about bringing out such potentials, their collaboration may result in biased representations in

classification schemes and indexing statements for the materials. It is therefore necessary to explore dimensions and concept formations in communicative interaction in information searching situations in order to determine how well collaborative indexing and classification practices bring forth the use potentials of collections and how such practices may inspire one another. Through the attempts at modelling the semantic territories of collection mediation, which has been discussed in this section, it can be assumed that it may be possible to identify what kind of structures of collection mediation semantics can be sustained in classification schemes. It can in addition be stated that the semantic structures may be recurrent as well as evolving in the domain. Yet the *stability* of semantic structures in this domain is a construct and relative to their representation in classification schemes and representation tools and their sustenance in actors' discourses.

4.2.11 Summary

This chapter has discussed the two research questions of the present research: classification schemes as sociological phenomena and advantages and limitations of cognitive work analysis for design of classification schemes. Section 4.1 addressed the first question through discussing the findings of the experience cases of the research in the light of selected literature. This part of the chapter discussed domain semantics of collection mediation and found that domain semantics in collection mediation is not equivalent to a static fusion between intersecting goals, interests and knowledge. The semantics emerges through ongoing translations and negotiations in mediating practices. It was concluded that if use potentials of collections shape the structures and contents of classification schemes and use potentials are not only given by existing features of the domain, then work centered design is faced with the requirement of analysing dynamic properties of collection mediation semantics. Section 4.1 addressed the second question of advantages and limitations of cognitive work analysis in development of classification schemes for collaborative collection mediation. This part of the chapter discussed the use of the means-ends abstraction hierarchy for modelling an example case of a common workspace for collaborative film indexing. Distinctions between physical resources in this workspace were introduced together with their affordances for action. A preliminary model of a common workspace was suggested. A classification scheme was introduced, whose purpose was to make recurrent semantic structures visible to the collaborating actors. The translation of task situation analysis to design of such a scheme by use of cognitive work analysis was illustrated and discussed. It was concluded that cognitive work analysis offers models and methods for analysing the territory of work and collaborative task situations in collection mediation. It was also concluded that cognitive work analysis' concept of invariant or stable structures should be understood as being equivalent to recurrent dimensions of the semantic territory, in that the structures are evolving in the

domain and that their stability is relative to their sustenance in classification schemes and intersubjective spheres amongst actors in the domain. In other words, the invariant structures are not essentially given, but only frozen in time in the schemes. 'Use potentials', then, appears to be a fruitful concept to define the design requirements for classification schemes for collection mediation within the methodology of cognitive work analysis.

Chapter 5

Research Methodologies

The research behind this thesis has been carried out in close collaboration with colleagues from different disciplines, especially computer science, library and information science and sociology. The collaboration has taken place in two settings. The first part of my study (from January 1998 to May 1999) was organised at the Royal School of Library and Information Science during my affiliation with the school in the Department of Information Studies. During this part of the study, I was enlisted as a part-time PhD student at the Department of Sociology and Social Anthropology, Keele University, now Lancaster University. In June 1999, I joined the Centre for Human-Computer Interaction (CHMI), as scientist and part-time PhD student in the Department of Systems Analysis at Risø National Laboratory. The second part of the study has been framed within two research projects in CHMI: Project 4: Theories and methods and project 1: Ecological Information Systems. In both these settings, the design and use of classification schemes for collection mediation has been explored through field studies.

Classification schemes are an object of research in many disciplines and cross-disciplinary communities. The present research acknowledges the interdisciplinary nature of the research object. Nonetheless, classification schemes in collection mediation are of a far lesser magnitude than ordering systems developed for theory-building in, for instance, the natural sciences and social sciences (eg., Theologus, 1969; Durkheim, 1963). Yet schemes developed for collection mediation often reflect the structures and concepts of larger schemes and may replicate epistemologies of larger parent schemes. Social constructivism attributes this quality of classification schemes to their interpretive flexibility and ability to travel between different use contexts as boundary objects (eg., Bowker and Star, 1999; Latour, 1988). Discourse theory points towards the existence of collective structures, determining the structures and contents of schemes (in particular, Hjørland, 1997). The first position implies that the schemes have social agency. The second position implies that 'outside' structures like the division of social labour into disciplines determine the structures and contents of schemes. Approaching a definition of classification schemes as sociological phenomena at work in collection mediation requires a close look at agency and structure in social theory. Habermas (1984; 1987) has, amongst other theoreticians in sociology,

notably, Bourdieu (1972) and Giddens (1984), promoted frameworks for synthesising previous social theories of agency versus structure.

Classification schemes are also an object of research in communities that deal with understanding and designing artefacts in work domains. Within the community of computer-supported collaborative work, Schmidt (1994) and others (Andersen, 1994) have explored classification schemes as artefacts for coordination of work. Cognitive work analysis has addressed design of classification schemes, based on work domain analysis (Rasmussen, Pejtersen and Goodstein, 1994). For the research in this summary and the five paper contributions I aim to contribute to research in social understandings of classification schemes and empirical work analysis and work centered design of schemes, based on cognitive work analysis. This means that I attempt to discover how empirical analysis of work domains can be informed by social theory and guided by cognitive work analysis towards bringing support to work centered design.

This chapter deals with the research conducted in this thesis and in the five contributions. Section 5.1 addresses main underlying assumptions in social theory that have been adopted for the research, that is, the theoretical basis. Section 5.2 addresses the approaches used in the empirical research cases.

5.1 Theoretical Foundations

A common distinction made in social science paradigms is between agency theory and structure theory. The first paradigm focuses on individual or collective action as the main factor shaping societies and domains (eg., Weber, 1962; Mead, 1932; Schutz, 1953). The second paradigm emphasises structures as the factor determining societies and domains (eg., Durkheim, 1963; Merton, 1972; Parsons, 1937). Additional distinctions are micro sociology, that is, the study of small groups, versus macro sociology, that is, the study of societies at large (c.f. eg., Ritzer, 2000). Methodological distinctions in sociology comprise normative versus descriptive social theory. Thus, Merton's sociology of science is normative in its stipulation of rules and laws for scientists' socialisation in research (Merton, 1972), whereas Latour's sociology of science is descriptive in its basis on empirical analysis of science in action in research laboratories (Latour, 1987). A simplified overview of sociological research paradigms concerned with agency and/or structure is illustrated in figure 5.1.

Structure theories presume that societies are put together of elements, whose relations shape a pattern that distinguishes them from their environments. Survival of society does rely on actions as long as the actions are instrumental and concern overall normative values. Recent theories within this paradigm, like Luhmann's theory of autopoietic systems (1987), introduce 'de-centred' views of societies as networks of self-organising systems with the ability to reproduce themselves and distinguish themselves from their environment. Central units of study are the

structures emerging from action and interaction. These structures are viewed as existing 'out there' as social facts, which can be readily captured by a social analyst.

Agency theories presume that human action is the main factor shaping the social. One foundation is Weber's theory of human action (Weber, 1962). Weber saw society as the sum of individual actions. This implies a mechanistic view of society, where the whole is to be understood from its parts. Central units of study are individuals and individual actions. The more recent approach of symbolic interactionism is rooted in American pragmatism (Mead, 1934; Schutz, 1953) and involves studies of human collective and coordinative practice. Modern agency theories are concerned with people's social acts and their interpretations and subject positions or roles as well as social worlds, teams and communities. The dominant methodological foundation is phenomenology.

| | Structure | Agency | Synthesis of structure/ agency |
|---|--|--|--|
| Shaping of societies and domains | Collective survival and evolution is made possible through structures | Sum of individual or collective actions constitute society | Action and structures are entangled in shaping of the social |
| Methodological foundation | Explanatory frameworks for identifying social-systemic 'laws' Normative | Phenomenology; Interpretation; Descriptive | Reflexivity; Interpretation Normative and descriptive |
| Research methods | Historical comparative; Surveys; Structured questionnaires | Field studies; Participant observation; Interviews | Field studies; Dialog research |
| Theories | Systems theory | Symbolic interactionism; Ethnomethodology; Social constructivism | Structuration theory; Praxeology; Critical theory |
| Exemplars | Parsons, Durkheim, Merton | Weber, Mead, Schutz | Giddens, Bourdieu, Habermas |

Figure 5.1 Overview of main research paradigms concerned with social structure and agency.

Synthesis theories attempt to bridge the dualism between structure and agency conceptions of society, that is, how agency is related to structure, but not determined by it or vice versa. Bourdieu (1990) focuses on the dialectical relationship between objective structures and subjective phenomena and explains practice as the outcome of the dialectical relationship between structure and agency = praxeology. In his structuration theory, Giddens (1984) pleads that social research must begin with 'recurrent social practices', which are not equivalent to

any form of social totality or individual experience. Giddens advocates the epistemological perspective of *reflexivity* in social analysis. That is, the perspective that the social analyst studies a society or domain, which is already interpreted by its actors. This is the basis for Giddens' concept of double-hermeneutics. The social analyst is not dealing with 'social facts', but with interpretations. Habermas (1984; 1987) deals with agency through the concept of life world and with structure through the concept of system world. Habermas' concept of life world departs from a hermeneutical tradition. Contrary to phenomenology, which is concerned with the life world of single subjects, Habermas tries to pin down agency or the subjective aspect of society through the notion of intersubjectivity in communicative action. From a method perspective this means to study the social 'from within'. However Habermas also claims that society cannot be understood solely 'from within', that is, in a hermeneutical perspective. He agrees with Bourdieu that the symbolic reproduction of society (eg., symbolic capital) requires a material foundation (economical capital). This means that the system world should be viewed from the outside as a self-regulating system, which comprises material production and reproduction. The action patterns in the system world are considered as strategically directed and rational. Habermas sees the dualism of agency-structure as mutually determined. However, Habermas regards the life world as primary, because language and culture are necessary conditions for communicative action as well as strategic action.

Paradigms that synthesise agency and structure shape the epistemic basis of this research, in particular Habermas' theory of lifeworld/system world and communicative action (Habermas, 1984; 1987; 1993). The research approach of this thesis is interpretive and reflexive. My contributions attempt to understand and interpret agency and structures, which shape the domain of collection mediation. The interpretive epistemology implies that I have attempted to understand recurrent activities in the studied domains through iterations of interviews, as needed, rather than surveys or experiments, through participation in collection mediation and design activities and through communication with informants, or rather, participants, about their interpretation of work. I have dealt with reflexivity through arranging workshops and meetings with participants in order to enrol them in the research process early on and share interpretations of findings with them.

5.2 Research Approaches and Methods

The primary research approach used for the empirical cases of this research is dialog research (eg., Giddens, 1984; Van Manen, 1990; Storgaard, 1998). Dialog research is founded on the hermeneutical research tradition, which aims towards understanding of domains from within. This implies continuous dialog between the parties involved in research – the informants and the researcher. Dialog research adopts the field study as a method for gaining insight into a domain. From the analyst's perspective, an important aim of dialog is to understand the interpretive

universe shaping the background for the informants' statements during field studies, including interpretation of data gathered in interviews, focus groups and participant observation. The major strength of dialog research is the integration and development of viewpoints between researcher and participants in the study and possibilities for opening up to collective reflection, action and possibilities for change. A major challenge is, in a positivist sense, that results may not be open to direct replication by other researchers in different settings.

Dialog research has been used in the field studies shaping the background for the contributions of this thesis. The contribution on collaborative design of a web-based multimedia database in a public library (Albrechtsen and Jacob, 1998), builds on an action-oriented approach to dialog research. The study of the public library is placed within the Scandinavian tradition of process-oriented evaluation research (eg., Baklien, 1985). Thus, formulation of claims, research questions, data gathering and data analysis has been conducted in cooperation with stakeholders in the project, from the local participants to the 'outside' stakeholders of grant holders for the project, system suppliers, data suppliers and policy makers in the local community. The contributions on work domain analysis of three film archives (Pejtersen et al, 2001; Albrechtsen, Pejtersen and Cleal, 2002) make use of a dialog research approach in the design and conduction of the field studies. These latter studies are focused on work domain analysis for system design, based on the methodological framework of cognitive work analysis (Rasmussen, Pejtersen and Goodstein, 1994). The following explains the use of research approaches and methods.

5.2.1 Action-Oriented Dialog Research: Application

Action-oriented dialog research defines a participatory research method, where the intentions of interventions are not only driven by outside stakeholders and the researcher, but are equally defined by the actors of a domain. Moser (1995) models the cycles of dialog-oriented action research as follows:

1. Gathering information about the context of practice for the study
2. Entering into a discourse amongst the participants about intentions and problems of the study in order to negotiate directions for action
3. Action

Dialog happens in all three cycles and progressively integrates the analyst and the participants of the study. Moser's model emphasises the relationship and mutual understanding that develops between the actors and the researcher. It does not account for outside factors, shaping the interaction, such as possible stakeholders in a research project. In evaluation studies, the researcher will often navigate through

cycles of dialog with local participants in the study and with actors representing outside stakeholders. The evaluation researcher will frequently act as a catalyst for communication between the local participants and outside stakeholders (eg., Launsø and Rieper, 1993). The study of the development of a web-based multimedia information system at a public library (Albrechtsen and Jacob, 1998; Albrechtsen, 1997) adopted the strategy of dialog with local participants as well as with outside stakeholders. The evaluation object was not only the technological system and the practices of the library, but also how teamwork and project-oriented mediation work impacted staff's competence development, and how this competence development could be related to new curriculum development in information science, grounded in real life mediation problems. Conflicts and communication breakdowns within and amongst the actors and stakeholders were studied in order to determine i) the role of communication in collaborative design; ii) strength and weaknesses of the socio-technical factors impacting the process and results. The study made use of Sonnenwald's theory of communication in design (Sonnenwald, 1993), as a hermeneutic perspective for understanding the domain 'from within', Law's methodology for process-oriented evaluation of socio-technical systems (eg., Law, 1994) and Star's concept of 'boundary object' (Star and Griesemer, 1989) to illustrate how the staff's development of classification schemes as design details provided the basis for an autonomous negotiation space for the team within a complex network of participants and stakeholders.

The study proceeded in cycles of information gathering from participation in project meetings, from semi-structured interviews and focus group interviews with participants (staff and users) and representatives of stakeholders, and from empirical evaluation of prototypes. Two research workshops were arranged during the study involving the participants and stakeholders, where preliminary research results were presented and discussed. In addition to data gathered, recorded and analysed from interviews and workshops, the evaluation is based on extensive documentation by the participants about the process, in particular a diary by the project group about events and decisions. The final report was presented and discussed with the participants of the study prior to its publication (Albrechtsen, 1997). The study was hence conducted in ongoing dialog with the library and the stakeholders of the project.

5.2.2 Work Domain Analysis: Application.

Work domain analysis differs from studies of practice-related research (eg., Mathiassen, 1998) and participatory evaluation research (eg., Baklien, 1985) in that the field of study is the work domain, i.e. work activities and the context of work, and that the explicit purpose of work domain analysis is system design (eg., Schmidt, 1990; Schmidt and Carstensen, 1990). Cognitive work analysis (Rasmussen, Pejtersen and Goodstein, 1994; Sanderson, in press) is a methodological framework for work domain analysis of the territory of work, of

actors' backgrounds and preference and their activities. The aim is to arrive at an understanding of the *system* of work, which will then shape the basis for design of information systems. Figure 5.2 (the onion model) illustrates the perspectives of cognitive work analysis:

1. Means-ends analysis of the territory of work
2. Work organisation
3. Task situation
4. Decision task
5. Mental strategy
6. Actors' knowledge

Cognitive work analysis characterises collection mediation as a loosely coupled domain (Rasmussen, Pejtersen and Goodstein, 1994), where actors can proceed fairly freely in activity and choices amongst alternatives. Cognitive work analysis separates analysis of the means and ends of the territory of work, analysis of actors' background and experience, and the activities in the domain. In this way, the domain analyst gains an overview of social and individual contributions to the domain as a *work system*.

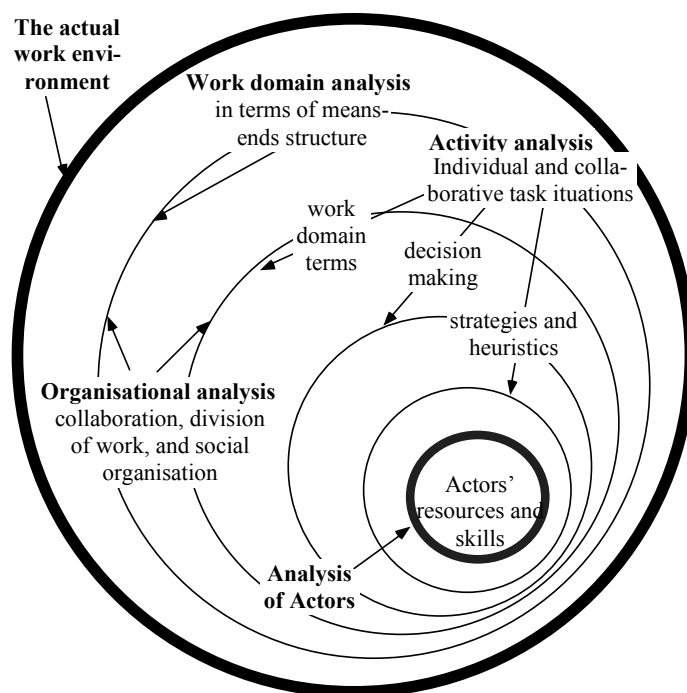


Figure 5.2 The different perspectives involved in cognitive work analysis (the Onion Model).

Cognitive work analysis is based on systems theory in an anti-reductionist sense (eg, Sanderson, in press). By anti-reductionist is meant that the work domain, as a system of work, is defined as more than the sum of its parts. In social science, systems theory is the theoretical foundation for social theory concerned with *structure* and relationships between elements in a social environment, rather than *agency* in the environment. Rasmussen, Pejtersen and Goodstein (1994) motivate an equal emphasis on structure and agency for empirical studies and analysis of work domains. The concepts of structure and agency are regarded as analytically distinct, but the aim is to arrive at an integrated structure/agency representation of the work system. Thus, structures of work territories and structures in actors' knowledge and communication are analysed and represented as means-ends representations. Agency is addressed through task situation analysis and analysis of decision tasks.

Cognitive work analysis characterises collection mediation as a loosely coupled domain (Rasmussen, Pejtersen and Goodstein, 1994), where actors can proceed fairly freely in activity and choices amongst alternatives. Cognitive work analysis also regards collection mediation as a dynamic work domain, where stable procedures are not the norm. In such domains, the objects of study are the factors shaping the work practice, like work territory and actors' backgrounds and knowledge, and recurrent task situations in the work domain. The field studies of the three film archives analysed and represented the work territory through means-ends analysis and modelling. The field studies also addressed task situations of collaborative indexing and collaborative information searching. Information about actors' background and knowledge was gathered through interviews and participant observation of mediation tasks. Information gathering about the territory of work was accomplished through participatory observation, through written documentation of the organisation and services of mediation work, and through analysis of data gathered from interviews with participants of the study. All data were recorded on audiotape and transcribed. Results from the empirical analysis of the work domains have been continually discussed with the participants in the study in follow-up interviews and through evaluation workshops.

Chapter 6

Conclusion

This chapter summarises the results of the research carried out in this summary and the five contributions. First, I summarise the results of work centered understandings and models of classification schemes in collaborative collection mediation, as discussed in chapter 4. Second, limitations of research approaches and results are identified and discussed. Finally, possible avenues for further research are proposed.

6.1 Work Centered Understandings and Models of Classification Schemes in Collaborative Collection Mediation

Work centered design of classification schemes for collection mediation involves analysis of work contexts and situational activity where classification schemes can be used to support work. Work centered design of classification schemes for collaboration in the collection mediation domain is an emerging area of research, which poses particular requirements and challenges to the design process. Two questions address the issue. They relate sociological understandings of classification schemes to cognitive work analysis for work centered design of schemes. The five individual contributions address the questions from different experience perspectives. The summary brings together the experience perspectives of the individual contributions through addressing the claim that use potentials of collections shape the structure and contents of classification schemes for collection mediation, defined as classification, indexing and searching of collections. The claim and two partial theses are addressed through challenging their ideas and implications for analysis and design in the light of additional literature. The primary results of this thesis are summarised in the following.

'Use potentials' of collections define the form and contents of classification schemes for collection mediation. 'Use potentials' address dynamic as well as more stable properties of actors' interpretations of deep semantics in the domain. These properties can be referenced by concepts or dimensions of knowledge or be referenced by classification schemes. 'Use potentials' is constructed in

communicative interaction amongst the actors in information searching, indexing and classification of collections. Collection mediation is a living system, which is open and maintains and reproduces itself within the constraints of the environment. It is a symbolically structured life world, where actors use, produce and interpret common symbolic resources, which structure and reproduce the life world of collection mediation. The life world is open to influences by the system world, i.e. to legislation and policies, which may give rise to specific priorities in mediation and production of codes of practice to reach such priorities. Communicative situations are characterised by the shaping of intersubjective spheres amongst users and staff. The intersubjective spheres contribute with implicit and often dynamic structures in collection mediation work. Mediation is also affected by more stable structures and affordances of the symbolic resources used to support mediation processes like subject analysis and formulation of search statements.

A key challenge in work centered design of classification schemes for collection mediation is that classification schemes are used in different work situations, and that their affordances are dependent on use situations. The research found that cognitive work analysis could be used to address this challenge through its models for analysis of prototypical task situations and through means-ends analysis of territories of work. The advantage of cognitive work analysis is that representations of work practice and structures can be created in a systematic and clear manner and be used as basis for design of information systems for work domains. Its focus on analysis and representation of context, i.e. work territory and practice, implies that the context for classification schemes can be made visible for users and staff and be adapted to their task situation and work process. Users and staff can perceive a familiar context for classification schemes in the interface and relate that perception to use potentials of schemes for collection mediation. The research proposes a general model of a common workspace for collaborative indexing, based on the framework for cognitive work analysis. In this model, a classification scheme mediates stable structures of the semantics of the workspace amongst the actors and between the actors and representation and data processing tools. This suggestion is made on the basis of a new unit of analysis for planning of classification schemes: collaborative task situations.

