Collaborative Knowledge Visualisation for Cross-Community Knowledge Exchange

Von der Fakultät für Ingenieurwissenschaften der
Universität Duisburg-Essen

zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften
genehmigte Dissertation

von

Dipl.-Ing. Jasminko Novak

aus Derventa

Referent: Prof.Dr-Ing. Jürgen Ziegler
Korreferent: Prof.Dr. Heinz Ulrich Hoppe
Tag der mündlichen Prüfung: 16. Oktober 2006

Abstract

The notion of communities as informal social networks based on shared interests or common practices has been increasingly used as an important unit of analysis when considering the processes of cooperative creation and sharing of knowledge. While knowledge exchange within communities has been extensively researched, different studies observed the importance of cross-community knowledge exchange for the creation of new knowledge and innovation in knowledge-intensive organizations. Especially in knowledge management a critical problem has become the need to support the cooperation and exchange of knowledge between different communities with highly specialized expertise and activities. Though several studies discuss the importance and difficulties of knowledge sharing across community boundaries, the development of technological support incorporating these findings has been little addressed.

This work presents an approach to supporting cross-community knowledge exchange based on using knowledge visualisation for facilitating information access in unfamiliar community domains. The theoretical grounding and practical relevance of the proposed approach are ensured by defining a requirements model that integrates theoretical frameworks for cross-community knowledge exchange with practical needs of typical knowledge management processes and sensemaking tasks in information access in unfamiliar domains. This synthesis suggests that visualising knowledge structures of communities and supporting the discovery of relationships between them during access to community spaces, could provide valuable support for cross-community discovery and sharing of knowledge. This is the main hypothesis investigated in this thesis.

Accordingly, a novel method is developed for eliciting and visualising implicit knowledge structures of individuals and communities in form of dynamic knowledge maps that make the elicited knowledge usable for semantic exploration and navigation of community spaces. The method allows unobtrusive construction of personal and community knowledge maps based on user interaction with information and their use for dynamic classification of information from a specific point of view. The visualisation model combines Document Maps presenting main topics, document clusters and relationships between knowledge reflected in community spaces with Concept Maps visualising personal and shared conceptual structures of community members. The technical realization integrates Kohonen's self-organizing maps with extraction of word categories from texts, collaborative indexing and personalised classification based on user-induced templates. This is accompanied by intuitive visualisation and interaction with complex information spaces based on multi-view navigation of document landscapes and concept networks.

The developed method is prototypically implemented in form of an application framework, a concrete system and a visual information interface for multi-perspective access to community information spaces, the Knowledge Explorer. The application framework implements services for generating and using personal and community knowledge maps to support explicit and implicit knowledge exchange between members of different communities. The Knowledge Explorer allows simultaneous visualisation of different personal and community knowledge structures and enables their use for structuring, exploring and navigating community information spaces from different points of view.

The empirical evaluation in a comparative laboratory study confirms the adequacy of the developed solutions with respect to specific requirements of the cross-community problem and demonstrates much better quality of knowledge access compared to a standard information seeking reference system. The developed evaluation framework and operative measures for quality of knowledge access in cross-community contexts also provide a theoretically grounded and practically feasible method for further developing and evaluating new solutions addressing this important but little investigated problem.

Acknowledgements

The work presented in this thesis has been undertaken during my time as a researcher at the Fraunhofer Institute for Media Communication. Parts of the reported research were conducted in the collaborative research project “AWAKE – Networked Awareness and Knowledge Discovery” led by the author and supported by the German Federal Ministry of Education and Research. (BMBF). The project “CAT/netzspannung.org – an Internet Media Lab for Communication of Art & Technology” (BMBF) also provided an important experimental context for testing the developed methods and solutions.

I would like to thank my supervisor Prof. Dr. Jürgen Ziegler for his support, critical feedback and stimulating discussions: they helped me in distilling the richness of the developed ideas and experimental findings into mature scientific work. His openness to new ideas paired with his great experience in the design and evaluation of interactive system were a valuable companion on the often unpredictable journey that has led to this thesis.

My thanks also go to Monika Fleischmann, Head of the MARS Exploratory Media Lab at the Fraunhofer IMK for her support during the course of this work. We have passed together many a storm and I was privileged to learn from her ceaseless enthusiasm in crossing the borders between different worlds of knowledge and fields of profession. I further thank Prof. Dr. Martin Reiser, the IMK Institute Director for his support of the project and of my work as well as the whole IMK staff. Many thanks to Prof. Dr. Gerhard Schwabe from the University of Zurich for his comprehension and the opportunity to deliver the finishing touches to this thesis after taking on my new position in his group.

This work would have not been possible without the help of many people. In particular, I have enjoyed the collaboration with many great people in the AWAKE project. I would like to thank the responsible partners Prof. Dr. Jürgen Ziegler (Fraunhofer IAO), Prof. Dr. Katharina Morik (University of Dortmund) and Prof. Dr. Bernd Freisleben (University of Siegen) for their great dedication and the belief in the project. During my stay at the University of Milano-Bicocca, I have also had the privilege to discuss and confront my ideas with Prof. Dr. Carla Simone, Prof. Dr. Giorgio de Michelis and Prof. Dr. Federico Butera.

My special appreciation goes to all the members of the project team who have contributed to the development of the AWAKE system and helped make the project the success that it has become. I have enjoyed many fruitful discussions with Michael Wurst (University of Dortmund), Stefan Paal (Fraunhofer IMK) and Christoph Kunz (Fraunhofer IAO) which provided inspiration for some of the important ideas of this work. Michael has also implemented the supervised classification modules and provided valuable advice on knowledge discovery methods. Stefan provided the basis for the system middleware and valuable help in keeping the system up & running.

Kresimir Simunic (Fraunhofer IMK) supported the implementation of the Kohonen map, Reni Banov (Db-IT Consulting), Christoph Seibert (Fraunhofer IMK), Roger Sennert and Jens Wagner (University of Siegen) contributed to the realization of system integration services. Early interface designs have benefited from the support by Danijela Djokic and Hartmut Bohnacker (projekttriangle). The final implementation of the Knowledge Explorer and underlying services was supported by the students Ansgar Himmel, Philipp Wever, Kerstin Schmidt, Stefan Winarzki and Malte Kloes. My long-time friends, Predrag Peranovic and Kresimir Simunic also provided valuable logistic assistance and were great hosts during my frequent Milano-Cologne travels.

I also thank all the colleagues from the Fraunhofer IMK and the Dept. of Interactive Systems at the University of Duisburg-Essen for their help with the preparation and execution of the laboratory experiments. I am indebted to all test persons for their participation in rather extensive experiments which made the empirical verification possible.

I am grateful to all my friends and colleagues at the Fraunhofer Campus and the MARS Exploratory Media Lab for their companionship in these years. I will kindly remember the many encounters and conversations with Gabriele Blome, Carla Valle, Martha Larson, Markus Klann and many others. I thank all the friends in Cologne, Bonn, Rome, Milan, Zagreb, Zurich, NYC and beyond - fortunately, they are way too many to be mentioned here by name!

Last but not least, I thank my parents for their unconditional and never ceasing support, my sister Maja for not letting me forget the lighter side of life, Francesco for the nightly scooter tours of Rome and Claudia, for her endless love.

Table of Contents

1. Introduction

Communities and knowledge. The notion of communities as informal social networks has been increasingly used as an important unit of analysis when considering the processes of cooperative creation and sharing of knowledge (Brown & Duguid, 1991; Wenger, 1998; Wenger & Snyder, 2000). Important types of such communities include communities of interest (Rheingold, 1993) and communities of practice (Lave & Wenger, 1991; Wenger, 1998). Communities of interest emerge as groups of like-minded people engage into communication and exchange of information around a common topic of interest (e.g. software development, health, hobbies etc.) through mailing lists, information repositories and discussion forums.

Communities of practice develop as people working on similar problems engage into helping each other, sharing experiences and discussing their work practice (Wenger, 1998). While they have been initially studied in shared physical workplaces (Wenger, 1998; Orr, 1996), with the increasing distributedness of modern organizations the concept of communities of practice has been extended to encompass informal networks which connect geographically distributed participants. Accordingly, much research has been devoted to the development of methods, tools and systems for supporting knowledge creation and sharing in distributed communities in fields as diverse as knowledge management (Gongla & Rizzuto, 2001; Wenger & Snyder, 2000), CSCW (Ishida 1998; Schlichter et al. 1998; Graether & Prinz, 2001) and Semantic Web (Davies et al., 2003).

Cross-Community interactions. Existing research has thereby largely focused on supporting knowledge sharing within relatively homogeneous communities that connect participants with similar backgrounds, practices and fields of expertise. At the same time, knowledge-intensive work in modern organizations increasingly involves complex tasks whose fulfilment requires the exchange of knowledge between participants from different fields of expertise. This typically involves both the need to take advantage of existing knowledge and experiences made in different places in an organization as well as the need to create new, previously not available knowledge (Nonaka & Takeuchi, 1995).

Studies in organization sciences, learning research and knowledge management have also pointed out significant shortcomings and limits of intra-community creation and sharing of knowledge (e.g. Swan, 2001). Just as the shared (local) language, interests and practices constitute a community and facilitate the creation and sharing of knowledge between its members, they also divide and close it from others. As a result, new knowledge and innovation in communities tend to be incremental improvements within the same knowledge perspective (Swan, 2001, 2003). Fundamentally new insights or knowledge which might be relevant both for the existing problems at hand as well as for long-term development, can only be acquired if ways are provided for the community to learn from experiences of other communities. Different theoretical and empirical contributions emphasise the importance of such cross-community interactions as a critical source of new knowledge required for solving ill-defined and complex problems (Brown & Duguid, 1991; Dougherty, 1992; Boland & Tenkasi, 1995).

Cooperative knowledge work. A critical problem for knowledge management has thus become the need to support the cooperation and exchange of knowledge between different communities with highly specialized expertise and activities (Purser & Pasmore, 1992; Brown & Duguid, 1991; Boland & Tenkasi, 1995). In scientific research this is particularly visible in efforts to connect expert knowledge from different research communities in order to address critical real-world problems (e.g. information society technologies, sustainable development, renewable energies and life science). In commercial organisations acting in highly turbulent environments, this is reflected in the need to balance two opposing requirements in developing new products and services: the need for high specialization of work with the need for flexible and effective inter-group knowledge exchange (Dougherty, 1992; Boland & Tenkasi, 1995; Mark et al., 2002; Fuchs-Kittowski et al., 2003).

Cross-Community portals. A common approach to this problem has been the establishment of shared community platforms and knowledge portals (Internet/Intranet) aiming at providing one central point of encounter and knowledge workspace for different communities. Examples are corporate knowledge portals in large commercial organizations and cross-community platforms in research settings, such as the ACM portal[1], the EU MOSAIC Network, the FONA platform[2] or the interdisciplinary knowledge portal netzspannung.org, to name but a few. The main idea of such shared platforms is to provide the missing common ground and an initial place of encounter for stimulating cross-community interaction and flows of knowledge. Providing cross-community access to community repositories externalizing community knowledge allows members from different communities to become aware of each others knowledge and learn by sharing information.

However, appropriate support for such platforms is still missing. They are typically based on a combination of centralized information repositories with standard community-tools that have been developed for supporting exchanges in teams and *within* communities (e.g. shared workspaces, awareness, online communication and messaging).

Research challenge. In contrast, different studies of the problem of knowledge sharing across community boundaries (Dougherty, 1992; Boland & Tenkasi, 1995; Wenger 1998, Brown & Duguid, 1998; Swan, 2001) identify special challenges and requirements that need to be considered in such heterogeneous situations. Similarly, the work on sensemaking in human-computer interaction has investigated the processes of knowledge construction during information seeking in unfamiliar domains and ill-structured problems (Dervin, 1992; Russel et al., 1993). While both of these lines of research highlight issues that are at the heart of the problem of supporting cross-community knowledge exchange, the development of appropriate systems and tools incorporating these insights is still lagging behind (Swan, 2001).

Against this background, this thesis addresses the following challenge: How can we design interactive systems and tools that incorporate these insights, in order to support and stimulate cross-community sharing and creation of knowledge?

In particular, the work presented in this thesis is based on two main observations:

1. The problem of developing technological support to specifically support cross-community knowledge exchanges has been largely under-investigated. Although theoretical foundations exist, current knowledge management and community-support solutions are neither theoretically grounded nor empirically evaluated with respect to the specific cross-community application context.
2. The potential contribution of the HCI perspective to developing methods and interactive tools for addressing the problem of cross-community knowledge exchange has been scarcely considered. Especially the potential of knowledge visualisation (Tergan & Keller, 2005) and sensemaking support (Russel et al., 1993, Qu & Furnas, 2005) for information access in such complex settings, have received comparably little attention.

Accordingly, this thesis proposes a novel method and an interactive system developed specifically for supporting cross-community knowledge exchange. Specific requirements that need to be satisfied for addressing this challenge are identified and a possible solution based on an integration of techniques from knowledge discovery, knowledge visualisation and human-computer interaction is proposed. In doing so, the thesis shows how the HCI perspective adopted by the proposed approach contributes to and enriches the existing methods and approaches to knowledge sharing from knowledge management and computer-supported collaborative work.

[1] http://www.acm.org
[2] http://www.fona.de

1.1. Problem Description: Supporting Cross-Community Knowledge Exchange

Collaborative information spaces and knowledge exchange. A major modality of knowledge exchange in distributed communities is the construction and use of shared information spaces (Borghoff & Schlichter 2000) such as shared document repositories and discussion forums. Applying Nonaka & Takeuchi's model of collective knowledge creation (Nonaka & Takeuchi, 1995) allows us to understand that effective knowledge exchange mediated through shared access to information resources involves both the exchange of explicit and tacit knowledge. Through externalization in information artefacts (documents, postings, emails) contributed to the collaborative information space, community members make available their explicit knowledge to each other. By accessing and interpreting the information contributed by others the gained insights are internalized as new knowledge (internalization).

Social theories of learning demonstrate that in order for groups of people to collectively construct and share knowledge, they have to establish a shared cognitive and social context against which they can construct shared meanings of information (Lave & Wenger, 1991; Luckmann & Berger, 1966; Bruner, 1990). Hence, sharing tacit knowledge is an important prerequisite for the exchange of knowledge mediated through sharing information. In communities, such a shared context is given on one hand by a shared domain of knowledge and interests, shared practices and a common community language. This provides members of a community with a common ground of understanding and a shared frame of reference for interpreting and expressing the meaning of shared information.

Thought worlds and interpretive perspectives. The core problem of cross-community exchanges lies in the absence of such a shared social and cognitive context. Different communities inhabit different "thought worlds" (Dougherty, 1992) or "interpretive schemas" (Boland & Tenkasi, 1995) which determine how their members interpret the meaning of information, artefacts and experiences. They are characterized not only by different terminologies but by different systems of meaning. In other words: knowledge artefacts produced by members of different communities (documents, emails, forum discussions) are not a neutral organization of information, but reflect knowledge structures and interpretative schemas of both their authors and the communities in which they have been created.

Different thought worlds and interpretive perspectives make it difficult for knowledge to cross community boundaries as information loses its original context. As a result, knowledge cannot be simply "shared" between different communities by exchanging information. In order to make sense of information and construct knowledge, one needs to contextualise it within one's own existing knowledge and thought world. To achieve this, members of different communities need to construct some form of understanding of an unfamiliar community context and its relationships to the familiar common ground of one's own community (ibid.).

Accessing explicit knowledge contained in the information space of an unfamiliar community is possible only if there exist ways for gaining insight into shared tacit knowledge of the community in question. According to Nonaka & Takeuchi's model in order for this to be established and maintained, either frequent interaction between community members (socialization) or engagement in following the community discourse reflected in the shared information space (internalization) are needed. In distributed communities socialization requires extensive informal online communication (chat, virtual spaces, online meeting rooms) while internalization requires frequent access to the community information space (reading documents and contributions in discussion forums, document repositories etc.).

Heterogeneous knowledge networks. While such processes normally occur within communities, they are not appropriate for cross-community contexts. Members of different communities rarely interact directly with each other and when they do, the interaction is very narrow, with the purpose of obtaining otherwise unavailable information (Alee, 2000). This occurs within heterogeneous knowledge networks that form around the intersections between the domains of problems and knowledge which are in parts shared by different communities. But such cross-community networks are invisible because there is no explicit membership as participants remain anchored in the different communities to which they belong.

They manifest themselves only as members of different communities engage in accessing information from each other's community spaces as a way of accessing complementary, otherwise unavailable knowledge (ibid.).

As a result the socialization processes responsible for the establishment of a common ground within communities cannot take place. Similarly, since the need for accessing information from unfamiliar communities occurs only very occasionally, based on a very specific need there is no lack of continuity of shared interests. Accordingly, members of one community do not have the motivation nor the resources for intensively following the information exchange in other communities. This makes it difficult to establish the shared context necessary for understanding the knowledge contained in information spaces of unfamiliar communities.

Ill-defined information needs and sensemaking. The need for accessing knowledge from unfamiliar communities occurs in practice when people face complex or ill-structured problems that do not fit within their known knowledge contexts. Such situations are very common in interdisciplinary work, scientific research and innovation processes in modern organizations (Swan, 2001; Fuchs-Kittowski et al., 2003). They require making sense out of an ill-defined problem context: understanding its structure and related knowledge domains in order to identify the knowledge gap and the relevant fields of expertise needed for a solution. This involves seeking information from different, often unfamiliar domains. Examples of such classes of tasks include strategy making, development of new products and project engineering (Fuchs-Kittowski et al., 2003;) , training and education as well as business intelligence (Russel et al., 1993) and self-directed learning (Qu & Furnas, 2005). All of these tasks require learning about potentially relevant but unfamiliar topics, knowledge domains and areas of expertise.

Consequently, in cross-community contexts the information need is very ambiguous and difficult to resolve by means of goal-directed search. Not only does the terminology problem play a critical role, but making sense of the problem space in order to translate it into a specific information need occurs during the information seeking process itself. This results in highly explorative nature of cross-community information seeking tasks for which the explorative access and the availability of appropriate knowledge structures to guide it (e.g. categorization, classification schemes) play an important role (cf. the sensemaking paradigm, Russell et al., 1993).

Perspective making and perspective taking. The basic requirements for supporting cross-community knowledge exchanges have been most succinctly described in the model of "perspective making – perspective taking" proposed by (Boland & Tenkasi, 1995). Perspective making refers to the processes through which members of each community express, develop, organize and exchange knowledge in the ways appropriate to the shared community perspective (interests, domain, vocabulary). This includes activities such as shared problem solving, document writing and sharing, cooperative gathering and structuring of information, discussions and narrative exchanges in online forums - supported by a number of existing community-support tools & systems (Ishida, 1998; Graether&Prinz, 2001).

The notion of perspective taking emphasises the need for developing an understanding of implicit knowledge perspectives underlying an unfamiliar community in order to be able to exchange knowledge with its members. Thus, constructing knowledge from information in unfamiliar community spaces requires that some support for understanding the inherent knowledge structures and the context within which the meaning of information in that community is defined, be provided. Subsequently, in order to internalize knowledge and make it one's own, one needs to contextualize it within one's own thought world and express it in one's own terms. This again is the process of perspective making. Hence, the exchange of knowledge between different communities can occur only if an interplay of the processes of perspective taking and perspective making can take place: understanding of "what and how the others know" (Boland & Tenkasi, 1995; Swan, 2001), finding out how this is related to one's own knowledge and translating it into one's own terms.

Shortcomings of existing community support. Perspective making is well addressed by a range of community support solutions: from discussion forums and shared workspaces (e.g. Bentley et al., 1997) to collaborative filtering and recommender systems (e.g. Glance et al., 1998; Kautz et al., 1997), to cooperative annotation and indexing (e.g. Gräther & Prinz, 2001) to community-based ontologies (e.g. Christophides et al., 2000; Davies et al., 2003). On the contrary, the support for processes of perspective taking has been very limited up-to-date. Common approaches aim at creating standardized taxonomies or organization-wide ontologies as a means for making knowledge accessible across the boundaries of individual departments and communities (Bonifacio et al., 2002). Such focus on the establishment of a shared, unified view comes short of the needs for supporting simultaneous co-existence and coordination between different local perspectives (ibid., Swan, 2001).

On the other hand, approaches to explicit construction of community ontologies (e.g. Christophides et al., 2000) are limited by the need for negotiating and achieving community consensus. Similarly, methods for ontology mapping allow the discovery of relationships between different conceptual schemes but require explicitly formulated and formally modelled ontological structures (see overview in Madhavan et al., 2002). They do not consider the highly implicit nature and social characteristics of community knowledge creation and sharing, nor provide means for understanding the mapping criteria needed for perspective taking to occur.

In summary, supporting cross-community knowledge exchange is a complex problem where the main difficulties are posed by the differences in the shared contexts of understanding of different communities and the absence of explicit knowledge structures representing community knowledge. In consequence, the lack of mechanisms for supporting simultaneous co-existence of different community perspectives, their natural evolvement and the discovery of relationships between them makes it difficult for members from different community to exchange knowledge by accessing unfamiliar community information spaces.

1.2. Solution Approach: Knowledge Visualisation for Multi-Perspective Access to Community Spaces

This work proposes an approach to addressing the described problem from the perspective of knowledge visualisation. The merits of information visualisation focusing on visualising structure, patterns and relationships in abstract data and information as a means for amplifying human cognition (Card, Mackinlay & Shneiderman, 1999a) have been long recognized. Knowledge visualisation extends this paradigm with looking specifically at how visualisation can be used to support the creation and exchange of knowledge between human users (Eppler & Burkhardt, 2004). The underlying paradigm is that cognitive processing of complex subject matter and access to information and knowledge resources can be enhanced by making the underlying structures behind ideas, knowledge and information explicit (Keller & Tergan, 2005).

The idea of using visualisation to support individual and cooperative knowledge management has been gaining attention in both fields, leading to a convergence and synergy of methods from information and knowledge visualisation (ibid.). On one hand, methods for visualising inherent knowledge structures from textual collections combine text mining methods such as topic extraction, text categorization and document clustering with visualisation methods such as document maps, tree maps and graph visualisation (Chen, 1999; Card, Mackinlay & Shneiderman, 1999b). Visual overviews of similarity structures of document collections have been successfully used for specialized knowledge management analysis tasks (Becks, 2001). Such visualisations provide means for easily identifying main topics, clusters and associations in a domain of knowledge represented by a given document collection.

On the other hand, knowledge visualisation approaches such as concept mapping (Novak & Gowin, 1984; Canas, Leake & Wilson, 1999) have demonstrated possibilities for eliciting and visualising conceptual structures of individuals and groups of users. Such methods visualise main concepts and relationships between them representing the way an individual or a group of users view a certain problem or structure a domain of knowledge. Their usefulness for facilitating self-regulated learning as well as for individual and collaborative problem-solving has been demonstrated by various studies (see Tergan & Keller, 2005). Increasing attention is being aimed at investigating how concept mapping can be used to support individual and cooperative knowledge management by providing contextualized access to information and knowledge resources (Tergan, 2005).

Related goals have been pursued in knowledge management by developing methods for the construction, visualisation and application of knowledge structure maps (Eppler, 2001; Probst et al., 1997) and ontologies (Gruber, 1993). By visualising meta-information of available knowledge resources (e.g. documents, people) knowledge maps can facilitate the identification of relevant new knowledge, its integration with existing knowledge and the contextualisation with respect to the needs of a given task (Probst et al, 1997). Although the usefulness of visualisation for supporting information seeking has been met with scepticism due to the mixed results of existing studies (Hearst, 1999), approaches focusing on specific application contexts where explorative access and discovery of information structures and relationships are more important than goal-directed search have provided promising results (Becks, 2001; Bruillard, 2000).

This thesis considers the use of knowledge visualisation for supporting the specific application context of cross-community knowledge exchange. It proposes that this application context is characterized by the specific kind of requirements that can be effectively supported by visual knowledge overviews. The exchange of knowledge between different communities is mediated by users' access to unstructured community information spaces representing unfamiliar knowledge domains. Cross-community information access is motivated by ill-defined problems which result in extremely ill-defined information needs (Section 1.1). Thus, visualising knowledge structures of different communities, providing ways for discovering relationships between and using them to contextualize information to a given tasks, promises to provide valuable support in such ill-defined situations.

Accordingly, this work proposes a method for eliciting and visualising implicit knowledge structures of individuals and communities of users in form of dynamic, personalised knowledge maps that make them usable for semantic exploration and multi-perspective navigation of community information spaces. The introduced knowledge map model combines Document Maps that visualise main topics, clusters and relationships between knowledge reflected in document collections comprising community information spaces with Concept Maps visualising personal and shared conceptual structures of community members. This is accompanied by a conceptual framework for applying such knowledge maps to support cross-community knowledge exchange and its realization in form of an interactive system and visual information interface.

The technical realization of the proposed approach integrates several existing methods into a novel solution. Automatic generation of document cluster maps (Lin et al., 1991; Honkela et al., 1997; Becks, 2001) and extraction of word categories from texts (Honkela, 1997) is integrated with collaborative filtering (Resnick, 1994) and personalised classification based on user-induced templates (Aha et al., 1991). This is accompanied by established visualisation and interaction techniques known for their ease of use and support of visual navigation (cluster maps, hierarchical trees, focus+context techniques). Finally, these methods are integrated into a visual information interface based on a multi-view paradigm (Baldonado et al., 2000) enabling simultaneous visualisation of different knowledge structures and their use for multi-perspective navigation of community information spaces.

A special aspect of the developed solution is a technique for eliciting personal knowledge structures based on the analysis of user actions occurring during information access. This follows the sensemaking paradigm (Russel et al., 1993) that views the process of information seeking as a rich source of knowledge-intensive activities reflecting users' personal knowledge (e.g. finding appropriate concepts for an information need, creating representation schemas for organising information, using structures of others for information classification). Information seeking is thus an inherently explorative process in which the construction of knowledge is facilitated by the creation and use of external representations which organize information into meaningful structures (ibid). Accordingly, supporting access to unfamiliar community information spaces through dynamic knowledge maps that visualise personal and shared knowledge structures of members from different communities, promises to provide valuable support for cross-community discovery and sharing of knowledge. This is the main hypothesis proposed by this thesis.

Several important characteristics support this hypothesis and differentiate the proposed solution from other approaches:

- *Visualisation of implicit knowledge structures of communities incorporating personal views.* The visualisation of conceptual structures of communities is based on personal structures of individual members reflecting their actual use of concepts in organizing information. This respects the informal social nature of communities whose knowledge structures are rarely explicitly expressed, consist of many different, but interrelated, personal points of view and evolve continuously in time. Relating the community concepts to concrete information resources enables semantic exploration of unfamiliar community spaces and learning about concepts describing the community knowledge.
- *Unobtrusive elicitation and application of personal knowledge for information classification.* Unlike other methods (e.g. concept mapping, collaborative filtering) the developed solution elicits users' personal points of view in a way that is unobtrusively embedded into their primary tasks. By making the creation of personal maps an intrinsic part of common activities of information access no additional effort for expressing their preferences is required from the users. Such maps function as semantic templates that can dynamically classify unknown information or structure entire information spaces based on user defined concepts. This makes them usable both for the map author and users from other communities. Such immediate benefits also alleviate the common cold-start and free-rider problems of classical collaborative filtering and community rating approaches (Herlocker et al., 2000).
- *Interactive visual exploration and manipulation.* The developed technique realizes light-weight knowledge maps as interactive visual artefacts that can be shared and interactively manipulated by the users in order to generate and explore different views of unfamiliar information spaces. This supports the inherently explorative nature of learning about unfamiliar knowledge. Sharing personal maps that classify information from the viewpoint of the map author, allows access to unfamiliar information spaces from different personal perspectives.
- *Contextualized recommendations and information filtering.* Personal knowledge maps not only reflect personal knowledge of the user but also the information needs of specific tasks. Automatically matching personal maps to the need of specific information tasks allows relevant knowledge from other users (concept structures and related documents) to be made available in the most appropriate context. In this way ill-defined information tasks are supported by recommendations of relevant information structures and personal knowledge contextualized to a specific need.
- *Multi-perspective navigation and access to community spaces.* Rather than trying to connect the structures of different communities in a single shared structure, simultaneous co-existence of individual local perspectives is supported by a multi-perspective visualisation method. This allows users from each community to access unfamiliar information spaces while retaining their own perspectives and discovering how they are related to unfamiliar community knowledge.

Overall, the outlined solution approach for collaborative elicitation and visualisation of personal and shared knowledge maps, their application and implementation in a multi-perspective visual information interface promises to provide valuable support for cross-community discovery and sharing of knowledge.

1.3. Research Goals, Methods and Contributions

1.3.1. Aims and Objectives

This thesis studies the problem of supporting the exchange of knowledge between heterogeneous communities through collaborative knowledge visualisation. The principal aim is to develop a method and an interactive system that enables the visualisation of implicit knowledge structures of individuals and communities, provides ways for discovering relationships between and makes them usable for supporting information seeking in unfamiliar community spaces.

More specifically, this comprises the following objectives:

- Theoretical grounding of the problem of cross-community knowledge exchange in the fields of knowledge management, computer-supported collaborative work and human-computer interaction. Identification of requirements for developing knowledge visualisation methods and interactive systems to support this problem.
- Development of a method for the construction of dynamic knowledge maps which visualise implicit knowledge structures of individuals and communities and make them usable for personalised structuring and contextualization of information,
- Development of a conceptual framework for the application of the developed knowledge visualisation method to supporting cross-community knowledge exchange and its prototypical implementation in a concrete system,
- Development of an interactive knowledge map tool enabling the use of the developed knowledge visualisation method for semantic exploration and multi-perspective navigation of community information spaces,
- Exploratory evaluation studies assessing the effectiveness of the developed method and interactive system for supporting cross-community knowledge exchange.

1.3.2. Research Method

As a starting point, a literature analysis of theoretical frameworks and empirical studies from knowledge management and CSCW regarding knowledge creation and sharing within and across community boundaries yields an overview of issues which are important for supporting the processes of cross-community knowledge exchange.

Contrasting this with existing approaches and technologies for supporting knowledge sharing within communities provides the motivation for a new approach. An analysis of HCI work on sensemaking and main classes of knowledge management needs and processes yields functional requirements for the design of the system and typical tasks for which such support is important in practice.

Based on the identified requirements a specific method for collaborative knowledge visualisation and a conceptual framework for its application to cross-community knowledge exchange are developed. To this end scientific analysis and engineering methods are used in order to analyse, combine and adapt existing methods and extend them with new functionalities. For example, the developed method for the construction of personalised knowledge maps combines well-known methods from knowledge discovery and information visualisation and integrates them into an intuitive interface and interactive tool for collaborative elicitation and cross-community exchange of knowledge.

Software engineering methods are used for designing appropriate system architecture and prototypically implementing the developed method in a concrete interactive system and visual information interface. The development of the interactive knowledge map tool and information interface is takes advantage of user-interface design and usability engineering methods such as iterative design, rapid prototyping and user involvement through focus groups. Finally, the last part of the work employs usability engineering techniques to evaluate the usability of the developed methods and prototypes.

Since the research subject has been little investigated in existing work and especially from the perspective of knowledge visualisation and HCI, the presented work is explorative research that integrates contributions from several different fields. Accordingly, the evaluation of the proposed approach is based on qualitative, exploratory studies in laboratory conditions aiming at verifying the viability of the developed solution, showing its strengths and limitations and identifying critical issues for further research.

1.3.3. Research Contributions

The contributions of this thesis fall into three complementary fields of research:

From the perspective of *knowledge management* the thesis

- develops a theoretically grounded approach to supporting cross-community knowledge exchange through the use of collaborative knowledge visualisation

From the perspective of *knowledge visualisation* it

- develops a novel method for collaborative elicitation and visualisation of dynamic knowledge maps incorporating implicit knowledge structures of individuals and communities of users
- shows the application potential of the developed method for a clearly defined application domain in knowledge management and information access,
- shows its strengths and weaknesses and thus may help to improve existing and develop new approaches to knowledge visualisation and their application.

From the perspective of *human-computer interaction* it

- develops a conceptual framework for applying the developed knowledge visualisation method to support facilitate and stimulate cross-community knowledge exchange by supporting information seeking in unfamiliar domains,
- develops an interactive system and visual information interface that implement the developed framework and make it usable for semantic exploration and multi-perspective navigation of community information spaces.

1.4. Thesis Outline

The thesis consists of three main parts. The first part presents a detailed problem analysis and requirements definition based on literature study of relevant theoretical frameworks, empirical case studies and relevant support tasks in practice. As a result the specific requirements for supporting cross-community knowledge exchange are formulated. This is followed by a survey of related work on support for community knowledge sharing and methods for knowledge discovery, knowledge visualisation and visual information interfaces, which yields the motivation for a new approach.

The second part discusses the methodical and technical contributions of the thesis. It presents the developed method for constructing knowledge maps that visualise implicit knowledge structures of individuals and communities of users, the framework for its application and its prototypical implementation in a concrete system. This part concludes with the presentation of the developed interactive tool supporting cross-community knowledge exchange through multi-perspective access to community information spaces. The third part presents the results of the empirical evaluation of the proposed approach by means of a comparative laboratory study investigating the task-adequacy and usability of the developed solutions.

Part I: Problem analysis and requirements definition

Chapter 2 Communities and Knowledge

Defines and describes specific types of communities relevant for this thesis. Analyses processes of knowledge creation and exchange within communities based on well-known theoretical frameworks. Introduces the need for cross-community knowledge exchange and describes the main contexts in which it occurs. Provides an overview of existing approaches and solutions for supporting knowledge sharing in communities. Analyzes their capabilities and limitations. Motivates the need for a new approach. The concluding analysis identifies principal barriers and shortcomings for cross-community knowledge exchange.

Chapter 3 Requirements for Cross-Community Knowledge Exchange

Identifies the requirements for developing a method and an interactive system supporting cross-community discovery and sharing of knowledge. Based on the analysis of an existing theoretical framework of cross-community knowledge exchange a set of high-level requirements and an idealized model of supporting cross-community knowledge exchange are identified. The high-level requirements are contextualized within the specific application context of knowledge management by applying the knowledge management process blocks model. The analysis of the sensemaking framework describing processes of knowledge construction during information access in unfamiliar domains relates the requirements to specific informational tasks that need to be supported in practice.

Chapter 4 Knowledge Visualisation and Visual Information Seeking

Provides an overview of relevant methods for knowledge visualisation, knowledge organisation and visual information seeking related to the thesis problem. This includes: 1) methods for discovery and visualisation of semantic structures of unstructured document collections, 2) methods for elicitation and visualisation of knowledge from human users, 3) methods for ontological structuring and organization of knowledge and 4) approaches to using visualisation for supporting information seeking. Points to possibilities and limitations of using existing techniques for addressing the research problem. Summarizes the potentials and shortcomings of existing solutions with respect to the requirements of the cross-community application contexts.

Chapter 5 Outline and Goals of the Own Approach

Synthesizes the results of the analysis in the previous chapter as a motivation for the new approach. Presents specific goals of the own approach and the basic idea for its realization with respect to the defined requirements for cross-community knowledge exchange.

Part II: Solution design and development

Chapter 6 Eliciting and Visualising Implicit Structures of Personal and Community Knowledge

Presents the developed method for the construction of dynamic knowledge maps that visualise implicit knowledge structures of individuals and communities of users and can be applied as visual information tools for contextualized access to unfamiliar domains of knowledge. Describes the basic concept and realization of the method based on the integration of techniques for visual document clustering and extraction of concept networks from texts with collaborative filtering and personalised classification based on user-induced templates. Presents the light-weight ontological structure of the developed knowledge map model. Summarizes how the presented method satisfies the specific requirements for cross-community knowledge exchange.

Chapter 7 A Framework and System Architecture for Cross-Community Knowledge Exchange

Presents a conceptual framework showing how the developed knowledge visualisation method for collaborative elicitation and visualisation of implicit knowledge structures can be applied to support cross-community knowledge exchange. Shows how the developed framework supports the specific requirements for cross-community knowledge exchange. Describes a system architecture enabling the framework implementation and its prototypical realization in a concrete system.

Chapter 8 Knowledge Explorer – An Interactive Tool for Multi-Perspective Access to Community Spaces

Describes the conceptual design and prototypical realization of the Knowledge Explorer, an interactive tool implementing the proposed knowledge visualisation method and framework for cross-community knowledge exchange introduced in previous chapter. Presents the iterative design cycle with a proof-of-concept prototype and the final implementation based on the results of a formative evaluation. Special attention is given to the developed model for multi-perspective knowledge visualisation and its realization in form of a visual information interface allowing semantic exploration and multi-perspective navigation of community information spaces.

Part III: Evaluation

Chapter 9 Empirical Evaluation in a Comparative Laboratory Study

Develops an evaluation framework for empirically validating systems aimed at supporting cross-community knowledge exchange. Presents the results of a comparative laboratory evaluation of the proposed approach and its prototypical implementation in a concrete system. Empirically verifies the task-adequacy of the developed method and its application to supporting cross-community knowledge exchange. Compares the adequacy of the developed solutions with respect to a standard information seeking reference system without knowledge visualisation. Assesses the overall usability and user satisfaction with the Knowledge Explorer.

Chapter 10 Conclusions and Outlook

Summarizes the main findings and results of this work. Discusses possibilities for improvement and outlines directions for further research.

2. Communities and Knowledge

In order to define the problem context more specifically, this chapter reviews the different definitions and types of communities and their main characteristics (Sections 2.1, 2.2). Particular consideration is given to specific community types relevant for this thesis, such as communities of interest, communities of practice and community-like networks of practice (Sections 2.3, 2.4). Different environments in which communities exist and the technologies used to support them are considered in Sections 2.5 and 2.6. The processes through which knowledge is created and exchanged within such communities are analysed and related to established theoretical models of collective creation and exchange of knowledge (Section 2.7.). Based on the presented analysis, the main context in which cross-community knowledge exchange occurs and the barriers impeding it are described (Section 2.8).

2.1. The Concept of Community

The notion of community as a sociological concept goes back to Ferdinand Tönnies[3] (Tönnies, 1887) . Tönnies used it to describe groups of people connected by a shared sense of togetherness which motivates behaviour based not only on self-interest but also strongly related to the common interests of the group as a whole. According to Tönnies such groupings develop based on shared social bonds (e.g. family, kinship), a shared geographical territory and shared beliefs.

The notion of community has later been used in sociology in many different ways to describe groups of people connected by a shared geographical space and social bonds (Hillery, 1955) [4]. Members of a community interact regularly with each other to share experiences and strengthen social ties. They voluntarily engage into acting collectively in order to achieve common goals. They share a set of values and symbols (e.g. language) which provides a common context for making sense of one's actions, desires and experiences (van Vliet & Burgers, 1987). A community provides a "place" for people to develop a sense of identity and their relation to the world in which they live. It provides conventions and norms for social behaviour and interaction with others (Schubert, 2000).

While the majority of early approaches emphasized the shared geographical space as the founding element of communities (Hillery, 1955), with the widespread use of computer networks (BBS, Email, Internet, WWW) the role of a shared physical space has lost importance. Accordingly, the notion of community has been extended to include groups of people whose shared sense of togetherness and social ties are created through social interaction and exchange of information via computer-supported systems and tools such as Email, Chat, mailing lists and discussion boards. The shared context of such communities is based on different kinds of common needs, goals or interests, regardless of geographical location and often without any previous familiarity between the participants (Wellman, 1998a; Wellman & Giulia, 1999). They are referred to as virtual communities (e.g. Rheingold, 1993), online communities (e.g. Preece, 2000) or communities of interest (e.g. Carotenuto et al., 1999).

On the other hand, research on social learning (Lave & Wenger, 1991) and knowledge sharing in organizations has investigated the role of community-like formations in professional and work-related contexts (Brown & Duguid, 1991; Wenger, 1998, Orr, 1996). Social bonds and a shared sense of togetherness develop as people working on similar problems engage into helping each other, sharing experiences and discussing their work practice. Thus communities evolve from the common need to perform everyday work and from the sense of shared identity that grows out of shared practices (e.g. common activities and ways of working, the shared domain of knowledge and ways of talking about it). Such communities that are based on social learning from actual work experiences have been discussed under the notion of communities of practice (Lave & Wenger, 1991; Brown&Duguid, 1991; Wenger,

[3] The original term is "Gemeinschaft", introduced by Tönnies in "Gemeinschaft and Gesellschaft" (translated as "Community and Civil Society"). In sociological work the term Gemeinschaft has often been used also in articles written in English.
[4] The analysis presented in (Hillery, 1955) differentiates between 94 different definitions.

1998). This approach has been followed by many researchers investigating how to support processes of collective creation and sharing of knowledge. Accordingly, different concepts and community definitions have been introduced to characterize communities in which the exchange of knowledge is their primary characteristics. These include notions such as knowledge communities (Erickson & Kellogg, 2001), communities of inquiry (Wellman, 1998a) and networks of practice (Brown & Duguid, 2000).

As we can see, it has been historically difficult to provide one all-encompassing definition of the concept of community. In order to more precisely identify the types of communities relevant for this thesis, a detailed overview of the different definitions and types of communities is presented in the next sections.

2.2. Community Definitions

The various approaches to defining and categorizing communities differ between each other mainly with respect to which characteristics they emphasise as the defining elements of communities and the related research issues on which they concentrate. A review of the existing literature in knowledge management, computer-supported collaborative work and HCI (e.g. Koch, 2003; Ishida, 1998; Wellman, 1998a; Davenport & Prusak, 1998; Preece, 2000) points to the existence of six broad clusters of community definitions:

- *Communities as shared places:* A shared geographical area (e.g. neighbourhood) has been historically seen as a defining characteristic of communities (Hillery, 1955). A shared place provides a context and the opportunities for people to meet, interact, get to know each other and hence develop social relationships. This creates a common context of needs and problems to be solved and motivates people to become involved in collective actions. As a result social ties between previously unfamiliar individuals are created which enables a shared sense of togetherness to develop. This definition of communities has been taken by virtual communities based on a shared virtual place (conferencing systems, 3D worlds) in which people can meet and interact, regardless of their geographical location. The virtual place is seen as a substitute for a shared physical place needed for a community to develop (e.g. Rheingold, 1993).
- *Communities as social networks:* A well-studied non-territorial definition of community is based on the concept of social network. The concept of social network usually refers to a set of people connected by a set of informal social relations such as friendship, co-working or information exchange (Garton, Haythronthwaite, & Wellman, 1999). By using electronic communication people create their social networks by looking for other people or other social settings based on their needs and desires rather than on their physical location (Wellman, 1998a,b). The shared feeling of togetherness and the community support (e.g. companionship, information, help, sense of belonging) is then a function of the quality of established relations with other people. Accordingly, this approach focuses its attention on investigating different kind of "community ties" existing in such networks (Wellman, Carrington and Hall, 1988).
- *Communities as identity:* The notion of identity refers to the "self experience" of an individual that is formed through his taking on of social roles that are recognized by specific social groups as belonging to their social context (Köhntopp, 2000). Identity provides orientation for one's behaviour and actions, needs and interests (Koch, 2003). Members of a community interact with each other in order to develop and affirm both their personal identity within the community as well as a sense of shared identity. Personal identity and reputation is an important source of reliability and trustworthiness of information (Donath et al., 1999). The awareness of a shared community identity by its members is a defining element of a community's existence (Ishida, 1998; Schlichter et al., 1998). Such community identity provides the basis for the willingness of its members to share information, knowledge and experience. The shared sense of togetherness, the reciprocity of social ties and the mechanisms of social reputation provide the necessary trust and motivation to contribute to community life: e.g. to answer questions, describe common problems and experiences, share news and opinions.

- *Communities as shared practices:* Another class of definitions emphasises the role of "shared activities" from which the sense of togetherness and the relationships between community members develop: "Members of a community are informally bound by what they do together, from engaging in lunchtime discussions to solving difficult problems and by what they have learned through their mutual engagement in their activities" (Wenger, 1998). An extended definition based on this concept has been proposed in (Koch, 2003): *"Communities [are] informal groups…of people that are based on shared interests and problems…in which…people communicate, cooperate, exchange knowledge and experiences, create new knowledge and in doing so learn from each other"*[5].
- *Communities as shared symbolic systems:* The combination of social interaction and a shared system of values and symbols as defining elements of communities has been proposed in (van Vliet & Burgers, 1987). Definitions based on this view introduce the concept of a "shared semantic space" that will play a central role in understanding knowledge sharing within and across community boundaries (Chapter 3.1): *"Communities are associations of members with corresponding goals and moral values…interacting…within a shared semantic space (shared system of values and symbols [i.e. language])."* (Schubert, 2000, pp. 72)
- *Communities as social value:* Another perspective on communities emphasises the notion of social value reflected in the mutual help network, shared knowledge and identity that the community provides for its members (de Michelis, 2001). According to this definition, communities are voluntary networks that provide social channels for information exchange and mutual help. They are based on a shared sense of togetherness, shared place and a shared language, which provide a sense of community identity and borders to others (ibid). This provides the basis for shared activities and interaction through which members fulfil their personal needs, contribute to the community and learn from each other (Koch, 2003, p.21).

As we can see most community definitions have in common the following basic understanding: communities are groups of people connected by a shared sense of togetherness which is based on some kind of a shared context and developed through regular interaction that supports evolvement of social ties (Koch, 2003). Beyond this least common denominator individual similarities and differences between different definitions exist along a number of additional dimensions. The different classes of definitions use rather idiosyncratic ways of defining community characteristics. Though a given class of definitions would often satisfy also some of the characteristics used by another definition they are not used by the authors following a given approach.

2.3. Types of Communities

In order to deduce requirements for technological support for particular community types different approaches for categorizing communities have been proposed (ibid.). Such approaches are often based on different definitions presented in the previous section. Different categorizations then take different subsets of those characteristics as main elements that differentiate specific types of communities between each other.

An attempt at providing a systematic overview of commonly used categorization criteria has been presented in (Lazar & Preece, 1998). This scheme distinguishes between the following criteria for classifying computer-supported communities:

- community attributes – such as shared goals or interests, the intensity of interactions and emotional ties, shared activities between community members, access to shared resources, support between community members, and social conventions, language or protocols (Whittaker, Isaacs, and O'Day, 1997).

[5] Translation from German.

- supporting software – ranges from chat systems and discussion boards to media spaces and virtual multi-user environments (e.g. Wellman, 1998a).
- relationship to physical communities - describes the degree to which a given community is based not only on online exchanges but also on off-line interaction and activities that take place in real, physical space.
- the degree of boundedness - describes the amount of social relationships that remain within a given community (Wellman, 1997).

The most common community categorizations distinguish different types based on the community attributes (e.g. Carotenuto et al., 1999) sometimes in combination with supporting software and the relationship to physical communities (Wellman, 1998a). For example (Carotenuto et al, 1999) classify communities based on the subject of community focus and its tightness into:

- community of interest (shared interests and diffuse focus),
- community of practice (both tight and diffuse focus on a shared activities or practice)
- community of purpose (shared interests and tight focus),
- community of passion (tight focus on shared interests, but small)

	Broad focus	Tight focus
Shared interests	**Community of Interest**	**Community of Purpose/Passion**
Shared practice	**Community of Practice**	**Community of Practice**

Table 2-1. Categorization of communities according to (Carotenuto et al., 1999)

In a similar way, considering the classification by community attributes and the role of online communication (Wellman, 1998a) distinguishes between:

- online virtual communities (online interaction, shared interests)
- integrated online/offline communities (online/offline interaction, shared interests)
- communities of practice (offline/online interaction, work-related)
- communities of inquiry (online interaction, shared interest with strong focus, loosely work-related)
- communities of consumers (online communication, formed around specific products).

Our concern in this thesis is to consider different types of communities with respect to the role that collective creation and exchange of knowledge exchange plays within them. From this perspective we subsume the results of the mentioned community classifications into the following four community types:

- virtual communities,
- communities of interest,
- communities of practice,
- networks of practice.

2.3.1. Virtual Communities

Virtual communities (also referred to as online communities) are characterized by online interaction which creates a sense of „sociability“ between the participants, in spite of geographical distance and the lack of any other previous social ties between them. The notion of “virtual community” has been introduced by Rheingold:

"Virtual communities are social aggregations that emerge from the Net when enough people carry on... public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace." (Rheingold, 1993, 5).

Rheingold based the notion of virtual communities on his experiences with a computer-based discussion system (the WELL), where previously unfamiliar people engaged in discussions around common topics of interest. The interactions between participants were asynchronous and took place through posting and reading contributions to online discussion boards, dedicated to specific topics.

In technical fields of research (e.g. CSCW, KM) the notion of virtual community has been generalized by different authors to refer to all communities in which the major part of interaction and communication between participants occurs through computer-based media (e.g. Carotenuto et al., 1999). As a consequence such approaches treat all communities making use of electronic media as subtypes of virtual communities and differentiate between them with respect to other characteristics such as being based on common interests (communities of interest) vs. being based on shared activities (communities of practice).

In this thesis we adopt a somewhat different perspective in order to differentiate more specifically between the nature of different community types and their main characteristics. The notion of virtual communities is restricted to real-time gatherings of people in “virtual places” such as chat rooms, conferencing systems or virtual 3D worlds. The concept of “virtual place” refers to the ability of participants to be aware of each other’s presence and communicate in real-time with each other by means of text chat, audio or sometimes video.

In case of 3D virtual worlds, the illusion of “sharing the same space” is also supported visually by providing a graphical representation of the virtual place (computer-generated 3D scene) and of the participants’ presence within it (so-called avatars) [6]. Through such virtual personas participants are able to move through the virtual place, interact and communicate with each other. A particular version are textual virtual environments of so-called MUDs and MOOs[7], in which both the virtual place and the participants’ virtual personas (Turkle, 1995) are represented by purely textual descriptions. The latter have been the focus of much research regarding the development of a shared sense of belonging and social identity in such environments.

In fact, in sociological research the notion of virtual communities has been used primarily to describe virtual gatherings of people in which social interaction, personal engagement and the meaning that participants bring to their online interaction make out the sense of shared togetherness and thus the essence of community existence (Szczepanska, 2001). Accordingly, most researchers using the term “virtual communities” focused on investigating issues such as the dynamics of social interactions, social norms and conventions, trust and the development of identity in such communities (e.g. Turkle, 1995).

2.3.2. Communities of Interest

As observed in the previous section, the notion of communities of interest is sometimes subsumed as a subtype of virtual communities (Carotenuto et al., 1999). In our approach we distinguish it as a separate type of community. The basis of our distinction is the focus of communities of interest on the exchange of information and sharing experiences around a common topic of interest or shared needs as opposed to focus on real-time interaction in virtual places.

Communities of interests emerge as groups of people engage into communication, exchange of information and experiences around a common topic of interest (e.g. patient forums, health, hobbies, book

[6] E.g. Active Worlds http://www.activeworlds.com/ , Blaxxun http://www.blaxxun.com/

[7] MUD stands for “Multi-User Dungeons and Dragons”, and MOO for “MUDs Object Oriented” i.e. a specific technological realization of MUDs. Originally MUDs were conceived as online adventure and role-playing games, whereas with the advent of MOOs a number of environments with the educational or purely socializing purposes emerged (e.g. MediaMOO, http://www.media.mit.edu).

fans etc.). Such communities are commonly composed of individuals who share common backgrounds or interests (Carotenuto et al., 1999) and communicate largely through mailing lists, online forums and discussion boards (Ishida, 1998).

Hence, in communities of interest the primary means of community constitution and interaction is asynchronous exchange of information as opposed to direct real-time communication and interaction in virtual communities (chat, avatar worlds etc.). Seen from this perspective, a special kind of communities of interest are also communities of inquiry (Wellman, 1998a). Communities of inquiry are composed of "people interacting to learn more about an area" (ibid.) and are most often found in academic and research contexts (e.g. special interest groups in professional associations etc.):

"Communities of inquiry now take a variety of online forms such as scholars exchanging ideas via distribution lists, researchers posting their results to Web sites, or a dispersed set of researchers working together and sharing their results via email and mutually-accessible databases" (ibid.)

Such communities could also classify as communities of practice (see below) but since they tend to have somewhat broader focus "learning more about an area of interest" (ibid.), not necessarily conditioned by a concrete work context or shared practice we consider it as a subtype of communities of interest.

2.3.3. Communities of Practice

Communities of practice are defined as "informal aggregations of people who share work practices and common experiences" (Wenger, 1998). They are self-organized and emerge spontaneously as people working in similar contexts help each other, share experiences and discuss their work practice (ibid). A crucial element characterizing such communities is the emphasis on learning which occurs through concrete experiences from everyday work in which members engage into shared activities and practice their profession (Brown & Duguid, 1991; Wenger, 1998). These shared activities and the resulting social learning represent the essence of such communities. They form the basis of our distinction of this concept from online gatherings that characterize virtual communities.

According to Wenger's definition, the shared context in communities of practice is based on three main characteristics:

- the shared domain,
- the sense of community,
- the shared practice.

The community domain is represented by the common problems, needs and knowledge shared by community members. The sense of community is based on members' shared sense of identity i.e. of belonging to a same social group. It develops through participation of joint activities of the members, communication and sharing of experiences around common problems and information social interaction (Wenger, 1998). The notion of shared practice refers to shared activities performed collaboratively by the community members, as well as to common tools and artefacts used by community members. This also includes ways of working and doing things that are characteristic of member's activities even when they work individually (ibid).

Such shared practices provide an orientation as to what are appropriate ways of doing/of participating in community life. They represent a shared perspective that provides members with implicit "guidance" for interpreting the relevance and the meaning of information, for appropriate ways of contributing to the community and of using community knowledge. Hence shared practices represent implicit filters which determine what knowledge and experiences are relevant for the community discourse.

In such communities shared mental models or cognitive frameworks emerge that determine the boundaries of relevant knowledge and enable efficient sharing between the members (Boland & Tenkasi, 1995). But such shared cognitive spaces also impose limits to the creation of new knowledge. As we will see later, this is the reason why the concept of communities of practice is one of central concepts considered by the approach proposed in this thesis.

While the original definition of communities of practice includes very different kinds of informal groups of people “who share an interest in a domain of human endeavour and engage in a process of collective learning” (Wenger, 1998), most approaches focus on communities of practice in work-related contexts. Communities of practice in organizations typically involve small groups of specialists (from one or several different professions) who work on same kind of jobs (e.g. sales representatives), collaborate on shared tasks (e.g. software developers) or work together on a product (e.g. engineers, marketers). The members are bound by a common responsibility of “getting work done” which motivates knowledge sharing and learning from mutual experience through collaboration on shared tasks and frequent informal interaction (Sharp, 1997).

Such communities have been originally observed in shared physical contexts (Wenger, 1998; Orr, 1996) where direct face-to-face collaboration and socialising have been the primary means of sharing information and developing new knowledge (e.g. the notorious sharing of “war stories” or “coffee machine conversations”). Through such interactions members of a community develop a common, shared understanding of information, events and experiences. This shared understanding provides orientation and knowledge for dealing with new situations. A much quoted example is the study of Xerox copy-machine repair representatives (Brown & Duguid, 1991; Orr, 1996). It describes how successful technicians learned about problem solutions in informal social occasions (such as lunch room conversations) by exchanging stories from their work experience about ways to repair certain machine problems (“war stories”) which were not documented in official sources of information.

Due to large distributedness of work processes and organizational structures in modern organizations, the concept of communities of practice has been extended to encompass also informal work-related networks which connect geographically distributed participants. In such communities, direct face-to-face interaction is replaced by online information exchange and computer-mediated communication (chat, audio/video-conferencing, awareness). Hence, the development of appropriate technologies and tools to support the sharing of knowledge and experiences within distributed communities of practice has been a major research topic in knowledge management and computer-supported collaborative work (Chapter 4).

2.3.4. Networks of Practice

As we have seen, the concept of communities of practice refers to both physically co-located as well as to distributed communities that are both based on strong social ties which develop from mutual cooperation and learning through problem-solving in everyday work. However, with the ubiquitous use of electronic communication and information exchange, social networks emerged in which social ties between participants are weak and the shared place is virtually non-existing (both virtual or real).

Furthermore, it has become clear that people increasingly belong to several different communities of practice which go beyond their immediate work setting. As a result the possibilities for interaction and communication across boundaries of individual communities open up, mediated by members who participate in activities of different communities.

Accordingly, the concept of networks of practice (Wenger, 1998; Brown & Duguid, 2000) has been introduced to refer to social networks that connect people working on similar practices but who do not necessarily know each other. Participation in such networks is based on some extent of shared practice and domain of knowledge much like in the communities of practice. But the interaction and the social relationships between members are much weaker and the domain is broad in scope with many different aspects, addressed in various ways by different members. The relationships and communication tend to be indirect and depend strongly on information exchange by means of websites, electronic repositories and discussion forums (Swarbrick, 2002). Often, shared cross-community portals provide access to information spaces of different communities involved in the network.

Such loosely connected, community-like networks, play an important role in the processes of knowledge sharing and knowledge creation in organizations. Since they are loosely connected and do not rely on intensive interaction nor on personal familiarity, they can both reach large groups of people and transmit knowledge and information across organizational boundaries.

The strength and advantages of weak-tie networks have been emphasized by different studies of social networks that examined the role of weak ties for accessing privileged information and resources (Granovetter, 1973; Boissevain, 1974). As noted in (Swarbrick, 2002) other examples of increased access to knowledge and resources through weak-tie networks of practice have been documented in studies of large international organisations, dispersed occupational communities (e.g. oceanographers) and in academic research communities.

Of particular importance for this thesis is the role of such networks in involving people who simultaneously belong to several different closely-knit communities of practice. The wide scope, the weak ties and the low interaction of the network result in a low treshold for participation which supports the involvement of participants who do not necessarily belong to the core of the shared practice of the network. Such individuals can then act as knowledge brokers (Wenger, 1998) supporting the flow of information and knowledge between different communities in which they are originally anchored.

2.4. Communities Relevant for this Thesis

It is important to note that the distinction between the described community types is based only partly on community attributes as the classification criteria (e.g. shared goals, intensity of online interactions). Another strong criteria of the distinction are the phenomena that are the primary research focus in itself.

More specifically, we can observe that:

- the term "virtual communities" is most frequently used by researchers focusing on the dynamics of social interaction between groups of people communicating by means of computer-based networks. The research focus is thereby on the influence of online communication on the evolvement of social relationships, social norms or identity building (e.g. Turkle, 1995.)
- the terms communities of interest and online communities are used by researchers focusing on online exchange of information and experiences between loosely connected participants and the conditions under which such exchanges occur (e.g. Lazar & Preece, 1998; Ishida, 1998; Carotenuto et al., 1999)
- the concept of communities and networks of practice is used by researchers focusing on understanding and supporting the creation and sharing of knowledge in communities acting in professional and work-related contexts (e.g. Wenger, 1998; Brown & Duguid, 2000, Davies et al., 2003). Communities in this context are sometimes also referred to as knowledge communities (Erickson & Kellogg, 2001).

In particular, the concepts of communities of interest and of networks of practice emphasises the importance of knowledge flows through information sharing in loosely connected communities that are based on weak ties with low levels of direct interaction. Such community-like networks exhibit larger degrees of heterogeneity and commonly connect people belonging to different communities of practice and different domains of knowledge.

With this in mind, for the purposes of this thesis we adopt a somewhat relaxed community definition from (Koch, 2003). In this approach communities are:

- informal groups or networks of people that are based on shared interests, problems or practices,
- in which people exchange knowledge and experiences through mostly electronically mediated interaction and information exchange
- through which they learn from each other, exchange and create new knowledge.

With respect to the analysis in the previous chapters, this means that we are referring to communities whose main characteristics are:

- the shared context is based on shared interests, needs or activities,
- intensity of real-time communication between members is low,
- the exchange of information and knowledge is occurs primarily through access to shared information spaces and asynchronous communication (e.g. document repositories, knowledge bases, discussion forums).

Table 2-2 shows a characterization of the main community types along these dimensions. Thereby, the distinction between community types is based on the previously described distinction of primary research phenomena addressed by the researchers using a given community type as main reference of the concept of community. As the table shows, communities of interest, communities of practice and networks of practice share important defining characteristics: they are based on shared needs and/or interests with access to shared information resources playing a major role for the exchange of knowledge. The importance of community-wide asynchronous communication is in general high and often dominant with respect to direct real-time communication. In co-located communities of practice direct social interaction still plays an important role, but is difficult to maintain in distributed conditions. This implies that the design of systems and tools supporting collaborative construction and use of shared information spaces is of high relevance for supporting the sharing and creation of knowledge in such communities. Accordingly, when using the term community in the rest of this thesis we will be referring to such kinds of communities and community-like knowledge networks.

	SHARED INTERESTS OR NEEDS	SHARED GOALS	SHARED PRACTICE	INTENSITY OF REAL-TIME COMMUNICATION	INTENSITY OF ASYNC. COMMUNICATION	ACCESS TO SHARED RESOURCES
VIRTUAL COMMUNITIES (VC)	+	-	-	HIGH	LOW	-
COMMUNITIES OF INTEREST (CoI)	+	-	-/+	LOW	HIGH	HIGH
COMMUNITIES OF PRACTICE (CoP)	+	-/+	+	HIGH/LOW	HIGH	HIGH
NETWORKS OF PRACTICE (NoP)	+	-/+	+	LOW	HIGH	HIGH

Table 2-2. Characterization of community types relevant for this thesis

2.5. Example Communities and Community Environments

Communities of interest, communities of practice and networks of practice can be found in different contexts and environments. This section briefly outlines some of the most common contexts and identifies those of special relevance for this thesis. To this end an adapted view of the classification of community environments presented in (Lacher, 2003) is proposed.

2.5.1. Personal Environments

Communities in personal environments are considered as personal social networks of individuals based on different kinds of social relationships related to one's personal life (Wellman, 1998a). This includes common kinds of ties such as friends, work colleagues or acquaintances. In this view members of such "personal communities" do not actively decide to become part of a community but do so "naturally" as an effect of interactions with other people which they encounter in their everyday life (Koch, 2002).

Another kind of personal communities are online communities formed around shared interests in which participants satisfy their personal needs and interests, such as patient forums where people exchange information and experiences on treatments for their illnesses as well as find social comfort and emotional support by others experiencing the same problems[8] (Preece, 1998). In contrast, in online socializing communities people meet in virtual space to informally communicate, entertain and establish social relationships with others, without a specific goal, purpose or predefined topic of interest.

A special kind of personal communities are also the increasingly popular consumer communities in which people come together based on a shared need to find information, exchange experiences or share a passion on specific products. All of the above kinds of communities develop and function in a fundamentally personal context defined by personal interests, needs or specific goals that are satisfied through social interaction and collective action with other people in real-world or computer-mediated environments.

2.5.2. Scientific Research Environments

In scientific research communities have long represented a common form of knowledge production, exchange and preservation. Such communities are formed both around common domains of knowledge and scientific disciplines as well as based on more specialized specific research problems. They typically combine both extensive use of shared information spaces and intensive direct communication and social interaction in physical meetings at workshops and conferences. In contrast to communities in personal environments, besides their informal character scientific communities often explicitly define specific norms and procedures for participation in the community life and for the production and exchange of knowledge (e.g. structure and format of scientific papers, peer-review procedures for validating and assessing quality of contributions). In research settings both communities of interest and communities of practice can be observed in a number of different forms and structures. Examples include more structured communities such as special interest groups of scientific associations (e.g. ACM SIGCHI – Special Interest Group on Computer-Human Interaction[9]) as well as more loosely organized communities around specific issues such as e.g. the community of researchers on HCI design patterns.

2.5.3. Organisational Environments

Modern organizations have also become an important locus where communities of practice and sometimes communities of interest are found. Some authors go as far as to conceive of entire organizations as "communities of communities" (Brown & Duguid, 1991). In addition to the value for individual community members (learning, mutual support, higher motivation) such approaches emphasise the importance of voluntary knowledge sharing in communities for organizational performance.

The value for organizations has been seen in the communities' ability at solving unstructured problems, sharing knowledge across different organizational departments and preserving long-term organizational memory (Lesser & Storck, 2001). As a result, there is often significant top-down influence on communities in organizations. They are not only hosted by an organization but their creation is often stimulated by special incentives. Sometimes they are even explicitly created by managerial decision.

While this is highly problematic with respect to the inherently self-organized nature of communities, experience shows that such communities can function as long as the influence is not perceived to be interfering with the goals and activities of the community (Gongla&Rizzuto, 2001). Table 2-3 shows several different examples of concrete communities in organizations and their value for organizational performance, as identified in an empirical study by (Lesser&Storck, 2001).

[8] E.g. http://www.thirdaid.com/, http://brain.hastypastry.net/forums/
[9] http://www.sigchi.org

Organization	Community	Objectives	Community Activities	Key Value Outcomes
Multinational lending institution	Urban services specialists	Share experience and expertise across similar projects	Held informal lunchtime seminars Conducted formal training sessions Facilitated Web site repository Produced CD of relevant intellectual capital Captured experiences of retiring practitioners	Faster project delivery Greater reuse of intellectual capital developed by projects
Pharmaceutical firm	Research chemists	Share knowledge about a new industry development	Held face-to-face discussions and meetings to share insights Used video-conferencing to connect research labs Maintained Web site	Development of a new business capability based on advanced research techniques
Software development company	Programmers	Respond to needs for customization of a standard product	Maintained internal listservs for individuals to post comments about modifications Maintained Web site to support sharing of software components Provided access to “spearhead” experts around the company	Greater reuse of existing software assets Increased innovation around new software products
Specialty chemical company	Researchers	Share and innovate new solutions to satisfy customer needs	Maintained extensive discussion database where individuals can post and seek answers to customer problems Employed knowledge brokers and editors to cull through discussion databases and identify frequently asked questions and other knowledge needs Held informal “breakfast seminars” to share discoveries and engage other researchers in problem solving	Faster response time to customer problems Greater linkage between customers and research staff in developing new solutions
Telecom company	Project managers	Transfer experience and techniques across industry groups	Held initial face-to-face meeting with community members to outline community objectives and opportunities Developed e-mail-based expert access/question-and-answer system to post and distribute inquiries	Faster response to project bids and request for proposals Greater reuse of existing knowledge assets

Table 2-3. Example communities in commercial organizations and their value for the organization. Selection from (Lesser & Storck, 2001)

As we can see from these examples, such communities tend to be rather homogeneous involving participants from similar fields of work and expertise (e.g. software developers, project managers, research chemists). We can also observe that the modalities of knowledge exchange in such communities include a combination of direct social interaction (geographically co-located communities) and the use of shared information spaces augmented with online communication (distributed conditions). All examples also confirm that an important role of communities is the evolution and maintenance of a domain of knowledge that occurs through “gathering, evaluating, structuring and disseminating knowledge…among community peers” (Gongla & Rizzuto, 2001). The main value of such communities was found to lie in the more efficient reuse of existing knowledge and resources, faster responsiveness of the organization to customer needs and the creation of ideas for improvement and creation of products and services.

2.5.4. Business Environments

In the business environment, companies have used the concept of communities as part of e-business strategies for improving relationships with their customers, employees and business partners. Such communities are referred to as business communities (Bullinger et al., 2002) and are used by companies as a means for optimising business processes, increasing brand recognition, improving customer binding and product quality. Typical examples are customer communities whose goal is to transform potential customers into highly loyal customers by providing different channels of information and communication about the company's products and services, often accompanied with the possibility for commercial online transactions (online-shops). In contrast to classic web information portals, the focus of such communities is on enticing and supporting customers' active contribution of information and feedback about the products and the exchange of experience with each other (opinions, ratings, advice). Well-known examples of such business communities include mySap[10], brigitte.de or smart forum[11].

The benefits of business communities for the customers and business partners lie in better information and transaction support for their purchase or business choices, while the community creator company benefits by gaining access to customer and partner preference profiles, accounts of product use experiences and stronger loyalty relationships. This provides valuable sources for improvement of marketing and sales strategies and ideas for new product development. In this way business communities attempt to create and strengthen mutually benefiting business relationships between different actors involved in a companies' business.

In contrast to previously described communities, business communities are always created in top-down fashion, are very instrumental in scope (improving profits of the community creator and sponsor) but still preserve self-organized characteristics of voluntary community participation and exchanges between the participants.

2.6. Community-Support Technologies

Many different tools and technologies have been developed for supporting communities. Typical functionalities include online communication for supporting direct social interaction (chat, multi-user environments), mailing lists and forums for information exchange, and shared workspaces for cooperative gathering, structuring and access to information. In addition, tools for visualisation of work activities and online presence are commonly used for providing mutual awareness and a sense of shared social context in distributed conditions. Integrated community-support systems support different combinations of these functionalities (Ishida, 1998).

Different overviews and classification of community-support technologies have been proposed by different authors. Koch divides the required functionalities for a community support system into the following classes (Koch, 2003):

- user profile management
- content management
- category and buddy-list management
- messaging management
- awareness & matchmaking
- personalisation and recommendation (for content and messages).

Bullinger et al. distinguish between three broad classes of support: information, communication & cooperation and transaction (Bullinger et al., 2002). They are refined into seven classes of functional

[10] http://www.mysap.com/community
[11] http://www.smart-forum.de

design by distinguishing between communication and cooperation, and including participation, navigation and administration. A synthesis of these two approaches providing a summary of community support technologies and their application classes relevant in the context of this thesis is given in Table 2-4.

INFORMATION	Functionalities such as newsletters, mailing lists, forum discussions and shared document repositories support the distribution and exchange of information between community members
COMMUNICATION AND SOCIAL INTERACTION	Synchronous online communication is enabled by chat, instant messaging and multi-user environments. Sometimes audio/video conferencing is also used. Awareness tools enable visualisation of online presence and shared work activities of community members (e.g. ICQ). An advanced example of an online community chat and awareness system is Babble (Erickson & Kellogg, 2001).
COOPERATION	Cooperation between community members working more closely together is supported by shared information spaces, document management systems and shared workspaces (Borghoff & Schlichter, 2000). An example is the BSCW system (Bentley et al., 1997). This also includes support for cooperative gathering, structuring, rating and annotation of information. An example is the SocialWebCockpit system (Gräther & Prinz, 2001).
RECOMMENDATION	Based on ratings of information in the community space and on personal interest profiles, recommendation and matchmaking systems provide proactive identification of relevant items or people of interest. Examples include systems like the Referral Web (Kautz et al., 1997) or the Knowledge Pump (Glance et al.,1998).
ADMINISTRATION	Administration of content and member profiles includes roles and access rights, site management, document repository configuration etc.

Table 2-4. Main classes of community-support functionalities. Based on (Koch, 2003; Bullinger et al., 2002).

2.7. Knowledge Exchange in Communities

For the scope of this thesis the main relevant aspect of communities are the processes of knowledge exchange that take place within them. A thorough understanding of the main characteristics, enabling factors and mechanisms through which such processes occur is needed as a background against which the problem of knowledge exchange across community boundaries can be defined and explained. An important point of departure is thereby clarifying the notion of knowledge and relating the processes of individual and collective knowledge construction to sound theoretical frameworks which can guide the analysis of specific aspects important for this work.

To this end two main premises are adopted in this thesis. The first is a social constructivist approach that considers all knowledge to be essentially socially constructed through human interpretation of information, events and experience and their interaction with other people. The second is the organizational theory of knowledge creation introduced by (Nonaka & Takeuchi, 1995) that explains collective knowledge construction through transformation processes between explicit and implicit types of knowledge. Such an approach provides a way for understanding how the exchange of different kinds of knowledge is mediated through specific community tools and technologies. The same framework also provides the basis for analysing the problem of cross-community knowledge exchange and the shortcomings of existing community-support solutions for addressing that problem.

2.7.1. Characterization of Knowledge

In the context of knowledge management, a commonly quoted definition of knowledge stems from (Davenport & Prusak 1998). According to this definition, knowledge is a combination of contextual information and domain expertise with past experiences and individual or shared values. This combination provides a framework within which new experiences and information are evaluated and integrated as new knowledge. Such creation and application of knowledge is a cognitive activity intrinsically connected to people as "knowledge carriers".

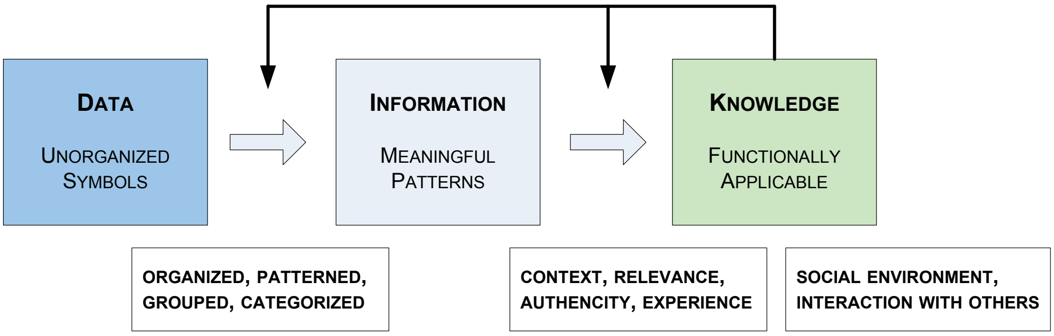


Fig. 2-1. Relationship between data, information and knowledge. An adaption from (Davenport & Prusak, 1998) extended with the recursive interdependency loop and social context.

In order to describe the differences between the process of knowledge construction by people on one hand, and the transmission and use of information as knowledge representations on the other, a commonly used distinction in literature is that between data, information and knowledge (e.g. Probst et al., 1997). The term data is referred to factual, "objective" representations of real or imaginative phenomena, often without a relation to any specific context (Heinrich & Roithmayr, 1998). Putting data into context and adding meaning to it through human interpretation results in information (Davenport & Prusak 1998). Finally, the transition from information to knowledge occurs through comprehension and integration with previous experiences and previous knowledge (Fig. 2-1).

It is often acknowledged that the transition between these different structural elements – data, information, knowledge – is not discreet but continuous i.e. it is difficult to keep the single entities clearly apart. In particular, the constructivist point of view argues that the facts represented by "data" and conveyed as information, must have been created or recognized as such by someone. As such they exist only against a particular frame of reference provided by background knowledge. Hence, the three constructs, data, information and knowledge are mutually interdependent and cannot be clearly separated from each other.

This point of view is further reinforced by the distinction between explicit and tacit knowledge. Nonaka & Takeuchi refer the notion of tacit knowledge (Polanyi, 1958, 1983) to highly personal knowledge, which is derived from experience and embodies beliefs and values. Explicit knowledge is easily detachable from people as knowledge carriers and can be codified and conveyed through information (e.g. documents, language). In contrast, the subjective and highly personal, partly unconscious nature of tacit knowledge makes its communication, externalization and representation in form of structured information very difficult.

In their theory of organisational knowledge creation Nonaka & Takeuchi (1995) describe the processes of the conversion between explicit and tacit knowledge, and their importance for creating collective knowledge. They highlight the critical role of tacit knowledge that functions as "background knowledge" which provides an underlying context and framework against which people interpret the meaning of information and internalize explicit knowledge.

Another important differentiation is that between individual and collective (group) knowledge. Table 2-5 presents a framework for classifying knowledge which incorporates also this aspect alongside the classic explicit/tacit distinction (Cook & Brown, 1999). This framework distinguishes between individual-explicit and group-explicit as well as individual-implicit and group-implicit knowledge. All four types are seen as equally important with each performing a special function that others are not capable of. From our point of view particularly important is the notion of implicit knowledge of groups of people that provides a background against individuals within a given community can construct and share meanings. Furthermore, the relationship between individual tacit knowledge and the group tacit knowledge is important, since it determines ways in which individuals construct knowledge from information.

	INDIVIDUAL	GROUP
EXPLICIT	**Individuals can know, learn and express explicitely** (e.g. concepts, rules and equations)	**Known explicitely but used, expressed or transferred in groups** (e.g. stories about work, successes and failures, myths, metaphors, language)
IMPLICIT	**Tacit knowledge of individuals** (e.g. skills, ways of using concepts, a "feel" for a proper use of tool	**Tacit knowledge of a group** (e.g. distinctive meanings that groups attribute to language, social and physical artefacts)

Table 2-5. Forms of knowledge according to (Cook & Brown, 1999)

Particularly helpful insights for understanding how such processes of knowledge construction occur are provided by social theories of learning (constructivism and social constructionism). They integrate both the individual and social aspects in describing how people construct knowledge through the exchange of information and how this is related to social interaction with other people. For example, (Berger & Luckmann, 1966) describe how people interacting in a certain historical and social context share information from which they construct social knowledge as a reality, which in turn influences their judgment, behaviour and attitude. (Bruner, 1990) shows how the construction of meaning can be related to cultural experiences, in a similar way as (Vygotsky, 1986) has explained how thought and language are connected and framed by a given socio-cultural context of the learner. The studies of (Lave & Wenger, 1991) emphasise the role of immediate social context for learning a body of implicit and expert knowledge through a kind of apprenticeship they call „legitimate peripheral participation". Similarly, (Orr, 1996) demonstrates how knowledge is socially distributed across a network of experts and is shared through processes such as storytelling.

All these studies demonstrate how the construction of knowledge (learning) is an inherently social process in which people actively construct meanings (of information, events and experiences) through a process of information exchange and social interaction with other people. Furthermore, both the personal tacit knowledge of an individual (his previous knowledge, interests, values and beliefs), his current context of intention (e.g. a problem or task at hand) and the social and cultural context in which learning takes place (e.g. community of practice) fundamentally determine the possible meanings that a person can/will construct in this process.

The primary implication of these studies is that in order for groups of people to collectively construct and share knowledge, they have to establish shared cognitive and social context, against which they can negotiate shared meanings of information.

2.7.2. Enabling Conditions: Common Ground, Trust, Social Reputation

Communities represent a special form of a such a shared social and cognitive context. Several specific characteristics of communities facilitate the described processes and allow community members to construct shared meanings from information which enables them to share existing and collectively create new knowledge:

- **Common ground:** The shared domain of interest, shared practices and the common use of language provide the community members with a common ground of understanding and a frame of reference for interpreting the meaning of information. This enables community members to construct knowledge out of shared experiences, mutual communication and exchange of information. The maintenance of the common ground in communities is often facilitated by the collaborative creation and use of a shared information space. As community members contribute and access knowledge artefacts (papers, articles, questions and answers) in the shared space this reinforces the visibility of the shared context that connects them. The shared information space assumes the role of a collective memory. In particular, the use of shared taxonomies which are sometimes created by communities for organizing information in such repositories additionally reinforces the visibility and perception of the shared context (Lesser & Storck, 2001).

- **Trust, social norms and protocols:** Another important enabler of knowledge sharing in communities is the high level of trust between community members. The existence of mutual trust favours the readiness of community members to exchange information and knowledge, that they would not make accessible to strangers. The development of trust is supported by the shared feeling of togetherness and belonging to the community. In particular, trust is developed through social interaction between the members and through participation in community activities that provide the basis for the formation of group identity (Wenger & Snyder, 2000). Social relationships that develop between members as they engage into exchanging information further strengthen the feeling of mutual reciprocity. Furthermore, the existence of implicit social norms and protocols that honour members active involvement and contribution to the community favour mutual reciprocity and readiness to help each other. This provides a fertile ground for sharing otherwise inaccessible knowledge.

- **Reputation economy:** An additional factor very characteristic for knowledge sharing in communities is reputation. The identity of individual members plays an important role in community exchanges. When judging the quality and relevance of contributed information and advice, seniority is a critical factor. Seniority hereby refers to the extensiveness of participation in community. As a member participates in community activities and exchanges, his belonging to the community is recognized and legitimated as such by others (Wenger, 1998). Since membership in communities tends to be open, regulated only by implicit social norms and protocols, it is through this implicit acceptance by other members that one becomes a member himself (ibid.).

  The more active a member is in community life – by contributing to community discussions, answering requests for help of other members or posting information to the shared community repository – the more he is recognized by others and his contributions are higher valued. The member's status within the community is raized and his expertise is acknowledged. In this way a member's seniority in terms of his relevance for the community is established – he has achieved community reputation. Community members often go to great lengths in order to achieve and safeguard such hard-won reputation by maintaining the high level of involvement and contribution in community exchanges. The higher the level and acceptance of a member's contributions, the higher his/her reputation. This results in higher personal satisfaction as the sense of recognition by others and hence of community membership is strengthened. Such a mechanism of identity and reputation provides an important stimulus for voluntarily sharing one's knowledge with other community members (Wenger et al., 2002; Bobrow & Whalen, 2002).

2.7.3. Knowledge Exchange Model

In order to describe more precisely the processes of knowledge exchange in communities we shall apply a well-known model proposed by (Nonaka & Takeuchi, 1995). This comprehensive model describes processes of the creation and exchange of knowledge from the individual level to the level of groups and organizations. In their theory of organizational knowledge creation Nonaka & Takeuchi propose that knowledge is constructed through social interaction and information exchange between individuals through different forms of conversions between tacit and explicit knowledge. Their model describes the following processes of knowledge transformation:

- **Socialization** refers to the process of sharing tacit knowledge through direct interaction between individuals (e.g. informal encounters) and participation in shared activities (e.g. learning by working with more experiences colleagues). The critical aspect is the sharing of experiences through which tacit knowledge such as shared mental models or technical skills is created.

- **Externalization** is a process through which tacit knowledge is articulated into explicit concepts. Typical examples include writing a document, describing a problem through metaphors and analogies, or building models.
- **Internalization** describes the processes through which explicit knowledge is internalized by individuals as tacit knowledge, for example through "learning by doing" or training.
- **Combination** is a process through which new explicit knowledge is created by combining pieces of existing knowledge available in explicit form. Nonaka and Takeuchi refer to this as "systemizing concepts into a knowledge system" such as such as "reconfiguration of information through sorting, adding, combining and categorizing explicit knowledge" (Ibid.). Examples are literature-based knowledge discovery or data-driven new product development (e.g. combining market research statistics with company knowledge base such as patent directory, (Borgoff & Pareschi, 1998)).

According to this model, knowledge is exchange and created through a continuous interplay of all of the four processes, in a loop referred to as the "knowledge spiral" (Nonaka & Takeuchi, 1995; Fig. 2-2).

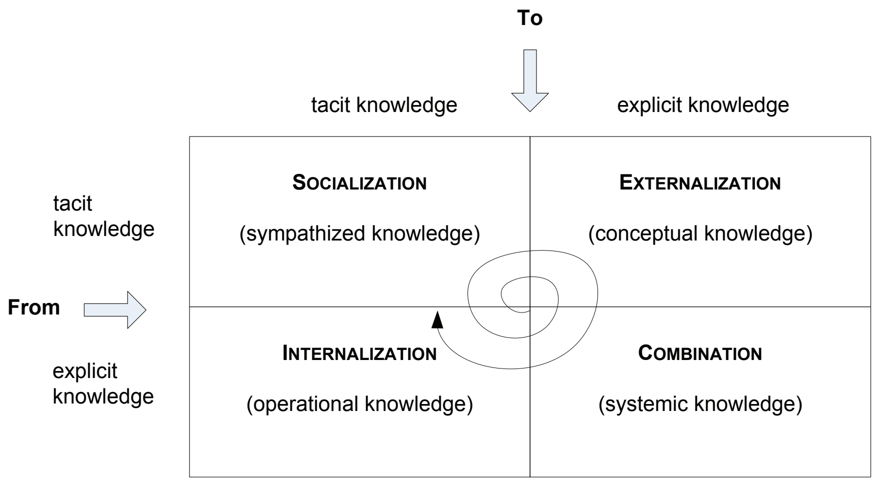


Fig. 2-2. The Knowledge Spiral (Nonaka & Takeuchi, 1995)

This model has been originally developed to describe the processes of collective knowledge creation in teams in the following way. First, in order to achieve some shared common ground implicit knowledge between different individuals needs to be shared through socialization. Based on sharing experiences and mental models underlying their existing knowledge the members establish an implicit common ground against which they can create explicit conceptualisations of the knowledge relevant for the problem at hand (externalization). Having externalized their shared implicit knowledge into a form that allows it to be communicated, the knowledge can flow beyond the team boundaries into the broader context of the organization. Through the combination mode the existing knowledge from other parts of the organization further contributes to the development of explicit knowledge. Finally, from their experiences in going through the process and working on the task at hand, team members are able to internalize the created and exchange knowledge which becomes their personal and highly tacit again.

In this way, the knowledge spiral model also describes the processes of knowledge transfer between individual and collective levels. Individual knowledge is constructed through internalization or combination of explicit knowledge which is acquired through access to shared resources or the exchange of information with others. At the same time, processes such as socialization and externalization enable the sharing and propagation of personal knowledge at the level of social collectives (teams, communities, organization). In socialization this occurs in implicit form (tacit-to-tacit) whereas through externalisation explicit conceptual structures translate tacit knowledge to an explicit form that can be accessed by others. It is internalized into personal tacit knowledge through learning or used for creation of new explicit knowledge through combination.

2.7.4. Mechanisms of Community Knowledge Exchange

As already noted in Chapter 2.3 common mechanisms of knowledge exchange in communities include:

- social interaction through participation in shared activities, informal encounters and online communication (Wenger, 1998; Erickson & Kellogg, 2001)
- the exchange of information through shared information spaces, mailing lists and discussion forums (Koch, 2003).

Applying the above knowledge exchange model to such community mechanisms allows us to understand which kind of knowledge can be exchanged in which of these ways.

KNOWLEDGE EXCHANGE MODE	COMMUNITY MECHANISM
Socialization	Shared activities, informal encounters, online communication
Externalisation	Contributing documents to shared information spaces, posting contributions to online discussions, creating classification schemes
Combination	State of the art reports, FAQs, best practices, design patterns
Internalization	Passive access to mailing lists, forums ("lurking")

Table 2-6. Classification of community knowledge exchange means based on Nonanka & Takeuchi's model.

The basis for participation in shared activities and direct communication is the existence of a common ground provided by a shared domain of interest and shared practice. This motivates the members of a community to take part in intensive interaction and hence share tacit knowledge through the process of socialization. This includes informal and chance encounters in which sharing of experiences through conversations about the experiences from one's own work and shared community practice occurs (cf. war-stories, Orr, 1996). In cases where such shared context is weak or intensive interaction is difficult due to geographical distance or large number of participants, awareness technologies are used for strengthening the shared sense of togetherness by providing mutual notification and visualisation of work activities (Bly et al., 1993; Dourish, 1997; Schmidt, 2002). Hence, through socialization members of a community reinforce their common ground and share otherwise inaccessible personal tacit knowledge.

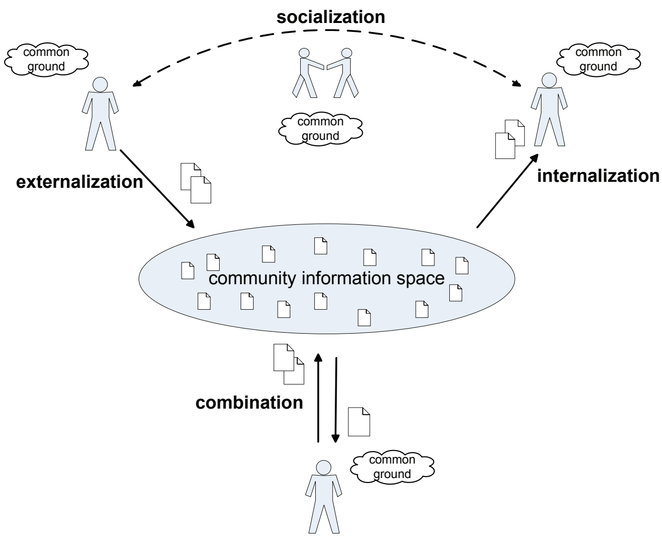


Fig. 2-3. The knowledge spiral model applied to knowledge exchange in communities

Explicit knowledge is exchanged in communities through the collaborative construction and use of shared information spaces. Such repositories commonly contain documents referring to topics of shared interests and discussions about solving common problems. Access to such repositories is a major modality of knowledge exchange in distributed communities, where occasions and motivation for direct communication are rare due to geographical distance and the broad scope of shared domain of interests.

An important means for supporting exchange of explicit knowledge is the construction of shared classification schemes which represent community knowledge. Although the importance of such externalized conceptualizations of knowledge structure is often emphasized, they are rarely explicitly constructed by the community members. In consequence, the exchange of explicit knowledge in communities is commonly reflected in an unstructured repository of interrelated but isolated information artefacts (e.g. document repositories, discussion forum and mailing list archives).

Another very common mode of knowledge exchange in communities is the internalization mode. In this mode community members develop a tacit understanding of the shared community knowledge by passive participation in community discussions: reading and following up the contributions of other members. This mode has been traditionally observed in communities of interest from their very beginning and is based on an inherent asymmetry between the small number of active participants and large numbers of interested but passive readers (cf. "lurkers", Nonnecke & Preece, 2000). Finally, the combination mode is reflected in activities that systematize fragments of knowledge expressed in community documents. This includes collaborative efforts in creating state of the art overviews of the shared domain of interest, compiling FAQs or distilling community knowledge into best practice repositories.

2.7.5. Technological Support for Knowledge Sharing in Communities

The main classes of technological support for knowledge sharing in communities include:

- discussion forums and shared workspaces (e.g. Bentley et al., 1997)
- shared repositories and collaborative information spaces (e.g. Leuf & Cunningham, 2001)
- online-chats and awareness-support (e.g. Erickson & Kellogg, 2001)
- cooperative annotation and indexing (e.g. Gräther & Prinz, 2001)
- collaborative filtering and recommender systems (e.g. Glance et al., 1998),
- community-based ontologies (e.g. Christophides et al., 2000; Davies et al., 2003).

In this section we consider how the different knowledge sharing technologies support the mechanisms of community knowledge exchange described in the previous section.

2.7.5.1. Internalization

Internalisation is the only mode supported by basic community technologies such as mailing lists and discussion forums. They allow exchange of information and experiences between the participants through asynchronous exchange of messages that are archived and organized chronologically and by topical subjects created by the authors. The basic organization principle is thereby the reply to a previous message. The development of a shared context based on such technologies requires members' extensive and active participation in the community discussions. There is no mode for the shared understanding of the community to be expressed, and the repository of the collective memory is an unstructured space of many interrelated but rather isolated pieces of information. Context is very difficult to establish.

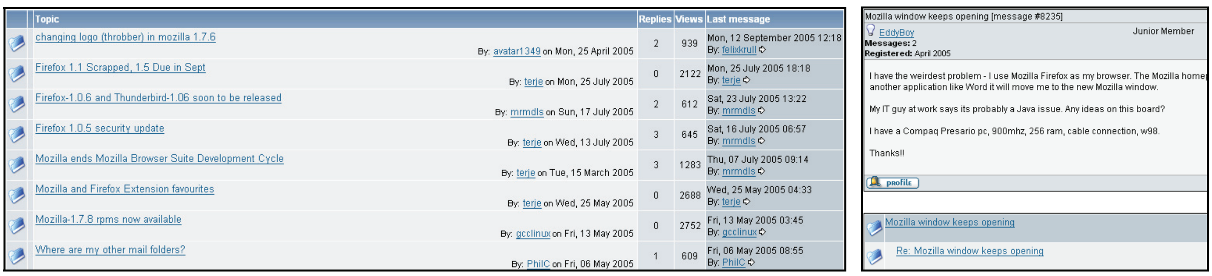


Fig. 2-4. Example of an online discussion forum interface: overview and details of forum contributions.[12]

[12] http://supportforum.sun.com/sjds/index.php?t=thread&frm_id=4&rid=0

2.7.5.2. Socialization

The socialisation mode is addressed by approaches that aim at supporting the sharing of tacit knowledge through a shared virtual space. This includes the awareness and knowledge socialisation approaches, which can be related to two basic premises. The first is that by providing mutual awareness of spatially distributed but contextually related users (e.g. working on same task, belonging to the same community) by means of a shared virtual space, the cognitive distance between them is bridged. The second is that once this cognitive distance is bridged, the conditions are established for the users to enter into conversations through which they exchange otherwise inaccessible personal knowledge. An example, of the knowledge socialisation approach is the Babble system proposed by Erickson et al. (1999). Babble provides an online environment in form of a multi-channel chat which allows both synchronous and asynchronous communication between different groups of users. In contrast to conventional chats, the messages exchanged by the users are persistently stored in an archive organized into hierarchical conversation topics (defined by the users who initiated the conversations). At the same time, user presence is visualised by circles designating active conversations and dots representing active participants involved in a given conversation (Fig. 2-5).

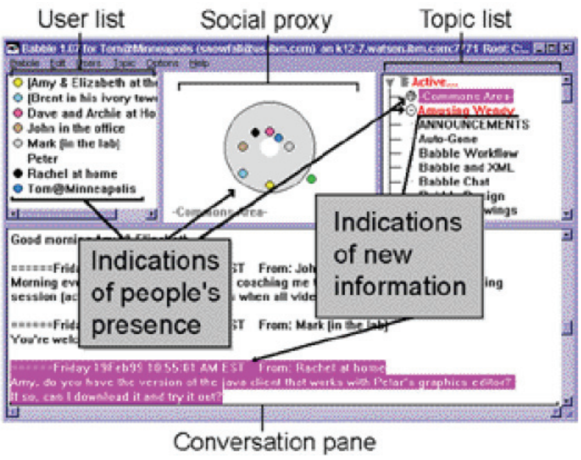


Fig. 2-5. The Babble interface (Erickson & Kellogg, 2001)

Another class of approaches that can be related to the socialisation model has investigated the possibilities of using textual virtual environments of MUDs/MOOs as a kind of online learning labs. Here knowledge is exchanged through shared design practices in building and programming the virtual world (e.g. Bruckmann, 1993). Such approaches are often related to the constructionist theory of learning (Papert, 1990) which emphasises the role of artefacts. This can also be compared to the approach of „learning by doing“ and to situated learning through „legitimate peripheral participation“ as studied and described by (Lave & Wenger, 1991). Other investigations on communities in MUDs focused on patterns of social interaction with respect to issues such as construction of identity and the self-organising establishment of social norms (e.g. (Turkle, 1995). Yet other approaches have explored the use of MUDs as social information spaces, in which social interaction is embedded within a concrete informational context. Related approaches include social navigation such as collaborative web browsing, populated web pages and collaborative histories. There are several variants of this basic model.

2.7.5.3. Externalization

The externalisation mode is addressed by approaches aiming at supporting the explicit formulation of shared conceptualisations in form of structured information repositories, knowledge taxonomies and community ontologies. This often also includes support for cooperative gathering and management of information collections (structuring, rating, annotation) as well as collaborative construction of community dictionaries (e.g. Gräther & Prinz, 2001). The C-Web project (Christophides et al., 2000) proposed methods and tools for explicit formulation of community knowledge in form ontologies and their application to building knowledge-based community portals facilitating e.g. semantic navigation and information classification.

Ontologies represent formal descriptions of concepts and named relationships between them that describe how a given individual or a group of people understand a particular domain of knowledge (Gruber, 1993). They commonly have to be created explicitly by hand and require a process of explicit community negotiation for achieving a consensus about the shared understanding that is to be expressed (see Chapter 4.4.2 for a more detailed discussion). Once created they can be used to access and navigate the community information pool, as well as to visualise the semantic structure of the shared community understanding. An example of existing efforts for building such ontologies in different disciplines but interrelated to each other is the DublinCore initiative (http://www.dublincore.org). The Open Directory Project aims at a collaborative definition of a somewhat simpler taxonomy for manually mapping the content of the whole Web (http://dmoz.org). A hybrid approach has been taken by the OntoShare system that combines manual creation of an initial ontology in workshop session with dynamic evolvement by allowing the users to manually add concepts or relationships between them, as part of their activities in using the ontology to search for information in the community space (Davies et al., 2003).

An increasingly popular technology for communities are Wikis. Wikis are systems for collaborative writing on the web[13], which provide a simple way for a group of users to collaboratively create, edit and structure web pages on the fly. Each member can edit or contribute information to a web page, regardless of whether he is the author of original content. Thus, not only is knowledge shared by exchanging information but pieces of information are weaved together by different contributors in a single "document" i.e. webpage and inter-linked with other pages. The structuring of the pages into subsections and the links between different pages represent a semantic structure of the shared community space represented by the Wiki. In this way, a collaborative structuring of a community knowledge domain occurs. Obviously, this requires great motivation and explicit effort by the users into authoring both content and structure with the purpose of communicating specific knowledge to others. Sometimes editorial procedures (special roles for selected community members) are introduced in order to ensure the consistency of produced content.

2.7.5.4. Combination

The most typical case in practice is the combination of the internalisation mode based on information exchange through mailing lists and bulletin boards, with the socialisation model supported through textual chat. The main problem of such approach is that the sharing of knowledge requires extensive interaction within the community. Recently, approaches have been developed that try to combine all three modes. A well-known solution is the BSCW system (Bentley et al., 1997) which provides a collaborative document workspace based on shared folders metaphor with awareness functionalities and integrated discussion forums. BSCW is used both by teams and communities with a relatively strong focus on solving a specific problem or project.

A more recent example is the Social Web Cockpit system (Gräther & Prinz, 2001). This solution combines a shared workspace for building up a collaborative information repository with socialisation mechanisms such as awareness and textual communication, and with the possibilities to explicitly build up and externalise a shared vocabulary without explicit negotiation. It allows users to add concepts into a shared concept repository and link them to documents currently being inspects. The concepts created in this way can then be used to support search and navigation in the community space.

[13] See http://en.wikipedia.org/wiki/Wiki. The most famous example is Wikipedia, an online encyclopedia written collaboratively by thousands of contributors with sub-communities formed around specific topics and sections: http://www.wikipedia.org

Fig. 2-6. User interface of the Social Web Cockpit (Gräther & Prinz, 2001)

Another line of work has considered the use of collaborative filtering techniques for recommending items of relevance to a community of users based on their previous choices and ratings of information (Resnick, 1994). The basic idea of this method is that items that have been highly rated by other people with similar preferences as a given user are likely to be also relevant for the user in question. Systems such as Knowledge Pump (Glance et al., 1998) use explicit expressions of user interest profiles and their ratings and assignment of information to specific topics to generate recommendations of relevant information. Furthermore, the choice of users from which to receive recommendations can also be restricted to a specific group of trusted experts i.e. “advisers” (ibid.).

Applying this method to members of a given community is a way of supporting both the exchange of explicit knowledge expressed through explicit user preferences (e.g. indication of topics of interest) as well as incorporating implicit knowledge reflected in users choices and ratings of specific information items. Systems such as ReferralWeb (Kautz et al., 1997) or Yenta (Foner, 1999) apply this technique to explicitly match community members with similar interests (also referred to as expertise location). In this way, the exchange of knowledge through externalization (expressing preferences) and internalization (following recommendations) is accompanied by socialization (supporting awareness of relevant experts facilitates the initiation of communication between other community members).

While the aim of ontologies and other forms of knowledge externalisation usually is to create a formalized common understanding, a different approach is to allow different knowledge structures to co-exist and to mediate between them automatically by means of a mapping between different taxonomies or ontology schemes (see (Lacher, 2003) for a survey). These approaches offer the benefits of allowing a decentralized creation and maintenance of knowledge (and thus personal views on a domain) with little explicit coordination. But finding an intentional mapping between conceptualisations is far from being trivial and usually depends on a logical description of concepts. Thus mapping ontologies or document catalogues also depends on the assumption that the meaning of concepts and thought worlds of communities can be codified in a formal representation and therefore suffers from the same basic problem as the other knowledge externalisation approaches.

The Macadam system (Dourish et al. 1999) allows users to create and use personalised views on a central document repository based on an initial centralized categorization scheme. Automatic mediation between different personal views is then possible since all personal views are initially related to a common categorization scheme. In the context of this thesis, a particularly interesting example is the CAIMAN system for mapping between personal knowledge structures and shared community structures (document catalogues) proposed in (Lacher, 2003). In order to minimize the effort of organizing community knowledge into centralized catalogues identification of related categories in different document catalogues is undertaken by means of statistical text analysis of underlying documents. Documents are encoded based on the frequency of word occurrence into a vector space model which allows calculation of inter-document similarities (see Chapter 4.2). Categories which contain similar documents are then considered to be related.

2.7.5.5. Summary

Approaches supporting only internalization (e.g. online forums, blogs) require members’ extensive and active participation in the community discourse. The shared understanding of the community cannot be explicitly represented, relationships between different views and the necessary context for interpreting information is very difficult to establish. The main shortcoming of computer-mediated socialisation approaches is that the sharing of implicit knowledge requires extensive interaction between individual members, and the resulting exchange still resides only in individual users. There is no possibility to visualise the resulting structure of shared understanding. On the other hand, existing approaches to creating externalized representations of a shared conceptual structure require explicit negotiation for achieving consensus between the members. There is no or little support for expressing the personal points of view of individual users and putting them in relation to the shared structure. Combination approaches integrating awareness and information sharing with collaborative community dictionaries as well as recommender systems suffer from an imbalance of effort vs. benefit for the users. The need for explicit expression of user preferences and ratings tends to suffer from cold-start and free-rider problems. Moreover, since collaborative filtering recommends items judged highly by other similar users, the recommended information tends to be narrowly focused within a domain already well known by the user.

Approaches supporting multiple conceptual schemes and automated mapping of relationships between them are a feasible solution for mapping between different personal catalogues within a community as there will be a significant degree of similarity in term usage in community documents. In contrast, in catalogues representing members of different communities, little overlap in concepts and terms used by a community is to be expected. There are likely to exist potentially related categories that contain documents expressed in different terms or containing same terms used in very different ways. In such cases, this approach is applicable to a rather limited extent.

Furthermore, the task of determining “similar” categories is not the same as the task of discovering categories potentially related to query terms or concepts describing a specific problem domain. Relevant categories in this sense may contain documents which consider much different terms than documents related to a category related to the problem in the user’s catalogue structure because the required knowledge is complementary to the user’s knowledge domain. In this case, there will be little or no intersection in terms of term usage between documents from different categories, but there may be “semantic relevance” between them. Establishing such relevance requires supporting the discovery and expression of cross-community relationships by human users.

2.8. Cross-Community Knowledge Exchange

Up to now we have defined the types of communities relevant for this work and the examples of concrete communities and environments in which they are found in practice. We have shown how the knowledge exchange within communities has been considered of great importance in different kinds of settings ranging from research environments to organisational and business environments. We have identified the mechanisms of knowledge exchange within them and analysed the existing approaches and technologies used to support them. We are now ready to analyse the barriers for cross-community exchange determined by the nature of intra-community processes and to introduce the notion of heterogeneous, cross-community knowledge networks as the main context within which cross-community knowledge exchange occurs.

2.8.1. The Need for Cross-Community Knowledge Exchange

As already noted in the introduction, with the increasing complexity of knowledge-intensive work the need for supporting knowledge exchange between different communities both in research environments and commercial organizations has been acknowledged. In particular, it has been empirically confirmed by different case studies of innovation processes such as the development of new products and services or organizational process innovation (e.g. Dougherty, 1992; Swan, 2001). These studies identify the need for integration of knowledge from different fields both for addressing a concrete problem at hand in its local context, as well as for supporting the identification of potentially relevant ideas and experiences in the global organizational context.

A well-known example from knowledge management is Swan’s longitudinal study of knowledge flows in a number of innovation projects involving very different organizations (Swan, 2001). The findings of this study point out two important aspects: 1) the knowledge relevant for the innovation processes in question was typically distributed across very different social groups and communities within and outside of the organization, and 2) the ability to integrate this knowledge from different communities was a key requirement for success of the innovation projects.

A similar situation is also visible in scientific research where increasing importance is given to the need to connect expert knowledge from different research communities in order to address critical real-world problems or fast growing fields of new technologies (e.g. information society, sustainable development, renewable energies, life science). Both traditional professional associations such as ACM and IEEE as well as more recent research networks exemplify this (e.g. the fona - network for sustainability research[14]).

Another typical context is technology transfer, an area of work concerned with supporting the application of results of scientific innovation into industrial development of concrete products and services. An important challenge of technology transfer is identifying potential fields of application and contexts in which a given technology or scientifically developed method could be of relevance (Garnsey & Wright, 1990). Achieving this clearly requires the ability to integrate knowledge across different fields of expertise and identify relationships between different context of research and application. Common activities in technology transfer include technology monitoring and idea scouting undertaken by specialized innovation-support agencies. In contrast, more recent approaches recognize the limits of such centralized “knowledge brokering” agencies and aim at supporting self-organized discovery and exchange of knowledge between different communities both from research and industry, through so-called collaborative innovation networks (Gloor, 2002; Coin, 2004).

2.8.2. From Communities to Cross-Community Networks

The notion of “knowledge networks” has been used to describe such contexts in which the exchange of knowledge across community boundaries occurs. In contrast to teams which are based on specific shared tasks and communities which are formed around shared interests or practices, knowledge networks are social formations which connect individuals from different disciplines and organizational structures who have no continuity of shared interest or common agenda (Merali & Davies, 2001). Their primary purpose is to collect and transmit information across boundaries of different professions, formal (teams) and informal structures (communities) (Alee, 2000). They are loose because participants do not share a specific task, and often belong to different communities which are based on very different interests and worlds of knowledge (Swan, 2001, 2003). Participants frequently change, activities are instrumental, social interactions are rare, ties weak and unstable. Such dynamics results in high heterogeneity but also provides access to otherwise unavailable knowledge resources from different domains and organizational structures. Low regularity of interaction and low commitment implies low intensity of “social maintenance” which allows large-scale participation and resistance to changing levels of activity, oscillations in size and shifts of focus between different knowledge domains over time.

Some authors describe specific subtypes of knowledge networks such as networks of practice (Brown & Duguid, 2001) or nets of experts (Fuhr & Fuchs-Kittowski, 2004) which do share some kind of a more specific common domain or purpose. As discussed in Section 2.3.4 networks of practice are a transitional form from communities to networks. They are wider in scope of the shared domain and larger in size than communities, with weaker ties and much lesser intensity of direct interaction between the members.

The notion of nets of experts introduced in (ibid.) considers rather small nets which are formed voluntarily by experts from different fields in order to solve a common task. Such nets are similar to cross-disciplinary teams as they exist only for the duration of the shared task and dissolve after its completion. They differ from teams in that they are self-organized and emerge from voluntary engagement of the members much like in communities**.** Since they are created when existing team

[14] http://www.fona.de

knowledge does not suffice for solving the problem at hand, their primary scope is the creation of new knowledge. Furthermore**,** the nature of the task and the division of work is not well-defined and as a result, members freely join and leave as the needs of the task require. However, this dynamics is characteristic not only of such specific type of network but rather a rule of the network dynamics at large. Different members continuously join and leave the network in search of knowledge resources and expertise for meeting their current need regardless whether this is motivated by a specific shared task, a problem at hand or by long-term interest in a specific domain of knowledge. At different points in time, different reasons and domains provide shared contexts for different sets of participants within a network.

In other words knowledge networks are not static but dynamic structures, which at any given point in time can be considered as a combination of smaller subnets or communities connected by some shared context (e.g. problem, need, task or practice). Each new participant identifies specific micro contexts within the network at large, that are relevant for his current need (Fig. 2-7). While at one point in time different participants will have identified different contexts, in another occasion they may connect through some shared need. Due to such heterogeneity and continuous dynamics participants leaving the network do not dissolve it, as well as they do not really "leave" the network.

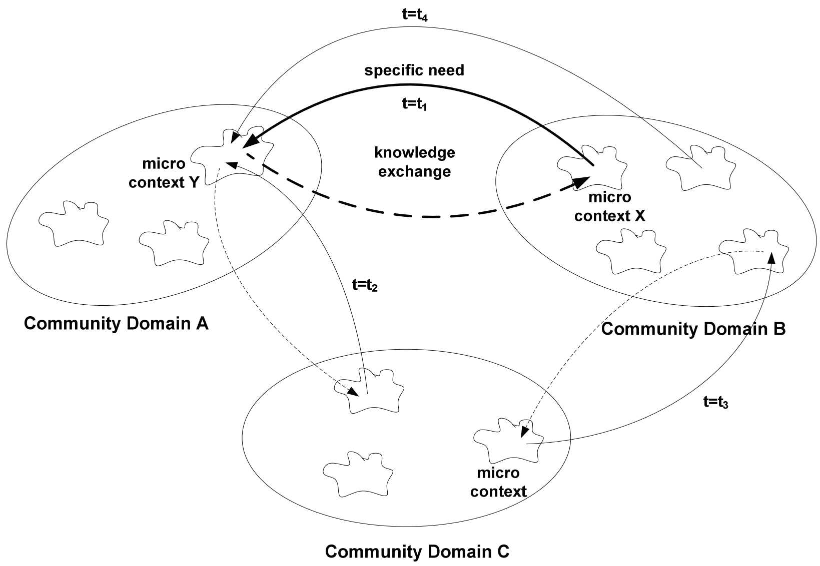


Fig. 2-7. Dynamics of knowledge exchange in cross-community networks

As a result knowledge networks are basically "sets of relationships" (Allee, 2000) that emerge from temporarily shared needs, stabilized through perceived benefits of past exchanges and cooperation between members from different organizational structures and domains of knowledge. They are characterized by a high volatility of needs, low regularity of interaction and weak ties between members. This makes the "shared" domain of the network very heterogeneous and constantly changing.

Due to the broad domain of different needs, interests and expertise of the participants such cross-community networks have been increasingly recognized as important configurations for the creation and preservation of knowledge that crosses the limitations of teams and communities. The important characteristics of such networks with respect to the exchange and creation of knowledge in teams and communities are depicted in Table 2-7 and can be summarized as follows.

As there is no explicit membership in such networks, participants remain anchored in their original communities and use the possibilities of information exchange in the network as a way of accessing complementary, otherwise unavailable knowledge. Since the threshold of access to knowledge from unfamiliar domains is high, the resources of the network are used only when the known contexts are insufficient and the need is critical enough. This typically occurs when existing knowledge within one's own common knowledge contexts does not suffice for satisfying ill-defined information needs and solving complex problems (Fuchs-Kittowski et al., 2003; Allee, 2000; Swan, 2001).

As a result there is higher engagement and openness to learning, partly induced by the awareness of one's own limitations. Similarly, the high ambiguity of context has been recognized to stimulate openness to new ideas and different points of view, which may not come to bearing in more homogeneous contexts. In consequence, the exchange of knowledge through information sharing in such cross-community networks results in high levels of learning and creation of new knowledge (Cohendet et al., 2001).

PROPERTY	TEAM	COMMUNITY	KNOWLEDGE NETWORK
FOCUS	task	interest/practice	task variety/interest
GOAL	realise task	increase knowledge in domain / skills in practice	access complementary knowledge
COMPOSITION	heterogeneous	homogeneous	heterogeneous
CONNECTEDNESS	high	high/low	low
TIME SPAN	short	long	persistent dynamics
MEMBERSHIP	formal	informal	informal
FLUCTUATION	fixed	variable	variable
SIZE	small	large/small	large
OPENNESS	closed	open	open
KNOWLEDGE ACTIVITY	knowledge integration & application	knowledge exchange	knowledge exchange & new knowledge creation
LEARNING MODE	unintended learning by interacting	unintended learning by exchange / by working	learning by exchange (intended/unintended)
PRIMARY SUPPORT	groupware, meeting-support	shared information spaces, online communication, awareness	cross-community information portals, expertise location

Table 2-7. Comparison of main social forms of cooperative creation and exchange of knowledge. Based on integration of findings in (Fuchs-Kittwoski et al., 2003) and (Cohendet et al., 2001).

2.8.3. Barriers to Cross-Community Knowledge Exchange

Based on the above analysis we can identify several important characteristics and barriers to cross-community knowledge exchange. The first barrier to cross-community flows of knowledge is the absence of a focused shared domain and continuity of shared interests that would provide motivation for participants to enter into freqeuent interaction and exchange of information. A related problem is the absence of a shared point of reference for becoming aware of unfamiliar communities and their potential relevance for one's own need. While the latter is commonly addressed by providing a shared point of access to information spaces of different communities, the problem of the missing common ground becomes the biggest barrier to cross-community knowledge exchange.

Missing shared place. Unlike communities, the described knowledge networks within which cross-community knowledge exchanges occur have no visible representation of their existence in form of a shared community space, membership lists or online communication spaces. The information spaces of different communities are isolated from each other and the members of different communities are often not aware of each other's existence. Hence, a prerequisite for cross-community exchanges to take place is the provision of means for the awareness of other communities, their domain and its potential relevance for one's own work. A common approach to achieving this is the establishment of cross-community portals that provide a shared point of access to information spaces of different communities (Fuhr & Fuchs-Kittowski, 2004; Mack et al., 2001; Reynolds et al., 2004). This realizes a necessary precondition for the exchange of knowledge mediated through access to community information spaces to occur in the first place.

Missing common ground. As shown by the analysis in Section 2.7.4, the existence of a shared common ground is a critical prerequisite for creating and exchanging knowledge. But rather than just being given by a shared domain of interest, the common ground of a community develops through the interaction and exchange of information between community members. As a result, it contains more than just a description of a domain of interest or a set of shared practices. Moreover, each community develops its own social and interpretative context (genres, repertoires, perspectives), which in turn determines its interpretations of the world (Boland & Tenkasi, 1995). Different communities inhabit different “thought worlds” (Dougherty, 1992) which determine how their members interpret the meaning of information, artefacts, procedures, events and experiences. Some authors refer to this as “readiness for directed perception” (Fleck, 1979), others use concepts such as cognitive paths (Weick, 1979). Different thought worlds then have not only different terminologies but very different funds of knowledge and systems of meaning. As a result, knowledge cannot simply be “passed” on by exchanging information between members of different communities.

Low intensity of direct interaction and participation. Accessing explicit knowledge from the information space of an unknown community is possible only if there exist ways for gaining insight into shared tacit knowledge of the community in question. According to Nonaka & Takeuchi’s model in order for this to be established and maintained, either frequent interaction between community members (socialization) or engagement in following the community discourse reflected in the shared information space (internalization) are needed. In distributed communities socialization requires extensive informal online communication (chat, virtual spaces, online meeting rooms) while internalization requires frequent access to the community information space (reading documents and contributions in discussion forums, document repositories etc.).

But while such processes normally occur within communities, they are not appropriate for the described cross-community contexts. Members of different communities rarely interact directly with each other and when they do so the interaction is very narrow, with the purpose of obtaining otherwise unavailable information (Allee, 2000). Hence, the socialization process cannot take place. Similarly, accessing information from information spaces of unfamiliar communities occurs only very occasionally, based on a very specific need. Due to the absence of a shared domain of interest, members of one community do not have the motivation nor the resources for intensively following the information exchange in other communities. As a result, tacit knowledge between different communities cannot be shared through common modalities of internalization.

Ill-structured problems and information exploration. As noted in the previous section, the need for cross-community exchange occurs when people face complex or ill-structured problems that do not fit within their known knowledge contexts. Such situations are common in knowledge-intensive work and innovation processes in modern organizations (Fuchs-Kittowski et al., 2003; Swan, 2001). They require making sense out of the problem context: understanding its structure and related knowledge domains in order to identify the knowledge gap and the relevant fields of expertise needed for a solution. This involves seeking information from different, often unfamiliar domains and communities that do not make part of one’s common knowledge context. As a result the information need is very ambiguous and difficult to resolve by means of goal-directed search. Not only does the terminology problem play an important role, but developing an understanding of the problem space in order to translate it into a specific information need occurs during the information seeking process itself. This results in highly explorative nature of corresponding information seeking tasks and processes, in which the availability of existing knowledge structures (e.g. categorization, classification schemes) plays an important role in the process (cf. sensemaking, Chapter 3.4.).

Lack of explicit representations of community knowledge structures. This problem is further aggravated by the absence of explicit representations of shared knowledge structures of different communities. Externalizations of tacit knowledge have been recognized as an important means for supporting the building of common ground and collaboration between heterogeneous actors in computer-supported collaborative work (Arias & Fischer, 2000). While such externalizations are commonly created in teams (e.g. collaborative design, classification schemes for shared project repositories) communities rarely engage into explicit creation of externalizations representing their shared understanding. And even when they do so, the results tend reflect only a limited and biased part of the shared community knowledge – either due to the involvement of a small group or due to the need for achieving community wide consensus upon a representation scheme (Bonifacio et al., 2002, 2002b). In consequence, shared information spaces of communities provide little points of reference that would support non-members in understanding the shared community context within which the accessed information makes sense. The community knowledge structures are highly tacit and invisible from the outside, acquired only through intensive participation in the community discourse. This makes the access to information from unfamiliar communities that mediates cross-community knowledge exchange very difficult.

In synthesis, methods aiming at supporting the exchange of knowledge between different communities should address two fundamental requirements:

1. They shouldn't depend on the processes of socialization and internalization in the form in which they occur within communities. Instead of intensive social interaction or participation in the community discourse, different means for gaining insight into shared implicit knowledge of communities are needed. Since the main means of knowledge exchange in cross-community networks is information exchange, the apropriate method should support the processes of internalization and externalization.
2. They should support explorative access to unfamiliar information spaces and the use of community knowledge structures in that process. This must provide ways for users to relate the unfamiliar information space to their own knowledge, in order interpret the meaning of unknown information.

The analysis of these two main fields of requirements and their further specification in a theoretically-grounded and practically feasible form for guiding the development of a concrete solution are the subject of the next chapter.

3. Requirements for Cross-Community Knowledge Exchange

In this chapter we identify the requirements for developing a method and an interactive system supporting cross-community discovery and sharing of knowledge. The adopted requirements definition model is depicted in Fig. 3-1.

First, a theoretical framework provided by the "perspective making – perspective taking" model (Boland & Tenkasi, 1995) and related contributions are analysed. This results in high-level requirements and an idealized model of supporting cross-community knowledge exchange.

In order to transfer the idealized model into guidelines for developing a practical solution, the high-level requirements are contextualized within the specific application context of knowledge management.

The knowledge management process blocks model (Probst et al., 1997) describes knowledge management processes in organizations. This provides the basis for identifying specific knowledge management application requirements. The requirements are further narrowed down by applying the sensemaking framework (Russel et al., 1993; Qu & Furnas, 2005) which describes the process of knowledge construction during information access in unfamiliar domains.

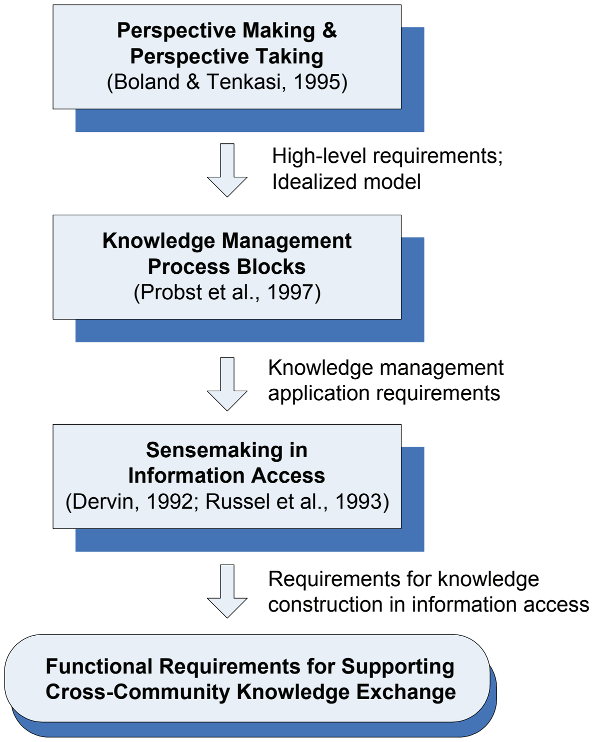


Fig. 3-1. Requirements definition model for cross-community knowledge exchange support

In this way theoretically-grounded high-level requirements of the idealized model are related to a concrete application context and narrowed down to specific informational activities relevant for cross-community knowledge exchange. This ensures that the heuristically defined functional requirements are based on sound theoretical frameworks as well as related to specific, practical application contexts.

3.1. Perspective Making and Perspective Taking

A well-known theoretical framework addressing the problem of cross-community knowledge exchange is the model of "perspective making and perspective taking" proposed by (Boland & Tenkasi, 1995). This model describes the processes of the creation and exchange of knowledge within and between different "communities of knowing" (ibid.) within organizations. The notion of communities of knowing refers to groups of actors with similar expertise or shared perspective on a domain of knowledge. This relates it to our definition of community introduced in Chapter 2.4.

As we have seen in the previous chapter, different communities are characterized by different thought worlds and different shared semantic contexts (common ground) which make it difficult to exchange knowledge between them. In order to describe processes through which such barriers to cross-community knowledge exchange can be overcome, Boland and Tenkasi focus on the relationships between knowledge and the language used by a community to develop, express and communicate this knowledge. This focus makes their model especially suitable for our scope of identifying requirements for supporting cross-community knowledge exchange mediated through the exchange of information. Their main proposition is that in order for knowledge to be exchanged between different communities, the shared semantic contexts constructed through intra-community activities (perspective making) need to be made visible and accessible to each other (perspective taking).

3.1.1. Perspective Making

In this approach, a community perspective is defined by a shared domain of knowledge and a specific vocabulary of expressing this knowledge as well as by methods, values and reasoning patterns accepted by the community members. Perspective making is then the process through which such perspectives are formed. A crucial aspect of perspective making is the development and refinement of the shared community vocabulary: by explicating the meaning of existing concepts, creating new concepts and establishing relationships between them, members of a community develop shared knowledge. Whereby in the initial stages of a community concepts tend to be broad in scope, as the community knowledge evolves their meaning is defined more and more precisely with more and more sophisticated distinctions and descriptions. Such "complexification" (ibid.) increases the ability of community members to use, share and create new knowledge within the community perspective.

In emphasising the importance of the local scope of meanings of community concepts, Boland and Tenkasi relate the process of perspective making to the language game model of Wittgenstein (Wittgenstein, 1974). According to this model, words, sentences, symbols and forms of speech do not have a fixed, objective meaning, but are fundamentally socially constructed. That is, their meaning is defined always with respect to a specific social context (e.g. community in our case) and occurs through actions, conversations and social activities between people taking part in this context. In this sense, the knowledge of a community and its use of language are intrinsically connected. The shared vocabulary, its concepts, their meanings and the relationships between them are not merely "descriptions of community knowledge" but *are* the knowledge itself. Rather than creating descriptions of an "objective" reality, through language games communities develop their own worlds of meaning and knowledge as a form of their own reality.

As a result of such processes, different communities are characterized by very different perspectives. Knowledge from different communities is not only expressed in different terminologies, but the meaning of concepts and their use is also strongly context dependent. Different communities may use the same concepts but they will use them to see and reason about things in different ways (ibid.). Members from one community may be interested in the same phenomena as another community but will see them in light of completely different problems or opportunities. As a result, similar wordings can be used to refer to very different concepts in different contexts. Such specialized language games and highly implicit context-dependency of shared community vocabularies make it difficult for knowledge to be shared by simply passing information or ideas from one community to another.

An important implication of this is that knowledge of a given community - the concepts and meanings through which it is expressed - cannot be easily understood or evaluated outside of their specific contexts of use i.e. is without gaining an understanding of the underlying community perspective.

3.1.2. Perspective Taking

In order to overcome this barrier and exchange knowledge, members of a given community need to develop an understanding of the context in which the knowledge from an unfamiliar community makes sense. Developing such an understanding is requires the ability to take on the perspectives of others and imagine their point of view – to see a given problem or domain of knowledge from through the eyes of another. Accordingly, this process is referred to as perspective taking (ibid.). Perspective taking then involves developing an understanding of the shared domain of knowledge of a given community and the specific vocabulary used by its members in expressing and sharing this knowledge. Such an understanding is needed in order to interpret the meaning of information, in the way it is understood within a given community.

But as we have seen in Chapter 2.7 shared community perspectives are highly implicit and not readily visible from outside. They are constructed and shared between members through processes of socialization (interaction between members) and internalization (consuming information created in the community). This includes activities such as shared problem solving, document writing and sharing, cooperative gathering and structuring of information, discussions and narrative exchanges in online forums.

Since these processes do not naturally occur between members of different communities, making perspective taking possible requires that the perspectives of individual communities be externalized in form of artefacts that can be made available to non-community members. Such artefacts need to enable members of different communities to “recognize and accept the different ways of knowing” (ibid., p. 358). This involves gaining insight into the topics comprising the shared community knowledge and the specific vocabulary used by the community in expressing and reasoning about them: the concepts, their meanings and ways of use and the relationships between them. One way of supporting this is by providing externalized representations of a community’s domain of knowledge in a way that reflects the use of language in the community in question.

But perspective taking is a difficult process. The assignment of meaning to artefacts, events and information is a strongly automatic process in which people normally do not notice the interpretation that is involved (Boland & Tenkasi. 1995, p. 362). The existing knowledge of a community member functions as an automatic template against which new information is compared. Information more closely related to existing knowledge will be more readily accepted than information coming from unfamiliar domains. Concepts from unfamiliar domains will be assigned familiar meanings from one’s own perspective first.

As a result, an important aspect for supporting perspective taking is stimulating “reflectivity”: the ability and willingness to become conscious of how one’s own interpretations of the meaning of given information are determined by one’s own underlying perspective: e.g. the importance of certain topics, the meanings of specific concepts and relationships between them. One way of supporting this is providing externalized representations of one’s own knowledge in a way which reflects one’s own perspective.

Hence, supporting perspective taking requires us to provide externalizations of implicit knowledge structures of communities both as a way for supporting insights into shared vocabularies of unfamiliar communities as well as for stimulating the awareness of one’s own interpretative assumptions, when accessing information from unfamiliar domains.

3.1.3. Multiple Views and Knowledge Integration

The problem of cross-community knowledge exchange has also been considered under the notion of knowledge integration. Empirical studies of the processes of innovation (e.g. development of new products or services) demonstrate how an important requirement for successful innovation is the ability to integrate knowledge from different communities and corresponding thought worlds (e.g. Swan, 2001; Dougherty, 1992). A critical problem in such processes is the establishment of a “shared context of knowing” as a way of “locating one form of knowledge in the context of another” (Swan, 2001).

A frequently cited example illustrating this is a field study of new product development processes by Dougherty (1992). In unsuccessful cases of new product development she investigated, the failure was due to the actors’ inability to acknowledge and reconcile the different interpretations of market-technology relationships (stemming from different community perspectives) and their consequences for desirable features of new products. A specific problem were the different understandings of the concept of “market orientation” that meant different things for different groups. In the thought world of research and development, it referred to product specifications and technical features: product functionalities define the market. For manufacturing, market orientation meant durability and reliability, which could be achieved more easily with a lower number of features. In marketing, what mattered were customer needs defined on an individual customer basis, while for the planning group identifying the appropriate market niche was the core problem of market orientation (see also the discussion in Boland & Tenkasi, 1995). Thus, although the basis for communication has seemingly been provided by a standardized, shared concept (market orientation), the lack of awareness about differences in implicit understanding characteristic for members of different communities disabled effective exchange of knowledge.

This example illustrates not only the point that different thought worlds can refer to same concepts with different meanings, but that the insistence on a single, unified view attempting to bridge the differences between perspectives is not an appropriate solution. A similar conclusion is supported by (ibid.) with respect to the difficulties of perspective taking. Making the thought worlds of different communities visible and accessible to each other must not aim at the development of the same "world view" but rather provide means for becoming aware of and understanding the views of others (ibid.)

3.1.4. Boundary Objects

A special form of artefacts that can stimulate and create the conditions for perspective taking to occur are boundary objects. Boundary objects are artefacts that connect different perspectives of heterogeneous actors on a given problem or a domain of knowledge, withouth requiring the establishment of one shared perspective (Star, 1989; Star & Griesemer, 1989; Star, 1993). Several important properties of boundary objects allow them to be used by different communities without the need for shared understanding between them to be established:

- *Common point of reference:* The existence of a shared object provides a common point of reference between otherwise distinct domains of shared knowledge, interests or a specific problem.
- *Multiple perspectives:* Due to the simultaneous availability of different perspectives members of different communities can interpret the information contained in the shared object in different ways, appropriate for the specific needs of each community. This allows members of different communities to cooperate, exchange and collaboratively develop of new knowledge without having to give up their own specific perspective.
- *Means of translation:* The visibility of different interpretations increases the awareness of the existence of different perspectives. Their relation to a common point of reference (shared object on a given topic, problem or domain of knowledge) provides a way for discovering relationships between them. In this way members of one community can develop an understanding about the perspectives of the others (e.g. the important topics for that community and the meaning of concepts used to describe it).
- *Adaptability:* Boundary objects are "working arrangements", created and adapted as needed (Bowker and Star 1999). They evolve with the needs for which they were created and cannot be imposed neither by a single community nor by referring to some external authority or standard (ibid.).

A number of different forms of boundary objects have been discussed in literature, ranging from project reports, to shared models to entire physical environments. The creation and use of such boundary objects has been demonstrated as an important means for supporting cooperation and communication in different contexts, such as collaborative design and decision making (Arias & Fischer, 2000; Fischer, 2001) or the coordination and exchange of knowledge between members of heterogeneous communities working on a specific shared problem (Star, 1989; Bowker & Star, 1999).

Specific types of boundary objects that could be used to support perspective taking are classification schemes and cognitive maps (Boland & Tenkasi, 1995). Different kinds of classification schemes (e.g. taxonomies, ontologies) are commonly used to represent and visualise knowledge structures by organizing relevant topics, concepts and documents and specifying different kinds of semantic relationships between them. Cognitive maps are graphical representations of different aspects of a given problem and the causal relationships between them seen from a specific individual or group perspective (see Chapter 4.3). Accordingly, supporting cross-community knowledge exchange requires the construction of knowledge representation models that provide common points of reference for knowledge from different communities, in a way which allows each community to retain local perspectives and yet these perspectives to become interconnected.

An additional kind of boundary objects are artefacts that make visible implicit knowledge structures of individual community members. Constructing artefacts that provide a visible representation of an individual's knowledge allows the representations of knowing from one community to be compared to and exchanged with members of another. Since such artefacts allow members of a different community to gain insight into a part of an unknown community perspective, they also support the basis for perspective taking and cross-community exchange of knowledge. This makes such artefacts also a kind of boundary objects (Boland & Tenkasi, 1995, p.362).

The role of such boundary objects becomes especially important in the light of the fact that perspective taking can never be reduced to a one-to-one mapping of meanings. Not only is it impossible for members of different communities to simply adopt each other's meanings, but even within the same community, different members will have different personal points of view. Even within strong communities full consensus on the meaning of words, information and artefacts is never given. Hence, constructing artefacts that represent personal points of view of individual community members is an important prerequisite for constructing boundary objects that incorporate different community perspectives.

3.1.5. High-Level Requirements and Idealized Model of Cross-Community Knowledge Exchange

According to the above analysis, the exchange of knowledge between different communities can be described through an interplay of perspective making and perspective taking supported through boundary objects that make visible both personal and shared community perspectives. The role of boundary objects is to represent personal and shared community knowledge structures in a way that allows members from different communities to develop an understanding of "what and how the others know" (Swan, 2001), discover how this is related to their own knowledge and translate it into their own terms (Boland & Tenkasi, 1995). Table 3-1 gives a summary of the main elements of this model.

COMMUNITY	• Group of actors with a shared domain of knowledge, interests or practice, connected with a shared sense of togetherness.
PERSPECTIVE	• Local reality meaningful to a particular community. • Defined by a shared domain of knowledge and vocabulary expressing it, reasoning patterns, values and beliefs accepted by community members.
PERSPECTIVE MAKING	• Development of a community's local reality (knowledge domain, shared practices) through externalization and internalization. • Refinement of the shared vocabulary: the meaning of concepts and relationships between them and creation of new concepts. • Narratives and information exchange.
PERSPECTIVE TAKING	• Internalization of local reality of another community. Gaining insight into its common ground reflected in its use of language. • Understanding the context in which knowledge from an unfamiliar community makes sense. Relating unfamiliar knowledge to one's own personal knowledge and the community domain.
BOUNDARY OBJECT	• Externalized construction of individual and community knowledge. Incorporating different perspectives and implicit knowledge. Meanings of concepts are ambiguous and depend on the interpreting actor. • Exchange of knowledge between different contexts without enforcing single, unified meaning of information.

Table 3-1. Summary of the perspective making – perspective taking model. Adaptation from (Totland, 1997).

In this way, perspective taking (understanding knowledge from unfamiliar communities) and perspective making (developing and expressing knowledge within one's own community context) are linked to each other: the exchange of existing knowledge from different communities results in the creation of new knowledge within a specific community perspective. The corresponding cross-community knowledge exchange model is depicted in Fig. 3-2. Supporting cross-community knowledge exchange then requires representing and visualising knowledge structures of individuals and communities in ways that allow members from different communities to:

- Gain insight into implicit knowledge structures of other communities different from their own (*perspective taking*),
- Discover how the knowledge reflected in information artefacts from other communities relates to their own knowledge and contextualize unknown information within their own thought world (*perspective taking -> perspective making*)
- Express the new knowledge in terms appropriate for their own personal and community perspectives (*perspective making*).

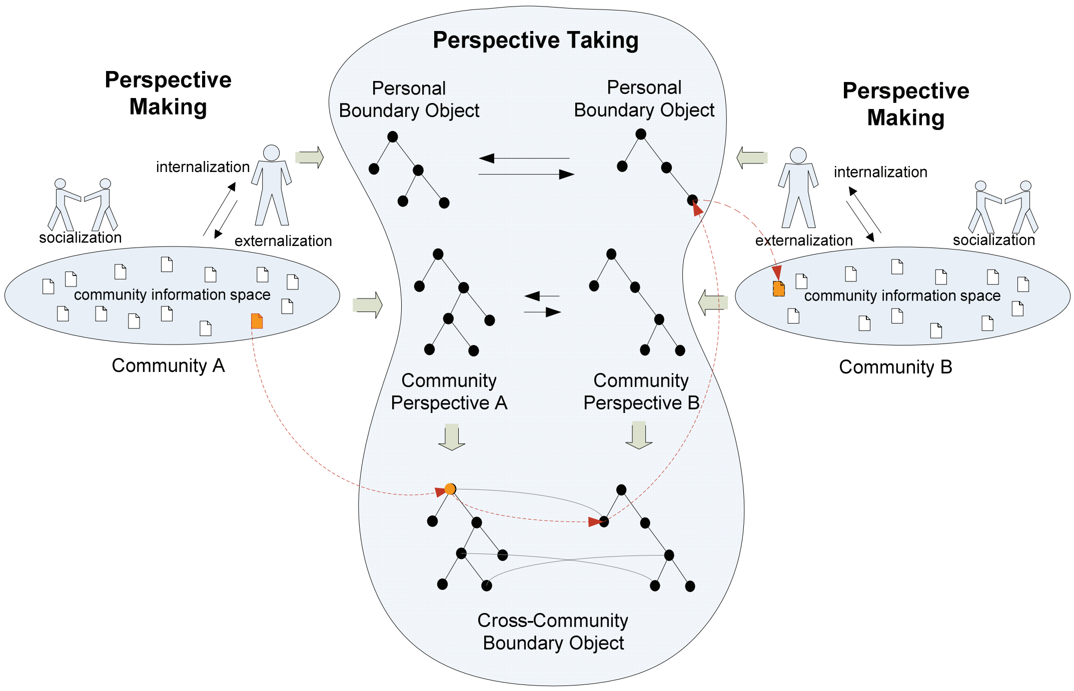


Fig. 3-2 Cross-community knowledge exchange model based on perspective making and perspective taking through boundary objects. Based on theoretical analysis from (Boland & Tenkasi, 1995).

Supporting these processes requires satisfying three main requirements:

- *Reflectivity* - people need to be willing to engage into reflective thinking and temporarily suspend one's own interpretive perspectives in order to accept the other,
- *Externalisation* - perspectives of individuals and different communities need to be externalized and represented in form of artefacts that make them visible and accessible to others.
- *Multiple perspectives* –different interpretations of the meaning of information need to co-exist, rather than requiring consensus on unified shared meanings.

According to (Boland & Tenkasi, 1995) artefacts satisfying such requirements include different forms of classification schemes such as taxonomies, ontologies and knowledge maps. They can represent knowledge structures by displaying the main concepts and relationships between them (semantic, causal etc.) describing a given perspective on a domain of knowledge.

A special way of realising such artefacts is by making them satisfy the properties of boundary objects. Boundary objects represent perspectives of different communities in a way which allows them to be used by members of different communities without the need for establishing a unified, shared understanding. In our case boundary objects must not only represent shared community structures but should also represent personal points of view of individual members. Creating such artefacts that express personal points of view is both a way of stimulating self-reflection upon one's own knowledge, as well as a way of providing insights into specific parts of community knowledge to others (ibid.).

The corresponding requirements for technologically supporting the creation and use of such artefacts for cross-community knowledge exchange can be summarized as follows:

- Support the creation of artefacts that provide externalized representations of a community's domain of knowledge and the shared vocabulary used by community members (concepts, their meanings and relationships). Such artefacts must reflect tacit knowledge, be realized in a form which allows them to be accessed by non-community members and reflect the shared community language as it evolves in community life.
- Provide means for expressing and representing personal points of view on community knowledge. This requires the elicitation and construction of artefacts that reflect personal (implicit) knowledge of individual members and can be exchanged with others.
- Provide means for locating knowledge from one community in the context of the other. This requires the development of methods and tools that support the discovery of relationships between different community perspectives.
- Allow the co-existence of multiple points of view and interpretation of information. This requires us to enable the evolvement and construction of artefacts that connect different community perspectives. Such artefacts must allow the co-existence of multiple perspectives on community knowledge and be usable by members of different communities, without requiring the establishment of shared meaning or one unifying point of view (boundary objects).
- Provide means for such artefacts to be created and used in a self-organized and adaptable manner. In order to function as boundary objects they must provide shared points of reference for members of different communities, must not be imposed by outside authority and must be adaptable to a given need.
- Integrate the support for perspective taking between different communities with perspective making within them. This requires us to embed the creation and use of boundary objects into processes of intra-community knowledge exchange.

The described requirements for supporting cross-community knowledge exchange point to several important challenges from the technical point of view. In order to develop of a method to realize them the following difficult problems need to addressed:

- How can we construct artefacts that elicit and visualise the existing, but not explicitly formulated knowledge of a community?
- How can we do so in a way which incorporates personal points of view of individual members?
- How can we make such artefacts usable for discovering relationships between perspectives of different communities and their domains of knowledge?
- And how can we make them part of common knowledge exchange tasks and processes occurring in users everyday practice?

As a first step in addressing these questions, we shall relate the high-level requirements to knowledge management tasks as a specific application context. To this end we consider a well-known knowledge management process model (Probst et al., 1997). This model describes the main building blocks of knowledge management based on a number of empirical studies of real-world organizations. By relating this model to our cross-community knowledge exchange context we can identify application requirements which need to be considered in order to develop a practical solution. This shall allow us to relate the described high-level requirements to more specific classes of functionalities that need to be supported by a cross-community knowledge exchange method and system.

3.2. Knowledge Management Process Blocks

The notion of knowledge management refers to processes through which organizations manage the creation, capturing, distribution, exchange and use of knowledge (Merali & Davies, 2001). Originally, the concept of knowledge management emerged in the context of business administration and organisation science as a response to the growing importance of knowledge in increasingly globalized, highly turbulent environments (Nonaka & Takeuchi, 1995; Davenport & Prusak, 1998; Wiig, 1997). In this view, knowledge is treated as an organizational asset in its own right that needs to be managed, just as other organizational assets (financial, human etc.).

Many different approaches and related definitions of knowledge management have been proposed, ranging from organizational learning (Argyris & Schön, 1997; 1996; Senge, 1991; Brown & Duguid, 1991) to the theory of organizational knowledge creation (Nonaka & Takeuchi, 1995) to the social capital approach emphasising the importance of informal social networks (Prusak, 1997; Cohen & Prusak, 2001; DeMichelis, 2001). In addition to these approaches focusing on organizational measures and social aspects of knowledge management, numerous work has considered the development of tools, systems and technologies supporting specific knowledge management aspects (for an overview see Borghoff & Pareschi, 1998; Marwick, 2001).

The most suitable approach for our purpose is the knowledge management process model proposed by Probst et al. (1997). According to this frequently-cited approach knowledge management comprises the processes that guide the formation of the organisational knowledge base and direct the processes of its change and evolvement (Table 3-2).

ORGANISATIONAL KNOWLEDGE BASE	Individual and collective knowledge resources available to the organization.
ORGANISATIONAL LEARNING	Change processes of the organisational knowledge base.
KNOWLEDGE MANAGEMENT	Formation and direction of the change processes of the organisational knowledge base.

Table 3-2. Knowledge management definition by Probst et al. (Probst et al., 1997)

Based on a number of empirical case-studies, Probst et al. further proposed a model for describing main classes of knowledge management efforts in organizations. According to this knowledge management process model, the main knowledge management needs and general approaches to satisfying them can be described by six core processes that ensure the distribution, creation, application and preservation of knowledge (Fig. 3-3, lower frame). This is accompanied by two additional processes of strategic intervention for ensuring that the direction of the core processes is aligned with specific organizational goals and that their success can be measured and evaluated (Fig. 3-3, top frame).

Our interest in this model is to use it as a structure against which we can heuristically define requirements for ensuring the practical application of the presented cross-community knowledge exchange model (Fig. 3-2) to support the typical needs and processes of knowledge management in organizations.

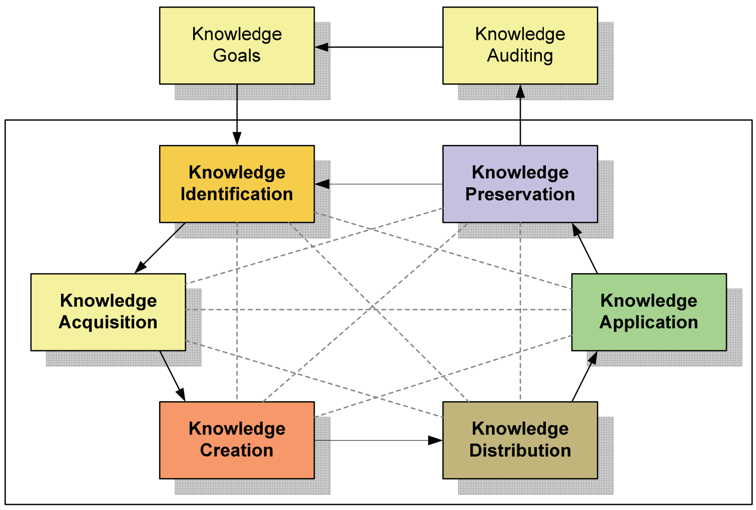


Fig. 3-3 Knowledge Management Process Blocks Model (Probst et al., 1997)

As depicted in Fig. 3-3, the core knowledge management processes include:

- Knowledge Identification and Awareness – achieving the transparency of available knowledge resources,
- Knowledge Acquisition - acquiring external knowledge (recruiting, consultancies, strategic alliances) in order to fill known gaps in the organizational knowledge base,
- Knowledge Creation and Development – the intentional production of new capabilities in an organisation,
- Knowledge Distribution – making existing knowledge available throughout the organisation,
- Knowledge Application – ensuring the productive use and application of available knowledge in contexts where it is required,
- Knowledge Preservation – ensuring the preservation of relevant knowledge and experiences created by the members of the organization.

Ensuring the alignment of the core processes with organizational goals is the task of two additional processes of strategic intervention:

- Knowledge Goals – identifying the strategic, operative and normative goals that are to be achieved through the knowledge management efforts,
- Knowledge Auditing – measuring the success of knowledge management efforts with respect to the specified knowledge goals.

The focus of this knowledge process model are formal organizational entities (e.g. teams, departments) which can be guided by managerial actions such as explicit definition of goals to be pursued and knowledge areas in which competencies need to be developed. In contrast, we consider the exchange of knowledge between communities which are informal and self-organized structures that refuse outside influence and control. Taking communities to be special kinds of collective entities within organizations requires us to consider which of the above processes are applicable to cross-community knowledge exchange.

As we have seen in Chapter 2, communities are not based on well-defined goals but on a shared domain of interests and practices that are defined in a self-organized manner. The development of community knowledge is not guided by external authority but takes place on a voluntary basis through informal social interaction and exchange of information. The evaluation of the created knowledge occurs through community-specific processes of social filtering and peer-review or implicitly through the amount of collective acceptance and use. Accepted community practices and forms of knowledge also evolve through collectively defined social norms and procedures, rather than being imposed by an outside authority. Membership is voluntary and based on adherence to the shared domain and participation in community practice.

As a result, the knowledge goals, knowledge auditing and knowledge acquisition processes as they are defined in the above model are not applicable to community contexts. In order to determine the relationship of the remaining processes to our specific problem we can summarize the most important aspects of the relevant process blocks as follows:

- *Knowledge identification and awareness* can be supported through activities such as the systematic analysis and description of the existing organisational knowledge. This includes the creation of classification structures (taxonomies, ontologies) that describe the content, location and means of access to available knowledge resources. Common examples are yellow pages, expert directories or graphical knowledge maps which depict the knowledge structure (knowledge structure maps) or the distribution of knowledge across human carriers (expertise location maps).
- *Knowledge Distribution* processes are especially dependent on appropriate technological support. The ultimate goal of knowledge distribution is to make existing knowledge available where and when it is needed throughout the organisation. Successful strategies require the connection of organisational measures (trust-building measures, incentives for knowledge sharing, self-organisation) and technological support. The latter ranges from basic support for cooperative work (information gathering, communication, awareness) to intelligent tools for contextualized information access and shared terminologies (for a detailed overview see (Marvick, 2001)).
- *Knowledge creation and development* require on one hand measures for establishing an open organizational culture with tolerance to failure, trust-building measures and self-organisation. Such conditions are commonly found in communities. An important aspect is the creation of new knowledge in everyday work. This requires tools for supporting the capturing and sharing of knowledge and experiences from everyday work (e.g. lessons learned) as well as the discovery of different viewpoints on a problem and relationships between them. Supporting creativity requires support for personal and collective reflection and tools that allow seeing familiar knowledge in new contexts. Finally, supporting shared problem-solving and development of new ideas requires methods for externalising personal, implicit knowledge in ways that makes it visible and usable for others.
- *Knowledge Application* processes refer to the actual use of knowledge that has been made available through knowledge distribution processes. For effective knowledge application the barriers to accepting the knowledge of others (e.g. the "not invented here" syndrome) need to be removed by appropriate organisational measures. The access to knowledge needs to be contextualized within specific work processes, in order to ensure that relevant knowledge can be accessed just when needed and that the users can easily connect it to their existing knowledge. These requirements are summed up as "ease of use", "just in time" and "ready to connect" (ibid., p. 277).
- *Knowledge preservation* processes need to support three main activities: the selection of relevant knowledge and experiences, its effective representation and storage in some externalized form and the continuous updating of the organisational knowledge base. While the preservation of strongly structured knowledge is currently well supported (databases, document management systems, expert systems), the management of highly unstructured and context-dependent knowledge is still a challenge. A typical task is the preservation of knowledge and lessons learned in the process of accomplishing different tasks. To this end, collective knowledge needs to be structured through cooperative activities such as the development of shared terminologies (ibid.., p.309). Establishing a shared terminology for the whole organization thereby carries an inherent

trade-off: while different groups within the organization use specialized languages appropriate for their domain of work, ensuring effective reuse of available knowledge throughout the organization requires a limited set of standardized concepts. But standardization leads to de-contextualisation: the generalized concepts are most often not appropriate for the specific work contexts of any particular group. Designing appropriate solutions for this trade-off is still an open research issue.

3.3. Application of the Process Blocks to Cross-Community Knowledge Exchange

The process of cross-community knowledge exchange can be described through the interplay of the described knowledge management process blocks, in the following way:

- Knowledge Distribution –knowledge from different communities is made accessible to all users across community boundaries,
- Knowledge Awareness – members from different communities are made aware of the knowledge from other communities and its potential relevance for their own needs,
- Knowledge Application - members of one community access knowledge from another community and apply it for the accomplishment of a task within their own context (knowledge reuse),
- Knowledge Creation – members of one community access knowledge from another community and combine it with their existing knowledge in order to create new knowledge required for solving a problem or satisfying an information need (knowledge creation),
- Knowledge Preservation – members from one community express the applied knowledge from another community (or the newly created combined knowledge) in their own terms and store it in the shared community space.

Framing our problem from this perspective allows us to relate the concepts from knowledge management process blocks to the cross-community knowledge exchange model based on the theory of perspective making and perspective taking presented in Section 3.1. These relationships are illustrated in Fig. 3-4.

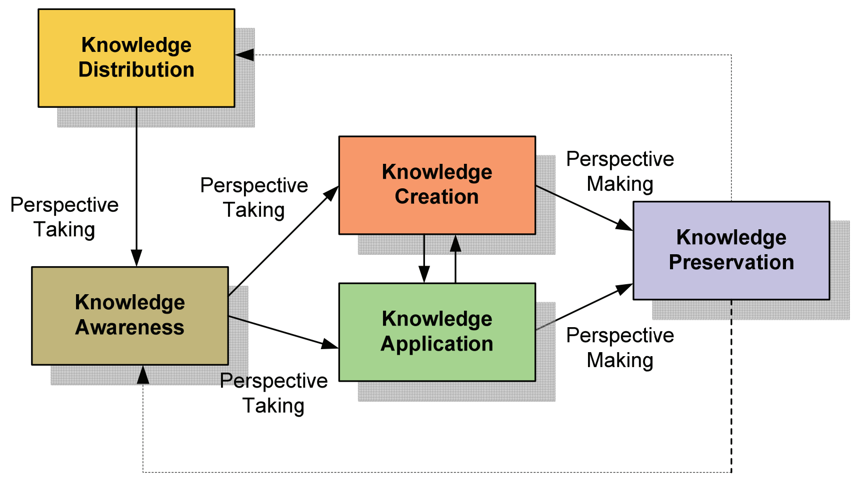


Fig. 3-4 Relationship between knowledge management process blocks and the cross-community knowledge exchange model of perspective making – perspective taking.

Becoming aware of the knowledge from an unfamiliar community and understanding its relevance for one’s need involves understanding the context of the unfamiliar community within which this knowledge makes sense (Section 3.1). This relates knowledge awareness to perspective taking. Understanding a given community perspective is also a prerequisite for being able to productively use knowledge from that community: by applying it to a given task or creating new knowledge. Accordingly, in the cross-community context, knowledge application and knowledge creation are rocesses of perspective taking.

Expressing the applied or newly created knowledge based on input from another community in one's own terms contributes to further development of the knowledge in one's own community. In this way knowledge preservation is related to and occurs through perspective making. Ensuring the practical application of a cross-community knowledge exchange support method to knowledge management thus requires us to realize the high-level requirements (Section 3.1.5) in a way that supports the knowledge management process blocks depicted in Fig. 3-4. Specific requirements for supporting these processes for knowledge exchange *within* communities proposed in (Lacher, 2003) are given in Table 3-3.

KNOWLEDGE DISTRIBUTION	– Knowledge should be uniquely identifiable, accessible and represented in a processable format. – Information pushed to community members should be highly relevant to the user's information requirements to ensure acceptance of the such distribution. – Information pull should incur very low effort for the user. – Personal and organizational privacy aspects should be respected.
KNOWLEDGE AWARENESS	– The awareness for available knowledge can be increased through the provision of knowledge resource maps as well as knowledge structure maps. – Increased awareness leads to a greater subjective information requirement and demand.
KNOWLEDGE CREATION	– Provision of knowledge should be personalised to the recipient's personal context to increase the acceptance of new knowledge. – Knowledge required for a task should be supplied pro-actively in time.
KNOWLEDGE APPLICATION	– Knowledge sources must be easy to use and ready to connect. – Knowledge must be available just in time and in an immediate acting context, i.e. just when the knowledge is required for the accomplishment of a task.

Table 3-3. Knowledge management process requirements for supporting knowledge exchange *within* communities (Lacher, 2003)

Due to the established relationship between the process blocks and our cross-community knowledge exchange model, we can now transfer these requirements to the cross-community context.

3.3.1. Requirements for Cross-Community Knowledge Distribution

Supporting knowledge distribution requires us to provide ways for making knowledge from different communities accessible to each other. In order to achieve this we need to solve two main problems:

1) ensuring the willingness of individual members to make their knowledge available to members of other communities,

2) establishing channels for information exchange between different communities.

The willingness to share knowledge within communities is based on the shared sense of community membership, mutual trust developed through participation in community activities and the desire to achieve expert reputation and social recognition within the community (Chapter 2.7.2). In contrast, in cross-community networks interactions are rare, social relationships are weak and trust is low (Chapter 2.8.2). In such conditions, establishing willingness to share knowledge can be based either on promoting the establishment of cross-community expert reputation (reputation economy principle) or on providing immediate personal benefits as a direct result of making one's own knowledge available to others (reciprocity principle). In order to achieve this, the shared knowledge must have high perceived relevance for members from other communities and accessing it must require low effort.

As discussed in Chapter 2.8, the primary means of knowledge exchange between different communities is shared information access. The establishment of cross-community portals is a common way for providing access to shared information spaces of different communities. In this context, the requirement for ensuring the relevance of shared knowledge for other communities means that information must be provided in a contextualized manner which allows it to be easily comprehended by members outside of the community. The challenge here is how to provide contextual description of information without incurring additional effort to the author or at keeping it sufficiently low with respect to the perceived benefits. At the same time, the recipient of the information should able to access it with respect to his own context of needs. These two requirements point to the need for personalised information access based both on the contextualisation by the author and the personal and community contexts of the recipient.

Realising such personalisation services requires that the contextual information be provided in a machine processable format that can be automatically evaluated by cross-community matchmaking or recommendation agents. In order to respect the requirement of privacy and trust within communities, cross-community information access must distinguish between private parts accessible only to community members, and the public parts which can be accessible also from the outside. Cross-community distribution then requires flexible and unobtrusive means of specifying publicly available knowledge that don't incur extra effort to the author.

KNOWLEDGE DISTRIBUTION
• **The willingness of individual communities to make their knowledge available to members of other communities must be established.** The emergence of natural, self-organized incentive mechanisms should be supported such as: o Reputation economy – promoting cross-community expert reputation. o Reciprocity incentives – providing immediate personal benefits as a direct result of making one's own knowledge available to others. • **Knowledge made available by members from individual communities must have high perceived relevance for members from other communities and accessing it must require low effort.** Knowledge distribution based on information sharing requires that information be provided in a contextualized manner, which allows it to be easily comprehended also by members outside of the community. This requires: o Easy creation of context information without incurring additional effort to the author o Representation of contextual information in a machine processable format that can be automatically evaluated by personalisation and recommendation services o Personalisation and recommendation services based both on the contextualisation by the information author as well as the personal and community contexts of the recipient • **Privacy and trust-boundaries should be respected by providing means to distinguish between private knowledge accessible within a community and public knowledge available outside community boundaries.** This requires unobtrusive means of specifying publicly available information which don't incur extra effort to the author.

Table 3-4. Summary of requirements for cross-community knowledge distribution

3.3.2. Requirements for Cross-Community Knowledge Awareness

Supporting knowledge awareness requires the construction and use of knowledge structure maps and knowledge resource maps that provide an overview of the content, location and means of access to available knowledge resources. In order to be usable for cross-community knowledge exchange such maps must provide an overview of knowledge from different communities and support the discovery of relationships between them.

But knowledge structure maps of individual communities are not readily available. As observed in Chapter 2.7 communities rarely create explicit representations describing the structure of their knowledge. Furthermore, due to the self-organized and informal nature of communities community knowledge is not uniform but incorporates many different, individual points of view. Since different communities use very different terminologies, knowledge awareness cannot be supported by the provision of isolated knowledge maps of different communities, even if they would exist. According to the requirement of perspective taking, in order for members of one community to recognize the relevance of knowledge from another community, they need to be able to locate knowledge from unfamiliar communities, within their own community perspective (Chapter 3.1.2).

In other words, the concept of knowledge structure maps provides an operationalization of the notion of boundary objects described in the high-level requirements for cross-community knowledge exchange (Section 3.1.5). Since the requirements of the high-level model distinguish between boundary objects representing personal, community and cross-community perspectives, this implies that knowledge structure maps should support all those three levels.

KNOWLEDGE AWARENESS
• **Knowledge structure maps should provide an overview of knowledge from different communities and support the discovery of relationships between them.** They should be provided insight into personal, community and cross-community perspectives. Such maps should visualise the existing but not explicitly formulated knowledge structures of individuals and communities of users in a way that: • Reflects the shared vocabulary used by community members (concepts, their meanings and relationships). • Incorporates individual points of view of individual community members, • Provides ways for contextualising knowledge from one community in the context of the other.

Table 3-5. Summary of requirements for cross-community knowledge awareness

As a result, supporting cross-community knowledge awareness requires the construction of knowledge structure maps that visualise the existing but not explicitly formulated knowledge structures of individuals and communities of users in a way that:

- Reflects the shared vocabulary used by community members (concepts, their meanings and relationships).
- Incorporates individual points of view of individual community members,
- Provides ways for contextualising knowledge from one community in the context of the other.

According to the findings of (Probst et al.,1997) the provision of knowledge structure maps increases the awareness of available knowledge which in turn leads to greater subjective information requirement and demand (Table 3-3). Thus, the provision of cross-community knowledge maps satisfying the above requirements should lead to increased demand for cross-community information access which in turn should result in increased cross-community knowledge exchange.

3.3.3. Requirements for Cross-Community Knowledge Creation

According to both Probst's findings knowledge creation is not only a result of a conscious deliberation aimed at creating new knowledge but an intrinsic part of every information access. Information is always interpreted with respect to a given problem context or information need and with respect to one's existing knowledge. In the perspective making - perspective taking model, knowledge from an unfamiliar community is internalized by the user only through translation (implicit or explicit) into his own terms.

This process is particularly important when knowledge exchange is mediated through the exchange of information in documents. The re-creation of knowledge is then highly dependent on the recipient's interpretation of the context within which the meaning of information contained in the document is defined. Within communities this re-contextualisation of information is relatively unproblematic since community members share the same common ground. But when accessing information from an unfamiliar community the provision of contextual information is of critical importance.

This requires that contextual information from the source community needs to be available and put in relation to the users familiar context. This context is thereby given by the personal knowledge of the user and by the context of the community to which he belongs. In this way new information can be related to the user's existing knowledge. By expressing the understanding of this knowledge in his own terms the user explicitly establishes relationships between his existing knowledge and the new knowledge created as a result of the information access. In this sense, the requirement for personalised access from Table 3 can be translated to the cross-community context in a twofold manner:

- Providing contextualized information access within the knowledge structure of the information source community,
- Supporting the discovery of relationships to the user's personal and destination community knowledge structures.

This kind of contextualized information access considers user's personal knowledge structures that reflect his general or long-term needs. Another important requirement for supporting knowledge creation in Probst's model is the provision of access to shared knowledge and experiences contextualized with respect to a given task from everyday work. This points to the need for contextualisation support with respect to a short-term need represented by a given task. This need is commonly expressed through search queries in everyday information access. Accordingly, supporting knowledge creation requires providing personalised information access that combines long-term need reflected in user's personal and community knowledge structures with short-term needs expressed through search queries reflecting the needs of a given work task.

With respect to knowledge-intensive tasks that involve deliberate creation of new knowledge Probst's model additionally points out several requirements for supporting creativity. They include:

- Supporting the discovery of different viewpoints on a problem and relationships between them,
- Methods for externalising personal, implicit knowledge in ways that makes it visible and usable for others.

This corresponds to the high-level requirement for the construction of personal boundary objects that reflect personal knowledge of members from different communities and can be exchanged with others. In the context of the above requirements such artefacts can be realized as knowledge structure maps that describe the knowledge resources of a given user and the concepts he uses in referring to them. Making such personal knowledge structure maps accessible to users from different communities should then be realized in ways that allow them to serve as means for contextualized access to information from different points of view. Finally, such a variety of different ways for contextualising and inspecting information and relationships between the different contexts in which it makes sense points to the greater importance of explorative access vs. goal-directed search for cross-community knowledge creation.

KNOWLEDGE CREATION
• **Access to knowledge from unfamiliar communities should be contextualized within the recipients existing knowledge as well as within the context of the given task.** Such contextualization should support understanding the unfamiliar knowledge and its internalization by the user. This requires: o Providing insight into the original context of the information o Relating unfamiliar information to familiar personal and community knowledge structures o Contextualising retrieved information to the needs of the given task, o Supporting the discovery of relationships between the original information context (source community) and the user's personal and community knowledge structures, o Representing personal knowledge of members from different communities relevant for a given task. These functionalities should be integrated with the creation and visualisation of personal, community and cross-community knowledge structure maps used for knowledge awareness. • **The new knowledge internalized by the user should be made explicit and incorporated into personal and community knowledge structures** This requires unobtrusive mechanisms for easily expressing the relationship of unfamiliar information to the concepts compatible with one's own knowledge and the context of the given task. • **Explorative information access should be combined with goal-directed search.** This should support the discovery and understanding of different contexts of meaning of the retrieved information.

Table 3-6. Summary of requirements for cross-community knowledge creation

3.3.4. Requirements for Cross-Community Knowledge Application

The notion of knowledge application emphasises the difference between making it possible to identify and access relevant knowledge (knowledge awareness, knowledge distribution) and putting this knowledge to actual use. Organisational and social factors can have a major impact on whether newly identified knowledge is accepted and actually used (e.g. the "not invented here" syndrome). According to Probst et al., the role of technology is to increase the probability of knowledge use by making it easily accessible by contextualising it within the given knowledge need. Allowing users to easily connect it with their existing knowledge is another important enabling factor.