An additional important challenge in work centered design of classification schemes is that the contents and structures must be rooted in the semantics of the work domain. This challenge demands an analysis of the semantics in a domain, i.e. concept formations and structures of knowledge. In addition, a translation of analysis to design is required. Cognitive work analysis comprises two design models for work centered classification schemes. One model is applicable for representing analytical search strategy and indexing. An additional model can be used to represent users' navigation in concepts. Analysis of deep semantics in the domain of collection mediation is a comprehensive task, because the semantics is dynamic and diverse. Difference in concept formation and understandings can be readily represented through the model for navigation in concepts. This model has

been developed for task situations and processes where users need a variety of different concepts to inspire mediation and understanding of collections. In contrast, some task situations and processes require a higher degree of commonality and stable structures. For such situations, the model for analytic search strategy and indexing can be applied. The framework suggests the method of means-ends analysis of properties of work domain semantics, including means-ends analysis of the semantic territories of collection mediation. This thesis proposes a preliminary method for translating analysis of semantics in collection mediation to a classification scheme with a stable structure. Yet the research of this thesis also indicates that actors involved in collection mediation may view semantic structures differently, depending on their relative socialisation in the sphere of production, and depending on the degree of conceptual variety and the need for sustenance of such variety within knowledge domains.

6.2 Limitations

This thesis investigates two research questions in order to address work centered design of classification schemes for collection mediation. The research has a number of limitations.

A first limitation is the lack of an acknowledged conceptual framework for work centered design of classification schemes. The present research is framed within work centered approaches to HMI. This implies the danger of simplification of the fine-grained concepts and understandings of classification schemes, developed within sociological research and information science research. In addition, the thesis has a focus on cognitive work analysis for work centered design of classification schemes. I have attempted to counteract simplification through applying and discussing key concepts developed outside the traditional HMI community to frame the experience cases of this thesis within different epistemologies. In addition, I have attempted to explain key concepts of cognitive work analysis for addressing collection mediation vis-a-vis a larger body of research on the social factors shaping classification schemes.

A second limitation has to do with the research process. Although the approach of dialog research has been applied in all empirical studies of this research, the frameworks for analysis of results from the studies have been different. The research process has shifted from action-oriented dialog research in the first phase of the study, to a phase where the process focused on work domain analysis, guided by the framework of cognitive work analysis. The challenge is methodological and epistemic pluralism. I have attempted to address the challenge through retaining the approach of dialog research, throughout all empirical studies in this research. In addition, I have formulated an overall epistemic perspective for this thesis within Habermas' social theory of communicative action in order to provide a handle for revealing strengths and limitations in concepts developed for classification schemes as symbolic artefacts in work domains.

A third limitation has to do with the lack of implementation of the models developed by use of cognitive work analysis. This includes the proposed model for a common workspace as well as suggested classification schemes for collaborative indexing and information searching.

6.3 Future Research

The findings of the thesis make for several future research paths within work centered design of classification schemes. In the following, I suggest potential paths for future research. The suggestions are motivated by the above-mentioned limitations and by key conceptual findings of this thesis.

The first suggestion is to explore design of common workspaces for collection mediation by use of cognitive work analysis. The research in construction of common workspaces and context-aware computing is substantial, and future development of common workspaces for collection mediation by use of cognitive work analysis will need to be informed by the research done in this area. Further development of concepts and models for analysis of collaborative work by use of cognitive work analysis may be needed to address the design of common workspaces. The thesis has proposed that distinctions between affordances of elements in a common workspace may be fruitful for modelling the situational contexts of classification schemes; but design of common workspaces for collaborative collection mediation demands further exploration of use situations.

A second suggestion is to examine how to deal with diversity vis-à-vis commonality in actors' negotiations and understandings of deep semantic structures. This thesis promotes the idea that 'deep semantics' should not be understood in an essentialist sense as universal structures, shared by all actors in the domain. The notion of 'use potentials' is introduced to address diversity as well as commonality in actors' interpretations of semantics. Yet the degree of socialisation and participation of actors in collection mediation impacts the actors' understanding of meanings of the schemes. The research in classification schemes as symbolic resources for coordination in work is substantial; yet very little is known about how to design classification schemes that reference the evolving comprehensive semantics of 'use potentials', except through the application of generic guidelines for detailed design. Further work is needed on how to transform models of semantic structures in work domains to overall design of classification schemes, based on empirical analysis and modelling of collaborative task situations.

Bibliography

Albrechtsen, H., Pejtersen, A.M. and B. Cleal (2002). Empirical Work Analysis of Collaborative Film Indexing. *Proceedings of the 4th International COLIS conference, edited by H. Bruce, R. Fidel, P. Ingwersen and P. Vakkari*. Greenwood Village (CO): Libraries Unlimited, pp. 85-107

Albrechtsen, H., Andersen, H.K.H, Bødker, S. and A.M. Pejtersen (2001). *Affordances in Activity Theory and Cognitive Systems Engineering*. Roskilde: Risø National Laboratory, Centre for Human-Machine Interaction (Risø-R-1287)

Albrechtsen, H. and E. Jacob (1998). The Dynamics of Classification schemes as Boundary Objects in Electronic Libraries. *Classification in the Electronic Environment, edited by S.L. Star and G.Bowker. Library Trends 47 (2), pp. 293-312*

Albrechtsen, H. (1997). *Database 2001: Design af en Multimedie Web Katalog på Ballerup Bibliotekerne*. Copenhagen: Royal School of Library and Information Science. 67 p. [in Danish: Database 2001: the Design of a Multimedia Web Catalog at Ballerup Public Libraries]

Albrechtsen, H. (1993). Subject Analysis and Indexing: from Automated Indexing to Domain Analysis. *The Indexer 18 (4)*, p. 219-224.

Albrechtsen, H. (1992). *Domain Analysis for Classification of Software*. Copenhagen: Royal School of Library and Information Science.

Andersen, H.K.H. (1996). *Cooperative Documentation Production in Engineering Design. The 'Mechanism of Interaction' Perspective*. Taastrup: Danish Technological Institute. (PhD thesis).

Andersen, P.B.; Carstensen, P.H. and M. Nielsen (2000). Dimensions of Coordination. *LAP 2000*, 14-16 September 2000 (preprint; available at <http://www.cs.auc.dk/~pba/Preprints/DimCoor.pdf>)

Bannon, L. and S. Bodker (1997). Constructing Common Information Spaces. In Hughes et al: *Proceedings of ECSCW97*, Dordrecht: Kluwer, pp. 81-96.

Berlin, B. (1974). Folk Systematics in Relation to Biological Classification and Nomenclature. *Annual Review of Ecology and Systematics 4*, pp. 259-271

- Bishop, A. and S. L. Star (1996). Social Informatics for Digital Libraries. *Annual Review of Information Science and Technology (ARIST)*, 31, pp. 301-403.
- Bowker, G.C. & S.L. Star (1999). *Sorting Things Out. Classification and its Consequences*. Cambridge (MA): MIT Press.
- Bærentsen, K. and J. Trettvik (2002). An Activity Theory Approach to Affordance. In: Bødker et al (eds.) *Proceedings of the 2nd Nordic Conference on Human-Computer Interaction (NORDICHI)*, held 19-23 October 2002, Aarhus, Denmark, pp. 51-60.
- Bødker, S. and K. Grønbæk (1991). Cooperative Prototyping. Users and Designers in Mutual Activity. In: Y. Engeström and D. Middleton (eds.). *Cognition and Communication at Work*. Cambridge: Cambridge University Press, pp. 130-58.
- Carstensen, P. H. (1997). Towards Information Exploration Support for Engineering Designers. In: S. Ganesan (ed.). *Advances in Concurrent Engineering*. Lancaster (PA): Technomic, pp. 26-33.
- Cooper, L. (2002). A Study of the Relationships between Categories of Library Information as Typified by Young Children. In: Bruce, H. et al (eds.). *Emerging Frameworks and Methods: Proceedings of the International Conference on Conceptions of Library and Information Science, 4 (CoLIS)*. Greenwood Village (CO): Libraries Unlimited, pp. 17-31
- Cooper, W. (1969). Is Interindexer Consistency a Hobgoblin? *American Documentation* 20 (3), pp. 268-278
- Dahlberg, I. (1974). *Grundlagen universaler Wissensordnung. Probleme und Möglichkeiten eines universalen Klassifikationssystems de Wissens*. Pullach bei München: Verlag Dokumentation Saur KG
- Douglas, M. (1966). *Purity and Danger*. London: Routledge and Kegan Paul.
- Durkheim, E. and M. Mauss (1963). *Primitive Classification*. Chicago: University of Chicago Press.
- Ellis, D. (1989). A Behavioural Model for Information Retrieval System Design. *Journal of Information Science*, 15 (4/5), pp. 237-247.
- Gaver, W.W. (1996). Affordances for Interaction. The Social is Material for Design. *Ecological Psychology* 8 (2), pp. 111-129.

Gaver, W.W. (1991). Technology Affordances. In: Proceedings of *CHI 91*, New Orleans (LA), 28 April - 2 May 1991). New York: ACM Press, pp. 79-84

Gerson, E. and Star, S.L. (1986). Analyzing Due Process in the Workplace. *ACM Transactions on Information Systems (TOIS)* 4 (3), pp. 257-270.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Hillsdale (NJ): Lawrence Erlbaum.

Giddens, A. (1984). *The Constitution of Society. Outline of a Theory of Structuration*. Cambridge: Polity Press.

Habermas, J. (1993). Ein anderer Ausweg aus der Subjektphilosophie. In: *Der philosophische Diskurs der Moderne. Zwölf Vorlesungen*. Frankfurt/Main: Suhrkamp [in German: An Alternative to the Subject Philosophy].

Habermas, J. (1987). *The Theory of Communicative Action. Lifeworld and System: A Critique of Functionalist Reason*. Boston (Mass.): Polity Press/Beacon Press.

Habermas, J. (1984). *The Theory of Communicative Action. Reason and the Rationalization of Society*. London: Heinemann/Beacon Press.

Habermas, J. (1972). *Knowledge and Human Interests*. London: Heinemann.

Hertzum, M. (2003). Hertzum, M. (2003). Requests for Information From a Film Archive: A Case Study of Multimedia Retrieval. *Journal of Documentation*, 59 (2), pp. 168-186.

Hertzum, M.; Cleal, B., Pejtersen, A.M. and H. Albrechtsen (2002). An Analysis of Collaboration in Three Film Archives: A Case for Collaboratories. In H. Bruce, R. Fidel, P. Ingwersen, and P. Vakkari (eds.), *CoLIS4: Proceedings of the Fourth International Conference on Conceptions of Library and Information Science* (Seattle, WA, July 21-25), Libraries Unlimited, Greenwood Village (CO), pp. 69-83.

Hirschheim, R., Klein, H.K. and Lyytinen, K. (1995). *Information Systems Development and Data Modelling. Conceptual and Philosophical Foundations*. Cambridge: Cambridge University Press.

Hjørland, B. and H. Albrechtsen (1999). An Analysis of Some trends in Classification Research. *Knowledge Organization*, 26 (3), pp. 131-139

Hjørland, B. (1998). The Classification of Psychology: A Case Study in the Classification of a Knowledge Field. *Knowledge Organization*, 24(4), pp. 162-201.

Hjørland, B. (1997). *Information Seeking and Subject Representation. An Activity-theoretical Approach*. London: Greenwood Press.

Hjørland, B. and H. Albrechtsen (1995): Toward A New Horizon in Information Science: Domain Analysis. *Journal of the American Society for Information Science*, 46(6), pp. 400-425.

Ingwersen, P. (1992). *Information Retrieval Interaction*. London: Taylor Graham.

Jacob, E. and D. Shaw (1998). Socio-cognitive Perspectives on Representation. *Annual Review of Information Science and Technology*, vol. 33, edited by M.E. Williams. Medford(NJ): Information Today for American Society for Information Science. , pp. 13 -185.

Kirsch, D. (2001). The Context of Work. In: T.P. Moran and P. Dourish (eds.): Context-Aware Computing. Special issue of *Human-Computer Interaction 16 (2-4)*, pp. 305-322.

Kuhlthau, C.C. (1993). *Seeking Meaning: A Process Approach to Library and Information Services*. Norwood (NJ): Ablex.

Kwasnik, B. H. (1989). How a Personal Document's Intended Use or Purpose Affects its Classification in an Office. *SIGIR'89, Proceedings of the 12th International Conference on Research and Development in Information Retrieval*, Cambridge, Massachusetts, USA, June 25-28, 1989. Cambridge (MA): ACM Press, pp. 207-210

Lancaster, F.W. (1999). *Indexing and Abstracting in Theory and Practice*. London: Library Association.

Latour, B. (1988). Drawing Things Together. In: M. Lynch and S. Woolgar (eds.): *Representation in Scientific Practice*. Cambridge (Mass.): MIT Press, pp. 19-68.

Law, J. (1994). *Organising Modernity*. Oxford: Blackwell.

Leach, E. (1976). Social Anthropology. A Natural Science of Society? In: *Proceedings of the British Academy*, 62. Oxford: Oxford University Press.

Leonard, L. E. (1977). *Inter-indexer Consistency Studies, 1954-1975: A Review of the Literature and Summary of Study Results*. Champaign (Ill.): University of Illinois, Graduate School of Library Science (technical report).

Leont'ev, A.N. (1978). *Activity, Consciousness and Personality*. Englewood Cliffs (NJ): Prentice-Hall.

Levy-Strauss, C. (1966). *The Savage Mind*. London: Weidenfeld and Nicolson.

Lind, M. (1999). Making Sense of the Abstraction Hierarchy. *Proceedings of the Seventh European Conference on Cognitive Science Approaches to Process Control, 21-24 September 1999*. Lez Valenciennes: Presses Universitaires de Valenciennes, pp. 195-202.

Mathiassen, L. (1998). *Reflective Systems Development*. Aalborg: Aalborg University (Dr. Techn. Thesis).

Mead, G.H. (1932). *Mind, Self and Society from the Standpoint of a Social Behaviorist*. Chicago: University of Chicago (1932).

Miksa, F.L. (1999). The DDC, the Universe of Knowledge, and the Post-Modern Library. 1998. Dublin (OH): OCLC Forest Press.

Miksa, F.L. (1992). The Concept of the Universe of Knowledge and the Purpose of LIS Classification. In: N.J. Williamson and M. Hudon (eds.): *Classification Research for Knowledge Representation and Organization. Proceedings of the 5th International Study Conference on Classification Research*. Amsterdam: Elsevier, pp. 101-126.

Monteiro, E. and V. Hepsø (2002). Purity and Danger of Information Infrastructure. *Systemic Practice and Action Research* (former Systems practice), 15(2) <http://www.idi.ntnu.no/~ericm/subm.htm>

Nielsen, M.L. (2001). A Framework for Work Task Based Thesaurus Design. *Journal of Documentation* 57 (6), pp. 774-797.

Norman, D.A. (1999). Affordances, Conventions and Design. *Interactions* 6 (3), pp. 38-43

Olson, H.A. (2002). *The Power to Name. Locating the Limits of Subject Representation in Libraries*. Amsterdam: Kluwer.

Pejtersen, A.M. and H. Albrechtsen (2002). Models for collaborative integration of knowledge. In: Lopez-Huertas, M. (ed.). *Challenges in Knowledge Representation and Organization for the 21st Century: Integration of Knowledge across Boundaries*. Würzburg: Ergon Verlag, pp. 412-421.

Pejtersen, A.M.; Albrechtsen, H.; Cleal, B.; Hansen, C.B. and M. Hertzum (2001). *A Web-based Multimedia Collaboratory. Empirical Work Studies in Film Archives*. Roskilde: Risø National Laboratory - CHMI (Risø-R-1284 (EN) CHMI-02-01

Pejtersen, A.M. and H. Albrechtsen (2000). Ecological work-based classification schemes. *Dynamism and Stability in Knowledge Organization: Proceedings of the International ISKO Conference, 6*/edited by C. Beghtol & N. Williamson, pp. 97-110.

Pejtersen, A.M. and J. Rasmussen (1997). Ecological Information Systems and Support of Learning. In: *Handbook of Human-Computer Interaction*, edited by M. Helander, T.K. Landauer and P. Prabhu. North-Holland: Elsevier, pp. 315-346.

Pejtersen, A.M. (1994). A Framework for Indexing and Representation of Information Based on Work Domain Analysis: a Fiction Classification Example. In Albrechtsen and Ornager (eds.): *Knowledge Organization and Quality Management. Proceedings of the 3rd International ISKO Conference, 20-24 June 1994*, Copenhagen, Denmark. Frankfurt/Main: Indeks Verlag, pp. 251-263.

Pejtersen, A.M. (1991). *Interfaces Based on Associative Semantics for Browsing in Information Retrieval*. Roskilde: Risø National Laboratory (Risø-M-2883).

Pejtersen, A.M. (1989). *The Book House. Modeling Users' Needs and Search Strategies as a Basis for System Design*. Roskilde: Risø National Laboratory (Risø-M-2794).

Petersen, J. (2002). Levels of Abstraction in Maritime Maneuvering Operations. In: *Organizational Semiotics - Evolving a Science of Information Systems*. Amsterdam: Kluwer. (in press)

Petersen, J. and M. Nielsen (2001). Analyzing Maritime Work Domains. In: Onken, R. (Ed.) *Proceedings of the 8th International Conference on Cognitive Science Approaches to Process Control*. Sept. 24-26, 2001, Neubiberg, Germany.

Pinch, T. and W.E. Bijker (1987). The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. In W.E Bijker and T.P. Hughes (Eds.). *The Social Construction of Technological systems: New Directions in the Sociology and History of Technology*. Cambridge, MA: MIT Press, pp.17-50.

Ranganathan, S.R. (1967). *Prolegomena to Library Classification*. Bombay: Asia Publishing House.

Rasmussen, H. (1992). *Emneindeksering - specielt med sigte på Miljøstyrelsens behov*. Copenhagen: Royal School of Library and Information Science (master's thesis) [In Danish: Subject Indexing, with special emphasis on the needs of the Danish Environmental Protection Agency]

Rasmussen, J., Pejtersen, A.M. and L.P. Goodstein (1994). *Cognitive Systems Engineering*. New York: Wiley.

Rasmussen, J.; Pejtersen, A.M. and K. Schmidt (1990). *Taxonomy for Cognitive Work Analysis*. Roskilde: Risø National Laboratory (Risø-M-2871).

Rasmussen, J. (1986). *Information Processing and Human-Computer Interaction*. New York: North-Holland.

Rasmussen, J. (1983). Skills, Rules and Knowledge; Signals, Signs and Symbols, and Other Distinctions in Human Performance Models. *IEEE Transactions on Systems, Man and Cybernetics* 13(3), pp.257-267.

Ritzer, G. (2000). *Modern Sociological Theory*/5th edition. Boston: McGraw-Hill.

Ritzer, G. and P. Gindoff (1992). Methodological Relationism: Lessons for and from Social Psychology. *Social Psychology Quarterly* 55 (2), p. 128-140.

Robinson, M. and L. Bannon (1991). Questioning representations. *Proceedings of ECSCW 2*/edited by L. Bannon, M. Robinson & K. Schmidt. Dordrecht: Kluwer, 219-233

Salton, G. and M.J. McGill (1987). *Introduction to Modern Information Retrieval*. Auckland: McGraw-Hill.

Sanderson, P. M. (in press). Cognitive Work Analysis. In J. Carroll (Ed.), *HCI Models, Theories, and Frameworks: Toward an Interdisciplinary Science*. New York: Morgan-Kaufmann. (To appear in 2003)

- Schmidt, K. (in press). Remarks on the Complexity of Cooperative Work. In: H. Benchekroun and P. Salembier (eds.): *Cooperation and Complexity*. Paris: Hermes [in press] (available from the author at schmidt@it-c.dk)
- Schmidt, K. and I. Wagner (2003). Ordering Systems: Coordinative Practices in architectural design and planning. In: M. Pendergast et al. (eds.), *GROUP'03: International Conference on Supporting Group Work, 9-12 November 2003, Sanibel Island, Florida*, ACM Press, New York, 2003, pp. 274-283
- Schmidt, K. (1994). *Modes and Mechanisms of Interaction in Cooperative Work. Outline of a Conceptual Framework*. Roskilde: Risø National Laboratory (Risø-R-666-EN).
- Schmidt, K. and L. Bannon (1992). Taking CSCW seriously. Supporting articulation work. *Computer Supported Cooperative Work (CSCW): An International Journal*, 1 (1/2), pp. 7-40.
- Schmidt, K. (1990). *Analysis of Cooperative Work. A Conceptual Framework*. Roskilde: Risø National Laboratory. [Rise-M- 2890]
- Schmidt, K. and P. Carstensen (1990). *Arbejdsanalyse. Teori og Praksis*. Roskilde: Risø National Laboratory (Risø-M-2889). [In Danish: Work Analysis. Theory and Practice]
- Schutz, A (1953). Common-sense and Scientific Interpretation in Human Action. *Philosophy and Phenomenological Research*. 14, pp. 1-38
- Sonnenwald, D. (1993). *Communication in Design*. New Brunswick (NJ): Rutgers University, Graduate School of Communication, Information and Library Studies (PhD thesis).
- Star, S.L. and G.C. Bowker (1998). Classification in the Wild. *Classification in the Electronic Environment*, edited by S.L. Star and G.Bowker. *Library Trends* 47 (2), pp. 185-189
- Star, S.L. & J.R. Griesemer (1989). Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19, pp. 387-420.
- Storgaard, K. (1998). Dialogue Between Research and Development, In Hetland & Meier-Dallach (Eds.). *Making the Global Village Local*, Vol 7, Cost A4, Social Science, European Commission.

Svenonius, E. (2000). *The Intellectual Foundation of Information Organization*. Cambridge (Mass.): MIT Press.

Talja, S.; Keso, H. and T. Pietiläinen (1999). The Production of 'Context' in Information Seeking Research: a Metatheoretical View. *Information Processing and Management* 35, pp. 751-763

Thorhauge, J.; Larsen, G.; Thun, H.P. and H. Albrechtsen (1997). *Public Libraries and the Information Society*. Luxembourg: Office for Official Publication of the European Communities (EUR 17648 EN)

Vicente, K. (1999). *Cognitive Work Analysis. Toward Safe, Productive, and Healthy Computer-Based Work*. Hillsdale: Lawrence Erlbaum.

Wand, Y. and R. Weber (1995). Towards a Theory of Deep Structure of Information Systems. *Journal of Information Systems*, pp. 203-223.

Weber, M. (1962). *Basic Concepts in Sociology by Max Weber*. Translated and with an introduction by H.P. Secher. New York: The Citadel Press.

Wilson, T. (1999). Models in Information Behaviour Research. *Journal of Documentation* 55 (3), pp. 249-270.

Ørnager, S. (1997). Image retrieval. Theoretical analysis and empirical user studies on accessing information in images. In: *Proceedings of the 60th ASIS Annual Meeting*. Washington, DC, November 1-6, 1997, pp. 202-214

.

The Dynamics of Classification Systems as Boundary Objects for Cooperation in the Electronic Library¹

Hanne Albrechtsen
The Royal School of Library and Information Science
Copenhagen, Denmark

Elin K. Jacob
Indiana University
Bloomington, Indiana, USA

Abstract

The notion of the classification scheme as a transitional element or "boundary object" (Star 1989) offers an alternative approach to the more traditional approach that views classification as an organizational structure imposed upon a body of knowledge to facilitate access within a universal and frequently static framework. Recognition of the underlying relationship between user access and the collective knowledge structures that are the basis for knowledge production points up the dynamic role of classification in supporting coherence and articulation across heterogeneous contexts. To this end, it is argued that the library should be an active participant in the production of knowledge and that this role can be effected by the development of classificatory structures that can support the needs of a diverse information ecology consisting of a complex web of interacting agents, users and technologies. Within such an information ecology, a classificatory structure cannot follow a one-size-fits-all paradigm, but must evolve in cooperative interaction between librarians and their user groups.

¹ Classification in the Electronic Environment/ edited by Geoff Bowker and Susan Leigh Star. *Library Trends* 1998 47 (2), pp. 293-312. Reprinted with the permission of the editors and the co-author.

1. Introduction

A bibliographic classification system is intended to provide both an overall structure for a document collection and a set of concepts that will guide the information searcher into the knowledge domains encompassed by the collection. Traditionally, classification research has approached these objectives by developing schemes based on a one-size-fits-all-searchers paradigm: "We have created a standard system, because, deep down, all users are the same!" Such classificatory tools often fail to fulfill their function of supporting the searcher's access to and navigation through the domain structure. In most databases, including catalogues on the Web, the searcher may find it difficult to comprehend the organizational structure that has been imposed upon the materials. This is not due simply to the often-exotic notations of a scheme or to the surface characteristics of the classificatory data. Rather, the problem is often a product of a lack of match between the structure imposed upon the retrieval system by the classification scheme and the user's individual knowledge structures and search strategies. Classification research has responded to this problem by collecting the terminology of individual users and compiling the results to generate larger, broader, and hopefully more successful sets of access points for users: "If we design an end-user thesaurus, that should do the trick." In his recent book on information seeking and subject representation, Hjørland (1997) argues that such endeavors to compile end-user vocabularies are generally conducted without recourse to an underlying theory of knowledge. Because failure of the classificatory structure to support user access is generally interpreted as a mechanical question of matching between different individual knowledge structures -- between those of the searcher, the author, and the librarian as mediator (cf, for instance, Ingwersen, 1992), the underlying relationship between user access and the collective knowledge structures that are the basis for knowledge production has not been widely recognized.

From the perspective of sociology of science, Star (1989) has argued that the Turing test, which is intended to measure the degree to which an expert system is able to perform as a human expert in its interaction with individual users, should be replaced by a "Durkheim test", where the system is evaluated on its ability to support the goals of a specific community of users. Star points out that scientific work is not all of one piece, but is distributed and heterogeneous, with differing viewpoints emerging only to be reconciled within the existing knowledge base. In her view, information systems should not be designed simply to represent consensus but to accommodate the dissent that can be expected to appear among the various communities participating in their use. To this end, she brings forward the concept of boundary objects as a method for resolving problems of heterogeneity in knowledge production and use -- or, in terms of library and information science [LIS], problems of variation or inconsistency in the

representations of information producers, information mediators and information users.

In this paper, we will investigate how classificatory structures can act as transitional elements, or boundary objects (Star 1989), to support coherence and articulation in the heterogeneous and sometimes distributed contexts where knowledge is produced and mediated. In particular, we will review, within the context of the library, two perspectives, put forward by Hjørland (1997) and by Star (1989), that analyze information systems as dynamic social constructs. We will build an analogy between a scientific enterprise and the library as an active participant in the general production of knowledge and use this analogy to bring forward a view of modern classification research that engages the library directly in the development of classificatory structures that can accommodate information searching by heterogeneous user groups. Following Nardi & O'Day (1996), we regard the library as a diverse information ecology, comprising a complex web of interacting human agents, users, and technologies. And we will argue that, within such an information ecology, a classificatory structure cannot follow a one-size-fits-all paradigm, but must evolve in cooperative interaction between librarians (and other information intermediaries) and their user groups. In this context, we draw on examples of information systems in Danish public libraries: the Book House (Pejtersen, 1980); and Database 2001 (Albrechtsen, 1997).

2. Classification Systems: From Rationalism And Empiricism To Social Constructivism.

Hjørland (1997) argues for a philosophical and sociological orientation for classification research. In his view, the problem of the searcher's uncertainty is a function of relative task uncertainty in the user's problem domain. Because information searching takes place within a particular social framework -- for example, an academic discipline -- task uncertainty in searching is often the result of the relative task uncertainty within the discipline itself. Albrechtsen & Hjørland (1994) have earlier shown how such task uncertainty within knowledge domains may be a function of various social factors involved in the production of knowledge, such as the degree of interdisciplinarity or maturity within a domain. Such uncertainties will not only be manifest in the searchers' difficulty in formulating queries for IR-systems, but will also be inscribed in the relative plasticity and variety of the concepts and terminology applied within the domains.

Classification research has too often neglected such broader social backgrounds that inform information searching and knowledge organization, and has relied, more or less implicitly, on either a one-size-fits-all paradigm (rationalism) or on the accumulation of data about user behavior (empiricism). While the rationalist approach argues that "We just need to get everyone to understand this", the empiricist counters that "We just need to get more data about users" and proceeds

to collect more or less meaningful sets of "facts" on the individual user's relative success measured as the number of "hits" resulting from a series of search queries.

| | Rationalism/empiricism | Historicism/ Social Constructivism |
|---|---|--|
| Basic view on knowledge in information systems: | Knowledge is infallible and objective. | Knowledge is historically, culturally and socially determined. |
| View of concepts in information systems: | Concepts are objective and exists as modules of knowledge or universal facets. | Concepts are culturally determined, domain-dependent and evolve from experience and use. |
| View of language and dialogue: | Dialogue is secondary to objective knowledge and can be controlled through standard classifications. Example: DSM-IV | Dialogue is central to knowledge production and mediation and should be facilitated, not controlled. Example: HIV/AIDS vocabulary |
| View of information systems and their designers: | Information systems are value-free gateways to knowledge. Designers are engineers whose primary function is to exert control in support of performance. | Information systems are meaningful historical products – social and cultural constructs. Designers are epistemic engineers and knowledge catalysts whose primary function is facilitation. |
| View of mediating tools, such as classification systems: | <i>What</i> is a classification? | <i>When</i> is a classification? |

Figure 1. Epistemologies for development of classification systems

Figure 1 divides the different approaches to classification research and practice into two broad epistemological categories: Rationalism/Empiricism on the one side and Historicism/Social Constructivism on the other. Both rationalism and empiricism are based on assumptions regarding the nature of truth and the objectivity of knowledge. From the empiricist approach, knowledge is reduced to sensory observations or facts. In classification research, empiricism is the prevalent epistemology in bottom-up thesaurus construction based either on user warrant or on terminology warrant, particularly when the process lacks grounding in a theory of knowledge. In contrast, rationalism strives to reduce knowledge to an all-embracing structure of concepts that is intended to be universally comprehensive. It is, for example, the epistemological foundation for Ranganathan's notion of universal facets. Rationalism is also closely related to more sociopolitical actions undertaken by a particular agency or from a specific disciplinary viewpoint – actions which are intended to impose one view of knowledge on all research and

practice within that domain. In a paper discussing the role of dialogue in the development of classificatory structures, Jacob & Albrechtsen (1997) have shown how the American Psychiatric Association's construction of DSM-IV (Kirk & Kutchins, 1992), the international classification for mental disorders, created a device for marginalizing and eliminating the viewpoints of competing professions such as psychology. In short, both empiricist and rationalist approaches to classification are primarily looking for invariant *structures* that can be imposed on encyclopedic knowledge (universalist approaches) or data compiled from local observations (eg. grounded theory approaches).

In contrast to these more formalized, structure-seeking approaches to classification, social constructivism, or historicism, offers a view of knowledge as a product of historical, cultural and social factors, where the fundamental divisions and the fundamental concepts are products of the divisions of scientific/cultural/social labor in knowledge domains. According to a social constructivist epistemology, the concepts and the structures are inseparable in a classification system, and hence the schemes must reflect the development, variety, plasticity and use of both within a particular knowledge domain. This implies that scheme designers are not primarily looking for ways to impose one single structure on knowledge, including one set of all-embracing facets. Rather, the designers should operate as "epistemic engineers", attempting to articulate and re-present the dynamics of knowledge in such a way that the searcher can proceed from the topic of his initial query to other related perspectives on the same topic or to related materials within the same knowledge domain. In this manner, epistemic engineering of classificatory schemes can provide for multidimensional classification schemes where the concepts are re-presented in a variety of different conceptual structures, functioning to articulate the multiple discourses performed in different domains. In the role of epistemic engineer, then, the scheme designer operates as an active participant in the process of knowledge production and mediation.

Such involvement on the part of the classificationist is particularly evident in areas of interdisciplinary research, which engage participation from many different professions. The HIV/AIDS vocabulary, developed by Huber and Gillaspay (1996), provides an illustrative example of such involvement on the part of the scheme designers. This system, which was not intended as a classification, per se, but as a mediating vocabulary, was developed to support dialogue between the different communities involved with the HIV/AIDS epidemic, including clinical and medical researchers, practitioners of alternative medicine, nutritionists, psychotherapists and other professionals, as well as those individuals who are either living with the disorder themselves or caring for someone who has contracted the disease. The HIV/AIDS vocabulary is built on a theory of knowledge generation that explicitly eschews the standard life cycle for knowledge production in medicine -- a knowledge cycle that proceeds in a top-down fashion from theory developed at universities and other research institutions, to applied clinical research, to daily

clinical application. Rather, according to the epistemological view driving the HIV/AIDS vocabulary, research in lived experience must necessarily feed into basic clinical research. Accordingly, this scheme was not developed solely as a tool for retrieval of information in the database of the local community, but also as a tool for facilitating communication both within and across diverse interest groups, from the so-called layman to the cloistered scientist. In its role as communicative facilitator, the scheme is also hospitable to adaptations and extensions as an indexing language in local contexts. For instance, specific drug names are not articulated in the scheme, but are left to local instantiations of the indexing language. In Star's (1989) terms, the HIV/AIDS scheme serves as a boundary object precisely because it supports cooperation and common understandings between the various interest groups touched by this epidemic.

3. Classification and Boundary Objects

The notion of "boundary objects" was developed by Star (1989) as a structure for coordinating distributed work, such as may occur with a scientific enterprise that not only involves heterogeneous actors, elements, and goals but also incorporates different research methods, values and languages. From field studies of scientific communities Star has found that scientists are able to cooperate without consensus or shared goals. They can work together successfully because they create objects that function in the same way as a blackboard in a distributed artificial intelligence system:

"I call these *boundary objects*, and they are a major method of solving heterogeneous problems. Boundary objects are objects that are both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. They are weakly structured in common use, and become strongly structured in individual-site use. Like a blackboard, a boundary object "sits in the middle" of a group of actors with divergent viewpoints. Crucially, however, there are *different types of boundary objects depending on the characteristics of the heterogeneous information being joined to create them.*" (Star, 1989, 46-47. Emphasis in original.)

Accordingly, Star (1989; Star & Griesemer, 1989) has identified different types of boundary objects in her various case studies, including:

- *repositories* -- databases, libraries, or museums;
- *ideal types or platonic objects* -- diagrams, atlases or abstract concepts such as, for example, the concept of "species" used by both the creators of a zoological museum and other interested parties involved in its construction;

- *coincident boundaries* -- common objects with the same boundaries but having different internal contents, such as maps of a geographical area that cover the same terrain but are outlined according to different knowledge interests such as, for example, the life zones identified by biologists contrasted with the trails and collection sties identified by museum conservationists;
- *standardized forms* -- forms created as methods of common communication across distributed work groups such as, for example, the forms filled out during field work or the cataloguing formats used for cooperation and networking between libraries which may, or may not, be part of each repository's database.

Unlike the model of the ideal universal computing machine whose goal, as proposed by Turing, is to emulate individual human mental capacities in all domains, boundary objects are advanced as an ecological concept -- that is, a concept that respects local contingencies and the viewpoints of different knowledge interests. In a case study on the formation of Berkeley's Museum of Vertebrate Zoology (Star & Griesemer, 1989), a classificatory structure of the species and subspecies of mammals and birds in California constituted an important boundary object. Thus the scientific classification scheme served as a shared conceptual structure and provided a shared vocabulary that facilitated communicative exchanges and cooperation across the different social and intellectual worlds represented by the scientists and the groups of amateurs who were involved in building the museum's collection.

Although they approach the problem of classificatory structures and knowledge access from two different angles, Star's exposition of the communicative and integrative functions of classificatory structures in the general knowledge production of the sciences is closely related to Hjørland's discussion of the epistemological positions adopted in classification research and his argument for following a more pragmatic philosophy of classification (Hjørland, 1997). Star builds on case studies and theoretical work in scientific communication and knowledge production, while Hjørland builds on case studies and theoretical work in the area of information searching for knowledge production. Both argue that classifications always serve pragmatic purposes, in the same way that science serves human action. According to Hjørland's theory, scientific classifications reflect a highly abstract and generalized method of knowledge organization, in contrast to classifications with more local contingencies, such as categorizing fruit and vegetables in a supermarket or the amateur horticulturist's categorization of plants by use or cultural preferences. Such variations in taxonomic structure could be argued to reflect different levels of ambition among the interested parties and thus to function as boundary objects, created and negotiated by different social worlds, with the scientific structure functioning as a more specific taxonomy dictated by the needs of the scientific community itself. However, with respect to its specific role within the praxis of a formal disciplinary community, the scientific taxonomy is just as concrete as pragmatic systems of classification that reflect local

contingencies. Indeed, when viewed from a broader sociological perspective, these latter systems may actually be interpreted as more abstract or generalized.

4. The Role of Classifications in Diverse Information Ecologies.

American anthropologists Nardi and O'Day (1996) have introduced the concept of "diverse information ecology" to describe the socio-technical network of heterogeneous materials, people and practices that constitutes a modern library:

"What we learned in the library suggests the possibility of a socio-technical synthesis, an opportunity to design an information ecology that integrates and interconnects clients, human agents and software agents in intelligent ways congenial to extending information access to, potentially, all of humanity. As we design the global information infrastructure, the ultimate goal should be to design an ecology, not to design technology." (Nardi & O'Day, 1996, 83).

Because information ecologies are situated within human practice, they are dynamic and constantly changing. An information ecology cannot be controlled by any one single agency, but evolves through the collaboration of heterogeneous socio-technical networks, whose elements strive constantly to achieve coherence and wholeness. The notion of an information ecology also implies a collective view of information systems as striving to meet heterogeneous community goals rather than the goals of a single agency or individual. In their study of two research libraries in software companies in the United States, Nardi and O'Day (1996) explored how the work practices and expertise of librarians can serve as a model for the design of computerized information services. They found that librarians are exemplary agents who evince particular expertise not only in communicating with users but also in searching for information. These two skills are closely interrelated in that the librarian's search strategy tends to evolve in collaboration with the user's project. Nardi and O'Day propose to extend this working relationship between the librarian and the user to the collaborative design of information ecologies.

In an information ecology, a classification system should function as a boundary object, supporting coherence and a common identity across the different actors involved. In its role as boundary object, a classification would be weakly structured in common use, while remaining open to adaptation in individual communities. Across diverse information ecologies, classification schemes would function as discursive arenas or public domains for communication and production of knowledge by all communities involved. This approach to the development of classification schemes also implies that the task of constructing such a scheme would no longer be invisible work. This view of classification systems is in line with the concept of "coordination mechanisms" in distributed collaborative work, as put forward by Schmidt & Simone (1996). More importantly, the understanding and appreciation of classification schemes as boundary objects and discursive

arenas, in cooperation with heterogeneous user groups and technology, engages the library as a facilitator of connections and ensures its continuing participation as an active contributor in the general process of knowledge production.

In the following we will illustrate how the role of classification systems is changing within the information system that is the library, shifting from reliance on a single, standardized or mainstream view of order, where classification is the invisible precursor to the organization of a collection, towards the creation of more diverse information ecologies, where the development of a classification schemes takes place within an arena of discourse to create a shared order across heterogeneous social worlds.

5. Something Old, Something New, Something Universal, Something Local

As illustrated in Figure 2, classification systems have served different pragmatic purposes in the history of libraries and information retrieval systems. In a recent European study of public libraries in the information society (Thorhaug et al., 1997), it was demonstrated that public libraries have progressed through three distinct stages, evolving from manual, paper-based services, via the automated library, to the current phenomenon of the electronic, multimedia library. This progression should not be understood to imply that the current status of libraries has been driven entirely by technology. Rather, the electronic, multimedia library must be understood from a more integrated socio-technical point of view, where the various actors, including librarians, computer suppliers, and researchers in computing and information science, constitute a heterogeneous network of agencies that bring certain technologies to the foreground while marginalizing others. In the recent development and use of communication technology, for example, there is a convergence of hitherto separate, even disparate, media and activities. This is apparent in the development and application of Web technology, which integrates text-based materials, graphic illustrations and audio materials with interactive features such as online conferences and e-mail. It is characteristic of this development that the technology is not only plastic and customizable to almost any context of use, rather like a boundary object, but is constantly re-negotiated and re-developed through such use.

In the recent past, manual, paper-based libraries focused on collection building. Intermediaries, or librarians, served both as collection builders and as agents controlling and interpreting the order of the libraries. Classification systems were frequently standardized in order to support interlibrary cooperation with the result that classification research was itself dominated by the development of universal schemes, which could be adopted by central agencies to control the organization of knowledge across libraries. As a result of such standardization, classification became invisible work performed without regard to the needs of the local community of users. And, because maintenance and development of these

classification schemes was often based on literary warrant, reflecting only those subjects represented in large, national collections, they can be interpreted as imposing an implicitly empiricist view of knowledge. There was, then, at this stage in the library evolution, a mix of rationalist and empiricist epistemologies underlying classification research and development.

| | Manual paper-based services | The automated library | The electronic library The digital multimedia library |
|---|---|---|--|
| Primary means of access to knowledge: | Collection building. | Circulation, acquisition, stock control etc. | Local access to global information. Networking. |
| Technology: | Cards, phone, fax. | IT for housekeeping functions OPACs. | Internet Multimedia Web catalogues. |
| Organizational culture: | Introvert and bureaucratic. | Some change in attitude towards users. | Project oriented culture. |
| Role of classification systems: | Order and control of collections. Invisible work. | Order and control Subject access via OPACs. Some experiments with thesauri. | Support of dialogue in information services. Integration and infrastructure in diverse information ecologies. |
| Examples from Denmark: | DDC is adapted and maintained in Denmark by central agency. | Verbal indexing in Danish national catalogue. | Local experiments with classification and indexing in Danish public libraries. |
| Dominating classification research approach: | Development of standard, universal classifications. | Indexing, thesauri. OPAC R&D. Automated indexing. | Communication studies. Domain analysis and science studies. Social construction of classificatory structures. |

Figure 2: Classification research and use in different stages of public library development

The role of the librarian as intermediary was challenged during the 1980s by the development of online retrieval systems, and, in particular, by the introduction of online public access catalogues [OPACs] for end-user searching. During this decade, classification research was dominated by work on thesauri and indexing

systems. There were numerous experiments with automated indexing, including the application of text analysis techniques developed in computational linguistics. OPAC development was often based on studying users, sometimes in naturalistic settings, but generally without prior analysis of their different social worlds or the functional role of libraries in knowledge production and mediation. Research in information retrieval systems was very much oriented by a mechanistic conception of human competence in information searching, indexing and classification, thereby neglecting the variety and heterogeneity with which human agents (both librarians and users), information sources and technology interact in different settings. Furthermore, as technological fixes were thrust to the foreground, displacing the search competence of the librarians, the librarian's role as intermediary between the searcher and the collection was gradually becoming marginalized as invisible work -- the preliminary work of representing and organizing the collection that occurs in isolation from and therefore without recognition by the users.

During the 1990s, the library has increasingly switched its service emphasis from building and guarding the collection or offering users access to the collection through the local OPAC to providing local access to global information resources available on the World Wide Web. This represents a shift from a closed to an open system. In some European public libraries, for example, traditionally introverted and bureaucratic organizations have migrated toward a project-oriented culture, where librarians and users cooperate on the development of new services, using the interactive affordances of Web technology and the Internet. In general, such projects have not involved the library schools in Europe, the traditional research communities in the library and information sciences. Close cooperation between libraries and the community of LIS researchers in Europe has yet to be manifested (Albrechtsen & Kajberg, 1997). In the United States, communities of LIS researchers have come together in workshops and research projects related to the social informatics of what are called "digital libraries", but could equally well be termed "electronic libraries" (Bishop & Star, 1996). In this research area, major topics include how knowledge is structured in digital libraries, including cataloguing and classification, and how digital libraries are used -- how knowledge is produced, communicated, applied and recycled in distributed social worlds. Research methods comprise ethnographic studies of communication and knowledge production in [digital] libraries as well as comprehensive sociological studies of professional classification schemes in medicine (Bowker & Star, 1994) and nursing (Bowker, 1996). Thus it seems apparent that classification research is gradually evincing a more sociological, historical orientation.

6. Classifications as Boundary Objects in Libraries: Librarians and Users In Mutual Design Activity

Ballerup, a Danish public library, is a medium-sized library on the outskirts of Copenhagen. There is, in this library, a tradition of direct collaboration between the librarians and their users. Until recently, a majority of the librarians regarded themselves as cultural workers – as intermediaries between collection and user, very much in line with the traditional perspective described above for libraries in the manual stage. In 1995, the library started a new project called Database 2001. This project, which was evaluated by Albrechtsen (1997), involved the development of an enriched multimedia catalogue on the Web. In addition to the evaluation researcher, the project group for Database 2001 included six librarians with different areas of expertise: several in the group were experienced intermediaries and online searchers, while others were specialists in catalogue design and in management of the library's technological resources. However, none of the librarians had experience with Web design or Internet browsing.