KNOWLEDGE APPLICATION
• **Accessing knowledge from unfamiliar communities needs to be contextualized within the needs of a specific task at hand.** This requires information contextualisation services that support: o Quick discovery and selection of information from unknown domains relevant for a given task, o The discovery of relationships between retrieved information and the concepts describing the short-term need of the task

Table 3-7. Summary of requirements for cross-community knowledge application

The focus of knowledge application is location and reuse of existing knowledge and experiences of others to the task at hand. This places the focus on the short-term need defined by the task at hand, rather than the user's general interest profile. While this leads to similar requirements for contextualized information access as for knowledge creation, the major difference is the stronger emphasis on convenience. Critical requirements for knowledge reuse are the ease with which this knowledge can be located, selected and transferred into the context of the given task. In Probst's model this is described with the concepts "ease of use", "just in time" and "ready to connect". According to (Markus,. 2001) knowledge reuse can be described more precisely by following activities:

- Location of information containing relevant knowledge,
- Selection of relevant items from a set of search results,
- Applying the knowledge in a particular work context.

While location and selection activities are common information seeking activities, applying the identified knowledge in a particular work context is a more complex activity. It involves interpreting the meaning of information and relating it to existing knowledge relevant for the task at hand, in order to identify to which part of the problem it can be readily applied. This means that the located and selected information needs to be put in relation to concepts describing the information need of the task. The latter include search terms, the concepts used by the user to describe the selected information as well as the concepts used by other users in a similar context. Accordingly, the main requirement for cross-community knowledge application is the availability of contextualisation services that allow quick discovery and selection of information from unknown domains relevant for the task at hand, and the discovery of their relationships to the concepts describing the short-term need of the task at hand

3.3.5. Requirements for Cross-Community Knowledge Preservation

Supporting knowledge preservation in the cross-community context requires that the discovered relationships between different community domains should be made explicit and available for future use. Accordingly, users should be supported in expressing the discovered insights in a way which related concepts from different communities to each other and makes them usable for other users e.g. for navigating unfamiliar community spaces. Finally, these structures should evolve with the dynamics of community knowledge and reflect the patterns of actual use

KNOWLEDGE PRESERVATION
• **The relationships between knowledge from different communities established by the users should be made explicit and available for retrieval by other users.** This requires: o Provision of conceptual structures that relate the personal concepts created by users from different communities, to each other. o Mechanisms for assigning information from a given community space to concepts created by different users. o Mechanisms for updating these structures based on their evolvement in actual use.

Table 3-8. Summary of requirements for cross-community knowledge preservation.

3.4. Sensemaking and Information Access

In the above analysis we have identified a set of requirements describing different aspects of the need for making the implicit knowledge structures of individuals and communities of users visible and usable for contextualized access to information from unfamiliar communities. We have considered the requirements imposed by the needs for supporting typical knowledge management processes in organizations. In doing so, the basis of our analysis has been the process of knowledge exchange taking place through the exchange of information i.e. accessing information from different community spaces. To understand how such requirements can be turned into a specific solution, we now analyse a particular class of information access tasks that are specific to the context in which the need for cross-community exchange occurs.

The need for accessing knowledge from unfamiliar communities occurs in practice when people face complex or ill-structured problems that do not fit within their known knowledge contexts. Examples of such classes of tasks include strategy making and development of new products or project ideas (Fuchs-Kittowski et al., 2003) as well as training, education, business intelligence (Russel et al., 1993) and self-directed learning (Qu & Furnas, 2005). All of these tasks require learning about potentially relevant but unfamiliar topics, domains of knowledge and areas of expertise. In particular, they require users to actively construct new knowledge by developing an understanding of the meaning of information from unfamiliar contexts with respect to the task or problem at hand.

This relates the problem of supporting cross-community knowledge exchange very closely to the problem of sensemaking investigated in the field of human-computer interaction (Russell et al. 1993; Baldonado & Winograd, 1997; Qu & Furnas, 2005). Sensemaking is a process that occurs when people face unfamiliar problems, situations or tasks where their existing knowledge is insufficient (Dervin, 1992). An important activity in sensemaking is finding and constructing representations of information that are appropriate for the task at hand. Such representations are used for organizing information in a way which relates the retrieved (unfamiliar) information to the problem or task at hand (Russel et al., 1993). The process of accessing information from an unfamiliar community and relating it to the concrete task and one's own knowledge can thus be understood as a sensemaking process. Hence, understanding how sensemaking occurs can provide us with important requirements for supporting cross-community knowledge exchange.

3.4.1. The Sensemaking "Learning Loop"

Based on a number of case studies Russel et al. (ibid.) proposed a model of sensemaking in information rich tasks. Information rich tasks are tasks which involve large amounts of unfamiliar information that is too difficult to be structured by hand. Sensemaking is an iterative process in which people successively develop structural representations of information and use them to organize information for a given task.

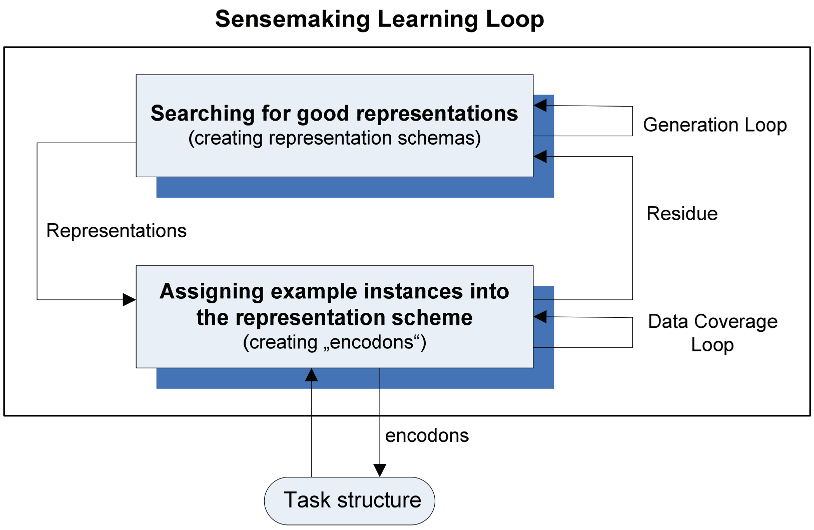


Fig. 3-5. Sensemaking model according to Russel et al. (1993)

The representations evolve based on how suitable they turn out to be for contextualising information retrieved during a search process. As a result, sensemaking involves a cyclic process of developing information representations, using them as a structure into which newly found information can be categorized and modifying the representations in order to better fit the retrieved information with respect to the needs of the specific task (Fig. 3-5).

Since this is an iterative process in which people develop new knowledge through making sense of unfamiliar information and its relationship to the given task, this sensemaking cycle is referred to as the "learning loop" (Fig. 3-5).

As we can see in Fig. 3-5 the described sensemaking cycle involves several processes that are interconnected with each other:

3. *Search for representations.*

   In order to make sense of a complex task the user ("sensemaker") creates representations of information structure (schemas) that capture the salient features of information needs of the task and can be used as templates into which relevant information can be categorized. This cycle in which the user searches for appropriate representations by considering the problem and information regarding it is called the "generation loop".

4. *Instantiation of representations.*

   After having developed an initial representation, the sensemaker continues to search for relevant information and contextualizes it into the representation created in the generation loop. The resulting artefacts - templates filled-in with specific information – are considered specific instantiations of a schema. In the sensemaking terminology this process of gathering and "encoding" information into a given schema is called the "data coverage loop". The resulting artefacts representing information classified according to the categories relevant for the task at hand are called "encodons".

5. *Representation shift.*

   Based on the extent to which retrieved information can be fitted into a given schema a "residue" can be identified. Residue includes ill-fitting information that cannot be classified into a given schema or missing information which leaves parts of the schema unused. When relevant information occurs that doesn't fit existing schema categories, the schema can be expanded or modified: categories need to be merged, split or new ones added. If parts of the schema are never filled-in they may be recognized as obsolete and removed. Thus, the emergence of residue triggers the search for more appropriate representations. This is the "representational shift" loop. The new representations are then used for encoding new information as long as sufficient residue is not built up (requiring a better representation) or the task is accomplished.

6. *Consumption of encodons.*

   The created instantiations of a schema (encodons) are used for accomplishing specific information processing steps of a task. They provide contextualized access to information required for a task. The schema representations guide this process by showing what to look for and how to organize the required information, while the embedded content provides the necessary information in the appropriate context.

According to case studies undertaken by Russel et al., the interplay of these processes represents a common pattern that people follow when making sense of a complex body of information (ibid.). An illustration of how this occurs in concrete examples of information rich tasks is given in Fig. 3-6.

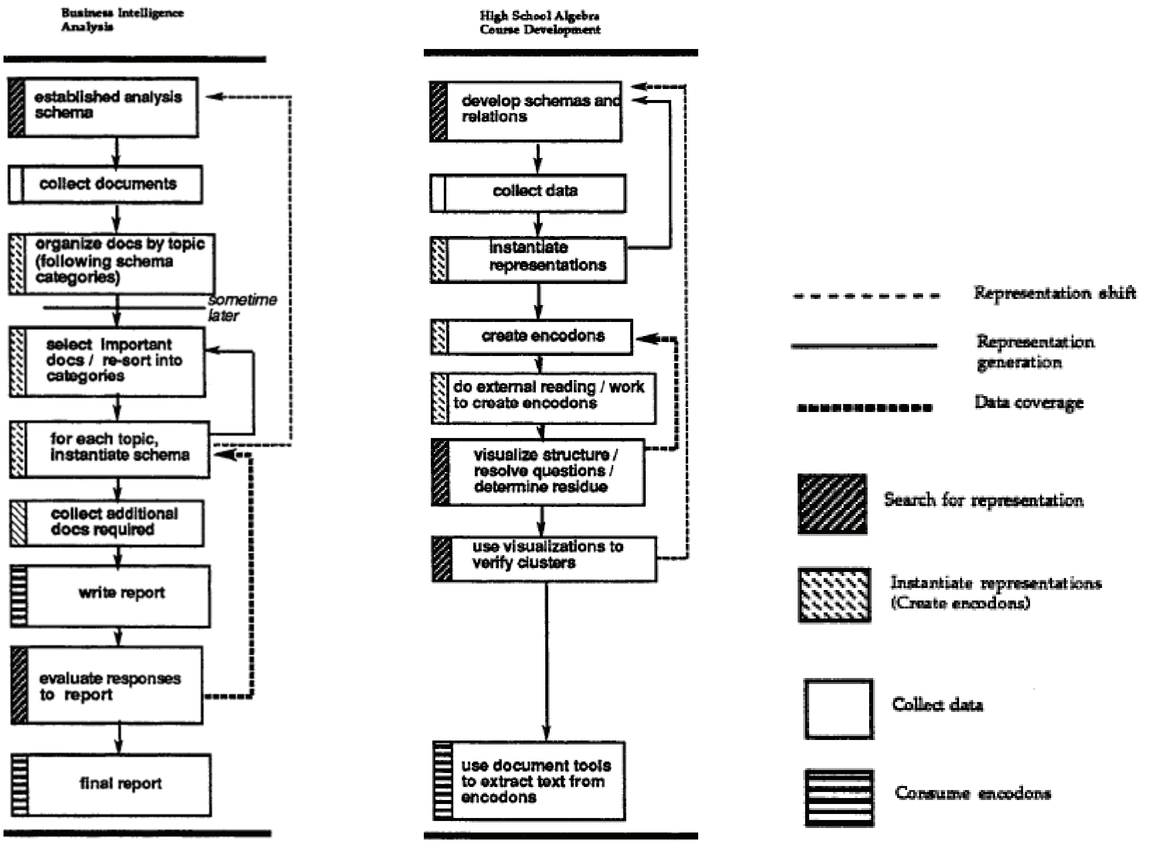


Fig. 3-6 The sensemaking process in concrete examples of information rich tasks (Russel et al., 1993).

An important insight of the described model is the critical role of external representations of information structure in supporting the activities performed to make sense of unfamiliar information and its relationship to a given task. In the approach of Russel et al., such representations are considered as resources that reduce the cognitive load of the user and make the process of identifying and interpreting relevant information more effective. In particular, they are seen as essential means for supporting three main classes of subtasks occurring in sensemaking:

- Finding the appropriate concepts characterizing the information need for a given task,
- Organizing the concepts into a form suitable for the task,
- Identifying and applying relevant information to accomplish the task.

3.4.2. Sources of Representations in Sensemaking

A critical problem in sensemaking is finding the suitable information representations for a specific task (ibid.). It involves a highly iterative process of gradually improving initial (low quality) representations with successively more appropriate ones. Thereby both top-down (applying an initial schema to new information) and bottom-up strategies (changing a schema due to breakdowns caused by newly encountered information) are used interweavingly.

In other words, the structure and the associated information need of a complex task becomes clear only as a result of having engaged in searching, sighting and organizing a certain amount of potentially relevant information. The more complete the body of the information considered so far, the higher the chance for finding a more appropriate information representation suitable for the task.

Hence, the creation of information representation structures and information seeking are tightly coupled. Results of empirical studies such as (Qu & Furnas, 2005) confirm this and point to three main ways through which people identify and construct structure in making sense out of unfamiliar information:

- Deducing from their existing knowledge (top-down),

- Inducing from retrieved information (bottom-up),
- Borrowing from structures created by others in their sensemaking efforts.

A more detailed distinction of the different modalities of representation construction identified by Qu and Furnas is presented in Table 3-9.

DATA-DRIVEN REPRESENTATION CONSTRUCTION	GENERATING STRUCTURE FROM OWN KNOWLEDGE
Adopting other's representations. Picking up interesting ideas. Re-evaluating and adjusting existing representations. Re-evaluating the sensemaking task.	Using existing own knowledge schemas related to the task. Top-down, step-by-step reasoning based on own knowledge.

Table 3-9. Representation construction strategies. Based on findings from (Qu & Furnas, 2005)

These findings point to two important implications. The first is that available information structures created by other users play an important role in the sensemaking process. This includes both directly adopting other's representations (e.g. classifications in literature reviews, bookmark structures and categories used to describe a given concept) as well as the triggering of representation ideas based on unusual, interesting aspects of information content (picking up an interesting idea). Similarly, the mere amount of available information on a topic of interest can lead to re-evaluation of existing representations or even of the sensemaking task itself (e.g. large availability of information on a certain category can favour its use with respect to another).

Hence, it is not only the retrieved information that influences the development of representational schemas but also the available structural information attached to or embedded into information content. This implies that the encoding activity (assigning information to a schema) and the representation search activity (creating a schema) are much more interconnected than suggested by the Russel et al. model.

The second implication is that in addition to retrieving relevant information another very important part of sensemaking is locating already available representations of information structure (e.g. bookmarks or outlines created by other users). These implications point to the need for a tight integration of structuring and information seeking activities in sensemaking support systems. More specifically, according to (Qu and Furnas, 2005) the following aspects should be given special consideration when designing sensemaking support systems:

- The availability of structured information resources created by other people offers great advantage for representational search in sensemaking. This should be incorporated into information seeking systems, e.g. by higher ranking of structured information elements or sources.
- The use of external representations should be more actively incorporated into search systems such as proposing appropriate representation structures for a given query.
- Representation comparison tools are required in order to allow users to compare the representations of other users to their own.

3.5. Summary of Requirements for Supporting Cross-Community Knowledge Exchange

The analysis presented in this chapter shows that supporting cross-community knowledge exchange requires three main sets of requirements to be considered. The high-level model of perspective making – perspective taking requires that implicit knowledge structures of communities be elicited and visualised in a way that reflects intra-community dynamics (include personal points of view, evolve over time) and supports the discovery of cross-community relationships. Visualisation of such knowledge structures needs to support three main classes of high-level tasks: 1) gaining insight into implicit knowledge structures of unfamiliar communities (*perspective taking*), 2) discovering how knowledge reflected in information artefacts from unfamiliar communities relates to own knowledge and internalizing the insights in own structures (*perspective taking -> perspective making*), and 3) expressing the new knowledge in terms appropriate for one's own personal and community perspectives (*perspective making*). The practical requirements of knowledge management process show how the interplay of perspective making and perspective taking should be related to knowledge awareness and distribution as well as to knowledge creation, application and preservation. They point to the need to consider social aspects of these processes (e.g. privacy and social reputation) and relate these knowledge processes to information access and sharing which is the primary way of knowledge exchange in the cross-community context. In particular, they highlight the need for making personal and shared knowledge maps usable for awareness and contextualization of unfamiliar information, for discovery of relationships between unfamiliar community domains and the needs of a given task as well as for personalised information filtering and contextualized recommendations of knowledge relevant for a given information need (knowledge distribution and application).

	Knowledge Awareness	Knowledge Distribution	Knowledge Creation	Knowledge Application	Knowledge Preservation
Representation Search	identifying appropriate concepts from unfamiliar community	finding existing schemas in unfamiliar community	creating new schemas (organizing concepts, creating new concepts)		creating new schemas (organizing concepts, creating new concepts)
	Perspective Taking				
Instantiation of Representations	finding relevant information	finding relevant information	assigning / classifying information into own schemas (creating encodons)	using existing schemas	assigning / classifying information into own schemas (creating encodons)
				Perspective Taking → Perspective Making	
Representation Shift			modifying a schema (removing, adding, re-organizing concepts)		
Consumption of Encodons				contextualised information access & processing for the task	
				Perspective Making	

Table 3-10. Relationships between the different dimensions of requirements for cross-community knowledge exchange

To support knowledge creation, the discovery of relationships between different personal points of view and the needs of a given task is of critical importance and may enable seeing a given problem in fundamentally new ways. This should be accompanied by unobtrusive and natural ways of expressing personal knowledge discovered in the process in one's own terms and showing how it relates to a specific need. Eventually, the relationships between knowledge from different communities discovered by the users in using the maps need to be externalized and made usable for others (knowledge preservation)

Finally, the sensemaking model reveals the typical tasks that need to be supported during information access in unfamiliar domains. It emphasises the importance of external representations of information structures that people develop in the process of making sense of unfamiliar information and ill-defined tasks. The sensemaking learning loop describes how new knowledge is constructed in an interplay between seeking and identifying relevant information and searching for appropriate representations for organizing it into meaningful structures. Making personal information structures of other users available during information access can greatly support this process as they can be used as semantic templates for identifying and contextualizing information from an unfamiliar domain. Table 10 shows how typical sensemaking tasks during information access in unfamiliar domains connect the different dimensions of requirements for cross-community knowledge exchange with each other.

4. Knowledge Visualisation and Visual Information Seeking

The presented requirements for supporting cross-community knowledge exchange (Chapter 3) place great importance on the elicitation and visualisation of implicit knowledge structures of individuals and communities and on their use to supporting specific sensemaking tasks during access to unfamiliar community spaces. In addressing these requirements in this thesis we propose an approach based on a combination of knowledge visualisation and visual information access interface. Accordingly, this chapter gives an overview of existing methods of knowledge visualisation and knowledge organisation and their use for supporting information access.

The chapter starts by more closely defining the notion of knowledge visualisation, in particularly the difference and relationships to information visualisation (Section 4.1). Since cross-community knowledge exchange is essentially mediated through access to unfamiliar community information spaces (Chapter 2.8) we then consider existing approaches to extracting and visualising semantic structures of document spaces (Section 4.2). This is followed by an overview of methods for eliciting knowledge from human users (Section 4.2) and methods of knowledge organisation and representation (Section 4.3). Finally, examples of approaches to using visualisation specifically for supporting information seeking are discussed (Section 4.4). Each section concludes with a discussion of advantages and shortcomings of presented solutions with respect to our problem context.

4.1. Knowledge Visualisation vs. Information Visualisation

In the HCI field, knowledge visualisation has often been used interchangeably with information visualisation. The research on information visualisation has a long tradition and is concerned with "the use of computer-supported, interactive, visual representation of abstract non-physically based data" with the goal of "amplifying cognition" (Card, Mackinlay & Shneidermann, 1999a). In contrast, the concept of knowledge visualisation is defined as being concerned with the "use of visual representations to improve the creation and transfer of knowledge between people" (Eppler & Burkhard, 2004). Such a definition makes clear a fundamental difference between the two fields of work: while information visualisation addresses problems of graphically displaying complex structures of abstract data, knowledge visualisation is intrinsically concerned with using visualisation to address problem of knowledge transfer in social structures.

The view that "knowledge visualisation aims to improve the transfer and creation of knowledge among people by giving them richer means of expressing what they know" (ibid.) puts emphasis on the relationship between knowledge and human actors - even when designing and developing visual knowledge artefacts, which will inevitably be based on some form of visual presentation of information. As a result, knowledge visualisation research often investigates the use and application of existing information visualisation methods for solving or supporting specific knowledge transfer and knowledge management problems (Burkhard, 2005a).

From this perspective, the concept of knowledge visualisation is inherently related to the perspectives of knowledge management and computer-supported cooperative work. The work on technological support for knowledge management has been concerned with supporting the organization, storing, access, sharing and creation of knowledge in organizations. In this field, the notion of knowledge visualisation has traditionally been used interchangeably with the concept of knowledge representation. Correspondingly, knowledge visualisation approaches in KM have largely been concerned with the development of formal (machine readable) models for representing the semantics of individual pieces of information and relationships between them. This includes various forms of semantic networks such as ontologies, taxonomies and topic maps (Section 4.4). Information visualisation methods such as hierarchical graphs, hyperbolic views or graph-layout algorithms are commonly used to represent such semantic structures (Chen, 1999). The combination of these methods has been largely adopted as standard solutions for generating and representing knowledge structures, regardless of the specific social context which they aim to support.

On the other hand, the main perspective of the CSCW research field is that of understanding and supporting communication, cooperation and coordination of groups working on shared tasks. Thus, although cooperative knowledge creation is a core concern in CSCW (e.g. de Michelis, 2001*)* research in this field has been largely concerned with structures, processes, models and functionalities supporting social processes and outcomes of cooperation, rather than with the role of specific visualisation models and interfaces. The question of how different methods and kinds of knowledge visualisation can be developed and used to specifically support different forms of cooperative knowledge creation and sharing have been only very partially addressed. Most closely related is the work on supporting group decision making (Conklin & Begeman, 1998) and the work on awareness (if the notion of knowledge is interpreted as the "knowledge of the activity of others" (Dourish, 1997; Schmidt, 2002).

An exception are approaches to concept mapping in learning research and CSCL which can be more specifically related to the presented concept of knowledge visualisation (Section 4.2). However, this work has mostly focused on concept mapping in the context of learning of individuals and groups of learners in classroom settings or collaborative problem-solving (Bruillard and Baron, 2000; Cañas et al., 2001). Similarly, the work on cognitive mapping in knowledge management considered mainly physically co-located group work settings concerned with collaborative problem-solving and decision making (see Huff & Jenkins, 2002).

Thus, while the concept of knowledge visualisation is interrelated with different research perspectives and approaches in HCI, KM, CSCW and CSCL the existing research perspectives suffer from two main problems. Firstly, their usage of the notion of knowledge visualisation is not properly defined and is used interchangeably with concepts such as information visualisation, knowledge representation or cognitive mapping. This is due to the fact that only recently a systematic attempt at providing a proper definition of the concept of knowledge visualisation and its relationships to different research fields has been undertaken (Eppler & Burkhard, 2004; Burkhard, 2005a). Secondly, besides concept mapping applications in the domain of education, little work provides theoretically founded and empirically evaluated approaches to developing knowledge visualisation methods and tools aimed at supporting knowledge exchange in specific social configurations.

In order to lay the ground for our approach to using knowledge visualisation to supporting cross-community knowledge exchange we distinguish between several classes of existing methods for eliciting and visualising knowledge structures and using them to support information access. On one hand, implicit knowledge structures of communities could be extracted from community information spaces to the extent to which they are reflected in the document contents. To this end, methods for extracting and visualising semantic structures of document spaces are a part of the design space of possible solutions to our problem and application context (Section 4.2). On the other hand, implicit knowledge structures of human users can be extracted from texts only to a limited extent. In particular, with respect to the requirement of eliciting and visualising personal points of view requires us to consider how we can elicit implicit knowledge from human users. Thus, relevant methods and approaches to knowledge elicitation (such as cognitive mapping and concept mapping) are discussed in Section 4.3.

A related problem is how knowledge structures can be represented and/or created manually. While not our main focus, a brief outline of common approaches to creation and visualisation of semantic structures such as ontologies, topic maps and knowledge maps does provide us with useful insights. Finally, this overview concludes with approaches to using visualisation and graphical interfaces to support information seeking, in order to ensure that our solution of applying knowledge visualisation to support information access is informed by rich experiences available in that field.

4.2. Visualising Semantic Structures of Document Spaces

Approaches to generating semantic overviews of unstructured document spaces aim at unveiling the intrinsic structure of implicit semantic relationships in a document collection. This includes structural properties such as main topics and clusters of documents and relationships between them. Semantic visualisations of document collections position documents in a 2D or 3D space in such a way that spatial positioning of individual documents reflects inter-document similarities: similar documents are positioned close to each other. The representation of such semantic structures inherent in a given document collection is also referred to as visualisation of semantic spaces (Chen, 1999).

The idea is that mapping and visualising characteristics such as the distribution of documents into topical clusters, their semantic density and relationships between them is a way of visualising the structure of knowledge contained in a given document collection. Presenting such overviews allows the users to quickly grasp the overall structure of a knowledge domain represented by a document space.

Based on the criteria used to determine the inter-document similarity two main classes of approaches can be distinguished: approaches based on content-similarity and approaches based on context-similarity such as co-citation and link-analysis. Once the inter-document similarity structure is computed, different visualisation methods can be used for displaying the document space. Most commonly, multi-dimensional scaling methods in combination with scatter plots are used to create spatial visualisations that preserve similarity relationships. Other popular alternatives include force-directed placement for drawing graphs and networked structures (e.g. Eades, 1984, Fruchterman & Reingold, 1991) as well as the use of self-organized maps for topology preserving visualisation of high-dimensional spaces on a two-dimensional grid (e.g. Lin et al., 1991; Honkela et al., 1997).

4.2.1. Content-Similarity Analysis

The two most prominent approaches to calculating inter-document similarities include the vector space model (Salton et al., 1994) and latent semantic indexing (LSI) (Deerwester et al., 1990). In the vector space model, each document is represented by a vector of terms with the weight of a given term represents the significance of that term for a document in question. The term significance is calculated by taking the frequency of term occurrence in the document and multiplying it with the inverse document frequency, in order to lower the impact of terms occurring in many documents: the *tf x idf* scheme[15] (Fig. 4-1).

$$w_{ik} = \frac{tf \times idf}{lengthnorm} = \frac{tf_{ik} \times \log\left(\frac{N}{n_k}\right)}{\sqrt{\sum_{j=1}^{T}\left(tf_{ij} \times \log\left(\frac{N}{n_j}\right)\right)^2}}, \qquad sim\left(D_i, D_j\right) = \sum_{k=1}^{T} wik \times wjk$$

$D_i = (w_{i1},\ldots,w_{iT})$ – document vector, w_{ik} - weight of term T_k, tf_{ik} – frequency of term T_k in D_i.

n_k - number of documents containing term T_k, N - total number of documents in the collection.

sim (D_i, Dj) - similarity between document vectors D_i and Dj.

Fig. 4-1. Computation of term weight and inter-document similarity in the vector-space model.

The similarity of two documents can now be calculated by taking a similarity measure of the corresponding document vectors. Frequently used similarity measures include the scalar dot-product i.e. the cosine similarity measure and Euclidean distance. A specific characteristic of the vector space model is that both documents and queries are represented by vectors. This allows the relevance of a document for a given query to be calculated by the same measure of vector similarity used to compare documents

[15] This is a popular variation of the basicc "bag of words" method for encoding semantic properties of documents into numerical representations based on a set of indexing terms and their frequency of occurrence in a given document.

between each other. The calculation of inter-document similarities based on the vector space model provides the basis for spatially organized visualisation of the document space where similarity is reflected in spatial distance.

The technique of latent semantic indexing has been designed to overcome the vocabulary mismatch problems (Deerwester et al., 1990): related documents may be indexed by different terms due to the different use of vocabulary by the document authors. The underlying assumption of LSI is that due to a certain degree of randomness of word choice (in authoring and retrieval) the inherent semantic structure of a document space can be better estimated with statistical techniques. Accordingly, LSI introduces a method based on mathematical decomposition of the document-term matrix into a lower-dimensional concept-document space approximating the original matrix. To this end a document-term matrix containing the frequency of occurrence for each term in each document is constructed and then decomposed into a set of orthogonal dimensions (less than the total number of terms) by means of singular value decomposition. The result is a vector space in which each vector represents a concept reflecting the correlations in the use of different terms across different documents. Inter-document similarity can then be calculated by taking the dot-product (cosine similarity measure) between corresponding concept vectors. The main claim of the LSI method is that thanks to such concept-based approximation of original document-term matrix, related documents can be determined even if they do not contain the same words.

4.2.2. Co-Citation and Link-Analysis

Another way of determining the similarity between documents is by taking advantage of explicit inter-document links. An example, is the citation analysis for scientific publications (Chen, 1999a). Here, bibliographic references provide explicit links between documents and the similarity between pairs of documents is measured by the number of publications in which both documents are cited. Another variation of this method, is author co-citation analysis (ibid.). Instead of inter-document similarity, the citations are used to determine similarity between document authors. The measure of author similarity is the number of documents citing publications of both authors. Applying methods such as multi-dimensional scaling (see next section) clusters of authors can be visualised presenting most important topical areas of the document space. This approach is also referred to as knowledge domain structure visualisation (ibid.). Due to its dependence on the convention of systematic creation of bibliographic references, this approach is limited only to application fields where the convention of such structures knowledge production exists and plays an important role, so that the quality of links is ensured (e.g. scientific research).

A representation of explicit relationships between documents, are also hypertext links created by the document authors. Accordingly, citation and co-citation analysis can be applied also to such links as a means of measuring inter-document and inter-author similarity, respectively. Thereby a distinction between semantic links reflecting content similarity and structural links is needed. A solution for combining similarity information from both kinds of links was presented by a generalized similarity framework proposed in (Chen, 1997).

4.2.3. Clustering

Based on the inter-document similarity information produced by the above methods, the groups of most related documents can be explicitly computed by the use of clustering methods. Such methods divide a data set into a number of sub-sets (clusters) based on a given similarity measure. Three main categories of clustering techniques include: graph-theoretical, single-pass and iterative algorithms. The following brief outline is based on (Chen, 1999). Graph-theoretical algorithms take advantage of an inter-similarity matrix and form clusters of closely related documents based on a similarity threshold. Each cluster is represented as a connected graph and depending on how the documents are separated the method is referred to as “single-link”, “group average” or “complete link” clustering.

An example of single-pass clustering is seed-oriented clustering. Here, clusters are “grown” from a individual data points (cluster seeds) by assigning documents to the cluster with a most similar seed. Such methods require a number of clusters to be defined in advance. Iterative algorithms employ heuristics in order to iteratively optimize the cluster structure, starting from an initial configuration (random or produced by another algorithm).

Popular clustering methods include K-means and hierarchical clustering. The first is fast but requires the number of clusters to be defined in advance. The second produces hierarchical structures by merging smaller clusters into larger ones. The number of clusters must not be defined in advance but the user can select the desired level of clustering interactively after the cluster structure has been produced, by simply moving down the individual hierarchical levels. However, rather than explicitly identifying clusters a more common approach to visualising the semantic structure of document spaces is to directly visualise the similarity structure and allow users to visually detect the existence of clusters based on the spatial arrangement of documents.

4.2.4. Visualisation Methods Based on Similarity Relationships

In order to realize a spatial visualisation metaphor, a coordinate-based representation of the document space based on the computed inter-document similarities needs to be produced. This section gives a brief outline of main techniques used for producing spatial visualisations of similarity data. It is based mainly on the overviews presented in (Becks, 2001) and (Chen, 1999).

Since the document term vectors are typically high-dimensional, methods for reducing the number of dimensions to produce 2D or 3D visualisation based on the proximity of document vectors, are commonly applied (e.g. multi-dimensional scaling, Section 4.1.3.1). The similarity structure of the document set can then be produced directly by using scatter plots (Section 4.1.3.2). The resulting visualisation shows the distribution of the document set in a way that preserves distance i.e. absolute similarity relationships between individual elements. A different method is the use of self-organized neural networks to map high-dimensional data on a low-dimensional grid based on non-linear similarity relationships (Section 4.1.3.4). Such visualisations preserve topology but not distance, thus reflecting relative similarity relationships between elements in the document space.

4.2.4.1. Multi-Dimensional Scaling

Multi-dimensional scaling (MDS) is a collection of multivariate statistical analysis techniques which are used to transform a high-dimensional vector space into a low-dimensional space such that the distances among the vectors match the original distances as closely as possible (Davidson, 1983). For visualisation purposes MDS is typically applied to scale a high-dimensional vector space representing a set of objects to a two-dimensional or three-dimensional space in which each point represents an object, such that the distances between the points match the distances between the object vectors as close as possible.

To this end a proximity measure δ defines the degree to which two objects are similar and represents either a similarity or dissimilarity measure. MDS then constructs a coordinate-based representation of the object space based on coordinate-free proximity information (δ). Given a set O of n objects and a proximity measure δ that represents dissimilarities for each pair of objects $i, j \in O$, MDS constructs a set of vectors $x_i = (x_{i1}, ..., x_{im})$, such that the error

$$s^2 =_{def} \sum_{i,j} \left(d_{ij} - \delta_{ij}\right)^2 ,$$

is minimized (Tulsa, 1997). In the above formula $d_{i,j}$ is the distance between objects *i* and *j*, measured by a given metric in the resulting *m*-dimensional space, δij is the observed dissimilarity of objects *i* and *j*, and error *s* is called stress. The resulting object vector x_i represents the “coordinate estimates” (Becks, 2001). Applying MDS to scale the original object space to 1D, 2D or 3D-space allows us to directly visualise the object set in such a way that distances between the objects reflects inter-object similarity relationships.

MDS is commonly applied in social sciences in order to characterize and distinguish objects by some features. In marketing it is used as a means of perceptual mapping for visualising and distinguishing customer preferences on products based on a set of features. In our application context, applying MDS to a document collection allows the computation of a spatial arrangement of documents based on inter-document similarities calculated by one of the methods described in the previous chapter. As we will see later, several existing systems employ this method for providing semantic overviews of document spaces.

The main disadvantage of MDS is its complex analytic procedure requiring high computation time (*O*(n^3) for *n* objects). In addition, adding new objects to the set requires the re-computation of the whole arrangement. A modified version of the MDS that solves these problems has been presented under the name "FastMap" (Faloutsos et & Lin, 1995; discussed in Becks, 2001). Besides significantly reducing computation time (*O(kn)* for *n* objects and *k* dimensions) this method allows fast positioning of new objects without re-computation of the entire space. Though the proximity relationships are preserved less well, the much faster performance and ability of incremental object positioning makes it much more suitable than MDS for producing fast arrangements of visualisations supporting information access.

4.2.4.2. Scatter Plots

Scatter plots are used to represent a data set available in a one, two or three-dimensional coordinate representation as points at corresponding object coordinates (Fig. 4-2). They are commonly used to display previously scaled data with MDS techniques. Scatter plots allow visual exploration and discovery of basic structural relationships in the dataset, such as clusters of objects or relationships with respect to specific dimensions. In information visualisation applications of scatter plots the dimensions (axes) often have no explicit semantics but indicate the relative relatedness of objects based on a given similarity measure (e.g. inter-document similarity in document collections).

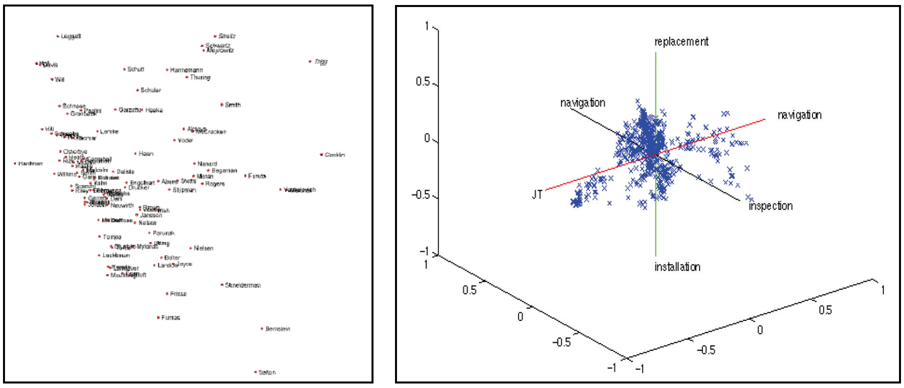


Fig. 4-2. Examples of 2D and 3D scatter plots visualising similarity data: an author co-citation map (Chen & Carr, 1999) and semantic axes of a document space (Booker et al., 1999)

4.2.4.3. Graph-Drawing by Force-Directed Placement

Another possibility for computing spatial arrangements of sets of highly-interconnected objects are graph-drawing algorithms based on the force-directed placement technique (Fruchterman & Reingold, 1991; Davidson & Harel, 1996). Such techniques model the graph as a physical system of particles attracted and repulsed by physical forces. A classical example is the spring-embedder model (Eades, 1984) in which the graph is modelled as consisting of steel rings (nodes) connected by springs (vertices). The similarity relationships between nodes determine the spring forces. The idea is that by modelling a graph with such physical forces, the objects (rings) will move based on force and energy effects. Eventually, the system will come to a halt when a state of minimal energy is reached. In order to find such an optimum state, the system is started from an initial random configuration and iteratively optimized towards lower energy levels. When the minimum energy state is found the placement of rings in this state gives the spatial arrangement of objects and connections between them. An overview of graph drawing methods can be found in (Chen, 1999).

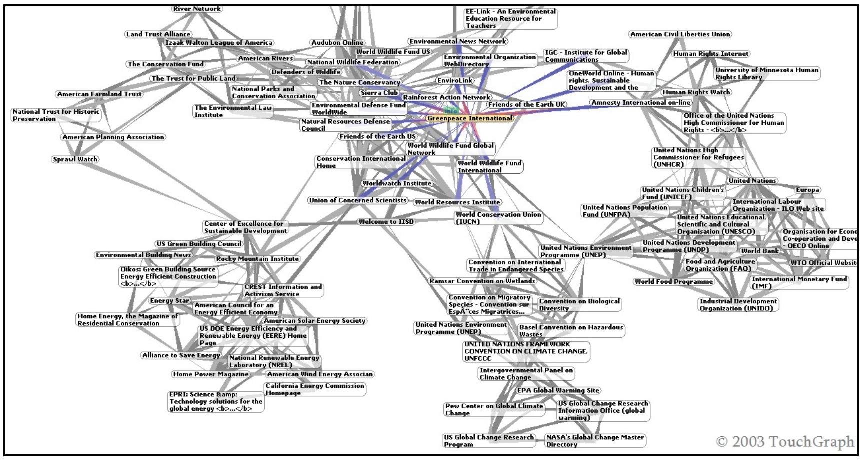


Fig. 4-3. Example of a graph-layout produced by a force-directed placement technique (touchgraph.com)

4.2.4.4. Self-Organized Maps (SOM)

Kohonen's self-organizing map is a neural network performing a mapping of high-dimensional input space on a two-dimensional grid (Kohonen, 1990, 1995). The SOM is based on an unsupervised learning algorithm producing a mapping which preserves the topological structure of the input space. The vectors close in the original high-dimensional space will also be positioned close to each other on the two-dimensional grid of neurons. The SOM performs a dimensionality reduction based on inter-document similarity and it realizes a learning function. During the learning process, each input vector is mapped to the neuron unit with the currently most similar weight vector. Accordingly, the weight vector of the winning neuron unit and of the neurons in a certain neighbourhood of the winning unit, are adjusted towards the given input vector. The amount of this shift depends on parameters such as the (time-dependent) learning rate, the distance between the weight vector and the input vector, and the unit's position in the neighbourhood of the winning neuron. The learning rate and the neighbourhood area influenced by the shift, decrease with time.

The result of training the SOM with a set of objects represented by numerical vectors is a two-dimensional grid of neurons with corresponding weight vectors, such that the vectors are ordered by their similarity: neurons that are close to each other on the grid will have similar weight vectors and clusters close to each other on the grid will have cluster centres close to each other in the input space as well. Visualising the structure of the SOM directly allows the basic structural properties of the document space, such as density of document distribution and cluster areas to be displayed. Different methods for generating such a visualisation exist, such as visualising the distances between neighbouring units in shades of grey (higher distance corresponds to higher shade of grey, Fig. 4-4, left) or colouring similar weight vectors with similar colours (Fig. 4-4, right).

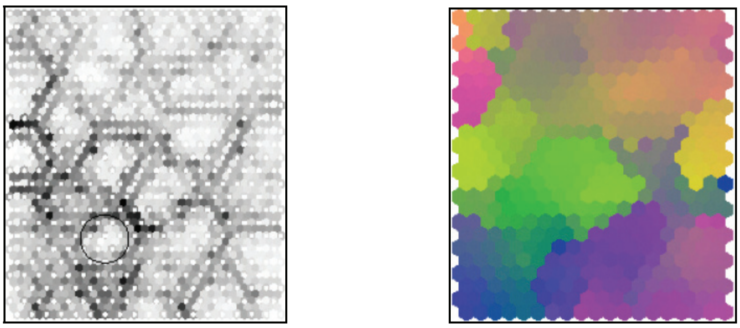


Fig. 4-4. Visualisations of the SOM based on neuron distance matrices: U-matrix (left) and similarity colouring (right). Source: (Vesanto, 1999 referenced in Becks, 2001).

The trained SOM grid can be used to map the objects from the input space by assigning each object to the grid unit containing the most similar weight vector, after the learning process. In this way a spatial distribution of a given object set on a two-dimensional surface can be created and visualised. Thereby, due to the described property of the underlying grid network, similar objects will be positioned on the map close to each other. Such a mapping preserves the topology of the input space which but it does not preserve distance. This is the basic difference to MDS which realizes a distance preserving mapping. In our application context, using the SOM to map the structure of a document space allows us to create a spatial positioning based on the preservation of the structural density of the document space and relative inter-document similarity relationships. This supports visual clustering and abstraction of the properties of the document space (Kohonen, 1995).

4.2.5. Example Systems Realizing Document Maps and Landscapes

Document maps and landscapes have been extensively used to provide visual overviews of inherent semantic structure of document collections. Since community information spaces commonly consists of largely unstructured document collections applying such methods could provide a point of departure for visualising semantic structures of community knowledge, as reflected in the documents created and exchanged by community members. The next sections give an overview of different approaches to constructing and visualising document maps.

4.2.5.1. Scatter/Gather

The Scatter/Gather technique (Cutting et al., 1992) clusters documents into topical groups and presents the result in form of textual summaries containing a list of most relevant terms and a list of typical document titles for each cluster (Fig. 4-5). Based on this information the user to select a subset of most interesting clusters and recluster only the document set contained in this selection. In this way, the user iteratively specifies sub-collections that capture his interest ("gather") and lets the system produce a new topical arrangement ("scatter"). Since different topical clusters will be produced in each step depending on a given sub-collection, the user can inspect different levels of detail of the collection structure based on a particular focus of his interest. An example of a typical Scatter/Gather interface is depicted in Fig. 4-5

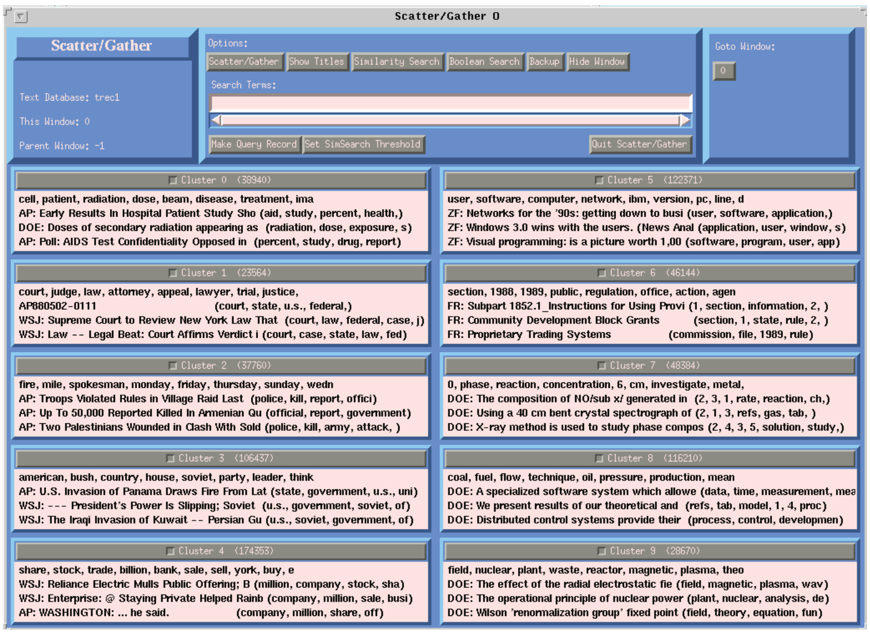


Fig. 4-5. Interface of the Scatter/Gather System (Pirolli & Card, 1995)[16]

[16] Different interfaces based on the Scatter/Gather technique exist. The depicted interface version has been developed by Marti Hearst at Xerox PARC.

4.2.5.2. SPIRE and Starlight

A well-known system for visualising relationships in textual document spaces is SPIRE – Spatial Paradigm for Information Retrieval and Exploration (Hetzler et al., 1998a,b). SPIRE includes two visualisation methods for presenting semantic overviews of document spaces (both based on term-vector calculation of document similarity), Galaxies and Themescape. Galaxies presents a document space in form of a 2D scatter plot where related documents (based on the similarity of their content) are positioned close to each the (Fig. 4-6, left). The visual metaphor is based on stars on a sky, with closely related documents representing star clusters i.e. galaxies.

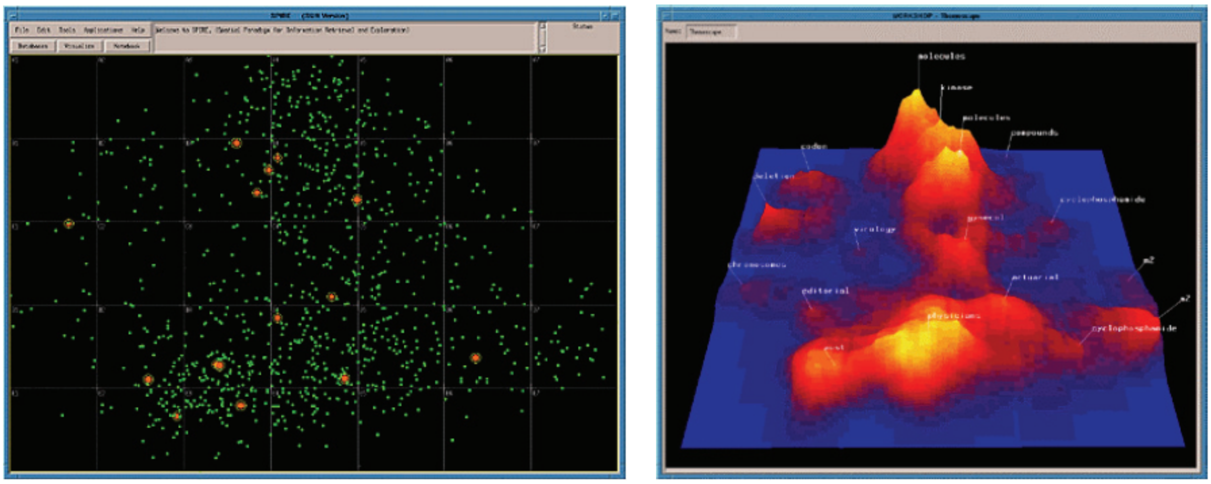
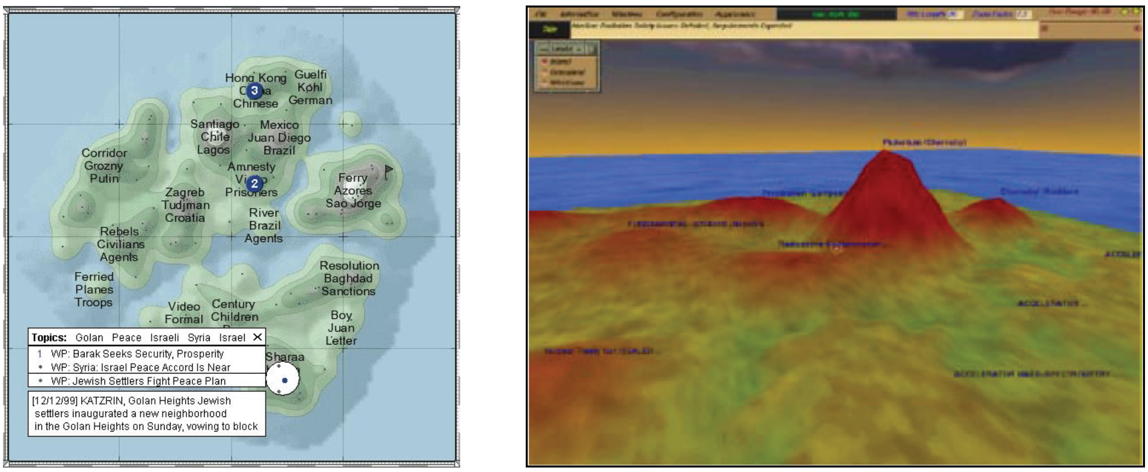


Fig. 4-6. SPIRE: Galaxies document scatter plot (left), Themescape visualisation of topic distribution (right). Source: http://www.cartia.com.

Based on the distribution of documents, Themescape visualises the distribution of main topics in the collection in the form of a mountains and valleys (Fig. 4-6, right). Both solutions use multi-dimensional scaling (MDS) to produce the basic visualisation i.e. the spatial positioning of documents. But instead of displaying individual documents, in Themescape the heights of peaks represent relative strengths of topics in the overall document set with similar topics being located close to each other. In this way, the user can quickly identify main topics of a given document space and relationships between them. Originally developed at Pacific Northwest National Laboratory (PNNL) for the U.S. intelligence community, document landscape visualisations enabled by SPIRE are produced by a commercial company (Cartia). A newer Themscape visualisation solution produced by Cartia is depicted in Fig. 4-7a.



a) Cartia Themescape **b) VxInsight Landscape**

Fig. 4-7. Document landscape visualisations of Cartia (www.cartia.com) and VxInsight of Sandia Laboratories (Source: Chen, 2002).

Another information visualisation system developed at PNNL is Starlight (Risch et al., 1996) aimed at visualising multimedia databases (also for the intelligence analysis application). It provides 3D scatter plots representing proximity of a data set based on textual descriptions of text and multimedia objects. The document similarity encoding is based on the vector space model. The 3D projection is created by means of 3D multi-dimensional scaling while an additional clustering algorithm is used for explicitly determining most closely related groups of objects presented scatter-plots within spheres delimiting cluster boundaries.

4.2.5.3. VxInsight

VxInsight is an information visualisation tool also based on a landscape metaphor, developed by Sandia National Laboratories (Davidson et al., 1998). It presents large data sets grouped into clusters of related objects based on similarity, visualised in form of a three-dimensional mountains and valleys landscape (Fig. 4-7b). The height of the mountain reflects to the density of the cluster that it represents. The peeks are labelled with titles and the results of user search queries can be visualised in the landscape. The system provides both content-similarity measures (keyword analysis) as well as citation and link-analysis. The visualisation is created in real-time based on a combination of multi-dimensional scaling and force-directed placement algorithms. The development of VxInsight has been motivated by the need to support the analysis of research literature as a means of facilitating decision-making regarding new scientific projects and strategic partnerships.

4.2.5.4. Knowledge Garden

The Knowledge Garden implements a collaborative virtual environment visualising a shared document space created collaboratively by the users. The users can meet in the virtual space and exchange information by collecting documents from the WWW, enriching them with meta-information and adding them as bookmarks into the shard information space. The shared document space visualises groups of similar documents based on hierarchical clustering (Crossley et al., 1999). The visualisation uses a garden metaphor with flowers representing groups of related documents. Single documents are represented by stalks of a flower with coloured icons at the end of each stalk indicating document status.

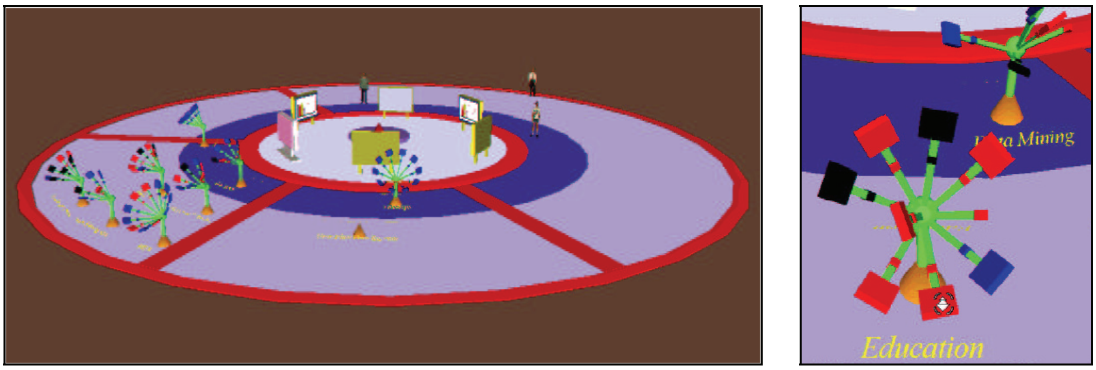


Fig. 4-8. The Knowledge Garden collaborative 3D information visualisation environment (Crossley et al., 1999): overview (left) and document cluster representation close-up (right).

4.2.5.5. Category Map

Probably one of the first applications of the SOM to visualising document spaces in order to support information access has been presented in (Lin et al., 1991). In this work, the SOM is used to create a two-dimensional category map showing main topical areas and their neighbourhood relations inherent to a document collection. The main application goal was to create a visual interface for an online bibliographic system providing access to articles in a given research domain. Accordingly, the metaphor used is a "graphical table of contents". The original solution encodes document similarities by applying a vector space model to a collection of research abstracts and providing the result as input for training a SOM. After the SOM has been trained, individual neurons are grouped into topical regions by labelling each unit with the most dominant term and visualising all units with the same label as one conceptual area. Furthermore, concept areas with frequently co-occuring terms are merged into one, thus resulting in areas labelled by multi-word concepts. An example of such a category map is depicted in Fig. 4-9: it shows a map of a collection of 140 articles from an AI research paper database, presented in (Lin et al., 1991)[17].

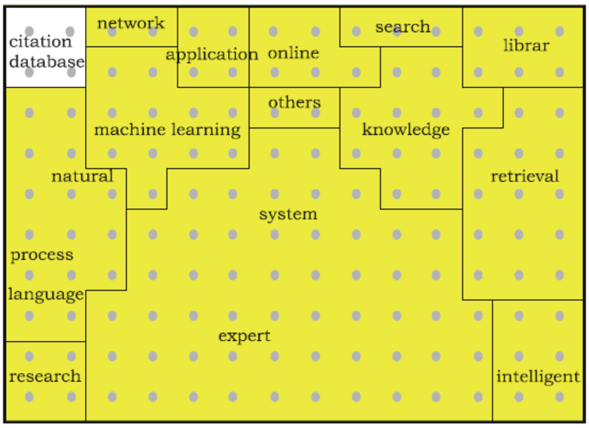


Fig. 4-9. Category map based on Kohonen SOM proposed in (Lin et al., 1991)

4.2.5.6. WEBSOM

The WEBSOM-project has applied the SOM to create a visual interface for browsing newsgroup collections (Honkela et al., 1996, 1997). In contrast to the Category Map which focuses on topical navigation, the WEBSOM system aims at visualising the overall similarity structure of document collections, based on a landscape metaphor. In this approach, the documents are encoded based on word categories rather than single terms. Words from the collection vocabulary are first clustered into word categories by using the SOM to map word vectors encoding the co-occurrence of a given word with other words, within a specified short context (Honkela, 1997). The assumption is that words appearing in similar short contexts will have similar meanings an can thus be replaced by a word category. The resulting word categories can then be used to encode large document spaces with a relatively small number of dimensions (in comparison to the whole vocabulary set).

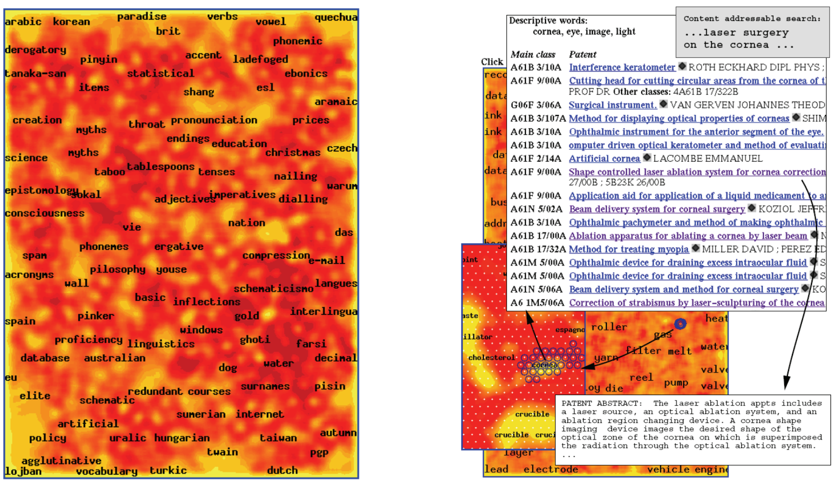


Fig. 4-10 WEBSOM map of sci.lang newsgroup (left, http://websom.hut.fi/websom/sci.lang-new/html/root.html) and interactive interface of the map of a patent database (right, Kohonen et al., 2000)

Since this method aims at mapping very large document sets, a SOM is then trained with a sample of the document set only and the remaining documents are positioned on the trained map by a best-matching neuron assignment. Due to the topology preservation properties of the SOM, similar documents will be

[17] This approach was later also applied to larger collections and full-text indexing (Lin, 1997).

mapped close to each other on the map and groups of nearby documents will likely be related to similar topics. The WEBSOM has been developed as an interface for explorative access to large document collections, such as newsgroups. An example of the map generated by this method is depicted in (Fig. 4-10, left). The visualisation model reflects distances between neighbouring nodes: nodes containing similar document sets are coloured in bright shades (e.g. grey or orange), whereas large distances (dissimilarities) are coloured in dark shades (of the same colour).

The interface accommodates point and click selection of documents (clicking on a node of the map grid presents a list of documents assigned to the node in question) and allows the user to zoom into selected areas of interest, at several zoom levels (Fig. 4-10, right). Later versions also accommodated visualisation of keyword search results (Kohonen et al., 2000). The main advantage of the WEBSOM method is its ability to scale to extremely large collections (e.g. several millions of documents, reported in ibid.).

4.2.5.7. DocMINER

DocMINER (Document Maps for Information Elicitation and Retrieval) is an interactive tool for generating, visualising and inspecting document maps presenting the overall semantic structure of medium-sized document collections. It has been developed by Andreas Becks and collaborators at the Computer Science Department, Information Systems Research Group of the RWTH Aachen. The main goal of development of DocMINER has been the evaluation of the application of document maps for specialized corpus analysis tasks in knowledge management (Becks, 2001).

The DocMINER system provides a modular framework for creation of document maps based on the Kohonen SOM. The calculation of inter-document similarity is based on a vector space model. After encoding the original vector set is scaled down to a lower dimension using MDS. The SOM is then used to create a document map based on such reduced dimensional representation of the original document space. The visualisation of the map is based on the “mountains and valleys” model following the landscape metaphor. The similarity information is displayed by colouring similar neighbouring nodes with bright shades while more distant nodes are coloured with corresponding dark shades of grey The interactive visual interface (Fig. 4-11) offers a range of interactive possibilities for exploring the structure of the document space that allow the inspection of document clusters at different granularity levels (zooming, scaling, sub-maps).

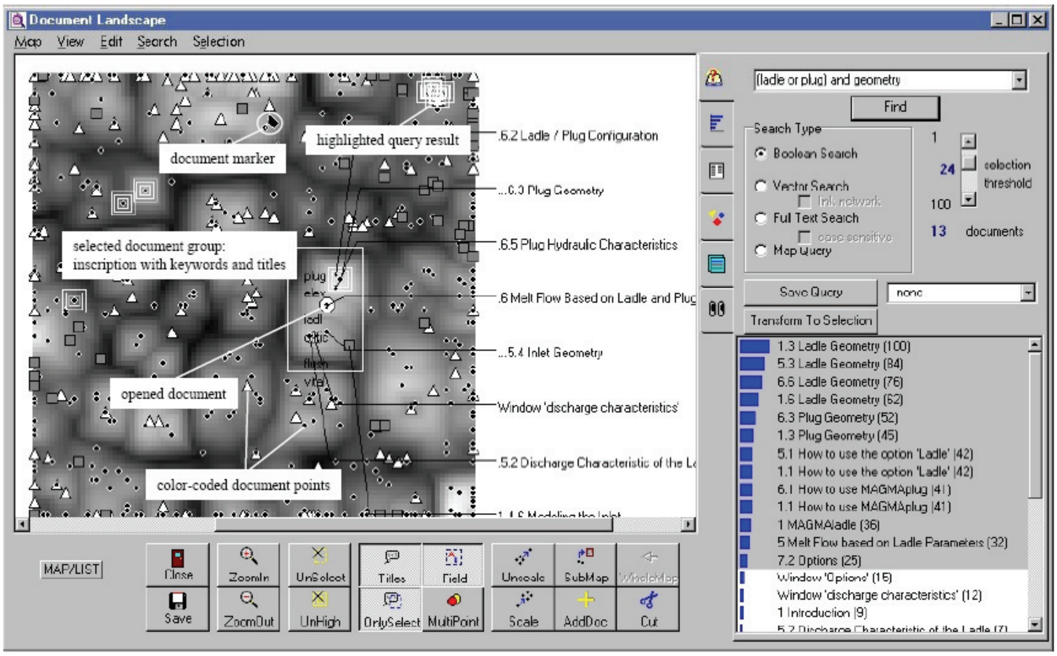


Fig. 4-11. DocMINER interface for visualisation and inspection of document maps based on Kohonen SOM (Becks, 2001).

Each document group is characterized by a set of most relevant key terms. The groups to be inspected can be defined arbitrarily and the characteristic key set is interactively determined by the system. The identification of relevant document groups is further supported by search queries, the possibility to assign documents to categories and the visualisation of important terms and sentences. Fig. 4-11 depicts the visual workspace of the DocMINER system. The document map is displayed in the left-window, the right window contains tabs with meta-information display for selected documents and groups of documents alongside with search queries and different selection possibilities. The tools for visual inspection of the map (such as zooming) are accommodated below the map visualisation.

4.2.5.8. Conversation Map

The Conversation Map system visualises the semantic structure of newsgroups archives based on statistical text-analysis of the messages in the archive (Sack, 2000). It has been conceived as a newsgroup browser and developed by Warren Sack at the MIT Media Lab. The system computes and visualises several different aspects of a newsgroup message space: 1) a social network of message authors based on co-citation analysis and analysis of links established by send-reply relationships, 2) a list of most frequently discussed themes and 3) a semantic network of main terms used in the discussions and the relationships between them. The java-based graphical interface (Fig. 4-12) allows the user to navigate through the message space by interacting with the visualisations and discover “who is talking to whom”, “what they are talking about” and “central terms…and metaphors of the conversation” (ibid.).

The social network is determined by co-citation and link-analysis methods such as those discussed in Section 4.1.2. The main themes and the semantic network of terms are determined using linguistic techniques such as part-of-speech tagging, morphological analysis and lexical cohesion (taking advantage of WordNet database to determine groups of related words). The graphical layout of the networks is based on a “spider web” algorithm, where initial elements (e.g. posts or terms) are placed in the middle, the subsequent elements are organized around it in concentric circles and related elements are connected by lines (ibid.).

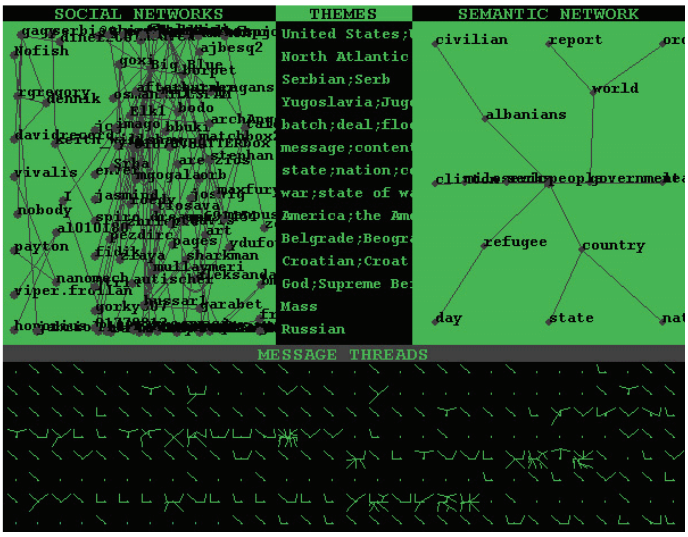


Fig. 4-12. The Conversation Map (Source: http://web.media.mit.edu/~lieber/IUI/Sack/Sack.html).

Although the Conversation Map is not a classical approach to visualising a semantic structure of a document space, it is closely related to our application context of visualising semantic overviews of community information spaces. In this case, the message archives represent rather general communities of interests where the purpose of communication is not the exchange of knowledge but social discourse, and which are thus not the main focus of our application (see Chapter 2). However, the map provides an interesting example of how the structure of a collaboratively created information space could be visualised also in the case of more knowledge exchange oriented situations.

4.2.5.9. Pathfinder Maps of Knowledge Domains

Another well-known approach to visualising structure of a knowledge domain reflected in its document space is based on graph-based techniques for extracting the salient structure of a collection based on either content-similarity of documents or author co-citation analysis (Chen, 1999a). To this end a generalized similarity analysis framework (GSA) was developed that includes a set of modelling and visualisation methods integrating content-based similarity, link-based similarity and usage-based similarity. The central element of this framework proposed by Chen (1997; 1999a) is the use of the pathfinder scaling technique (Schvaneveldt et al., 1989) for reducing redundant links in a graph to the most salient structure which can then be easily displayed by force-directed placement (Section 4.1.2).

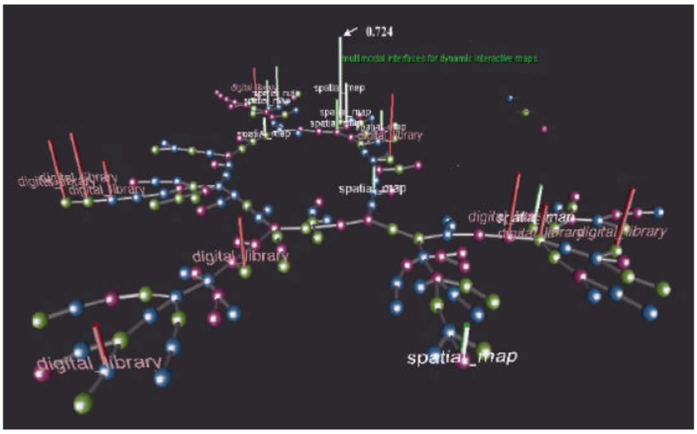


Fig. 4-13. Using Pathfinder scaling to visualise the semantic structure of the ACM SIGCHI collection (1995–1997). Source: (Chen, 1999a).

An example of using pathfinder scaling to visualise a graph representing inter-document similarity relationships based on content-analysis is given in Fig. 4-13. The inter-document similarity matrix is generated by LSI, based on document abstracts, titles and meta-information such as authors' affiliations and document keywords provided. The Pathfinder network scaling is applied to this similarity matrix and the result is rendered in a 3-dimensional space.

The visualised structure displays a central ring with a number of branches, each containing documents similar to each other with colour indicating the year of publication. The central ring is the only connection between the documents in different branches due to the applied pathfinder method for pruning redundant paths (only shortest paths between two documents are included). More frequently, this approach has been used to visualise author co-citation maps (Section 4.1.2) as a way of mapping the structure of a scientific domain. In this case, the inter-author similarity matrix is visualised by applying the same method (Fig. 4-14). The labels for topic areas (in blue) are determined from keyword-based author profiles.

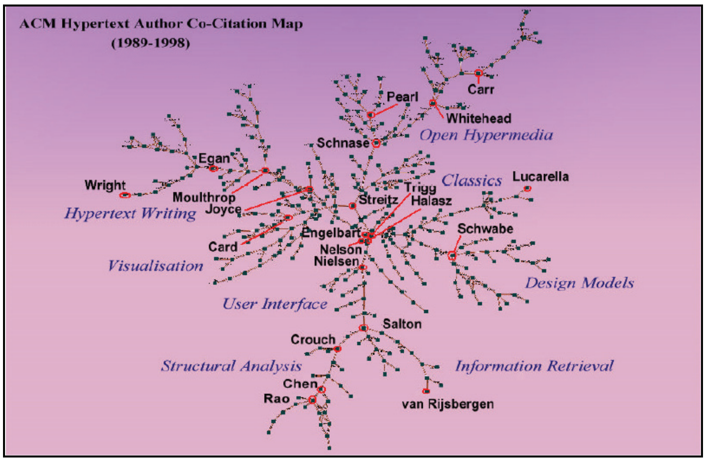


Fig. 4-14. Using Pathfinder scaling to visualise the author co-citation map of the ACM Hypertext collection (1989-1998). The topic labels have been manually added. Source: (Chen, 1999a).

4.2.6. Summative Discussion

A variety of existing methods and experiences with generating and visualising semantic overviews of unstructured document collections suggest this as a good point of departure for uncovering implicit knowledge structures contained in a community information space.

The visualisation model of document maps thereby seems the most promising one for supporting information access. Especially the use of self-organized maps has been well-documented and extensively studied and evaluated in a number of approaches (Lin et al., 1991; Honkela et al., 1997; Kohonen et al. 2000; Becks, 2001). It provides an intuitive visualisation metaphor that provides quick insight into structural poperties such as man topics, document clusters and relationships between them. The presented document map solutions mostly focus on specialized information analysis tasks (SPIRE, Starlight, DocMINER) or on supporting strategic decision making (VxINSIGHT). In contrast, the Category Map and WEBSOM were conceived as explorative information interfaces to large collections.

While the use of co-citation an alysis is limited only to special kinds of communities in which citation is a strongly present social convention, content-similarity analysis can be easily applied to most cases. Link-analysis could be reinterpreted in terms of context-analysis such that documents are considered related if they are placed by users in similar contexts (e.g. folders). This would resolve the problem of the lack of explicit links and enrich the semantic information otherwise not expressed by content-similarity alone. Another alternative are graph-based visualisations such as those proposed by Chen (1999a) based on the use of minimal-spanning trees and pathfinder networks for extracting salient semantic structure. They provide immediate insight into most important relationships between individual topics of a corpus and greatly simplify the graph structure thus reducing problems associated with standard force-directed graph visualisation techniques such as visual clutter for large and highly interconnected graphs and overview of relations limited to a given focus.

4.3. Eliciting and Visualising Knowledge from Human Users

Different techniques have been developed in order to elicit and visualise knowledge of human users. Since their goal is to make the otherwise invisible, implicit structures of knowledge of individuals or groups of people, they commonly combine both methods for eliciting and for visualising knowledge. This section summarizes the most important techniques such as concept mapping, repertory grid elicitation, cognitive mapping and social data mining.

4.3.1. Concept Mapping

Concept mapping is a technique for visual representation of personal knowledge in which subjects construct two-dimensional, spatially-organized representation of concepts and relationships between them. The technique has been originally developed in the field of education and learning research by Novak and Gowin (1984) as a means for assessing students' understanding and supporting their learning about an unfamiliar subject-matter or a domain of knowledge. While concept maps of experts aim at representing expert knowledge of a topic or knowledge domain, to be used as guide for others, concept maps created by learners themselves are cognitive tools which support learning through externalization and self-reflection.

4.3.1.1. Concept Map Models

Concept maps represent knowledge in form of a network structure containing nodes representing concepts, beliefs or ideas and directional named links representing relationships between them (Fig. 4-15).

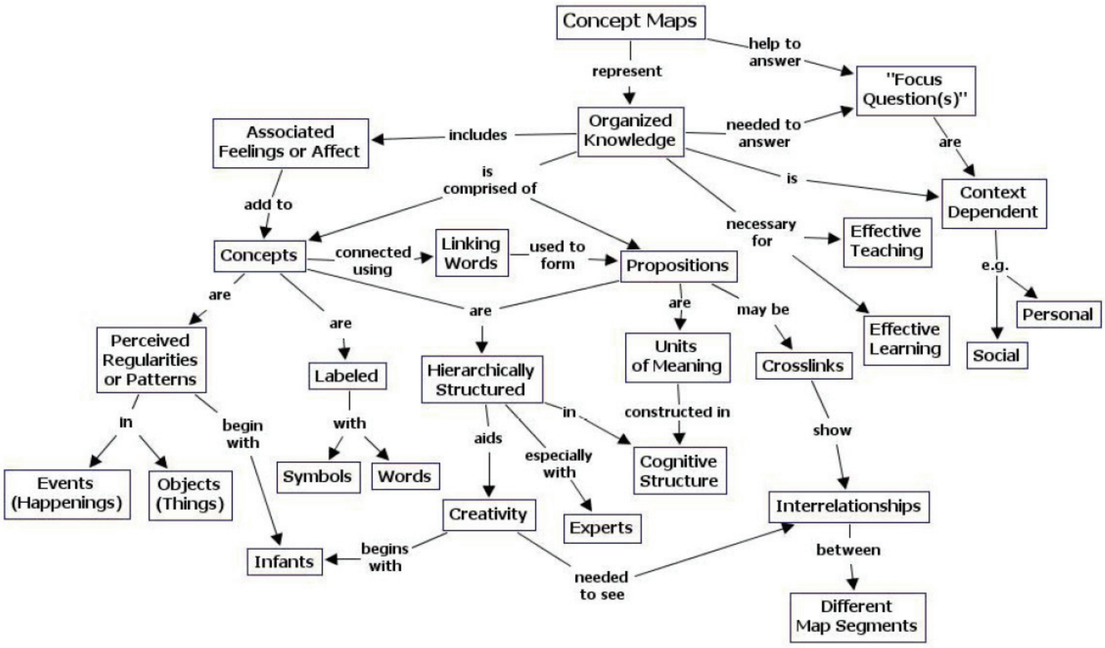


Fig. 4-15. A canonical concept map about concept maps according to Novak (depicted in Cañas et al., 2005).

Concepts are visually represented by boxes or circles and connected with lines, labelled with a descriptive characterization of the kind of relationship. Inter-concept relationships are expressed by verbs, hence forming propositional phrases between the connected concepts (e.g. "learning *affects* cognitive structure" in the map depicted by Fig. 4-15). In some approaches, the links are non-directional and unnamed (Ahlberg, 2004). By convention, the vertical arrangement of concepts expresses hierarchical relationships between them: more inclusive concepts are located at higher levels with more specific concepts arranged below them. Since multiple links between concepts commonly connect concepts in different regions of the map, concept maps exhibit a special property of integrating hierarchical and network forms of organization (Tergan, 2005).

Mind mapping is sometimes considered as an analogy to concept mapping. But researchers have been careful at pointing out the inherent differences of the two approaches (e.g. Jonassen et al., 1993). Two basic differences are that mind mapping is not concerned with the meaning of relationships between concepts and that it uses a hierarchical form of representation based on a central node with exploding branches as opposed to the combination of hierarchical and networked relationships of concept maps. Finally, mind mapping is based on a "brainstorming paradigm": it focuses on externalizing a broad range of issues and resources related to a given problem. This also differentiates concept mapping from the notion of "knowledge mapping" used in knowledge management, which aims at mapping the available cognitive and information resources related to a given issue or domain. In this sense, knowledge mapping is concerned with mapping meta-information i.e. structures of knowledge resources rather than knowledge structures itself (see discussion on knowledge maps in Section 4.4.4)

The aim of concept maps is much more specific: to elicit and represent conceptual structures which express an individual's personal knowledge and understanding of a given topic in a specific context (e.g. problem, task). The process of creating concept maps is seen as a learning process in which the "learner" creates new meaning out of processed information by externalizing the acquired knowledge in the form of concepts and propositional relationships between them. (Dansereau, 2005) distinguishes between three different types of concept maps with respect to the procedure of their creation and context of use: 1) information maps are concept maps produced by experts and made available to non-expert users or learners, 2) guide maps are structured templates guiding a learner in constructing an own map as a result of a learning process and 3) freestyle maps produced by a learner or user in order to express personal knowledge of a domain or problem at hand.

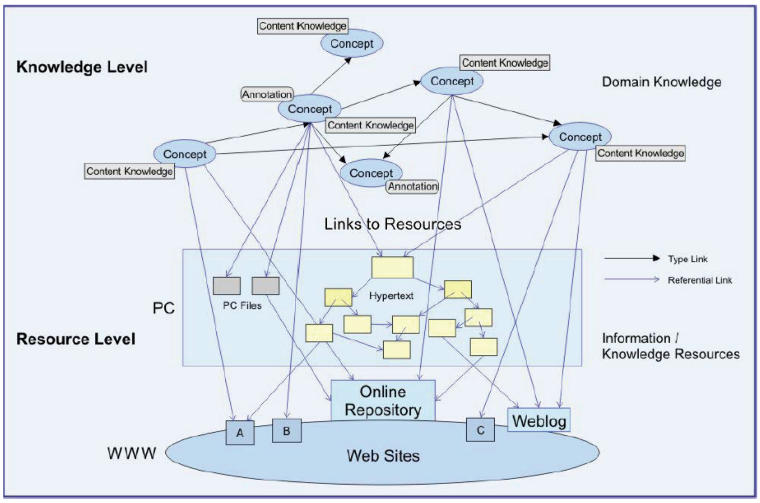


Fig. 4-16. Extended concept map model for representing and accessing domain knowledge and related information (Tergan, 2005).

Originally, concept mapping was developed as a pen-and-paper technique aimed at providing students with a tool to support their learning of new subject domains or problem-solving tasks. This was soon complemented by computer-supported concept mapping tools as well as by the exploration of the use of concept mapping in different domains. In particular, the use of computer-supported concept mapping tools emphasized the need to extend the representation of personal conceptual knowledge with content knowledge about a domain and knowledge about relevant information resources (Tergan, 2005).

In this view, three different kinds of knowledge that can be represented by or linked to a concept map include: concept knowledge, content knowledge and resource knowledge. While concept knowledge represents the conceptual macrostructure of a domain, content knowledge reflects the microstructure of an individual's declarative knowledge and is externalized different forms, such as personal notes, sketches, summaries or examples illustrating the meaning of concepts (ibid..). This is complemented by resource knowledge (Siemens, 2005) that refers to knowing about the location and content of task-relevant information as well as its relationship to the abstracted conceptual structure represented by the map.

According to this approach, in order to fully define a personal knowledge perspective on a subject-matter, concept maps should externalize the structure of all these three different kinds of individual knowledge, conceptual, content and resource knowledge (ibid.). Fig. 4-16 depicts such an extended concept map model. Thereby, since concept maps always focus on topic knowledge relevant for a specific task, accessing and locating information resources through concept maps promises to be much more effective than using information visualisation that visualises the overall semantic structure of an information collection.

4.3.1.2. Applications of Concept Mapping

The adequacy of concept maps in representing individual knowledge structures has been demonstrated in numerous empirical studies (e.g. Jonassen et al., 1997). The effectiveness of their application in supporting learning and knowledge acquisition has also been studied and confirmed in their original application areas such as self-regulated learning (Novak, 1990; Bruillard and Baron, 2000) and instruction support (Novak, 1990) as well as in collaborative work on distributed cognitive tasks and (Zhang & Norman, 1994; Cañas et al., 2001). The main rationale of using concept maps to support learning is that constructing concept maps fosters the construction of meaning as a learner makes an effort in formulating concepts and propositional relationships between them. Thus, the resulting propositional structure both reflects the learners understanding of a domain i.e. his personal knowledge and supports the process through which this knowledge is acquired. The availability of such knowledge representations constructed by domain experts has proved a useful means for capturing and sharing personal expert knowledge (Coffey et al., 2002, Cañas et al., 1999).

Newer approaches to concept mapping (following the extended model presented in Fig. 4-16) have been investigating the application of concept maps to supporting individual and cooperative knowledge management (e.g. Cañas et al., 1999; Tergan, 2005). In particular, Tergan (2005) has argued that such concept maps are well suited for supporting the entire spectrum of the knowledge management process blocks model (Probst et al., 1997, Chapter 3.2). The use of concept maps to support knowledge organisation and knowledge representation is the most obvious one, due their inherent nature of representing conceptual structures. Knowledge identification support is based on the possibility of using concept maps as navigational tools. Thereby, most existing approaches focus on the use of concept maps for supporting navigation in large data collections (e.g. Bruillard & Baron, 2000; Shen et al., 2003).

Using concept maps as navigational tools has proved particularly useful for supporting students in accessing large information collections of unfamiliar knowledge domains (Kommers & Lanzing, 1997; Lee, 2005). An application of concept maps to supporting web searches has also been presented in (Carvalho et al., 2001). This approach is based on taking a the concepts in a given concept map (or a user-selected subset of concepts) and links between them, as a context against which the documents retrieved by a web search can be matched and ranked. The ranking is based on the similarity comparison between propositional phrases represented by concept-link-concept combinations (see example in the previous section) and the phrases contained in the retrieved documents.

In contrast, although the use of concept maps as personal knowledge repositories and interfaces for identifying relevant expert knowledge has been proposed (Neumann et al., 2004) it has still been little investigated. One major problem is the observed high cognitive demand that self-constructed concept maps pose on the users, making privately constructed maps less rich than those intended to be shared with others (Cox, 1999). This questions the suitability of existing concept mapping approaches as tools for unobtrusively expressing personal knowledge structures as part of an information seeking process and using them to support the identification of relevant and generation of new knowledge.

The assumed usefulness of concept maps in supporting knowledge generation is based on the intrinsic correspondence between the processes of knowledge construction involving the expression, elaboration, organization, integration and linking of existing and new knowledge, with the cognitive activities taking place during concept map construction and use. Although studies in instructional settings confirm the adequacy of concept maps for facilitating knowledge generation (Bruillard & Baron, 2000) their specific application to typical knowledge management tasks has been relatively little investigated.

The suggested adequacy of concept maps to facilitate knowledge distribution is based on the analogy with instructional settings (knowledge communication between teacher and students) in which concept maps have been extensively studied (Tergan, 2005). Similarly, the collaborative construction of concept maps may contribute to the creation of common ground and establishment of shared knowledge between the participants, by visualising the domain knowledge shared by the group (Coffey et al., 2002; Fischer & Mandl, 2001). Finally, knowledge application

The use of concept mapping is often integrated into cooperative learning activities in which groups of participants collaboratively work on constructing a shared concept map (Chiu et al., 1999). While this largely occurs in physically co-located settings, sometimes systems for computer-supported collaborative mapping are used which allow multiple users to synchronously or asynchronously edit and extend a shared concept map.

Studies of collaborative mapping largely emphasise the effects of knowledge creation facilitated through discussions occurring between participants in the cooperative creation process (ibid.). In contrast, the use of the collaboratively created maps for knowledge exchange and navigation in unfamiliar knowledge domains by other users who didn't participate in the process, has been little investigated. Some studies report on learning effects in achieved by allowing different users to produced individual maps representing their personal perspectives and then exchange, discuss and collaboratively reflect on them with others (Gaines & Shaw, 1993). The latter also describe both the use of collaborative concept mapping for achieving consensus on a shared knowledge representation as well as situations in which the absence of consensus exhibited the need of providing collections of personal concept maps representing multiple perspectives on a knowledge domain (ibid..).

4.3.1.3. Concept Map Tools

A widespread suite of tools for the creation, use, management and sharing of concept maps is Cmap Tools. Cmap Tools is a computer-supported knowledge modelling and sharing environment developed at the Institute for Human and Machine Cognition (Cañas et al, 2004). It supports individual and collaborative construction of concept maps with linked media resources, as well as sharing and publishing maps for use by others.

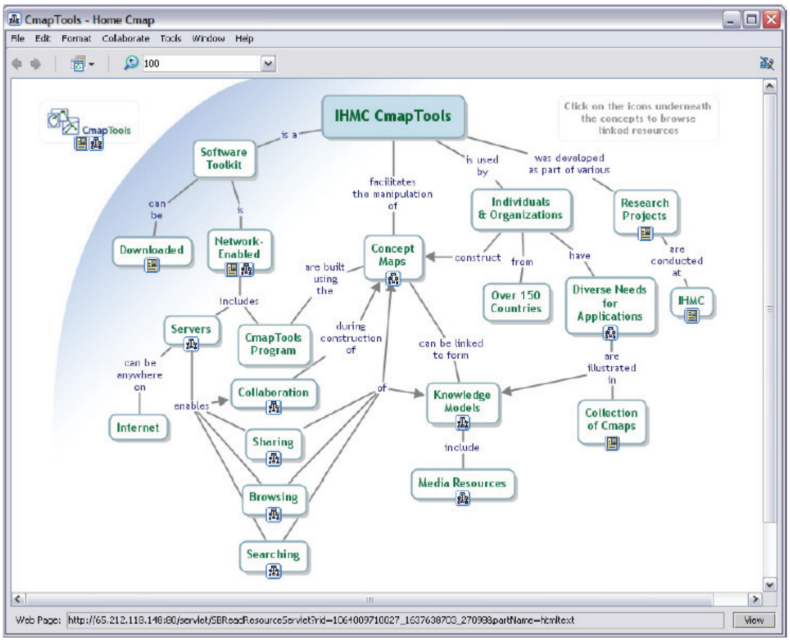


Fig. 4-17. User interface of Cmap Tools (Cañas et al., 2004b).

The basic functionalities of the user interface of CmapTools (Fig. 4-17) allow the user to create concept nodes and connect them with named links by simple point & click and drag & drop operations. The interface functionalities aim at being both simple enough for primary school students and naïve users with little technical skills, as well as powerful enough for the needs of expert knowledge engineering users.

One of the important design goals of Cmap Tools has been to enable the creation and use of concept maps not only as means of knowledge visualisation but also as visual interfaces for browsing a domain of knowledge and related information resources (Cañas et al., 1994). Besides standard functionalities needed for constructing concept maps, such as creating and naming concept nodes and links between them, each concept can be connected to multimedia information resources which further explain, illustrate or complement the information provided by the map.

Cmap Tools also supports the creation of representations of large domain of knowledge by allowing users to create linked collections of individual concept maps, each representing a specific aspect or topic of the domain. Such collections of concept maps and linked information resources representing a domain of knowledge are referred to as “knowledge models” (Fig. 4-18.) The collaborative use of concept maps is supported by a client-server architecture that allows users to access, publish and share concept maps located on different Cmap Tools servers on the Internet. Collaborative creation of maps is supported by asynchronous and synchronous annotation, discussion and editing of the same map by different users.

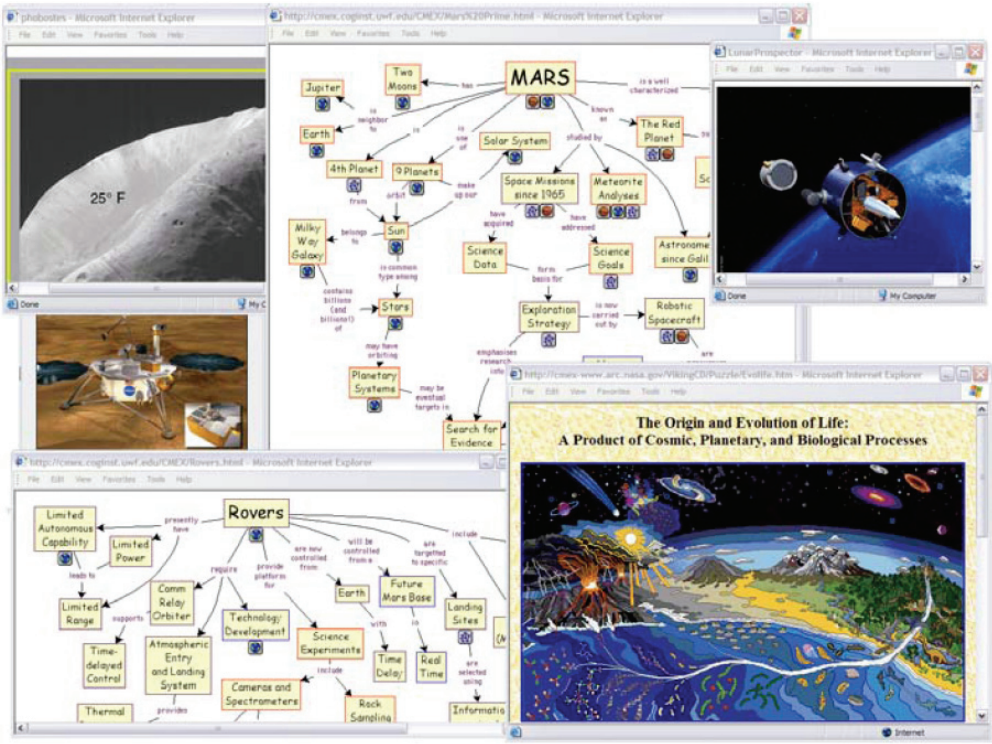


Fig. 4-18. A knowledge model of MARS exploration created by Cmap Tools in the NASA's Center for MARS Exploration (Cañas et al., 2004b).

Cmap Tools is a frequently used and referenced set of tools in concept mapping literature. Other similar tools offering different subsets of functionalities of Cmap Tools include Webster (Alpert & Grueneberg, 2001), Inspiration[18] (focusing on use by children and primary school students), SemNet[19] (outdated, one of historically first concept mapping tools), SMART Ideas[20] and Concept Systems[21].

4.3.2. Repetory Grid Elicitation

A closely related technique to concept mapping is the repertory grid method for elicitation of conceptual structures of individuals and groups of persons. Repertory grid elicitation has been originally developed by a psychologist George Kelly as part of his personal construct psychology theory (Kelly, 1955). The main assertion of this theory is that individuals interpret the world through personal conceptual models composed of a finite number of dichotomous concepts used to distinguish between similar and different elements in the world (Gaines & Shaw, 1992)[22].

Accordingly, the repertory grid method is based on asking a person to compare and contrast a set of examples along a set of dichotomous conceptual dimensions. The purpose is to build a conceptual structure without people having to directly formulate structures representing concepts and relationships between them. The assumption is that it is easier for a person to provide examples in a given domain and then describe how it would distinguish them by a set of conceptual properties, than having to explicitly create an abstract conceptual structure describing her personal view on a domain of knowledge from scratch.

[18] http://www.inspiration.com/

[19] http://trumpet.sdsu.edu/semnet.html

[20] http://www2.smarttech.com/st/en-US/Products/SMART+Ideas/

[21] http://conceptsystems.com

[22] Instead of concepts, the original term used is "constructs" in order to signify both a personal character of concepts as being actively "constructed" by different persons as well as the abstract (more general) "conceptual" properties of such personal constructs (ibid..).

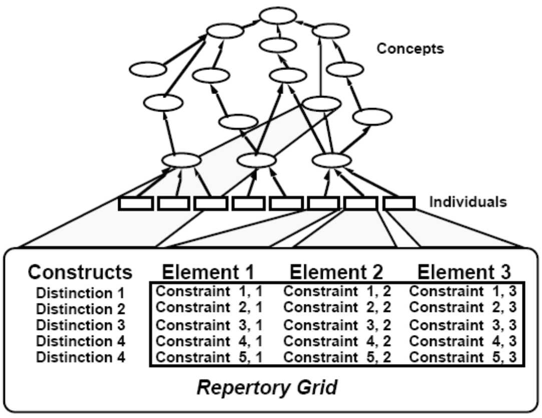


Fig. 4-19. Structure of the repertory grid (Gaines & Shaw, 1992)

Eliciting, visualising and comparing conceptual systems of individuals can then provide insights into the systems of meaning behind the reasoning and behaviour of a given person. The extent of similarity between conceptual systems of different individuals can be taken as an indication of the extent to which these individuals understand a specific context (situation, topic, knowledge domain) in similar ways. Based on the original paper-an-pen elicitation technique, a computer-based system for elicitation of repertory grids, their analysis and transformation into visual representation of personal and group conceptual structures has been developed by Shaw (Shaw, 1980; Gaines & Shaw, 1992).

4.3.2.1. The Repertory Grid Method

The basic form of a repertory grid is a matrix containing elements (rows) representing examples relevant in a specific thematic context and a set of constructs (columns) formulated by a person in order to describe different dimensions of similarity and difference between the elements (Fig. 4-20). The intersections of rows and columns represent a person's ratings of the extent to which a given element is related to a specific conceptual construct. The construction of such grids occurs in two phases: the knowledge elicitation phase and the grid rating phase.

In the knowledge elicitation phase the personal characterization of the example set relevant for a given application domain (e.g. topic, issue, situation or domain of knowledge) is obtained from the person whose knowledge structure is being elicited (e.g. a domain expert). To this end, the person is presented with sets of three examples at a time (the "triads") and asked to describe how two of them are similar to each other and different from the third by stating the most important attribute distinguishing the two most similar examples from the remaining one. In this way personal characterizations of the examples by dichotomous concepts ("bipolar constructs) are created (Kelly's triads method, see also Bradshaw et al., 1993).

Additional conceptual dimensions distinguishing the examples are obtained by repeating the process until the person has no more distinctions to make (or at a predefined maximum number of iterations). After the set of personal constructs has been created the person rates all examples in terms of how well they are characterized by each bi-polar construct. As a result the basic form of the repertory grid presenting examples, elicited personal constructs and user ratings is produced (Fig. 4-20).

Based on the grid results, a personal conceptual structure is generated and visualised by applying multivariate statistical analysis (e.g. principal component analysis) to produce a map showing relations between personal constructs and examples (Fig. 4-21, left) or by applying hierarchically clustering to determine groups of similar concepts and similar domain example elements (Fig. 4-21, right). The latter is used also to generate hierarchical relationships between the user-defined concepts elicited by the grid. Based on the grid data implicit relationships between concepts and rules can also be induced and represented in the form of a conceptual graph.

		1	2	3	4	5	6	7	8	9	10	11	12		
qualitative and quantitative	1	8	8	6	9	1	9	8	8	9	4	4	4	1	quantitative
local	2	9	9	6	3	5	1	9	9	1	4	4	4	2	global
autocorrelation not considered	3	1	1	5	4	7	2	3	1	2	9	9	9	3	autocorrelation considered
doesn't honour data points	4	2	1	3	5	9	9	1	1	9	8	7	7	4	honours data points
multiple variables considered	5	9	9	9	9	9	9	9	1	9	9	9	9	5	usually one variable considered
mathematical curve fitting	6	4	4	8	9	9	1	2	4	9	5	8	8	6	doesn't fit a mathematical curve
nonparametric	7	9	8	6	1	1	3	6	8	3	8	8	1	7	parametric
interval or ratio data	8	9	1	1	5	5	1	1	1	1	1	1	1	8	nominal data
requires periodicities	9	6	6	9	9	9	9	1	6	9	9	9	9	9	doesn't require periodicities
doesn't fit a trend	10	9	9	1	1	8	1	7	9	1	6	1	1	10	fits a trend to the data
heavy computing load	11	7	6	7	8	9	4	4	5	3	1	2	3	11	no computing load
assumes isotropic surface	12	1	1	4	3	8	9	1	1	9	7	6	6	12	assumes anisotropic surface
not as susceptible to clusters	13	8	8	6	9	3	4	7	8	5	1	2	2	13	estimates susceptible to clusters
doesn't incorporate geologic model	14	2	2	3	1	9	1	2	2	1	6	6	6	14	incorporates geologic model
interpretive	15	9	9	5	9	1	7	9	9	7	3	4	4	15	representative
not very important	16	4	5	2	1	9	1	1	6	7	8	9	9	16	very important
not very effective	17	4	5	3	1	9	2	4	6	7	9	8	8	17	very effective
not widely used	18	3	8	7	5	9	3	3	4	6	2	2	1	18	widely used
		1	2	3	4	5	6	7	8	9	10	11	12		
		Probability mapping	Trend surface analysis	Distance weighted averaging	Proximal mapping	Hand contouring	Bicubic splines	Double Fourier series	Most predictable surface	Triangulation	Universal kriging	Punctual kriging	Nonparametric kriging		

Fig. 4-20. A repertory grid representing an experts view on spatial mapping technologies (Gaines & Shaw, 1992). Columns are specific entities relevant for a domain, rows are user-defined personal constructs and values are user ratings of the degree to which a given entity can be characterized with a specific construct.

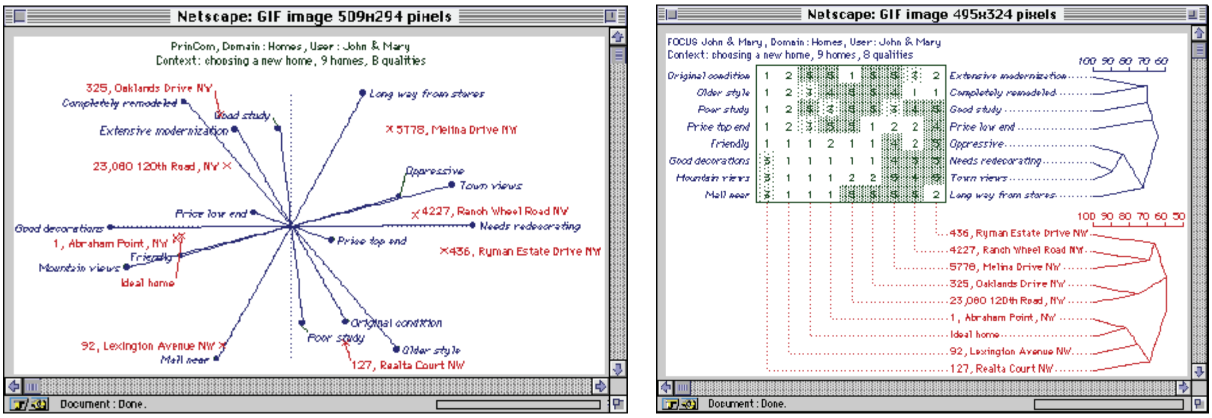


Fig. 4-21. Visualisation of the conceptual structures elicited by the repertory grid: relationships between concepts and example entities (left), clustering of similar concepts and of similar examples (right). Source: (Shaw & Gaines, 1996)

The elicitation of group structures follows the same principle extended with an additional phase for identifying and resolving conflicts. After a set of entities relevant for a given application context is defined by the group of experts, each expert elicits personal constructs on the entities following the repertory grid method. The resulting personal grids are then exchanged between the experts and each expert scores example entities based on those attributes of others that he agrees with. The result provides the basis for identifying identifying relationships between concepts of individual users and merging them into a shared structure by differentiating between consensus, conflict, correspondence and contrast relationships (Fig. 4-22).

		Terminology: Same	Terminology: Different
Attributes	Same	***Consensus*** Experts use terminology and concepts in the same way	***Correspondence*** Experts use different terminology for the same concepts
	Different	***Conflict*** Experts use same terminology for different concepts	***Contrast*** Experts differ in terminology and concepts

Fig. 4-22. Classes of relationships between concepts in different conceptual systems (Shaw & Gaines, 1992).

4.3.2.2. Applications of Repertory Grids

While the original application of this technique stems from psychotherapy (Kelly, 1955) the computer-based method has been largely applied in education, management and the development of knowledge-based expert systems (Gaines & Shaw, 1993a; Bradshaw et al., 1993). The repertory grid method has also been frequently used as a means for eliciting and comparing individual conceptual views on a given issue or domain of knowledge in situations requiring collaboration between different participants. For example, to reveal the differences of subjective judgments of managers on specific company policies or to reveal and concile terminological differences in scientific collaboration in groups of scientists (Shaw, 1980).

Related applications include the facilitation of matchmaking of similar experts in social networks based on the ability of one expert to understand the concepts of another (Shaw & Gaines, 1991) and the visualisation of terminological structures of scientific communities to support participation of newcomers in online community exchanges such as forums and newsgroups (Gaines & Shaw, 1994; Shaw & Gains, 1999). The overview presented in (Bradshaw et al., 1993) further mentions the use of knowledge representation in repertory grids as a foundation for knowledge-databases (based on the correspondence to the entity-relationship model) and for supporting the creation of database interfaces. Finally, a spatial emphasis has been given to the use of grid techniques for supporting decision-making under high uncertainty and complex trade-off situations in commercial organization (ibid.).

4.3.2.3. Example Systems

The most wide-spread repertory grid system is WebGrid and its successors WebGrid-II and WebGrid-III[23], developed by Shaw & Gaines at the Knowledge Science Institute in Calgary (Gaines & Shaw, 1996; Shaw & Gaines, 1998; 1999). This is a web-based implementation of an older system, Knowledge Support System Zero (Gaines & Shaw, 1992; Fig. 4-23) realizing the repertory grid technique and a set of additional tools for visualising the grid results (principal component maps, hierarchical clusters, concept graphs), for comparing different personal grids and merging them into a shared structure.

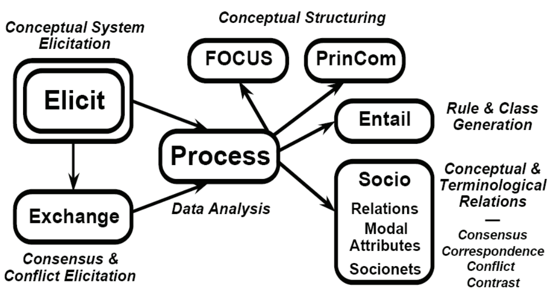


Fig. 4-23. Functional blocks of the repertory grid elicitation system Knowledge Support System Zero providing the basis of the newer WebGrid-II system. Source: (Gaines & Shaw, 1992).

[23] http://repgrid.com/RepIV/

In addition WebGrid-II & III support integration with multimedia information resources and include an inference engine for case-based reasoning. A server for public use of system functions allows the sharing of grids in communities (RepGrid-Net).

4.3.3. Cognitive Mapping

Cognitive mapping is a well known technique for eliciting and visualising implicit knowledge structures of individuals and groups of users. Cognitive maps are a form of visual representation of a subjective belief structures originally developed by political scientists to represent belief and argument structures of decision makers (Wellman, 1994). To this end cognitive maps are modelled as directed graphs in which nodes represent concepts and edges causal relationships between concepts, often signed to express the positivity of negativity of the relationship (ibid.). This very structured model has been largely used in political science (ibid..). In this field, cognitive maps have often been constructed by scientists by manual analysis from verbal protocols and decision documents in order to capture the structure of a decision problem and the relationships among different arguments related to a given policy.

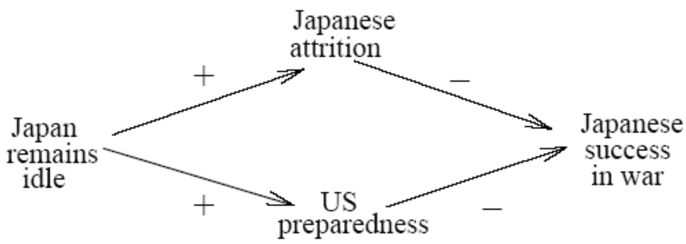


Fig. 4-24 Example of a simple cognitive map (Wellman, 1994).

4.3.3.1. Methods of Cognitive Mapping

Cognitive mapping has also been taken on in organizational studies, management and marketing where it has been used for a variety of purposes such as uncovering implicit beliefs behind managerial decisions, identifying organizational routines or mapping the belief structures of customer groups or customer-segmentation with respect to product preferences (see Huff & Jenkins, 2002 for an overview). Here different techniques for eliciting cognitive maps from direct user feedback have been used and developed.

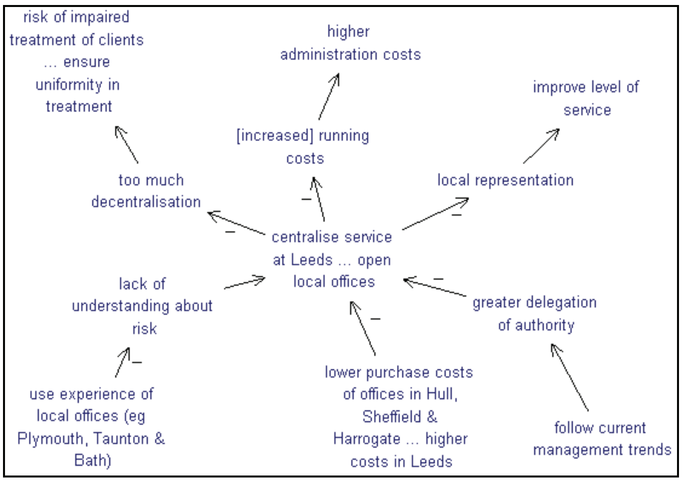


Fig. 4-25. Example of a cognitive map for a decision-making process (Ackermann et al., 1996).

One set of techniques is based on a process similar to that of concept mapping where a person or a group of persons construct conceptual structures through an interview or discussion mediated by a special facilitator. First, a set of conceptual entities relevant for a topic, problem or issue at hand is elicited in a brainstorm fashion, the concepts are then clustered according to similarity and causal relationships are established between them through a mediated discussion and iterative analysis of the structures elicited so far. An example of such kind of cognitive mapping is the oval mapping technique (Bryson et al., 1995; Eden & Ackermann, 1998) often also reffered to as knowledge mapping and frequently used for decision-making and strategy making support. The technique is performed either in the pen-and-paper version or by using computer-supported tools (Eden & Ackermann, 2002).

Another form of cognitive mapping is based on obtaining user responses to a set of selected stimuli which represent concrete examples of concepts to which user attitudes are to be uncovered. User responses are analysed by statistical techniques such as multi-dimensional scaling or hierarchical clustering in order to determine groups of related stimuli from a person's point of view as well as relationships between attributes established by users choices. The results are commonly visualised in form of scatter plots, cluster maps and graphs.

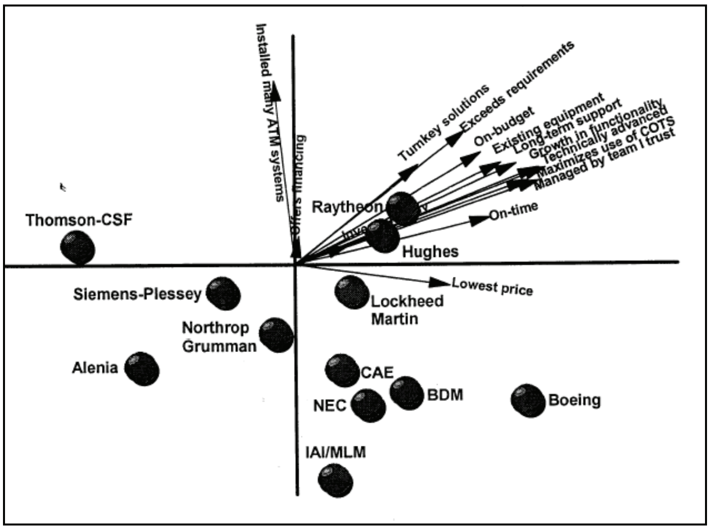


Fig. 4-26 Example of perceptual map displaying relationships between different companies and their product attributes. (Wittenschlaeger & Fiedler, 1997)

Instead of asking users about similarities between stimuli directly they are asked to rate them on a specific set of attributes. The similarities are then inferred from users' ratings. In this way, any set of subject can be rated on a set of attributes and then mapped to show the relative position and relationships between subjects as well as the relationships between the subjects and the characteristic attributes. The validity of the resulting maps thereby depends both on the set of stimuli and the predefined set of attributes rated by the users.

The theoretical basis of this method is the assumption that the obtained similarity data about the user-perceived relationships between a set of stimuli reflects the organisation of these concepts in user's memory (e.g. Fillenbaum and Rapoport, 1971). An example is the elicitation of relationships between declared features of a product (e.g. functionality aspects) and product preferences expressed by a group of people (e.g. style, personal status, perceived utility). For example, the stimuli represent products and ranking dimensions are the perceived and declared product characteristics. The stimuli can also be represented by a problem statement and the ranking dimensions by possible actions or issues related to the problem, which are elicited in a group session. Since it gathers users' subjective perceptions in order to identify implicit belief structures on a given set of concepts or properties, this technique is also referred to as perceptual mapping (Wittenschlaeger & Fiedler, 1997).

4.3.3.2. Applications

Causal cognitive mapping (Fig. 4-24, Fig. 4-25) is used in scientific research in political sciences and organizational studies with a descriptive purpose. A lot of attention has also received its application in organizational and managerial studies as a means of facilitating collaborative problem solving or mapping implicit assumptions and supporting strategy development and decision-making (see Huff & Jenkins, 2002 for an overview).

Perceptual mapping has been used as a tool for pro-actively supporting strategic management since its inception. Typical examples include the discovery and visualisation of relationships between marketplace competitors, relationships between different product lines or the customers purchase decision-making criteria (Wittenschlaeger & Fiedler, 1997). The obtained insights can then be used to support market segmentation, the development and evaluation of new product concepts or to track changes in marketplace perceptions, among others (ibid.).

Techniques of cognitive mapping have been also applied to support the design of information systems and user requirements capture, such as the information architecture of tourism guide maps (Lokuge et al., 1996) or of online grocery shopping systems (Pallant et al., 1996). The underlying rationale is that organising information in a way compatible with the structure of users' mental models may enhance information retrieval through more effective navigation. In particular, the use of perceptual mapping methods such as card-sorting have become a part of standard repertoire of methods for capturing both specific information architecture and more general interface design requirements in usability engineering (Rosson & Carroll, 2001).

4.3.3.3. Tools for Cognitive Mapping

Different cognitive mapping tools exist providing the functionalities for creating causal cognitive maps manually or for constructing perceptual maps by eliciting user feedback on a presented set of stimuli. Examples include the Banxia Decision Explorer (http://www.banxia.com/) for creating and analysing causal cognitive maps, the Sensor tool for story-based strategic analysis and customer segmentation (Rughase, 2002) or the Fountain Park Weak Signals Toolset (http://fountainpark.com) for collaborative elicitation and discovery of ideas and tacit knowledge in strategy making. An overview in the field of management studies can also be found in (Huff & Jenkins, 2002).

4.3.4. Social Data Mining

Social data mining (Terveen et al., 2001) focuses on elicitation of knowledge of human users based on analysis of their usage and interaction with information. Rather than visualising knowledge structures social data mining aims at enabling people to share knowledge by "using" each other's experiences as a basis for identifying relevant information items for a specific information need. Thereby the information need can be expressed explicitly by a search query or topical selection by the user or it can be inferred from a user profile or task context (specified manually or determined from the results of user's previous actions and information selections). To achieve this social data mining techniques analyse histories of information use and social activity such as newsgroup messages, web usage logs, bookmarks, citations or hyperlinks (Amento et al., 2003).

The main idea behind such approaches can be related to the metaphor of "social navigation" (Munro et al. 1999): making available the "paths through the woods" followed by other users is a means of supporting navigation through information collection based on experiences of others. Social data mining analysis then uses computational analysis of accounts of user information activities to make such paths visible or usable for information access. Since they exploit user preference information reflected in records of existing activities (e.g. bookmarking), in contrast to collaborative filtering systems based on explicit user ratings of information or personal preferences (e.g. Resnick et al. 1994; Shardanand & Maes 1995) social datamining systems do not require users to accomplish any additional activities (Amento et al., 2003) and hence don't distract them from their primary task.

Most prominent examples of such approaches include text-analysis of online conversations such as usenet newsgroups. A concise but informative overview can be found in (ibid.). For example, the PHOAKS system (Hill and Terveen 1996) identifies mentions of web pages in newsgroup messages and categorizes them into groups with ranked lists of most frequently mention pages in each group. Another line of approaches analyses newsgroups or chats in order to create visualisations of thread structures or conversation profiles of individual users based on properties such as frequency of posting or topical correspondence (Smith, 1999; Xiong & Donath, 1999; Fig. 4-27). Such visualisations can then be used to identify users with similar interests or high engagement in specific topics. Well-known examples of web-history analysis and visualisation of resulting navigational patterns (e.g. clusters of potentially related documents) include the Footprints system (Wexelblat & Maes 1999) and the activity path model proposed by (Chalmers et al., 1998). The latter allows recommendation of web pages browsed by others based on similarity between navigational paths represented as sequences of visited pages and/or documents.

Fig. 4-27 People Garden newsgroup visualisation (Xiong & Donath, 1999). Each flower shows messages of a given user, indicating also time, intensity of replies and length of time a user has spent in the newsgroup.

Finally, a number of approaches investigated the categorization of web pages according to similarity based on hyperlink-analysis and co-citation analysis[24] such as (Pirolli et al., 1996; Kleinberg, 1998). Probably the most famous approach is the PageRank algorithm developed by the founders of the Google search engine (Page et al. 2002). In this method, documents which are linked to by other documents with high ranking scores will receive a higher rank score themselves. A hybrid approach has been successfully demonstrated by the Topic Shop System which combines content-similarity analysis and link-analysis to determine a set of web pages similar to an initial set of pages the user has visited so far (Amento et al., 2003). A particular characteristic of this system is a visual information exploration interface for presenting the results of the underlying system. This feature also classifies it into the category of information workspaces that use visual information presentation to support user information seeking activities. We consider these aspects in a separate part in Section 4.5.

4.3.5. Summative Discussion

The strength of concept mapping is its ability to elicit and visualising personal knowledge of human users. Rather than a general conceptual structure, concept maps successfully represent personal knowledge about a given topic in a specific context. Collaborative use of concept maps has been less well studied and supported to a limited extent. Current research challanges include the process of collaborative creation (one shared map vs. multiple maps) and the identification and retrieval of maps relevant for the needs of a given task.

On the other hand, creation of concept maps is a time consuming process requiring significant cognitive effort of the users and the willingness to explicitly engage into it. It has been reported that motivation for creating private maps is low, resulting in less rich conceptual structures than when maps are produced with the purpose of being shared with others (Cox, 1999). Rather than abstracted conceptual structures, for practical concerns in knowledge sharing linking relevant information resources to concepts has become of great importance – e.g. when using concept maps to support the identification of relevant information for a task. Similar remarks apply to cognitive maps and the repertory grid elicitation: the construction of both requires explicit effort on behalf of users, not related to their primary task. The resulting maps provide static artefacts which have been largely used as means for individual reflection, group discussion facilitation and decision-making support rather than for supporting knowledge exchange mediated through information access.

[24] See Section 4.2.2 on basic techniques of co-citation and link analysis.

At the same time, social data mining methods suggest ways for eliciting semantic profiles of human users from their information seeking actions and applying them to pro-actively support the identification of relevant information. This is also an implicit way of supporting knowledge exchange between different users: recommendations of information items are based on the elicited knowledge models (user profiles) both of the knowledge seeker(s) and of information authors. However, the existing methods do not provide ways for explicitly visualising the elicited knowledge structures. This places them as a borderline class of approach between knowledge elicitation, knowledge visualisation and information seeking support interfaces.

The above analysis suggests that the most promising point of departure in our context could be the combination of social data mining as elicitation technique and concept maps as knowledge visualisation model. In order to make concept maps more usable for knowledge exchange and contextualized access to information resources, the complexity of their creation process needs to be lowered. As emphasized in (Zhang, 1997), the cost of cognitive effort involved in creating concept maps must be superseded by the benefits of their use. Furthermore, the modalities in which concept maps could be applied not only for individual but also for collaborative visualisation and sharing of knowledge should be developed. Currently, manual elicitation of collaborative structures is made possible by some tools but applications of collaborative knowledge elicitation for supporting specific knowledge management tasks have been relatively little addressed.

Accordingly, an important challenge in our problem context will be the development of methods allowing unobtrusive creation of lightweight concept maps as a natural element of user's information seeking and knowledge construction activities, non-distracting from the user's primary task, possibly based on the patterns of usage and interaction with information.

4.4. Knowledge Organisation and Representation

A number of different methods for knowledge organisation and representation have been developed in fields such as knowledge-based systems, artificial intelligence and knowledge management. This section gives a brief overview of the main approaches in the context of knowledge management and information access, since they are most related to our problem context. This includes thesauri, ontologies, semantic nets, topic maps and knowledge maps.

4.4.1. Thesauri

The simplest models of knowledge organization include classification schemes, taxonomies and categorization schemes. Though these terms are often used interchangeably there are differences between them (Hodge, 2000). Classification schemes define ways of organizing information into different topical categories based on well-defined criteria and classification structure (e.g. Dewey Classification Scheme). Taxonomies may be less-structured and provide groupings of objects into often broad topical categories. Both may include hierarchical structuring of topics. The basic class of formally defined knowledge organization systems with richer expressive power and structures of relationships between concepts are thesauri (ibid.).

In contrast to categorization schemes and taxonomies that focus on assignment of objects to concepts, thesauri focus on mapping the space of available concepts and relationships between them. To this end thesauri are formed as collections of terms, concepts or phrases with relationships between them and descriptions of their meaning. Their goal is to represent as completely and specifically as possible a vocabulary used in a specific domain. To this end the concepts are comprised either of single or multiple terms carrying significant and well-defined meaning with respect to a domain.

An important aspect of the thesauri is the specification of different kinds of relationships between individual concepts. The basic relationships commonly expressed in a thesaurus include: hierarchy, equivalence and associative or related (Soergel, 2000). The hierarchy relationship is used to represent broader term or narrower term relationships. The equivalence relationships indicates synonyms whereas near-synonyms and related concepts are connected by the associative or related term relationship (ibid.). In knowledge management and information retrieval applications the natural language descriptions of the meaning of concepts may be sometimes omitted. Whereas in everyday contexts general thesauri exist (e.g. Roget's "Thesaurus of English Words and Phrases") in knowledge management thesauri are largely constructed for a specific subject or application domain.

The main goal of thesauri is to support the authors in the knowledge creation process and mediators in the knowledge organization processes by ensuring both the versatility of choice as well as the precision of use of terms and concepts. In connection with a categorization scheme this also supports better user access to previously indexed information and knowledge resources. Two main kind of access to thesauri include alphabetical lists with descriptors (concepts or phrases) and indexes which list single terms with references to complex concepts in which they are used. Thesauri are commonly created by hand and in a collaborative effort of knowledge organization specialists and domain experts (ibid.).

4.4.2. Ontologies

The term ontology has been borrowed from philosophy where it refers to a fundamental branch of metaphysics concerned with the study of "being" or existence and its properties (e.g. categories, entities and types of being). In computer science the notion of ontology is most commonly defined as "a formal explicit specification of a shared conceptualization" (Gruber, 1993). Often, the notion of a formal and explicit specification of a shared conceptualization is referred to represent a shared understanding of a domain of knowledge or interest (e.g. Uschold & Gruninger 1996). Accordingly, in more descriptive terms ontologies can be described as formal semantic models of concepts and named relationships between them expressed in a machine-readable form which explicitly describe how a given individual or a group of people understands a particular domain of knowledge.

4.4.2.1. Ontology Modelling and Representation

The main constructs used to model an ontology include concept classes, relations (also referred to as properties), axioms (logical rules) and instances (Gruber, 1993). An ontology model organizes concepts are into a hierachical taxonomy with inheritance mechanisms and well-defined relationships and constraints between the concepts as well as logical rules allowing automatic logical inferencing. Several different representation languages for implementing abstract ontological models into a form usable by concrete software systems exist, largely based on predicate-logic such s CYCL (Lenat & Guah, 1990) , KIF – Knolwedge Interchange Format (Genesereth & Fikes, 1992), Ontolingua (Gruber, 1993) and OIL – Ontology Inference Layer (Fensel et al., 2000)[25]. Within the Semantic Web project a standardization framework for XML-based ontology representation languages has been developed based on the RDF metadata description standard[26] (Resource Description Framework) and OWL - web ontology languages specification[27].

RDF allows the expression of concepts and semantic relationships between them. Relations are expressed in form of (subject, object, predicate) tupples that can be represented in a graph form by translating subjects and objects to nodes, whereas the predicates become directed vertices. The extension to RDF Schema (RDFS) provides the basic vocabulary for the construction of classes (concepts), their hierarchical organization and association of instances to classes with the property that classes can also be instances of other classes. Class relationships are defined as property subclasses with scope restriction by means of domain and range specification. OWL (McGuiness & Van Harmelen, 2003)[28] is a standardization of the DAML+OIL language which allows the distinction between class relations (object

[25] For a comparative discussion of characteristics of individual ontology languages see (Kunz, 2005).
[26] http://www.w3.org/RDF
[27] http://www.w3.org/TR/**owl**-ref/
[28] http://www.w3.org/TR/owl-features/

properties) and attributes (data type properties), supports transitive, inverse and equivalence relation declarations as well as union and intersection. OWL is organized into several different abstraction levels allowing different complexities of user. OWL Lite allows the realization of applications with simple classification without complex inference mechanisms. OWL DL allows formal inferencing through descriptive logic systems whereas OWL Full further supports meta-modelling and full RDF vocabulary. The inferencing is based on axioms, if-then expressions cannot be modelled.

Depending on the abstraction level modelled by a given ontology, several different ontology types can be distinguished (Guarino, 1998): top-level or generic ontologies, domain ontologies, task ontologies and application ontologies. Generic ontologies (also referred as foundational ontologies) describe general concepts such as space, time or events independently of a particular domain and are thus applicable across different domains. Domain ontologies describe the knowledge structure of a specific domain whereas task ontologies describe a domain-independent but task-specific vocabulary as a basis for describing task types and related problem-solving concepts (Studer et al., 1996). The most specialized ontologies are application ontologies whose concepts most closely model specific entities and their roles for a concrete application purpose in a specific domain. An example of an ontology model is given in Fig. 4-28.

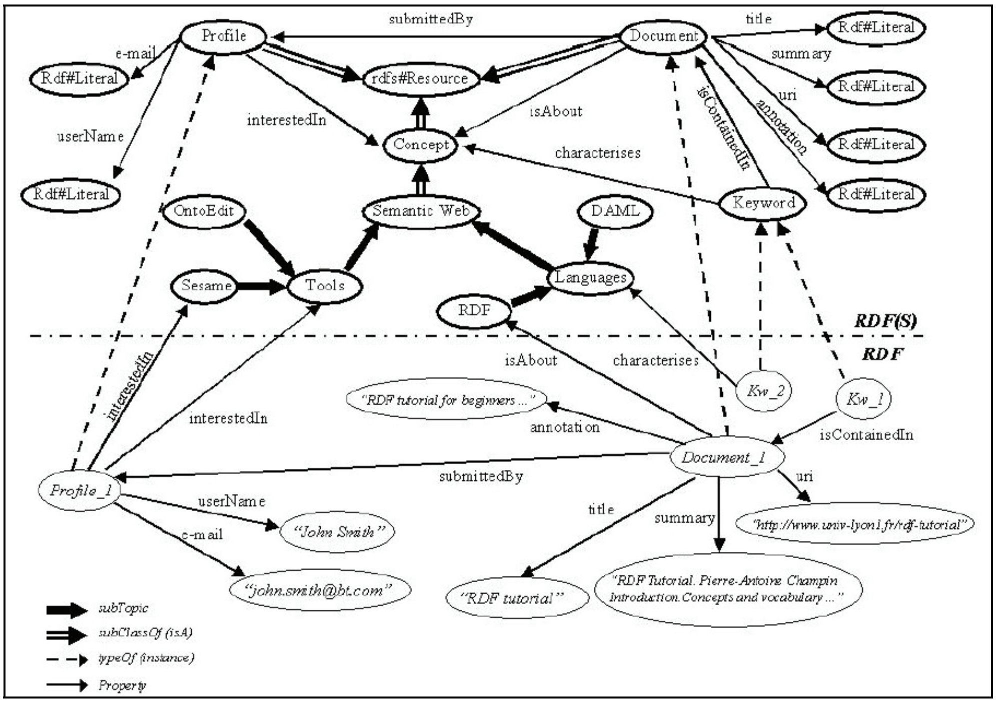


Fig. 4-28. Example of a simple ontology model in RDFS (Davies et al., 2003)

4.4.2.2. Ontology Design and Generation

Building and maintaining ontologies are time-consuming and difficult tasks addressed by ontological engineering, a specialized scientific field in its own right (Fernandez et al., 1997). It is concerned with issues ranging from ontology design and evaluation to ontology deployment, maintenance, integration and sharing. Such a wide spectrum of tasks that need to be supported as part of an ontology design and use life-cycle shows that the development and application of ontologies is a complex undertaking. This complexity is one of the main reasons why ontologies have limited applicability in our problem context.

One of the main challenges in building ontologies for a specific domain or application is the need to achieve consensus across a community of people concerned with the domain in question. More likely than not, different members of a community (especially domain experts) will have very different views on the conceptualization of domain knowledge. In order to deal with this problem different strategies are employed. (Gruninger & Lee, 2002) differentiate between two principal classes of solutions. One approach is to develop small and simple personal ontologies by a large number of people which are then merged into a common one. Another, more often practiced solution is the development of standardized formal ontologies by consortia or standards organizations. While the former requires sophisticated methods and effort for ontology mapping and merging (including conflict resolution) in the latter case well-defined coordination mechanisms and support for collaborative ontology design are needed. Such shortcomings of these two approaches provide the main reasons of their limited applicability to our problem context.

Another problem in ontology design is that due to the complexity of the task ontology designers are often not the end-users. Accordingly, the effective construction of shareable and easily reusable knowledge bases by means of ontologies is constrained by the problem that users and designers may do not always share the same assumptions about the actual meaning of concepts and descriptors used to characterize them, the implemented hierarchical structure etc (ibid.).

This problem further limits the applicability of ontology use both in intra-community contexts and even more so in the cross-community application contexts. In the former case a diversity of views in a community exists and gaining insight into implicit structures of community knowledge is subject to extensive community participation. Thus designers would have to "become" community members. In the latter case, the difference of meanings between different communities is the basic problem that needs to be overcome. Hence, the designers would either need to become expert members in both communities or laborious negotiation of meaning and conflict resolution in collaborative heterogeneous design sessions involving groups from different communities would need to occur.

Another basic problem of ontology design is also manual generation. The four main phases of ontology building include problem analysis and requirements specification, conceptual modelling, formalization and formulation of axioms and evaluation (Guarino, 1995) Thus, on one hand, a set of relevant and well-defined concepts needs to be determined, typically for a specific problem or knowledge domain. They need to be organized into well-defined hierarchies with valid inheritance mechanism. On the other hand, in order to support logical inferencing, sophisticated structures of logical relationships need to be determined.

While all of these clearly require manual effort, attempts have been made in automating parts of ontology design such as a generation of initial concept sets or inference of semantic relationships from texts which are then validated and edited by human expert ontology designers. A prominent example of such approach is the Text-to-Onto framework (Maedche & Staab, 2000) depicted Fig. 4-29. It is based on a combination of shallow text-processing techniques (morphological analysis, recognition of name entities, retrieval of domain-specific information, part-of-speech tagging) and machine learning for discovering rules and relationships. Lexical analysis is used to determine word forms and word types as well as to distinguish between different word meanings based on the local phrase context. This process is supported by thesauri and general language semantic networks such as WordNet[29].

The extracted words are collected in a domain specific dictionary from which the concepts for the ontology are selected. Syntactical analysis is used to determine subject-object relationships between concepts whereas multi-word concepts can be extracted by means of a collocation analysis (concepts frequently co-occuring in pairs). Term relevance and clustering techniques are used to group concepts into hierarchical levels whereas non-hierarchical relationships and associations between ontology classes can be discovered by applying collocation analysis to local contexts (e.g. phrases). A comprehensive review of different approaches and attempts of partially-automated ontology design and generation has been presented in (Ding & Foo, 2002)

[29] http://wordnet.princeton.edu/

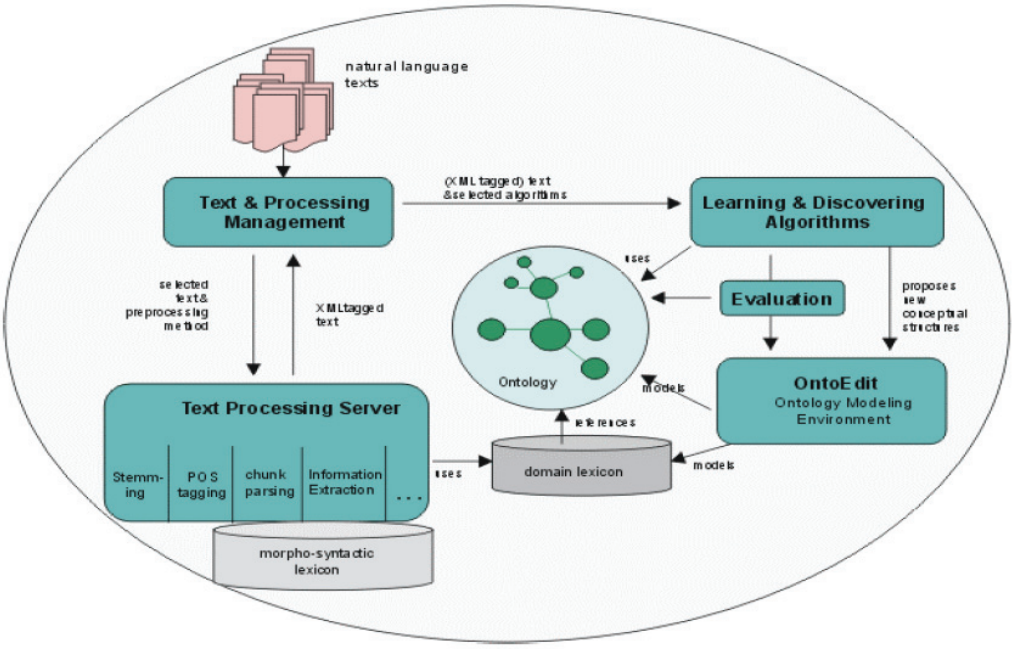


Fig. 4-29. Structure of the Text-to-Onto ontology generation framework (Ding & Foo, 2002)

4.4.2.3. Ontology Applications

As a form of both machine-readable and human-understandable knowledge representation, the originally main indented use of ontologies was semantic knowledge organization facilitating knowledge reuse and sharing. More recently, a distinction between three main classes of ontology applications has been proposed (Gruninger & Lee, 2002): using ontologies for communication, for computational inference and for knowledge organization.

Uses of ontologies
For communication
• between implemented computational systems. • between humans. • between humans and implemented computational systems.
For computational inference
• for internally representing and manipulating plans and planning information. • for analysing the internal structures, algorithms, inputs and outputs of implemented systems in theoretical and conceptual terms.
For organization and reuse of knowledge
• for structuring or organizing libraries or repositories of plans and planning and domain information.

Table 4-1. Uses of ontologies (Gruninger & Lee, 2002)

Within these broad categories the great diversity of existing concrete ontology applications can be organized into four main categories of more specific ontology application scenarios: neutral authoring, ontology as specification, common access to information and ontology-based search (Uschold & Jasper, 1999). A characterization of the individual application scenario classes is given in Table 4-2.

Categories of Ontology Application Scenarios
Neutral Authoring:
• An information artifact is authored in a single language, and is converted into a different form for use in multiple target systems. • Benefits of this approach include knowledge reuse, improved maintainability and long term knowledge retention.
Ontology as Specification:
• An ontology of a given domain is created and used as a basis for specification and development of some software. • Benefits of this approach include documentation, maintenance, reliability and knowledge (re)use.
Common Access to Information:
• Information is required by one or more persons or computer applications, but is expressed using unfamiliar vocabulary, or in an inaccessible format. • The ontology helps render the information intelligible by providing a shared understanding of the terms, or by mapping between sets of terms. • Benefits of this approach include inter-operability, and more effective use and reuse of knowledge resources.
Ontology-Based Search:
• An ontology is used for searching an information repository for desired resources (e.g. documents, web pages, names of experts). • The chief benefit of this approach is faster access to important information resources, which leads to more effective use and reuse of knowledge resources.

Table 4-2. Main categories of ontology application scenarios (Source: Uschold & Jasper, 1999).

4.4.3. Semantic Nets and Topic Maps

Semantic nets represent knowledge structures as graphs of concepts and relationships between them (Soergel, 2000). They have been originally developed as a means of allowing richer relationships between concepts in thesauri. To this end they structure concepts in form of a network rather than hierarchies. In contrast, the described constructs used for modelling ontologies a combination of hierarchical organization and network relationships. Semantic nets have a well-defined set of specific types of relationships such as “whole-part”, “cause-effect”, “parent-child” etc. In general, semantic networks can be considered as a rudimentary form of ontological representation with less expressive flexibility and without logical inferencing. However, due to the absence of formal semantics and related inferencing capabilities they don’t qualify as ontology models. The already mentioned WordNet system is an example of a semantic network of English language, used in a number of information retrieval, artificial intelligence and ontology applications.

A related form of knowledge organisation are Topic Maps. Topic Maps represent collections of topics or concepts and specific kind semantic relationships between them. The three main structural elements comprising a topic map include topics, associations and occurrences (Pepper, 2000). The topics represent conceptual aspects of a subject-matter or domain and are connected to each other through multiple, undirected associations. The occurrence relations connect external resources (e.g. documents) to topics. In addition, each topic can be referred to with multiple names (organized within scopes) and can have multiple role relations to other topics. Such a structure makes Topic Maps are especially suited to organizing and navigating information collections. The Topic Map format has been standardized by the ISO/IEC Standard 13250 and later defined in a corresponding XML Topic Map format (XTM; Park & Hunting, 2002) which provides the basis for representing and exchanging Topic Maps on the Web (Fig. 4-30).

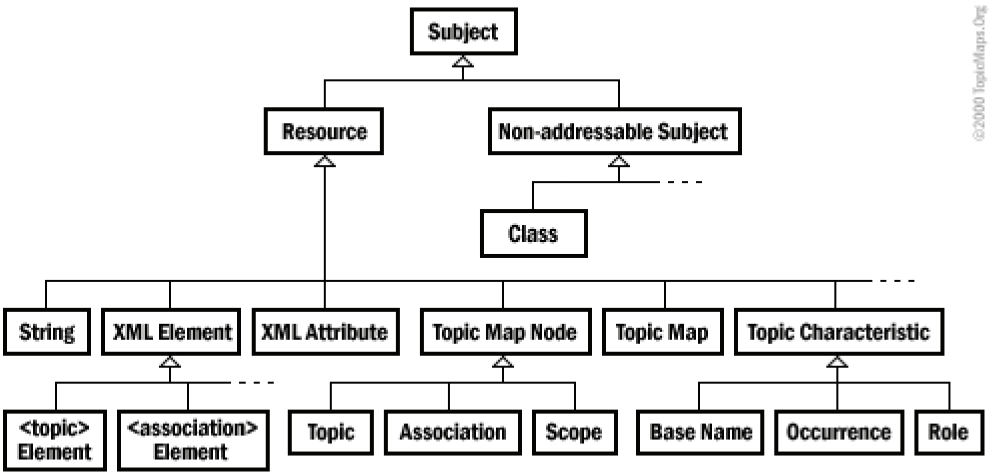


Fig. 4-30. Class hierarchy of the XML Topic Map format[30].

Due to their intuitive concept, Topic Maps had received a lot of attention as means for organizing and navigation information resources. But due to several important difficulties related to Topic Map creation they have never achieved wide-spread use. One important idea of Topic Maps was that end-users would be able to create own personal Topic Maps organize information resources on a specific topic. Based on the XTM standard they would be able to easily exchange maps with others and in this way, share knowledge.

(Lin & Qin, 2002) observe several main difficulties related to the creation and use of Topic Maps by end-users as compared to trained professionals such as: creating knowledge structures is difficult for many users, adding occurrences to topics is too labor-intensive, working with XML-coded topic maps is tedious, due to their highly inter-connected network structure topic maps are not easy to display visually (force-directed graph placement). As a solution to these limitations they propose the development of practical tools for easy creation and editing of topic maps, possibly based on predefined templates or on link-copying from other resources as well as adding hierarchical relationships to ensure simpler visualisation of resulting navigation structures. Examples of such attempts include their Topic Map Repository (ibid.) as well as commercial tools for knowledge management applications such as topic map engines, editors and viewers (e.g. Ontopia Knowledge Suite)[31].

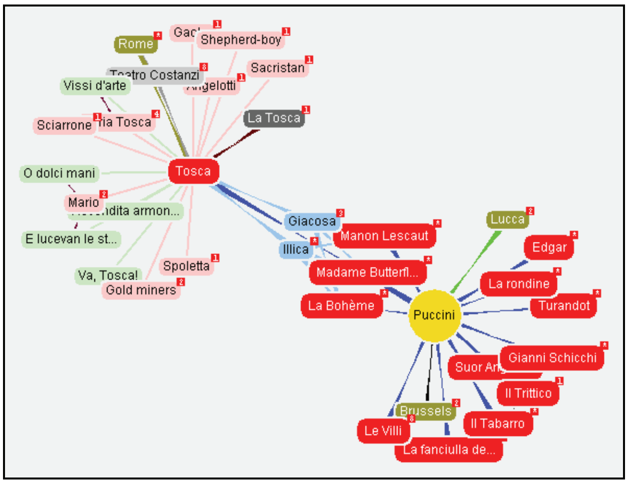


Fig. 4-31. Example of a Topic Map visualisation (Omnigator viewer)[32].

30 http://www.topicmaps.org/xtm/1.0/
31 http://www.ontopia.net
32 http://www.ontopia.net/omnigator/models/index.jsp

4.4.4. Knowledge Maps

Knowledge maps are a hybrid technique between knowledge organisation and knowledge visualisation. On one hand, the notion of knowledge maps has been used to describe quite different knowledge visualisation metaphors. (Eppler, 2001) refers the notion of knowledge maps to all graphic formats that are based on cartographic conventions for representing and referencing knowledge. Accordingly, knowledge maps combine an application specific context as the ground layer on which the individual knowledge elements are mapped.

This can include e.g. organizational competences, organizational structures, business areas or product development processes. Individual elements mapped on such contexts include experts, project teams or communities as carriers of knowledge, as well as patterns of relationships between topics, concepts and knowledge resources (documents, papers, emails). Commonly, the individual elements are grouped or connected with each other based on semantic relationships that are defined either manually or based on similarity data extracted by different knowledge discovery techniques (clustering, classification). This class of approaches focuses on the use of knowledge maps to supporting knowledge communication and transfer between individuals and groups of people (Eppler & Burkhard, 2004).

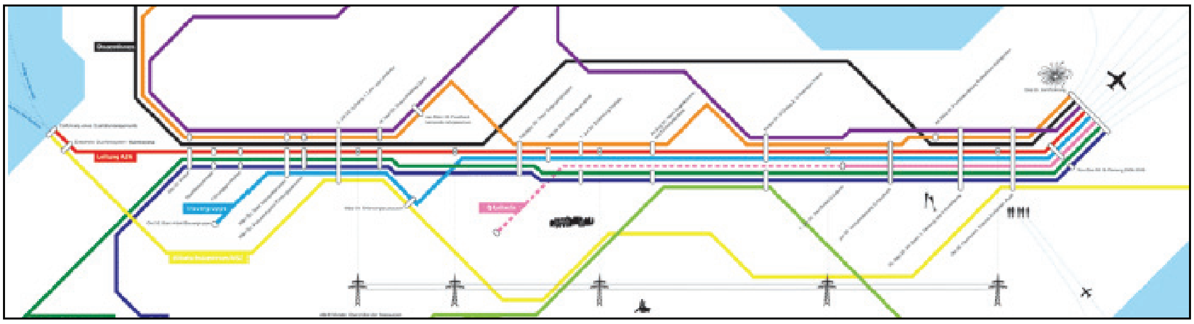
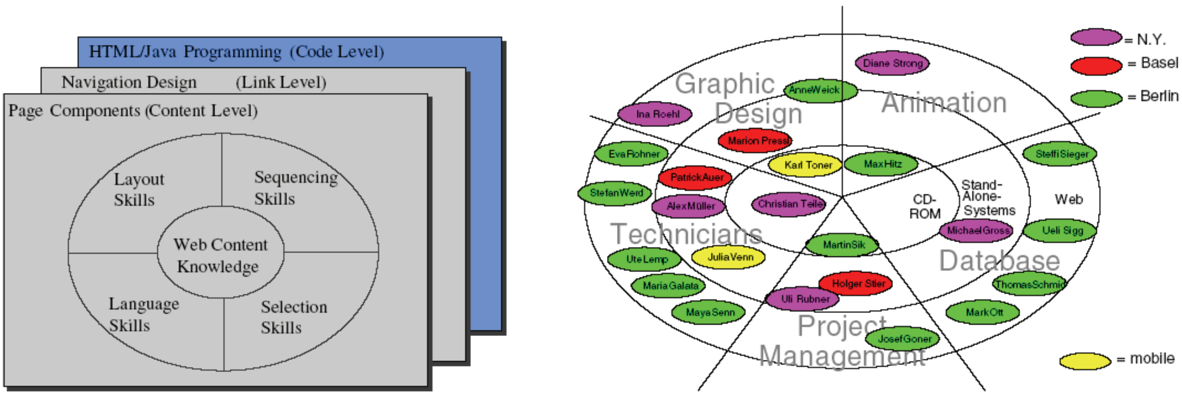


Fig. 4-32. Example of a knowledge map supporting knowledge transfer in project teams (Eppler & Burkhard, 2004). Tube map visualisation metaphor serves to present both an overview and details of a project. Lines represent target groups and stations project milestones. The lines and stations cross where a given target group is involved in a milestone. Original size of the visualisation is 1,2 x 2,4 meter.

In mainstream knowledge management, knowledge maps have been considered primarily with respect to their function of providing semantically organized overviews of available information and knowledge resources (e.g. Probst et al., 1997). Here, representation of knowledge maps is often reduced to standardized models of hierarchical trees presenting relatively static taxonomies. In this approach knowledge maps are concerned with organizing and presenting meta-information of different kinds of available knowledge resources with the primary purpose of facilitating knowledge awareness, identification and distribution.



a) knowledge structure map **b) knowledge carrier map**

Fig. 4-33. Examples of knowledge maps in knowledge management (Eppler, 2001)

Rather than focusing on the visual metaphor (e.g. cartographic maps in Eppler's taxonomy) the notion of knowledge maps is defined by their function of realizing structured overviews of knowledge, regardless of the means employed for visual presentation. In this sense, such maps are a means of knowledge organisation and knowledge structure representation. This view is reinforced by approaches that view knowledge maps as a means of knowledge codification aimed at supporting knowledge flows and organizational memory (e.g. Grey, 1999; Lehner, 2000). Table 4-3 lists the different kinds of knowledge maps in the context of knowledge management according to the taxonomy introduced in (Probst et al., 1997).

KNOWLEDGE CARRIER MAPS	mapping relationships between specific knowledge domains and experts possessing this knowledge (e.g, competence maps)
KNOWLEDGE STRUCTURE MAPS	representing the conceptual structure of a knowledge domain (e.g. concept maps, cluster maps)
KNOWLEDGE RESOURCE MAPS	representing overviews of all available knowledge resources in an organization
KNOWLEDGE FLOW MAPS	linking knowledge and organizational processes
ARGUMENTATION MAPS	representing causal schemas for decision-making (e.g. cognitive maps)
LOCAL THEORY MAPS	representing causal structures of reasoning of different local contexts in an organization

Table 4-3. Taxonomy of knowledge maps in knowledge management (Probst et al., 1997)

In contrast, research in the field of knowledge visualisation has been increasingly considering the need of systematic research and evaluation of the potential of different methods and knowledge map metaphors for supporting interactive and dynamic visualisation of knowledge to support knowledge transfer in specific application contexts and social structures (Eppler & Burkhard, 2004). In fact, knowledge carrier maps, knowledge structure maps and knowledge resource maps distinguished in Table 4-3 are the classical examples provided in different knowledge management systems and applications. More specialized is the application of knowledge flow maps, argumentation and local theory maps which are more likely to be found in Epplerian approach to visual knowledge transfer facilitation and the previously considered use of knowledge visualisation for mediating group discourse and decision making (Section 4.3).

This field of research places more emphasis on the role of knowledge maps in supporting knowledge exchange through knowledge communication, creation and application. Accordingly, of critical importance are criteria such as the choice of appropriate visual metaphors on which such knowledge maps are based, the ways in which individual elements can be exchanged, investigated and modified by the users to reflect their own interpretation and the results of insights gained in group interaction.

4.4.5. Summative Discussion

In summary, we can observe that for all described forms of knowledge representation (onotologies, semantic nets, topic maps) the main difficulty consists in the need for explicit (manual) creation of knowledge structures, their formalization and their population i.e. the linking between defined concepts and relevant information resources. These reasons make such techniques difficult to apply in highly dynamic and unstructured collections and informal social contexts such as communities. In particular, the many different points of view and decentralized dynamics of information exchange and knowledge creation make structuring processes requiring explicit user effort and centrally coordinated, highly structured procedures impractical for in such environments.

However, some form of semantic specification of conceptual structures is needed in order to elicit and visualise structures of personal and community knowledge. The described methods and existing experiences in their application allow us to determine the right level of formalism appropriate for our application context that can be incorporated in an unobtrusive method for revealing implicit knowledge structures of human users. In fact, while in our application context light-weight ontological structures such as a hybrid of concept maps and semantic networks seem adequate for representing structures of personal and community knowledge, developing complete formally modelled ontologies for different communites between which logical inferencing and mapping mechanisms can be established are not a feasible solution.

As highly informal structures community are rarely able to generate to necessary effort and coordination needed for designing, building and implementing an ontology. Furthermore, the required establishment of consensus makes such procedures inherently inappropriate to diversity of views and evolutionary dynamics typical for communities as informal social networks. As a result approaches to community systems and knowledge exchange which are based on the assumption that formally defined community ontologies are built (e.g. Christophides et al., 2000) will have limited relevance for practical community application.

Even the individual ACM communities which are well organized and include procedural authorities as well as members concerned with ontology-development itself, have not developed community ontologies. The only explicit conceptualization is a meta-classification scheme for the ACM network as whole, which suffers from being too generic for specific needs of individual communities and is thus not widely used by researchers (Carotenuto et al., 1999). Though its use for indexing can be enforced by authority rules (e.g. assigning ACM classification categories when submitting conference papers) the intended appplication for supporting information access and retrieval fails since the concepts are not congenial with users needs and specific community language.

An alternative are knowledge maps combining knowledge organisation and visual representation. They are commonly used as static conceptual structures providing an overview of available knowledge resources in a given domain and facilitating navigation in complex information spaces. However, most solutions provide only one, unified knowledge structure to which all users and communities must adhere in order to access the shared information space and to make their knowledge accessible by others. The connection of personal knowledge structures with the shared taxonomy or ontology must be created manually by the users through assigning the information to the centrally defined concepts (Bonifacio et al., 2000). While personal knowledge organization takes place naturally as a result of any problem solving or information seeking, assigning documents to another scheme requires additional effort that users are often not willing to undertake (Lacher, 2003).

4.5. Information Visualisation for Information Seeking

Traditionally, information retrieval has been concerned with the development of search algorithms and tools for supporting the finding of relevant information in large information collections (databases, web, information spaces). The notion of information seeking describes a more user-centred approach concerned with supporting the different aspects of the information search process, from the user perspective. Information visualisation has been used to support information seeking in three principal ways: by visualising concept spaces for navigation through information collections, by using visualisation to contextualise search results in topical categories and by using visualisation to support query formulation. The presentation of these approaches in the sections 4.5.1-5.5.3 is partly based on reviews presented in (Hearst, 1999) and (Kunz, 2005).

Another class of approaches, so-called information workspaces consider the use of visual interfaces to support the entire information seeking *process* i.e. the formulation of queries, interactive visualisation and inspection of selected search results and storage and organization of seeking results into personal collections. This work is summarized in Sections 4.5.4 and 4.5.5 and is mainly based on (Card, Mackinlay & Shneidermann, 1999), (Hearst, 1999), (Amento et al., 2003) and (Qu, 2003).

4.5.1. Visualisation of Concept Spaces for Navigation

A common technique for supporting navigation in complex information spaces is the visualisation of networked concept structures describing the semantic space of a domain or document collection. The conceptual structures can be obtained with different techniques discussed in the previous sections. The most frequently used ones for supporting navigation are topic maps, semantic nets and ontologies although also approaches to applying concept maps for navigating information collections exist (Section 4.3.1.2). Besides for navigation, information networks combining concepts and information items as their concrete instances are also used to contextualise search results (see Section 4.5.2). The goal of the visualisation is to provide an overview of the conceptual structure in order for the user to identify concepts that best express his information need. This approach is often used when information need is unclear or difficult to express due to the unfamiliarity of the domain or terminological differences between user's background knowledge and target collection.

Some approaches use classical graph visualisation techniques based on the force-directed placement techniques (Section 4.2.4.3). The disadvantage of such solutions is that only a part of the entire conceptual structure can be displayed at a given moment, which greatly reduces the overview and thus the effectiveness of navigation (e.g. the topic map visualisation depicted in Fig. 4-31, Section 4.4.3). As conceptual structures are often complex containing many different concepts and connections between an effective visualisation needs to solve two main problems. The first is how to provide a detailed display of the concept network in focus of user interest while at the same time allowing him/her to keep an overview of the entire space. The second is how to display potentially large numbers of relations between the concepts making up the vertices of the graph i.e. the links of the network.

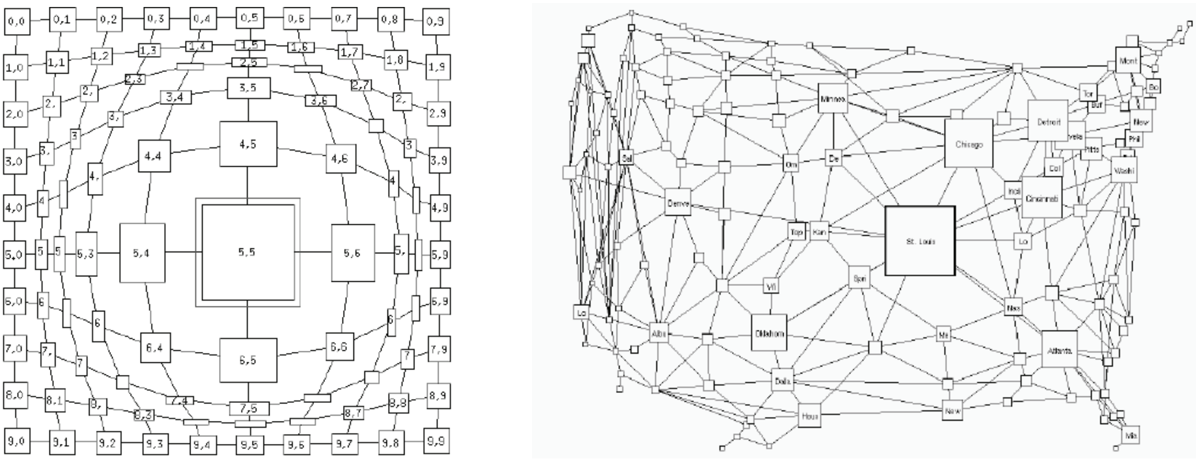


Fig. 4-34. Focus and context visualisation based on graphical fisheye views (Sarkar & Brown, 1992; 1994)

The two main solution approaches to this problem include the detail+overview and focus+context techniques. The detail+overview solution is based on using two separate views: while the main portion of the screen is used to display the enlarged view of the part of the network currently in user's focus, an additional small window provides a miniaturized overview of the entire network. This solution is intuitively understood by the users since the small overview behaves as a cartographic map with respect to the detail view. But it poses problems regarding the placement of the second view and incurs additional cognitive demand on the users who need to mentally coordinate and integrated the two views into one.

The focus + context techniques keep both views in one surface by using different kinds of non-linear distortion algorithms to assign the area closer to user focus more space and display enlarged, more detailed information while information farther away is assigned less space and de-magnified (Fig. 4-34). This allows interactive navigation of the concept network with detailed inspection of user-defined focus while continuously keeping an overview of the entire structure. A popular solution is the so-called fisheye technique originally proposed by Furnas (1986) and developed into a graphical fisheye view by (Sarkar & Brown, 1992).

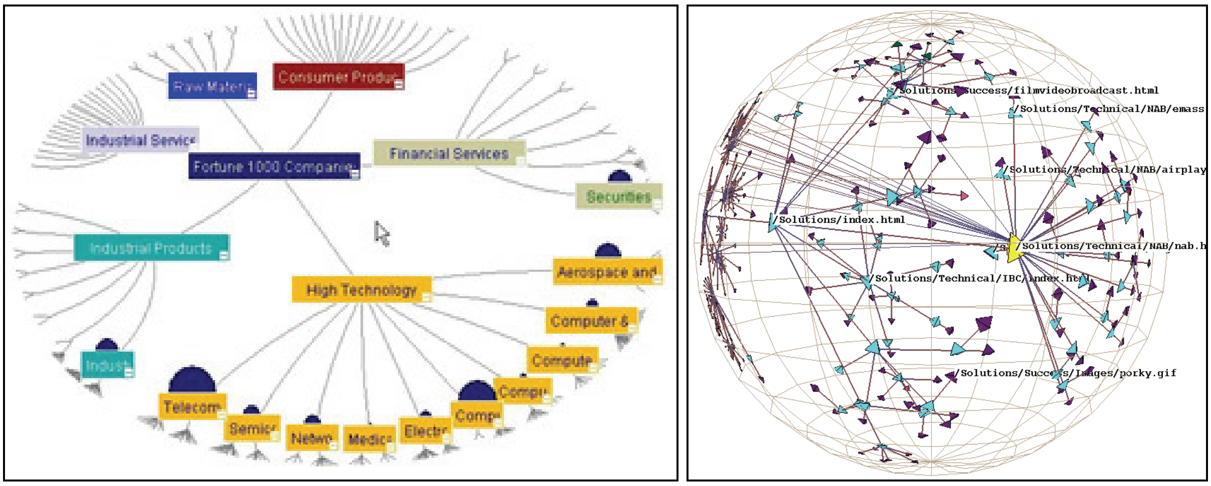


Fig. 4-35 Hyperbolic tree visualisation in 2D (inxight.com) and 3D (Munzner, 1997)

Other widespread techniques include the cone-tree (Robertson et al., 1991) and hyperbolic tree (Lamping et al., 1995) used for visualising large hierarchical category structures (Fig. 4-35). In the cone-tree model each tree node is represented by a three-dimensional cone which can contain sub-nodes or leaf nodes. This method has been used both for visualising hierarchical concept structures such as Topic Maps (LeGrand & Soto, 2000) as well as for visualisation of multiple category sets in visual information workspaces (Fig. 4-44, Section 4.5.5).

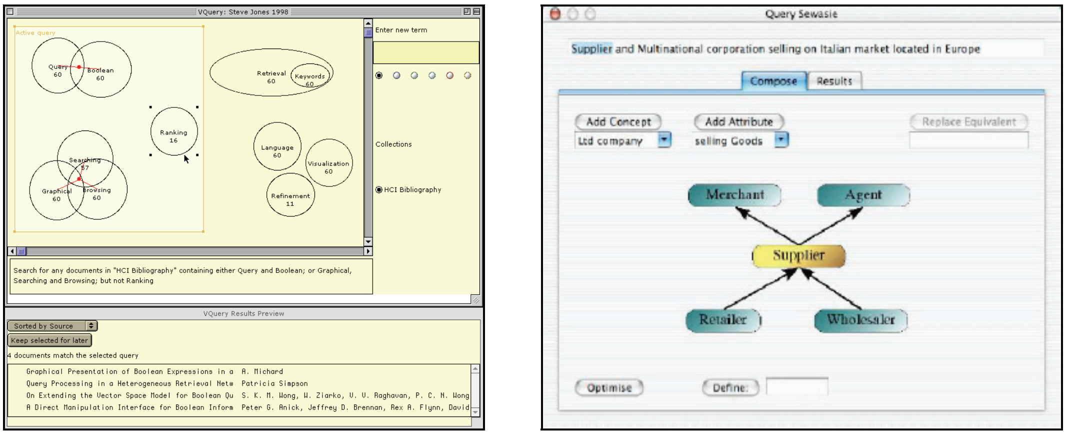
The hyperbolic tree techniques first transforms the tree structure onto a hyperbolic plane before projecting it on a Cartesian 2D (inxight.com) or sometimes 3D surface (Munzner, 1997). The main disadvantage of this technique is a limited visibility of non-hierarchical inter-connections since only relations between nodes in the vicinity of a given focus are visible. This makes it inappropriate for visualising concept graphs and highly inter-connected hierarchies.

4.5.2. Visualisation for Query Formulation Support

A straightforward approach to supporting keyword-based query specification is to present the user with a list of concepts relevant for a given domain based on a domain-terminology model (e.g. thesauri). However, such approaches are limited only to specialized searches where the target domain is well defined. In this case, even specific categories of predefined query types can be constructed and used to support both the user expression of his information need as well as for improving retrieval.

Such an approach has been realized in the DynaCat system (Pratt et al., 1999) for supporting information access to large medical databases (see Section 4.5.4). In (Fowler et al., 1991) a thesaurus has been used to visualised and extract search keywords from a natural language query. Both the search keywords, the retrieved documents and related concepts from the thesauri are visualised in form of graphs based on inter-document similarities and related inter-concept relationships.

The associated thesauri terms can be interactively added or removed from a search query. Visualisation of query terms based on Venn-diagrams for supporting query specification has been proposed by systems such as VQuery (Jones, 1998). Query terms are represented as rings containing sets of documents relevant for a given term. The intersection of the rings represents documents containing both terms (Fig. 4-36, left). The term rings can be interactively manipulated so to express logical combinations of terms (intersecting rings for logical AND, disjunctive rings for logical OR) and thus specify a query.



a) VQuery b) SEWASIE

Fig. 4-36. Visual and semantic query specification: a) Venn-diagrams and b) Semantic nets

One of the main applications of ontologies is supporting knowledge-based information access through semantic query formulation. Querying an ontology is based on identifying instances of specific concept classes related to a user's information need, by formulating logical conditions based on the specific logical formalism of a given ontology implementation. To this end the users need to be provided with insight into the existing classes, properties and relations describing a given domain model and they need to be familiar with the appropriate logical formalism. Systems such as SHOE (Heflin & Henler, 2000) and SEWASIE (Catarci et al., 2003) aim at reducing the user's cognitive burden by providing overviews of ontological classes and means for formulating conditions through graphical user interfaces.

SHOE uses a classical table and forms interface similar to those of bibliographic information systems. The available classes are represented in a hierarchical list accompanied with form fields representing kinds of conditions (attributes and class relations) that can be set by the user for a given class. The user selects a set of classes and formulates conditions by simply entering an attribute value in a given condition field (e.g. "Smith" for "publicationAuthor") while the results are presented in a table format, much like tables in relational databases. In the Ontobroker system (Fensel et al., 1998) this approach is accompanied by a visualisation of the class hierarchy in form of a hyperbolic tree (see Section 4.5.1).

A simplified, more visually oriented approach has been realized in the SEWASIE project (Semantic-Webs and Agents in Integrated Economies). As a starting point the user is presented with a choice of predefined query scenarios accompanied by a diagrammatic interface for query composition (Fig. 4-36, right). By choosing a prototypical scenario the set of relevant ontology classes is narrowed down only to those related to the scenario in question. Selecting a given concept (ontology class) puts it into a focus of visual diagram showing other related concepts (more specific and more general). In this way, the presentation of available semantic constructs and relations is always contextualized to the current focus of user interest.

4.5.3. Visualisation for Contextualisation of Search Results

The contextualisation of search results aims at satisfying two main user needs in the information seeking process: 1) understanding how the retrieved documents are related to the formulated query in order to assess their relevance and 2) discovering new terms helpful for reformulating a query in order to better express a given information need.

The simplest contextualisation method for supporting relevance assessment is "keyword-in-context" (KWIC) which presents excerpts from document phrases in which the search terms occur and highlights or colour-codes the occurring terms. Most current Internet search engines use this technique (e.g. Google). Another related technique are tilebars (Hearst, 1995). Here each document is accompanied by a rectangular bar which marks relative location and intensity of term occurrence in the document (one bar per search term). The length of the bar is proportional to the document length and the shades of gray indicate the match frequency (Fig. 4-37). Both of these methods provide a micro-level contextualisation of search results providing insights into term distribution at the level of individual documents.

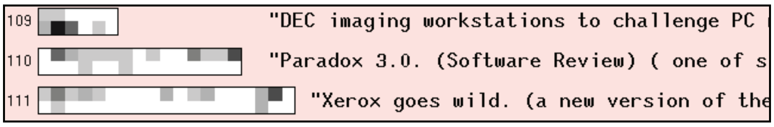


Fig. 4-37. Tile bar visualisation of term distribution in search results (Hearst, 2000)

In our context, more relevant is the macro-level contextualisation of documents with respect to their semantic context. The main classes of approaches here include: a) contextualisation in predefined categories, b) dynamically generated categories for a given search result set (based on clustering or co-citation analysis) and c) implicit contextualization based on overall inter-document similarity relationships in a given collection.

4.5.3.1. Using Predefined Categories for Result Contexts

A representative example of the approaches using a predefined set of categories is the DynaCat system (Pratt et al., 1999). This system uses a domain-specific terminology model to map most typical terms characterizing a set of search results to predefined set of more abstract categories appropriate for a specific domain (medicine in this case). In addition, the queries are also mapped to a predefined set of typical query categories relevant for this domain (e.g. treatment, prevention or action queries) which are linked to corresponding category types relevant for a given kind of query[33]. This is an example of a knowledge-based search result visualisation where a predefined ontology-like conceptual structure is used to contextualise documents retrieved by a search query. The interface of the DynaCat system presents the results in a ranked list alongside a hierarchical list of topical categories that can be navigated to display only documents related to a specific topic (Fig. 4-38).