During the development of Database 2001, the project group collaborated with user groups and colleagues in the library to identify different kinds of materials, including books, musical recordings on CD, CD-ROMs, and audiotapes of books. Text, pictures and sound were selected as enrichment for the database, the idea being to emulate a kind of virtual library on the Web. The menus were designed as graphical layers of icons representing both user groups and the kinds of materials available. The subject icons in Database 2001, which represent the subject content of materials in the database, went through several iterations. In addition, the interface designed for browsing the menus was customized for both children and adults. The librarians arranged evaluation session with users who represented different user communities and their evaluations were very positive: users with different interests were able to use the icon-based interface for browsing in the database even though they had very different interests and different goals for searching.

In the database, documents were indexed using standard call numbers from the Danish variant of the Dewey Decimal Classification [DDC]. Even though indexing by class number would take advantage of the hierarchical structure of DDC and thus would be potentially useful for browsing by users, the librarians knew from their practice as intermediaries that users found it very difficult to understand the standard classification. They experimented with a more pragmatic and much weaker structured classification, which could reflect the kinds of questions actually posed to library staff by the different user groups. For instance, for the children, they worked with the seven categories listed below and designed for each of these categories an icon for subject browsing on a Web page:

1. Computers
2. Astronomy, nature, animals, environment

3. First love, star signs, being young today
4. Horses
5. Excitement, humor
6. Fantasy, science fiction
7. Books that are easy to read

From a semantic or disciplinary point of view, the separation of subjects like animals and horses would appear to be "incorrect" or "illogical". For the children, however, this classification worked very well. Category 2 [Astronomy, nature, animals, environment] was intended for a broad group of interests, including fact literature, whereas Category 4 [Horses] was intended, in particular, for girls interested in novels about horses. There is, in Denmark, a special research tradition within children's librarianship, based on Wanting's research on how children ask questions in libraries (Wanting, 1984), that advocates mediating literature according to the different user interests of children. Pejtersen (1994) has also studied children's use of libraries in Denmark and their communication with librarians. In her development of the Book House system in the 1980s, Pejtersen used a collaborative prototyping approach, engaging librarians, information scientists, and users in Danish public and school libraries, and subsequently designed a special interface of subject icons for browsing of the Book House system by children. Database 2001 took advantage of both of these research approaches to children's information searching.

The Book House (Pejtersen, 1994) is a retrieval system for fiction and is based on a general conceptual model that seeks to surround users with an adequate resource space within which to situate their own search spaces. The design involves multidimensional representations of different kinds of user needs, search strategies and literary paradigms as well as authorial intentions. This multidimensional structure for subject access is intended to match the different levels of user interest. The system interface is constructed around the metaphor of a "house of books", guiding the users through the rooms of a library where they can browse the collection. Users can also switch between different search strategies, including analytical search in the multidimensional database structure, visualized as icons for each dimension, and browsing of subjects, visualized as icons in a picture gallery. The design of these icons involved classification experiments using both word association experiments and evaluations of suggested icons in Danish public libraries.

The icons for browsing subjects in the Book House and in Database 2001 serve similar functions: to provide the users with an overview of the subjects included in the databases. Because The Book House system builds on the central design metaphor of rooms in a library, and it provides a single, uniform interface. Database 2001, in contrast, is realized as a mixture of interfaces which include the Web layer of icons, designed by the librarians; a more or less standard search client offering conventional text-based searching; and a database structured according to

a standard cataloguing format that uses traditional call numbers to represent the subject content of the documents. While the Book House is a general system for fiction retrieval, which, in its present form, cannot be customized by individual libraries to support the idiosyncratic needs of specific user communities, Database 2001 is a localized experiment with system design and classification drawing upon a range of technologies that reflect the heterogeneity of tools used in today's libraries, from conventional customizable applications such as the closed systems of the database and the search client, to the open systems supported by interactive Web technologies.

7. Collaborative Development and the Agency of Libraries

Both the Database 2001 project and the Book House system were realized using a collaborative approach between librarians, users, researchers, and technicians. In this way, users were involved in negotiating classificatory structures and the design of subject icons in the interfaces of the two systems. Because the Book House was a new approach for interface and database design in the 1980s, it had to be developed technically from scratch. Database 2001, on the other hand, was able to take advantage both of the design ideas generated during development of the Book House system and of the possibilities for integrating modern Web capabilities within existing technology. The process of designing an interface adapted for local needs quite naturally involved local experiments with classification. In Database 2001, the graphic Web layer and its icons were intended to represent both the users' needs and the existing technology. Decisions regarding the subject icons, as well as those pertaining to the search client and the database, were determined as much by the users as by the demands of the Web technology itself. Thus, the icons employed in the graphic interface constitute an integrated system of boundary objects that mediate between the library, the users, and the technology. In this way, Database 2001 exists as an open system in that it makes the library available not only to local users but to other users as well through the medium of the technology. Without the interface of icons, the system would have been technically open but conceptually closed.

Design of the Book House and Database 2001 involved heterogeneous human actors, elements and goals, which are also found in Star's (1989) description of a scientific enterprise. Star draws upon the example of a scientific enterprise to put forward a more collective concept of design than the psychological approach generally employed for the design of AI systems. Traditionally, design of library systems is based on a consensus model, or a one-size-fits-all approach. Multidimensional classifications providing different views of concepts in IR systems are still the exception (Albrechtsen & Hjørland, 1994; Jacob, 1994). But in the Book House system and in Database 2001, classificatory structures can perform as boundary objects by accommodating both the heterogeneous

information needs and the various search strategies of different user interests as well as different knowledge communities.

Figure 3 juxtaposes some important boundary objects developed in the Book House and Database 2001 with Star's typology in order to illustrate the analogy between boundary objects in a scientific enterprise and the creation of a library system. Obviously, this analogy between the library and a scientific enterprise, even when supported by parallel structures, does not establish that what goes on in a library is isomorphic to what goes on in a scientific enterprise. Hjørland (1997) has proposed a theory of classification at multiple levels, from specific classifications developed in accordance with local contingencies to those general classifications developed by the so-called "hard" sciences, such as biology and medicine. However, analysis of the role of dialogue in the creation of classificatory structures indicates that traditional classification schemes frequently function as control structures that forestall an interpretive approach to scheme design through the imposition of controlled vocabulary that limits the impact of dissonant view points (Jacob & Albrechtsen, 1997). In this manner, current developers of classification systems do not function as epistemic engineers, creating a discursive arena or forum for multiple views of knowledge, but rather as engineers of one episteme or worldview, seeking to control the flow of knowledge production within a given domain by systematically legitimizing a single, universal classificatory scheme, thereby disenfranchising those researchers and practitioners, who do not participate in the resulting structure.

| Types of boundary objects/Applications | Book House | Database 2001 |
|---|---|--|
| Repositories: | Database with multiple orderings of information | Enriched database. Multiple kinds of materials |
| Ideal type: | Icons for subjects and ordering dimensions | Icons for subjects and target groups |
| Coincident boundaries: | Design metaphor Rooms in library | Library without walls metaphor Interactive features (e-mail etc.) |
| Standardised forms: | Database structure | Database structure |

Figure 3: Boundary objects in the Book House and Database 2001 viewed in relation to Star's typology of boundary objects

In general, however, the library and its organizational structures must be viewed as important actors in the general process of knowledge production because their primary goal is to mediate between knowledge producers and users. This role is generally realized through the provision of information services to users and producers who are very often members of the same knowledge communities. Although the scenario sketched here is traditionally understood as a closed world, where librarians mediate between documents and users, it could equally well be

described as an information ecology -- as a practice that builds environments by bringing together heterogeneous materials and actors.

The librarians' practice of building information ecologies is based on both explicit and tacit knowledge. The explicit knowledge is typically based on principles and formalisms for presenting classificatory structures in the form of universal classifications or faceted thesauri. The tacit knowledge includes knowledge of the interests of their user communities, the users' levels of computer and information literacy, and preferred tactics for "mediated" versus "unmediated" information services. In mediated services, the librarians communicate with the users, either directly or on e-mail, and guide them to relevant information sources. In unmediated services, such as the Book House system or Database 2001, the users may search a card catalog, a database or a contingent, local classification scheme, prior to browsing the conceptual space within a domain. Such "unmediated" services are, in fact, "silently" mediated by the librarians or other information professionals, who designed or customized a conceptual space for end-user searching. The librarian's service to the users has been "translated" or formalized through the classification scheme. It has been abstracted or "disembedded" from the work context of a human intermediary interacting with a user.

Following Star and Strauss (1997), much of the mediating practice of librarianship may be considered "invisible work". Even though the librarian as human intermediary is visible within the traditional library setting, his work is frequently considered to be "background work" involving the identification and delivery of books or other materials in support of the "real thing" -- the user's immediate work task or particular interest. When the work of the intermediary is abstracted from the work setting, this "invisible work" may become "visible" in that the task now falls to the user himself, but the dialogue between the user and the intermediary is effectively silenced. No longer is there a human intermediary to inform the user and ensure equality of services.

Gross and Borgmann (1995. Cited in Bishop & Star, 1996) point out that "Even home shopping requires informed consumers". When the librarian's mediation work is silenced in the high-tech home-shopping environment of electronic libraries -- when the task of the user is no longer supported by or facilitated through dialogue with the human intermediary -- some users will not be informed but will be reduced to mere consumers of standardised information services. Obviously, then, the information ecology of the electronic library cannot be responsive to the needs of the individual user without achieving a balance between visible and invisible work. As Star and Strauss (1997) point out, "Making visible can incur invisibilities; obscuring may itself become a visible activity". In "unmediated" information services, cooperation between librarians and users in the design and maintenance of classificatory structures may be one method for achieving this balance between the visible and the invisible and for ensuring the

evolution of an information ecology that is contingent upon the needs of an informed public.

8. Conclusion

Classification systems and indexing languages have been constructed as organizational tools in order to provide structure to a body of knowledge, but they frequently have the effect of limiting or re-structuring individual conceptual structures during a process of information searching (Tang & Solomon, in press). Established approaches to classification research and development appear to suffer from a fear of touching the real thing -- the social worlds constituting an information system and the collective conditions for knowledge production. However, in LIS and the sociology of science, new approaches to classification research are emerging, approaches that build on the idea of information systems as open and collaborative systems. A similar trend toward development of open systems has been identified in the public libraries in Europe, from manual, paper-based services to the electronic multimedia library. In the modern electronic library, classification is similarly transformed from a tool for establishment of order and control over the collection to a boundary object functioning to create cohesion across diverse information ecologies. The modern information ecology is a socio-technical network comprised of heterogeneous materials, people and practices. Within this emerging network, the classification scheme constitutes a discursive arena facilitated by the library and functions as a boundary object for the various interests that exist among users and librarians. Such an information ecology is at the same time a situated network and an open system wherein the classification scheme supports coherence and articulation across the domains encompassed by the network, both locally and globally.

The practice of classification is changing from invisible work carried out in centralized agencies to articulation work emerging within socio-technical networks. As the role of the library evolves from collection guardian to facilitator of connections, the role of classification is similarly transformed from control of collections to facilitation of communication, maintenance of coherence, and establishment of a shared conceptual context. From this perspective, then, the intelligent intermediaries of today are the human agents in diverse information ecologies who facilitate the process of knowledge production by collaborating with communities of users in the creation and use of boundary objects such as classification schemes.

Acknowledgements

We gratefully acknowledge the inspiration and comments by Susan Leigh Star for this paper. We also want to acknowledge the Danish National Library Authority for support of the project Database 2001 in Ballerup Public Library. An earlier version

of this paper was presented and discussed at the ASIS SIG/CR workshop in Washington, 1997.

References

Albrechtsen, H. (1997). *Database 2001: design af en multimedie Web katalog på Ballerup Bibliotekerne*. Copenhagen: Royal School of Librarianship. [In Danish: English title: Database 2001 - The Design of a Multimedia Web Catalog in Ballerup Public Library]

Albrechtsen, H. & Kajberg, L. (1997). Professional development: education, training and continuing professional education. In J. Thorhauge et al (eds.) *Public Libraries and the Information Society* (p. 82-91). Luxembourg: Office for Official Publication of the European Communities. EUR 17648 EN

Albrechtsen, H., & Hjørland, B. (1994). Understandings of language and cognition: implications for classification research. In R. Fidel et al (eds). *Proceedings of the 6th ASIS SIG/CR Classification Research Workshop, Alexandria, VA, 16 October 1994* (p. 1-16).

Bishop, A.P., & Star, S.L. (1996). Social informatics of digital library use and infrastructure. In M.E. Williams (Ed.), *Annual review of information science and technology*, vol. 31 (p. 301-401).

Bowker, G. C. (1996). *Lest we remember: organizational forgetting and the production of knowledge*. Available at <http://alexia.lis.uiuc.edu/~bowker/pubs.htm>

Bowker, G.C. & S. L. Star (1994). Knowledge and infrastructure in international information management: problems of classification and coding. In L. Bud (ed.) *Information Acumen: The Understanding and Use of Knowledge in Modern Business*. London: Routledge.

Gross, R.A. & Borgman, C.L. (1995). The incredible vanishing library. *American Libraries* 26 (9), p. 900-904.

Hjørland, B. (1997). *Information seeking and subject representation: an activity-theoretical approach to information science*. Westport, CN: Greenwood Press.

Huber, J. T., and Gillaspay, M. L. (1996b). Controlled vocabulary for HIV/AIDS: an evolving nosological record of a diseased body of knowledge. In R. Green (Ed.). *Knowledge organization and change: Advances in knowledge organization*, vol. 5 (pp. 293-301). Frankfurt/Main: Indeks Verlag.

Ingwersen, P. (1992). *Information Retrieval Interaction*. London: Taylor Graham. x,246 p.

Jacob, E. K. (1994). Classification and crossdisciplinary communication: breaching the boundaries imposed by classificatory structure. In H. Albrechtsen and S. Oernager (Eds.), *Knowledge organization and quality management: Advances in knowledge organization*, vol. 4 (pp. 101-108). Frankfurt/Main: Indeks Verlag.

Jacob, E.K., & Albrechtsen, H. (1997). Constructing reality: the role of dialogue in the development of classificatory structures. In I.C. McIlwaine (ed.) *Knowledge Organization for Information Retrieval: proceedings of the 6th international study conference on classification research, held at University College London, 16 - 18 June 1997*. (pp. 42-50). The Hague: FID.

Kirk, S.A. & Kutchins, H. (1992). *The selling of DSM : the rhetoric of science in psychiatry*. New York : A. de Gruyter.

Nardi, B., & O'Day, V. (1996). Intelligent agents: what we learned in the library. *Libri* 46 (2), pp. 59-88.

Pejtersen, A.M. (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 O. Harbo & L. Kajberg (Eds.), *Theory and application of information research* (p. 146-159). London: Mansell

Pejtersen, A.M. (1994). A new approach to design of document retrieval and indexing systems for OPAC users. In D.I Raitt & B. Jeapes (Eds.), *Proceedings of the 17th International Online Information Meeting* (p. 273-290). London: Learned Information.

Pejtersen, A.M.; Albrechtsen, H.; Lundgren, L.; Sandelin, R. & Valtonen, R. (1996). *Subject access to Scandinavian fiction literature: index methods and OPAC development*. Copenhagen: Nordic Council of Ministers (Temanord).

Schmidt, K. & Simone, C. (1996). Coordination mechanisms: towards a conceptual foundation of CSCW systems design. *Computer supported cooperative work. The Journal of Collaborative Computing*, 5 (2-3), p.155-200.

Star, S.L. & Strauss, A. (1997). Layers of silence, arenas of voice: the ecology of visible and invisible work. To appear in *Computer Supported Cooperative Work. The Journal of Collaborative Computing*.

Star, S.L. (1989). The structure of ill-structured solutions: boundary objects and heterogeneous distributed problem solving. In L. Gasser & M.N. Huhns (Eds.), *Distributed artificial intelligence* (p. 37-54). London: Pitman.

Star, S.L. & Griesemer, J.R. (1989). Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19, 387-420.

Tang, R. & Solomon, P. (1998). Toward an understanding of the dynamics of relevance judgment: an analysis of one person's search behavior. *Information Processing & Management* 34 (2/2), pp. 237-256

Thorhauge, J; Larsen, G.; Thun, H.-P. & H. Albrechtsen (1997). *Public Libraries and the Information Society*. - Luxembourg: Office for Official Publication of the the European Communities. EUR 17648 EN.

Wanting, B. (1984). How do children ask questions in children's libraries? Concepts of visual and auditory perception and language expression. *Social Science Information Studies* 4, p. 217-234.

Ecological Work Based Classification Schemes²

Annelise Mark Pejtersen & Hanne Albrechtsen
Centre for Human-Computer Interaction
Risø National Laboratory Denmark
P.O.Box 49, DK-4000 Roskilde Denmark

Abstract

This paper introduces a new approach to the design of classification schemes for complex work domains to help structure the knowledge domains in databases for single users and multiple users in co-operative work. Ecological work based classification schemes are designed on the basis of an empirical analysis of the invariant structures of the work domain and of the information needs of its actors. Invariant structures of a work domain can be explicit or implicit (hidden structures). The invariant structures are identified through empirical analysis of field studies in work domains, guided by the use of a means ends abstraction hierarchy. This hierarchy provides a model for analyzing, organizing and relating the different levels of properties within a work domain. The resulting structure is an ecological classification scheme, comprising the different dimensions or categories of domain information that needs to be available for an actor to make a decision. Contrary to traditional classification systems which usually are designed from one particular point of view (a single discipline, paradigm or purpose), ecological classification schemes provide a transparent and structured information environment in which actors can navigate freely according to their current perspectives of work and subjective preferences.

1. Introduction

During the last decade, a number of scientific studies within Information Science, Social Studies of Science and Social Anthropology have explored the construction and use of classification schemes in professional work domains. Information

² Proceedings of the 6th International ISKO Conference, held July 10-13 2000, Toronto, Canada, edited by C. Beghtol & N. Williamson, pp. 97-110. Reprinted with the permission of the editors of the conference proceedings and the co-author of the paper.

Science studies include Bowker & Star (1999); Davenport (1998); Huber & Gillaspay (1998); Pejtersen & Albrechtsen (2000); Albrechtsen & Jacob (1998). Studies within Social Anthropology and Social Studies of Science comprise i.a. Suchman (1994); Andersen (1994); Cole & Engeström (1993); Star & Griesemer (1989). These studies are important contributions to a cross-disciplinary social theory of classification schemes at work. In addition, a few of these studies have discussed and assessed recent cases of classification schemes developed for specific professions, like nursing and psychiatry (see for instance, McCloskey & Bulechek, 1995; American Psychiatric Organization, 1994; Bowker & Star, 1999; Davenport, 1998; McCarthy & Gerring, 1994). For the medical field, the schemes are found to promote various implicit socio-political agendas within the explicit motivation of “standardization”. Implicit agendas range from managed health care (Davenport, 1998), towards global co-operation and knowledge sharing within medical research (Bowker & Star, 1991) and the development of a new self-understanding within a particular health care profession (Berg, 1999; Davenport, 1998; McCarthy & Gerring, 1994).

These recent studies of work based classification schemes are important for an articulation and broader understanding concerning the diversity of socio-technical functions, intentionalities and constraints of classification schemes within a professional work context. However, so far, no studies have addressed the problem of how to design work based classification schemes that go beyond local work places and/or professions. This paper describes the cognitive systems engineering approach to the design of ecological work based classification schemes based on field studies. The design of such schemes involves a means ends analysis of the work domain and the work content required in decision making during work (Rasmussen, Pejtersen & Goodstein, 1994). Special emphasis is on capturing recurrent themes and topics functioning as invariants (Gibson, 1979) in the domain in order to structure the concepts in such a way that the schemes will function as stable conceptual structures for the actors’ information searching.

1.1. Ecological work based classification schemes

Ecological work based classification schemes are schemes that are designed on the basis of an empirical analysis of the invariant structures of the work domain and of the information needs of its actors. Invariant structures of a work domain can be explicit or implicit (hidden structures). These invariants can serve as affordances for action. They are identified through the use of the means-ends abstraction hierarchy providing a generic model for analysis and description of work domains and the actors' information needs. This means ends analysis represents the invariants and affordances of work domains as a category separate from the actors, and in addition represents the invariants of the actors' information needs as they are formulated in their decision making during work.

The design of ecological classification schemes is a fundamental problem of structuring knowledge base contents to suit the actors' information needs during their decision-making. Thus, ecological schemes require field studies of the search questions and the need formulations as they occur in the actors' work situation. The first step is to make an analysis of the workspace that is studied for the purpose of system design. The next step is to conduct an analysis of the decision situations and the actors' formulations of their information needs during their decision-making. These analyses guide the identification of domain knowledge and design of a work based classification system.

This paper explains our approach to the design of ecological work based classification schemes by a conceptual introduction to the means ends model for analysis of a work domain, and by an introduction to J.J. Gibson's theory of direct perception together with implications for scheme design. The two-step approach to scheme design is illustrated by two examples: The design of an ecological classification scheme for the full-scale library system called the Book House and by a field study of engineering design. This latter study covers a means ends analysis of the design workspace and the related information needs derived from interviews and questionnaires. In contrast to the library studies, the engineers' search questions during actual decision-making were not collected, and therefore, ecological classification schemes have not been developed for this work domain.

2. The means-ends abstraction hierarchy

The means-ends abstraction hierarchy provides a model for analyzing, organizing and relating the different levels of properties within a work domain. The resulting representation of domain properties is a map of the overall conceptual territory, or in Gibson's terminology (1979), the *affordance space* within which the actors will navigate. Domain properties are represented in a means-ends structure that relates multiple goals and constraints with a variety of resources. The analysis of domain properties addresses *why* a work activity exists (goals and constraints), *what* activities are performed or should take place in the domain, and *how* these activities can be implemented in tools and work processes (see figure 1). The analysis gives further structure through its decomposition into elements along a part-whole dimension and through its identification of the potential means and ends at several levels of functional abstraction. These levels include representations of physical work resources, work processes, general work functions, abstract value functions and, finally, goals and constraints with reference to the environment outside the organization (see figure 2). These levels will also guide the identification of domain knowledge and development of a work based classification system.