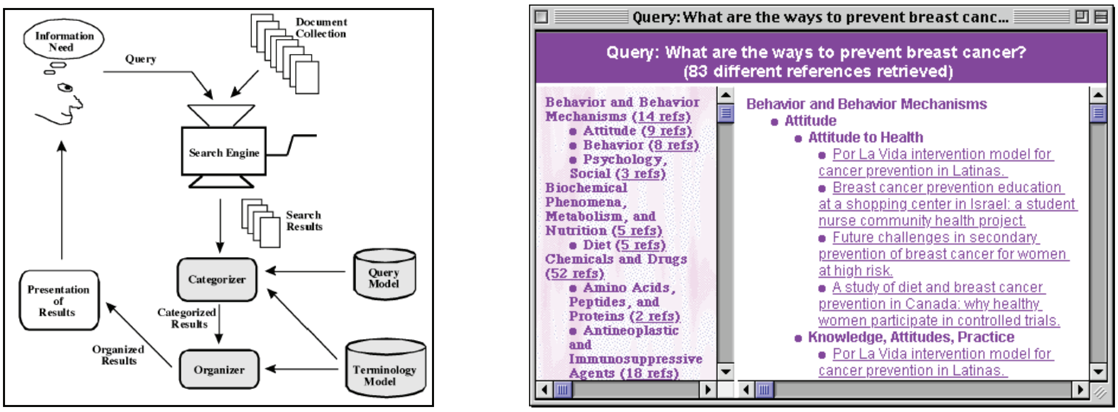


Fig. 4-38. List-based search contextualization in predefined categories: DynaCat system (Pratt et al., 1999).

In recent years different commercial systems providing graphically oriented visualisation of search results in predefined topical categories have appeared. They take advantage of existing topical directories on the Web[34] to categorize search results. The visualisations adopt either the information landscape metaphor in the style of SOM-based document maps (Section 4.2.4.4., 4.2.5) or graph-based visualisations. Two well-known examples include The Brain[35] and Antarctica[36]. The Brain interface accompanies a search result list with a graph-based visualisation of a related topical hierarchy that can be interactively navigated by the user.

[33] Obviously, such mapping is applicable only in domains where with a set of very specific frequently occurring query types. For details on query mapping and practical limitations see the original article.
[34] E.g. Open Directory Project, http://www.dmoz.org
[35] http://www.thebrain.com
[36] http://maps.map.net

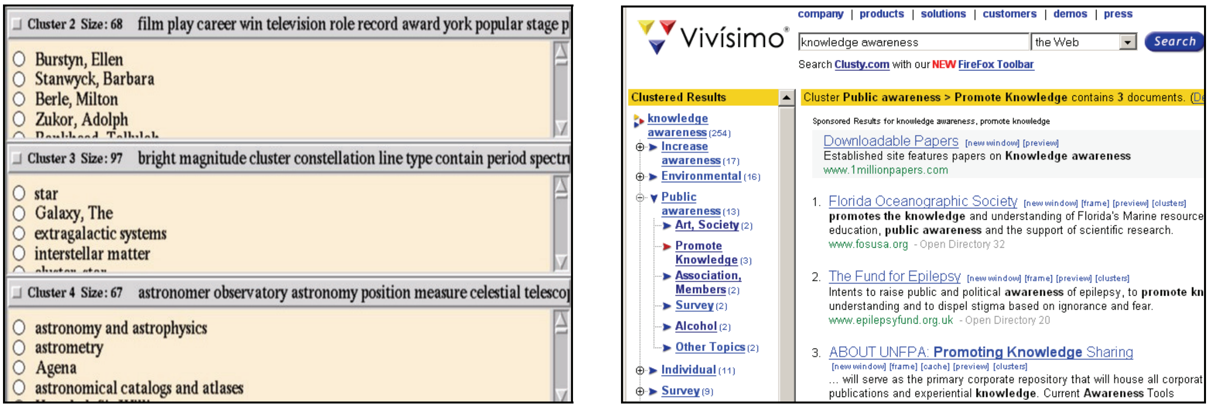
Each topic is represented with a node while lines represent hierarchical relationships between them. Selecting a topic places it in the centre of a radial visualisation that surrounds the main topics with its sub-topics. The search result list is updated accordingly to display newly retrieved documents for the selected topic. Invoking a manual search query causes the topic with most hits to be positioned in the centre. The Antarctica system implements the landscape metaphor showing topics as countries on a continent and document as cities. Search results are highlighted on the map and can be explored by zooming into topic regions.

a) The Brain b) Antarctica

Fig. 4-39. Visual search contextualisation in predefined categories: a) topic graph and b) info landscape.

4.5.3.2. Dynamically Generated Categories

One way of generating thematic contexts for search results "on-the-fly" is to cluster retrieved documents into topical groups based on inter-document similarity (see Section 4.2 for different methods). The search results can then be presented either in a list-based overview of topical groups or assigned to hierarchical categories. An example of the former is the application of the Scatter/Gather technique (Section 4.2.5.1) to clustering search results (Hearst, 1999). Instead of applying the technique to the whole collection, only the retrieved documents are clustered and the presented as a list of groups with topical summaries in form of most typical terms (Fig. 4-40, left). The user can then iteratively choose clusters of interest and let the system re-cluster only that part of the retrieved document set.



a) Scatter/Gather b) Vivisimo

Fig. 4-40 Search result contextualization in dynamic categories with list-based interfaces: a) flat categories and b) hierarchical categories.

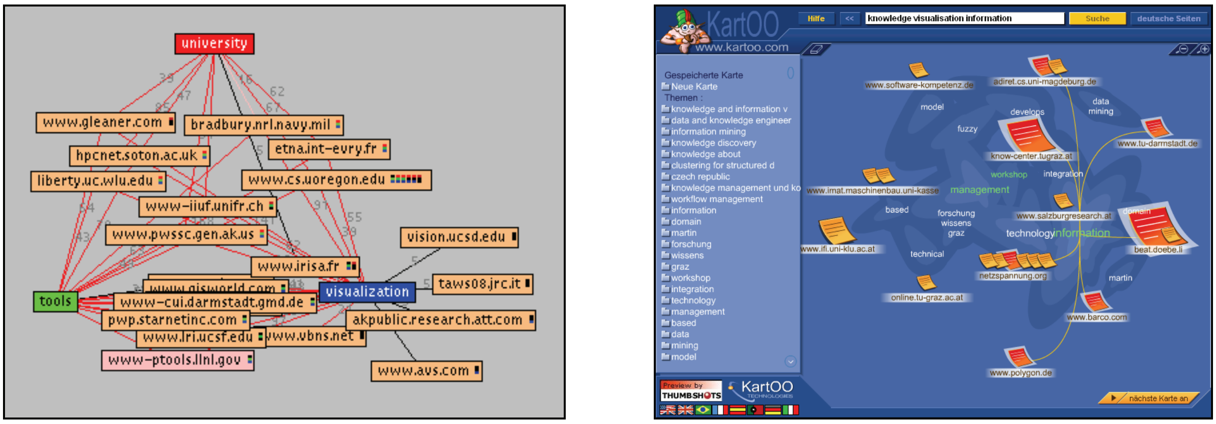
Another approach is to organize documents into hierarchical topics based on inter-document similarity by means of hierarchical clustering techniques (Section 4.2.3). The topic names are typically extracted by taking most relevant terms of cluster centroids. The interface can then present the user both with a complete list of retrieved results as well as with a set of navigable topical hierarchy: choosing a topic displays the list of documents that have been grouped together and characterized with the topic in question. An example of this approach is the Web meta-search engine Vivisimo[37] that clusters search results into groups based on information from titles and excerpts of sentences typically returned in search engine lists (Fig. 4-40, right).

In contrast to solutions that allow the assignment of documents to only one category, a frequent case in practice is that documents can belong to several different categories simultaneously. A straightforward solution is to present a category structure and visually mark topics related to the set of retrieved documents or a user-defined subset. Such solutions have been adopted in more complex interfaces aiming at providing an integrated environment supporting the entire search process. They are presented in Section 4.5.4. Another possibility of showing multi-topic relationships is to refrain from explicit assignments of documents to single topics and provide a visualisation that enables the user to infer relationships to different thematic contexts. Such approaches are outlined in the next section.

4.5.3.3. Implicitly Defined Contexts

Instead of explicitly clustering documents into topical groups, thematic contexts can be depicted by using a similarity-based visualisation which positions documents in a spatially-organized coordinate-system based on their semantic distance to a set of semantic coordinates. We have seen examples of such visualisations for depicting the overall semantic structure of entire document collections in Section 4.2. Given such a similarity based document space visualisation, the search results can be contextualized by visually marking documents from the result set in the map of the document space. The location of the documents will indicate their semantic context by the vicinity to one or different characteristic terms used to label areas of the map (e.g. Themescape, Section 4.2.5.2).

Since the clusters are not explicitly formed but only visually implied no exclusive categorization is undertaken and multiple topical relationships can be induced based on the distance of the document to different topics. Another contextualization method based on visualising implicit semantic relationships has been developed in the SQWID (Search Query Weighted Information Display) system for visualisation of Web-search results (McCrickard & Kehoe, 1997). This method generates graphs of search results consisting of web page nodes and most important term nodes, with links between them expressing the relevance of a given term for a page. The graph is visualised by taking the three most important terms and placing the web pages in such a way that their distance from individual terms reflects the relevance of each term for a given page: the more relevant the term, the smaller the page distance from that term (Fig. 4-41a).



a) SQWID **b) Kartoo**

Fig. 4-41. Visualising search results based on implicit semantic contexts allowing multi-topic relationships.

[37] http://www.vivisimo.com

As a result, pages located in the middle of the term triangle will be related to all three terms, pages along the edges to two terms and pages around the borders of the triangle indicate a relation to one term only. Pages matching no terms float away of the term nodes to the edge of the screen. Accordingly, the thematic context of a given page is only implicitly indicated by its distance to one or several terms, characterizing the result set. The user can change the choice of terms for the term nodes and determine the number of pages displayed with a slider.

A similar principle has also been realized in the kartoo[38] meta-search engine that combines the landscape metaphor of document maps with term-page relationships to characterize the relations in a given search result set. Pointing on a term on the map highlights pages for which this term has been determined relevant (Fig. 4-41, b) while selecting it initiates a new search. In addition a list of most relevant terms (possible topics) is also displayed outside of the map and can be navigated in the same way. The visual simplicity of the interface is maintained by greatly limiting the number of pages visualised in one map and dividing the search result into a set of individual maps.

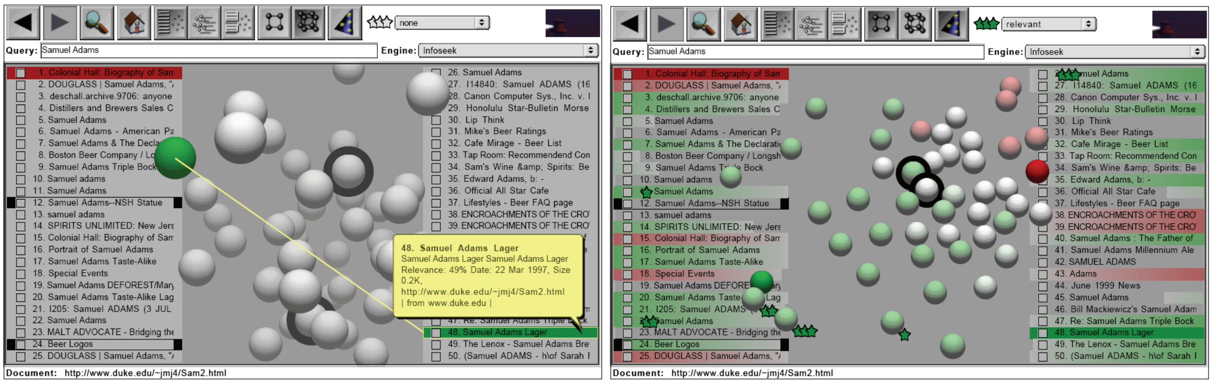


Fig. 4-42. Combined visualisation of search results in a ranked list and visual clusters based on inter-document similarity in the Lighthouse system. Based on user ratings of inspected documents the system (left) visualises the assessed relevance of the remaining documents (right).

The Lighthouse system (Leuski & Allan, 2000) integrates a ranked list display of search results with a visual clustering based on inter-document similarity. Documents are represented as spheres in 3D or 2D space such that spheres representing similar documents are close to each other, whereas spheres representing very different document contents are far apart (Fig. 4-42). The two different views provide different types of semantic information: the list shows the relevance of documents to the query while the visualisation shows how the documents are related to each other.

Having inspected individual documents the user can express his assessment of document relevance by clicking the checkbox attached to the document title: one click – non relevant, two clicks -relevant, three clicks – unmark (Fig. 4-42, left). Spheres of documents marked as relevant are automatically coloured in green and the document titles are assigned a green background in the list. Documents marked as irrelevant are marked with red. Based on this information the user can ask the system to assess the relevance of the remaining documents not inspected by the user. The system assessment is based on the comparison of content-based similarity of un-inspected documents to documents rated by the user. Documents similar to those rated as relevant are marked in shades of green (intensity reflecting degree of relevance) while documents assessed as non-relevant are marked in shades of red (Fig. 4-42, right).

4.5.4. Summative Discussion of Visualising Concept Spaces, Search Results and Queries

Graph visualisations of search results are very popular but suffer from several usability problems. One is the constantly changing appearance of visualisations produced by force-directed placement following changes in user focus. This results in cognitive strain and confusion, distracts the user from his/her primary task and makes visual search through exploration difficult. In case of large or highly-

[38] http://www.kartoo.com

interconnected graphs the visual representations also become overly complex. Due to space limitations, representing context in graph-based visualisations of networked concept spaces is often limited only to the immediate neighbourhood of a selected concept. Larger patterns of relationships and inter-connections between concepts are difficult to detect.

Visually supporting query specification improves effectiveness and precision of search queries (Hearst, 1999). However, providing sophisticated interactive visualisations for query formulation such as Venn-diagrams leaves little space for simultaneously visualising search results as well. An integration of into interfaces offering integrated support for the entire search flow of typical information seeking tasks is difficult to achieve. The presented visualisation approaches focus on isolated, specific problems of different phases occurring in an information seeking process (e.g. querying, navigation, search result inspection).

More suitable in this context is the use of thesauri or category trees and maps which can be used both for supporting query formulation or direct topical navigation. Such visualisations can be used both to visualise query terms as well as search results in the same model (see e.g. Cat-a-Cone interface in the following section). The use of hyperbolic trees is an appealing solution for presenting hierarchical trees but it only supports contextualisation of search results in one category at a time. On the contrary, a visually more demanding (3D) cone-tree visualisation allows multiple-category search result visualisation.

Existing approaches to using predefined (hierarchical) category lists allows only step-wise insight in the retrieved document set: only one category can be selected at a time and only the portion of the documents assigned to that category is displayed. This makes obtaining an overview and insight into the topical structure of the retrieved documents very difficult. A better solution is the use of visual document maps based on the landscape metaphor such as in the Antarctica system (Fig. 4-39, Section 4.5.3.3).

Another solution is offered by the scatter-gather approach since all topics are visible at once and can be iteratively refined based on user focus. However, the display is overloaded with textual information causing great visual clutter as no other means than text are used to present topical clusters and their content: clusters are described by sets of terms rather than labels and sets of document titles (see also Fig. 4-5, Section 4.2.5.1). Other examples of using implicitly defined categories (sets of keywords instead of single labels) demonstrate that a more suitable solution is to use document maps based on the landscape metaphor (e.g. kartoo, WEBSOM).

The generation of explicit topical labels by extracting most characteristic terms from the retrieved document set bears both advantages and disadvantages to the above methods. On one hand, predefined categories need to be established by hand for each domain are too generic if defined at a general level. On the other hand, the automatically extracted topical labels tend to represent single-term concepts which do not directly represent thematic contexts. Furthermore, even if collocation analysis methods are applied to extract multi-word concepts they are often likely not to identify correct topical descriptors since they may not be explicitly used in the retrieved documents.

However, the use of maps based on the landscape metaphor seems to be most suitable for contextualising search results since it allows a quick overview both of the retrieved documents and their thematic contexts as well as of the overall topical structure of a given domain. In this way, the users can easily switch between browsing and goal-directed search which has been reported as a preferred strategy in different user studies (Hearst, 1999; Lin, 1995). One problem is the visualisation of large sets of search results, since the map space can become quickly overloaded. This can be solved by zooming and focus+context techniques such as graphical fisheye views (Section 4.4.1).

In our application context the latter method seem most promising for providing visual overviews and contextualized information access to community spaces. The critical issue thereby is the generation of topical structures as a basis for contextualization since they have to reflect personal and shared structures of knowledge of community members. A related discussion has been presented in previous sections on knowledge elicitation and representation (Section 4.3, 4.4).

Finally, an important observation is that though different visualisation solutions for supporting information access have been proposed, there exist few evaluations of their use for specific application contexts. Most existing studies and reviews of empirical findings (see overview in Hearst, 1999; Chen & Yu, 2000) are based on generic information tasks and task models, such as (Wehrend & Lewis, 1990; Belkin et al., 1995; Shneidermann, 1996). Such studies provide important contributions and valuable insights into the intrinsic capabilities and limitations of a specific method for informing the designs of specific solutions. However, the results of general studies such as those whether graphical overviews do or do not support information access better than list-based interfaces[39] may highly depend on the nature of a specific application context. In particular, in our problem context the theoretical analysis presented in previous chapters and the requirements for its practical evaluation, all suggest that visualisation of knowledge structures and their use for supporting information access to unfamiliar domains should provide valuable support to cross-community exchange of knowledge. Proposing a concrete solution for this line of approach and submitting this hypothesis to an application-specific empirical evaluation is an important goal of this thesis (see Chapter 5).

4.5.5. Information Workspaces

The concept of information workspaces has been introduced in (Card et al., 1991) do describe environments which support the entire information seeking process with special focus on the capabilities of not only visualising but also manipulating, evaluating and organizing information into personal collections. The simplest and most popular way of organizing information into personal workspaces are bookmarks containing hierarchical organized references to web pages (URLs). This basic technique has been expanded into more sophisticated information workspaces taking advantage of visualisation and automated retrieval techniques. An example is the Butterfly interface to citation databases (Mackinlay et al., 1995) which contextualizes access to individual articles in a visual space representing the article references and its citers. This allows users to easily browse related articles and groups of articles as well as to generate queries for retrieving related articles with respect to a particular kind of relationship.

Other systems realize visual workspaces that use the object-oriented metaphor in combination with direct manipulation to provide cognitive shortcuts to more complex information seeking services. The DLite system (Cousins et al., 1997) represents queries, document sources, documents and search results as graphical objects that can be manipulated on a two-dimensional surface: moved around and dropped on each other in order to invoke a specific action. For example, by filling out fields in the query constructor object the system creates a query object that can be dropped on a document source object to initiate a search in the corresponding document collection. Several search result sets can be displayed at once and individual document icons can be dragged out and dropped into the personal pool of relevant articles or on other objects representing different services, such as document summarizers or language translators. Document details are displayed in a separate browser window which can also show Scatter/Gather style clustered overviews of the search result collection. In this way the control space is separated from the space used for scanning and reading actions in the information seeking process.

A somewhat simpler set of services but more consistently spatial visual organization has been realized in the VIKI system (Marshall et al., 1994) which allows users to organize personal collections by arranging documents on a two-dimensional plane. It also supports hierarchical organization and sophisticated visual layouts for large information spaces such as graphical fisheye views (Shipman et al., 1999). The DAFFODIL[40] system provides a sophisticated information workspace focusing on “strategic support” for information search in digital libraries (Fuhr et al., 2002). The idea of strategic support is to combine complex sets of actions into easy-to-use high-level search functions (so-called stratagems) exploiting not

[39] See (Hearst, 1999) and (Becks, 2001) for reviews of results and experiences gained in such studies.
[40] DAFFODIL = **D**istributed **A**gents **f**or User-**F**riendly Access **o**f **Di**gital **L**ibraries.

only explicit user queries but also the properties of the information structure of a specific domain. To this end the interface adopts a similar metaphor as DLite and represents major elements of the system as graphical objects whose functions can be activated by simple drag&drop. For example, users can associate personal profiles with digital library objects such as journal or conference proceedings. As a result the meta-information from these objects (e.g. authors, titles, topics etc.) is automatically associated with the user's profile and can be used to invoke different classification and recommendation services for finding relevant documents.

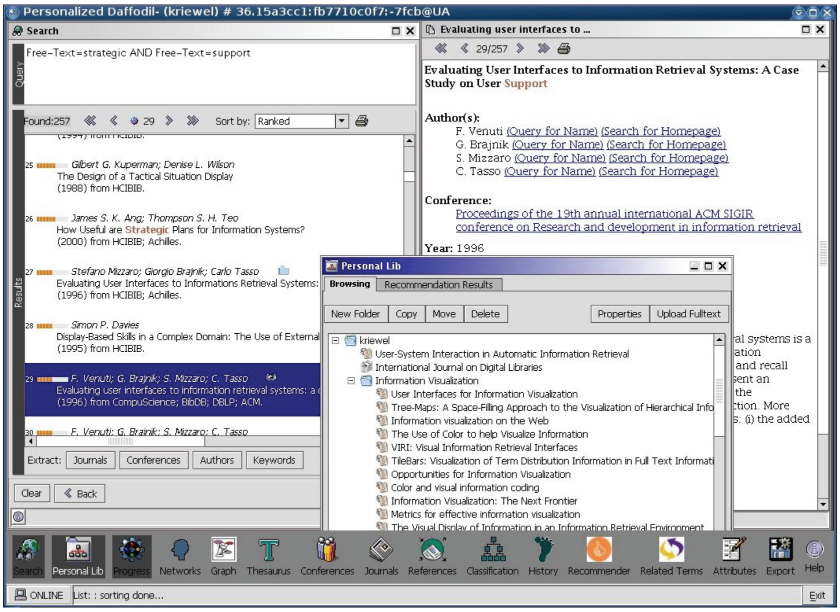


Fig. 4-43. The DAFFODIL information workspace: supports browsing, search, classification, collation and organization of documents in personal and shared folders (Source: Kriewel et al., 2004)

DAFFODIL's graphical interface provides an integrated environment supporting browsing and goal-directed search as well as collation, organization and re-use of retrieved documents in personal and shared folders. The main application context is search in scientific collections. To this end the system also supports citation analysis and the visualisation of relationship graphs displaying collaboration networks (based on co-author relationships) for a given author relevant for a specific search. A classification tool allows browsing of classification schemes of a given domain (such as the ACM Computing Classification System) and query formulation is supported by a thesauri tool. The system also provides user with information on new or changed objects related to previous searches and supports collaborative recommendations based on content-similarity of user's personal folders.

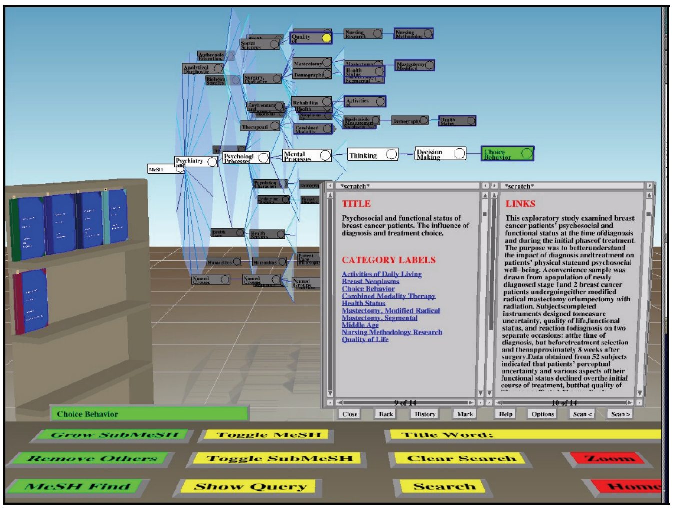


Fig. 4-44. The Cat-a-Cone interface: integrated querying and browsing through hierarchical categories with multiple document-topic relationships, contextualisation of search results and query formulation support.

In the Cat-a-Cone interface (Hearst & Karadi, 1997) a combination of explorative and goal-directed access to large document collections is supported by integrating the querying of text collections with browsing of associated hierarchical category trees. A predefined set of categories is visualised by a cone-tree model (Section 4.5.1) and a "virtual book" metaphor is used for representing the document set retrieved by a search query. Users can access to collection through several different interaction modalities. Entering a category name visualises related categories in the alphabetically organized category tree which supports the selection of terms for query formulation. The results of a query are displayed both as a collection of pages in a virtual book: each document being represented with a content page and a page of links to corresponding categories.

Opening the book to a specific document page causes the related categories to be highlighted in the category tree. The cone-tree visualisation allows simultaneous disply of multiple hierarchies in the category tree. This makes it possible to simultaneously visualise multiple document-category relations thus allowing the user to see how a specific document relates to different thematic contexts. Furthermore, the user can view different concepts at different level of description, e.g. more familiar concepts in more detail and less familiar at a more general level (Hearst, 1999). The book containing the search results can be saved for later use and users can also assign documents to categories by themselves, hence expressing the insights discovered in the search process.

4.5.6. Interfaces for Sensemaking-Support

A special class of information workspaces is based on the paradigm of information access proposed by Pirolli & Card (1995) and the sensemaking model of Russel et al. (1993), discussed in detail in Chapter 3.4. This paradigm acknowledges the need to understand the process of information retrieval as a knowledge acquisition and sense-making process – i.e. a process in which people through their interaction with information develop and internalize new knowledge.

This has been addressed by information interfaces aiming specifically at providing sensemaking support. Such tools incorporate support for multiple perspectives and for interactive manipulation of criteria that determines how information is visualised and put in relation to each other. They aim at supporting the discovery and visualisation of contexts and relationships between information, as integral part of the process of information seeking and information access. For example, the SenseMaker (Baldonado&Winograd, 1997) provides interactive views of search results that can be organized along different dimensions and supports various operations on the results to accommodate the evolution of users' information need. The gIBIS system (Conklin & Begeman, 1998) supports collaborative argumentation and sensemaking in design problems by using a hypertext network to facilitate the discovery and visualisation of relationships between individual design issues, personal positions on an issue and related arguments.

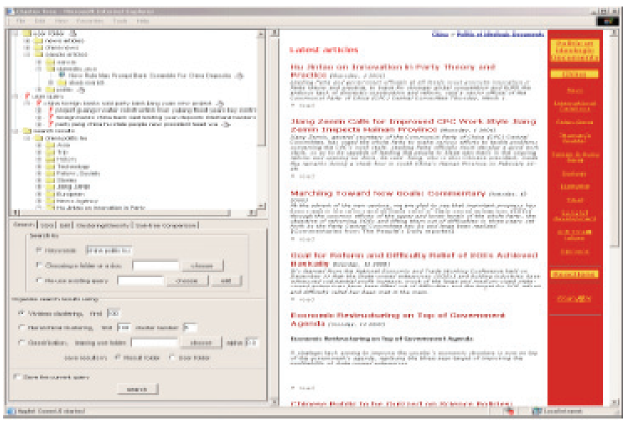


Fig. 4-45. User interface of the Sensemaking-Supporting Information Gathering System (Qu, 2003)

The Sensemaking-Supporting Information Gathering System (SSIGS) proposed by Qu (2003) most closely considers the requirements identified by the Russel et al. sensemaking model. It focuses on the role of external representation of implicit information structures and provides a workspace with integrated support for information search and for identifying and developing information representation schemes appropriate for a given task. The SSIGS interface (Fig. 4-45) allows the user to submit queries to web search engines, to organize them into personal hierarchical folders and to use automatic clustering of search results in order to identify possible structuring schemes for organizing the retrieved information.

The user can also submit a so-called “structured set of queries” which is defined by selecting a sub-set of user’s own information tree. The system then classifies retrieved documents into categories defined by the submitted template. A tree comparison tool further supports the user in discovering correlations between two different categorization schemes of the same document set. In this way the user is provided with a rich set of possibilities for exploring and organizing retrieved information from different points of view.

Similar functionalities are provided by the well-known eClassifier system (Cody et al., 2002) extended with visual clustering displays. eClassifier uses clustering to generate an initial hierarchical categorization of a large document collection which can then be refined by the user. Techniques such as multi-dimensional scaling and scatter plots provide similarity-based visualisations of the document collection and the underlying concept space. The collection can be explored from different viewpoints based on different taxonomies and several kinds of analytical summary reports are provided.

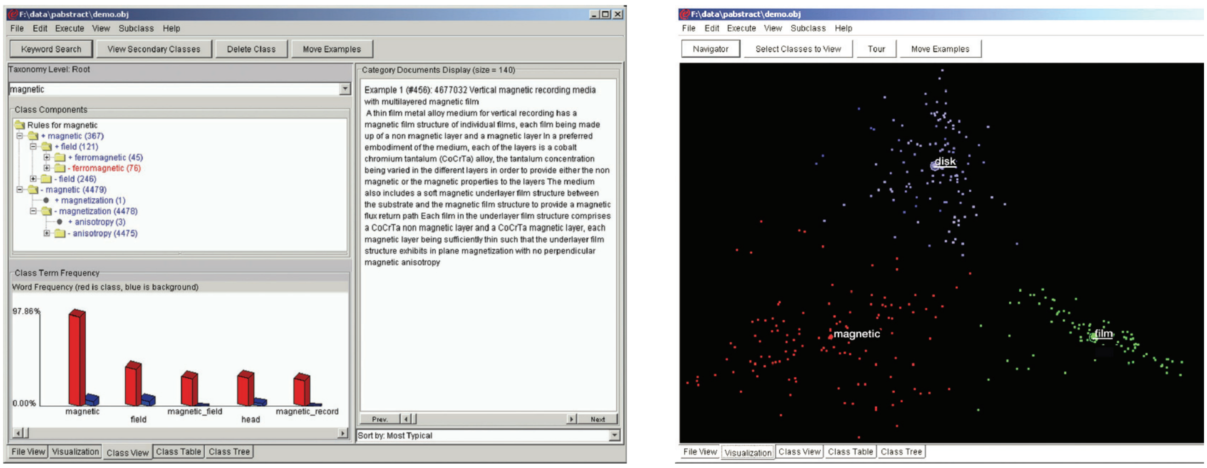


Fig. 4-46. The eClassifier system: hierarchical taxonomy and bar chart of term frequency (left), similarity-based scatter plot of the document collection (right).

In contrast to previously mentioned systems, e-Classifier is an even more analytically oriented solution. The main focus is on supporting specialist users such as analysts closely inspecting a document collection. In practice of general users this is not the case; though in sensemaking users face complex tasks they do not adopt such strongly analytical approaches. Sensemaking information access strategies typically include a combination of browsing and goal-directed search, with scanning, skimming and filtering out portions of the collection based on relevant concepts. But support for such tasks is expected to be available in a easy to use intuitive manner, not requiring the familiarity with specialized concepts and techniques such as clustering etc. The tools must provide shortcuts not involving the user into procedures requiring a comprehensive understanding of the underlying processing and visualisation models. For this reason tools such as eClassifier are appropriate for specialized users only i.e. analysts concerned with highly analytical tasks and familiar with the basic ideas of text analysis and word frequencies, hierarchical and interactive clustering, data analysis and visualisation techniques.

In summary, existing sensemaking-support interfaces facilitate contextualized access through the discovery and visualisation of different perspectives on information as an integral part of the process of information seeking. A significant critical aspect is that they provide complex analytical tools requiring significant level of expertise of the user and a rather analytical approach to accessing information. They also do not directly elicit implicit knowledge structures of human users, but support the user to become indirectly aware of implicit patterns and relationships between his information need and the information presented by the system. Another essential shortcoming in our application context is that existing approaches address the problem of sensemaking only at the level of the individual user. The social context of the user and the socially constructed nature of knowledge (communities) as well as the elicitation and application of collaborative knowledge structures of human users are not addressed. In contrast, the addressed problem and application context investigated in this thesis require us to consider support for developing an understanding of knowledge as social construct of a community of users, which is only partially reflected in the document collection.

4.5.7. Summative Discussion of Information Workspaces and Sensemaking-Support

The model of information workspaces provides a good point of departure in considering how to embed knowledge visualisation to support information access in community spaces in our problem context. The integration of explorative and goal-directed access supported by a combination of browsing and search, accompanied with visual search result contextualisation in thematic contexts has been shown to be well-received by users and helpful in accomplishing complex information tasks (Robertson et al., 1998), Russell et al., 2005). The idea of using direct manipulation and object-metaphors to provide cognitive shortcuts to more complex information seeking and processing services (e.g. classification, recommendation) certainly has great potential for reducing complexity in an already complex application context given in the access to unfamiliar knowledge domains represented by unfamiliar community information spaces.

In particular, the sensemaking-support interfaces relate closely to the requirements identified in Chapter 3.4 by facilitating contextualized access through the discovery and visualisation of different perspectives on information as an integral part of the information seeking process. However, the existing solutions cannot be readily applied to our problem. One weakness is that they provide complex analytical tools requiring significant level of expertise of the user and a rather analytical approach to accessing information (Russell et al., 2005). They also do not directly elicit implicit knowledge structures of human users, but support the user to become indirectly aware of implicit patterns and relationships between his information need and the information extracted from the documents.

Another essential shortcoming in our application context is that existing approaches address the problem of sensemaking only at the level of the individual user. The social context of the user and the socially constructed nature of knowledge (communities) as well as the elicitation and application of collaborative knowledge structures of human users are not addressed. In contrast, the addressed problem and application context investigated in this thesis require us to consider support for developing an understanding of knowledge as social construct of a community of users.

5. Outline and Goals of the Own Approach

Based on the theoretical analysis of knowledge processes in communities (Chapter 2), the specific requirements for cross-community knowledge exchange (Chapter 3) and the state-of-the art approaches in related fields and techniques (Chapter 4) this chapter summarizes the main idea and the goals of the proposed own approach.

The basic hypothesis is that cross-community knowledge exchange can be supported by dynamic knowledge maps which visualise implicit knowledge structures of individuals and communities of users and make them usable for contextualized access to information from unfamiliar community knowledge domains. Accordingly, the principal goals of this work are the development of a method for the construction of such dynamic knowledge maps (Chapter 6), development of a framework for its application to cross-community knowledge exchange (Chapter 7) and the prototypical realization in form of an interactive system and visual information interface (Chapter 8). This is accompanied by exploratory evaluation studies assessing the effectiveness of the developed solutions for supporting cross-community knowledge exchange.

Since the research subject has been little investigated in previous research and the presented work is one of the first contributions from the HCI perspective, this thesis presents inherently explorative research which integrates contributions from several different fields in order to identify important issues and propose ways for addressing them. Accordingly, the evaluation method is based on qualitative exploratory studies aiming at verifying the viability of the proposed approach, identifying its limitations and issues for further work.

5.1. Rationale of the Own Approach

As demonstrated by the analysis in Chapter 2 and Chapter 3, the core problem of cross-community exchanges lies in the absence of a shared social and cognitive context which makes it difficult to identify, understand and construct knowledge through accessing information from unfamiliar community spaces. Different through worlds and interpretive perspectives make it difficult for knowledge to cross community boundaries as information loses its original context and the recipients and authors do not share common vocabularies.

Existing solutions for knowledge sharing within communities do not consider this problem which strongly limits their applicability to supporting cross-community exchanges (Chapter 2.7, 2.8). The model of perspective making – perspective taking (Boland & Tenkasi, 1995) provides a theoretical basis for addressing the problem of supporting cross-community knowledge exchange (Chapter 3.1). Its analysis leads to the need for visualising knowledge structures of different communities and supporting the discovery of relationships between them. Contextualising this model within the knowledge management process model (Probst et al., 1997) identifies specific processes and application requirements for realizing such support in a way that is relevant for practical knowledge management needs (Chapter 3.3).

Finally, the work on sensemaking from the perspective of human-computer interaction emphasises the importance of external representations of conceptual structures that people develop in the process of making sense of unfamiliar information and when working on ill-structured problems (Chapter 3.4). Since cross-community knowledge exchange is largely mediated by accessing information from unfamiliar community information spaces in order to satisfy an ill-defined information need, the requirements for sensemaking support systems are of great relevance to our problem.

This concerns in particular the importance of making available personal knowledge structures of other people as a means for supporting information seeking (Qu & Furnas, 2005). Finally the comparative studies of explorative access vs. goal-directed search show that visual information tools can offer important benefits for supporting information access when the information need is ill-defined or unclear

(Becks, 2001). The interrelation of these findings suggests the need to incorporate the elicitation of and access to personal and community knowledge structures into visual tools for explorative information access.

The approach proposed in this thesis focuses on the possibilities for supporting cross-community knowledge exchange by providing ways for visualising knowledge structures of different communities and using them as tools for explorative access to community information spaces. More specifically, to develop a practically feasible solution for cross-community knowledge exchange three main sets of requirements need to be considered:

- Since communities rarely engage into creating explicit representations of their shared knowledge, a method for capturing and visualising implicit knowledge structures of groups of users needs to be developed. Such structures should incorporate personal views of individual members, reflect shared community vocabularies and evolve with community use (Chapter 3.1.5).
- The resulting knowledge structures need to be made usable for supporting information access in unfamiliar domains. This requires them to be realized as a basic element of a visual information tool supporting a combination of explorative and goal-directed access to unstructured information spaces. Such a tool should provide multi-perspective views on shared information spaces, support the contextualization of information from different perspectives and enable the discovery of relationships between them (Chapter 3.1.5, 3.4).
- The main application context for supporting cross-community knowledge exchange is given by knowledge management needs in knowledge-intensive organizations and heterogeneous knowledge networks (Chapter 2.8, Chapter 3.2). In order to ensure the relevance of the above methods to such application contexts they need to support the knowledge management process blocks relevant for the cross-community context (cf. Chapter 3.3). This includes providing mechanisms for fostering the willingness of different communities to make their knowledge available to members of other communities and supporting the contextualization of unfamiliar information with respect to the needs of a given task.

As we have seen in the analysis of related work existing approaches do not address such requirements (Chapter 2.7, 4). Their main shortcomings are summarized in the next sections.

5.1.1. Capturing, Representing and Visualising Implicit Knowledge Structures of Individuals and Communities

The first problem is defined by the highly implicit nature of shared knowledge structures of communities. Communities rarely engage into explicit formulation of shared taxonomies or ontologies representing their knowledge. Furthermore, the shared community perspective is not monolithic and unique: it is represented by a number of different perspectives that are related to each other and based on a shared common ground (Chapter 2). Community information spaces are commonly unstructured repositories of many isolated but interrelated information artefacts (documents, emails, newsletters, forum postings).

Approaches to knowledge sharing are based on processes appropriate within communities (e.g. internalization, socialization) and do not consider specifics requirements of cross-community knowledge exchange. Although recommender systems and collaborative filtering support information access based on implicit sharing of personal knowledge, they do not capture nor visualise the underlying knowledge structures of individuals and groups of users.

Furthermore, the lack of feedback about the criteria on which such recommendations are based has been diagnosed as the main shortcoming of such approaches (Herlocker et al., 2000). As a result, the straightforward use of such methods to provide cross-community recommendations (e.g. Glance et al., 1998, Merali & Davies, 2001) is of very limited practicability. It doesn't satisfy any of the above requirements and provides no means for members of different communities to develop an understanding of each other's perspectives, within which the meaning of respective information items is defined. The same applies to social data mining approaches that which provide more unobtrusive ways of inferring personal points of view on semantic relationships between information based on users analysis of information usage and patters of navigation (Chapter 4.3.4).

But we cannot straightforwardly employ neither the existing methods for automatic extraction of concept networks and ontologies from texts (Chapter 4.4). Their application to overall community spaces will only provide a limited set of (mostly single-word) concepts and it will not reflect the personal, implicit knowledge of individual members. At the same time, taking only specific personal collections raises the problems of the choice of personal document set to be inspected and its relation to the cross-community context.

Existing methods for collaborative elicitation and visualisation of implicit knowledge structures (cognitive mapping, concept mapping and repertoire grids) require great additional effort on behalf of the users and have not been applied as tools for information access. They are largely conceived as artefacts supporting goal-directed analysis and reflection for problem-solving in individual or group sessions. Finally, collaborative annotation and rating schemes have historically failed at producing shared structures even within communities, due to the cold-start problem and low user acceptance, based on the lack of immediate benefits for compensating the extra effort[41] (e.g. Fuhr & Fuchs-Kittowski, 2004).

5.1.2. Supporting Multi-Perspective Access to Community Spaces

The second problem concerns the discovery of relationships between knowledge structures of members from different communities and its application to sensemaking during information access. On one hand, the access to and use knowledge structures from different communities needs to be made an integral part of the process of information seeking. In particular, in order to support the interplay of perspective taking and perspective making during information access, we need to provide a way for dynamically contextualizing information from different points of view.

This requires close integration of knowledge organisation and visualisation models with methods for personalised categorization and classification of information, that can reflect perspectives of individuals and groups of users. As we have seen in Chapter 2.7.5 and Chapter 4.4, existing methods either focus on the discovery of mappings between different formally expressed conceptualisations (e.g. ontology mapping, see Madhavan, 2002) or employ statistical text analysis in order to discover related concepts in less structured representations such as taxonomies and document catalogues (Lacher, 2003).

But they do not consider the need for human users to understand the underlying criteria of their creation and the meaning of the discovered relationships. In consequence, they facilitate the identification of potentially relevant topics but lack the support for constructing new knowledge by understanding and expressing the meaning of different views and relationships between them.

On the other hand, as demonstrated by contributions such as (Becks, 2001) supporting strongly explorative access to information is very important in ill-defined situations, where the information need is highly unclear, broad in scope or cannot be formulated in an appropriate way (e.g. terminology differences between communities in our case). In particular, interactive manipulation of criteria determining relationships between unknown information and different knowledge structures becomes a critical means for discovering the appropriate contexts of meaning .

This is exactly the situation at the core of the problem of cross-community knowledge exchange. The information need is very unclear and cannot be appropriately formulated. The available information is

[41] The imbalance between required effort and perceived benefit is a classical problem of collaborative systems causing problems in adoption of existing technologies based on collaborative filtering, ranking and annotation of information (e.g. Herlocker et al., 2000; Resnick & Varian, 1997).

heterogeneous and underlying interpretive perspectives determining the meaning of information are highly implicit. This makes it difficult to represent them in complete formal models. Rather, in order to support perspective taking in such heterogeneous settings, formal models need to be replaced with lightweight conceptual structures that can be visualised intuitively.

Methods for semantic structuring of document spaces (Chapter 4.2) and methods for knowledge visualisation and representation (Chapters 4.3, 4.4) thus need to be complemented with interactive tools allowing the users to visualise and actively manipulate the extracted structures, in order to discover possible relationships between different community perspectives. As an assistive technology, such tools must actively involve the users in the process of identifying relevant knowledge, discovery relationships and constructing new knowledge. This can also serve as a means for compensating the limits of existing methods for automatic extraction of semantically related information that work well in homogeneous collections but not in heterogeneous situations (Wurst & Novak, 2004).

Finally, the findings of Boland & Tenkasi (1995) point to the need to support perspective taking and perspective making as an integrated process. To achieve this, access to knowledge structures of unfamiliar communities needs to be integrated into modalities of information access which are congenial with processes of intra-community exchange of knowledge. This includes the need to simultaneously support the ability for expressing personal views reflecting knowledge developed during the information seeking process and relating it to shared community structures.

5.2. Basic Idea

As the point of departure for the own approach, we adopt the knowledge map metaphor as a visual structuring of information in a way which provides insight into contexts and relationships between semantically related items. As we have seen in Section 4.4.4 knowledge maps are commonly used as static conceptual structures providing an overview of available knowledge resources in a given domain or facilitating navigation in complex information spaces.

Most solutions provide only one, unified knowledge structure to which all users and communities must adhere in order to access the shared information space and to make their knowledge accessible by others. The connection of personal knowledge structures with the shared taxonomy or ontology must be created manually by the users through assigning the information to the centrally defined concepts (Bonifacio et al., 2000). While personal knowledge organization takes place naturally as a result of any problem solving or information seeking, assigning documents to another scheme requires additional effort that users are often not willing to undertake (Lacher, 2003).

This work proposes to address this by developing a method for the construction of dynamic knowledge maps that visualise implicit knowledge structures of individuals and communities and can be applied as visual information tools for contextualized access to unfamiliar information spaces. The main hypothesis is that using such knowledge maps to provide access to shared information spaces from different points of view, representing personal and shared perspectives of different communities, can provide valuable support for cross-community discovery and sharing of knowledge.

The crucial difference to other approaches from knowledge management and knowledge visualisation is that the proposed method aims at enabling the construction and visualisation of knowledge maps which reflect implicit knowledge structures of individuals and communities of users. The implicit nature of socially constructed knowledge in communities suggests that such knowledge maps cannot be static representations which “codify” knowledge.

Rather they need to be conceived as interactive visual artefacts that can be manipulated by the community members in order to get an understanding of different interpretative schemas underlying different communities and how they are related to one’s own knowledge: e.g. by exploring maps of different users, applying them as templates for contextualising unknown information or comparing one’s own personal concept structures to those of others.

The idea of using such maps to support the exchange of knowledge between different communities is to apply them as a means for “seeing through the eyes of the other” (metaphorically speaking). The elicited knowledge structures of individuals or groups of users should allow a given user to apply them for structuring information according to concepts representing a particular point of view. In this way the user should be able to take on the perspective of members from other communities and discover relationships to his own structure in order to develop an understanding of the meaning of different concepts. In order to support the sensemaking paradigm, the knowledge maps that reflect specific perspectives of different communities or their members, should be applicable as templates for organizing information into meaningful contexts during access to unfamiliar information spaces.

Finally, such dynamic knowledge maps should also allow the user to express the newly developed understanding by relating the discovered information to concepts from his own perspective. In this way the maps would support an interplay between perspective making and perspective taking while being embedded into the sensemaking process occurring during information access. Accordingly, developing a method for constructing and using knowledge maps in such a way promises to satisfy the requirements identified in Chapter 3 and thus provide valuable support for cross-community sharing and creation of knowledge.

To construct knowledge maps in a way which incorporates personal knowledge structures of individuals and groups of users we consider the process of information access as a context in which users actions reflect their interpretation of the meaning of information. Such an activity is bookmarking (Amento et al., 2003): when accessing information spaces people select relevant information, organize it into groups of related items and label the groups with concepts representing their meaning. The resulting structures of the relationships between document groups and labels which have been assigned to them reflect the existing personal knowledge of the users and the new knowledge created in the information seeking process. Considering the structures created by a group of users provides us with a basis for discovering common concepts and connections between them, which characterize a given community and its domain of knowledge.

According to this approach the resulting knowledge maps would reflect personal viewpoints of a user or a specific community with the ability to generate personalised semantic perspectives on an information space. Hence, they are “personalised maps”. Since they are not static but can evolve based on user actions i.e. “learn” his point of view based on his interaction with information they are ”learning maps”. These two important aspects of our approach are referred to with the notion of personalised learning knowledge maps. Finally, since the maps are based on patterns of collaborative information structuring revealed from personal structures of a community of users and are used to support exchange of knowledge this is a way of collaborative knowledge visualisation.

Another major task is the design of an interactive system and interface that enables the use of such maps for information access to shared community repositories. In order to make the maps usable for cross-community knowledge exchange this system needs to support the interplay of perspective taking and perspective making (Chapter 3.1.5).

This calls for developing a multi-perspective visualisation model and an intuitive interface enabling interactive exploration, manipulation and comparison of different semantic perspectives on an information space. In particular, this concerns the discovery of relationships between unfamiliar perspectives and one’s own, familiar knowledge structures. To this end results of user’s actions need to be automatically contextualized within both perspectives: for example, displaying search results or navigating unfamiliar concepts should be related to known concepts from one’s own structure. In this way the understanding of unknown knowledge contexts their relationship to one’s own knowledge could be supported.

5.3. Goals of This Work

The discussion in the previous chapters as summarized in Section 5.1 has shown the limitations of existing methods and the main idea of the proposed approach based on a developing a method for construction and visualisation of knowledge maps reflecting personal and shared community knowledge structure and their application to supporting cross-community knowledge exchange.

Therefore, the main goals of the approach presented in this thesis are:

- development of a method for the construction of dynamic knowledge maps which visualise implicit knowledge structures of individuals and communities and make them usable for personalised structuring and contextualization of information (Chapter 6)
- development of a framework for applying the developed method for collaborative elicitation and visualisation of personal and shared knowledge maps to supporting cross-community knowledge exchange (Chapter 7).
- development of an interactive visual information interface supporting multi-perspective access to community information spaces (Chapter 8)
- exploratory evaluation studies assessing the suitability and effectiveness of the application of the developed method and interactive system to support cross-community knowledge exchange (Chapter 9).

6. Eliciting and Visualising Implicit Structures of Personal and Community Knowledge

This chapter introduces the developed method for the construction of dynamic knowledge maps that visualise implicit knowledge structures of individuals and communities of users and can be applied as visual information tools for contextualized access to unfamiliar domains of knowledge.

The method integrates well-known techniques from visual document clustering (Lin et al., 1991; Honkela et al,, 1997) and extraction of concept networks from texts (Honkela, 1997) with collaborative filtering (Resnick, 1994) and personalised classification based on user-induced templates (Aha et al., 1991).

Such a combination of unsupervised knowledge discovery methods with techniques involving the knowledge of human users allows the elicitation of semantic structures unavailable in other approaches. The application of visualisation techniques for representing similarity structures and relationships in complex information spaces makes the extracted structures usable for contextualized access to information from different points of view.

This integration is the main difference and special added value to other approaches from knowledge discovery and knowledge visualisation.

It allows us to elicit and visualise implicit knowledge structures of individuals and groups of users in a way which makes them usable for collaborative discovery and sharing of knowledge. It also solves the cold-start and free-rider problems of other approaches.

The overall structure of the developed method for achieving this is depicted in Fig. 6-1. The developed approach and different aspects of its realization have been presented in (Novak et al. 2002, 2003a, 2003b, 2003c, 2005a).

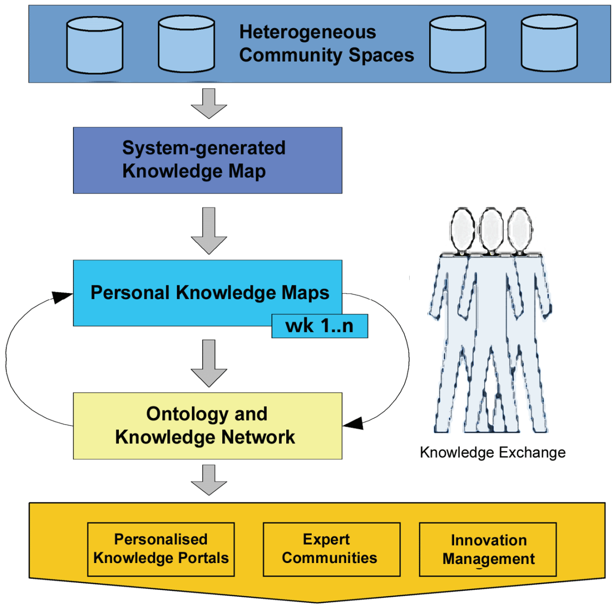


Fig. 6-1. The overall model for eliciting personal and community Knowledge Maps.

This chapter is organized as follows: First, the basic concept of the overall solution approach is introduced (section 6.1). This is followed by the description of the specific methods through which the main elements of the solution model have been realized (section 6.2-6.4). The ontological structure of the developed knowledge map model is presented in Section 6.5, while Section 6.6. summarizes how the presented method satisfies the identified requirements for cross-community knowledge exchange.

6.1. Basic Concept

As outlined in the previous chapter, the main idea for constructing knowledge maps that elicit and visualise implicit knowledge structures of individuals and groups of users is to consider how users' interaction with information reveals their personal, implicit knowledge structures. This approach is related to the sensemaking paradigm that views information seeking as a process in which users actively construct and develop knowledge through their interaction and use of information (Chapter 3.4). In particular, the users' actions in selecting, labelling and organizing items into meaningful structures reflect their interpretation of the meaning of information, their existing personal knowledge and the evolvement of new knowledge in the process.

But in order to develop a practically feasible solution for capturing and visualising implicit knowledge structures of human users based on their interaction with information, several difficult problems need to be addressed:

- First, an initial context for user actions has to be created in order to solve the cold-start problem and to relate user actions to a shared context. The lack of a clear interaction context is the main difficulty of general "user-tracking" and interaction-mining approaches.
- Second, a form of visual representation has to be found that communicates to the user both the semantics of the information space in itself (content, structure and relationships) and relates this to the meaning of his actions.
- Third, an appropriate method for calculating inter-document and inter-concept similarities based on user actions needs to be selected.
- Finally, an unobtrusive model for eliciting user actions is needed, such that the user's effort in using the system is balanced by an immediate personal benefit (e.g. discovering previously unavailable knowledge). The problem of the imbalance between user effort vs. perceived benefit of expressing information preferences is a well-known free-rider and cold-start problem in classical collaborative filtering and community rating approaches[42].

As a basic element in addressing these requirements we introduce a special knowledge map model aiming at providing both a visual representation of users' knowledge structures as well as a context within which the meaning of their interaction with information can be interpreted. In our approach a knowledge map is composed of two closely coupled elements: a Document Map and a Concept Map.

The Document Map (Fig. 6-2, left) presents information items from the community space (e.g. documents, forum postings) structured into clusters of semantically related objects. Its purpose is to provide an overview of the semantic structure of the community information space: the main topics, the distribution of documents and relationships between them. The Concept Map (Fig. 6-2, right) displays groups of different words used in similar contexts and relationships between them. It aims at providing insight into the patterns of language use by a given user or community and into the implicit criteria determining the semantic structure of the Document Map.

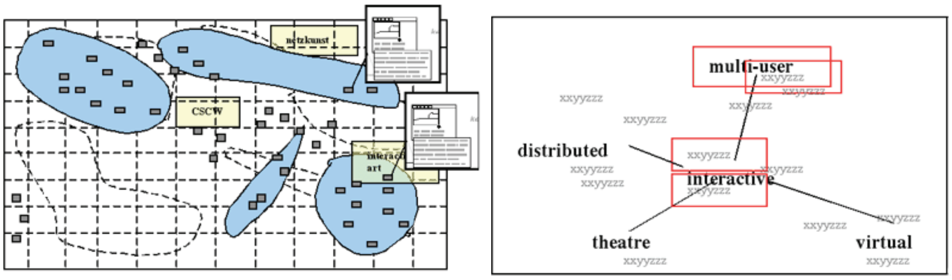


Fig. 6-2 Schematic illustration of the proposed Knowledge Map model

Such a knowledge map model allows us to both present a specific semantic structuring of an information space to the user (Document Map) and to provide an "explanation" of the meaning of this structure - by displaying the main concepts and relationships between them, which describe a specific personal or shared perspective (Concept Map).

To construct such maps based on user interaction with information and hence to incorporate implicit knowledge revealed by user actions, the following method is proposed:

- *System-generated overview:* First, the user is presented with a system-generated structure that provides an initial context for accessing the community information space (e.g by means of

[42] For example in (Resnick, 1994) or (Graether & Prinz, 2001). See also (Herlocker et al., 2000) for an overview of critical issues in collaborative filtering.

methods for unsupervised document clustering). By providing an overview of the main topics, groups of semantically related documents and the concepts describing community knowledge, the map serves as a guide for users' exploration. In doing so, the user discovers information items corresponding to his interest, develops a personal interpretation of their meaning and their relationships to the context of his information need.

- *Expressing personal points of view:* The user can now express this personal view by creating a personal map which allows him to re-arrange the system-generated structuring as part of his normal access to information (e.g. by selecting relevant information items, moving objects between clusters, renaming or creating new clusters). By embedding the possibility for expressing a personal structure in the natural flow of actions performed by users in the information seeking process (e.g. bookmarking) we realize an unobtrusive way of eliciting user knowledge. As a result, by performing natural, non-distracting actions that are embedded into his primary task (finding relevant information), the user creates a personal Document Map that reflects his personal point of view and the insights he discovered and internalized as knowledge.
- *Learning the user-defined structure:* The results of user actions can now be taken as a point of reference for creating a dynamic knowledge structure that reflects the user's personal point of view. The user-defined map is learned by the system (e.g. employing supervised learning methods) and formalized in a way which allows it to be applied as a user-defined „template" for semantic structuring of an arbitrary information space from a specific point of view. Furthermore, all Document Maps belonging to the user are analysed in order to extract implicit relationships between different concepts used by the user. To this end, the user-defined concepts (cluster labels) from all personal maps are put in relation to most relevant terms appearing in the corresponding clusters. Different concepts describing similar groups of documents are also related to each other. This results in a personal Concept Map representing the most important concepts for a given user and the relationships between them. In this way the implicit knowledge of a given user that has been revealed through his actions on the initial semantic structure has been acquired, represented and visualised in form of a personalised knowledge map that can be used for contextualized access to information reflecting a user's personal point of view.
- *Aggregating personal views into shared community structures:* By extending the described analysis to maps of all users from a given community, a shared concept map representing a specific community perspective can be created. The cluster names used by members of a community will describe the most important concepts for the community in question. The association of concepts to documents illustrates the meaning of the concepts by providing concrete examples of the contexts in which a given concept is used. The degree of relatedness between different concepts is inferred from the similarity of different clusters. This is achieved by a combination of a text-based similarity measure with a measure considering co-occurrences of documents in similar contexts. Thus, personal views are connected into a shared community-specific conceptual structure that can be used both for navigating the community information space and for gaining an understanding of the shared community vocabulary.

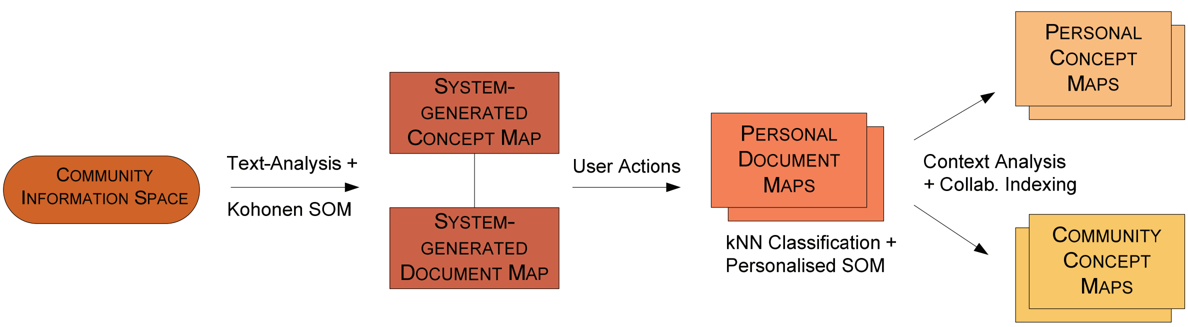


Fig. 6-3. Overall diagramme of the proposed method for personal and community knowledge maps

The diagramme in Fig. 6-3 depicts the overall structure of the proposed method. The basic assumption of this solution is that by creating personal maps users express their knowledge about a specific topic, domain or information space. The reflected knowledge is of two-kinds. On one hand, personal maps reflect the knowledge of a given user about the mutual relevance or similarity of documents. On the other hand, grouping documents in named clusters reflects the knowledge about document-concept relationships. Since, the process of creating the personal maps is not the primary goal of the user, but a natural activity supporting his primary task (information search), the reflected knowledge is of implicit nature.

Rather than being concerned with “explaining” his knowledge to someone, the user creates maps as kind of mental notes: they serve as memory aids in the search process and externalize the important insights the user has discovered regarding the relationship of identified information to his given need (sensemaking paradigm, Chapter 3.4). At the same time, the decisions about which documents are grouped together and the concepts used by the user to assign cluster names reflects also the users background knowledge.

In this sense, it can be said that personal maps reflect implicit knowledge structures of the user as they are revealed in his information seeking process and interaction with relevant information items. Accordingly, analysing and aggregating the personal maps of groups of users in the proposed way, can reveal implicit knowledge structures characterizing of that group. Applying this principle to users from a same community thus allows us to reveal implicit knowledge structure of a community in question, as it is reflected in the information seeking behaviour of its members. Extending the analyses across community boundaries provides a means to discover otherwise unavailable cross-community structures.

6.2. Realization Method

According to the described solution approach, knowledge maps are constructed at several different levels in order to support different stages and aspects of the knowledge exchange process (Chapter 3.1, 3.3). This includes:

- Automatically generating knowledge maps that present overviews of the implicit semantic structure of community information spaces,
- Constructing knowledge maps that represent personal points of view of individual users (based on user interaction with information),
- Constructing knowledge maps that represent implicit knowledge structures of different communities (based on the aggregation of members’ personal views),

Developing a concrete realisation of the described model involves integrating and extending methods from several fields. Overall, for the automatic construction of system-generated maps methods for visual clustering of document collections such as (Lin et al., 1991; Honkela et al., 1997) provide a good point of reference. For creating personal points of view based on user-defined templates, methods for supervised learning (e.g. Joachims, 1998) or personalised classification (e.g. Aha et al., 1991) offer possible solutions.

Extracting community-specific conceptual structures and relationships between them based on the analysis of personal user templates, requires the use of context-based similarity methods for determining inter-document and inter-concept relationships (e.g. Resnick et al., 1994). For intuitive visualisation and interaction we can employ methods for visualising complex information spaces such as two-dimensional cluster maps, spring-based graph layouts or hierarchical trees (Card, Mackinlay, Shneidermann, 1999b).

The next sections consider how these methods have been applied and combined with each other, in order to realize the specific aspects of the presented overall solution method.

6.3. Generating Overviews of Implicit Semantic Structure of Community Spaces

The initial knowledge map presenting an overview of the implicit semantic structure of a community information space is constructed by combining methods for self-organizing document maps (Lin et al., 1991, Honkela et al., 1997) with a related method for extracting concept maps from texts (Honkela, 1997). The first allows us to generate the system-generated document map while the second is used for producing the concept map element of our knowledge map model. Both of them are based on the Kohonen's self-organizing map (SOM), an unsupervised neural network which performs a topology preserving mapping of high-dimensional vectors to a two-dimensional grid (Kohonen, 1990, 1995).

The main idea of applying the Kohonen SOM to an information space is to generate a two-dimensional document map that presents a semantic structuring by grouping related documents close to each other. Since the SOM maps a high dimensional space on a two-dimensional grid while preserving topological structure, applying the SOM to a document set will result in the distribution of documents on a two-dimensional grid with similar documents lying close to each other. This is a frequently used method of generating document maps (Chapter 4.2), since the two-dimensional map generated by the SOM readily lends itself as the basis for an intuitive and visually appealing visualisation of the implicit semantic structure of an information space. Fig. 6-4 illustrates the process of using the Kohonen SOM to generate document maps.

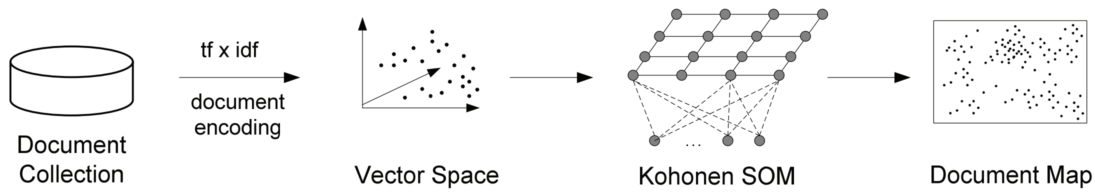


Fig. 6-4. The principle of using the Kohonen SOM to generate document maps

The distribution of the document set on a 2D grid provided by the SOM can be enriched with information about clusters of especially related documents and the main concepts describing their content (e.g. Lin et al., 1991). Since similar clusters will be found close to each other, similarity relationships between related concepts describing the clusters content can also be established. In this way, an initial overview of the structure of shared community knowledge, as it is reflected in the documents created by community members can be provided. The successful use of document maps based on the Kohonen SOM for explorative information access and for gaining insight into the semantic structure of document collections has been demonstrated by studies such as (Becks, 2001). Using such a map the user can get a quick impression of what topics can be found in the community space, see which ones are related to each other based on spatial proximity as well as distinguish their relative importance based on the size of the corresponding clusters.

6.3.1. Basic Method for Generating Document Maps based on Kohonen SOM

The basic method for generating document maps based on the SOM consists of following steps:

1. *Encode documents in a vector-space model*: In order to pass them as input to the Kohonen SOM documents need to be translated into vectors representing their semantic properties. To this end the document set is represented in a vector-space model (Salton et al., 1994). A set of terms (words) is selected as dimensions along which the document space shall be represented. This term set is selected automatically by removing common words (stop words) and eliminating the most and least frequent words. In order to additionally reduce the number of dimensions, the words can be reduced to their stem which produces terms that represent groups of words sharing the same stem. Each document is then described by a term vector whose components represent the significance of each term for the document in question. The significance of each term is calculated based on the frequency of its occurrence in the given document (ibid.). Such term

frequency (TF) is often corrected by an inverse document frequency (IDF) which calculates the inverse of the number of all documents in which a term occurs.

2. *Train the SOM with the generated document vectors:* The document vectors are then provided as inputs for training the SOM. In this process the map repeatedly maps similar term vectors to neighbouring neurons on the two-dimensional grid, and adjusts the term weights of the neuron correspondingly. As a result of the training a two-dimensional grid of neurons has been created, where each neuron is represented by a characteristic weight vector. The topological properties of the document space (similarity relationships) are encoded in the weight vectors of each neuron.

3. *Map the document vectors on the resulting network:* In order to produce a visible map of the document collection, the document vectors are again applied to the SOM which now maps each document vector onto the most similar neuron. Since neighbouring neurons have similar weight vectors, similar documents will be mapped on neurons that are close to each other. In this way, a two-dimensional map of documents is created where spatial distance reflects the relative degree of semantic similarity based on the document contents. Fig. 6-5 shows the basic model of the SOM neural network and an example of the distribution of documents on a two dimensional grid produced by it.

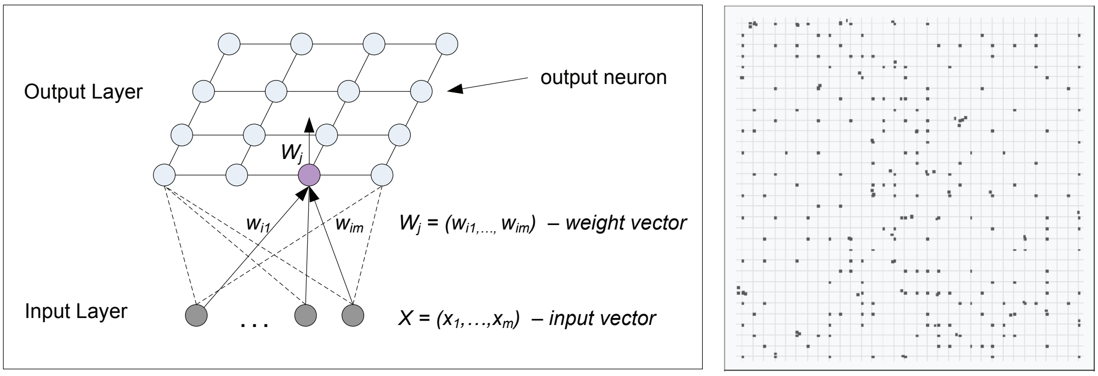


Fig. 6-5 Architecture of the SOM and example of the two-dimensional mapping produced by it.

4. *Visualise the structure of the resulting map:* A common method for visualising the resulting map is to display the 2D grid of the underlying neural network and assign each neuron cell a colour that reflects its average semantic distance from the neighbouring neurons. This is the so-called U-Matrix method (Ultsch, 1993). Using light shades of gray for large and dark shades of gray for small distances to neighboring units reveals boundaries between clusters of related documents (Fig. 6-6, left). Another approach is to determine the clusters explicitly by labelling each neuron with its highest-weighted term. All neurons with the same label are then grouped together into one cluster. The resulting set of labels then represents the categories reflecting the main topics found of the document collection (Fig. 6-6, right).

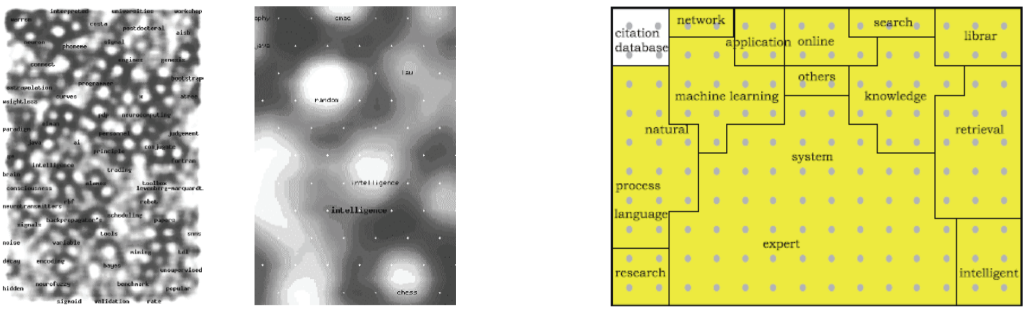


Fig. 6-6. Different visualisation methods for document maps based on the Kohonen SOM: U-matrix visualisation used in WEBSOM[43] (left) and category map from (Lin et al., 1991, right).

[43] http://websom.hut.fi/websom/

6.3.2. Separating Visualisation and Clustering for Interactive Document Maps

In our approach, constructing Document Maps of community information spaces aims at presenting the implicit structure of community knowledge. In order to apply the Kohonen SOM for generating such maps we need to consider two special needs. First, the map needs to provide a basis for realizing an intuitive but powerful visual information tool (see requirements in Chapter 3 and Chapter 5). It must allow users to obtain insights into the implicit knowledge structure of a given community as it is reflected in its document space. To this end it needs to present the main topics of the community knowledge, the distribution of documents within them and the relative relationships of their importance to the community. Since the maps have to be embedded as visual information tools into the process of information seeking, they need to present this information in an intuitive and clearly visible way.

Second, the map needs to be realized as an interactive artefact that can be manipulated by the users. This means not only using it as a tool for explorative access and navigation of an information space but also influencing the criteria determining its behaviour. Such criteria include the granularity of the map (number of clusters) and the most important terms used to encode the document vectors (vector space dimensions). Finally, being able to explicitly determine the number of clusters is also important for the design of a usable interface. There are both visual perception and cognitive limits on the number of categories that can be easily perceived by the users. This becomes even more critical when the visual display contains different layers of information. The design of the map solution should reflect these concerns.

In order to satisfy these criteria we modified the method presented in Section 3.6.2 in a way which allows us to independently influence the clustering and the visualisation aspects. The modified method allows us to clearly present important semantic information such as the number of clusters, the main topics and the criteria used for the clustering.

In the approach described in the previous section the SOM is used as a visualisation method that distributes a document collection on a 2D map such that similar documents are mapped to neighbouring positions. Determining groups of similar documents (clustering) is a side effect of this mapping: clusters are formed by similar documents being mapped to the same neuron or by related documents being spread across a group of neurons with similar characteristic vectors. The problem of such implicit clustering is that for maps with many neurons (which is usually the case) we get too many clusters.

As a solution to this problem, in (Lin et al., 1991) all neurons are first labelled and then all neurons with the same label collected into one cluster. But this clustering method is difficult to influence: neither the system designer nor the end-user can define the categories and the number of clusters. The number of clusters is still intrinsically connected to the resolution of the SOM grid. Decreasing the resolution will produce a lower number clusters, but this will also reduce the granularity of the 2D mapping. Many more documents will be mapped onto the same neuron, reducing the precision with which they can be distributed on a two-dimensional surface. In consequence, the semantic relationships between documents and groups of documents can be less clearly expressed and visualised.

To overcome these limitations, we introduce a method that decouples the two-dimensional similarity visualisation from the clustering. The SOM is used only as a visualisation method for mapping a high-dimensional document space to a two-dimensional grid with preservation of relative semantic distance. Clustering is obtained in parallel to the SOM mapping which offers the possibility for interactive clustering without loosing the precision of the SOM visualisation. This method has been published in (Simunic & Novak, 2005c).

The concrete procedure for constructing the system-generated Document Map based on this method is depicted in Fig. 6-7. It consists of two main phases:

- In the first phase most important categories describing the contents of the document pool are identified by selecting N most important terms (words) from all documents (words are stemmed and the stop words, least and most frequent words are eliminated beforehand). These terms represent labels that will be used to categorize documents into clusters. Initially, they are selected automatically but they can also be influenced by the users. After vectorization the documents are mapped on a two-dimensional grid by means of the Kohonen SOM.

- In the second phase, the documents mapped on the SOM are assigned to the best matching category. This occurs by comparing neuron vectors with unit vectors representing each label and assigning the unit vector with least distance as the neuron label. All neurons with the same label are then grouped together to form a cluster category represented with that label.

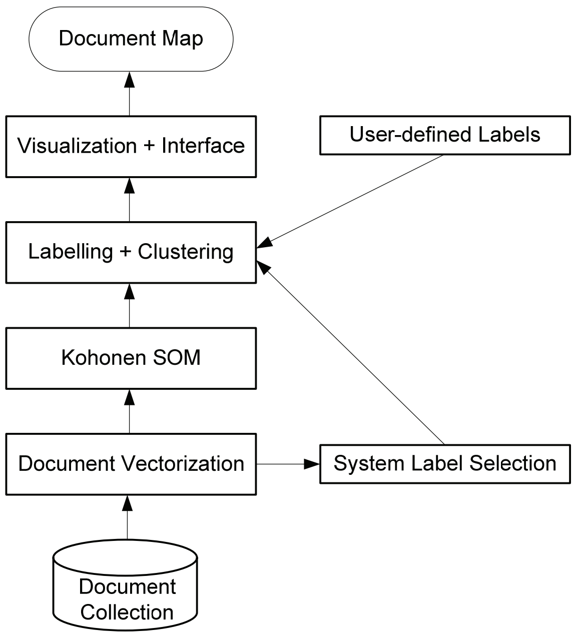


Fig. 6-7 Method for constructing the system-generated Document Map

In other words, the SOM is used only to position the documents on a two-dimensional grid in a way which preserves their similarity relationships and the topological properties document space. Similar documents will be placed close to each other and areas with higher document density in the high-dimensional space will occupy larger areas of the two-dimensional map as well. The labelling is then used to cluster documents in groups based on the best matching label from the set of most important terms identified in step one. Since similar documents are positioned close together the resulting groups of documents with the same label will represent cluster areas on the map[44]. Fig. 6-8 and Fig. 6-11 in the next section show two different visualisations of maps created by this method.

The described separation of two-dimensional mapping for visualisation and of clustering allows us to independently influence the resolution of the Kohonen SOM and the desired number of clusters. This provides us with two important benefits. The first is that we can set the SOM parameters in the way most suitable for the needs of the visualisation (e.g. adjusting the number of neurons such that on average one document is mapped on one neuron). We can set a maximal number of clusters such that the resulting map provides a sufficient level of granularity while considering the limitations of the number of categories that can be simultaneously visually perceived and cognitively processed by the users.

The other is that we can add functionalities that allow users to interactively influence the parameters of the clustering (e.g. the number of clusters or the most important labels) while using the map as an interface for accessing the community information space. Such a possibility to interactively manipulate map criteria for discovering relevant topics, items and their relationship to one's information need is an important requirement when accessing information in unfamiliar contexts (cf. sensemaking, Chapter 3.4).

[44] When clusters split across different areas on the map occur, they are divided into separate clusters with the same label.

6.3.3. Visualising the Document Map

There are several problems to readily applying existing visualisation models for Kohonen SOMs to Document Maps intended as an information interfaces to community information spaces. In purely visual methods (e.g. Becks, 2001; Honkela et al., 1997, Kohonen et al., 2000) the number of clusters is very large and difficult to control (see previous section). The focus is on visually representing the degree of (dis)similarity between the neighbouring neurons as a means for identifying possible clusters of related documents visually (e.g. WEBSOM in Fig. 6-6). The characterization of the map content is supported by labelling neurons rather than specific document clusters (since they are not explicitly defined).

On the other hand, methods such as the category map (Fig. 6-6, right) present clearly delineated clusters labelled with the most representative term. However, this is only a very general description of the cluster contents. Providing additional labels to better characterize individual areas could provide a better overview. Furthermore, both methods display the structure of the underlying network rather than the the distribution of the actual document set. This may be of interest for special analysis tasks (Becks, 2001) but for realising an information access interface we need to display the distribution of concrete documents contained in a given collection[45]. In our application context, the focus of the user's interest is not the structure of the information space in its own right, but only as a means for supporting his understanding of the context within which the meaning of information he is seeking is defined.

In our application context, the map should support user access to information spaces of unfamiliar communities by providing insight into the structure of that community's knowledge reflected in its information space. The focus of the user's interest is not analysing the structure of the information space per se, but grasping the structure of an unfamiliar area in order to be better able to locate and understand which information is potentially relevant for his need. In other words, gaining insight into the structure of community knowledge represented by the map is only a means for supporting the user's understanding of the context within which the meaning of information he is seeking is defined. A well-known method for supporting orientation in unfamiliar spaces is the use of landmarks that support visual navigation (Dillon et al., 1990). Thus, visualising the structure of an information space represented by a Document Map should be realized in a way that offers clear visual landmarks for immediate orientation by the user.

In order to address these concerns we have developed two different basic visualisation models for displaying document maps which provide overviews of the semantic structure of community information spaces. They combine the most important elements of the approaches proposed in (Lin et al., 1991) and (Honkela et al., 1997), adapted to our application context. The first visualisation model is based on the idea of distinguishing between individual topical clusters as clearly as possible in order to support easy visual perception of the most salient semantic structure.

Documents are displayed as points on the two-dimensional grid[46] and they are grouped in specific topical clusters with the labels characterizing the cluster content. Cluster boundaries are drawn as convex and non-convex hulls around documents belonging to a cluster[47]. The corresponding cluster label is displayed below each cluster. Such clearly identifiable cluster areas and the associated labels serve as visual landmarks that can be immediately grasped by the user. Making the cluster hulls fit as closely as possible around the document points allows us to create the necessary white space for clearly displaying different cluster arrangements and avoiding visual clutter. Fig. 6-8 depicts an example map visualised with this method.

[45] Displaying the distribution of documents naturally poses problems with screen space for large document sets. See Chapter 8 for the implemented solution based on focus+context zoom.

[46] Being able to set the grid size in the map generation process independently of the maximum number of clusters allows us to choose the grid size corresponding to the number of documents to be displayed. This positions one document in each cell on average, hence limiting visual clutter.

[47] First convex hulls are calculated for all clusters and tested for convexness. For clusters failing the tests a non-convex hull is calculated.

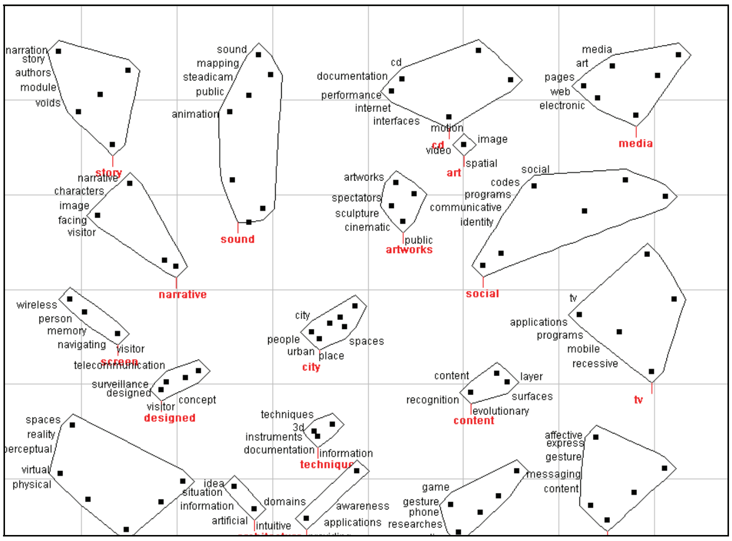


Fig. 6-8 Visualisation model for the Document Map in own approach (Version 1)

In addition, for each cluster a set of most representative keywords is calculated and positioned around the cluster outline (Fig. 6-8). Besides keywords extracted by the text-analysis a selection of keywords originally assigned to documents by their authors can also be displayed (in another colour). Providing such additional meta-information about cluster contents aims at better characterizing individual landmarks. Individual labels can hardly provide satisfactory overview in this case, especially when the users are unfamiliar with the terminology in question (e.g when accessing unfamiliar community spaces). Further levels of detail such as titles, document abstracts and full content details have been incorporated by different modalities of selection, filtering and zoom techniques. Fig. 6-9 illustrates this schematically. Since this concerns a specific interaction design for turning the document map into an intuitive visual information interface, a detailed description of these aspects is given in Chapter 8.

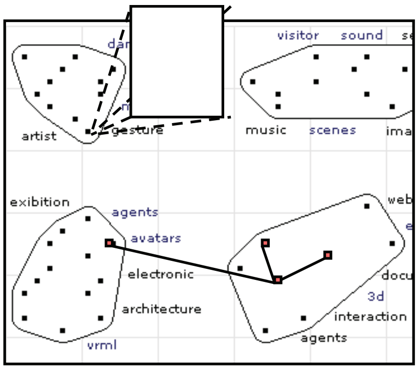


Fig. 6-9 Schematic illustration of different levels of detail for displaying the map meta-information

To ensure a good visual distribution, the number of neurons of the SOM is set to approximately equal the number of documents. In general, this results in each document being mapped on one neuron. Thus documents can be assigned to screen cells representing appropriate neurons straightforwardly by placing each document in the centre of the cell. If several documents fall on the same neuron (e.g. very large maps) the documents are positioned within the cell according to the procedure illustrated in Fig. 6-10.

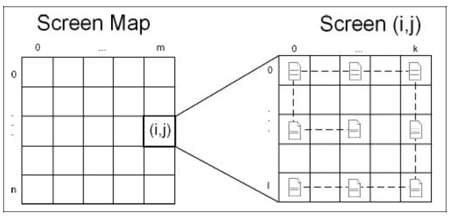


Fig. 6-10 Screen placement procedure for multiple documents assigned to the same neuron (Wever, 2005).

The second visualisation model has been developed as a result of user feedback in the formative evaluation (Chapter 8.5). It retains clearly identifiable clusters as important visual landmarks but softens the boundaries between them. The individual clusters should not be understood as clearly defined, separated entities but as possible suggestions. In particular, the neighbouring clusters often have much in common and the semantic transitions between them are rather fuzzy than clear. Especially interesting documents can often be found in map areas between clusters, where one cluster transits into another. Also, neighbouring clusters might represent two parts of a complex concept (e.g. “visualising knowledge” in Fig. 6-11). This has also been observed in (Lin et al., 1991). While acknowledging the importance of clear visual landmarks, the user feedback in the formative evaluation pointed to the need that the visualisation also better expresses this fuzziness of cluster boundaries and the interrelatedness between neighbouring clusters.

To achieve this, the first visualisation model in the following way. Instead of drawing clear cluster boundaries by polygonal shapes, the clusters are coloured with a linear decreasing fill from the cluster centre. The cluster centre is calculated as the centroid of the polygon formed by the boundaries of outer cells of the cluster. The intensity of the colour for each cell is then calculated based on the cell distance from the cluster centre. In this way, an intuitive and visually appealing presentation of individual clusters is achieved where borders are soft and possible interconnections between neighbouring clusters are perceivable. The display of additional cluster keywords has been removed from the overview level. Instead the cluster keywords are displayed when the user focus is directed to a specific area (for details see the discussion on interaction design in Chapter 8).

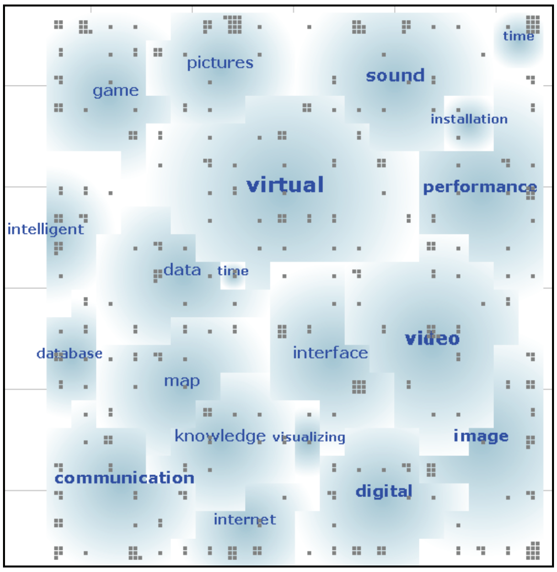


Fig. 6-11 Document Map visualisation model (Version 2). The example is based on the Document Map of the community information space of netzspannung.org[48].

The reason for not using the U-matrix method and colouring the cells based on their average distance form the neighbours is straightforward. Such a visualisation would suggest a different cluster structure than the structure produced by the labelling method. Hence, the visualisation would claim something else than the underlying method for generating the map in the first place (section 6.3.2). Adopting the centroid-distance measure as the basis for expressing the degree of relatedness of a given cell (and hence the documents positioned within it) to a given cluster allows us to balance two opposing requirements: on one hand, the grouping of documents into specific clusters needs to be clearly expressed, on the other hand the possible relationship of documents in areas closer to neighbouring clusters needs to be communicated as well.

[48] http://netzspannung.org

6.3.4. Generating the Initial Concept Map based on Kohonen SOM

An important requirement identified in the analysis of the problem of supporting cross-community knowledge exchange is the need to provide users with insight into shared vocabularies of unfamiliar communities. The sensemaking model further requests that this be provided as part of users information seeking process in unfamiliar domains (Chapter 3.4). The knowledge management process blocks model extends this with the need to incorporate such support into knowledge structure maps supporting the processes of knowledge awareness and knowledge distribution (Chapter 3.3). Supporting these requirements is the rationale for accompanying the Document Map with a related Concept Map in our knowledge map model (Fig. 6-2, Section 6.1).

The main idea is to construct a Concept Map that visualises patterns of language use by a given community. In this section we present a solution for constructing such maps based on the analysis of word usage in the documents from the community space. This follows the same purpose as the system-generated Document Map described in the previous section. Namely, providing an initial overview to motivate user actions from which Concept Maps based on personal, implicit knowledge of human users can be constructed.

To realize the system-generated Concept Map we employ an approach similar to the method for generating word category maps from document collections proposed in (Honkela, 1997). The aim is to identify groups of words that are used in similar contexts and the relationships between them. To this end, as the measure of relatedness the frequency of co-occurrence of different words in the same context is taken (ibid.). In our case the context is given by the document abstract. Based on this measure, for each word from a set of M most frequent words a corresponding vector is created, describing the relatedness of a given word to each of the remaining ones.

The resulting vectors are then submitted to the SOM in the same procedure as for constructing the system-generated Document Map (Fig. 6-7). In brief:

- In the first phase the set of most significant words to be mapped is determined. They are used for creating the term vectors for each word as described above.
- In the second phase, after the training with the word vectors the resulting map positions words on a two-dimensional grid, such that words frequently co-occurring in the same context are positioned close to each other. The words are clustered in groups by determining the most significant term of the corresponding vector. Accordingly the cluster is labelled by the same term.

As a result, a two-dimensional Concept Map is generated that presents groups of words frequently occurring in the same context. Each group is characterized by a most significant word and the relationships between both individual words and different word clusters can be inferred based on the closeness of their location on the two-dimensional grid. We can refer to individual words as rudimentary concepts. Since the clustering groups together words used in similar contexts, these groups will represent words not necessarily having the same meaning but rather being used complementary. As a result, the combinations of individual words in the same group and in particular with the group label, are likely to yield complex concepts[49]. In this way, a representation of a shared vocabulary used by a given community, as reflected in the documents from the community space has been created.

In order to provide more explicit information about inter-cluster relationships the map is extended with a similarity relation describing the relationship between individual concepts, regardless of the clustering produced by the described method. For each concept a set of most significant terms from its term vector is extracted and included in the meta-data layer of the map objects (Section 6.6). While the map considers global similarity patterns, this relationships provides a measure of 1-to-1 relationships between

[49] Beyond providing initial insights into “real” concepts significant for a community vocabulary, this fact can also be used to provide concept based navigation (see Section 6.3.6. on coupling the document map and the concept map and the interface and interaction design in Chapter 8).

individual concepts which may be located in different clusters, and still relatively frequently co-occur. This relationship allows us to enrich the map with an additional dimension that turns it into a network of groups of concepts and relationships between them. Fig. 6-12 shows the visualisation of a system-generated Concept Map.

6.3.5. Visualising the System-Generated Concept Map

The Concept Map generated in the described manner is a reversed Document Map. Instead of documents, the objects positioned on a two-dimensional grid are the terms from the document vectors themselves. A natural way for visualising this map is to show the distribution of concepts on the map and display the label of the each cluster (the most significant concept) in its centre. Inter-cluster relationships between individual concepts are shown by lines connecting the related concepts with each other. Fig. 6-12 (left) shows an example Concept Map visualised in this way.

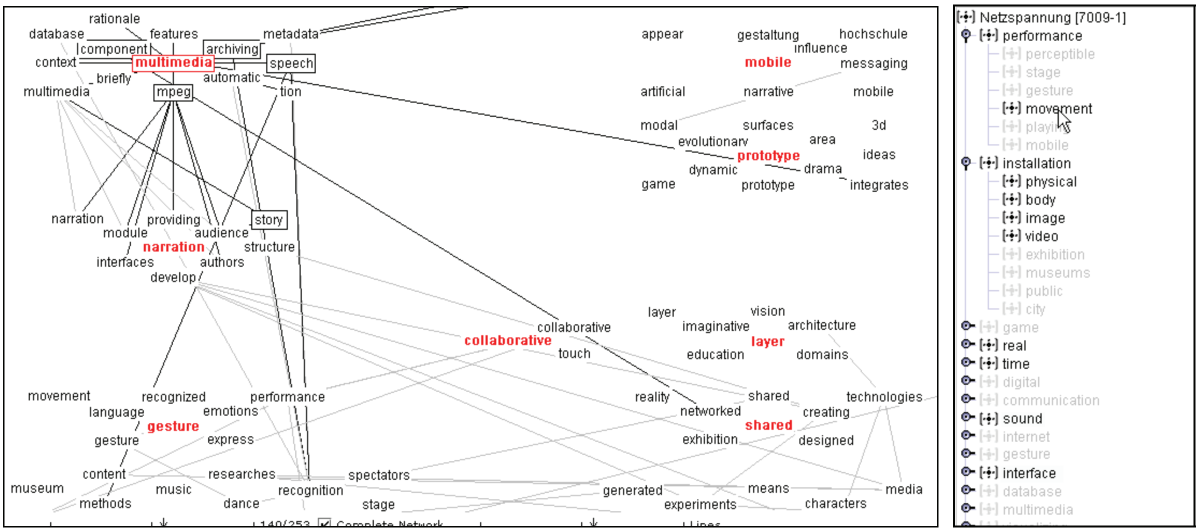


Fig. 6-12. Visualisation of the system-generated Concept Map: two-dimensional layout based on Kohonen SOM (left) and tree-folder visualisation (right).

While this method provides a good visual overview of groups of words and relationships between them as they are used in a given community information space, it is also very visually intensive. The results of the formative usability evaluation (Chapter 8.5) confirmed the usefulness of such Concept Map visualisation on its own, but pointed to strong increase in complexity when it is displayed alongside a related Document Map. To accommodate this feedback an additional, less visually demanding visualisation has been developed. To this end the familiar folder tree visualisation has been chosen, where each concept cluster is represented as a folder with an appropriate label (main concept). The inter-cluster relationships are not shown at the overview level, but by a selection mechanism when the user's interaction with the map places his focus of attention on a given concept (see interaction design in Chapter 8). This alternative visualisation method is depicted in Fig. 6-12 (right).

6.3.6. Coupling the Concept Map with the Document Map

The Document Map and the Concept Map are connected with each other in two different ways:

- First, during the generation process for the Document Map, for each document a set of most relevant terms[50] is extracted. This relationship between a document and most significant terms is represented by a "key" relation that is added to the meta-information layer of the map (Section 6.7). Based on this information, for each term in the Concept Map a set of related documents is identified and represented by the appropriate "key" relation in the Concept Map. In this way a

[50] The number of terms to be selected can be parameterized. We have experimented with sets between 5-10 terms for different indexing granularities, depending on the size of the document collection.

coupling between the Document Map and the Concept Map has been established. The users can now select concepts from the Concept Map representing the community vocabulary to identify related documents and vice versa, selecting a set of documents can identify most relevant concepts describing them. A detailed description of the ways this coupling has been used to support user interaction and navigation is given in the section concerned with specifics of the developed interaction design (Chapter 8).

- Second, the concepts displayed in the Concept Map can be selected by the user as the main labels according to which the Document Map should be reclustered. This is possible because the realized method for generating the Document Map separates the visualisation (SOM) and clustering. Invoking the clustering phase of the Document Map with the selected concepts will result in a reordered Document Map, displaying clusters of documents based on user-defined labels.

6.4. Eliciting Personal Points of View

The elicitation of personal knowledge structures of individual users is based on a two-stage processs. First, the described system-generated knowledge map of a community space is presented to the user. This provides an initial context for the user's access to information and a clear context against which the personal point of view of the user can be expressed and interpreted. Second, through his interaction with the information presented in the map the user reveals his personal point of view on the meaning of information and its relevance to a given need.

As the user explores the information presented in the map, he identifies relevant topics and documents and assesses their relevance both for his information need as well as for his more general interests. In doing so, he learns about both the specific knowledge of the community relevant for his need and its more general context represented by the map. This allows the user to develop a better understanding of an ill-defined information need during the process and considers his information seeking as a dynamic, non-linear process in which a user's need develops in the process of spotting various information that captures his attention ("berry-picking") (Bates, 1989).

The user's insights developed in the process can be expressed by selecting individual items into personal collections and by (re-) arranging them according to his personal understanding of their meaning. This includes creating groups of documents, moving objects between groups and adding group labels. In other words, the user creates bookmark-like collections of named groups of documents that reflect his personal point of view and the insights he discovered and internalized as knowledge (Fig. 6-13).

The promise of this idea is that although users do not explicitly describe their interpretation of the content of a given document they do so implicitly in two important ways: 1) by assigning a document to a specific concept, and 2) by grouping similar documents together. In the first case, an explicit concept represents a part of the implicit knowledge structure of a given user and relates it to a concrete instance. In the second case, grouping different instances under the same concept expresses the relationship of similarity. This can be used to classify unknown documents to a given concept, based on their comparison to the documents already assigned to the concept by the user. In this way, named document groups are used as a means for reflecting implicit knowledge in a way that can be made usable for contextualizing unknown information from a user's personal perspective.

6.4.1. Creating Personal Document Maps

Two different methods for creating personal Document Maps are provided. The first is based on the idea of two-dimensional bookmarks. The user is presented with a 2D visual workspace on which he can position selected documents and arrange them in named clusters (Fig. 6-13). Such a solution supports the highly explorative nature of accessing unfamiliar community spaces in an intuitive and visually appealing way. It is also a natural was of keeping the visual representation of the personal maps consistent with the way the system-generated Document Maps are displayed. This will play an important way when visualising the results of applying personal maps to dynamically generate personalised views of unfamiliar community spaces.

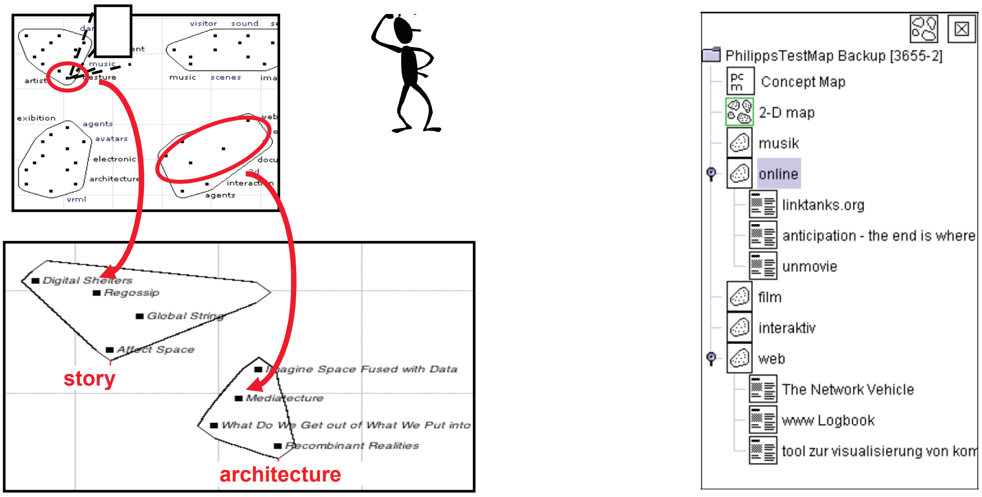


Fig. 6-13. Expressing personal points of view through personal Document Maps: two-dimensional workspace model (left) and one-dimensional bookmark folders model (right).

The second solution is based on the common bookmark-folders model (Fig. 6-13, right). In this case the user assigns documents to clusters represented by folders in a one-level folder hierarchy. The advantage of this method is that it uses a well known representation model that is visually undemanding and can be used for quick, unobtrusive access when users focus is on exploring the initial system-generated map. The limitation to a one-level folder hierarchy is sufficient for our case since the focus is on explorative access. Effectively, in our model the user works with a two-level hierarchy: the root map folder is also named and the individual maps represent themselves specific information seeking topics for which the user is gathering documents. Such a flat model also allows us to easily maintain a correspondence with the flat two-dimensional map structure. These turn out to be important issues when developing an appropriate interaction design for incorporating the maps into an interactive information interface (Chapter 8).

6.4.2. Learning Personal Document Maps

The user-defined maps are turned into dynamic templates for structuring information from a user's personal point of view in the following way. By creating a personal document map, the user defines a set of document clusters. The idea of learning a personal document map refers to finding a function which can autonomously assign new objects to the appropriate user-defined cluster. Realizing such a decision function allows the map to be applied for classifying any single object or information source into the user-defined structure.

An appropriate method for realizing such a function that assigns documents to clusters is the nearest neighbour method (Aha et al., 1991). This method assigns a given object to the cluster in which the most similar object from all existing objects in the personal map is located. It is especially suitable for our purpose since it can be applied in a way that satisfies two important requirements. The first is the efficiency and fast response time needed for real-time application. The second is good performance with few training data which is important since the user created personal maps will contain relatively small numbers of documents.

These criteria can be satisfied by applying the method in combination with a predefined document similarity matrix based on which the decision about the most similar object can be taken by a simple lookup. Furthermore, instead of performing similarity calculation only based on the objects selected by a given user so far, we can introduce a global measure. To this end, two different aspects for determining document similarity are combined: structural information based on content analysis and collaborative information based on context analysis from document maps of all users[51].

The basic concept of the solution for applying the nearest neighbour method to using user-defined maps as templates for personalised classification of information has been published in (Novak, et al. 2003a). The overall solution can be summarized as follows:

- *Generate and maintain a global similarity matrix:* A global similarity matrix representing the degree of similarity between document pairs is pre-computed and refreshed periodically. The document similarity measure is defined by a combination of content and context analysis. Content analysis uses text properties of documents to measure similarity between them. The calculation is based on a Euclidean similarity in a vector space model (i.e. calculating the distance between two document term vectors). Context analysis takes into account the user's personal interpretation of the similarity of two documents that is implicitly expressed by grouping them in a same cluster. If two documents appear together in many document clusters, it can be assumed that they are in some way similar. This similarity can be calculated by using the "Dice" – coefficient (McGill, 1979) that puts in relation the number of clusters in which two documents co-occur with the number of clusters in which the objects occur individually:

$$sim_{context}(D_i, D_j) = 2\frac{|C_{Di} \cap C_{Dj}|}{|C_{Di}| + |C_{Dj}|},$$

  where C_{Di} is a set of clusters containing document D_i and C_{Dj} is a set of clusters containing document D_j. By using this measure clusters which do not contain any of both objects are not considered whereas co-occurrences get double weight.

- *Apply the k-nearest neighbour method using the similarity matrix:* For a given personal map and a document collection to be classified, the k-nearest neighbour method (Aha et al., 1991) is applied. For each document in the document collection the most similar object in the given personal map is determined, by using the pre-computed global similarity matrix. The unknown document is then assigned to the map cluster in which the most similar object is located. The kNN-algorithm finds the k-nearest neighbours and weighs their categories (i.e. clusters in which they are located in the given personal map) based on a similarity to the document to be classified. The weights of neighbours sharing a same category are added together to form the probability weight of the given category for the document in question. The resulting set of candidate categories is ranked by weights and the document is assigned to the category with the highest score. The classification decision function for the kNN-algorithm is given with:

$$f(W_x, C_j) = \sum_{Wi \in Cj} sim(W_x, W_i) * f(W_i, C_j) - b_j,$$

  where $f(W_i, C_j) \in \{0,1\}$ is the classification function for document vector W_i with respect to cluster C_j (f = 1 for document belonging to cluster C_j and f = 0 for document not belonging to cluster C_j), $sim(W_x, W_i)$ is the similarity between the unknown document and the k-nearest neighbour documents and b_j is category specific threshold for the decision function (Yang & Liu, 1999).

Using content-based similarity bears the advantage that it is always applicable and doesn't rely on user generated data. However, using content-based similarity would lead to poor results in the cross-community context, where underlying objects are heterogeneous: little intersection between different

[51] The application of the nearest neighbour method and the similarity measure have been developed in collaboration with Michael Wurst, University of Dortmund, Dept. of AI, as part of his participation in the project AWAKE led by the author (http://awake.imk.fraunhofer.de).

community domains exists and different communities are using different terminologies. Using context similarity avoids these problems.

On the other hand, context similarity is problematic when a given object has been assigned to personal maps only by few users, since there is no reliable evidence on the similarity of this object to other objects. In order to balance this trade-off the two measures are combined in a way which allows us to detect automatically when the context-based measure is statistically significant and when the content-based measure only must be used. A statistical chi-square test is used to examine if co-occurences of two items in personal maps are statistically significant. If this is the case context-based similarity is used whereas otherwise only text-based similarity is considered by the system (Wurst, Novak, Schneider, 2002). The theoretical validity of the described method for generating personalised classification functions based on user-defined personal maps has been demonstrated in an evaluation on synthetic data by Wurst and published in (Wurst & Novak, 2004; Wurst, 2005).

6.4.3. Visualising Personalised Document Maps

Based on the described method personal Document Maps can be used to generate semantically structured views of the community information spaces in a way that reflects a specific point of view of a given user. Applying the personalised classification function of a given map (based on the nearest neighbour method) results in a a new personal Document Maps that contains relevant documents assigned to user-defined clusters. One way of visualising the resulting map is by displaying the original personal map created by the user and placing the classified documents in the appropriate clusters (Fig. 6-14).

In this case, the relative positioning of the document clusters on the two-dimensional surface correspond to the positions determined manually by the user when adding documents and clusters into the original personal map. In order to accommodate the newly classified documents, the cluster areas are automatically expanded such that clusters with more documents occupy a larger area (in a proportion relative to the size of other clusters).

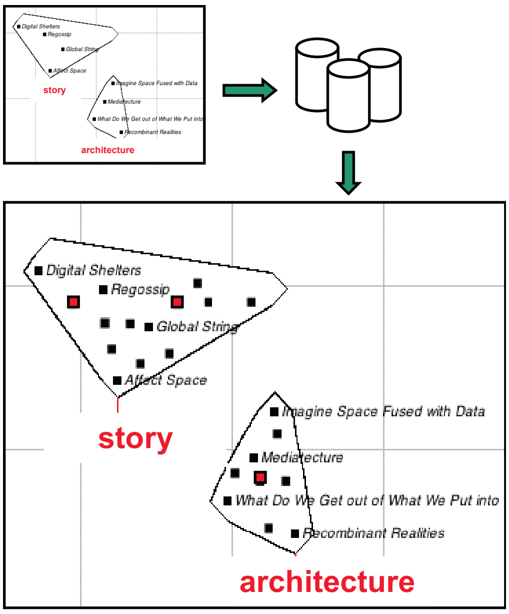


Fig. 6-14. Applying a personal Document Map to classify an unfamiliar information space from a personal point of view (no inter-cluster similarities).

While this solution is fast and immediate it, is not fully consistent with the visualisation model introduced by the system-generated map based on the Kohonen SOM. The assignment of documents to clusters indicates a certain degree of similarity between them, based on the user’s explicit grouping for the original documents and inferred by the classification function for the new ones. The original documents have positions assigned to them by the user while the newly classified one’s can be positioned in the same cell with the most similar document based on which they have been assigned to the cluster. In this way, a certain degree of spatial similarity relationships can be achieved within individual clusters.

But beyond this, the spatial arrangement of the documents doesn't provide any information about the semantic relationship between them. Hence, unlike in the case of the system-generated maps spatial distance does not imply semantic similarity. This is due to the fact that the classification function provides no two-dimensional mapping and offers no information about the degree of relatedness of documents to clusters to which they have *not* been assigned. This problem is even more critical in the case when the personal map has been defined as one-dimensional bookmarks only. In this case, the two-dimensional positions of the clusters are not available at all and must be generated in some sensible way. This requires a solution for generating a visualisation that displays both the results of the personalised classification into user-defined clusters while positioning the documents on a map in a way that reflects the semantic relationships between them.

To achieve this, a method has been developed that combines personalised classification based on user-defined templates with the Kohonen SOM. The main idea is to use the Kohonen SOM for positioning the documents on a two-dimensional map based on inter-document similarity but in a way which will retain user-defined document clusters. Since the SOM is an unsupervised network a way has to be found that will ensure that documents assigned to the same cluster by the user and by the classification function will remain close to each other after the SOM mapping.

For each document a similarity vector is computed representing how well a given document fits each of the user-defined clusters (Fig. 6-15). Each component of a document vector represents the degree of membership of that document to a specific cluster. Similar vectors then imply documents which have similar degrees of membership to different clusters. In other words, instead of using word frequency in texts, the semantic character of documents is encoded based the degree of their relationship to example documents in user-defined clusters. Passing such vectors to the SOM results in a map where similar documents are close to each other if they have similar degrees of similarity to user-defined clusters.

	D_1	…	D_m
W_1	0.2		0.3
W_2	0.1		0.5
…	…		…
W_n	0.6		0.4

DOCUMENT-TERM MATRIX

$D = \{D_1,\ldots,D_m\}$ – set of documents
$W = \{W_1,\ldots,W_m\}$ – set of document word vectors
$W_i = (w_{i1},\ldots,w_{in})$ - word vector for document Di

	D_1	…	D_m
C_1	0.2		0.3
C_2	0.1		1.0
…	…		…
C_l	1.0		0.4

DOCUMENT-CLUSTER MATRIX

$C = \{C_1,\ldots,C_l\}$ – set of clusters
$C_j = \{D_i\}$ – set of documents in cluster Cj
$Qi = (c_{i1},\ldots,c_{il})$ – cluster membership vector for Di

Cluster centroid: $Zj = \{z_{j1},\ldots,z_{jn}\}, \quad z_{jk} = \frac{\sum_{Di \in Cj} w_{ik}}{|C_j|}$

Membership of D_i to cluster Cj: $c_{ij} = 1.0\ ,\ D_i \in C_j$

$$c_{ij} = d(W_i, Z_j) = \sqrt{\sum_{k=1}^{n} (w_{ik} - z_{jk})^2}\ ,\ D_i \notin C_j\ ,\ c_{ij} = \frac{c_{ij}}{|Q_i|}$$

Fig. 6-15. Calculation of inter-cluster document membership for personal Document Maps

To ensure that the documents that have been assigned to specific clusters manually by the user or by the personalised classification function remain close to each other, the following procedure for calculating document-cluster membership is applied:

- The degree of membership of documents to clusters to which they have been assigned by the user is set to maximum weight,
- The degree of membership of documents to clusters to which they have been classified by the personalised classification function corresponds to the value calculated by the classification function,
- The degree of membership of documents to other clusters to which they don't belong corresponds to the Euclidean distance between the document term vector and the centroid of the cluster in question. The centroid vector is computed by taking the average of all term vectors representing documents in a given cluster.

The basic information about the document-cluster membership is provided by the result of applying the personalised classification function for a given user-defined map to a given document collection. The resulting map contains clusters of documents with corresponding weights describing the degree of membership to the cluster into which they have been classified. The exact calculation of the membership weights of documents to all other clusters is given in Fig. 6-15.

The final document-cluster matrix includes relationships computed by the the personalised classification function extended with information on inter-cluster relationships. The rows of the matrix represent document vectors that encode document properties based on the degree of their membership to different clusters in a way that reflects the user's personal point of view.

Passing such document vectors as the training set to the Kohonen SOM produces a two-dimensional map on which similarity of documents with respect to their membership to different clusters is reflected in their relative spatial positions. Thanks to the intrinsic property of the SOM to dynamically classify documents that have not been used in the training, the resulting map can also dynamically position new documents without repeating the training process. This further allows the contextualisation of search results containing documents that are not explicitly contained in the map.

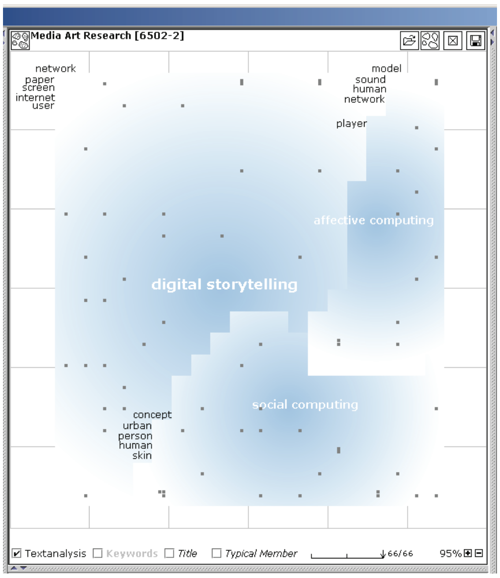


Fig. 6-16. Visualisation of personalised classification map with spatial distance reflecting inter-document similarities from a user-defined point of view.

The described method enables two-dimensional visualisation of personalised document maps in a way that retains the topological properties and spatial similarity relationships provided by the SOM mapping while reflecting user-defined criteria for determining the document clusters. This is special aspect of our solution, not available in other approaches. An example visualisation of a personalised Document Map created by this procedure is given in Fig. 6-16.

A specific aspect of the presented solution is the two-stage model which allows us to balance the inherent trade-off between the requirement of fast real-time response and the semantic accuracy of the visualisation. In the first step an immediate, real-time capable personalised view can be provided to the user: the classification results are displayed within the original spatial arrangement defined by the user (2D or 1D). The price for this is the missing relationship between spatial distance and semantic similarity of the document placement. Introducing this information, requires an additional step for which the user must take into account additional time for processing as a result of obtaining more precise semantic relationships.

6.4.4. Generating Personal Concept Maps

Personal Document Maps are accompanied by corresponding personal Concept Maps that represent most relevant concepts and relationships between them reflecting the conceptual structure of the map author. The main idea is that the relationships between document groups and concepts established by user-defined cluster labels in personal Document Maps can be used to infer groups of related concepts as seen from a user's point of view. Since cluster labels are defined by the user in the process of interpreting and organizing information they represent concepts used by the user to contextualize retrieved information with respect to his specific need and personal knowledge (Chapter 7.1.3).

Besides the user-defined labels each document cluster can be characterized by a set of most relevant terms from the document term-vectors created by statistical analysis of word occurrence in document contents (Section 6.3.1). Based on this information individual user-defined concepts can be related to groups of terms that represent a richer context of their meaning. The relationship of user-defined concepts between each other can be established by considering the degree of similarity between document groups to which they are referring.

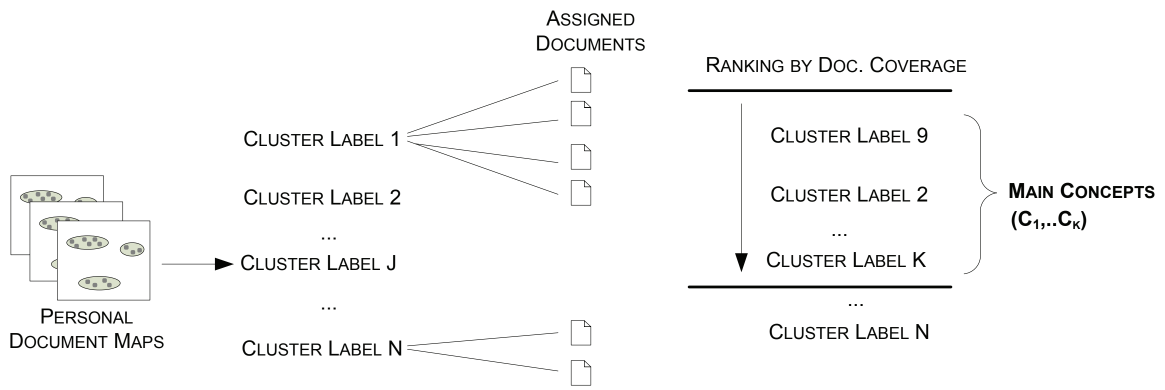


Fig. 6-17. Constructing Personal Concept Maps: Identifying main concepts

The basic method for constructing personal Concept Maps according to this idea is illustrated in Fig. 6-17 and Fig. 6-18. First, a set of main concepts (topics) is identified by selecting a set of most relevant user-defined cluster labels. To this end, the cluster labels occurring in all personal maps of a given user are ranked based on the number of documents assigned to each label. For labels occurring in different maps, all label instances are counted as one label. All documents assigned to different label instances are counted as belonging to this one label. In this way, labels covering larger numbers of documents are considered to be representing more general areas of user's interests. Labels covering smaller numbers of documents can be interpreted as referring to more specific or less frequently occurring topics.

The set of documents associated with each main concept is used for finding relevant terms and other labels as sub-concepts related to the main concept in question. To this end, the standard TFxIDF measure is applied for calculating the relevance of a term or label to individual documents (Section 6.3.1) and then averaged across the given documents set. An equal number of most relevant terms and most relevant user-defined labels is taken as sub-concepts for the main concept in question (Fig. 6-18).

TERM | CONCEPT

DOCUMENT

	W_1	…	W_m	C_1	…	C_i	…	C_q
D_1	0.2		0.3			0.1		0.2
D_2	0.1		0.5			0.5		0.1
…	…		…			…		…
D_n	0.6		0.4			0.3		0.4

TERM-DOCUMENT MATRIX | CONCEPT-DOCUMENT MATRIX

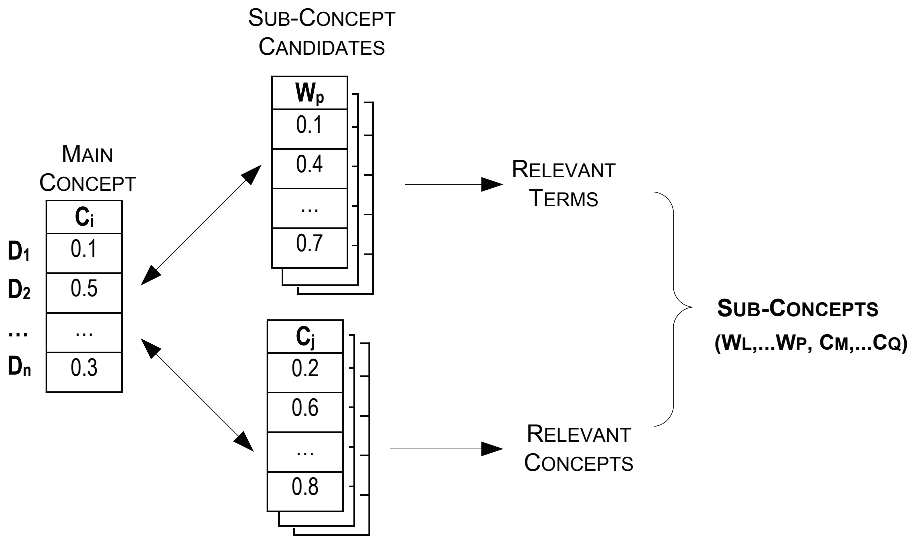


Fig. 6-18. Constructing Personal Concept Maps: Selecting sub-concepts

The degree of relatedness between sub-concepts assigned to different main concepts is determined by taking a cosine measure between the corresponding document vectors found in the term-document and concept-document matrices respectively (Fig. 6-18, top). Finally, the degree of relatedness between the main concepts themselves is calculated by taking an average of the pair-wise similarity between documents referenced by the concepts being compared.

In this way a personal knowledge structure of the user is represented by a network of concepts and relationships between them, based on their usage in the user's personal maps. The concepts are linked to documents as concrete instances of their meaning which makes the Personal Concept Maps usable for navigating an information collection from a personal perspective of a specific user.

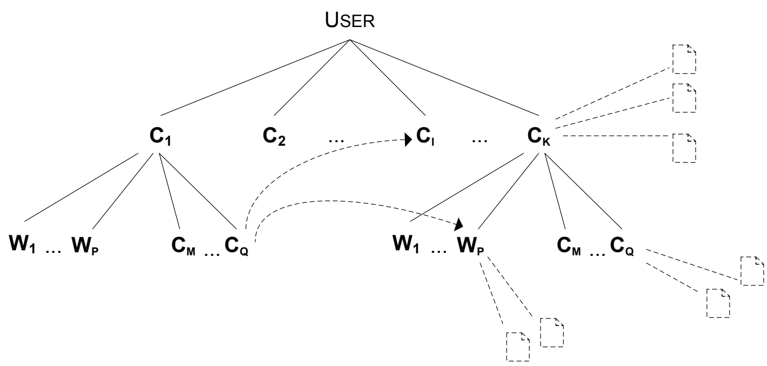


Fig. 6-19. Constructing Personal Concept Maps: Structure of the resulting concept network

6.4.5. Visualising Personal Concept Maps

The concept network represented by a Personal Concept Map can be visualised in several different ways. A straightforward approach is to apply graph-visualisation methods based on the spring-embedder model used for creating 2D graph layouts. In this approach the links between the nodes of a graph are modelled as physical springs where the weight of the link describes the distance between the nodes and in effect the energy state of the spring. The spring-embedder algorithm then tries to find an optimal spatial distribution of nodes that results in the minimum energy level of the system (see Chapter 4.2.4.3)

While this is a very popular method for visualising complex graphs it bears several disadvantages in our application context. Besides being relatively slow in general (though there are exceptions for small graphs) the biggest disadvantage is its iterative procedure. Until the algorithm has converged, the shape of the graph changes in front of the user. Even worse, the same algorithm applied to the same graph can produce two very different layouts (Hermann et al., 2000).

Thus, the user could be confronted with a new visual layout every time when opening the Concept Map even when there were no changes in its contents. Last, but certainly not the least, introducing a different visualisation model for Personal Concept Maps than for the system-generated Concept Maps increases cognitive complexity for the user. Such inconsistency makes the interaction with spring-embedder models unsuitable for visualisation and navigation of Concept Maps in our application context.

For these reasons the visualisation methods for the Personal Concept Maps are the same as the visualisation models for the system-generated Concept Map introduced in Section 6.3.5. The first is based on the Kohonen SOM (Fig. 6-12, left) while the second provides allows a simpler folder-tree visualisation (Fig. 6-12, right). In order to visualise the Personal Concept Map by means of the SOM the same method as the one developed for visualising the Personal Document Map can be applied (Section 6.4.3).

Both maps have an analogous structure, with concepts in the Personal Concept Map taking the place of documents in the Personal Document Maps. For each sub-concept the degree of relatedness to different main concepts can be determined by calculating the distance of that concept's document vector to the centroid vector of the concept cluster in question, following the same calculation given in Fig. 6-15 (replacing document clusters with concept clusters and word vectors with concept document vectors). The sub-concept document vectors can be taken directly from the concept-document matrix depicted in Fig. 6-18. Accordingly, for each a sub-concept a vector is created whose components represent the degree of relatedness to different concept clusters.

Passing these concept vectors to the SOM results in a two-dimensional map where spatial closeness reflects semantic similarity between concepts, while retaining the original concept group memberships. The relations between concepts from different clusters are displayed by lines connecting appropriate concept nodes with each other (Fig. 6-12, left).

The second visualisation model represents each concept cluster by a folder labelled by the corresponding main concept. Sub-concepts are contained in the concept folder while the display of inter-concept relationships follows the interactive user focus selection. This corresponds fully to the folder-tree visualisation model of system-generated maps depicted in Fig. 6-12 of Section 6.3.5.

6.5. Extracting Conceptual Structures of Communities

The described method for determining relationships between concepts used by an individual user in his personal Document Maps is based on the fact that by assigning cluster labels the user draws a connection between a concept and documents contained in the cluster. Expanding this analysis to personal Document Maps of all members of a given community allows us to establish relationships between concepts used by different users and create a Concept Map representing a shared community perspective. The basic premise is that relationships between different concepts used by different users can be established by considering the relatedness of the document sets indexed by the concepts in question. In a similar way, documents indexed by many different concepts are likely bridges between different personal perspectives on a shared topic of interest. Collecting the main concepts defined by the community members in their information seeking activities and putting them in relation to each other thus allows us to create a representation of a shared conceptual network representing implicit knowledge structure and shared vocabulary of a community of users.

6.5.1. Generating Community Concept Maps

To construct a Community Concept Map representing most relevant concepts and relationships between them as used by a community of users, we can take advantage of the same method introduced in Section 6.4.4. To this end, the user-defined cluster labels from all Personal Document Maps created by the members of a given community are collected and merged into a ranked list, based on the number of documents indexed by each label (Fig. 6-20). The main concepts used by a community are those who cover the largest number of documents while the remaining concepts are clustered with respect to the main concepts according to degree of relatedness calculated by the same procedure as for the Personal Concept Map. The relations between sub-concepts and different main concepts as well as the relations between main concepts themselves are computed in the same way as presented in Section 6.4.4.

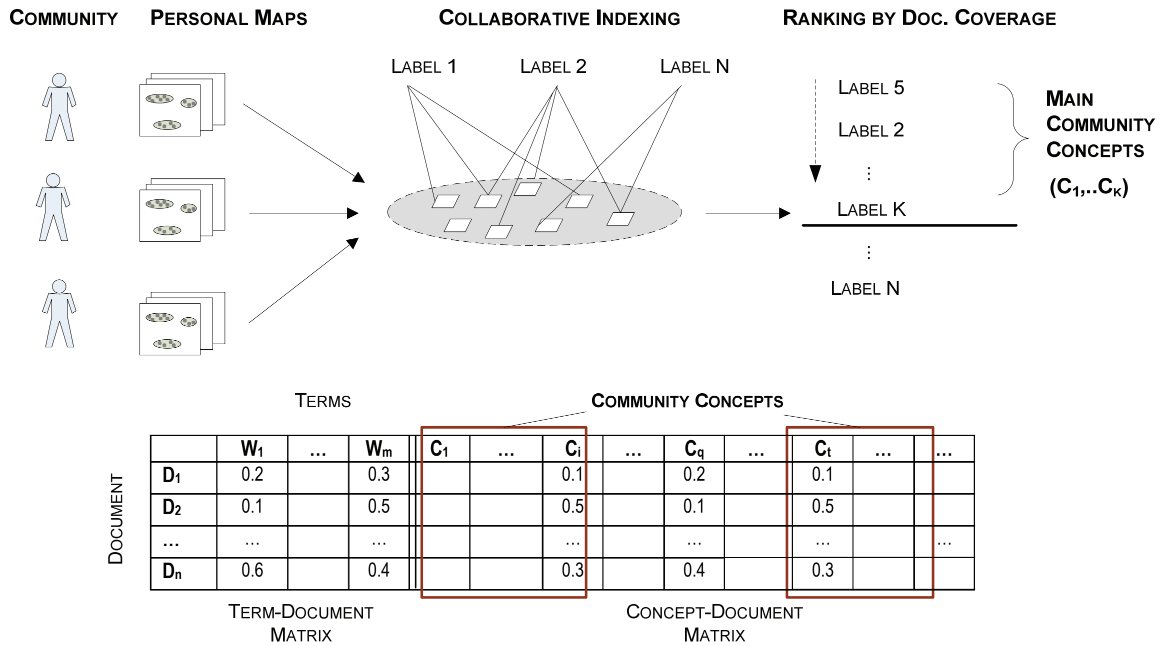


Fig. 6-20. Constructing Community Concept Maps: Identification of main concepts (top) and evolution of the Concept-Document Matrix (bottom).

An important effect of using a combined term-document and concept-document matrix is that the weight of collaboratively defined concept-document relevance will gain more importance as the number of maps created by the community evolves (Fig. 6-20, bottom). In this way, the relevance measure used for determining concept-document and concept-concept will naturally switch from the initial text-based measure to the collaborative measure with the increasing number of user-defined concepts. In this way a smooth transition from the cold-start solution is achieved. At the same time, since the sub-concepts are explicitly selected both from the text-analysis terms and user-defined concepts (Fig. 6-18), there will still be a balance between inherent content-based and user-defined aspects of community knowledge.

The conceptual structure extracted in this way presents relationships between concepts used by different members of a community and relates them to specific documents that serve as bridges between the different personal perspectives.

6.5.2. Visualising Community Concept Maps

The visualisation of the community Concept Maps follows the same model as the visualisation of the personal and system-generated Concept Maps. We can either use the procedure based on applying the Kohonen SOM to position the concepts on a 2D grid (see Section 6.4.5). Or we can take advantage of force-directed placement methods for two-dimensional graph visualisation, with all the disadvantages discussed before (Section 6.4.5, Chapter 4.2.4.3). Fig. 6-21 (left) shows an example of a community Concept Map visualised by a two-dimensional graph layout based on a spring-embedder model. However, after the results of a formative usability evaluation, due to the visual complexity of two simultaneous spatial visualisations (document map and concept map) this was replaced by a much simpler tree-folder visualisation with interactive user focus (Fig. 6-21, right), as described in Section 6.4.5 (for details see also interaction design in Chapter 8).

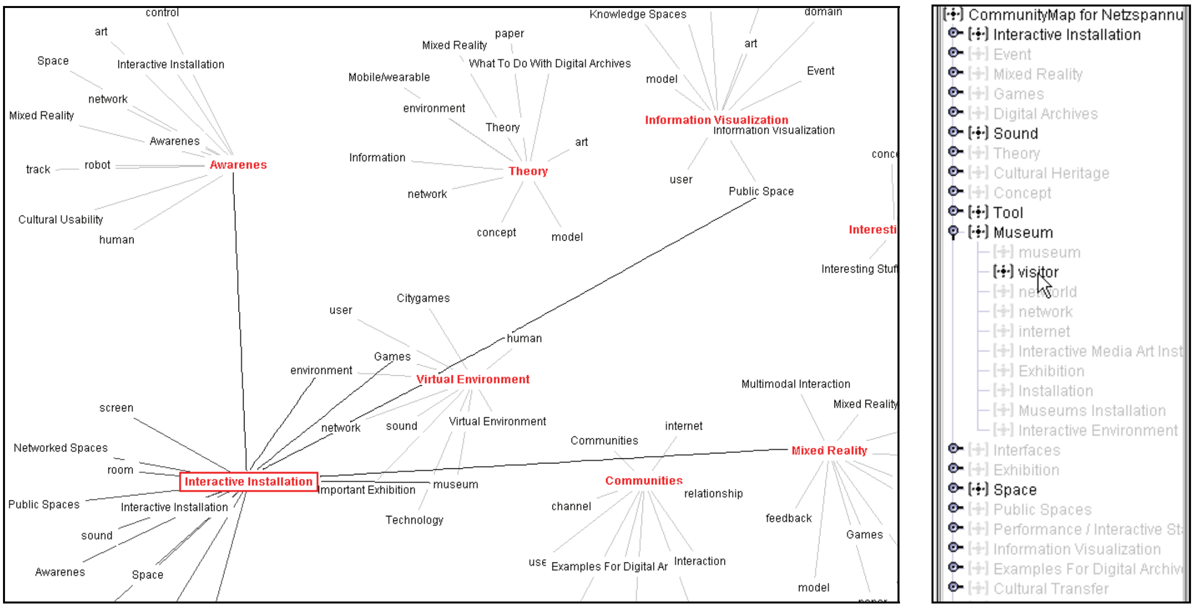


Fig. 6-21. Visualisation of community Concept Maps: two-dimensional graph layout based on spring-embedder technique (left), tree-folder visualisation (right).

6.6. Ontological Structure of the Developed Knowledge Map Model

The described method allows us to elicit conceptual structures representing knowledge of individuals and groups of users in a way that relates it to concrete information items and entire information spaces. This provides the basis for using them as tools for information access that contextualizes information from specific viewpoints defined by different users and communities.

A specific aspect of our approach is the focus on using the knowledge maps as visual information tools allowing interactive use and discovery of relationships between different perspectives. Furthermore, rather than formulating one comprehensive ontology, the creation of many different (but related) knowledge structures must be supported (different users creating different personal maps, different community maps being extracted for different communities).

Personal maps are intended as a medium of exchanging personal knowledge and must be shareable between users. Different community perspectives should be represented independently of each other support the discovery of relationships between them. Achieving high-level of interactivity of the maps requires significant metadata information to be locally available to the user.

Hence, rather than modelling a complex ontology into which all maps and relations between them are incorporated we adopt a distributed approach with a light-weight formalisation structure. Each personal map is considered as an aspect of a semantic structure of a larger whole, reflecting a specific view based on a short-term need of a given user. In this sense, each map can be considered a partial ontology.

However, two important characteristics distinguish the ontological map structure from classical ontology approaches. First, the main semantic structure can be expressed by few different but powerful kinds of relationships represented by the maps, such as “is_member”, “is_similar”, “is_key” and “is_representative” (Fig. 6-22, Fig. 6-23). Second, the inference of relationships which are not explicitly expressed is based on assisting human users in discovering implicit patterns through visualisation and interactive manipulation (e.g. spatial visualisation of similarity structure) rather than machine-based reasoning. In this sense, simple but powerful visual metaphors and the ease of interaction with the displayed structures are more important than complex inference mechanisms provided by ontology-engines. Against this background a light-weight ontological structure has been chosen for a formal knowledge map representation and serves only as means for supporting the visualisation and interactive manipulation by the user (exploration, navigation etc.). No complex inferencing mechanisms and ontology querying are required.

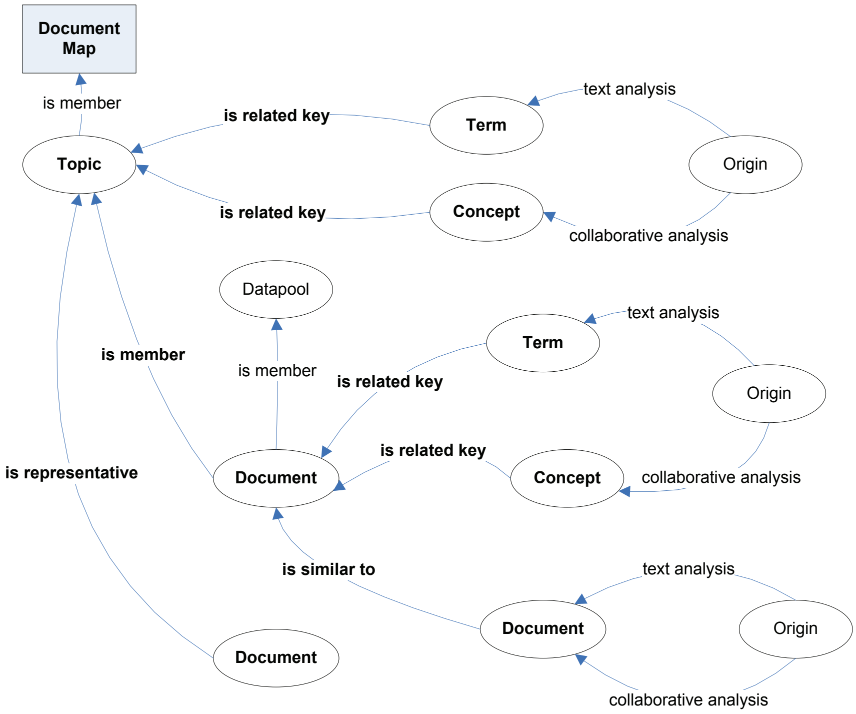


Fig. 6-22. Basic ontological structure of the Document Map

The ontological structure of individual document maps is depicted in Fig. 6-22. In order to connect such individual document maps with each other, instead of merging them into one ontology we can add to each map enough semantic information about its relations to other maps describing other aspects of the knowledge structure and its relationship to a given information space. According to the presented method, this is done in two ways:

- by collecting relationships between concepts in personal document maps of a given user into a Personal Concept Map representing the user’s personal knowledge structure (Section 6.4).
- by extracting relationships from personal document maps of all users of a given community into a Community Concept Map representing the shared knowledge structure of the community in question (Section 6.5).

Since in our knowledge map model the concept map serves as an “explanation” of the related document map, its formal representation follows the same structure. Whereas in the Document Map the main relation is the association of documents to topics representing clusters of related documents, the Concept Map structure represents the association of terms and concepts to topics representing clusters of related concepts. Fig. 6-23 depicts the ontological structure of such a concept map in our model.

The connection between the overall Personal Concept Map of a given user and the Document Maps belonging to him is provided by the "key" relation connecting topics, concepts and terms to documents. In this way, for any given personal document map a link between a user-defined concept or a significant term extracted by the text-analysis and related documents can be established (Chapter 6.3.5). Additional relations such "is_similar" and "is_representative" provide insights into the relationships between different concepts within the user's conceptual structure and the degree of membership of a given concept to its topic. This allows the realization of simple but powerful visualisation and interaction modalities such as semantic zoom and cross-concept navigation that will be introduced in Chapter 8.

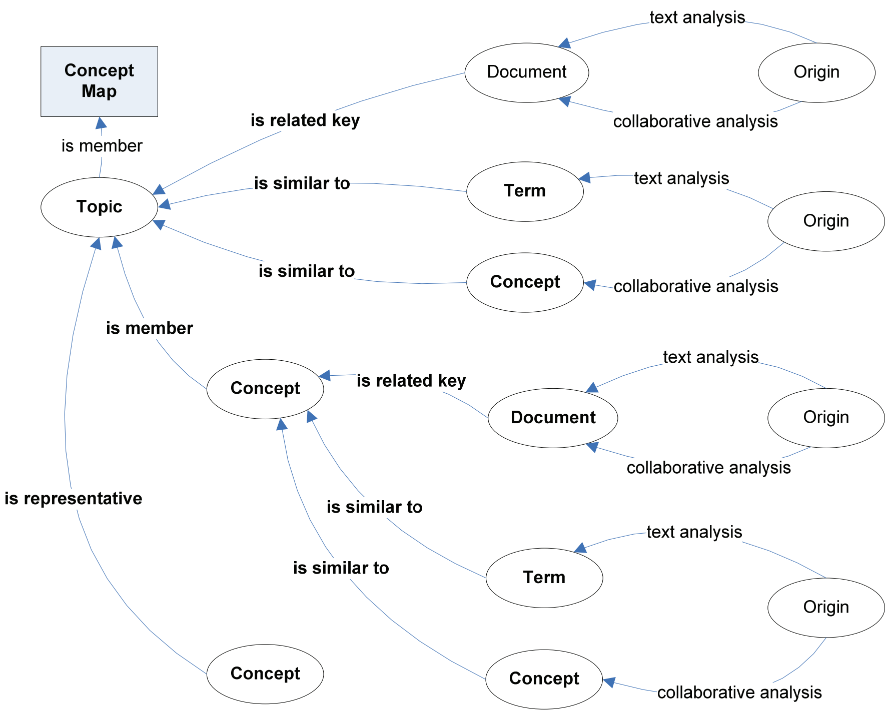


Fig. 6-23. Basic ontological structure of the Concept Map

The Community Concept Maps are represented in much the same way, since the method of their generation is the same as for the Personal Concept Maps (Chapter 6.4.4). The only is their extension over all personal maps of a community of users and the assignment to a specific community. The latter is accommodated by an additional "community" attribute in the map-metadata attributes. Fig. 6-24 shows the full ontological structure of our knowledge map model including relations connecting the Document Map and Concept Map with each other (dotted lines in bold blue colour).

6.7. Summary of Supported Requirements for Cross-Community Knowledge Exchange

The knowledge maps constructed by the described method satisfy several important requirements for cross-community knowledge exchange identified in Chapter 3. They are constructed in a way that visualises implicit knowledge structures of individual users and communities and makes them usable for contextualized information access. Providing knowledge maps of community information spaces allows users to gain an overview of unfamiliar knowledge domains. The related conceptual structures support insight into the shared community vocabulary and allow members of different communities to develop an understanding of the meaning of community concepts by exploring the relations of concepts to documents that exemplify their use. In this way both the requirements of knowledge awareness (Chapter 3.3.2) and perspective taking are supported (Chapter 3.1.5).

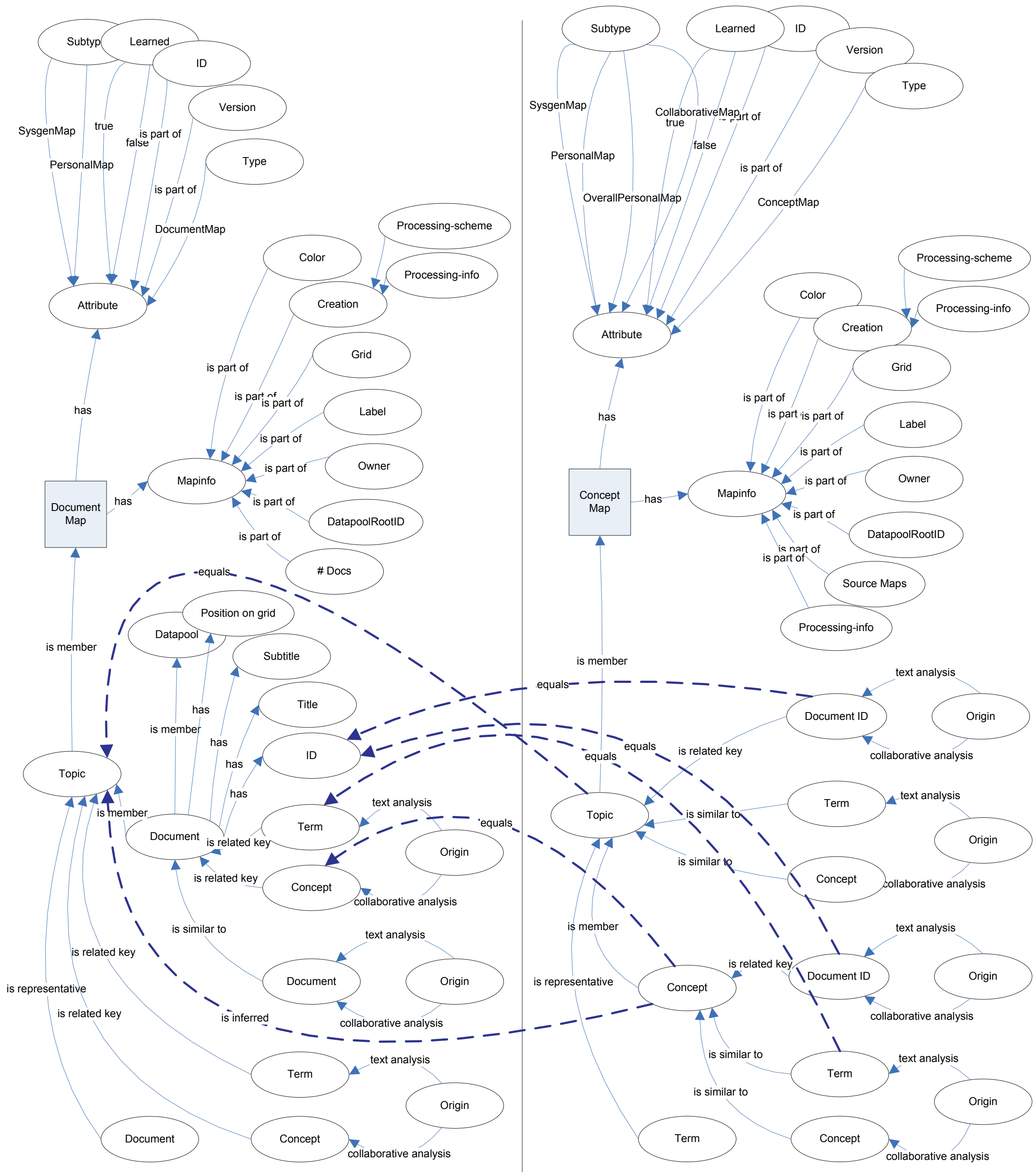


Fig. 6-24. Full ontological structure of the proposed knowledge map model with metadata details and inter-map relationships.

The two-stage model of transition from community maps based on visualising the semantic structure of community information spaces to collaborative elicitation of conceptual structures expressed by human users through their interaction with information provides two important benefits. On one hand, there is an immediately available solution which solves the cold-start problem until enough user created maps are available. On the other hand, extracting conceptual structures from user-defined maps involves otherwise unavailable personal knowledge of user's regarding their interpretation of documents in the community space. Moreover, instead of simple term-document relationships both more complex and more natural conceptual structures involving multi-word concepts can be extracted.

Finally, since such maps reflect the interests of community members expressed through their activities of information seeking aimed at satisfying different information needs they will evolve in time as the community interests evolve. Thus, they will reflect the implicit changes in community knowledge in an immediate way, even before this is explicitly externalized in new documents created by community members. In this way, the community maps are a living mirror of community knowledge as it develops over time. This respects two important requirements. First, the personal views and implicit knowledge of community members are incorporated in such maps (cf. perspective taking, Chapter 3.1.2) . Second, the applied procedure of knowledge elicitation respects the social processes through which knowledge is created and shared within communities (cf. perspective making, Chapter 3.1.1).

Personal knowledge maps provide artefacts that reflect specific personal knowledge of individual users and can be shared with others. They can be applied as dynamic semantic filters for personalised classification of information and for structuring unknown information spaces from a personal point of view. This motivates the construction of such maps by providing immediate benefits to the map author, which alleviates the free-rider problem.

Moreover, personal maps are realized in a form that allows them to be shared between users. A user can apply a personal map of another user in order to classify information into knowledge structure representing that user's point of view. In this way the personal knowledge of the map author is made usable to other users. They can discover information relevant from the authors' point of view, that otherwise they may not be able to identify themselves (cf. knowledge distribution requirements, Chapter 3.3.1). Applying the maps in this way is useful both within and across community boundaries. This supports the requirement for integrating processes of perspective making and perspective taking (Chapter 3.1.5).

Applying personal maps of other users as templates for information filtering is also a special way of providing personalised recommendations. By choosing a specific personal map as the classification template the user expresses a specific focus of interest (represented by the cluster topics and original documents contained in the map). The willingness to share personal maps with others is based on the principle of reputation economy. In classical collaborative filtering techniques the criteria and knowledge of individual users used for providing recommendations is invisible to the users. In contrast, in the described solution the author of a given personal map is always identifiable. This respects both intra-community mechanisms of social reputation and ensures credibility for other users (Kautz et al., 1997). These are important ways of satisfying the requirements of knowledge distribution (Chapter 3.3.1).

Making the maps usable for intra-community knowledge exchange ensures that enough maps will be produced within communities in order to extract concept networks representing shared community perspectives. This solves the cold-start and free-rider problems of collaborative filtering techniques. It also increases the probability that there will be enough different maps that may be relevant for different needs of members from other communities. The privacy requirements can be respected by introducing a publishing functionality that allows users to explicitly distinguish between private maps and maps accessible to others.

The described method differs from other existing methods in several important aspects. First, it allows us to construct personalised and shared knowledge maps in a way which does not require explicit negotiation and interaction between individual community members. Furthermore, the proposed technique is unobtrusively embedded into users' access to information: it doesn't require any special additional effort, it provides immediate benefits to the user and offers a solution even during initial phases of the system when small number of personal maps is available. Thus, it promises to overcome the bootstrapping and free-rider problems of collaborative filtering[52] and existing solutions for cooperative concept indexing (Gräther & Prinz, 2001; Voss et al., 1999).

Secondly, in contrast to static ontologies which aim at "codifying" knowledge, the proposed technique aims at realizing knowledge maps as dynamic visual artefacts that can be interactively manipulated by the users in order to generate and explore different views of the information space. This includes dynamic contextualization of unfamiliar artefacts into familiar knowledge structures (personal or shared). In this

[52] See (Herlocker et al., 2000) for analysis of advantages and shortcomings of collaborative filtering.

way, the implicit knowledge of communities of users that is reflected in the elicited knowledge maps is made usable for contextualized information access. While other approaches based on document maps exist, they are conceived as tools for supporting detailed analysis of patterns and relationships in specialized document collections (Chapter 4.5.6). They do not incorporate personal knowledge of individuals and groups of users and are not intended as means for supporting information access.

Finally, sharing and applying personal maps allows users to classify an information space from the viewpoint of different users or contextualize search results from unknown domains. This is both a way of accessing otherwise unavailable knowledge of other users and discovering different contexts and meanings of information. Similarly, the existence of visual representations of knowledge structures of different communities provides the basis for visual discovery of relationships between them. A concrete model for achieving this based on a cross-community knowledge exchange application framework and a multi-perspective knowledge visualisation interface that provide a way to relate unfamiliar concepts to one's own personal knowledge structure is presented in the next chapters (Chapter 7-8).

7. A Framework and System Architecture for Cross-Community Knowledge Exchange

This Chapter introduces the basic framework for applying the described method for collaborative elicitation and visualisation of implicit knowledge structures to supporting cross-community knowledge exchange. A corresponding system architecture enabling the realization of the framework is also presented. The design of the framework (Section 7.1) considers two main sets of requirements: 1) supporting typical sensemaking tasks during information access in unfamiliar domains (Chapter 3.4) and 2) supporting knowledge management process requirements relevant in the cross-community context (Chapter 3.3).

The developed system architecture implementing the services required by the framework gives special attention to performance capabilities required for an interactive system (Section 7.2). Specific aspects of the conceptual model, framework and related system designs have been published in (Novak et al., 2003c, Novak et al., 2004a,b,c; Novak & Wurst, 2005a).

7.1. Application Framework for Cross-Community Knowledge Exchange

In the previous chapter we have introduced a method for constructing dynamic knowledge maps that reflect implicit knowledge structures of individuals and communities of users. In this section we propose a more precise conceptual framework for using such knowledge maps to support cross-community knowledge exchange (Fig. 7-1). We do so by relating the functional properties of the described knowledge map method to specific tasks relevant for cross-community knowledge exchange identified by the requirements analysis in Chapter 3.

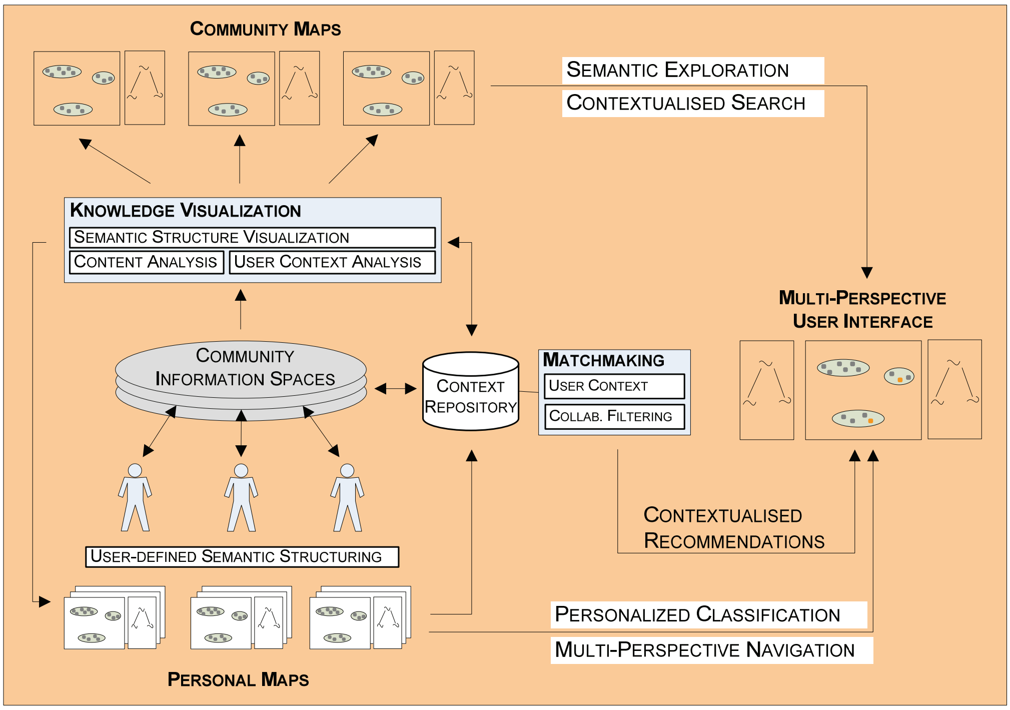


Fig. 7-1. Conceptual Framework for Cross-Community Knowledge Exchange

The application framework demonstrates ways in which the personal and community knowledge maps can be used for implicit and explicit exchange of knowledge between members of different communities and includes a set of services supporting such use cases. This includes the use of community maps for semantic exploration of unfamiliar community spaces, the use of personal maps for personalised classification and filtering of unfamiliar information, contextualized recommendations of maps relevant for a given information need as well as multi-perspective navigation for discovering how unfamiliar information is related to one's own knowledge (and vice versa).

7.1.1. Semantic Exploration and Contextualized Access to Unfamiliar Domains

Automatically generated semantic overviews of community information spaces in form of system-generated maps provide insights into the implicit structure of community knowledge as reflected in the documents contributed by community members. The Document Map presents a topical and similarity structure of the community information space. This allows users to gain a quick overview of the knowledge structure of an unfamiliar community: main topics, concepts, groups of related documents and relationships between them. Identifying main topics and exploring the character of the document groups they are composed of allows quick identification of potentially relevant areas and items of interest.

Visualising the shared community vocabulary in the related community Concept Map enables users to develop an understanding of the language used by the community in question. Since the Document Map and the Concept Map are linked to each other, the meaning of concepts used by the community can be understood by inspecting how they are used in specific documents assigned to them (Fig. 7-2). This supports perspective taking (Chapter 3.2).

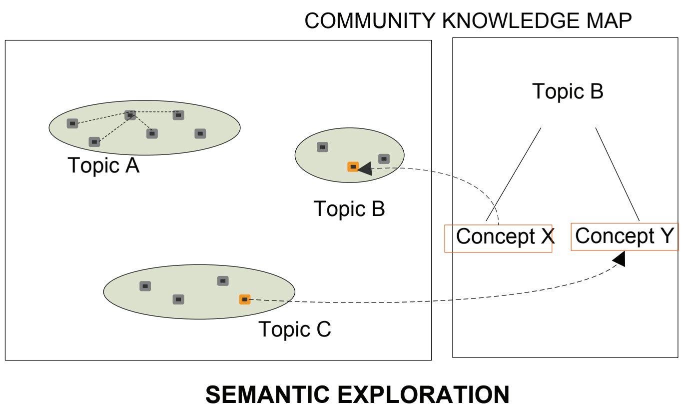


Fig. 7-2. Exploring the Community Knowledge Map for learning about the conceptual structure of an unfamiliar knowledge domain (schematic illustration)

When enough user-defined personal maps are available, the system-generated map can be replaced by the collaboratively elicited Concept Map containing concepts that have been used by the users to specifically characterize different document clusters. Unlike single word terms extracted by the system-generated map, the collaboratively elicited Concept Map contains multi-word concepts used by the community members. Since they stem from personal maps defined as part of user's information seeking activities they will reflect the evolution of user interests over time.

This ensures satisfying the requirement that shared vocabulary of the community be reflected in a way that incorporates personal viewpoints of individual users (Chapter 3.1.5). Providing such maps as information interfaces to community spaces allows members of other communities to gain insight into the structure of an unfamiliar knowledge domain of communities other than their own, as part of an information seeking process. This supports the requirement of perspective taking. It also provides support for knowledge awareness since by gaining insight into the structure of community knowledge users can become aware of topics potentially relevant for their information need.

Explorative access is important in ill-defined problems and unfamiliar domains, since the main difficulty lies in formulating the information need. An overview of available topics and concepts describing community knowledge supports the identification of appropriate concepts for describing the information need (cf. sensemaking, Chapter 3.4). The understanding of meaning of unfamiliar topics and concepts is supported by relating them to concrete documents.

SUPPORTED TASKS	COMMUNITY KNOWLEDGE MAPS
PERSPECTIVE TAKING	• Gaining insight into community knowledge structure: o topical structure o document-term relationships • Developing an understanding of the shared community vocabulary: o concept-concept relationships o concept-document relationships o concept-topic relationships
SENSEMAKING	• Finding appropriate concepts to express an information need
KM PROCESSES	
KNOWLEDGE AWARENESS	• Becoming aware of knowledge from unfamiliar communities
KNOWLEDGE APPLICATION	• Relating unfamiliar knowledge to the information need of the task at hand

Table 7-1. Tasks and requirements supported by Community Knowledge Maps

Supporting processes of knowledge creation and knowledge application requires that access to information be contextualized with respect to the need of a specific task (Chapter 3.3.3, 3.3.4). This can be supported by combining explorative access with goal-directed search in the following way. By formulating a search query the user expresses an information need based on his current task. Search results can be contextualized in the map of a community information space by visualising the corresponding documents (Fig. 7-3).

Using the map to localize the documents retrieved by the search query contextualises them within a knowledge structure of a specific community. Labels of clusters within which the documents appear put them in relation to main community topics. The spatial proximity to other documents within the same cluster or in neighbouring clusters allows the user to identify other potentially relevant documents that have not been retrieved by the query. In this way, relevant documents can be found although the query did not contain the appropriate search terms. When the user is unfamiliar with the specific terminology of the community in question or the information need is too unclear to be formulated well, this is of great importance (Chapter 3.4).

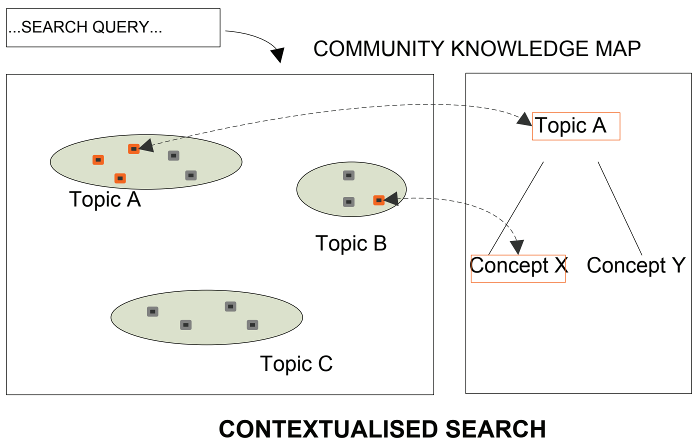


Fig. 7-3. Contextualising search results in a Community Knowledge Map and finding appropriate concepts for expressing a given information need (schematic illustration)

The documents retrieved by the search query can also be put in relation to the concepts displayed in the community Concept Map based on its connection with the Document Map (Chapter 6.3.6, 6.6). Each document in the Document Map points to a set of most relevant terms describing it. Each concept in the Concept Map points to a set of documents that have been assigned to this concept by different users. These relationships can be used to mark concepts that have been used within the community to refer to documents contained in the given search results. The user can now identify community concepts corresponding to his query and thus learn how to express his information need in terms more suitable for the community in question. This supports a typical sensemaking task of finding appropriate concepts for expressing an ill-defined information need (Chapter 3.4) Learning about the meaning of concepts from an unfamiliar community and their relationship to the information need of one's specific task, is also a way of learning about relevant knowledge from another community and applying it to the needs of the task at hand (cf. knowledge application, Chapter 3.3.4). Table 7-1 summarizes the main tasks and requirements of cross-community knowledge exchange supported by the described combination of semantic exploration and contextualisation of goal-directed access to community knowledge maps.

7.1.2. Contextualising Unfamiliar Information within Personal Knowledge Perspectives

The personal knowledge of a given user is expressed by his personal maps in two ways. On one hand, different personal Document Maps relate different aspects of his knowledge to specific information needs. On the other hand, the personal Concept Map represents the overall structure of the user's knowledge as it is reflected in all of his maps.

The capability of personal maps to dynamically classify a given set of documents into user-defined clusters allows them to be applied as semantic filters for categorizing unfamiliar community information spaces from the user's point of view. Applying a personal map to the information space of an unfamiliar community will produce a map that contextualises documents from the community space within clusters in the user's personal map. Since the assignment of new documents to clusters is based on their similarity to documents already contained in the user's map (Chapter 6.4.2) this effectively results in a categorization of an unfamiliar information space based on user's criteria of document membership to a given topic.

Such categorization of unknown information into categories defined by a user's personal point of view, allows the user to discover relationships between relevant parts of another community's knowledge and his own information need expressed by a given personal map. Furthermore, by exploring the relationships between clusters in a given personal Document Map and the corresponding concepts in his personal Concept Map (and vice versa) the user can discover how the unfamiliar information relates to a broader context of his personal knowledge (Fig. 7-4).

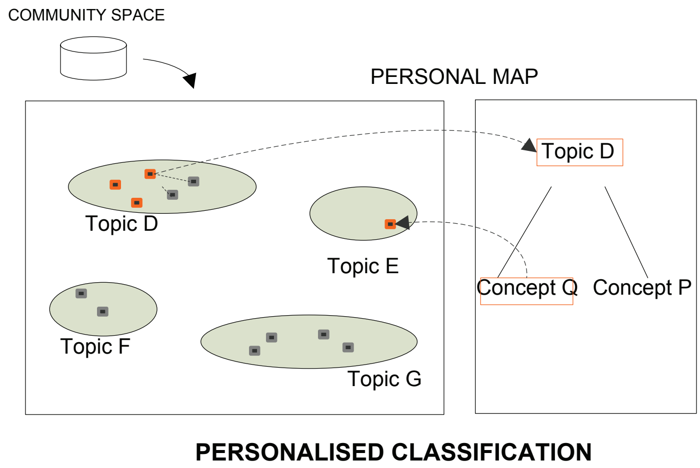


Fig. 7-4. Applying a personal map to categorize an unfamiliar information space (left) and discover relationships of unfamiliar information to personal knowledge (right).

Although the new documents are unlikely to be linked to user's concepts directly (due to terminological differences it is unlikely that there will exist matching terms) there are two important ways in which this relationship can be implicitly established. On one hand, documents which have already been assigned to some personal map by other users from the same community are likely to have some explicit links to some of the user's personal concepts as well. In case this may not hold, there is a straightforward way to infer this relationship. The vicinity of a new document to specific documents originally located in the user maps implies semantic similarity. Thus concepts to which those documents are linked can be considered to be related to new documents located near them as well. In this way the contextualisation of unfamiliar information within a personal knowledge structure can be achieved, even when no explicit relationships have been established by any other user. While such automatic classification is quite useful for finding initial documents and concepts clarifying the user's information need, there may still be too many categorized documents to be inspected in an explorative way.

At this point, the user can narrow down his exploration by issuing a specific search query. The results of this query can than be contextualized in the personal map in the same way as described in the previous section. Matching documents can be visualised in the Document Map and related concepts marked in the personal Concept Map. Documents that do not appear on the map can be dynamically classified into the appropriate cluster (Chapter 6.4.2) and the related concepts can be inferred based on their vicinity to original documents in the same way as described above. The relevant tasks and requirements in the cross-community context supported by such contextualisation are summarized in Table 7-2.

SUPPORTED TASKS	PERSONAL KNOWLEDGE MAPS
PERSPECTIVE TAKING → PERSPECTIVE MAKING	• Discovering how knowledge reflected in documents from other communities relates to one's own knowledge o document-topic relationships o document-document relationships o document-concept relationships
SENSEMAKING	• Finding appropriate concepts to express a given information need
KM PROCESSES	
KNOWLEDGE DISTRIBUTION	• Providing immediate personal benefits as a direct result of making one's own knowledge available to others
KNOWLEDGE CREATION	• Relating unfamiliar information to personal knowledge structures • Contextualising retrieved information to the needs of a given task
KNOWLEDGE APPLICATION	• Discovering relationships between retrieved information and concepts describing the short-term need of the task

Table 7-2. Tasks and requirements supported by contextualising unfamiliar information within Personal Knowledge Maps

Making personal knowledge maps useful for the map author himself provides immediate benefits which ensure intrinsic motivation of users to create personal maps and thus express personal knowledge (cf. knowledge distribution, Chapter 3.3.1). The combination of such immediate usefulness of personal maps to the author with a mechanism for their creation as a natural part of an information seeking process (Chapter 6.4.2) is an important aspect that ensures the feasibility of the proposed overall solution approach (Chapter 6.1).

7.1.3. Sharing Personal Knowledge through Contextualized Recommendations

Personal knowledge maps are structured information templates created by the users during the information seeking process. As such, they do not only reflect personal knowledge of the user but also the information needs of a specific task for which they have been created. At the same, time since the maps contain documents associated to specific concepts, they are also concrete instantiations of a given representation schema (Chapter 3.4.1).