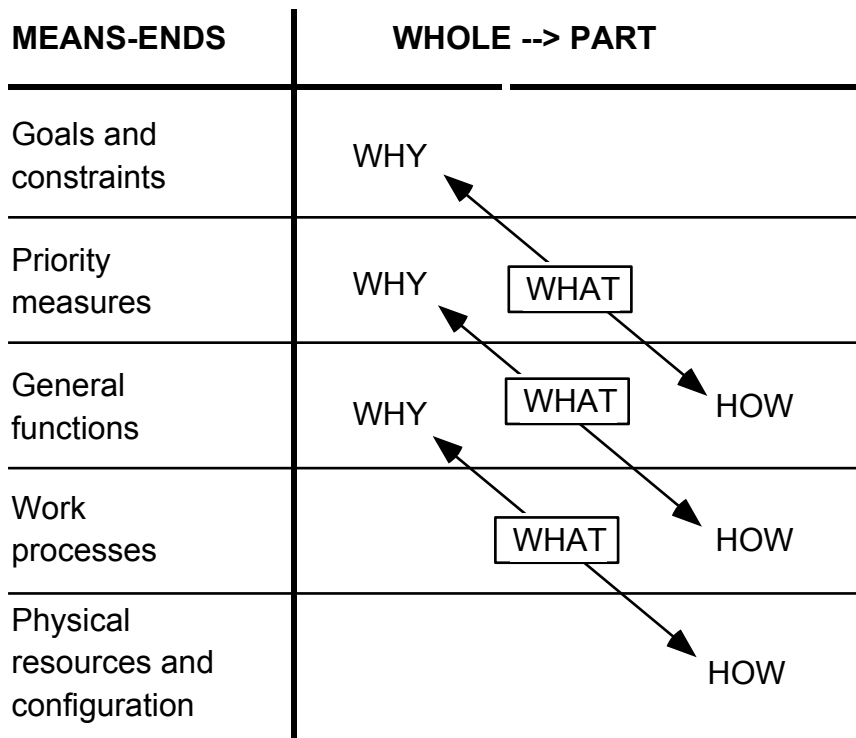


Figure 1. *WHY, WHAT and HOW in the means-ends space.*

| MEANS-ENDS RELATIONS | PROPERTIES REPRESENTED |
|---|--|
| Goals and Constraints | Properties necessary and sufficient to establish relations between the performance of the system and the reasons for its design, i.e., the purposes and constraints of its coupling to the environment. <i>Categories are in terms referring to properties of environment.</i> |
| Priority measures | Properties necessary and sufficient to establish priorities according to the intention behind design and operation: Topology of flow and accumulation of mass, energy, information, people, monetary value. <i>Categories in abstract terms, referring neither to system nor environment.</i> |
| General Functions | Properties necessary and sufficient to identify the 'functions' which are to be coordinated irrespective of their underlying physical processes. <i>Categories according to recurrent, familiar input-output relationships.</i> |
| Physical work Processes and Activities | Properties necessary and sufficient for control of physical work activities and use of equipment: To adjust operation to match specifications or limits; to predict response to control actions; to maintain and repair equipment. <i>Categories according to underlying physical processes and equipment.</i> |
| Physical resources and Configuration | Properties necessary and sufficient for classification, identification and recognition of particular material objects and their configuration; for navigation in the system. <i>Categories in terms of objects, their appearance and location.</i> |

Figure 2. The classes and properties of the means-ends hierarchy.

The lowest level of abstraction represents the physical anatomy of the system and the appearance of its elements, that is, its material configuration. The next higher level describes the physical activities and processes of the various elements in a language related to their specific material properties (e.g., physical, mechanical, electrical, or chemical processes). At the level above this, work functions are represented by more general concepts without reference to the physical processes or parts by which the functions are implemented. At the level of abstract function the functional implications are found which are used to set priorities and coordinate resource allocation to the various general work functions and to compare their

results with the goals and constraints formulated at the upper level. This level of abstract functions represents functionality / intentionality in terms of flow of values for which laws of conservation are valid, such as monetary values, energy, material, people, etc. (For a detailed discussion of the means-ends hierarchy, see e.g., Rasmussen, 1986; Rasmussen et al., 1991; Vicente, 1999).

In the work domain analysis, the substance matter of a work domain will be represented at several levels of abstraction which represent goals and constraints, general functions, physical work processes and activities, as well as physical resources and configuration. The need for human decision-making is present only because of the many-to-many mapping among the elements at the various levels. In any work domain, there are many action possibilities and options for choice in the means-ends abstraction hierarchy, which have to be eliminated by a decision, guided by the functional criteria related to the result of a decision as well as subjective preferences concerning the work process.

3. Ecological scheme design and invariants

According to American psychologist J.J. Gibson (1979), our perception and actions are dependent on the invariant properties of our environment (or, in Gibson's terms: ecology). These invariant properties are among the *affordances* we perceive in the ecology. According to Gibson, our perception is neither mediated (cf. Vygotsky, 1980) nor randomly picked up through our senses and subsequently processed towards "cognitive structures" (cf. eg. Lindsay & Norman, 1977). Gibson formulated an alternative theory of direct perception which is made possible by the affordances, that are the *action possibilities* of the "real world" (natural or artificial).

Invariant structures constitute a specific form of affordances providing us with a stable context within which we can situate our choices and actions (cf. eg. Gibson, pp. 72-73). Invariant structures can for instance be found as physical objects arranged together in a particular way, such as a block of buildings in a city, labeled with sequential numbers. Invariant structures can also be understood in more abstract terms as classificatory structures of concepts. However, as Gibson points out, "To perceive an affordance is not to classify an object" (1979, p. 134). Gibson's view of classification (invariant structures) is more in alignment with pragmatic epistemologies and philosophies, in particular with Wittgenstein's pragmatic philosophy of language: "The theory of affordances rescues us from the philosophical muddle of assuming fixed classes of objects, each defined by its common features and then given a name. As Ludwig Wittgenstein knew, you *cannot* specify the necessary and sufficient features of the class of things to which a name is given. They have only "family resemblance". But this does not mean you cannot learn how to use things and perceive their uses. You do not have to classify and label things to see what they afford" (Gibson, 1979, p. 134).

While this view by Gibson may seem to exclude the necessity, or even the “epistemological truth” of classification, then the underlying assumption is that invariant structures and objects *do* exist in the ecology either as “natural” or “artificial” (humanly constructed) structures. However, when we are situated in the ecology, our actions and understandings do not come from the structures per se, but from their perceived usefulness to us, relative to our specific purpose of action and decision in a particular situation. Thus, an ecology can be regarded as having stability and consistency of structure, while at the same time being plastic and changeable according to the perspective from which it is approached by us. Hence, the idea of invariant structures is not equivalent to essentialist, scientific views of a stable and fixed structure of our environment, held by for instance Aristotle and Linnaeus in their theories of natural classification. Rather, following for instance the social theory of classification formulated by Bowker & Star (1999, p 7), they are “spatial, temporal or spatio-temporal segmentations of the world that do some kind of work – bureaucratic or knowledge production.” This latter view by Bowker & Star together with Gibson’s view of invariant structures is in line with the underlying rationales for the means-ends approach to work domain analysis. Paraphrasing Bowker & Star’s definition of a classification, the means-ends approach promotes a segmentation of the work domain that is intended for a particular form of work: the design and use of an information system.

3. 1. Implications for ecological classification schemes

The challenge of developing a concrete classification for a particular domain is to combine the stability and consistency of the underlying structure of the work context at different levels with the subjective value structures and conceptions of the actors according to their individual and collective purposes for action. Hence, *an important challenge for ecological scheme design is to identify and articulate invariant structures of a work domain.* These invariant structures must be made visible as affordances in such a way, that they provide the actors’ information searching with a context for understanding different levels of possible actions when they use the system. For instance, the identified invariant structures of a work domain can be shown to the actors in the form of metaphorical mappings. These can range from simple representations such as the desktop metaphor, to more abstract representations, like a house with rooms of books (library) as a metaphor for a collection of databases.

Abstract representations can also cover knowledge representations, i.e. representations of deep semantic structures of a work domain, like classification schemes. Such representations can take many forms. The simplest form builds on one basic structure, implementing one single principle for dividing domain knowledge (eg. part-whole, universal facets etc.), or imposing one single professional point of view on a body of knowledge (eg. the viewpoint of one single paradigm). Conversely, knowledge representations can build on highly complex

and rich structures, implementing a diversity of work constraints and goals, and a diversity of expertise and interests among human collaborators.

According to for instance Svenonius (1992), classification research has so far to a large degree focused on developing principles for scheme design that support ease of design and modification (for instance, universal facets for subdivision of any body of knowledge). As Svenonius points out, such principles are primarily formulated in order to simplify the practice of information retrieval, scheme design and maintenance. In addition to such practical problems is the meta-problem of “what constitutes an appropriate knowledge representation, how is compatibility to be achieved among classifications, how are classifications to be evaluated and what general laws and principles underlie classifications” (Svenonius, 1992, p.17). Because classification schemes are designed to represent and mediate knowledge, an important meta-problem is semantics, i.e. a theory of meaning. Svenonius suggests that future classification research works with the idea of “meaning in use”, i.e. a context-dependent theory of meaning, or a pragmatic philosophy of representation (cf. also Wittgenstein, 1953). This suggested shift of focus in classification research from principles of universality towards principles of use and context-dependency is in line with the ecological approach to scheme design proposed in this paper.

4. Design case 1: Full-scale library system

During the 1970s and 80s, the model and method for means-ends analysis were applied for the design and evaluation of the icon-based Book House system for fiction retrieval (Pejtersen, 1994; Rasmussen, Pejtersen & Goodstein, 1994.). The means-ends analysis addressed the workspace of fiction retrieval and the questions and formulation of needs during decision-making.

4.1 Means ends analysis of the library domain

The work domain of a library includes the librarians and the users. The workspace to be addressed is the book stock and multimedia materials. Library users are largely autonomous even if they are somewhat constrained by the materials that are made available under the library laws and policies for library services. The information-seeking task therefore can be described as a multi-level representation of the invariants of users' needs and intentions on one hand, and a similar representation of the invariants of the contents of the available documents on the other hand. Since information seeking in libraries often takes place in co-operation with intermediaries, local organizational issues and library policies will impact the search process and will have to be taken into account in classification system design. The goals and intentionality of the library may constrain the original goals and intentionality of the users' domain of work or interest. Additional constraints comprise the general administration of the library, for instance, division of labor,

acquisition policies, tools etc. Invariants in all these domains are relevant for design of classification schemes, since they will influence the users' perception of affordances, and thus the performance of the decision task.

| | Document Content | Users' Needs | Library |
|---|--|---|--|
| Goals and Constraints | Author intention, Information, Education, Enjoyable experience | Readers' ultimate goal | Policies, goals cultural environment laws and constraints. |
| Priority Measures | Literary or professional quality, Paradigm, style, or school | Value criteria related to reading process and product | Flow of money, funding sources and drains. Book and user statistics |
| General functions | General frame of content, Cultural environment, Historical period, Professional context | General topical interest of historical or social setting. | Retrieval, document acquisition, indexing, cataloging, reference work, user education, web design |
| Physical work processes And activities | Specific, factual content, Episodic course of events, Factual descriptions | Topical interest in specific content | Using specific tools, Internet www, files; desk expedition of users answering request on phone or e-mail |
| Physical Resources and Configuration | Physical characteristics of document, Form, size, color, typography, source, year of print | Reading ability | Physical characteristics of shelves, documents computers, equipment, offices, buildings, Staff and end-users |

Figure 3. The workspace of the library work domain. In general, these workspaces are not formulated in compatible terms.

The field study of the work domain of libraries focused on the retrieval task and the users' query formulations during decision-making. Decisions in information seeking are for example analysis of information need, choice of information source, planning a search and evaluation of the relevance of search results. The aim of these field studies was to identify the invariants, or recurrent themes, for the design of a classification scheme that could represent the stock of documents in a way that converged with the invariants of the user needs and with the invariants of the library work domain. In general, the three task domains of the library (documents, users, library) are not formulated in compatible terms. The means ends

abstraction hierarchy provides a generic model that makes it possible to map the invariants in these domains.

The means-ends analysis showed that users tend to characterize the contents of documents at different levels of abstraction and from many different perspectives of use. These abstraction levels and perspectives were found to be recurrent themes, or invariants (Gibson, 1979), within the domain. Figure 3 shows the invariants based on the field analysis of libraries: the document content, user needs, and the library, all represented according to the five abstraction levels in the means ends representation (Pejtersen 1980, 1986). This mapping articulates that information retrieval is essentially an activity that attempts to achieve a mapping between three domains: (1) the documents with their goals/values and knowledge domains as expressed by authorships; (2) the user/reader with his/her knowledge domain derived from his/her leisure or work domains; (3) the organizational policies of the library domain. Each of the three domains has five mean-ends levels (see figure 3).

For the task domain of documents, the classification scheme identifies the content with respect to the author's goals when writing a document. These also include the authors' constraints from social, cultural, professional contexts, their subscription to and affiliation with paradigms, their choice of specific content, and their ways of composing a document to reach their target reading groups. The left column of the document content in figure 3 refers to the analysis of different invariant perspectives that was adopted during analysis and representation of documents to build a database content supporting multidimensional subject retrieval. The right column shows the invariants of the library domain that are related to the goals of libraries as expressed in laws to promote high quality of information dissemination, education, and culture. Limited budget resources often impacts the tools, resulting in limited information and classification schemes that do not reflect the users' perspectives nor the quality goals of library institutions.

4.2 Invariant structures and classification of the fiction domain

In order to provide support and guidance for the users' exploration of the Book House, an overall metaphor for ecological interface design was developed around the structure and space of a house of books. This metaphor is a mapping that articulates the invariant structures identified in the work domain analysis. Following Gibson (1979), such a metaphorical mapping constitutes an artificial environment whose invariants, just like the invariants of a natural environment, guide our direct perception and actions. Invariant structures in the Book House are found at different abstraction levels, from the analogy between rooms in a library and information databases to more deep semantic structures like the classification scheme for fiction. The deep semantic structures that are embedded in the constructed classification scheme are equivalent to the different knowledge levels

about fiction within the user community. Each invariant represents one way of organizing and retrieving information in the system, and thus, a specific discourse or knowledge level within the work domain/user community. Together, the invariant structures implemented in the scheme provide the users with a set of decision and action alternatives through its implementation of a dual principle of simultaneous semantic consistency and prediction of possible action. Thus, the design of the scheme is founded on a descriptive approach to complexity and dynamism of domain knowledge, while still implementing a core semantic stability.

A means/ends analysis of the library work domain and of the users' needs domain has been useful in order to develop a classification scheme for analysis of the document contents at the five different means-ends abstraction levels. In the present case, a means/ends analysis can only be made from the users' problem formulation and query statements in the library retrieval situations. For retrieval in a public library domain analysis beyond this representation of users' work domains will be difficult, but for information systems dedicated the service of the staff in a particular organization, a more thorough means/ends analysis of users' information needs will be useful and possible.

5. Design case 2: Engineering design

Increasingly, design work requires a multidisciplinary team of actors, viewing a product as an object belonging to different functional contexts. Multidisciplinary design teams therefore need to seek and integrate information from different domains in order to come to an understanding of how the product being developed will coexist with work activities and patterns of use. The consequence of this development is the need for classification schemes that support decision making during dynamic co-operation among actors with very different expertise and professional backgrounds, concerns and objectives, time horizons, and information needs within and across organizations. Such schemes can serve several different purposes if they provide a compatible and structured information environment for knowledge sharing and collaborative organisation and retrieval of information and knowledge. One purpose is to cope with the complexity of the co-operative decision-making in work situations that require complex information with a high amount of abstract attributes from many different domains. Another purpose is to structure and file the heterogeneous information that is produced and retrieved by the team members during a long term product development, which serves to document and reuse work information.

5. 1 Means ends in engineering design

The workspace of engineering design can be viewed as a means-end hierarchy representing the many to many mappings between the various levels of description bridging the many different work domains to be considered, when solving design problems (Pejtersen, 2000; Pejtersen et al. 1997; Carstensen, 1997). This is illustrated in figure 4, and by the following presentation of a project on the design and planning of the production of an advanced electronic control unit for hydraulic forklifts, trucks, and other power tools for construction work. The project lasted four years and the following presentation is a short synthesis based on interviews, questionnaires and studies of project files and work documentation. Our presentation of the engineering design project is structured according to the means ends abstraction levels, starting with the highest abstraction level of goals and constraints for the project and team, and concluding with the concrete physical resources etc. involved.

Goals: The goal of the project was to sell a new technology for electronic control units for hydraulic forklifts, trucks, and other power tools for construction work. Another goal was to conduct the work as joint venture and transfer knowledge on hydraulics from another industry to the company. It was important to maintain the image of the company as protective of the environment and of working life qualities.

Constraints: The choice of new materials to be used in the control unit was constrained by market conditions, environmental laws, work regulations and union agreements. The costs caused by a possible new manufacturing method constrained the choice of materials, since a new method could lead to the need for new production tools, reorganisation of work and training of the workers. Another constrain was the company's policies on budgets, time schedules etc.

Priority measures: Increased profit through a better market share depended on the costs of a new technology and on an improved quality, reliability and safety of the control unit in harsh environments. In addition to the functional specifications, very strict fail-safe characteristics were required due to reported hazards of loosing control with high power tools. Finding a new technical solution at a competitive price that met the required functions in use was given a high priority: a cheap, yet reliable control unit that ensured the end users' safety. Minimum waste disposal of the materials used in the product would protect the environment and reduced use of materials causing disposal problems had a high priority. The product was to be manufactured through available production tools, or it should be easy to specify and produce in new ways without increasing the production costs. Priorities had to be made among the costs of a new technology, the degree of safety required by the users and the costs of the disposal of the new material. It was obvious that a new

and unknown technology had to be invented and therefore the current competence of the design team members and the state of the current technology were other value criteria. In order to increase the speed of the design process and the quality of the product through concurrent engineering, it was important to gain experience in organising the work as concurrent teamwork with a flat management structure.

General functions: In order to achieve the goals of the project, a larger number of functions had to take place. Some of these are architecture design, module component design and implementation, interface design of modules and components, testing and problem diagnosing, production planning, assembling, quality control, waste management documentation, interaction with sub-suppliers, documentation of the project work, information seeking, management, coordination and planning the work.

Work Processes: The physical work processes involved in the general work functions and in the use of the physical resources were many-sided, such as: to draw product components, make experiments with hybrid technology such as chip-on-wire bonding and new production materials and methods such as pasting and soldering, seek information about new technologies and materials, write and read specifications, work documentation and reports, prepare meetings with team members, other colleagues and companies, talk to customers and sub-suppliers, lectures and presentations of the work, search for written and oral information, and people with relevant expertise.

Physical resources and configuration: These include both the manufacturing equipment and other physical entities such as: production tools, automatic assembling equipment, transistor tapes, cad/cam tools, test equipment, databases, engineers, marketing staff, materials, computers, Dialogue databases, Internet, archives, books, manual drawings, specifications, reports, minutes of meetings, release notes, authorizations, test samples, photos, work plans, financial status reports, calculations, overheads, in-coming and out-going mail, CAD drawings.

This means ends presentation of the workspace of the project is needed to bridge the space between the goals and constraints that the final product should meet and the physical resources of the technologies needed for manufacturing. Information seeking to satisfy information needs and enable decision making during the design process implies iteration between the means ends levels in the abstraction hierarchy in order to find a solution that meets the goals and constraints of the project. It provides the ground for a further analysis of the *invariants* and *affordances* of engineering design work that need to be made explicit in work based classification schemes. For example, the organisational work domain invariants such as the company's policy for running development projects, the requirement that the work should be conducted as a joint venture, and that the product should be produced

through existing production tools. Other recurrent themes include the environmental laws, the union agreements and work regulations imposed by society.

5.2. Bridging means ends levels and domains during co-operation

The co-operation in a library is a dynamic, coordinated activity that takes place in a single task situation between an expert and a novice and usually does not continue after the completion of the task. The co-operation in the engineering design team was a dynamic collaboration during weekly meetings among experts on technical design, customers and end users, marketing, manufacturing, and recycling and disposal. The decisions to be made during the co-operation involved exploration of the role of the product in different domains, analysis of product specifications and test data, generation and evaluation of ideas for error detection, and choice among alternative solutions to safe materials. These decisions are intertwined with the decisions involved in the information retrieval task similar to those in libraries. For example, analysis of information needs, choice of information source, planning a search and evaluation of the relevance of the results.

In figure 4, the design team members' information needs during their co-operative decision making are viewed as a means-end hierarchy representing the many to many mappings between the various levels of description bridging the many different work domains. It shows that decision making and information seeking during engineering design tasks concurrently iterates among several different work domains: Other companies on the market, marketing and sales, prior products of the company, the customer domain, the manufacturing domain, the sub-supplier domain and the disposal and recycling domain. The vertical bridging between abstraction levels and the horizontal bridging across domains are simultaneously involved in decision taking. For the sake of simplicity, not all levels involved in the co-operative decision-making are described within each design domain in figure 4. In what follows, a few examples of co-operative decision taking illustrate the common information needs that occurred during the project.

5.3. Examples of information needs

One of the goals of the project was to explore the feasibility of a "joint venture" co-operation with another company with expertise on hydraulics. The technical expertise of hydraulic valves was very likely to be available in another, not rivalling branch. Information was needed about *other companies on the market* that applied a similar technology. The automobile industry was found to have the necessary expertise on hydraulic valves. Once the industry had been identified, the co-operation and decision making also involved information about the goals of the

possible company candidates, their image, size, power, technological strategies, competence, other work functions outside the technical specifications of the project, their technical production conditions etc. The team developed a local classification scheme and classified the competence of possible company candidates for joint ventures (what size? what resources? openness to co-operation? etc). This supported their choice of priorities among company characteristics and helped to cope with complex information involving a high number of abstract attributes.