Different personal maps represent different representation schemas that can be applied to guide the information seeking process – by suggesting which concepts to look for, by providing examples of concrete documents relevant for a specific information need represented by the map or by serving as dynamic templates for classifying unknown information. All of these are typical sensemaking tasks that are difficult to accomplish when working on ill-structured problems requiring information from unfamiliar domains (Chapter 3.4.1). In the cross-community context, this means that accessing personal maps of users from an unfamiliar community but relevant for a given information need, could provide valuable support in accomplishing the above tasks.

A prerequisite for achieving this is making the maps shareable between users and available on request in the context of a given information need. The light-weight ontological representation of personal maps supports the sharing requirement since individual maps are small (compared to full-scale ontologies), can be easily transferred and contain all the necessary meta-information for being displayed and applied by the user (Chapter 6.6) For extracting the shared community concept maps all user maps can be considered without compromising individual privacy requirements (as the origin of concepts is not revealed).

On the other hand, sharing the original personal maps themselves requires a publishing mechanism that allows users to decide which maps they want to make publicly available and which not[53]. By making map authors clearly identifiable, the reputation economy principle (Chapter 2.7.2, 3.3.1) can be activated in order to provide the necessary motivation for making one's own maps available to others.

By sharing personal maps, users from one community can thus identify maps from other communities that might be relevant for the information need of their task. This can be done manually by browsing a map catalogue or it can be supported by an intelligent search functionality that identifies the most relevant maps for a given information need and delivers them to the user. The critical question thereby is what criteria shall be used for determining the relevance of a given map for a given user in the specific context of his need.

The information need of a given user can be identified based on the search terms of his query. But as observed before, in the cross-community context the information need is inherently unclear and will often be expressed in inappropriate terms since the problem is ill-defined and the information is being sought in an unfamiliar domain. Hence, taking the search terms as the sole criteria for matching maps to a user's information need is likely to deliver poor results. In order to overcome this problem, the general interest profile of the use - as expressed in his maps created so far - can be included in the matchmaking procedure. This approach is based on the following idea: while the search query expresses a current (short-term) information need related to a specific task, the user profile reflects his general (long-term) interests, inferred from the maps created in information seeking tasks accomplished so far.

Combining these two criteria also enables us to take into account both the user's implicit knowledge and the evolution of his knowledge of the unfamiliar domain. As the information seeking process proceeds, the user's search query formulation evolves in a way that reflects the insights gained in the process: e.g. the related topics and concepts used in a given community. This will be reflected in his personal maps created throughout this process (selected documents and concepts used to name document clusters). The recommender function can then select relevant maps in two ways:

- Maps relevant for a specific search query can be identified by calculating the relevance of the search result document set for each map. The relevance calculation is based on a TFxIDF encoding of documents (Salton et al., 1994) with respect to each map known by the system. In this way a Document-Map matrix is created that contains pair-wise similarities representing a relevance of each map for a given document.
- Collaborative filtering techniques can be applied to identify similar users by taking an average of pair-wise similarities between their personal maps, represented as sets of documents. Since the documents are encoded by vectors in the document-term and document-concepts matrices (based on a TFxIDF measure), the inter-document similarities can be easily calculated e.g. by taking a scalar product. Accordingly, the averaged similarity can be taken to represent inter-map similarity. Calculating this similarity pair-wise between all maps of two different users, provides us with measure of similarity between two user profiles. This information can be used to rank or filter maps identified by document match, as well as to introduce new maps not corresponding to query but having high general relevance for the user.

A realization of such a matchmaking and recommendation service is described in Section 7.2.2.2.

[53] A simple publishing scheme allows users to publish all personal maps by default or manually for selected individual maps.

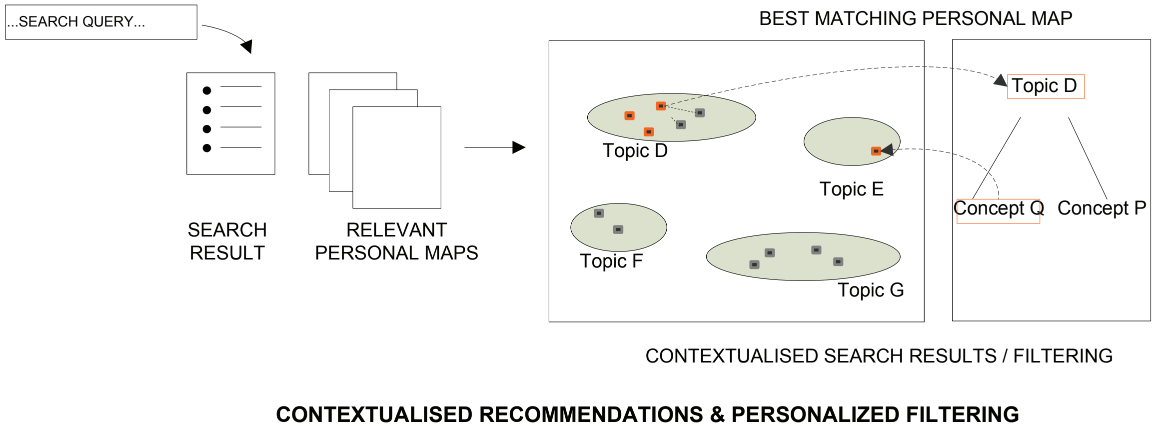


Fig. 7-5. Using personal maps for contextualized recommendations and personalised information filtering

Personal maps of other users represent both their past information needs and their personal knowledge developed in the information seeking process. As a result, contextualising results of a search query in a set of personal maps is both a way of providing personalised information filtering (delivering relevant documents for a given need) as well as a way of sharing implicit personal knowledge (delivering semantic contexts and their conceptual descriptions). Since the maps provide contextualized information related to a specific knowledge structure, the potential relevance and meaning of information is easier to establish than when accessing isolated information items. Rather than by abstract categories, the context is provided by specific concepts that have been assigned to documents by the map author as part of describing the relationship of documents to a specific information need. This makes it easier for users to recognize relationships of their information need to relevant concepts from different communities, which may be referring to similar problems in a different terminology.

The described model differs from existing approaches to using recommender systems for knowledge sharing in several ways. Most importantly, recommendations are contextualized with respect to a given information seeking process that reflects the needs of a specific user task. This respects the requirements of knowledge application and knowledge creation processes (Chapter 3.3.3, 3.3.4). Rather than recommending isolated information items, the user is provided with a set of semantic contexts that represent conceptual structures relevant for a given need and thus serve as an explanation of both the meaning of information items and the criteria by they have been chosen.

On one hand, this provides an explanation of the recommendations – a critical aspect missing in most existing approaches to recommender systems (Herlocker et al., 2000). On the other hand, the conceptual structures provided by the maps serve as an orientation in identifying better descriptions and representations of his information need (cf. sensemaking requirements, Chapter 3.4). This allows the user to identify relevant concepts in unfamiliar domains based on the previous efforts and knowledge of other users. Since maps contain concrete documents assigned to specific concepts they also provide immediate instances that may satisfy the user's information need. This supports the consumption phase of the sensemaking model (Chapter 3.4.1).

Finally, based on the personalised classification method described in Chapter 6.4.2 the maps can be applied as semantic templates for dynamically filtering relevant items from an unknown information space in a way that reflects personal preferences and implicit knowledge of the map author. In this way, personal knowledge of other users is provided in a contextualized manner that matches both the user's current need ("just-in-time" requirement, Chapter 3.3.4) and relates unfamiliar information to his own knowledge (finding maps of others related to his own personal maps). Such a way of sharing personal knowledge and making it usable for contextualized information access in unfamiliar domains respects the requirements of typical sensemaking tasks (Chapter 3.4) and the requirements for knowledge distribution, knowledge application and knowledge creation from Probst's knowledge management process model (Chapter 3.3).

SUPPORTED TASKS	PERSONAL KNOWLEDGE MAPS
PERSPECTIVE TAKING → PERSPECTIVE MAKING	• Discovering how knowledge reflected in documents from other communities relates to own concepts expressing a given information need o document-topic relationships o concept-concept relationships
SENSEMAKING	• Searching for representation schemas • Finding concepts to express a given information need • Recognizing the need for new structure • Finding an appropriate representation schema for the task • Using instantiations of representation schemas o to identify information relevant for the task o understand relationship between collected information and task structure • Applying a representation schema o to find information relevant for the task o to classify information into task structure
KM PROCESSES KNOWLEDGE DISTRIBUTION KNOWLEDGE CREATION KNOWLEDGE APPLICATION	 • Personalisation and recommendation based on information author, personal user and community contexts • Supporting reputation economy principle • Publishing service supporting privacy requirements • Providing insight into original context of information • Contextualising retrieved information to the needs of a given task • Representing personal knowledge of members from different communities relevant for a task • Discovery of relationships between retrieved information and concepts describing short-term needs of the task

Table 7-3.Tasks and requirements supported by contextualized recommendations and personalised filtering through shared Personal Knowledge Maps

7.1.4. Expressing Personal Knowledge and Relating it to the Knowledge of Others

Users express the knowledge developed in the information seeking process by organizing retrieved information into personal maps (Chapter 6.1, 6.4.1). By adding documents from maps of others to one's own personal maps and assigning them to one's own concepts, a user establishes relationships between his concepts and the concepts used by others. The personal concepts thereby reflect both the user's overall personal knowledge as well as the vocabulary appropriate for the needs of a given task.

When the maps from which documents have been taken come from the same community, this creates the necessary relationships between personal views of individual users in the community in question. Based on such relationships the Community Concept Map is constructed (Chapter 6.5.1) which presents a shared conceptual structure based on actual use of concepts from community vocabulary in organizing knowledge for accomplishing specific tasks.

In this way, expressing personal knowledge developed in the information seeking process through personal maps supports perspective making (Chapter 3.1.5). On one hand, the Community Concept Map can be used for semantic exploration and navigation of community information spaces. As discussed in Chapter 7.1.1., this is of particular importance for accessing information spaces of unfamiliar communities and the understanding of corresponding community vocabulary and knowledge structures.

More critical is the case when documents are added from personal or community maps of other communities. The discovered relationships between concepts from different communities are internalized into new knowledge by expressing it in one's own terms. In this way taking on perspectives of others is connected with perspective making within one's own community (Chapter 3.1.5). In this case, the documents contained in maps of different users become bridges between different concepts used in different communities (Section 7.1.5).

SUPPORTED TASKS	PERSONAL KNOWLEDGE MAPS
PERSPECTIVE TAKING → PERSPECTIVE MAKING	• Expressing new knowledge in one's own terms
SENSEMAKING	• Creating representation schemas • Assigning examples to representation schemas (instantiation)
KM PROCESSES	
KNOWLEDGE DISTRIBUTION	• Making personal knowledge available to others • Easy creation of context information without extra effort for the author
KNOWLEDGE CREATION	• Externalizing new knowledge and relating it to community knowledge structure
KNOWLEDGE APPLICATION	• Formulating concepts that describe the information need of the task • Expressing relationships between retrieved information and the information needs of the task

Table 7-4. Summary of tasks supported by the creation of Personal Knowledge Maps

As already discussed in the previous chapter, two concepts referring to the same document can be considered to be related from a specific point of view (Chapter 6.4.1). In principle, this information can be used to construct a cross-community concept network representing relationships between knowledge structures of different communities. This however would require sufficient amount of choices of the same documents by members from different communities in order to discover statistically significant relations as compared to the amount of assignments of documents to concepts from the same community.

A fundamentally different solution is to refrain from explicitly constructing the cross-community concept network and use interactive visualisation allowing the user to discover the relevant relationships himself relevant for the specific context of his current need. Instead of having the system choose which concepts should and which should not be included in the cross-community network, by using visualisation the user can be empowered to do so himself. This can be achieved by simultaneously displaying knowledge structures from different communities and interactively visualising relationships between them, based on the user's focus of interest. The relationships between concepts and documents established in the personal maps that have been created as part of addressing a specific information need enable such a solution.

The next section presents a method for using such multi-perspective knowledge visualisation as a means of supporting discovery of cross-community relationships and making them usable for semantic navigation and creation of new knowledge.

7.1.5. Multi-Perspective Navigation for Discovery and Creation of New Knowledge

An important requirement for cross-community knowledge exchange is the ability to locate the knowledge from one community in the context of the other (Chapter 3.1.5). Cross-community relationships established through personal maps of users from different communities provide the basis for supporting this (Section 7.1.4). But not all relationships are relevant in a given context. On one hand, a relationship between two concepts from different communities that makes perfect sense in one case may be completely irrelevant or even incorrect in another context.

Furthermore, the user's goal is to understand the meaning of unfamiliar information and its relationship to his knowledge and a specific information need. Considering all possible relationships between concepts used by the user and his community and those of others is not the subject of the user's interest. Thus, what is needed is providing a way for the user to discover the relationships between concepts used by the author (or community) of the personal (or community) map he is currently accessing and his own knowledge structure.

Addressing this need can be achieved by displaying the user's personal knowledge structure simultaneously with the visualisation of an unfamiliar knowledge map (personal or community one). The given knowledge map represents the current information context the user has selected so far - either manually or by means of the automatic matching and recommendation service (Section 7.1.1-7.1.3). The knowledge structure of the map author is presented by his personal Concept Map while the related Document Map contains the set of potentially relevant information items presented in their local semantic contexts (clusters). The personal knowledge structure of the user is represented by his personal Concept Map. It contains concepts used by the user to express and organize documents and knowledge gained in his previous information seeking tasks (Section 6.4.1, 7.1.3).

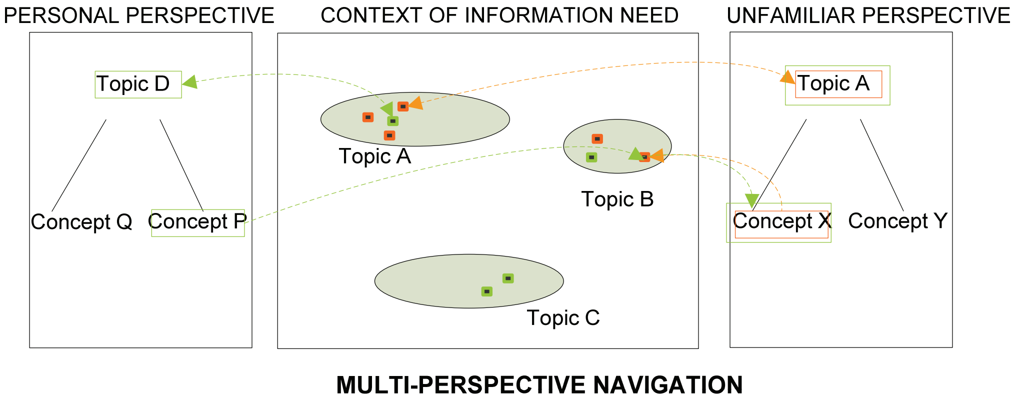


Fig. 7-6. Multi-perspective navigation for discovering cross-community relationships

The mapping of concepts between the unfamiliar perspective of the map author and the user's personal concepts with respect to a specific focus is achieved in the following way. Based on a selected set of concepts from the personal Concept Map of the map author, a set of related documents in the Document Map is determined. This set includes all documents found in the "key" relation of the selected concepts and documents which contain the selected concepts in their own "key" relations (Section 6.4.4, 6.6). A set of related concepts in the personal Concept Map of the user is then determined in two complementary ways:

- by finding which concepts from "key" relations of the selected documents also appear in the user's personal Concept Map,
- by finding which concepts in the user's personal Concept Map contain in their "key" relation any of the documents related to the selected concepts of the map author,

Establishing relationships between Concept Maps of two users from different communities in these two ways is subject to two main constraints. The user's personal Concept Map will contain links to documents from another map only when the user has already assigned some of the same documents into one his existing personal maps. This is not likely to occur very often at the individual level. The second choice compensates this by matching documents which need not have been known by the user before but that are characterized by some of the selected user concepts.

This choice however is subject to terminological differences between the two communities since the "key" relations in the Document Map characterize documents with concepts from their community of origin. Whereas in heterogeneous networks that provide the main context for cross-community knowledge exchange, some intersections are bound to exist (Chapter 2) they are likely to be very constrained at the individual level.

In order to compensate for these shortcomings, additional relations are inferred by taking advantage of the shared conceptual structure represented by the collaborative Concept Map of the user's community. Since this map connects the concepts used in the personal maps of different community members, the number of referenced documents greatly increases. The clusters of concepts in the Community Concept Map group together related concepts used in similar contexts by community members while the "key" relations of individual concepts point to documents that have been assigned to them by different users.

This information can be used to infer additional relationships between the unfamiliar concepts of the map author and the concepts in the user's personal structure in the following way:

- finding concepts in the shared Concept Map of the user's community whose "key" relations contain any of the selected documents,
- finding concepts in the user's personal Concept Map that are contained in the same clusters as the identified community concepts.

In this way, relating concepts from an unfamiliar community to user's personal knowledge takes advantage of the collaborative knowledge and previous information seeking efforts and relations established by other community members. The reverse direction of discovering relationships between concepts from the user's personal perspective and those of the map author functions in an analogous way.

Another possibility for discovering relations between the two different perspectives is based on taking advantage of the visual similarity structure of the Document Map. Selecting a set of concepts in both perspectives will result in two different sets of related documents corresponding to each set of concepts. In case when no matching concepts may be established directly by the above method they can still be inferred visually by the user, since spatial closeness reflects semantic similarity. Accordingly, the concepts related to documents positioned close to each other can also be considered to be related themselves. In other words, a range of possible relationships that have not yet been established by any other user can be discovered.

The described method of establishing a mapping between concepts mediated through documents as concrete examples of their meaning is based on the "prototype" theory of Rosch (Rosch, 1978). Rosch found out that people do not treat categories as abstract notions described by a set of rules and principles that determine the membership of an object to a category, but establish "prototypes" – concrete examples - which serve as their image of what a given category represents. An object is then assigned to a category based on how it relates to a given prototype. The mapping between perspectives of users from different communities in our approach functions in a similar way. The documents as concrete instances (prototypes) of a given set of concepts in the unfamiliar perspective of the map author determine the related concepts in the user's personal perspective.

The visualisation of such relationships supports several different tasks important for cross-community knowledge exchange. On one hand, the possibility of navigating an unfamiliar domain through one's own concepts has a high practical relevance for identifying relevant documents, which otherwise would have been much more difficult to locate. Developing an understanding of their meaning in their community of origin occurs as a natural side-effect of the process.

Selecting specific concepts reflects the focus of a user's current information need expressed in his own terms. Showing how the related documents are connected to concepts from another community allows the user to contextualise unfamiliar information with respect to the needs of his current task. Since the selected concepts are embedded into the user's overall knowledge structure, the unfamiliar knowledge is also related to his personal background. In this way the main requirements of knowledge application and knowledge creation processes are satisfied (Chapter 3.3.3, 3.3.4). On the other hand, by navigating through unfamiliar concepts of the map author the user can discover related documents from the author's point of view and in this way take advantage of his personal knowledge. The contextualization of unfamiliar concepts within the user's own perspective supports the translation of unfamiliar knowledge into his own terms. This supports the integration of perspective taking (understanding the perspective of other communities) and perspective making (translating new knowledge into one's own terms) identified as the essential process of cross-community knowledge exchange (Chapter 3.1).

The introduced approach of using multi-perspective visualisation to support the discovery and navigation of cross-community relationships offers several advantages for our specific application context with respect to other approaches. First, by focusing on a limited set of relationships relevant for a specific information need, it overcomes the complexities of explicitly constructing a shared cross-community network that relates all personal and community perspectives to each other. Second, instead of determining fixed mappings between conceptual structures based on an automated statistical analysis of document class similarities (e.g. Lacher 2003) it assists human users in discovering and expressing relationships relevant from their point of view. In doing so it incorporates both relationships expressed collaboratively by human users and supports the discovery of implicit relations based on visualising inherent similarity structures in document collections.

SUPPORTED TASKS	PERSONAL KNOWLEDGE MAPS
PERSPECTIVE TAKING → PERSPECTIVE MAKING	• Locating knowledge from one community in the context of the other • Discovering relationships between knowledge from other communities to one's own personal perspective o Topic-concept relationships o Concept-concept relationships o Document-concept relationships • Translating new knowledge into one's own terms
SENSEMAKING	• Finding relevant concepts • Using instantiations of representation schemas o to identify information relevant for the task o understand relationship between collected information and task structure
KM PROCESSES	
KNOWLEDGE CREATION	• Contextualising information to the needs of a given task • Relating unfamiliar information to personal knowledge • Discovering relationships between original information context and personal structure
KNOWLEDGE APPLICATION	• Discovering of relationships between retrieved information and concepts describing short-term needs of the task

Table 7-5. Tasks and requirements supported by multi-perspective visualisation and navigation of Personal and Community Knowledge Maps

7.2. System Architecture

Realizing the described conceptual framework requires us to implement both the methods for generating and visualising personal and community knowledge maps described in Chapter 6 as well as additional methods for connecting them into specific functionalities of the framework. This is achieved in form of a service-oriented architecture consisting of the following main layers (Fig. 7-7):

- Semantic structuring layer
- Context management layer
- Integration layer
- User interface layer

Each layer implements different services that realize specific functionalities of the framework. Using the service paradigm allows us to separate implementation and interface and thus achieve a high-degree of flexibility with regard to different methods that can be used for implementing individual functionalities of the framework. This makes the architecture easily customizable and expandable with new methods as they develop or are needed for specific application cases.

The system services are realized as Web Services which communicate between each other and with the user interface through XML-based messages, using http and SOAP as transport and communication protocols. The datapool service additionally implements also a Java RMI interface in order to minimize overhead processing for performance-critical data access, such as frequent on-the-fly fetching and display of document abstracts in the user interface triggered by user interaction and semantic zoom.

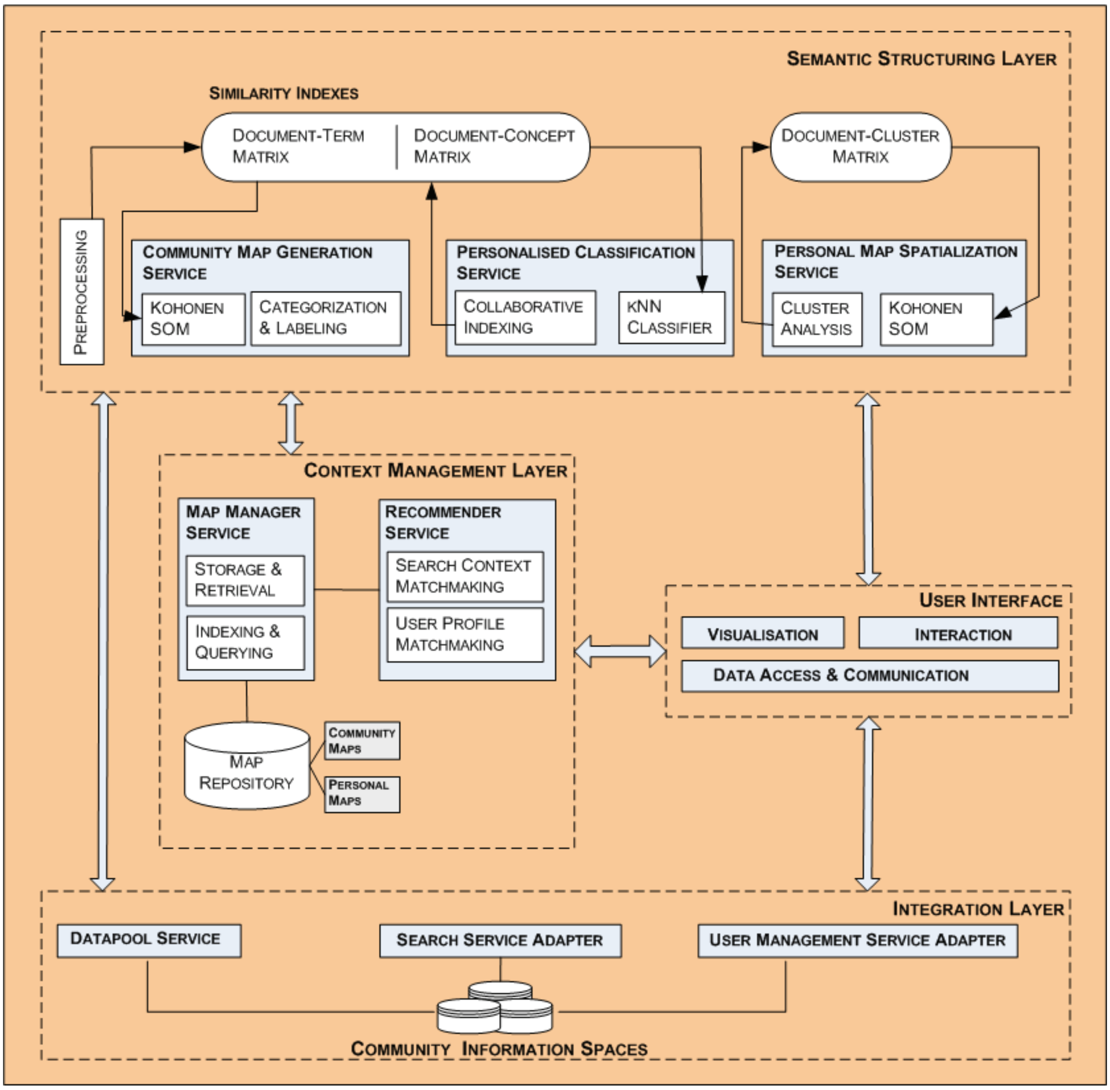


Fig. 7-7. System architecture implementing the proposed cross-community knowledge exchange framework

The implementation and deployment of the system services is based on the ODIX (On Demand Computing System) service-oriented integration middleware for cross-platform Internet application deployment (Paal et al., 2006[54]). The ODIX system is written in Java and supports multi-user and multi-tasking operation in a single JVM while synchronizing user profiles across distinct computing nodes; providing the illusion of a pervasive application environment to geographically distributed, nomadic users (ibid.). Though the actual system is Java-based it allows deployment of native executables in separate system processes, which in our case allows the system services realized in heterogeneous languages and environments (Java, C, C++, Perl) to run in an integrated application environment. The architecture of the ODIX system is depicted in Fig. 7-8.

In particular the ODIX Internet application factory (ibid., Fig. 7-9) is used for realizing the different service interfaces and composing the different module functionalities for deployment and communication with the client application (Knowledge Explorer). The client application is commonly deployed by means of Java Web Start whereby additionally using the cross-platform runtime environment provided by the ODIX Internet Workbench allows it to be seamlessly deployed and automatically configured for use on heterogeneous system environments without any user intervention.

[54] See also http://crossware.org.

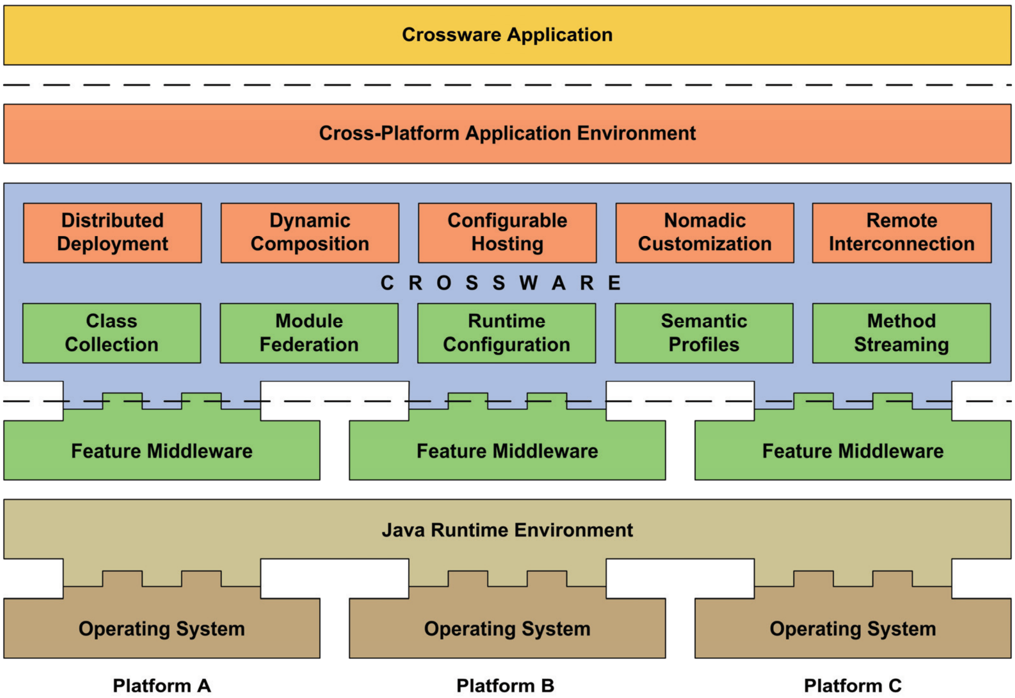


Fig. 7-8. ODIX System Architecture (Paal et al., 2006)

The Internet application factory is configured by the platform administrator by adding URIs of repositories of available modules, applications and their services alongside with runtime configurations. The system being deployed then references application descriptions for determining necessary modules and fetching them from the module repository. Finally, the Internet application factory configures an appropriate environment in which the system application is composed and executed. The client-side application (Knowledge Explorer, Chapter 8) accesses the application factory service broker for determining available services and their interfaces. Details of service management implementation provided by ODIX can be found in (Paal et al., 2006).

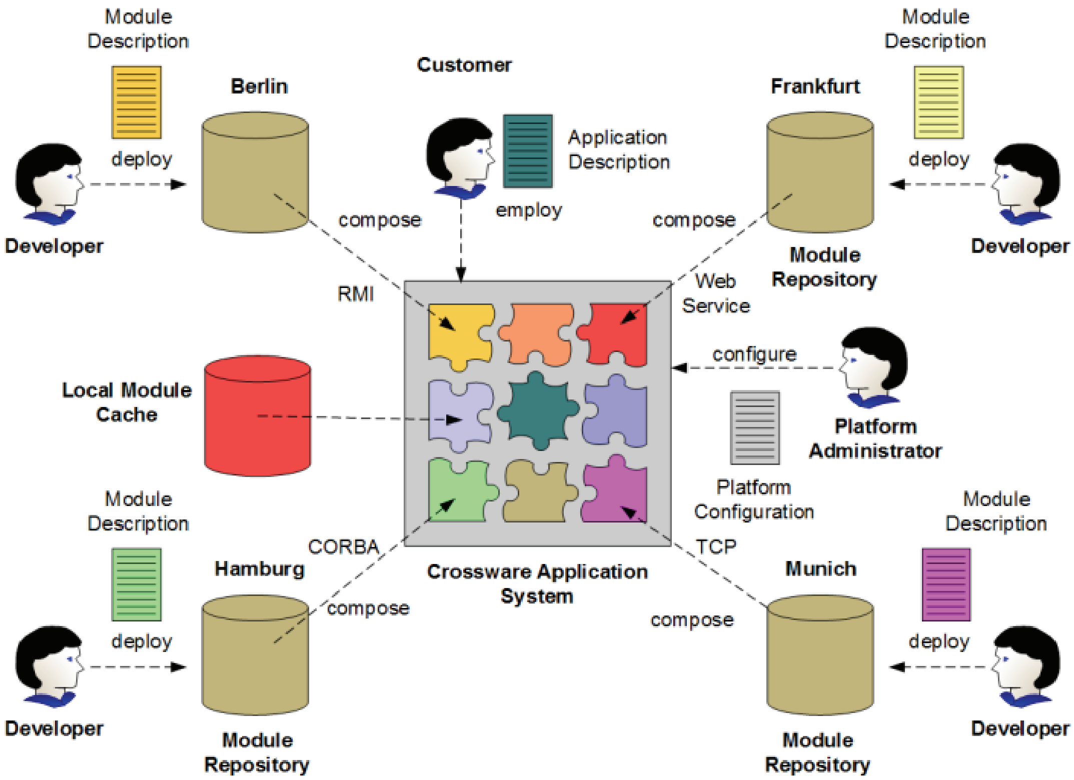


Fig. 7-9. ODIX Internet application factory operating principle (ibid.)

7.2.1. Semantic Structuring Layer

The semantic structuring layer consists of several services implementing the proposed method for automatic and personalised generation of knowledge maps representing implicit knowledge structures of individuals and communities of users[55] (Chapter 6).

This includes:

- Community Map Generation Service – implements the developed method for extracting and visualising knowledge structures based on document collections in community information spaces, such as system-generated community Document Maps and Concept Maps (Chapter 6.3),
- Personalised Classification Service – implements the developed method for learning user-defined Personal Maps and applying them as templates for personalised classification and semantic structuring of unfamiliar information spaces. It also implements the methods for extracting personal and community Concept Maps based on the analysis of personal Document Maps created by the users (Chapter 6.4)
- Personal Map Spatializer Service – implements the developed method for generating two-dimensional spatialisations of the personal Document Maps created by the classification service (Chapter 6.4)

7.2.1.1. Community Map Generation Service

This service implements the methods for generating automatic semantic overviews of community information spaces based on text-analysis and visual clustering by the Kohonen SOM (Chapter 6.3). The implementation includes a pre-processing module for encoding documents into a vector space representation, a Kohonen SOM module for creating the two-dimensional map of the document space and a categorization and labelling module that produces the final form of the Document Map containing clearly delimited clusters labelled with most characteristic terms (Chapter 6.3.2). The method for text-based automatic generation of community Concept Maps (Chapter 6.3.4) is realized analogously by the same modules. The individual modules are implemented in C++ with a service interface implementation in Java. The Kohonen SOM module is implemented by using the open source SOM_PAK package[56] (Kohonen et al., 1996).

The realized method for generating community Document and Concept Maps is an off-line method since the Kohonen SOM doesn't provide real-time performance for high-dimensional vectors used to encode the documents (for medium size collections of several hundreds to few thousands of documents the vector dimensions can range form several tens to a few hundred of terms). In our application context the generation time ranged from several minutes to half an hour, depending on hardware used (e.g. 2 GHz Pentium PC, 1 GB RAM), size of the document pool (400-2000 documents), the vector dimension and the parameters of the Kohonen SOM (resolution, number of training steps)

Accordingly, the community map generation service is initiated off-line when a community space is added to the system through a manual procedure (see Integration Layer, Section 7.2.3). The produced maps are stored in the Map Manager which delivers them to the interface and other modules per request. To accommodate for changes in the community space the maps can be periodically refreshed i.e. generated anew. If the number and size of community spaces connected to the system is contained (e.g. several medium-sized spaces), the refresh could well be synchronized with the periodic refresh of the search index. Otherwise, for a larger number of spaces the refresh period could be limited to once a month or less. The latter solution is especially attractive due to the real-time ability of the Kohonen SOM to dynamically position new entries on a previously trained map.

[55] Implemented partly within the AWAKE project, led by the author (http://awake.imk.fraunhofer.de): See acknowledgements.
[56] http://www.cis.hut.fi/research/som_lvq_pak.shtml

7.2.1.2. Personalised Classification Service

The implementation of the Personalised Classification Service according to the method introduced in Chapter 6.4.2 includes two main modules. Both modules as well as the service interface are implemented in Java. The first module implements the kNN nearest-neighbour classification algorithm (Aha et al., 1991) for using personal maps as templates for dynamic semantic classification of document collections. Its realization is based on a kNN classifier provided by the open source YALE machine learning toolkit[57]. The second module implements the developed methods for generating personal and community Concept Maps from user-defined personal document maps based on collaborative indexing and context-based similarity measures for document-document and document-concept relationships (Chapter 6.4.4, 6.5.1).

Both the personalised classification of document sets into user-defined classes (personal Document Map) by the kNN-classifier as well as the generation of personal and community Concept Maps are real-time operations. The classification takes less than a second, while the generation of Concept Maps needs a few more. This is ensured by generating the similarity indexes (document-term and document-concept matrices) once at initialization and keeping them in runtime memory. This provides a fast lookup of pair-wise similarity values used by the document classification and Concept Map generation operations.

The generation of similarity indexes is based on the analysis of personal maps contained in the map manager and the vector space encoding of the documents from the community space. This is an off-line procedure taking between several minutes and half an hour in our experimental setting with several medium-sized community spaces (400-2000 documents) and a collection of 20-100 personal maps (on a 2 GHz Pentium PC with 1 GB RAM under Linux). This initialization procedure is invoked once at system start-up. The similarity indexes are periodically refreshed whereby the refreshment of the document-concept matrix needs to be performed more often in the early stages (bootstrapping) due to the much higher dynamics and impact of the creation of additional personal maps on than the influence of additional documents in the community space. The service provides a near real-time update operation that can be invoked in runtime and takes only a few minutes for a medium-sized map collection.

7.2.1.3. Personal Map Spatializer Service

The described realization of the personalised classification service provides real-time performance of the basic personalised classification functionality, required by an interactive system. However, in order to generate an appropriate visualisation in accord with the defined Document Map visualisation model, the result of the classification service (logical assignment of documents into user-defined clusters) needs to distributed over a two-dimensional grid. This task is performed by the Personal Map Spatializer Service which implements the spatialisation method based on the Kohonen SOM, introduced in Chapter 6.4.3.

Accordingly, after the invocation of the personalised classification service the result is passed to the spatializer service, directly on the server-side (avoiding unnecessary network transport between the services and the client). The service then transforms the provided one-dimensional input (list of clusters with assigned documents) into a two-dimensional Document Map reflecting inter-document and inter-cluster similarities (Chapter 6.4.3). To this end the service implementation includes a cluster analysis module that encodes document vectors based on inter-cluster similarities and a SOM module realized by the SOM_PAK package (introduced in Section 7.2.1.1.). Although the SOM is not a real-time method, when the network resolution is small and presented with low-dimensional vectors for a limited number of documents it can achieve close to real-time performance. In the spatializer service, the SOM is presented with personal maps containing only a small number of clusters (3-9 on average, see Chapter 9) and a limited number of best-matching documents assigned by the classification service (e.g. a 100 document limit in our case). Accordingly, the dimension vector encoding the document membership to individual clusters will have only 3-9 dimensions and the SOM size of 10 is sufficient for mapping 100 items.

[57] YALE is publicly available at http://www-ai.cs.uni-dortmund.de/SOFTWARE/YALE/index.html. Implementation contributed by Michael Wurst within the AWAKE project, led by the author.

With such a configuration, the spatializer module exhibits performance of approx. 16 seconds for processing such maps (including the pre-processing of the input and post-processing of the output of the Kohonen SOM) on a 2 GHz Pentium PC, with 1 MB RAM, running Linux. While this is somewhat over the suitable performance limit required for interactive systems[58] it provides us with the basis for a hybrid solution combining direct response and caching. In addition, by reducing the number of classified documents returned per map, the performance can be significantly improved (e.g "step-by-step" model of search result deliver, where user is typically presented with 10-20 documents in each result screen, ranked by relevance) [59].

On one hand, the classification result can be presented to the user without the spatialisation by visualising the classified documents with clusters by user-defined two-dimensional placement on the screen (visualisation model one, Chapter 6.4.1) or sequentially for personal maps defined by the tree-folder model (visualisation model two, Chapter 6.4.1). In case the user desires a full visualisation which considers inter-cluster similarities he can invoke it explicitly in a next step but has to take into account the necessary waiting time.

On the other hand, the result of the spatialisation module can be cached after the first user request. To this end, the spatialisation service maintains an index of all processed requests. In this way, for the next request of the same apply operation the result can be delivered instantaneously. A reasonable assumption is that just certain maps will be called more often than others based on the aggregation of user interests in specific areas and on the amount of matching documents influencing the recommender service. This makes the caching principle especially attractive for ensuring fast response-times needed for an interactive system.

The choice of the two models can be set in a configuration of the user interface and be conveyed to the system. The first solution has been used in connection with the first Knowledge Explorer proof-of-concept prototype (Chapter 8.3), while the caching-based solution for displaying spatialized results of personalised classification service was used by the improved user interface implementation in Knowledge Explorer II (Chapter 8.5)

7.2.2. Context Management Layer

The context management layer is comprised of two services:

- Map manager service – manages the repository of personal and community maps. Provides functionalities for storing, retrieving, publishing and querying maps.
- Recommender Service – implements the methods needed for providing contextualized recommendation and personalised filtering services described in Chapter 7.1.3.

7.2.2.1. Map Manager Service

All maps created by using the system (both generated automatically and created by the users) are stored in the map repository. Different services from other layers use the map manager functionalities in order to store, retrieve and query maps required for their operations. The maps are exchanged between services in an XML-based representation format (Fig. 7-10, Fig. 7-11). The map manager service maintains a map index that allows maps to be queried by different map attributes such as user and community membership, public or private maps. The distinction between public and private maps is based on a user's decision to publish a given personal maps in order to make it shareable with others. An appropriate functionality is provided by the user interface and invoked through the map manager service.

[58] Studies of user-perceived tolerable wating times on the web report different scales of delays, ranging between 10-12 seconds (e.g. Nielsen, 1997; Hoxmeier and DiCesare 2000). Thereby, for first-time web page access longer times tend to be more tolerable. The tolerability can also increase depending on the nature of the task and degree of user interest e.g. if expected importance of information is high and the user can take care of other tasks in parallel while waiting.

[59] The processing time could also be further shortened by optimizing the process, see Chapter 11.1.

Maps in the map manager are persistent. This ensures consistency for the analysis operations required by the personalised classification service and the recommender service. It also allows the realization of the caching mechanisms ensuring real-time performance of the personalised spatializer service, described in the previous section. On the other hand, preventing unnecessary database bloat can be ensured by periodically purging the database from outdated maps (e.g. by comparing the map manager index with the personalised classification cache-index).

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<!ELEMENT map ((mapinfo, cluster+) | (cluster+, mapinfo))>
<!ATTLIST map
  learned (true | false)            #REQUIRED
  map_id  CDATA                     #REQUIRED
  subtype (SysgenMap | PersonalMap) #REQUIRED
  type    (DocumentMap)             #REQUIRED
  version CDATA                     #REQUIRED
>

<!ELEMENT mapinfo (colour | creation | grid | label | owner)*>

<!ELEMENT creation (processing-info)>
<!ATTLIST creation
  initiator_id      CDATA #REQUIRED
  processing-scheme CDATA #REQUIRED
  timestamp         CDATA #REQUIRED
>

<!ELEMENT processing-info (original-maps?)>

<!ELEMENT original-maps (mapref+)>

<!ELEMENT mapref EMPTY>
<!ATTLIST mapref
  map_id  CDATA #REQUIRED
  version CDATA #REQUIRED
>

<!ELEMENT grid EMPTY>
<!ATTLIST grid
  width  CDATA #REQUIRED
  height CDATA #REQUIRED
>

<!ELEMENT label (#PCDATA)>

<!ELEMENT owner EMPTY>
<!ATTLIST owner
  user_id CDATA #REQUIRED
>

<!ELEMENT cluster ((clusterinfo, (cluster+ | object+))
                   | ((cluster+ | object+), clusterinfo))>
<!ATTLIST cluster
  cluster_id CDATA #REQUIRED
>

<!ELEMENT clusterinfo (grid-position | icon | label | relations)*>

<!ELEMENT grid-position EMPTY>
<!ATTLIST grid-position
  x CDATA #REQUIRED
  y CDATA #REQUIRED
>

<!ELEMENT icon EMPTY>
<!ATTLIST icon
  file_id      CDATA #REQUIRED
  height       CDATA #REQUIRED
  width        CDATA #REQUIRED
>

<!ELEMENT relations (relation*)>
<!ATTLIST relations
  type (representative | key | similar) #REQUIRED
>

<!ELEMENT relation (#PCDATA)>
<!ATTLIST relation
  origin (collaborative_analysis | human_author | text_analysis) #REQUIRED
  val    CDATA                                                   #REQUIRED
>

<!ELEMENT object (objectinfo)>
<!ATTLIST object
  object_id CDATA #REQUIRED
>

<!ELEMENT objectinfo (grid-position | icon | label | relations)*>
```

Fig. 7-10. DTD of the XML-based Document Map representation used for map exchange between services

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<!ELEMENT map ((mapinfo, cluster+) | (cluster+, mapinfo))>
<!ATTLIST map
  map_id  CDATA                      #REQUIRED
  subtype (SysgenMap | PersonalMap | OverallPersonalMap | CollaborativeMap) #REQUIRED
  type    (ConceptMap)               #REQUIRED
  version CDATA                      #REQUIRED
>

<!ELEMENT mapinfo (colour | creation | grid | label | source-maps | owner)*>

<!ELEMENT colour (#PCDATA)>

<!ELEMENT creation (processing-info)>
<!ATTLIST creation
  initiator_id      CDATA #REQUIRED
  processing-scheme CDATA #REQUIRED
  timestamp         CDATA #REQUIRED
>

<!ELEMENT processing-info ((source-maps?) | (original-maps?))>

<!ELEMENT original-maps (mapref+)>

<!ELEMENT source-maps (mapref+)>

<!ELEMENT mapref EMPTY>
<!ATTLIST mapref
  map_id  CDATA #REQUIRED
  version CDATA #REQUIRED
>

<!ELEMENT grid EMPTY>
<!ATTLIST grid
  width  CDATA #REQUIRED
  height CDATA #REQUIRED
>

<!ELEMENT label (#PCDATA)>

<!ELEMENT owner EMPTY>
<!ATTLIST owner
  user_id CDATA #REQUIRED
>

<!ELEMENT cluster ((clusterinfo, (cluster+ | object+))
                   | ((cluster+ | object+), clusterinfo))>
<!ATTLIST cluster
  cluster_id CDATA #REQUIRED
>

<!ELEMENT clusterinfo (grid-position | label | relations)*>

<!ELEMENT grid-position EMPTY>
<!ATTLIST grid-position
  x CDATA #REQUIRED
  y CDATA #REQUIRED
>

<!ELEMENT relations (relation*)>
<!ATTLIST relations
  type (representative | key | similar) #REQUIRED
>

<!ELEMENT relation (#PCDATA)>
<!ATTLIST relation
  origin (collaborative_analysis | human_author | text_analysis) #REQUIRED
  val    CDATA                                                   #REQUIRED
>

<!ELEMENT object (objectinfo)>
<!ATTLIST object
  object_id CDATA #REQUIRED
>

<!ELEMENT objectinfo (grid-position | label | relations)*>
```

Fig. 7-11. DTD of the XML-based Concept Map representation used for map exchange between services

7.2.2.2. Recommender Service

The recommender service implements two modules. The search query matchmaking module is responsible for determining a set of relevant maps for a given search query, by implementing the method described in Chapter 7.1.3. The user profile matchmaking module is implements the methods for determining relevant maps for a given user, by computing the overall similarity between the set of his own personal maps and the sets of personal maps belonging to other users (Chapter 7.1.3). To this end, both modules have access to the similarity indexes (document-term similarity matrix and document-concept similarity matrix) maintained by the personalised classification service. Based on these similarities the document-map similarity indexes necessary for computing a set of most relevant maps for a given search result (document set) as well as for computing inter-map similarities are calculated and maintained by the recommender service. These indexes are initialized at service start-up and can be updated at runtime in the same way as the indexes of the personalised classification service described in Section 7.2.1.2. This ensure a fast, real-time computation of a set of relevant maps for a given search result or for a given user map collection.

In order to retrieve a set of personal maps relevant for a given search query, the set of documents returned by the search service for the query in question is passed to the search query matchmaking module. This module computes a relevance value for each personal map by applying the method described Chapter 7.1.3 to the information contained in the map-document similarity index. As a result, a list of personal maps ranked by their relevance is produced and returned to the caller (user client or search service). The service can be configured to return a list of all maps with non-zero relevance or a cut-off threshold can be set by defining a maximum number of maps to be returned.

Setting the appropriate service configuration parameter causes the user-profile module matchmaking module to be subsequently invoked. Having determined the list of best matching users, this module then extends the list of relevant maps with relevance information based on user profile matching. For each map a relevance value is added reflecting the general similarity between the personal profile of the map author and the personal profile of the user having issued the query request. This information can then be used by the client to re-rank or filter a subset of the maps identified by the search result document matching. The results are presented as a ranked list of maps in the user interface (Chapter 8).

7.2.3. Integration Layer

The presented framework is conceived as an addition to existing cross-community portals and community systems. In order for it to be applied it needs to be connected with the existing infrastructure and services. The basic functionalities that need to be provided include:

- Data services for accessing documents from community information spaces
- Search functionalities,
- User management.

To this end the integration layer implements service interfaces that must be connected with the available access points and services of a given system:

- Datapool service - provides functionalities for configuring access to external community information spaces and implements datapool handlers for access to each specific information space. Provides a unique interface to the rest of the system for document access (contents, metadata) from connected community information spaces (Fig. 7-12).
- Search service interface – provides a unique search interface for other system services. Implements search handlers for search functions of the connected community information spaces.
- User management service interface - provides a unique interface for user management functionalities to other system services. Implements handlers for user management functions of the connected community information spaces.

Integrating an existing community information space into the system then requires that three basic services be provided:

- A datapool handler for accessing documents from the information space needs to be written and integrated as a plugin into the datapool service,
- A search service wrapper for invoking the community system search function must be written,
- A user management service wrapper needs to be written that takes control of basic functionalities such as user login and access control.

In a prototypical implementation the cross-community Internet-platform netzspannung.org[60] has been connected into the system by realizing the appropriate set of services. For evaluation purposes additional community information spaces have also been integrated by providing appropriate datapool handlers and search wrappers (e.g. the ACM Digital Library archive of SIGCHI conference proceedings). The user management service provided by netzspannung.org has been integrated as the basic user management service of the system. In this way, a complete prototypical system could be realized based on the integration of the developed system architecture with a typical infrastructure provided by a real-world community system. This provided the basis for a user-centred evaluation of the developed methods and framework for cross-community knowledge exchange.

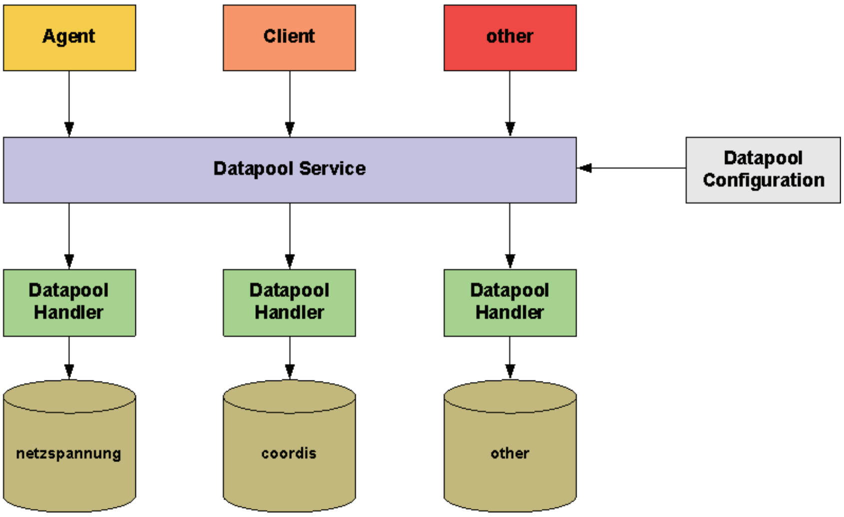


Fig. 7-12. Structure of the datapool service

7.2.4. User Interface Layer

The task of the user interface layer is to make the functionalities of the presented application framework and system architecture available to the user. At the core of our approach is the use of different kinds of personal and community knowledge maps with different modalities of visualisation and interaction to support explorative and contextualized information access. Hence, the development and design of an appropriate user interface realizing the proposed visualisation and interaction modalities and its connection with the capabilities of the underlying system architecture is one of the central issues of this work.

Accordingly, the realization of the user interface is presented in a separate chapter of its own. The next chapter shows how the functionalities of the application framework and services of the corresponding system architecture have been integrated into an intuitive, easy to use visual information interface supporting multi-perspective access to unfamiliar community information spaces.

[60] http://netzspannung.org

8. Knowledge Explorer – An Interactive Tool for Multi-Perspective Access to Community Spaces

This chapter describes a prototypical realization of the proposed method and framework for cross-community knowledge exchange in form of an interactive tool for multi-perspective access to community information spaces – the Knowledge Explorer. Specific aspects of the solutions described in this chapter have been published in (Novak et al., 2004a;b; Novak & Wurst, 2005a,b).

8.1. Design Rationale

The Knowledge Explorer is an interactive tool supporting multi-perspective access to community information spaces through personal and community knowledge maps. It has been developed to demonstrate the application of the proposed method for eliciting and visualising personal and community knowledge structures (Chapter 6) to supporting cross-community knowledge exchange (Chapter 7). The Knowledge Explorer takes advantage of a multi-view visualisation paradigm in order to make these functionalities available to users, in an intuitive easy to use, visual information interface.

Besides allowing an empirical evaluation of the proposed solutions, the Knowledge Explorer is also a special contribution of its own. It realizes an interactive tool supporting cross-community knowledge exchange which provides functionalities unavailable in other existing systems. This includes access to personal and shared knowledge structures enabling dynamic semantic structuring of information spaces from different points of view, multi-perspective visualisation enabling contextualized access to unfamiliar community spaces and the discovery of cross-community relationships.

A significant difference to other information and knowledge visualisation systems is that they focus on visualising data patterns, rather than personal and shared knowledge structures of human users, and are conceived as tools for specialized analysis tasks (e.g. data mining, analysis of specialized document collections; Chapter 4).

In contrast, the Knowledge Explorer enables the elicitation and visualisation of implicit knowledge structures of human users and focuses on its application to information access in unfamiliar domains. Thus, it is intended as a powerful and yet simple to use visual information access tool for normal users rather than specialized analysts. This poses special requirements on its design, requiring real-time performance, ease of use and interactive visualisation adapted to the needs of typical information access tasks in unfamiliar domains (cf. sensemaking, Chapter 3.4).

To achieve this, knowledge elicitation and knowledge discovery methods of the underlying method (Chapter 6) and application framework (Chapter 7) are tightly integrated with multi-view visualisation and navigation techniques in an intuitive user-centred interface. The development of an appropriate interface and interaction design has been informed both by insights from well-known information access task models (e.g. Belkin et al., 1995; Shneiderman, 1996) as well as by findings of studies on knowledge construction during information seeking in unfamiliar domains, such as (Russel et al., 1994; Qu & Furnas, 2005; see Chapter 3.4). In this way the vast amount of experience on developing visual information interfaces could be productively applied, while considering the particular needs of the specific requirements of the cross-community application domain considered by this thesis.

As a result, the Knowledge Explorer implements a specific blend of capabilities that are unavailable in other existing solutions (Chapter 4). A particular strength of its design is that it is both theoretically grounded as well as informed by practical application concerns, based on the combined analysis of theoretical frameworks and empirical studies from knowledge management and human-computer interaction (Chapter 2 and Chapter 3).

8.2. Multi-View Visualisation Model

The task of a visual information interface in our approach is two-fold. On one hand, it has to implement the necessary interface support for realizing the proposed method of eliciting and visualising implicit knowledge structures of individuals and communities of users (Chapter 6). On the other hand it needs to make those structures accessible to users and usable for supporting cross-community knowledge exchange based on the application framework introduced in Chapter 7.

The main requirements from Chapter 3 require that the elicited personal and community knowledge structures need to be presented to the user in a way which allows him/her to identify relevant knowledge from an unfamiliar community, understand its semantic context and relationships to his own knowledge and needs of a specific task. According to our application framework this requires visualising personal and community knowledge maps in a way that supports semantic exploration, contextualized access and multi-perspective navigation of unfamiliar community spaces.

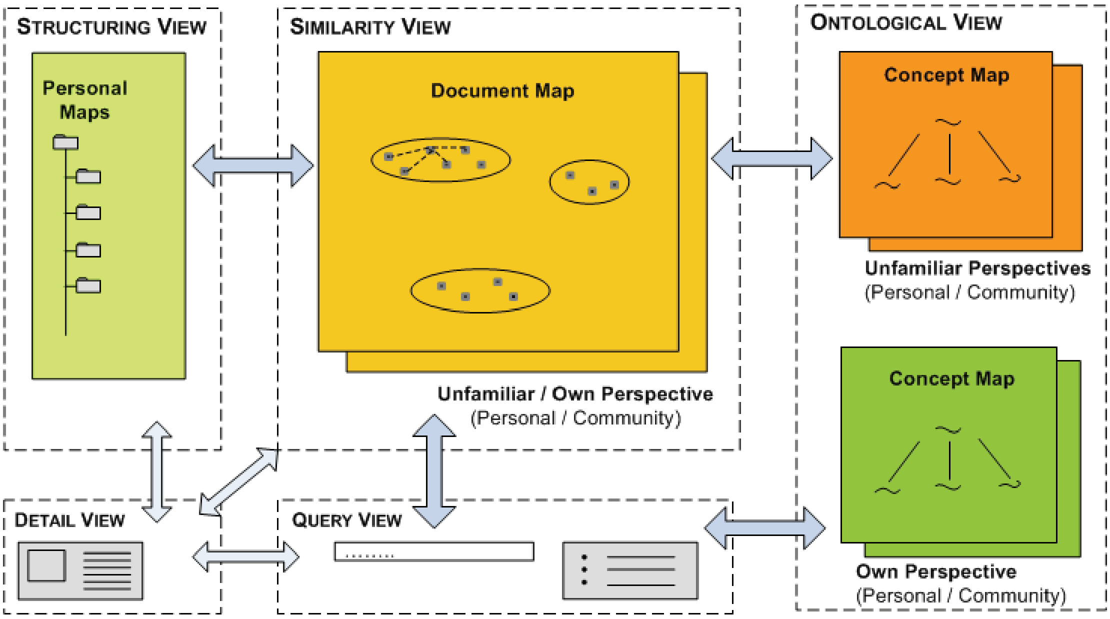


Fig. 8-1. Multi-view visualisation model for multi-perspective access to community information spaces.

To achieve this, we have developed an interactive visualisation model for representing different knowledge perspectives and relating them to each other. We employ a multiple coordinated views concept (similar to Becks & Seeling, 2004) for simultaneously providing different visualisations of personal and community knowledge structures and using them to structure, explore and navigate community information spaces from different points of view. The developed multi-perspective knowledge visualisation model (Fig. 8-1) consists of the following components:

- **Similarity View**: The inherent semantic structure of a community information space based on inter-document similarity relationships is presented in this view. Its purpose is to enable the user to gain a quick overview of main groups of semantically related documents and relationships between them. To this end the Similarity View displays Document Maps grouping documents into clusters of semantically related content while preserving global inter-document similarity relationships. This can be a system-generated map of the entire community space (Chapter 6.1) or a personal map representing a specific portion of documents reflecting personal knowledge of an individual user, relevant for a specific information need (Chapter 6.4). Furthermore, the maps can present the perspective of an unfamiliar community member (contextualized recommendations, Chapter 7.1.3), or they can present a personalised structuring of information from the unfamiliar space based on a user's own point of view (Chapter 7.1.2). Typical information access tasks supported by this view in terms of the taxonomy proposed in (Shneiderman, 1996) include overview, zooming, filter and details-on-demand.

- **Ontological View:** This view allows the user to visualise and explore relationships between knowledge structures of different individuals and communities of users represented by corresponding Concept Maps. This can be a community Concept Map displaying groups of concepts and relationships between them reflecting the shared community understanding (Chapter 6.5), or it can be a personal Concept Map representing a personal conceptual structure of an individual community member (Chapter 6.4.4). Depending on the community origin, maps displayed in this view will represent knowledge structures of an unfamiliar community or they will represent structures of user's own personal or familiar community knowledge. Through concept-based navigation in the ontological view users can identify and select related documents displayed in the similarity view and vice-versa: by selecting documents in the similarity view the user can identify related concepts in the ontological view. This allows users to navigate an information space based on a conceptual structure reflecting knowledge of human users (personal or shared) as well as to contextualize unfamiliar information into a familiar knowledge structure (Chapter 7.1.2). As different Concept Maps from different communities can be simultaneously displayed, the user can also discover how concepts from different communities are related to each other (7.1.5). Information access tasks supported by this view include filter, relate and extract.
- **Structuring View:** In this view the user can organize the results of his/her information seeking activities into personal maps. To this end it provides the necessary document and map manipulation functionalities such as: creating maps and named document clusters, adding documents from the Similarity View, moving documents between clusters, deleting and renaming documents, clusters and maps.
- **Query View:** The formulation of search queries and the list-based display of search results are accommodated in this view. This includes both the list of retrieved documents as well as the list of relevant maps produced by the matchmaking method and the recommender service (Chapter 7.1.5, 7.2.2).
- **Detail View:** This view implements the display of document details for a selected document (or set of documents) from the Similarity View.

The different views are coordinated between each other according to the "navigational slaving" principle (Baldonado et al., 2000): effects of user actions in one view are immediately reflected in dependent views. In our case, we have two main couplings of dependent views. First, the Similarity View is directly coupled to the Ontology View. As the user explores the document clusters and relationships presented in the Similarity View, the concepts relating to the set of documents currently in the users focus of attention are visualised in the Ontological View. And vice versa, a selection of a concept in a map displayed in the Ontological View results in the visualisation of related documents in the Similarity View.

The second coupling connects the Query View with the Similarity View and transitively also with the Ontology View. For each query in the Query View, the corresponding set of search result documents is visualised in the Similarity View and the set of related concepts in the Ontological View is updated as well. The same applies to the Structuring and Detail View: selecting a specific document in one of those views causes the corresponding document to be visualised in the Similarity View while its related concepts are shown in the Ontological View.

Conceptually, this coupling is represented by associations between documents and clusters, documents and concepts, and between concepts themselves. Technically, the coordination is modelled by the star-schema layout with the Similarity View acting as the coordinating view (e.g. Becks & Seeling, 2004). Whenever the user performs a selection in one view, the corresponding set of documents is determined and other views are updated appropriately. This is possible because the data-models underlying all views, associate documents as instances of knowledge resources to corresponding concepts, and vice-versa. The described model enables users to access a community information space from different perspectives. It allows simultaneous investigation of an unfamiliar community space on a document content level (similarity-based exploration), on a conceptual level (concept-based navigation) and based on relationships between unfamiliar and well-known conceptual structures (relation-based navigation)[61]. The user can choose to use any of the access methods individually and independently from the others or any given combination of simultaneous access based on the needs of a given context of use. This model allows us to realize different combinations of functionalities defined by the application framework.

8.3. General Architecture and Iterative Design Method

The general architecture of the Knowledge Explorer (Fig. 8-2) follows the model-view-controller paradigm for the development of graphical user interfaces. The data model is implemented by a data layer that connects the Knowledge Explorer to the data integration layer providing access to documents contained in the community information spaces. The data layer is also responsible for the communication with all other services implementing the functionalities of the application framework for cross-community knowledge exchange described in the previous chapter (e.g. data service, map manager, search service, personalised classification and recommender services).

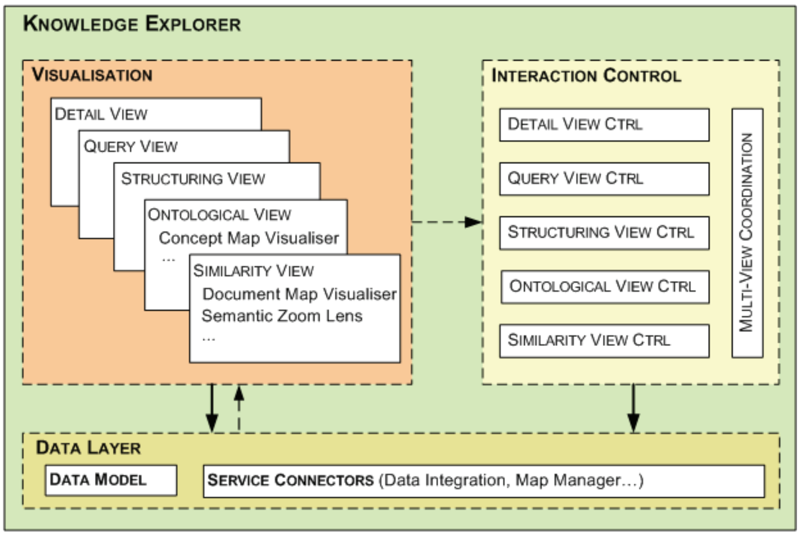


Fig. 8-2. General architecture of the Knowledge Explorer.

The visualisation layer implements different views responsible of presenting personal and community knowledge maps to the user. The abstract views of the multi-view visualisation model presented in the previous section can contain different specialized sub-views such as the Document Map visualisation (Similarity View), the Concept Map visualisation (Ontological Views), the semantic zoom lens (Similarity View) or the Personal Map Editor (Structuring View). The interaction layer implements the controllers managing the individual views as well as a specialized controller responsible for the described principle of multi-view coordination.

Based on this general architecture, two different Knowledge Explorer prototypes implementing two different realizations of the multi-view visualisation model presented in the previous section have been developed. The iterative design process of the Knowledge Explorer has been based on the user-centred design methodology (see e.g. Rosson & Carroll, 2001)[62]. The main principles of this method are:

- Active user-involvement in the design process,
- Clear identification of task requirements and user requirements,
- Iterative solution design cycle with usability testing in the process,
- Multidisciplinary approach incorporating human-factors and ergonomics techniques.

[61] Structurally, this resembles the multi-view model for exploration of relations between text repositories and databases proposed in (ibid.). Our focus is a different application domain and we implement different visualisations, interaction design and use flows.
[62] See also ISO Standard 13407.

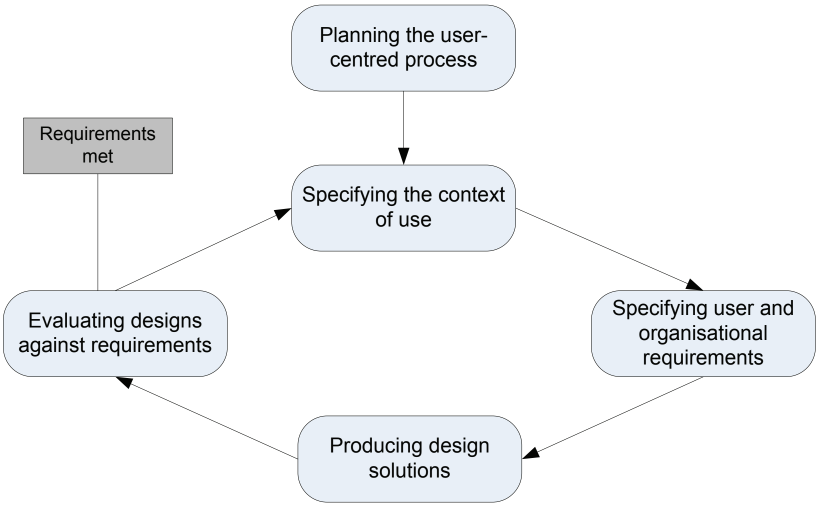


Fig. 8-3. User-centred design process according to ISO 13407

The next sections describe the two Knowledge Explorer prototypes developed in an iterative design cycle following the methodology of user-centred design.

8.4. Proof-of-Concept Prototype: Knowledge Explorer I

The main requirement for the development of a concrete interface realizing the describe multi-view visualisation model in our application context is that the use of the available functionalities need to be provided in an unobtrusive way which naturally accompanies users everyday activities of information access, without distracting them from their primary task. This is an essential difference to many other approaches that combine interactive interfaces with some subset of techniques described in our approach, but which regularly require dedicated attention of the user to the purpose of the system. Different classes of such approaches have been described in Chapter 4.5, but in the context of this chapter it is useful to remind us of some examples in order to highlight the difference of our context.

Most knowledge discovery tools with interactive visualisations are conceived and designed for users as analysts closely inspecting a given datapool to identify relationships and patterns of data as the primary task. Typical examples are interactive interfaces for clustering tools, document maps (e.g. Becks, 2001) or for combined analysis of textual clusters and relational data (e.g. Seeling & Becks, 2004).

In case of interfaces which take advantage of methods and techniques of collaborative filtering or cognitive mapping, the users either need to explicitly specify their preferences and then use the tool to access the recommendations of other users (collaborative filtering and recommender systems) or the users purpose of using the tools is to consciously map their conceptual and cognitive structures (e.g. Huff & Jenkins, 2002; see Chapter 4.3). An exception are information access interfaces with sensemaking support discussed in Chapter 4.5 (e.g. Qu & Furnas, 2005). But while these operate in the search modus, they have no connection to the broader social and task context of the user, such as the knowledge perspectives of other users and communities.

8.4.1. Organization of the Interface

The interface organization of the Knowledge Explorer proof-of-concept is based on a matrix of four multi-purpose windows with context-dependent menu buttons and a global tool panel (Fig. 8-4). There is no fixed configuration of the use of windows for accommodating the individual views of the multi-perspective visualisation model introduced in Section 8.2. The use of windows is completely flexible and defined by user actions. Each window contains a context-aware toolbar with a set of menu buttons activating the four basic functionalities: open Document Map, create Personal Map, open Concept Map, close Map. By choosing the appropriate function the user can open any of the main map types (Document Map, Concept Map, own personal Document Map) in any selected window and resize the windows at will. Opening a specific map type in a given window causes the set of menu buttons to be expanded with additional functionalities, available with the specific map type. For example, in case of the user's own personal map a save button appears in the context-aware toolbar, since this map type can be edited and saved by the user. The coupling between a Document Map and the corresponding Concept Map functions in the same way: the toolbar of a Document Map contains am icon that can be dropped in another window to display the corresponding Concept Map.

An example configuration created through such user actions is depicted in Fig. 8-4. The Similarity View containing a community Document Map is opened in the top-left window while the Ontological View with the corresponding Concept Map is displayed to its right. The lower-left window accommodates the Structuring View containing a user's own personal map with the Detail View displaying the abstract of a selected document to its right. The right-most panel is fixed and accommodates the Query View and a global toolbar for visual map interaction (panning, zooming etc.)

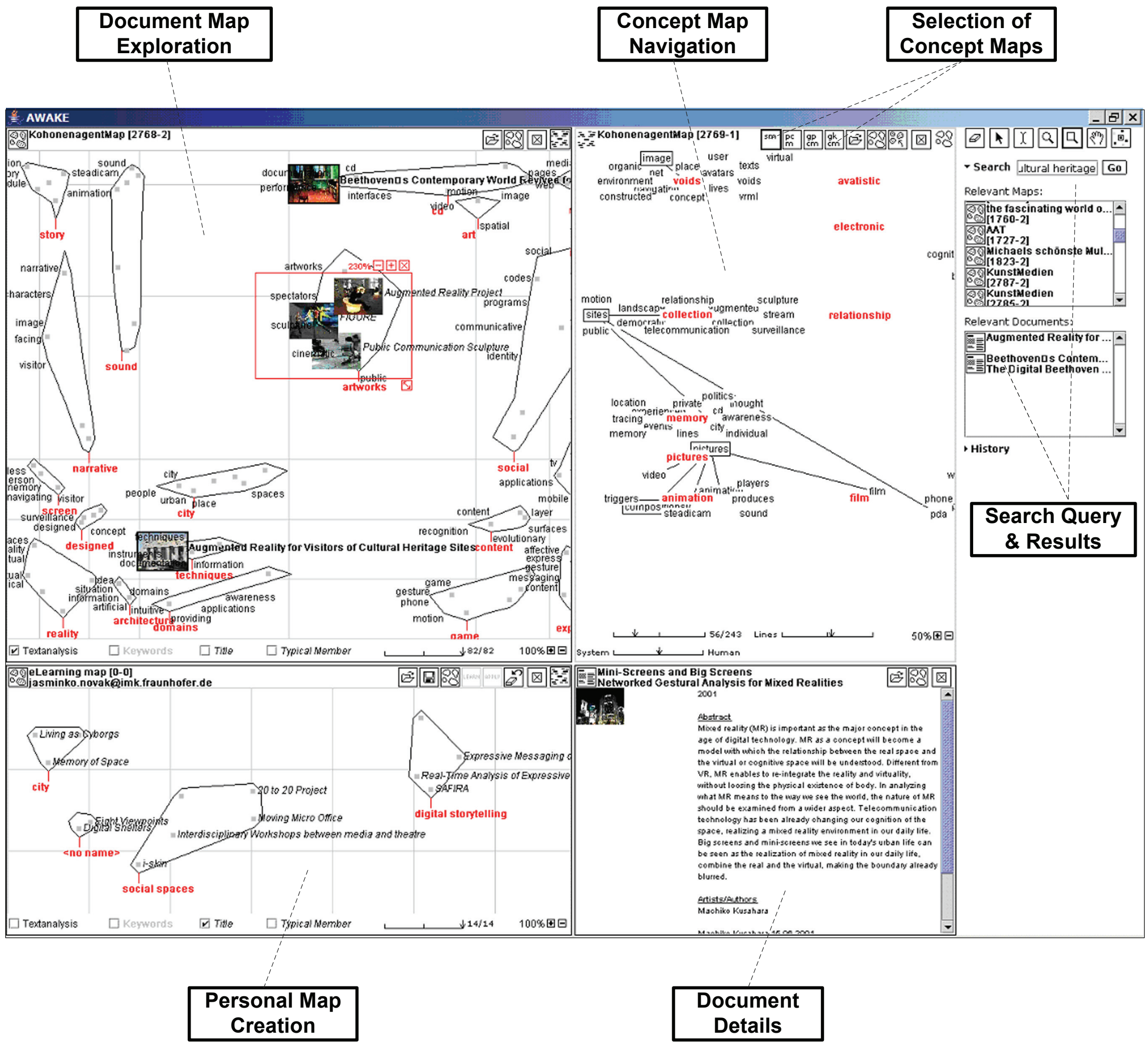


Fig. 8-4. Basic visual configuration of the Knowledge Explorer I.

8.4.2. Interaction Design and Use Flows

The interaction design in this implementation of the Knowledge Explorer provides two main use flows aimed at unobtrusively embedding the presented knowledge elicitation and visualisation model into an exploratory information interface. The overall interaction design is based on two principal modalities of information access: the 1) catalogue-browse point of entry and 2) the search-based point of entry.

Catalogue-browse point of entry. In the browse point of entry, the user opens the tool and expects to get an overview of the contents of the community space without expressing any specific context of his current information need. He browses through the catalogue of existing maps – both system-generated and created by other users – and selects an interesting one, based on the displayed map metadata. This metadata includes the map label, the names of thematic clusters (topics) contained by the map and the name of the map author.

Opening a map, a Document Map is displayed in the Similarity View, together with the associated Concept Map. The clusters of the document map give a quick overview of the topics presented by the map and relationships between them. The Concept Map visualises the conceptual structure of the knoweldge perspective of the map author. The user can now browse the topics of interest and create a new personal map into which he can bookmark the relevant documents found. The user can now inspect the search results within their thematic contexts in the Document Map, navigate by highlighted concepts or related terms in the Concept Map, He can bookmark the documents that he identifies as relevant by creating a new personal map in one of the four available display windows.

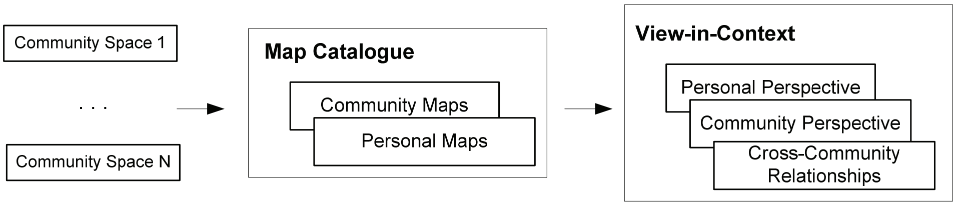


Fig. 8-5. Diagramme of the catalogue-based use-flow.

Search-based point of entry. In the search-based point of entry, the user opens the tool, logs into the system and enters a search query. The system conducts a full-text search in all available community pools and delivers a list of results. Based on the search result and on the user's personal profile reflected in his personal maps (if any exist), the system also returns a ranked list of most relevant maps from the pool of all available Document Maps (system-generated and user-generated).

The most relevant map from the list is automatically opened in the Similarity View and the search results are highlighted: the matching documents in the Document Map clusters and the corresponding concepts in the Concept Map. By default a system-generated Document Map of the entire community space is opened first. In addition the user can manually select another, personal map that he wishes to open, by simply dragging it from the list of relevant maps in any of the four windows. Such a series of actions would result in visual configuration of the Knowledge Explorer such as the one depicted in Fig. 8-4.

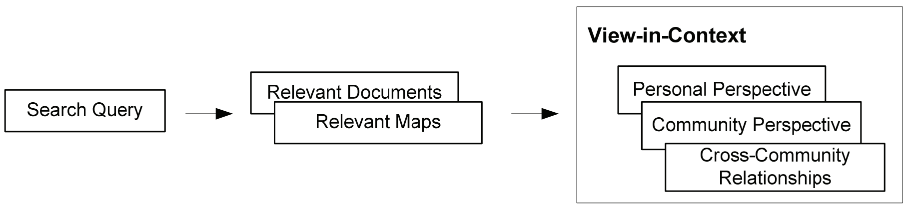


Fig. 8-6. Diagramme of the search-based use-flow

In this way, following the search query a contextual view has been created and presented to the user. It presents the search results in the context of the inherent semantic structure of a community space (system-generated Document Map and Concept Map) as well as in the context of potentially relevant knowledge perspectives of other users (personal maps). The insights discovered by the user are expresse in a personal map that he can parallely create in a free window. For each document bookmarked by the user into a named, personally-defined cluster, several relationships are established automatically: a relationship between the document and the concept characterizing the cluster, an implicit relationship between the document and search terms and an implicit relationship between a given document and other documents located in the cluster. The establishment of these relationships doesn't require conscious dedicated effort on part of the user. It is embedded in his natural task flow of using the tool as an information access interface, in a familiar modality and common metaphors, such as search and bookmark. At the same, the user actions contribute to the creation of conceptual relationships and visualisation of the underlying knowledge structures, which are commonly not expressed and thus unavailable during access to community spaces.

8.4.3. Visualisation of Document Maps and Concept Maps

The visualisation of Document Maps corresponds to the first spatial visualisation model presented in Chapter 6.3.3. Documents are displayed as points on the two-dimensional grid while cluster boundaries are drawn as convex and non-convex hulls (where needed) around documents belonging to a cluster. The corresponding cluster label is displayed below each cluster (Fig. 8-4, top-left window). This visualisation follows the idea of distinguishing between individual topical clusters as clearly as possible in order to support easy visual perception of the most salient semantic structure. The visualisation of Concept Maps follows a similar spatial visual model based on the Kohonen Map, described in Chapter 6.3.5. Instead of document points, the map displays concepts positioned on a two-dimensional grid and with main concepts displayed as labels in the centre of each concept cluster. Relationships between concepts in different clusters are shown by lines connecting the related concepts with each other (Fig. 8-4, top-right window).

8.4.4. User Interaction with Visualisations

This proof-of-concept realization of the Knowledge Explorer implemented all basic functionalities of the application framework described in Chapter 7. This include the visualisation and exploration of system-generated Document Maps of community information space, the creation and visualisation of personal Document Maps and the visualisation and exploration of personal and community Concept Maps. Applying personal maps for personalised classification and providing recommendations of relevant personal maps for a given query have also been incorporated.

This allowed all of the proposed functionalities of the application framework to be tested in an informal formative evaluation, before proceeding to the next step in the iterative development cycle and before the final formal evaluation. Since this prototype was only a first development step and the main evaluation focus was on the final version presented in Section 8.6. Accordingly, in the next sections we will briefly summarize the main interaction possibilities realized by this prototype version at sufficient level of detail for understanding the results of the formative evaluation later presented in Chapter 9. Most of the described interaction modalities are directly visible in Fig. 8-4, while others can be looked up in the description of the corresponding visualisation models in Chapter 6.

8.4.4.1. Exploration of Document Maps and Concept Maps

The main modalities of user interaction with the visualisations are selection, navigation and drag&drop. The exploration of the Document Map is supported at several different levels. The cluster labels provide a quick overview of main topical groups contained in the map. The most important terms for a given document cluster stemming from the text-analysis or from descriptions by human users can be displayed by checking the appropriate checkbox in the bottom of the Document Map (Fig. 8-4, bottom of the top left window).

The most representative document for each cluster can be displayed (Fig. 8-7, right) by checking the "typical member" checkbox located in the bottom of the Document Map. Information about an individual document can be obtained in two ways: 1) clicking on a document will cause the document abstract to be displayed in a pop-up layer directly over the document and 2) dragging a document in another window will display the abstract with additional meta-information (e.g. links to full-text and related media files).

A semantic zoom lens implements a focus+context algorithm (Sarkar & Brown, 1992) that allows a closer examination of a selected group of documents in the zoom focus, while keeping an overview of the remaining area of the map (Fig. 8-4, top left window). The lens can be resized at will and zooming-in causes subsequently titles and then document icons to be shown[63]. The default zoom behaviour can be overridden by selecting that titles be displayed at all times ("titles" checkbox) A detailed view of a specific cluster can also be gained by dragging and dropping a selected cluster in one of the other windows. The same zoom functionalities apply to the Concept Map as well. The results of a search query issued by a user are also visualised in the Document Map. The documents contained in the search results are marked as red dots with red titles (Fig. 8-7, left).