Another goal was to sell a successful product. Due to customers' complaints about instability of the control unit in extreme environments, it was necessary to improve the safety of the electronic control unit in the existing product. Information was needed about *the customers' attitudes to new technical solutions*. The *marketing and sales domain* needed information about the *customer domain* in order to match the properties of the product to user requirements, including the degree of safety and working quality required by the end users in different workplaces. Information was needed about customers' attitudes to new solutions and the costs that they were willing to pay for a high degree of safety and an improved quality of work. This also involved customer surveys, complaint reports, and materials on competitive products. The technical design solution chosen to fulfil this goal was the implementation of hybrid technology called chip on wire bonding in the electronic control unit. The result of this technical solution and choice of a new material was a change in the manufacturing process, where pasting replaced soldering.

Information about costs, the current capacity, competence of the staff and existing production equipment in the *manufacturing domain* was used to decide a new manufacturing process, where some components could be produced by the company itself, whereas other components would be purchased from sub suppliers. *Sub suppliers* contributed with information on the reliability and stability of a new production process using transistor tapes. Information was needed about sub suppliers and their competence and manufacturing expertise in soldering, pasting and transistor tapes. In order to provide access to information about people's expertise, a local classification scheme is needed that can be used to model and elicit experts' knowledge. The design of such a scheme can be based on the means ends abstraction hierarchy and the information needs and search questions expressed by design team members in a work situation, when they are looking for people with the expertise that they need to solve their problems. Likewise, the experiences gained from the marketing and purchase departments that have developed good skills in finding the proper expertise by applying a number of different criteria (Pejtersen, 1998, Hertzum and Pejtersen, 2000). The new materials to be used in the control unit should cause a minimum of disposal problems and therefore it was preferable to choose materials where test data were already available or could be easily gained at a low cost. Thus, in the *recycling*

domain test data were needed about different, optional materials that reduced disposal problems.

As exemplified above, the engineers' collaborative information needs can be organized in a means-end hierarchy reflecting the fact that the product involved can be described at several different levels of abstraction and within several different domains involved in the design work. It follows that in order to be useful in future work situations when the product is redesigned or when similar tasks are conducted, classification schemes for engineering design work should contain information about the product at all these levels and from all the different domain perspectives, as shown in figure 4. In order to address this requirement, it is necessary to identify the invariants of the questions that engineers ask to make decisions during team meetings in an ongoing project over a longer time span. A larger number of need formulations and a higher granularity of the information needs than those identified in this study will be necessary to represent the actors' formulations within the abstraction hierarchy.

6. Conclusion

The approach taken to the design of an ecological work based classification scheme for libraries was based on invariants identified in the policies of the library organisation and in the library users' information needs. If we generalize this approach to the engineering design domain, classification schemes should be based on the invariants found in the company policies (environmental and technical policies), in the laws of society (work regulations, environmental laws), and in the collaborative knowledge sharing and common information needs. Despite these similarities between the library work domain and the engineering design domain, identification of invariants in the design domain is a complex process because of the large number of knowledge domains that are involved. Generic and compatible classification schemes will have to cover the marketing domain, technical domain, the manufacturing domain, the disposal and recycling domain and so on. Further, they should be able to encompass the large-scale diversity of heterogeneous information sources.

An additional problem for the scheme designer is the dynamic and unpredictable nature of the design work. For the scheme designer, this for instance poses the challenge of how to cope with problems like invariant work domain structures versus changing information needs among actors due to changing constraints from the external environment. Another issue related to this problem is compatibility and integration of global schemes and local schemes and maintenance and re-design of schemes. It is likely that the means-ends abstraction hierarchy may serve as a generic, compatible model for analysis of information needs in co-operative work, but it is questionable whether one resulting classification system can encompass the totality of complexity in information

searching and knowledge sharing in co-operative work that is distributed in time and space.

Our approach for the design of ecological classification schemes is also intended as a contribution to the general discussion of future classification research. In particular, the ecological approach advocates priority to empirical and situated studies of activities in the work place. Additional sources to capturing domain knowledge are considered, for instance found in related domain studies by social studies of science. The uniqueness of the ecological approach is that it provides the designer with tools to combine conceptual analysis and empirical studies for domain analysis and scheme design.

Acknowledgements

We gratefully acknowledge the support of the Centre for Human-Machine Interaction (CHMI), Risø National Laboratory, Denmark and the Danish National Research Foundation who funded the research presented in this paper. We also want to thank our colleague at CHMI, Hans Andersen, for valuable comments of our presentation of Gibson's theory of visual perception. We are also grateful to Elaine Svenonius, University of California, Los Angeles, for valuable inspiration on classification philosophy. Finally, we want to thank Danfoss for hosting and participating in our field studies.

References

- Albrechtsen, H. & E. Jacob (1998). The dynamics of classification schemes as boundary objects for collaboration in the electronic library. In: *How classifications work: problems and challenges in an electronic age* edited by G. Bowker & S.L. Leigh Star. Library Trends 47:2, pp. 293-312
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* /4th edition. Washington, D.C.: APA. – 886 pp.
- Andersen, H.K. (1994). Classification schemes: supporting articulation work in technical documentation. *Knowledge Organization and Quality Management: Proceedings of the 3rd International ISKO Conference* edited by H. Albrechtsen & S. Ornager. – Frankfurt: Indeks Verlag, pp. 286-298
- Berg, M. (1999). Accumulating and Coordinating: Occasions for Information Technologies in Medical Work. *Computer Supported Cooperative Work (CSCW)* 8 (4), pp. 373-401
- Bowker, G.C. & S.L. Star (1999). *Sorting things out: classification and its consequences*. Cambridge (Mass): MIT Press

- Bowker, G.C. & S.L. Star (1991) Situations vs. standards in long-term, wide-scale decision-making: The case of the International Classification of Diseases. *Proceedings of the 24th Annual Hawaii International Conference on Systems Sciences, Kauai, Hawaii, January 7-11, 1991*. Edited by J. F. Nunamaker Jr. & R.H. Sprague. – IEEE Computer Society Press (volume IV).
- Carstensen, P.H. (1997). Towards information exploration support for engineering designers. *Advances in Concurrent Engineering/* edited by S. Ganesan. – Lancaster (PA): Technomic, pp. 26-33
- Cole, M. & Y. Engeström (1993). A cultural-historical approach to distributed cognition. In: *Distributed Cognitions/* edited by G. Salomon. Cambridge University Press, pp. 1-47.
- Davenport, E. (1998). Translating texts into care: classification issues raised by evidence-based practice in the UK health sector. In *Proceedings of the 9th ASIS SIG/CR classification research workshop/*edited by E. Jacob, pp. 53-69
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale (NJ): Laurence Erlbaum.
- Goodstein, L.P. & A.M. Pejtersen (1989). *The Book House: system functionality and evaluation*. Roskilde: Risoe National Laboratory. Risø-M-2793.
- Hertzum, M. & A.M. Pejtersen (1999). Information seeking practices of engineers: searching for documents as well as people. *Information Processing and Management* (forthcoming)
- Huber, J. & M.L. Gillaspay (1998). Social constructs and disease: Implications for a controlled vocabulary for HIV/AIDS. In: *How classifications work: problems and challenges in an electronic age/* edited by G. Bowker & S.L. Leigh Star. Library Trends 47:2, pp. 190-208
- Lindsay, P.H. & D.A. Norman (1977). *Human information processing*. New York: Academic Press
- McCarthy, L.P. & J.P. Gerring (1994). Revising psychiatry's charter document DSM-IV. In: *Written Communication* 11 (12), pp. 147-192
- McCloskey, J.C. & L.M. Bulechek (1995). Nursing intervention classification (NIC): development and use. In *Proceedings of the 6th ASIS SIG/CR classification research workshop/*edited by R. Schwartz et al, pp. 111-132

Pejtersen, A.M.; Andersen, H. & H. Albrechtsen (2000). *Cognitive Systems Engineering: development of theories and methods*. – Roskilde: Risø National Laboratory, Center for Human-Machine Interaction. (forthcoming)

Pejtersen, A.M. (2000): Design of Ecological Information Systems for Co-operative Work. *Proceedings of the xivth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society*. July 30-August 4, 2000, San Diego, California, USA (forthcoming)

Pejtersen, A.M. & Albrechtsen, H. (2000). Work domain analysis and the design of work-based classification schemes. In: M.E. Williams (ed): *Annual Review of Information Science & Technology (ARIST) 35*. – Medford (NJ): Information Today (in press)

Pejtersen, A. M.(1999): Icons for Organization and Representation of Domain Knowledge in Interfaces. In: *Lines of Thought in Knowledge Organization*. Festschrift in celebration of the ISKO's 10th anniversary/edited by W. Schmitz-Esser. Frankfurt: Ergon Verlag (in press). [Originally published in 1991, proceedings of the 1st International ISKO Conference]

Pejtersen, A.M. (1998). Semantic information retrieval. *Communications of the ACM* 41 (4), pp. 90-92

Pejtersen, A.M.; Sonnenwald, D.; Buur, J.; Govindaraj, T. & K. Vicente (1997). The design explorer project: using a cognitive framework to support knowledge exploration. *Journal of engineering design* 8 (3), pp. 289-301

Pejtersen, A.M. (1994). A framework for indexing and representation of information based on work domain analysis. In: *Knowledge Organization and Quality Management: Proceedings of the 3rd International ISKO Conference* edited by H. Albrechtsen & S. Ornager. – Frankfurt: Indeks Verlag, pp. 251-262

Pejtersen, A.M. (1986). Design and test of a database for fiction, based on an analysis of children's search behavior. In: *Information technology and information use. Towards a unified view of information and information technology*. Edited by P. Ingwersen, L. Kajberg and A.M. Pejtersen. London: Taylor Graham, pp. 125-146

Pejtersen, A.M. (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 O. Harbo & L. Kajberg. – London, Taylor Graham, pp. 146-159

Rasmussen, J.; Pejtersen, A.M. & L.P. Goodstein (1994). *Cognitive Systems Engineering*. New York: John Wiley. – 378 pp.

Rasmussen, J.; Pejtersen, A.M. & K. Schmidt (1991). *Taxonomy for cognitive work analysis*. – Roskilde: Risø National Laboratory. – Risø-M-2871

Rasmussen, J. (1986). *Information processing and the human-machine interaction*. – New York: North Holland.

Star, S.L. & Griesemer, J.R. (1989). Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19, 387-422

Suchman, L. (1994). Do categories have politics? The language/action perspective reconsidered. In: *Computer Supported Collaborative Work*, 2, pp. 177-190

Svenonius, E. (1992). Classification: prospects, problems and possibilities. In: *Classification Research for Knowledge Representation and Organization*. Edited by N.J. Williamson & M. Hudon. – Amsterdam: Elsevier, pp. 5-25 (keynote address)

Vicente, K. (1999). *Cognitive work analysis. Toward safe, productive and healthy computer-based work*. Mahrah (NJ): Laurence Erlbaum.

Vygotsky, L. (1980). *Mind in society*. Cambridge: Cambridge University Press.

Wittgenstein, L. (1953). *Philosophical investigations*. New York: MacMillan Company.

| Domain Level | Market: other Companies | Marketing & sales | Company's prior products | Customer domain (functions in use) | Manufacturing domain | Sub-Supplier domain | Recycling domain |
|---|---|---|--------------------------|--|--|--------------------------------------|--|
| Market, goals, agreements legislation | Product domains, their strategy, joint venture? Company image | Profit, company image, Board policies & commercial agreements | | Profit, company image, Working life qualities | Profit, company image, Union agreements, piecework, Working life qualities | | |
| Priority measures, Finance, safety, reliability | Technology, Competence, Size, Economy | Profit, market share, produce/purchase Recycling? | | Cost, safety, reliability, ease of use | Economy, waste disposal, capacity, safety, competence and production equipment | Reliability, Stability of production | |
| General functions, activities | Functions outside technical specifications | Information seeking Data collection Marketing survey | | Improved accuracy Improve safety | Production functions: assembling, testing, quality control, waste management | | Recycling of waste and scrapped products |
| Physical work processes | Similar production conditions in other domains / companies? Automobiles!! | Read reports on failures, complaints | | electronic controller fail-safe circuitry Improve reliability in harsh environment | Pasting replaces soldering | | Recycling processes for materials applied: - Hydraulic oil/water - Epoxy, solder - copper, silver |
| Physical resources & configuration | Technology applied Hydraulic valve | Knowledge transfer | | Hybrid technology: chip-on-wire bonding | Automatic assembling equipment, "transistor tapes" | "transistor tapes" | |

Figure 4: The cooperative workspace of engineering design

Empirical Work Analysis of Collaborative Film Indexing³

Hanne Albrechtsen, Annelise Mark Pejtersen & Bryan Cleal
Centre for Human-Machine Interaction
Risø National Laboratory
P.O. Box 49
4000 Roskilde, Denmark
hanne.albrechtsen@risoe.dk; annelise.m.pejtersen@risoe.dk;
bryan.richard.cleal@risoe.dk

Abstract

The present study explores document indexing as a collaborative work activity in the case of the creation of a filmography at a national film archive. The study, which was conducted according to the framework for cognitive analysis, identified two prototypical collaborative task situations in collaborative film indexing at the archive: screening meetings, where the subjects of films are analyzed in a collective, synchronous work arrangement, and ongoing consultation among filmography staff on entering information in a filmography database. The focus of the study is on how the current common conceptual tools like databases and cataloguing rules supported or failed to support the collaborating actors in contributing to the archive's overall priorities of mediating research to a diversity of user communities. A means-ends model for mapping the diverse perspectives is introduced. The paper contributes with ideas for the development of future common conceptual tools, based on the new means-ends model. Such common conceptual tools would provide a stable interpretive environment for collaborative film indexing, retaining the diversity of interpretive perspectives of films.

1. Introduction

The purpose of indexing is to facilitate subject access to documents in information retrieval systems through the creation of subject representations. Indexing is

³ In H. Bruce, R. Fidel, P. Ingwersen, & P. Vakkari, eds. (2002), Proceedings of the 4th CoLIS Conference. Greenwood Village, CO: Libraries Unlimited, pp. 85-107. Reprinted by the permission of the editors of the proceedings and the co-authors of the paper.

usually described in the Library and Information Science (LIS) literature as an individual work activity, which relies primarily on the indexers' experience and discretionary choices (e.g., Lancaster, 1991, Foskett, 1982). The indexing activity is explained as a stable procedure, comprising a number of steps, or task decisions: 1. interpretation or subject analysis of the object of work, the document; 2. translation of the interpretation to a linguistic statement or draft subject expression; 3. translation of the draft subject expression to subject expressions, using conceptual tools like a classification scheme or a thesaurus. However, recent contributions to indexing theory (notably, Hjørland, 1997, Pejtersen, 1994; Albrechtsen, 1993; Soergel, 1985) have argued that the first decision task of subject analysis is not necessarily dependent on stable procedures or on the indexer's cognitive structures or individual preferences. Rather it is argued that subject analysis and the subsequent steps of translating a subject analysis to a subject representation of a document are dependent on broader socio-cognitive factors. Such factors include, for instance, the indexers' knowledge of the socio-historical production context for a document and its relation to other documents (intertextuality), as well as the possible information needs of target groups that the document may serve.

The present study explores indexing as a collaborative work activity where socio-cognitive factors are continually articulated, discussed and negotiated in a professional work context. The study is based on empirical work analysis of collaborative film indexing at the national film archive, Národní Filmový Archiv (NFA), of the Czech Republic. The empirical work analysis of collaborative film indexing was carried out within current research on collaborative work contexts and activities in three European film archives (Pejtersen, Albrechtsen, Cleal, Buur Hansen & Hertzum, 2001). The overall aim of these work place studies is to contribute to the development of a future web-based research collaboratory for European film archives (Hertzum, Pejtersen, Cleal & Albrechtsen, 2002). The empirical work-studies of collaboration within and between the archives and their collaborating partners (individuals and institutions) have been conducted according to the framework for cognitive work analysis (Rasmussen, Pejtersen & Goodstein, 1994). The central emphasis of the present study is to contribute to a preliminary understanding of film indexing as a collaborative work activity, through an empirical analysis of its common workspace and activities, based on the framework of cognitive work analysis.

Section 2 introduces the methodology of cognitive work analysis, which is applied by the present study for empirical analysis of collaborative film indexing at a national film archive. Section 3 addresses the archive's current practice of collaborative film indexing in terms of the common work domain and its means and ends. Section 4 explores the collaborative indexing task situations of screening meetings and film description at the archive. Section 5 identifies the common resources applied by the archive for collaborative film indexing, with special emphasis on its common conceptual tools of database formats, cataloguing rules

and lists of indexing terms. The constraints of the current common conceptual tools applied by the archive are discussed in section 6. In section 7, a means-ends model of collaborative film indexing is introduced. The intent of this model is to provide a foundation for designing new common conceptual tools that can support continual joint access to film interpretations that are shared among collaborating indexers. Section 8 discusses the findings of the study in the light of recent contributions to indexing theory and makes some suggestions for application of the new means-ends model for constructing new conceptual tools for collaborative film indexing. Section 9 is the conclusion.

2. Methodology

The framework for cognitive work analysis is part of the Cognitive Systems Engineering (CSE) methodology for work-centered evaluation and design of information systems which was developed at Risø National Laboratory (Pejtersen, Sonnenwald, Buur, Govindaraj & Vicente, 1997; Rasmussen, Pejtersen & Goodstein, 1994). Figure 1 (the Onion Model) illustrates the overall units of analysis that are addressed in empirical work analysis, using the framework of cognitive work analysis. Following this model, *work domain analysis* identifies the overall *means and ends structure* of the work domain. This activity results in a means-ends representation of the work domain. According to the Onion Model (figure 1), collaborative work is analyzed in terms of the *collaboration, division of labor and social organization*. In this paper, the emphasis is not on a general analysis of collaborative work activities at the national film archive, but on analysis of the specific activity of collaborative film indexing. Consequently, this paper addresses collaboration primarily through *activity analysis* of collaborative task situations, in terms of collaborative work arrangements and collaborative task decisions in film indexing. In addition to analysis of decision-making activities in a work domain, the Onion Model addresses the actors' local competence to tackle situational work problems. Such competence can be found in the *strategies and heuristics* that the actors apply for coping with the work problems at hand. In collaborative work, the actors' strategies and heuristics are developed not only through individual work experience and formal professional background, but are also a function of collective factors, such as shared work problems and ongoing sharing of interpretations of work. In other words, the actors' competence to cope with situational work problems is also built up more informally through their choices of particular collaboration forms or work arrangements.

At the present stage, the empirical work analysis of collaborative film indexing at a national film archive primarily considers the following units of analysis:

- The common work domain content in terms of means and ends;
- Collaborative task situations, in terms of work arrangements and decision-making activities undertaken by the actors (situational dynamics);
- The information resources and tools required for the actors' coping with situational dynamics

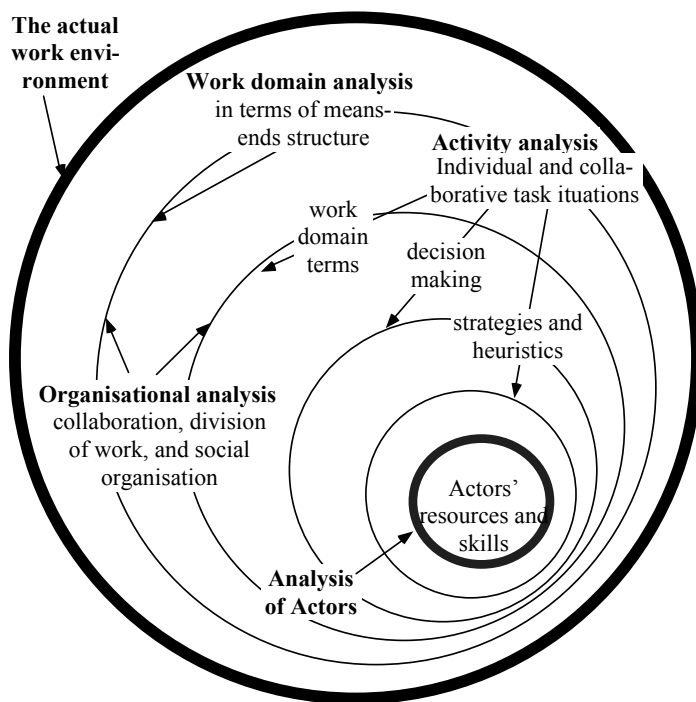


Figure 1: The Onion Model.

2.1 Addressing the common work domain

The present empirical analysis of collaborative film indexing builds on field studies at the Národní Filmový Archiv (NFA) (Pejtersen, Albrechtsen, Cleal, Buur Hansen & Hertzum, 2001). The field studies employed focus group interviews to gain knowledge about the common work domain and tasks of NFA, and the archivists' understandings of the means and ends of the common work domain. The study also draws on participant observation of work activity and interviews with archivists on collaborative indexing, with particular emphasis on the creation of a national filmography. The overall common work task, in which NFA is engaged, is research in Czech film history, based on a documentary research approach of collecting, storing, preserving and restoring national films. The research involves ongoing mediation of research results in the shape of a national filmography and through information services to and cooperation with other national research institutions. The work activity of collaborative film indexing at NFA is an important contribution to the creation of the national Czech filmography, and hence to NFA's ongoing film research activities.

The generic means ends model applied by the present study to analyze the common work domain of collaborative film indexing has five levels. The lowest level represents the resources and physical configurations that are involved in work. The next level describes the work processes in which the various resources and configurations are involved. Work functions are represented by more general concepts without reference to the work processes or resources that constitute a function. The fourth level presents the priorities that are used to coordinate resource allocation to the various general work functions. The highest level of abstraction contains the goals and constraints. In addition, the means-ends model of collaborative work domains in concurrent engineering (Pejtersen, Sonnenwald, Buur, Govindaraj & Vicente, 1997) has inspired the present study's analysis of the common work domain of collaborative film indexing. The means-ends model developed for concurrent engineering captures and represents the interpretive perspectives from different work domains that needs to be available for actors to cope with their knowledge and information exploration in engineering design decisions.

2.2 Addressing collaborative task situations

According to for instance Schmidt (1994), the purpose of collaborative work among actors in work domains is:

- augmenting of individual actors' capacity;
- providing for a differentiation and combination of specialties, through involving actors with different skills and knowledge; and/or

- opening up to mutual critical assessment or confrontation and combination of perspectives, where actors share multiple strategies and heuristics to solve a work problem or task decision

In complex and dynamic work domains, such as film archives, stable work procedures are not the norm. Work objects, like film materials, in such domains are amenable to multiple interpretations and applications. For instance, preservation of film materials requires assessment and description of their physical states. Information retrieval of films necessitates input of information in a film database, for instance addressing interpretations and descriptive attributes of films. Actors' collaborative work procedures vary, then, according to the work problem at hand, from preservation to information retrieval of films. Regardless of such situational dynamism, the general work problems addressed in a work domain can be contextualized within the stability of a means-ends representation of work, following the framework for cognitive work analysis (cf., for instance figure 2: means-ends representation of NFA). As a further contribution to analyzing stable patterns of collaborative work contexts, the framework for cognitive work analysis considers recurrent patterns of situational dynamics in terms of the following prototypical collaborative work arrangements (Rasmussen, Pejtersen & Goodstein, 1994):

Collective work in groups. This work arrangement is formed ad hoc to address a particular work problem or task decision. This work arrangement quickly dissolves into individual work.

Collaboration in teams. This work arrangement is formed in the long term and may address one or several work problems or task situations in collaborative work.

Consulting colleagues in individual work. This work arrangement is formed informally at a day-to-day basis among the actors. The criteria for choice of collaborating actor is either particular expertise or technical/domain knowledge or the closest colleague with respect to common experience with related work problems.

Analysis of decision tasks serves to determine the interpretive dynamics occurring during an individual or collaborative work task. Decision-making is more manageable when decomposed into possible subroutines shared among collaborating actors. For analysis of collaborative film indexing, the decision tasks identified may involve different interpretations of films among the actors and ongoing negotiation of interpretations of work. In addition, the actors' access to common work resources such as conceptual tools may constrain the decision-making activities in collaborative work.

2.3 Addressing common information resources and tools

Finally, cognitive work analysis considers the information resources and tools that the actors apply or develop for coping with work tasks in collaborative as well as individual task situations. Resources like human domain expertise on the one hand, and conceptual tools like cataloguing rules and database formats on the other hand, may shape as well as constrain the outcome of collaborative work, for instance the outcome of collaborative film indexing at NFA: the filmography. An important feature of work domains considered by cognitive work analysis is how such resources and tools support or fail to support the actors in coping with the work problems at hand, while simultaneously supporting the actors in meeting overall priorities of the domain. Hence, the purpose of the present study on collaborative film indexing at a national film archive is not only to analyze indexing as a collaborative activity, but also to explore how current resources and tools support collaborative film indexing so as to contribute to the archive's priority of documentary film research. Consequently, the study also investigates how new models or conceptual tools may support a higher degree of accordance between the goals of the overall work task of film research, and the outcome or product of the function related to this goal, collaborative indexing. The product of collaborative indexing is in the present case a national filmography. In order to explore new models and conceptual tools for collaborative film indexing, the present study applies the interpretive model for the related domain of fiction mediation developed by Pejtersen (Pejtersen, 1994, 1986). This latter model captures the stable semantic structure of the information required by actors to cope with their decision-making activities during indexing and information retrieval of fiction.

3. The common work domain of collaborative film indexing

The overall work problem addressed by NFA is film research. This overall work problem is reflected in the goals and priorities of the archive. The means-ends representation in figure 2 shows the five work domain dimensions (1-5) of NFA in terms of its (1) goals and constraints, (2) priorities, (3) functions, (4) work processes and (5) physical resources and configurations.

| | |
|---------------------------------------|---|
| Goals | Collect, archive, preserve and restore all film-related materials relevant for the Czech Republic. Ensure commercial viability, via exploitation of materials owned by NFA. National and international collaboration with archives and other professional institutions in and out of film domain. Develop international awareness of Czech film culture. |
| Constraints | Obligated to hold all films released in Czech Republic. International code of practice for cataloging. All income beyond 28% provided by state must be generated by NFA itself. Ownership of rights restricts commercial exploitation of archive information content. Annual audit by Ministry of Culture. Acquisition criteria. |
| Priorities | Dedication to collection building, documentation, preservation and publication of research results. Preservation and restoration of disappearing knowledge. Broaden appreciation for singularity of NFA collection. |
| Functions | Index and describe all materials of collection. Acquire new materials based on archive's own agenda. Publishing activities: production of annual Czech filmography , <i>Illuminace</i> (quarterly) and <i>Film Review</i> (monthly), film books. Collect material from amateur filmmakers, organize retrospectives nationally and internationally, and participate in international film festivals. Participate in international standardization activities in film indexing , restoration and mediation. Develop and implement new IT resources. |
| Processes | Two groups meet weekly to watch and analyze films. Input information in filmography database. Communication with three sites (buildings) via face-to-face contact, telephone, fax, and email. Cut, edit, and restore film material, restore and scan posters, identify and label new material. |
| Resources and physical configurations | Full length feature and short films , masters, print and negatives, videocassettes, books, journals, magazines, film scripts, digital laboratory, scanner and color printer, restoration equipment, workbenches and special chemicals, editing equipment. Cinema. Screening room. Computers, databases, local IT-network environment. Conceptual tools: database format, cataloging rules, list of indexing terms. Human resources: staff, colleagues outside the institution |

Figure 2: Means-ends representation of NFA; **Bold** refers to what is discussed in the text.

4. Collaborative indexing task situations

An important overall goal (1) of NFA is to collect, archive, and preserve and restore all film-related materials relevant for the Czech republic. This goal implies the priority of documentary research in terms of (2) collection building, documentation, preservation and publication activities. The production of a national filmography in Czech and English language is an important publication activity. This activity documents the archive's research activities and in addition contributes to the development of international awareness of Czech film culture, as reflected in NFA's overall goals (1). A number of functions (3) contribute to this priority, for instance indexing and description of films and film-related materials, acquisition of new material, participation in international activities in film indexing and description, and development and implementation of IT resources, such as databases. The work processes (4) of film indexing involve for instance weekly screening meetings and input of information in a filmography database. The physical resources (5) applied for the work processes involved in collaborative film indexing comprise full-length feature or short films, a screening room, computers, a database, a local IT-network environment and conceptual tools, such as a database format, cataloguing rules and lists of indexing terms. The common work domain for collaborative film indexing, then, consists of the above-mentioned elements in the goals, priorities, and functions, processes and physical resources of the overall means-ends representation of NFA (figure 2).

The objectives of the Czech national filmography (Opela, 1998) are to document research results by NFA and its collaborating national partners (universities, historical archives, film distributors, schools) on Czech film history. Thus, the production of the filmography is founded on a historical, documentary interest. Additional objectives include cultural mediation and dissemination of Czech film production at a national as well as international level. The primary target groups for the filmography are researchers, film students and film archivists. Important secondary target groups include film distributors, educators and other cultural mediators. The primary sources used to produce the filmography are the surviving film copies and negatives. Secondary sources comprise censorship documents, screenplays, advertisement materials and film periodicals. The information about primary sources (the films) is primarily obtained from discussions among NFA filmography staff and national film experts at weekly film screening meetings. The screening meetings constitute one of two prototypical work situations of collaborative film indexing, explored by the present study: (1) screening meetings, and (2) film description.

The overall justification for collaborative film indexing in NFA is to ensure that the production of the national filmography is carried out in compliance with the evolving body of knowledge and interests in Czech film production and mediation. Likewise, an important aim is to ensure that the film descriptions in the

filmography represent a balanced reconciliation and representation of such knowledge and interests. In the terminology of document indexing, the production of the filmography involves (i) capturing knowledge of the socio-historical production context for a film and its relation to other documents (secondary sources and other primary sources - 'film intertextuality'), and (ii) addressing the possible information needs of target groups that the film may serve. Currently, collaborative film indexing at NFA consists of two prototypical collaborative task situations:

(a) *Screening of film and collaborative subject analysis*: Weekly meetings at NFA where a film is screened, discussed and analyzed in a collective, synchronous work arrangement. A group of actors discuss and assess themes and socio-historical and -cultural contexts of the film.

(b) *Collaborative film description*: Entry of subject data and descriptive data by NFA's filmography staff, and carried out as distributed collaborative work. Each staff member functions as consultant for colleagues involved in the joint workspace.

4.1 Screening meetings and collaborative subject analysis

The prototypical task situation of film screening and collaborative subject analysis of films takes place at a screening room at NFA. This is an important event for collaborative film indexing, where national film experts outside NFA are invited as regular participants together with NFA's filmographers and database manager. The common workspace is the screening room, a film and an agenda for the meeting. The common work activity is to watch the film itself, participate in discussions, following a meeting agenda, and contribute with subject analyses of a film. There is a weekly screening meeting at NFA, and each screening meeting addresses one film only. The main purpose is to decide on the subjects of a film, through inviting contributions from all participants on the *overall theme(s)* as well as *interpretive contexts* of a film. Interpretive contexts addressed are the socio-cultural, -historical and technical backgrounds for the production of the film as well as subsequent reinterpretations and new versions of the film. Collaborative subject analysis of films is thus dependent on the collaborating actors' contributions on the contexts of a film across time, cultures, technologies and disciplinary paradigms and assessments of its potential for knowledge discovery, cultural mediation and contribution to new aesthetic forms. Further constraints for this work arrangement are the time schedule (one-day) and the reliability of experts involved. By the experts' reliability is meant: their ability to contribute to decisions on subject analyses and draft subject expressions that are balanced and robust enough to meet the diverse information needs of the filmography's target audiences.

The corresponding collaborative task decisions are 1. Subject analysis of a film
2. Formulating draft subject expressions.

4.2 Collaborative film description

Following the screening meeting, the staff communicates, mostly by phone, with the national film experts that participated in the screening meeting for further clarification of understandings, concepts, attributes and names of the film that was screened and discussed. The filmographers also coordinate data entry for the film with their colleagues within NFA. For this prototypical task situation, the common workspace involves the filmography, the archive's collections, a database structure, a list of indexing terms and the cataloguing rules of the International Federation of Film Archives (FIAF). Although the primary work object of this collaborative task situation is still the film itself, the attention focus of the staff shifts to the creation of filmography entries. In addition, there is a shift in the division of labor, in the sense that the decision tasks of film description are carried out by the staff only. Conversely, during the screening meetings, the decision tasks involve not only the staff, but also the invited participants in the meetings.

Corresponding collaborative task decisions are 1. Analysis of the draft subject expressions, formulated at the screening meeting, 2. Translation of this analysis into subject representations, 3. Planning filmographic entries, that is, decisions about what attributes of the database format to apply for describing a film, 4. Entering filmographic data and 5. Evaluating, reformulating and finishing the filmographic record.

5. Common resources applied

The common resources that are applied in collaborative film indexing at NFA comprise full-length feature and short films, screening room, computers, databases and local IT-network environment. In addition, the common resources comprise a set of conceptual tools for indexing: an *alphabetical list of indexing terms*, a *database format* and the *FIAF cataloguing rules for films* (Harrison, 1991). Furthermore, human expertise (staff, experts invited to screening meetings) constitutes an important common information resource in collaborative film indexing (cf. figure 2, means-ends representation of NFA, level 5: Resources and physical configurations).

NFA's filmography staff has created a *list of indexing terms*, using a bottom-up approach of ongoing indexing and cataloging of films. Management and development of the alphabetical list of indexing terms are negotiated between the database administrator and the filmographers to ensure uniformity and consistency of assigned subject representations. In addition to this list, NFA has compiled an overall list of 12 genres of films, reflecting NFA's overall typology of target groups. Presently, the semantic relationships between the indexing terms are not

made explicit through, for instance a linking to the list of genre terms. Nor are they linked through the structure of, for instance, a thesaurus or a clustering of semantically related terms.

| Descriptive levels (field segments): | Attributes (field names): |
|--|---|
| 1. Subjects of a film | Abstract Indexing terms Literary model Cast |
| 2. Communication of film formats | Types of film materials Censorship documentation Sound systems Sources |
| 3. Descriptive attributes, identifying the film as a unique work | Titles Names Location Year of production |
| 4. Provenance and mediation | Name of first distributor Production and film studies Premieres Awards |

Figure 3: Database format for NFA's filmography

Figure 3 shows the *database format* that is currently applied in task decisions for collaborative film description by the archive. This database format has been created to comply with the *FIAF cataloguing rules* (Harrison, 1991). The FIAF cataloguing rules provide much guidance on how to describe a film as a unique work, through its detailed examples for each possible descriptive aspect of a film. However, the FIAF cataloguing rules provide very little guidance on how to represent the *subjects* of a film. According to these rules, the decision task of creating subject representations is recommended to follow the structure and contents of universal classification schemes, in particular the Dewey Decimal Classification System, DDC (Miksa, 1998). Universal classification schemes, such as the DDC, are usually structured according to a fundamental division of scientific disciplines. Each discipline is subdivided in compliance with an assumed consensus of particular educational or disciplinary basic concepts or categories (Hjørland & Albrechtsen, 1999; Bliss, 1929). Presently, there are few standard classification schemes for the film domain (see for instance, Rasmussen, 1997; Turner, 1994; O'Conner, 1985). The classificatory structures of these schemes reflect similar divisions of knowledge, into basic concepts and categories of a particular scholarly consensus on division of film labor. Such classifications may fruitfully represent possible knowledge interests in films, from a scholarly or educational point of view (see e.g. Hjørland, 1997; Pejtersen, 1985). However, the films' potentials for cultural or emotional experience and understanding of

particular socio-cultural and -historical periods and contexts are rarely made explicit in the classifications through the presence of pertinent basic concepts. NFA's list of film indexing terms has been built to compensate for the current lack of standard conceptual tools like international classification schemes for films to support the decision task of subject representation of emotional, socio-cultural and -historical themes in films.

6. Constraints of current conceptual tools

NFA has been a major contributor to the development of the international FIAF cataloguing rules applied for film description, through its contributions to their creation and ongoing revision. Hence, the filmography staff at NFA is committed to particular work practices that can only be learned and developed within their particular field of work. The database format for the filmography has been designed to comply with the film attributes listed in the FIAF cataloguing rules. As standardized conceptual tools for filmography work, the film cataloging rules and the filmography database format applied by NFA could be said to reify particular professional practices and expertise developed within this field of work. Contrary to the stability and standardization of work practices and their reification in tools, the size of NFA's collections has grown rapidly over the last decade. Furthermore, NFA's services to their users are evolving from an almost exclusive focus on compilation and restoration of films and on-demand search requests from users, toward strategic research in the socio-cultural history of films and proactive collaborations that are targeted towards a diversity of professional user communities, at a national as well as international scale.

NFA's database format does not readily support subject representation of the diversity of film interpretations, brought forward by the participants at film screening meetings. By use of this conceptual tool for film description and information input, all interpretive contexts are currently merged into two database fields: a. indexing terms and b. abstract (figure 3). It should be mentioned that a merging of subject representations in database fields does not in general imply that interpretive context would be lost. Using an indexing approach of pre-coordinated or syntactic indexing, information about interpretive contexts can be expressed through the sequence of indexing terms or through the assignment of roles or links between the terms (Foskett, 1982). Alternatively, such contextual information can to some degree be made explicit through arrangement of indexing terms in term networks such as thesauri or clusters of associated terms. Presently, NFA applies a post-coordinate approach to indexing, which means that terms are combined in the search phase. This in turn means that information retrieval in the filmography database is currently constrained by the lack of explicit representations of the film interpretations brought forward and discussed during collaborative subject analysis at screening meetings. The current conceptual tools (database format, cataloging rules, list of indexing terms) applied by NFA for input of information in the

filmography database could be said to constitute a rather narrow conceptual cage for the rapidly growing and changing body of film knowledge. This makes it difficult to address the diversity of information needs among users searching the database.

The constraints of current conceptual tools applied for collaborative film indexing may be an important cause for a semantic drift observed in NFA's translation of subject analyses and draft subject expressions of films into subject representations of films. Semantic drift of interpretations of work is a major challenge in any collaborative work arrangement, involving actors with different skills and knowledge. In particular in loosely coupled work domains, where work objects are often characterized by high degree of interpretive flexibility or ambiguity (Pinch & Bijker, 1987). Examples in point are the design and implementation of large-scale computer-systems (Robinson & Bannon, 1991) and case handling (Gerson & Star, 1986). By semantic drift in the present case of collaborative film indexing is meant a shift in meaning, with respect to the extent to which subject representations (indexing terms) reflect or do not reflect particular interpretations of a film. Semantic drift in translating interpretations and subject analyses of documents to representations can be a consequence of a change in place of the object of interpretation, or work object (a film) and the interpretation itself (the subject analysis of a film). For instance a shift from the screening room at NFA to a filmographer's individual work place (see for instance, Robinson & Bannon, 1991; Gerson & Star, 1986). Likewise, a semantic drift can be due to a change of attention focus, in the present case, from a focus on a film as the work object to a filmography entry as the work object, or it can be due to a change in the division of labor (Schmidt, 1994). However, as will be discussed in the following, the semantic drift in the present case of collaborative film indexing is in particular due to the lack of affordances in existing applied conceptual tools to express an overall conceptual view of recurrent interpretive perspectives of a film.

7. A means-ends model for collaborative film indexing

Following the framework of cognitive work analysis, the common resources applied by the actors involved in collaborative work may constrain the actors' coping with a situational work problem. Conceptual tools like cataloging rules, database formats and classification schemes are a particular type of common resources, as symbolic artifacts (Schmidt, 1994) that provide collaborating actors with stable representations of shared work objects as well as shared interpretations of these objects. In the case of collaborative film indexing at NFA, the films and their interpretive contexts are the primary work objects for filmography production. Figure 4 illustrates the means-ends perspectives of collaborative film indexing (subject analysis and description), from original goals and values of a film (level 1) to archive attributes (level 5). The means-ends perspectives of collaborative film indexing (figure 4) are introduced here in a model that can be applied as a basis for

design of new conceptual tools for collaborative film indexing. A means-ends model that was previously created for the related common work domain of fiction mediation inspired the model. (Pejtersen, 1994; 1986). The compliance and difference in means-ends perspectives of subject analysis and description of films between NFA's current model, as implemented in NFA's database format for the filmography, and the means-ends perspectives of collaborative film indexing are also illustrated in figure 4. In this figure, the field attributes of NFA's database format are mapped towards the means-ends perspectives of the new suggested model for film indexing.

7.1 Decision tasks in collaborative subject analysis with reference to the means-ends model for film indexing

The collaborative decision tasks of film analysis during discussions at screening meetings at NFA currently address all levels of film knowledge, as represented in the means-ends model for film indexing (figure 4). The participating actors are experts representing a diversity of knowledge interests in film research and mediation. The collective and synchronous task situation of a screening meeting is intended to bring forward knowledge from all actors on overall theme(s) of a film as well as its interpretive contexts. While the actors at the screening meetings contribute with information on original titles, locations and production and film studios, a main emphasis is on exploring the overall theme of a film and the film's interpretive contexts over time. These perspectives or attributes of a film are in particular addressed in dimensions 1 and 2 of the means-ends model for film indexing (figure 4): 1: Goals and values of film, and 2: General and specific content in a film. Corresponding collaborative task decisions are (i) subject analysis and (ii) translation of subject analysis to draft subject expression.

7.2 Decision tasks in collaborative film description with reference to the means-ends model of collaborative film indexing

An important general hypothesis of the means-ends approach to conceptualization of work domains is that the objects and concepts of the work domain are ontologically distinct. The structure of knowledge is modelled in a stratified hierarchy, ranging from resources and physical configurations to high-level purposes that can be met by several resources and physical arrangements (Rasmussen, Pejtersen & Goodstein, 1994). The objects or concepts in one stratum are what can be applied as means to generate the end of a higher stratum. Such linkages can be identified in the means-ends model for collaborative film indexing. For instance, the physical characteristics of sound systems used in a film (level 3) can be regarded as means to give voice to the dialog among the characters of a film (level 4). Conversely, the promotion of particular ideas through the medium of a

film (level 1) can be realized through the choice of a particular time frame, subject matter content (plot), main characters and cast (level 2).

| | |
|---|---|
| 1. Goals and values of film | |
| Why? | (1) <i>Intentions and goals</i> : Communication of information; education; promotion and stimulation of ideas and emotional, socio-cultural and aesthetic experience; Literary model (novel, novella, short story, theatre play) <i>Constraints: censorship</i> , production means |
| Why? | (2) <i>Originators' affiliation and attitude</i> : Professional paradigms (narrative/aesthetic style, form, etc.) and their socio-historical context, political and cultural movements, or other value criteria |
| 2. General and specific content in film | |
| When? | (3) <i>General frame/time content</i> : Time, year, historical period |
| Where? | (4) <i>General frame/place content</i> : Place and setting, geographical, historical, socio-cultural contexts of the topic |
| What? | (5) <i>Subject matter content</i> : Specific topic and plots; psychological and social phenomena; ending of film; Contents (abstract of film), keywords (indexing terms) |
| Who? | (6) <i>Living beings, institutions and artifacts</i> : Main characters, persons, animals, plants, institutions that are involved in the topic Cast (names) |
| 3. Communication and presentation of film format | |
| How? | (7) <i>Film types and formats</i> Types of film materials |
| How? | (8) <i>Accessibility level</i> : Censorship documentation cards, contemporary documentation, sources, bibliography (reviews etc.),, indexes |
| How? | (9) <i>Physical characteristics</i> : Sound systems , etc. |
| 4. Filmography attributes | |
| Who? | (10) <i>Filmographic data and identification numbers</i> : Film titles, names of director, assistant director, story, screenplay, director of photography, set designer, art director, costume designer, editor, sound, introductory title designer, production manager, music and songs (composers, titles) |
| Where? | Locations |
| When? | Year of production , version |
| What? | Type of material |
| How? | Size and format |
| 5. Archive attributes | |
| Why? | (11) <i>Archive goals and policies</i> (12) <i>Local archival conditions</i> : Shelving, storage |
| When? | Preservation, restoration, sales and lending of films and videos |
| 6. Distribution, provenance, mediation | |
| Why? | (13) <i>Distributors', donators' etc. goals and policies</i> : Name of first distributor |
| Where? | (14) Production and film studios |
| When? | (15) Premieres (dates), awards |

Figure 4: Means-ends model for collaborative film indexing; italic/bold = attribute expression, applied in the NFA filmography

This quality of the means-ends model is applicable for the actors' decision tasks in collaborative film indexing of formulating subject representations and evaluation of a filmographic entry. For instance, an entry for the filmography database can be evaluated at any stage in the work process, through circumscribing a filmographic description to a linguistic statement. A means-ends representation of Gustav Machatý's film "Ecstasy" (English title), 1932, is provided in figures 5a-5b.

| 1. Goals and values of film | |
|--|--|
| Why? | (1) <i>Intentions and goals</i> : Women's sexual emancipation as a path to social and cultural liberation; personal freedom linked to gender equality in society; rebellion against previous models of relations between women and men; passion as original individual life force; the human being in a natural, instinctual state Literary model : drama Censored : 1933 (grounds: nudity, violation of traditional values of marriage and women's sexuality) |
| Why? | (2) <i>Originators' affiliation and attitude</i> : Cultural movement: Czech modernism; Socio-cultural background: youth rebellion; distancing from pre-nazi movement in Germany |
| 2. General and specific content in film | |
| When? | (3) <i>General frame/time content</i> : Time: 1930s |
| Where | (4) <i>General frame/place content</i> : Prague; bourgeois environment; country side; estate |
| What? | (5) <i>Subject matter content</i> : Women's emotions; marriage between young woman and elderly man; pedantry; divorce; young lover; passion; sex; nudity; husband's suicide; child birth Contents : see figure 5 b. |
| Who? | (6) <i>Living beings, institutions and artifacts</i> : Young woman; old man; father; young lover Cast/stars : Hedy Kieslerová (Eva); Zvonimir Rogoz (Emil Jerman); Aribert Mog (Adam) |
| 3. Communication and presentation of film format | |
| How? | Types of film materials : [omitted] |
| How? | (8) <i>Accessibility level</i> , Sources : [omitted]; Bibliography : [omitted] |
| How? | (9) <i>Physical characteristics</i> : Sound systems : n/a |
| 4. Filmography attributes | |
| Who? | (10) <i>Filmographic data and identification numbers</i> : Film title : Exstase (cz.); Director : Gustav Machatý; Assistant director : Alexander Hackenschmied; Original source : Gustav Machatý; Screenplay : Gustav Machatý; Frantisek Horký; Camera : Jan Stallich; Art director : Bohumil Hes; Editor : Antonín Zelenka; Sound supervisor : Josef Zora; Commentary written by : Frantisek Halas; Production supervisor : Frantisek Horký; Music composer : Guiseppe Becce |
| Where | Locations : n/a |
| When? | Year of production : 1932; Versión : Gustav Machatý, Elekta |
| What? | Type of material: n/a |
| How? | Size and format: n/a |
| 5. Archive attributes | |
| Why? | (11) <i>Archive goals and policies</i> (12) <i>Local archival conditions</i> Shelving, storage (n/a) |
| When? | Preservation, restoration, sales and lending of films and videos (n/a) |
| 6. Distribution, provenance, mediation | |
| Why? | (13) <i>Name of first distributor</i> : Slavia-film |
| Where? | (14) <i>Production and film studios</i> ; studios : Vinohrady External locations : Podkarpatská Rus; Prague; Brrandov; Slovensko |
| When? | (15) Premiere : 20 January 1933; Awards : n/a |

Figure 5a: Means ends representation of 'Ecstasy' by Gustav Machatý

| Re: Means-ends level 2, General and specific content of film |
|---|
| <p>Contents: Eva, the daughter of a landowner, marries the well off but many years older Emil Jerman. Upon returning from the marriage ceremony to their luxurious Prague apartment she realizes her mistake: her husband shows no interest in her as a woman. The sensitive girl suffers from his cold and bachelor's meticulousness so much that she returns to her father's estate and requests a divorce on grounds of insuperable revulsion. While bathing in the pond Eva meets Adam, a young civil engineer working close by on a highway construction site. The man is the exact opposite of her husband: though he looks robust, he acts gently. During a nighttime storm Eva goes after him to his house and becomes his lover. Only now she knows the passion of love. Her husband receives a court dispatch and sets off after Eva and beseeches her, to no avail, to return. On the way back he comes across Adam and realizes who he is. Both men are staying in a village inn where Eva comes to visit Adam. The husband commits suicide in his room. His death comes between the lovers. Adam returns to his work, Eva gives birth to a child.</p> |

Figure 5b: Content description of 'Ecstasy' by Gustav Machatý (from Opela, 1998, pp. 88-89)

Linguistic circumscriptions of means-ends film representations can iterate throughout the task decisions of collaborative film indexing. Thus, it can be assumed that a conceptual tool, based on the means-ends model for film collaborative indexing, can facilitate the work task, regardless of whether collaborative film indexing is carried out in a sequential or concurrent manner. The following example illustrates a linguistic circumscription of the means-ends representation of Gustav Machatý's film "Ecstasy" (English title), 1932 (see figures 4 and 5a-b):

"This film received its premiere on 20 January 1933, was produced at Vinohrady Studios and external locations of Prague etc., and is distributed by Slavia-film (level 6). Produced in 1932, its original title is "Extase" (cz.), director is Gustav Machatý, assistant director is Alexander Hackenschmied, etc. (level 4). The general and specific content of the film is the 1930s of bourgeois environment in Prague in conflict with the environments of an estate and a village on the countryside. The main characters are a young woman, elderly man, father, young lover (cast: Hedy Kieslerova, Zonomir Rogoz and Aribert Mog). The plot involves the topics of women's emotions, marriage between a young woman and an elderly man, pedantry, divorce, young lover, passion, sex, nudity, husband's suicide, child birth (level 2). The overall intentional value of the film is, through its genre as a drama, and coming from the director's affiliation with Czech modernism and sympathy with the youth rebellion, to promote the idea of women's sexual emancipation as a path to general social and cultural liberation. To show how personal freedom is linked to gender equality in society, to offer an alternative to previous models of relations between men and women, such as marriage, to show how passion is the original individual life force and to show the human being in its natural,

instinctual state. These intentional values were the major motivations for the censorship of the film in 1933 (level 1).”

In comparison, the corresponding entry for the film in the filmography (Opela, 1998, pp. 88-89) primarily addresses film content and attributes that are related to the levels 2-6 of the means-ends model (figures 4 and 5a-5b). For level 2 of the means-ends model (General and specific content in film), NFA's filmography entry neither addresses the general frame/time content, nor the general frame/place content. As a result it could be claimed that the socio-cultural and -historical contexts of the film's intentions and goals, which in the present case constitutes the overall theme of the film, are only marginally available as subject representations in the present filmography entry (Opela, 1998, pp. 88-89).

8. Discussion

The high degree of task uncertainty in subject analysis during the indexing activity in general has been discussed in the Library and Information Science (LIS) literature by in particular Hjørland (1997), Albrechtsen (1993) and Frohmann (1990). These contributions to indexing theory consider the socio-cognitive factors like production context of documents (e.g. scientific paradigms, discourse communities and disciplines) and the use of documents by particular disciplines or discourse communities as fundamental interpretive constraints and opportunities for the indexers' decision task of subject analysis. However, the decision task of translating a subject analysis to a subject representation will equally constrain or facilitate subject access. This latter problem has rarely been addressed in the LIS literature, except as generic recommendations (see for instance, Lancaster, 1991; Foskett, 1982). Recently, Mai (2001) has reintroduced the theme of translating a subject analysis and draft subject expression to a subject representation as a fundamental interpretive constraint for users' subject access to documents in information retrieval systems. Traditionally, the decision task of translating draft subject expressions to subject representations like indexing terms is considered an individual cognitive process of mapping between terms in linguistic draft subject expressions towards the conceptual tools of indexing languages like classification schemes, thesauri or alphabetical lists of indexing terms. According to Mai (2001), however, all decision tasks that are involved in document indexing, including the assignment of indexing terms, are primarily dependent on socio-cognitive factors such as the indexer's knowledge of a document's possible interpretive contexts. In other words, it is argued that interpretive contexts of a document may constitute a stable conceptual environment throughout all steps or decision tasks of document indexing. This contribution by Mai implies that regardless of an indexer's shift of attention focus, from a document during subject analysis, to application of indexing terms from indexing languages to represent the subject of a document, the *work object - the document and the indexer's interpretation of the document* - remains as

primary. In collaborative indexing, then, interpretive stability of indexing is not only driven by conceptual tools such as classification schemes or database formats, nor solely a function of consensus building among actors, but is equally determined by the indexers' access to a diversity of interpretations shared among collaborating actors.

The present study has addressed the theme of translating a subject analysis to indexing terms through an empirical analysis of collaborative film indexing activities at a national film archive. Using the framework for cognitive work analysis, the study identified two recurrent collaborative task situations: (a) screening of film and subject analysis, and (b) film description and information input. During the collaborative task situation of film screening, the film and its possible interpretations were a primary attention focus for the participants' joint subject analysis. During the subsequent collaborative task situation of film description, the actors' attention focus was found to shift to the filmographic entry. The actors' ongoing communication with colleagues inside and outside the archive during this latter task situation could, in the terminology of cognitive work analysis, be viewed as a recurrent strategy to retain the stability of a film analysis. Furthermore, because the task situation of film description involves clarification of analyses brought forward by participants at the meeting, it can be assumed that it may give rise to new possible interpretations of a film. The current practice of collaborative film analysis at the workplace studied was found to have the advantage of eliciting a diversity of interpretations of films. However, it was found that a semantic drift, or a shift in meaning, occurred when subject analyses and draft subject expressions were translated into subject representations. Such semantic drift may be due to the actors' ongoing reinterpretations of the film itself during their consultations with colleagues following a screening meeting. The present study has argued, however, that a primary constraint for sustaining interpretive stability in the empirical case of collaborative film indexing was the common conceptual tools currently applied for film description: the database format, the standard code of cataloguing practice and the list of indexing terms.

The study introduced a means-ends model for collaborative film indexing (figure 4). The intent of this model was to provide a foundation for designing new common conceptual tools for continuous support of collaborative film analysis and description. The current common conceptual tools for film indexing applied by the archive studied, in particular the database format and cataloguing rules, function as codes of practice for input of information in the archive's filmography database. Due to their lack of attributes to support the articulation of interpretive contexts of a film, these conceptual tools are not readily applicable for film analysis, for instance during discussions at film screening meetings. In other words, the affordances of current conceptual tools applied by the archive are not completely in alignment with the overall ambition of the archive's approach to documentary research, addressing a diversity of user needs. Alternatively, a new conceptual tool for collaborative film indexing, based on the means-ends model (figure 4), may

function not only to support the work processes of film interpretation and analysis as well as film description. A conceptual tool, based on the means-ends model, may also align the work processes of collaborative indexing with the archive's overall priority, or implicit indexing policy of addressing heterogeneous user needs. Possible applications and opportunities of the means-ends model include:

1. Making visible the state of work in collaborative indexing activities, ranging from film screening to production of filmographies. The conceptual tool can be introduced at screening meetings as a schema for taking notes on decisions made, and be used for iterations and evaluations of decisions if applied for minutes from the meeting. If the conceptual tool is made available for task decisions involved in collaborative film indexing, it can be applied for continuous backtracking and reassessment of filmographic entries.

2. Concurrent collaborative film indexing, where all actors participating in a screening meeting on a film have access to annotate filmographic entries with comments as the entries are developed locally in a filmography database, provided that the state of work is made visible to all actors throughout the production process.

3. Design of a classification scheme or a thesaurus, through using the perspectives of the means-ends model as a fundamental division of lists of indexing terms.

9. Conclusion

The empirical work analysis of collaborative film indexing has focused on the means and ends of the common work domain, the collaborative task situations, the conceptual tools applied and the decision activities of collaborative film indexing at a national film archive. Traditionally, indexing is considered an individual activity, relying on stable procedures and the indexer's preferences. Based on the framework for cognitive work analysis, the study addressed the indexing activity as a common work problem, shared by several actors within one work domain, engaged in the production of a national filmography. Several common points of reference and work practices were found to guide the actors' decisions. These common factors included the common work overall problem of national film research and mediation through the creation of a national filmography and the common resources applied. The study in particular focused on how the current common conceptual tools supported or failed to support the actors in coping with work problems at hand so as to contribute to the archive's overall priorities of addressing a diversity of user communities. An important aim of the empirical work analysis of collaborative film indexing was not only to understand the current work practice of the archive, but also to contribute with new ideas for the development of common conceptual tools that may articulate a stable interpretive environment for collaborative film indexing. It was argued that such conceptual tools could be developed on the basis of a means-ends model of the interpretive

perspectives of film mediation. Future work on development of conceptual tools for collaborative film indexing will focus on empirical evaluation of prototypes of conceptual tools, based on means-ends analysis of film mediation. This will look, in particular, at how annotations to filmographic data entries created locally at several film archives may involve collaborating actors across and outside archive sites and contribute to the creation of an open-ended and interactive film research collaboratory for European film archives.

Acknowledgements

This study was supported by the Danish National Research Foundation through its funding of the Centre for Human-Machine Interaction and by the European Commission's Information Society Technologies Programme through its funding of the Collate project on collaboratories for annotation, indexing, and retrieval of digitized historical archive material (IST-1999-20882). We thank our informants at Národní Filmový Archive, Prague, Czech Republic.

References

- Albrechtsen, H. (1993). Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, 18, 217-224.
- Bliss, H. (1929). *The organization of knowledge and the system of the sciences*. New York: Holt.
- Foskett, A.C. (1982). *The subject approach to information*. London: Clive Bingley.
- Frohmann, B. (1990). Rules of indexing: a critique of mentalism in information retrieval theory. *Journal of Documentation*, 46, 81-101.
- Gerson, E. & Star, S.L. (1986). Analyzing due process in the work place. *ACM Transactions on Office Information Systems*, 4, 257-270.
- Harrison, H. (1991). *The FIAF cataloging rules for film archives*. New York: K.G. Saur. (Papers and reference tools for film archivists dealing with audiovisual material, vol. 1).
- Lancaster, F.W. (1991). *Indexing and abstracting in theory and practice*. London: The Library Association.
- Hertzum, M.; Pejtersen, A.M.; Cleal, B. & H. Albrechtsen (2002). An analysis of collaboration in three film archives: A case for collaboratories. In H. Bruce, R.

Fidel, P. Ingwersen, & P. Vakkari (eds.), Proceedings of the 4th CoLIS Conference. Greenwood Village, CO: Libraries Unlimited.

Hjørland, B. & Albrechtsen, H. (1999). An analysis of some trends in classification research. *Knowledge Organization*, 26, 131-139.

Hjørland, B. (1997). *Information seeking and subject representation. An activity-theoretical approach to information science*. London: Greenwood Press.

Mai, J.E. (2001). Semiotics and indexing: an analysis of the subject indexing process. *Journal of Documentation*, 57, 591-622.

Miksa, F. (1998). *The DDC, the universe of knowledge, and the post-modern library*. New York: Forest Press.

O'Conner, B.(1985). Access to moving image documents: background concepts and proposals for surrogates for film and video works. *Journal of Documentation*, 4, 209-220.

Opela, V., ed. (1998). *Czech feature film II, 1930-1945*. Prague: National Film Archive.

Pejtersen, A.M.; Albrechtsen, H.; Cleal, B.; Hansen, C.B. & Hertzum, M. (2001). *A web-based multimedia collaboratory. Empirical work studies in film archives*. Roskilde: Risø National Laboratory (Risø-R-1284 (EN) - CHMI-02-01).

Pejtersen, A.M.; Sonnenwald, D.; Buur, J.; Govinderaj, T. & Vicente, K. (1997). The design explorer project: using a cognitive framework to support knowledge exploration. *Journal of Engineering Design*, 8, pp. 289-301.

Pejtersen, A.M. (1994). A framework for indexing and representation of information based on work domain analysis. Albrechtsen, H. & Ornager, S. (eds.): *Knowledge Organization and Quality Management: Proceedings of the International ISKO Conference, 3, held 20-24 June 1994, Copenhagen, Denmark* (252-262). Frankfurt: Indeks Verlag.

Pejtersen, A.M. (1986). Implications of users' value perception for the design of a bibliographic retrieval system. Agrawal, J.C. & Zunde, P. (eds.): *Empirical Foundations of Information and Software Science* (23-39). New York and London: Plenum Press.

Pinch, T.J. & W. Bijker (1987). The social construction of facts and artifacts: or how the Sociology of Science and the Sociology of Technology might benefit each

other." In: Bijker, W.E, Hughes, T.P. & T. J. Pinch (eds.): *The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology*. Cambridge (London): The MIT Press, 1987, 17-50.

Rasmussen, E. (1997). Indexing images, *Annual Review of Information Science and Technology (ARIST)*, 32, 169-196.

Rasmussen, J.; Pejtersen, A.M. & Goodstein, L.P. (1994). *Cognitive Systems Engineering*. New York: John Wiley.

Robinson, M. & Bannon, L. (1991). Questioning representations. *Proceedings of 2nd European Conference on Computer-Supported Cooperative Work, ECSCW 91, held Amsterdam, September 1991* (219-233). Amsterdam: Kluwer.

Schmidt, K. (1994). *Modes and mechanisms of interaction in cooperative work*. Roskilde: Risø National Laboratory (Risø-R-666-EN).

Turner, J. (1994). *Determining the subject content of still and moving image documents for storage and retrieval: an experimental investigation*. Toronto: University of Toronto (PhD thesis)

Summary in Danish

(Sammenfatning på dansk)

Arbejdsdomæneanalyse og arbejdsbaseret design er en ny tilgang til design af klassifikationsskemaer. Klassifikationsskemaer er symbolske artefakter, der viser og formidler begreber og strukturer i domæners semantik. Klassifikationsskemaer kan bruges til at søge og ordne fysiske og digitale samlinger. Forskning har vist, at arbejdsbaseret design af klassifikationsskemaer er vanskeligt, og at løsninger ofte er baseret på generiske konventioner for detaljeret design, der ikke bygger på behovsanalyse i arbejdsdomæner.

Denne afhandling omhandler kognitiv arbejdsanalyse og arbejdsbaseret design af klassifikationsskemaer til formidling af samlinger. Kognitiv arbejdsanalyse er en metodologi, der understøtter empirisk analyse af arbejdsdomæner og arbejdsbaseret design af informationssystemer. Kognitiv arbejdsanalyse sætter fokus på forståelse af strukturer og begreber i arbejdsdomæner og aktørernes praksis som designgrundlag. Forskning i klassifikationsskemaer til arbejdsdomæner omfatter også andre indfaldsvinkler, der enten lægger vægt på aktørernes begrebsdannelser i arbejdsdomæner eller på strukturer og begrebsdannelser i sociale fællesskaber som discipliner eller arbejdskollektiver.

Afhandlingen behandler to temaer indenfor arbejdsbaseret design af klassifikationsskemaer til formidling af samlinger: klassifikationsskemaer som sociologiske fænomener og kognitiv arbejdsanalyse, med særlig vægt på analyse af samarbejde. Temaerne definerer to forskningsspørgsmål der har dannet udførelsen af nærværende forskning. De to forskningsspørgsmål besvares af fem individuelle videnskabelige publikationer. Der findes ikke nogen anerkendt begrebsramme for arbejdsbaseret design af klassifikationsskemaer til formidling af samlinger. Derfor anvendes kognitiv arbejdsanalyses begrebsramme som udgangspunkt til arbejdsanalyse af formidling af samlinger og som afsæt til at kendetegne klassifikationsskemaers brug, indhold og strukturer i arbejdsdomænet. De fem videnskabelige publikationer bliver indplaceret i kognitiv arbejdsanalyses model for mål-middel analyse af arbejdsdomæner for at vise ligheder og forskelle i bidragenes fokus.

Fire empiriske kilder bidrager til resultaterne. Feltstudier af tre filmarkiver og et folkebibliotek adresserer samarbejdspraksis i formidling af samlinger. De primære metodikker anvendt i feltstudierne er deltagerobservation og interviews. Der er også brugt deltagende undersøgelsesdesign og empirisk evaluering af designresultater. Resultater fra feltstudierne diskuteres ud fra kognitiv arbejdsanalyse og inddrager også begrebsrammer udviklet indenfor forskning i datamat-støttet samarbejde og sociologisk teori om kommunikativ handling. Studierne om

formidling af samlinger i filmarkiverne fokuserer på kognitiv arbejdsanalyse af samarbejde og identifikation af empirisk grundlag for design af et web-baseret kollaboratorium for filmforskning, hvor klassifikationsskemaer kan indgå som støtte til formidling og udvikling af ny viden om film. Studiet af formidling af samlinger i folkebiblioteket fokuserer på samarbejdsbaseret design af klassifikationsskemaer som grænseobjekter i et web-baseret informationssystem til formidling af multimedie-materialer til børn og voksne.

De primære resultater af denne afhandling er:

1) Samlingernes brugspotentiale definerer form og indhold i klassifikationsskemaer til formidling. Brugspotentiale er ikke statisk, men konstrueres i samarbejde mellem brugere og personale i samlinger. Konstruktionen af brugspotentiale sker i kommunikative situationer som informationssøgning, indeksering og klassifikation af samlinger. Afhandlingen diskuterer forskellige teoretiske perspektiver på konstruktion af ordningssystemer i samarbejdspraksis og identificerer fordele og ulemper i perspektiverne for analyse og design af klassifikationsskemaer til formidling af samlinger.

2) En vigtig udfordring i arbejdsbaseret design af klassifikationsskemaer til formidling af samlinger er, at klassifikationsskemaer bruges i forskellige arbejdsprocesser, hvor deres affordances er afhængig af brugssituationen. Kognitiv arbejdsanalyse adresserer denne udfordring gennem modeller for analyse af prototypiske arbejdssituationer og mål-middel analyse af domænets strukturer. Afhandlingen stiller et forslag til en overordnet model, baseret på kognitiv arbejdsanalyses begrebsramme, for et fælles arbejdsrum for samarbejdsbaseret formidling af samlinger, hvor klassifikationsskemaer kan ses og anvendes af brugerne i kontekst.

3) En anden væsentlig udfordring i arbejdsbaseret design af klassifikationsskemaer er, at skemaernes indhold og strukturer skal have rod i arbejdsdomænets semantik. For at imødegå denne udfordring kræves analyse af domænesemantik, dvs. af domænets begrebsdannelser og strukturer, og en omsætning af analysen til design. Denne afhandling stiller forslag til en foreløbig metode til at omsætte analyse af domænesemantik i samarbejdsbaseret formidling af samlinger til et klassifikationsskema med en fast struktur.