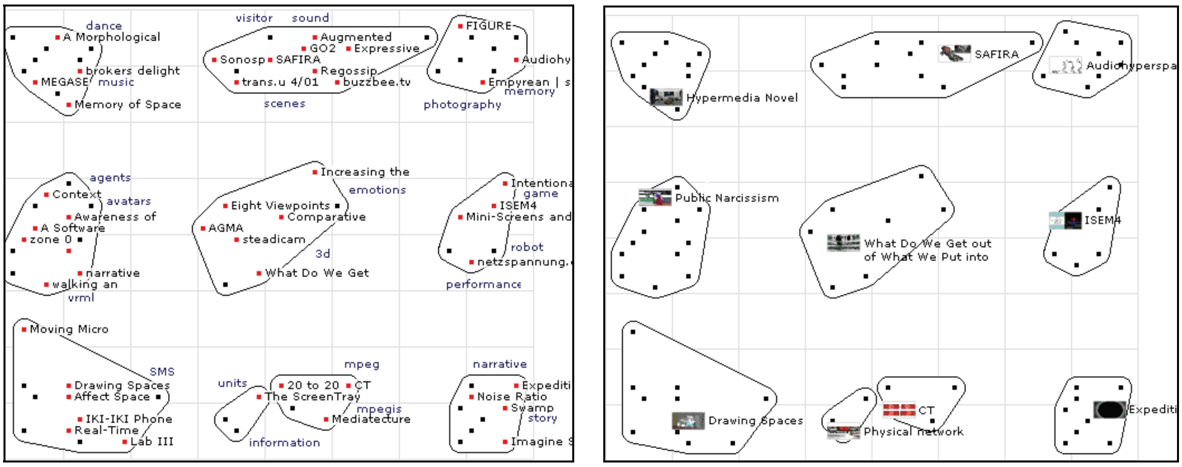


Fig. 8-7. Visualisation of search results (left) and typical members (right) in the Document Maps of Knowledge Explorer I.

(Note: In this version, the visualisation of search results did not automatically highlight also the related concepts in the Concept Map. In the formative evaluation the users explicitly remarked that they missed this functionality as they found out the possibility to manually invoke the visualisation of related concepts for a set of documents very useful (see 8.4.4.3). The subsequent version of the Knowledge Explorer presented implements this improvement)

8.4.4.2. Creation of Personal Document Maps

Relevant documents and insights on their relation to a specific semantic context can be expressed by the use through creating a personal map containing a selected set of documents organized into named clusters. By clicking the new map icon in the toolbar contained by each window (see Section 8.4.1) a new personal map is created in the given window. The user can now add individual documents or entire clusters by simply dragging them from the Document Map he has been investigating, into the newly opened personal map. In this version the documents and clusters can be positioned on a two-dimensional surface at the free will of the user (Fig. 8-4, bottom left window). The clusters can be (re)named by selecting the "text tool" from the toolbar and clicking on a cluster. Deleting documents and clusters is supported by a delete button in the map toolbar (see personal map in the lower left window in Fig. 8-4). The same toolbar contains the icon for saving the map.

8.4.4.3. Navigation of Document Maps and Concept Maps

The user can identify relevant documents in a Document Map also by navigating the Concept Map. Selecting a concept in the Concept Map marks the concept with a black border and highlights the related documents in the Document Map (grey document dots and displayed titles are coloured in red, as in Fig. 8-7). Analogously, having selected a document (or a set of documents) in the Document Map the user can invoke the function "related concepts" in the right-click context menu. As a result concepts related to the

[63] Though the idea of directly showing abstracts at further zoom levels is conceptually appealing, we found it unfeasible due to limited space and great amount of visual complexity on the screen.

selected set of documents will be marked in the Concept Map with a red border. The realization of this mechanism (matching concept key relations and document key relations) is based on the coupling between the ontological structures of Document Maps and Concept Maps (Chapter 6.3.6, 6.6).

This version of the Knowledge Explorer differentiates between four different kinds of Concept Maps: system-generated Concept Maps (produced by the text-analysis of the community space), personal Concept Maps (corresponding to specific personal Document Maps), overall personal Concept Maps (reflecting the conceptual structure for a specific user) and a cross-community Concept Map constructed through collaborative analysis of personal maps of users from different communities. The switching between the different types of Concept Maps is realized through the buttons in the top-right corner of the Concept Map toolbar. Activating a specific button causes the corresponding Concept Map to be displayed in the same window.

8.4.4.4. Using Personal Maps for Personalised Classification and Contextual Recommendations

The application of a given personal map to classifying documents from a community information space into user-defined clusters is enabled by the "learn" and "apply" buttons in the context-aware toolbar of the personal map (Fig. 8-4, lower left window). Activating the "learn" button causes the map to be "learned" by the system (Chapter 6.4.2). Invoking the apply map, the user is asked to select an information space from the list of available community spaces which should be classified by the map. The classification result is visualised in the same way as the original personal map. (Note: The formative evaluation has shown that the necessity of performing two steps for this operation was irritating for the users. This has been solved in the improved Knowledge Explorer II version).

Another way in which personal maps are made available to the user is by providing them as recommendations of relevant semantic contexts for a given search query (Chapter 7.1.5). To this end, the search panel accommodates two separate lists of results. One contains the list of relevant documents and the other contains the list of relevant maps (Fig. 8-4, rightmost panel) returned by the recommender service of the underlying system architecture (Chapter 7.2.2.2).

Having issued a search query, the user can select a personal map from the list of relevant maps to be displayed by dragging it into a free window. After the map has been opened, the user can either inspect the existing documents contained in the map or choose to apply it as a template for structuring the entire information space. In this way, when the user has identified a map of another user containing relevant documents and topics for his information need he can apply the "knowledge" of this user to identify more documents that may satisfy his need.

8.5. Informal Evaluation of the Proof-of-Concept Prototype

Following the methodology of user-centred design, the development and prototypical implementation of the proposed approach in form of an interactive system (Knowledge Explorer) has been evaluated at different stages of the design process through user workshops and usability studies. This section presents the results of an informal evaluation of the proof-of-concept prototype described in the previous section.

A practical development context and test-bed for evaluating the developed solutions through the development cycle has been provided by the Internet knowledge portal and community space for media art, design and technology netzspannung.org. Typical netzspannung users are artists, researchers, programmers, designers, curators and journalists interested in the fields of media art and design. Since the community of netzspannung.org is in itself a heterogeneous one (consisting of members from different fields of profession) good development test bed within which allows our concept to be tested even within this community itself. An informal evaluation of the Knowledge Explorer proof-of-concept prototype has been undertaken in this context.

8.5.1. Evaluation Objectives

The goal of this formative evaluation was to asses the overall suitability of the developed concept and the usability of the realized prototype, at a stage in which discovered insights could be incorporated as improvements into the development process. The main questions addressed by the evaluation were:

- How suitable is the overall concept of eliciting and visualising personal and community knowledge maps for supporting access to community information spaces?
- How useful are the individual functionalities and their implementation in the prototype?
- Which aspects of the developed solution need to be reconsidered? Are there additional, previously not identified issues that need to be addressed in order to achieve the stated design goals?

8.5.2. Test Design

A sample of 12 users was composed out of representatives of different communities found in netzspannung.org: designers, computer scientists and media art curators. They represent heterogeneous community with members from different professions where intersections between their knowledge domains exist but which have very different perspectives. Thus, we are simulating a cross-community context within the same community by taking two groups of users who share a similar field of profession. While this isn't the ideal way for formally evaluating the proposed method and approach, it does provide us with an excellent practical context for an informal evaluation of the Knowledge Explorer proof-of-concept prototype aimed at gaining input for the next development iteration.

The test users were divided in two groups of 6 users each and were asked to solve two information seeking tasks. One group was composed of users with a rather technical background in computer science (researchers, programmers, interface developers) while the other group was composed of users with a dominantly art and design related professional background (artists, curators, designers).

The test collection for solving the task consisted of two separate collections, each comprising a limited subset of the netzspannung information space. The first contained the online repository of the cast01 conference proceedings while the second contained the online repository of projects submitted to student competitions "digital sparks". Initially, each sub-collection was assigned to a different test group.

Both tasks were equally defined but the users were required to solve them in different collections. They were ill-defined information tasks with an ill-defined information need: "Create a guided tour of netzspannung.org comprising the most interesting documents from your point of view, organized into thematic contexts". Available time for each task was 1h.

First each group solved the task within the assigned sub-collection by using system-generated community maps and created a set of personal maps as a result of accomplishing the task. Then the groups changed document collections and solved the same information seeking task on the newly assigned collection. This time they used personal Document Maps created by the previous group and the resulting Concept Maps. Such setup allowed us to test both the adequacy of the functionalities available in the bootstrapping phase (no personal maps exist yet) as well as the use of collaboratively elicited structures (personal maps and shared concept networks).

8.5.3. Results of the Qualitative Questionnaire

The main goal of this test was to elicit user feedback on the suitability of the proposed concept and usability of the realized implementation in the Knowledge Explorer proof-of-concept prototype. To this end user feedback was gathered through a semi-structured questionnaire and informal interviews. The questionnaire was composed out of one 6 questions requiring qualitative response in order to elicit as much relevant aspects as possible from the users.

Table 8-1 shows the results of user responses to the question regarding the ease of use of the Knowledge Explorer. The majority of users found the Knowledge Explorer to be easy or very easy to use (61%) whereas one third of users found it difficult to use (33%).

Q: How easy or difficult was it to use the Knowledge Explorer?		
Response	**Feedback**	**User %**
Very easy	Very intuitive interaction concept	**17%**
Easy	Appealing map metaphor, intuitive interaction principle (drag&drop), visual explorative access	**42%**
Neutral	Some bugs, interface not very inviting.	**8%**
Difficult	Complex tool, missing help, bugs.	**33%**
Very difficult		**0%**

Table 8-1. User responses regarding the ease of using the Knowledge Explorer proof-of-concept prototype.

On the other hand, the majority of users who perceiving the Knowledge Explorer easy to use, emphasized the intuitive and consistent interaction metaphors (drag&drop), the appealing and easily understandable map metaphor and visually-oriented explorative access.

Q: Something you particularly liked about the Knowledge Explorer?		
Preferred Functionality	**Feedback**	**User %**
Parallel Views	Four windows practical	**42%**
Semantic Zoom	Zoom liked, very practical	**42%**
Concept Maps	Concept Map very good, helpful	**33%**
Personal Document Maps	Personal Maps and apply function practical, good, interesting	**25%**
Metadata Visualisation	Keywords, text-analysis and title visualisation practical	**25%**

Table 8-2. User feedback to question "Something you especially liked ?"

The user responses to the question on aspects they liked were evenly distributed (Table 8-2). The users especially liked the functionalities of the semantic zoom and parallel views (42% each). The usefulness of personal maps and the apply function was emphasized by 25%, the Concept Map by 33% and the integrated display of document metadata into the map visualisation by 25% of test users. This lack of predominance of any single functionality confirms the suitability of the overall system design. Especially important here is the emphasis of the multiple views concept by 42% of users which is the highest percentage of user agreement.

Q: What functions did you find useful for completing the assigned tasks ?	
Concept Map	**33%**
Apply Personal Map	**25%**
Semantic Zoom	**25%**
Search	**17%**
Parallel Views	**17%**
Abstract Visualisation	**17%**

Table 8-3. User responses to the question „What functions were useful for the tasks?"

Similar results can be observed by user responses to the question about functions they found useful for completing the tasks. This question was used to verify the results of the previous question and differentiate between users liking particular features and finding them useful for accomplishing a specific task. The overall distribution of answers shows large correspondence which cross-validates the results of both tables. The only notable difference is that much less users nominated the semantic zoom as having been useful than users who expressed it as a particularly liked functionality (25% in Table 8-3 vs. 42% in Table 8-2). The provided feedback offered no explanation for this effect which could be due to chance as is suggested by the high amount of users providing a positive assessment of the semantic zoom, when asked specifically about the usefulness of individual functions (Table 8-4).

Q: How useful did you find…?				
Function	**Positive**	**Negative**	**Neutral**	**Didn't try**
Personal Document Maps	**58%**	**25%**	**17%**	**0%**
Concept Maps	**33%**	**9%**	**16%**	**42%**
Semantic Zoom	**50%**	**33%**	**8%**	**8%**

Table 8-4. User feedback to "How useful did you find….?"

The positive user feedback to the overall concept is further strengthened by the rising agreement when explicitly asked about the usefulness of the three main classes of functionalities. Table 8-4 shows the distribution of answers clustered into categories of positive, negative and neutral answers. For example, while a special preference for personal Document Maps was emphasized by 25% users in the open-end question, when asked specifically about the usefulness of personal Document Maps, 58% gave a positive feedback. Somewhat more problematic has been the use of Concept Map. While more users judged it positively than negatively (33% vs. 9%) a very high percentage didn't use it at all (42%).

Full user responses and informal feedback explain why: these users either didn't try this function or provided no answer. Most pointed to the lack of time for learning to use all functionalities. If these users are not taken into account positive feedback rises to 80% This implies that users who passed the learning curve were satisfied with the use of Concept Maps. The overall positive result is also supported by user responses to the question whether they found relevant documents outside of their community. Here 67% of test users gave a positive response. Half of them attributed this to Concept Maps and the other half to personal Document Maps.

Q: Something you didn't like about the Knowledge Explorer?	
Feedback	**User %**
Non-intuitive cluster renaming tool	58,33%
Interface complexity / learning curve	33,33%
Bugs	33,33%
Abstract visualisation non-appealing	25,00%
User help missing	25,00%
Concept Map visualisation not practical	16,67%
Constant tool-switching	16,67%

Table 8-5. User points of critique of the Knowledge Explorer proof-of-concept prototype.

Such results are inline with user points of critique (Table 8-5). Interface complexity and learning curve have been noted by 33%. Other issues included a non-intuitive cluster renaming (60%), bugs (33%), visually non-appealing abstract display (25%), missing user help (25%) and the necessity for repeated tool-switching (16%). Finally, the users were asked whether in accomplishing the tasks they identified any projects outside of their common domain of interests and if so, by means of which functionality. The results in Table 8-6 show that two-thirds of users (67%%) declared to have found interesting information outside of their common domain of interests. One half of the users who gave a positive answer declared to have achieved this through the Concept Maps, almost entire other half of the users declared the Document Map visualisation of clusters as the primary means of such discovery with one users attributing his discovery to using the apply function of personal Document Maps.

Q: Did you find interesting projects outside of your common interests ? If so, how ?		
Response	**Function used**	**User %**
Yes	Concept Map	33,33%
	Document Map Clusters	25,00%
	Apply Personal Map	8,33%
		66,67%
No		**16,67%**
No Answer		16,67%

Table 8-6. User responses to the discovery of information outside their common domain of interests by using the proof-of-concept Knowledge Explorer prototype.

8.5.4. Design Implications

While the evaluation results confirmed the overall approach and application concept, they also revealed several important critical insights with implications for the design of the Knowledge Explorer. The proof-of-concept prototype was originally designed with a strong emphasis on a catalogue-based point of access. The assumption was that when accessing an unfamiliar community space, users would prefer an explorative access by opening the list of maps of available community spaces and choosing the one to open for browsing.

In contrast, informal user feedback and user observations have shown that users approach the problem by formulating a search query as a point of access, rather than starting with explorative browsing. Explorative behaviour ensues only after initial search results are contextualized in the map. This is in accord with the “berry-picking” model of information seeking behaviour (Bates, 1989). The interaction design for the user flow of the interface should be organized accordingly and place more emphasis on the search-based access.

With regard to individual functionalities, the Concept Maps required the most learning effort. Discussing this with the users revealed that the number of different Concept Maps accommodated all as equally important in one window (with button switching) was confusing: it wasn’t clear what kind map offered what kind of advantage and use with respect to each other. As a result the system-generated concept map based on the text-analysis was the most used one, for the simple reason of being the first Concept Map to be displayed.

Accordingly, the positive user feedback on the use of Concept Maps refers to this map and confirms that it is a good bootstrapping solution. The suitability of personal and collaborative concept maps could only be partly tested based on user feedback. Due to limited time, with respect to the learning curve, many users did not try the personal concept maps. This was a topic of more specific assessment in the summative evaluation (Chapter 9).

Especially important feedback regards the cross-community Concept Map which was also included test-wise in the tested prototype solution. This map was explicitly discussed in the follow-up group discussion to compensate missing feedback from direct usage. The obtained user feedback has shown that mixing two community vocabularies into a shared cross-community map is perceived more confusing than useful. Such a map suffered from too high ambiguity and with too many possible but not directly intuitive relations, since they connected conepts from different communities. The users missed a clear point of reference for navigating such a map and voiced a strong preference for separate Concept Maps of different communities and means for discovering relationships between them.

Overall, such feedback pointed to the need to introduce a more specific distinction and visual organization of the different Concept Maps (personal vs. community) and the respective benefits of using them for navigating an unfamiliar community space. This has been addressed in the improved Knowledge Explorer design by introducing two dedicated Concept Map views, with clear indication of their origin, the conceptual difference in their purpose and benefits provided to the user (Section 8.6).

The multi-view paradigm was emphasized as very useful by a number of users. But the possibility to use each view for display of arbitrary elements and kinds of visualisation has caused some confusion. Users expressed preference for a fixed-functionality layout, where each window has a specifically predefined purpose. This is also a natural way of reducing interface complexity.

The simultaneous display of different spatially-organized visualisations imposes an increased cognitive burden on the users: it decreases orientation and thus the usefulness of individual visualisations rather than supporting it. Users tended to focus on two visualisations at a time appropriate for a specific stage in the use flow: e.g. moving documents from system generated document map into personal maps focused on those two views, navigating the concept map placed focus on the concept map view and the associated document map. Several users proposed to use folder-based visualisation for organising personal maps in order to reduce visual complexity.

Accordingly, the visualisation model based on tree-folder visualisation for the concept maps (Chapter 6.3.4, 6.4.5) and for the creation of personal maps has been introduced (Chapter 6.4.1). The interface organisation of the Knowledge Explorer has been modified to allow only one spatially-organized visualisation at a time and individual views have been assigned fixed functions. The search bar has been moved from utmost right column to the foremost left column to emphasise search as the entry mode of using the system.

Overall, the interaction design has been much more streamlined, leaving less freedom for the user but providing more guidance and drastically shortening the learning curve (customizability vs. simplicity trade-off). The development process of the new design included cognitive walkthroughs with test users which have confirmed the suitability of these decisions. The final design of the Knowledge Explorer prototype incorporating the obtained insights is presented in the next section.

8.6. Final Prototype: Knowledge Explorer II

Based on the results of a formative evaluation of the proof-of-concept implementation of the Knowledge Explorer presented in the previous section, a new improved version of the prototype has been developed. A detailed discussion of the results and lessons learned from the formative evaluation is presented in Chapter 9. The improved design and implementation of the Knowledge Explorer II prototype are based on three main findings:

- the need to reduce the perceived visual complexity of the interface caused by too many simultaneous spatial visualisations,
- the need to increase “at-a-glance” comprehension of the overall interface organization and available interaction functionalities (felt to be diminished by the flexible possibility to activate any functionality in any window of the first prototype),
- the need to more clearly support the discovery of relationships between the unfamiliar community knowledge structure and user’s own familiar knowledge.

This version of the Knowledge Explorer provides the basis for the main evaluation of the proposed approach and specific hypothesis regarding the use of knowledge visualisation for supporting cross-community knowledge exchange. Accordingly, the following sections provide a detailed presentation of its design and implementation of the proposed method of collaborative knowledge visualisation (Chapter 6) and related functionalities for supporting cross-community knowledge exchange by facilitating access to unfamiliar community information spaces (Chapter 7).

8.6.1. Organization of the Interface

Fig. 8-8. depicts the organization of the interface in the Knowledge Explorer II prototype. The main principles of this implementation are the pair-wise simultaneous visualisation of two different kinds of perspectives on an unfamiliar community space and a simplified organization of different views in a predefined visual structure.

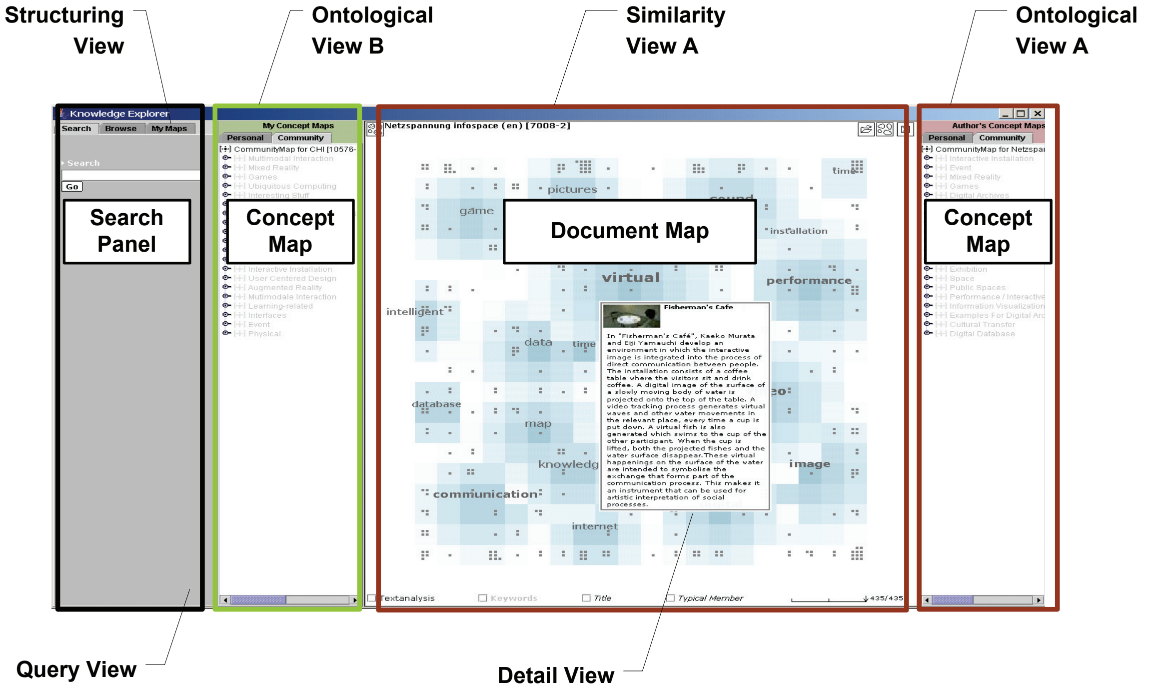


Fig. 8-8. Organization of the interface in the Knowledge Explorer II prototype.

The Similarity View A and Ontological View A located in the central and right-most panels (respectively) present a semantic structuring of an information space with respect to the personal and shared knowledge structures of the community to which the information space belongs. The Similarity View displays personal or community Document Maps whereas the corresponding Concept Maps are displayed in the related Ontological View.

The other ontological view (Ontological View B) located to the left from the Similarity View presents the knowledge structures familiar to the user, represented by his personal Concept Map or the community Concept Map of his community. In this way, the user can easily switch or combine different access modalities to an unfamiliar community space: exploring the viewpoint of the community in question, navigating the space from his familiar perspective or using simultaneous navigation in both structures to discover relationships between them.

The Query View and the Structuring View are located in the left-most panel with register tabs allowing seamless switching between them, based on the current context of user actions. The same register tab principle is used for the switching between personal and community Concept Maps within individual Ontological Views. The Detail View is realized as an additional layer incorporated directly over the Similarity View.

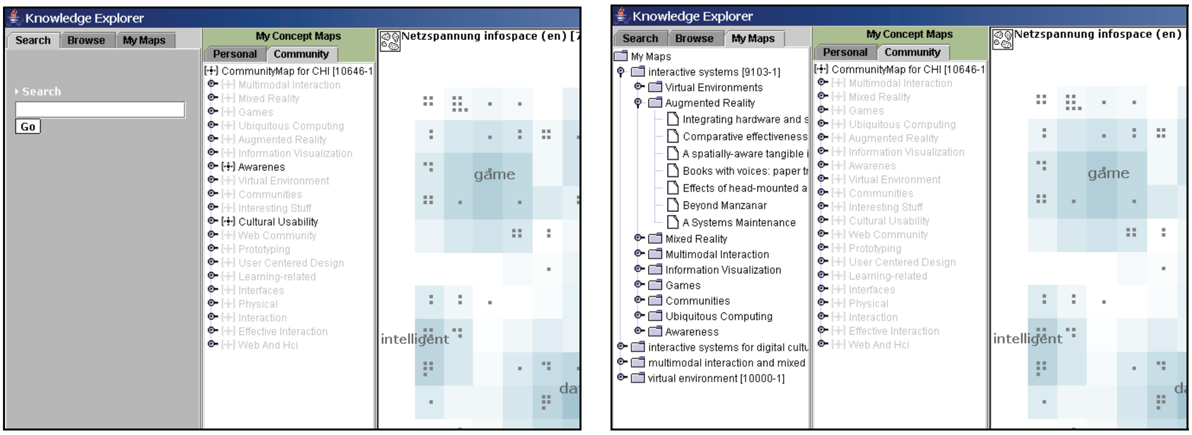


Fig. 8-9. Switching between the Query View and the Structuring View

8.6.2. Visualisation of Document Maps and Concept Maps

The Document Maps are visualised according to the second visualisation model presented in Chapter 6.3.3. Instead of drawing clear cluster boundaries by polygonal shapes, the clusters are coloured with a linear decreasing fill from the cluster centre. The cluster centre is calculated as the centroid of the polygon formed by the boundaries of outer cells of the cluster. The Concept Maps are visualised according to the tree-folder visualisation model introduced in Chapter 6.3.5 and 6.4.5: each concept cluster is represented by a folder labelled by the corresponding main concept. Sub-concepts are contained in the concept folder with interactive display of inter-concept relationships based on user focus selection. An overview of these two visualisation models is given in Fig. 8-10. As specific details of individual visualisations are related to the interaction and navigation modalities in using the maps, they will be discussed in the appropriate contexts throughout the following sections.

8.6.3. Main Interaction Modalities and Use Flow

Similarly to the realization of Knowledge Explorer I, the interaction design of Knowledge Explorer II also provides two main modalities of information access: 1) the search-based point of entry and the 2) catalogue-browsing point of entry. The catalogue browsing point of entry is the same as in the previous version of the Knowledge Explorer described in Section 8.4. The main difference of this version is in the simultaneously available visualisations of different kinds of Concept Maps and the interaction between them.

In the search-based point of entry, the user opens the tool, logs into the system and enters a search query. The system conducts a full-text search in available community spaces (which can be selected in a configuration file) and delivers a list of results consisting of a list of documents and document maps relevant for the users query. The maps are matched to the user's query and personal profile expressed through his own personal maps, by the matchmaking and recommendation service described in Chapter 7.1.3 and 7.2.2.1. The most relevant Document Map from the list is automatically opened in the Similarity View and its corresponding Concept Map is opened in the right-most Ontological View. The user's personal Concept Map is opened in the left Ontological View. The system configuration can be set to prefer personal or community Document Maps to open by default. The user can revert this choice by manually switching between the two or opening a specific map that grasps his attention.

The search results are then visualised in the active maps in the following way: the set of documents contained in the search results is highlighted in the Document Map and a set of related concepts is highlighted in both Concept Maps. The set of related concepts is determined by matching occurrences of the selected documents in key relations of concepts contained in a given Concept Map as well as by matching occurrences of concepts from a given Concept Map in key relations of the selected documents in the Document Map (see Chapter 6.3.5, 7.1.5).

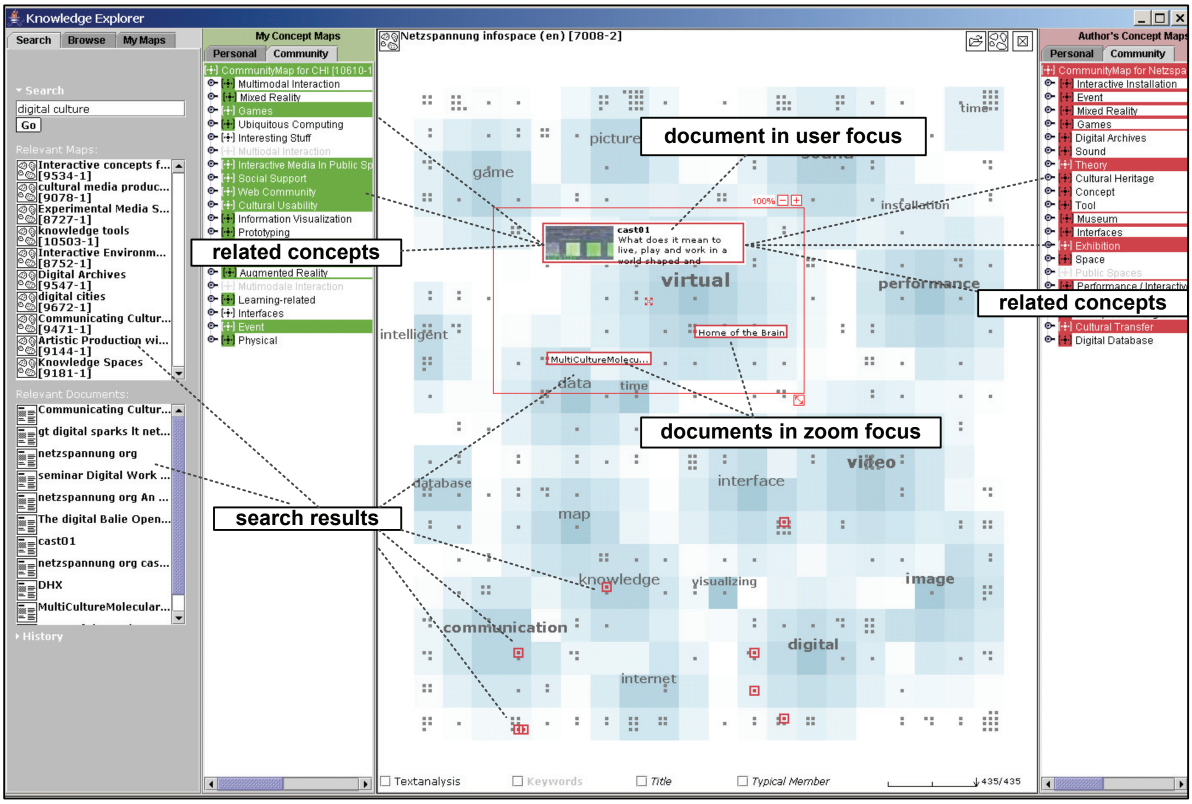


Fig. 8-10. Basic visual configuration of the Knowledge Explorer II: Contextualisation of an information need within multiple knowledge perspectives.

As a result, by issuing a single search query the user is presented with a rich semantic context. On one hand, the documents relevant for his expressed information need are contextualized within the inherent semantic structure of an unfamiliar community space as well as within the conceptual structures representing implicit knowledge structures expressed by members of the community in question. On the other hand, the visualisation of relationships to a user's own personal or community structure relates the information from the unfamiliar domain to a familiar domain of knowledge. The user can now inspect the search results within their thematic contexts in the Document Map navigate by highlighted concepts or related terms in the Concept Map or create a new personal map to save and organize relevant documents identified through his exploration. Fig. 8-10. shows this basic visual configuration of using this version of the Knowledge Explorer.

The next sections describe the specific modalities of interaction with individual visualisation functionalities and services of the underlying application framework in more detail. The realization of these functionalities follows the guideliness for the design of visual information interfaces proposed by (Shneiderman, 1996): "overview first, then zoom, filter, details on-demand, relate, history and extract". Overview, zoom, filtering and details on-demands functionalities are supported directly by the Document Map and its connection to the Detail View. The relationship discovery tasks are supported by the Document Map – Concept Map coupling. History is provided as part of the Query View. The "extract" tasks in our model refer to the creation of personal maps that externalizes discovered relationships and insights acquired by the users and are accommodated in the Structuring View.

8.6.3.1. Exploration of Document Maps

As can be seen in Fig. 8-10, the Document Map visualisation is the central visual element of the interface. It provides a quick overview of document clusters in the community information space (community maps) or a special portion of it (personal maps) and accommodates the main functionalities for accessing, selecting and inspecting individual documents and document clusters. The implementation of these functionalities follows the "overview – zoom – details on-demand" principle in the following way.

The overview of the topical structure of the document space is provided by cluster labels representing the most characteristic term for the cluster in question. The size of the clusters in combination with the display of document points gives an indication of relative cluster density and overall document distribution to different topical areas. The user can narrow a focus of his interest to a group of documents by using a focus+context zoom lens or by selecting individual documents (represented as grey dots) through point and click.

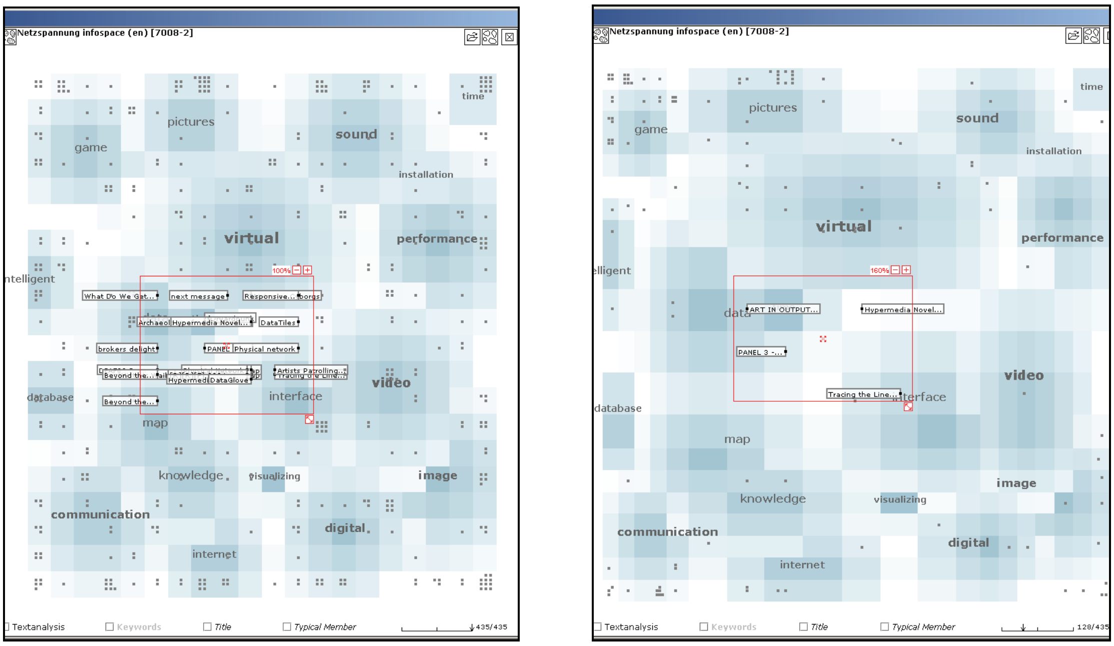


Fig. 8-11.Semantic zoom and semantic filtering in the Document Map.

Moving a zoom lens over a group of documents (red rectangle in Fig. 8-11, left) causes short titles (max. 20 characters) to be displayed for all documents in the zoom focus (Fig. 8-11, left). With increasing zoom the titles of documents with a greater semantic weight describing the degree of document membership to a given cluster.are displayed first. This mechanism is activated when a maximum number of documents that can be displayed simultaneously in the zoom rectangle is exceeded.

Accordingly, the zoom lens is referred to as the semantic zoom. The assignment of screen space between the zoom focus and the remaining area of the map is based on a variation of a well-known graphical fisheye technique (Sarkar & Brown, 1992). The selected zoom level is applied to the zoom focus area which can be resized at will. The rest of the map is applied lower zoom levels decreasing sharply with distance from the zoom centre. In this way, the user can focus on a specific area of interest while keeping an overview of the map as a whole (Fig. 8-11, right). Visual clutter is avoided by retaining the overview only for cluster labels, whereas titles are displayed only for documents in the zoom focus, regardless of the zoom level in neighbouring areas.

The user can also limit the number of documents to be displayed in the whole map by setting the semantic filter to a desired number of documents (Fig. 8-11, right; slider in lower right corner of the maps). Moving the slider to lower levels causes documents with a lower semantic weight to be removed from display first. In this way, the user can filter out potentially less relevant documents and focus on overall more characteristic documents first. Such a combination of zooming and filtering allows the user to easily specify and display meta-information for his focus of interest while avoiding visual clutter.

The display of further document information is accommodated by different levels of details on-demand. Pointing on a document displays a short preview with full document title, miniature icon and the first few words of the abstract. Clicking on a document displays the abstract in an additional layer over the map while double-click invokes the display of full document detail page in the external browser (Fig. 8-12, left to right).

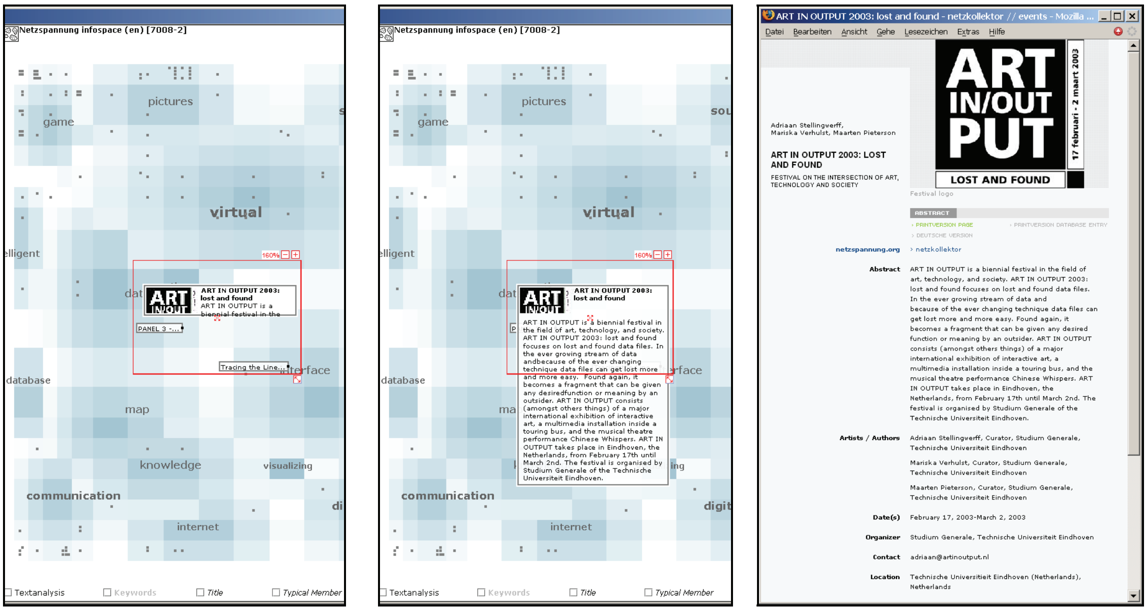


Fig. 8-12. Basic exploration modalities in the Document Map: overview, semantic zoom, filtering and different levels of document details on-demand.

8.6.3.2. Visualisation of Search Results in the Document Map

As already outlined in Section 8.6.3, the visualisation of search results occurs on two levels; the documents from the result set are highlighted in the Document Map whereas related concepts are visualised in the active Concept Maps. We thereby differentiate between different visualisation modalities depending on document relation to current user focus.

The visualisation in the Document Map follows the user focus defined by the zoom and filtering principle described in the previous section. First, all documents contained in the search results are displayed as red dots with a red border. Then, as the user positions his focus of interest by moving the zoom rectangle, the documents contained in the current zoom focus are displayed as short titles with a red border.

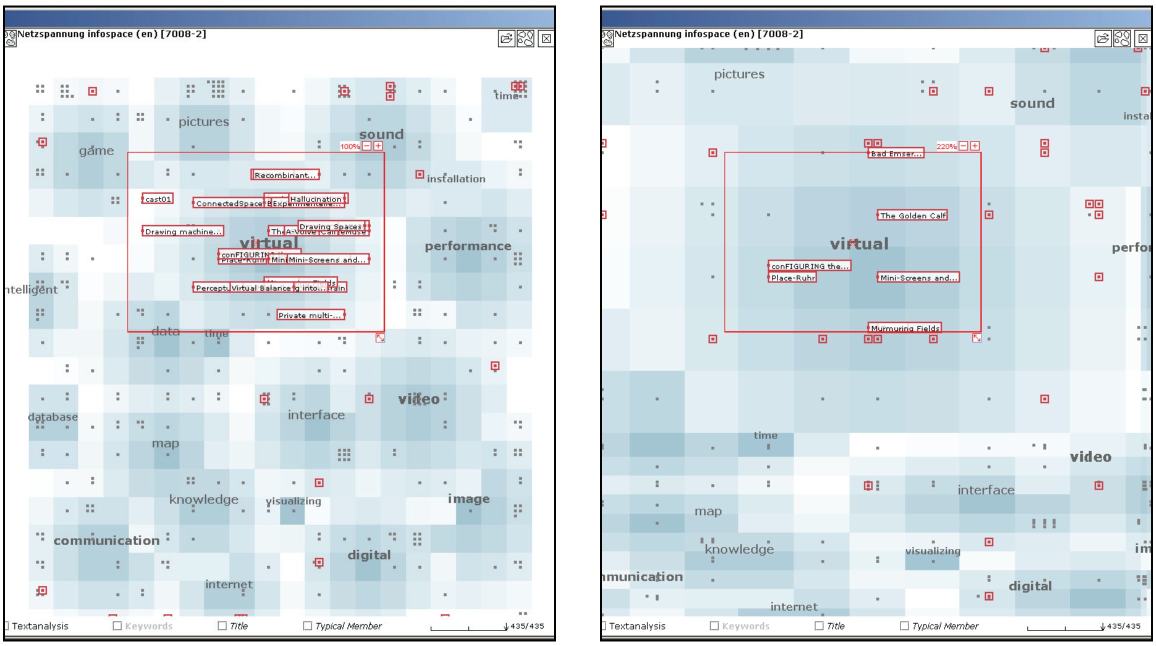


Fig. 8-13. Visualisation of search results and semantic zoom in the Document Map.

The visualisation of relevant documents in user focus (titles in white rectangle with red border) is thus clearly distinguished from all relevant documents on the map (red dots with red frames). An additional automatic filtering functionality constraints the display of document titles only to those documents in the zoom focus which are also part of the search result set (Fig. 8-13). In combination with the focus+context zooming technique such different levels of semantic visualisation of relevant documents with respect to user focus allows the user to keep both a semantic focus and overview of the entire set of search results.

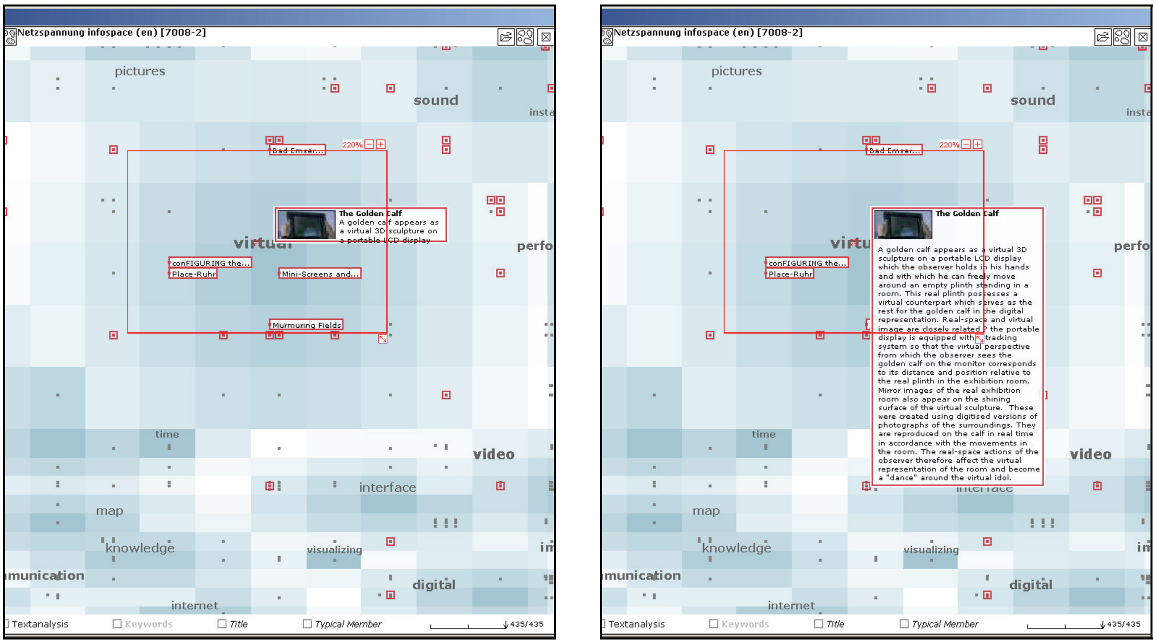


Fig. 8-14. Search result visualisation with semantic zoom and different levels of detail of document contents, in the Document Map.

The display of different levels of detail of document contents follows the same point and click principle as described in the previous section. Fig. 8-14 gives an overall visual impression of the relationships between different functionalities providing overview, zoom, filtering and document content level of detail applied to search result visualisation. An additional interaction possibility is provided by the coupling of the search result list in the Query View and the Document Map display. Selecting a document in the search result list causes the corresponding document to be visualised on the map alongside with the set of most similar documents (Fig. 8-15). The selected document is displayed in the described documents-in-focus visualisation mode (title in white rectangle with red border) while the similar documents are highlighted in the documents-out-of-focus search result visualisation mode (red dots with red frames).

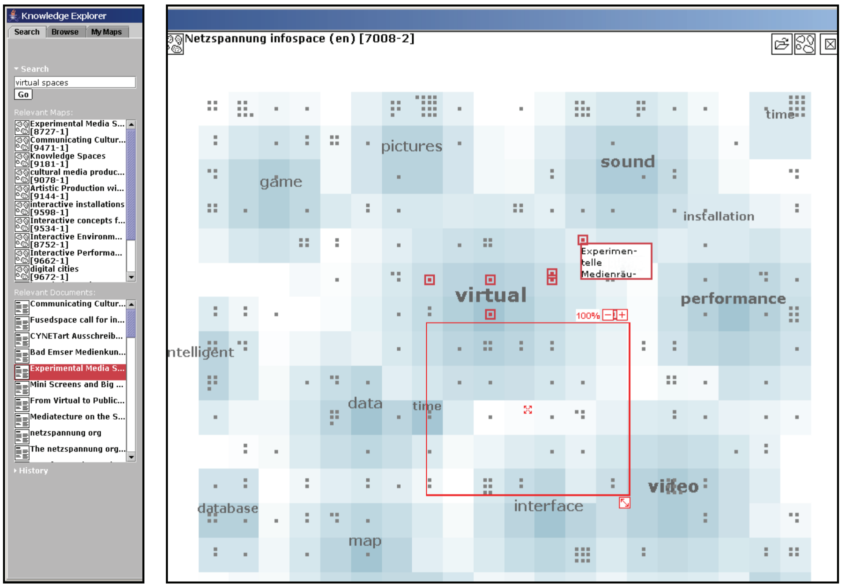


Fig. 8-15. Contextualistion of a selected document from search result with most similar documents on the Document Map.

8.6.3.3. Visualisation of Related Concepts in the Concept Map

Whenever a set of documents is selected in the Document Map, a set of related concepts is determined and visualised in the active Concept Maps. The matching procedure based on the coupling of the corresponding ontological structures of Document Maps and Concept Maps, (Chapter 6.6) has been described in Chapter 7.1.1. The visualisation of concepts in the Concept Maps follows the same visualisation modalities introduced for visualising search results in the Document Maps. First, a union of all concepts contained in the key relations of all documents in the search result is computed. The matching concepts on the Concept Maps are then marked by highlighting the corresponding concept icon with red colour. The resulting red rectangle icon resembles the framed red dots representing selected documents on the Document Map. This visualisation thus represents the set of all possible concepts related to at least one of the document in the search results (Fig. 8-16, top).

In the same way, the concepts related to the set of documents contained in the zoom focus are determined and marked by a red border around the concept label (Fig. 8-16, top). This visualisation is active always when there are documents from the search result set within the current zoom focus area. Finally, the concepts related to one specific document only are marked by a red fill. This visualisation is activated only when a user points to (mouse over) or clicks on a specific document (Fig. 8-16, bottom). This adds another level of semantic information to the displayed document information (short preview with full document title for pointing, full abstract for clicking).

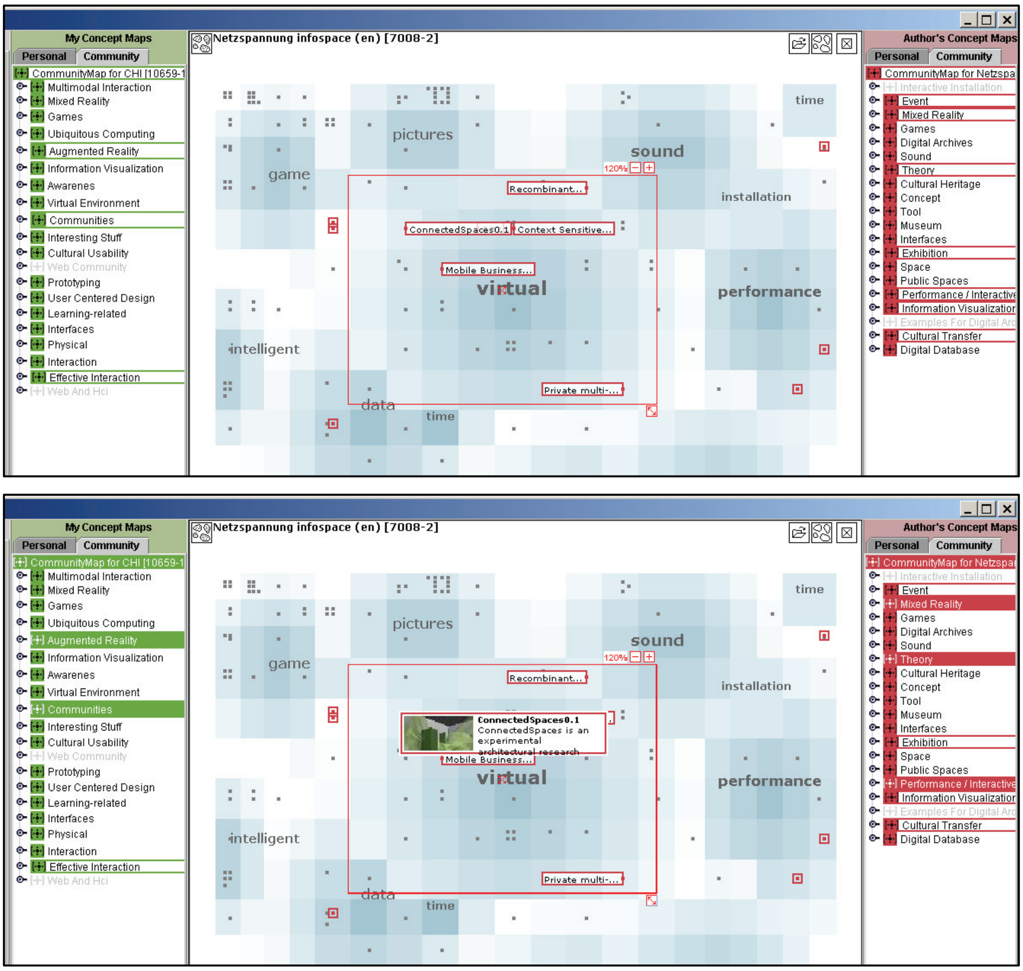


Fig. 8-16. Visualisation of related concepts in the Concept Map: concepts from the entire search result set are marked with a red icon (top), concepts related to the documents in the zoom focus only are marked with a red border (top and bottom) while concepts matching a specific selected document only are highlighted with a red fill (bottom).

Such visualisation of related concepts for a given document set determined by a search query allows the user to identify relevant topics or concepts for his information need from two different perspectives. On one hand, the concepts visualised in the right-hand Concept Map represent the knowledge structure of an unfamiliar community or one if its members (depending on the selected tab register and type of the Document Map opened in the central Similarity View).

Thus the user can discover previously unfamiliar concepts that may be relevant for his information need and use them for organizing documents into topical groups or for reformulating the search query in order to express his information need in a more appropriate way. On the other hand, the concepts visualised in the left-hand Concept Maps are related to the personal or community knowledge structure of the user and thus represent a familiar point of reference. This allows the user to discover how knowledge from the unfamiliar community is related to his own domain.

8.6.3.4. Concept Map Navigation

The Concept Map visualisation model is closely coupled with interactive navigation by the user in order to avoid visual clutter that is easily caused by a potentially large number of concepts (e.g. a total of 200 concepts for 20 main concepts and 10 sub-concepts per cluster). Each concept cluster is represented by a folder labelled by the corresponding main concept, containing several sub-concepts.

Initially all main concepts are greyed out (Fig. 8-17, left) and are painted in black only as the user scrolls over them with the cursor. When a cursor moves over a concept both the selected concept as well as the related concepts in the map are displayed (Fig. 8-17, middle). In this way, the user can quickly inspect all of the top-level concepts, identify the relevant ones for his current need and expand only those concepts he is interested in at the given moment (Fig. 8-17, right).

Fig. 8-17. Basic visualisation and navigation of the Concept Map in the Knowledge Explorer II.

Selecting a concept in one Concept Map causes a set of related documents to be visualised in the Document Map (Fig. 8-18) in the same way as the visualisation of search results described in Section 8.6.3.2. The set of related documents is determined by matching the key relations of concepts in the Concept Map with documents in the Document Map and vice versa (see Chapter 7.1.1, 7.1.5).

The selection mechanisms also allows several concepts to be selected simultaneously, in which case the intersection of documents related to all selected concepts will be computed and visualised in the Document Map.

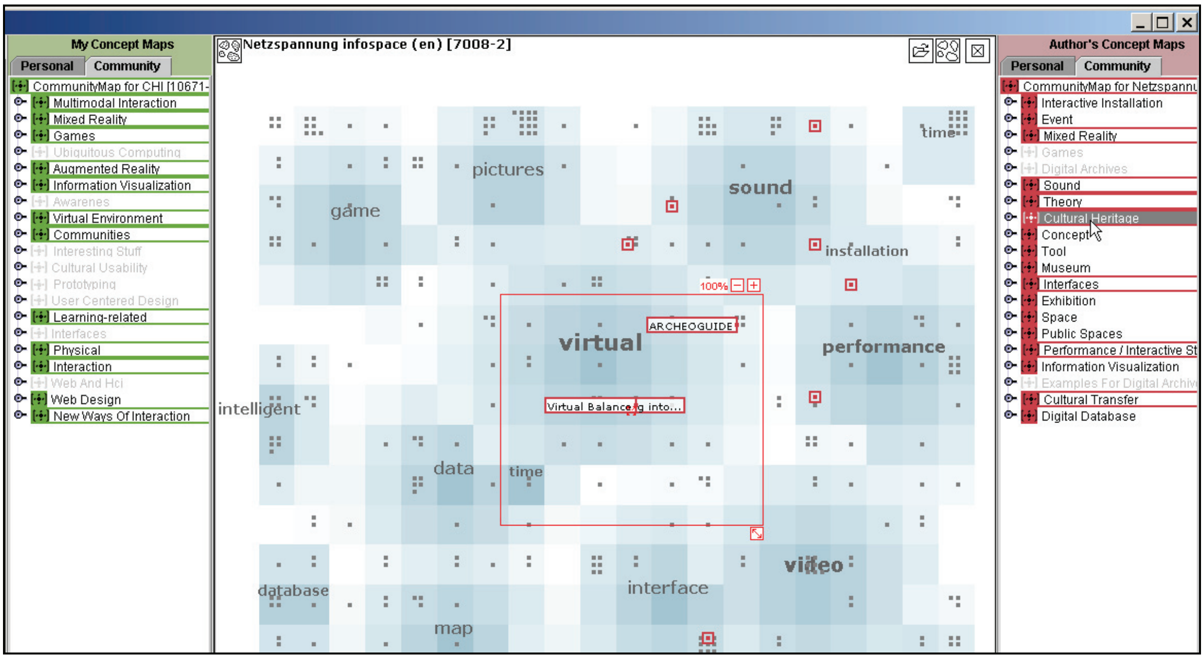


Fig. 8-18. Concept-based navigation in the Concept Map of an unfamiliar community perspective.

Following the computation of the document set corresponding to the selected concept, a set of concepts related to this document set will also be computed and visualised both in the Concept Map where the selection originated as well as in the Concept Map representing the other perspective (view coordination, Section 8.2). Such concept-based navigation allows the user to discover documents related to a specific topic, followed by a set of other potentially relevant topics.

Since the two displays are split in different views, the user can decide himself at which point he cares to consider or ignore which kind of information (documents or concepts in the familiar or unfamiliar perspective). Fig. 8-18 depicts the result of a concept-based navigation from the viewpoint of the unfamiliar community perspective with the set of related documents and related concepts both in the unfamiliar community (right-hand side) and the user's own community structure (left-hand side)

The described mechanism functions for both maps independently. The difference between the two views is realized by colour-coding: concepts selected in the right-hand Concept Map (unfamiliar perspective) and the resulting document set are visualised in red colour, while concepts selected in the left-hand Concept Map (users own personal or community perspective) and the resulting document set are visualised in green colour (Fig. 8-19, top).

The documents related to both selected concepts from the two different perspectives can now be identified in two ways. First, documents that are directly selected by both concepts will be marked in both colours, indicating a direct match by both concepts (Fig. 8-19, "concept intersection").

And second, documents marked in different colours but located close to each other are also likely to be related to both concepts: since their spatial closeness indicates semantic similarity a concept matching one document is likely to be related to the other document as well, and vice versa. The more different the two community perspectives, the higher the likeliness of absence of any direct inter-concept matches and the higher the importance of the later visual inference enabled by such multi-perspective visualisation.

Such simultaneous concept-based navigation from two different perspectives realizes the multi-perspective navigation functionality introduced by the application framework in Chapter 7.1.5. It allows users to both discover documents in an unfamiliar community space related to a specific topic from their own community and an topic from an unfamiliar community that seems to be relevant for their information need. In doing so, they discover cross-community relationships between different concepts as well as documents that exemplify their meaning.

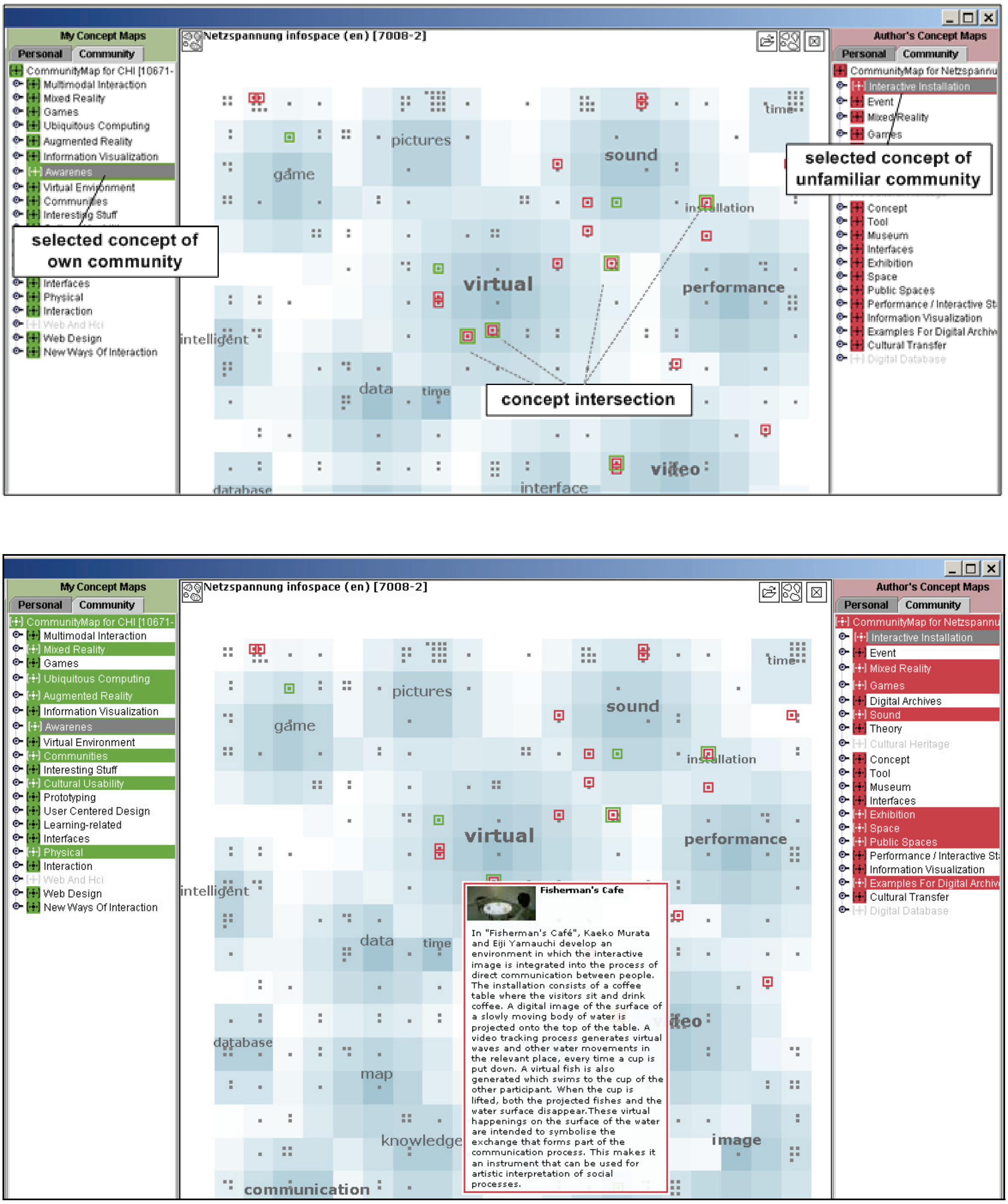


Fig. 8-19. Multi-perspective navigation by simultaneous concept selection in two different Concept Maps representing unfamiliar and own community perspectives: example shows the discovery of relationships between the concept "interactive installation" in media art community netzspannung.org and the concept "awareness" in the community HCI.

In the example depicted in Fig. 8-19, the multi-perspective navigation principle allows the user to discover that the topic "interactive installations" from the media art community netzspannung.org is related to the topic "awareness" in the HCI community. One of the documents connecting the two topics (Fisherman's Café) deals with an artistic approach to awareness of social communication and presence in shared physical space of public places.

8.6.3.5. Creation of Personal Maps

The creation of personal Document Maps in this version of the Knowledge Explorer is based on the common bookmark folders model. The user can organize the identified documents and express the knowledge gained in the information seeking process by organizing documents into named clusters represented by folders in a one-level tree hierarchy. This provides a well known representation model that is visually undemanding and can be used for quick, unobtrusive access without distracting the users from his primary activity of information seeking. As discussed in Chapter 6.4.1, the limitation to a one-level folder hierarchy is sufficient for the prototypical investigation since the focus is on explorative access[64]. Effectively, the user works with a two-level hierarchy since the root map folder is also named and the individual maps represent themselves specific information seeking topics for which the user is gathering documents.

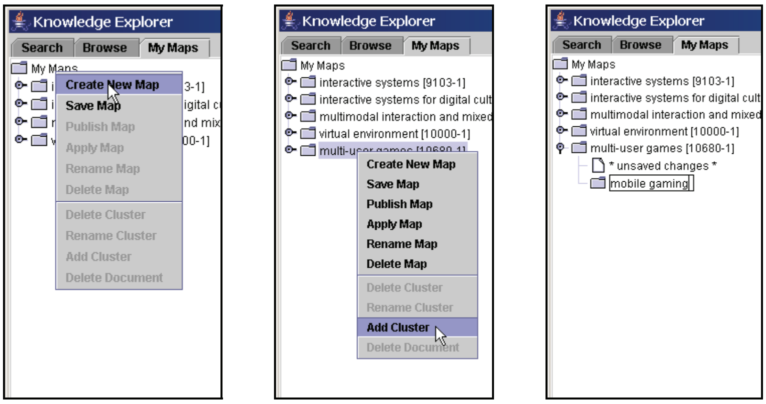


Fig. 8-20. Creating personal maps and named document clusters in Knowledge Explorer II.

The provided interface functionalities for creating personal maps, adding documents, creating and (re) naming clusters are illustrated in Fig. 8-20 and Fig. 8-21. Once a map is created by invoking the corresponding function in the context menu of the MyMaps tab new items can be added to the personal map by simply dragging & dropping individual documents or entire clusters from the Document Map opened in the Similarity View. Dropping a document outside of any existing clusters causes a new cluster to be created automatically which can then be named at will. The switching between the Search tab and the MyMaps tab also occurs automatically: moving a document over the search panel automatically opens the MyMaps tab. This avoids superfluous user actions and supports the simplicity of interface use.

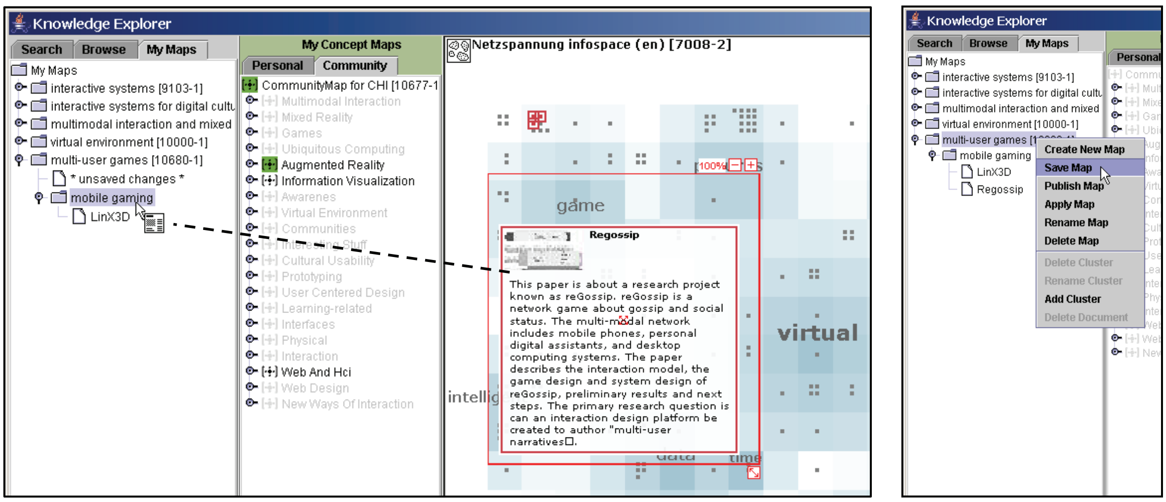


Fig. 8-21. Adding documents into personal maps in Knowledge Explorer II.

[64] A multi-level hierarchy could be incorporated without compromising the realized model e.g. by a multi-level SOM (Kohonen et al., 2000). The visual complexity of displaying multi-level clusters in the 2D Document Map could be solved by displaying multiple hierarchical levels in the folder-tree visualisation while retaining only the first level cluster in the 2D Document Map visualisation. Information on sub-hierarchies could be provided interactively on user-demand (e.g. by sub-cluster zoom).

8.6.3.6. Personalised Classification

The visualisation, interaction and navigation modalities described in previous sections apply both to community maps and personal maps without exception. The special functionalities made available by personal maps only include personalised classification and contextualized recommendations. These functionalities are realized by the corresponding services of the application framework and system architecture presented in Chapters 7.1.2 and 7.2.1.2. The personalised classification service capability of personal maps to dynamically classify a given set of documents into user-defined clusters allows them to be applied as semantic filters for categorizing unfamiliar community information spaces from the user's point of view. Applying a personal map to a community information space produces a map that contextualises documents from the community space into user-defined clusters in a given personal map. This functionality has been incorporated into the Knowledge Explorer II interface in the following way.

The user can select either an own personal map (MyMaps tab) or a published personal map of another user (Browse) tab to be applied as a template for classifying an information space. After invoking the ApplyMap function from the appropriate context menu (Fig. 8-22, left) the user is asked to choose the information space that is to be categorized by the selected personal map. As a result of user's selection the personalised classification service is invoked which returns a new map containing the clusters defined in the original map into which the best fitting documents from the selected information space have been assigned. Since the assignment of new documents to clusters is based on their similarity to documents already contained in the user's map (Chapter 6.4.2) this effectively results in a categorization of an unfamiliar information space based on user's criteria of document membership to a given topic.

To obtain a two-dimensional visualisation of the map the client then calls the spatial visualiser service which positions the map items on a two dimensional surface. The result is then presented to the user with the familiar Document Map visualisation (Fig. 8-22, right). The only difference is the additional visual differentiation between documents contained in the original map (assigned to named clusters by the map author) and the documents automatically assigned to the map by the personalised classification service. The documents of the original map author are displayed in bold, black dots whereas the automatically classified documents are displayed in grey, as usual. In this way the user can differentiate the results stemming from a human user from those classified by the system.

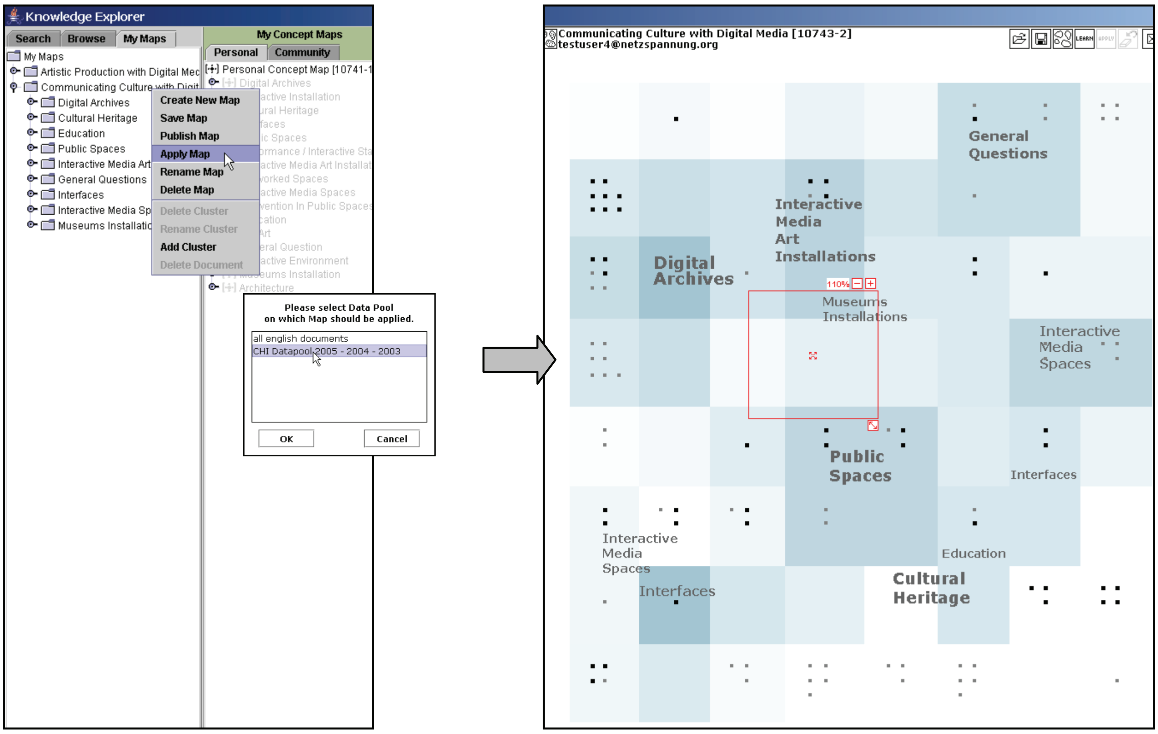


Fig. 8-22. Applying a personal map to classify information in an unfamiliar community information space.

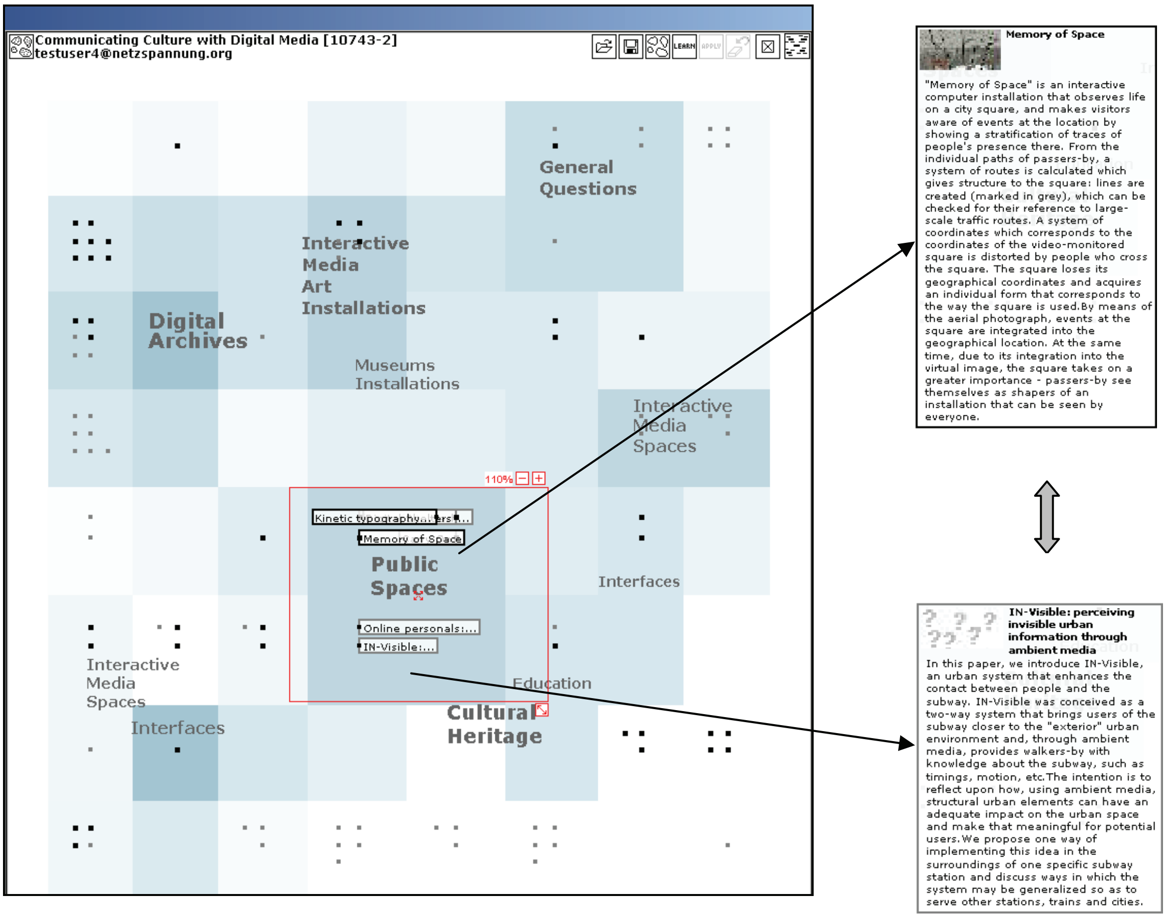


Fig. 8-23. Discovering related documents from unfamiliar community by personalised classification.

In the example depicted in Fig. 8-22, a user from the media art community netzspannung.org (a real-world testbed for our work, see Chapters 9 and 10) applies his personal map to classify the information space of the unfamiliar HCI community SIGCHI[65]. The resulting map displays information from the unfamiliar community space assigned to the topics defined by the clusters in the user's personal map. The user can now identify: which documents from an unfamiliar community are potentially related to the topics and documents from his own community that he is interested in. For example, inspecting the documents classified in the cluster "Public Spaces" in the vicinity of a document assigned to the map by himself, he discovers a document from the HCI community that is very much related to the same subject, although described in a different way (Fig. 8-23).

8.6.3.7. Contextualized Recommendations

Based on the results of a given search query a set of relevant maps from the collection of available personal maps of all users is determined by the matchmaking and recommendation service of the underlying system and application framework (Chapter 7.1.5, 7.2.2.2). The list of relevant maps is then displayed alongside the list of relevant documents in the Search Tab of the Query View of the Knowledge Explorer Interface. The personal Document Map with the highest relevance determined by the recommender service is automatically opened in the Similarity View followed by the personal Concept Map of the map author in the right-hand side Ontological View. Alternatively, the user can also open any other personal map form the list by dragging and dropping it from the list into the Similarity View. An example interface snapshot of this configuration is depicted in Fig. 8-24.

[65] See Chapter 10.7.7.1 on the construction of the HCI sample community space based on selected proceedings of ACM SIGCHI conferences.

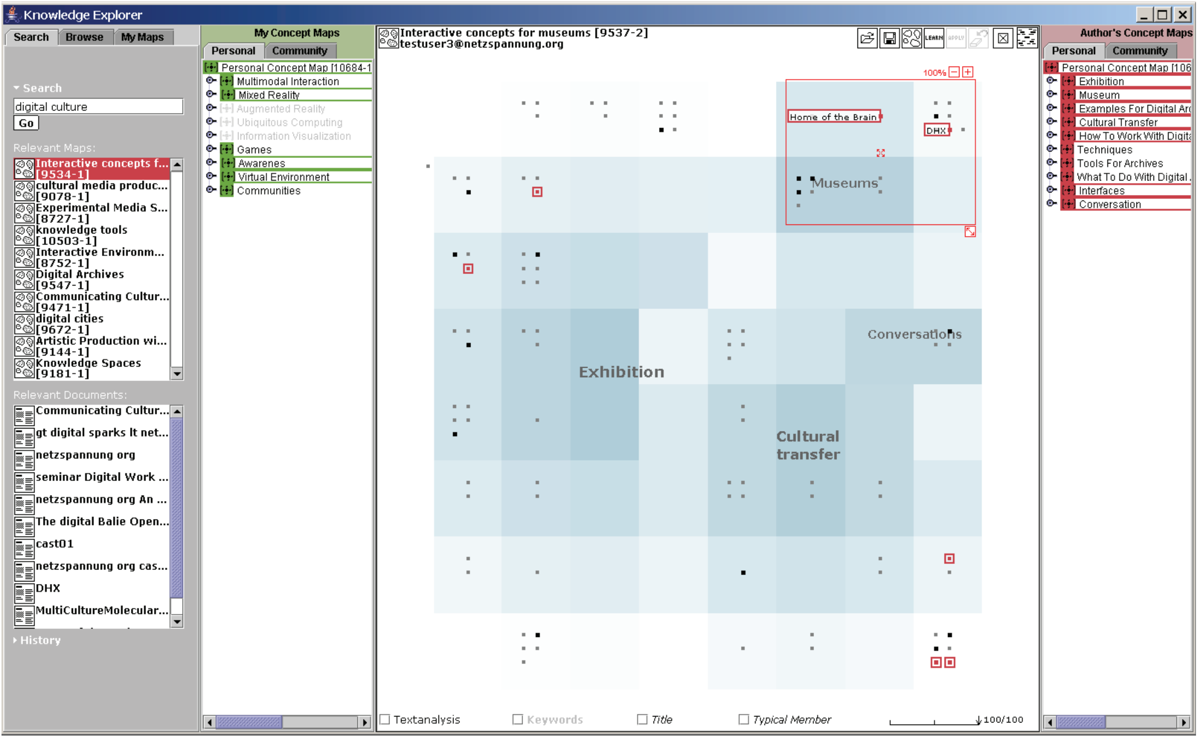


Fig. 8-24. Displaying personal maps as contextualized recommendations for a given search query.

The displayed personal Document Map contains the clusters and documents assigned tp them manually by the map author as well as a set of related documents which have been automatically classified from the entire community space. When opening a personal map in this mode, the map is by default automatically applied to the community space determined by the current search configuration context. Accordingly, a predefined number of best-fitting documents is classified into appropriate clusters defined in the given map. This is achieved by invoking the personalised classification service (Chapter 7.1.3, 7.2.1.2) and the the personal map spatializer service (Chapter 7.2.1.3) responsible for generating the corresponding two-dimensional visualisation of the Document Map model for personal maps (Chapter 6.4.3).

The visualisation and interaction with the maps is based on the same modalities described in the previous sections. The documents which have been organized in the map manually by the map author are visualised as bold, black dots, whereas documents classified automatically are displayed in grey colour. After the map is opened, the search results are also automatically visualised by highlighting the resulting document set in the personal Document Map (Section 8.6.2.2) and by visualising related concepts in the two Concept Maps (Section 8.6.2.3).

Fig. 8-24 shows an example result for a search query on “digital culture” returning a list of personal document maps with a map titled “interactive concepts for museums” on the first place in the list. The map shows clusters of documents dealing with “museums”, “exhibition” and “cultural transfer” containing documents from the search results (marked red). In addition, close to the search results other documents assigned by the map author as belonging to a given topic (e.g. “museum”) can be easily recognized as bold, black dots. This allows the user to identify documents that are relevant for a given topic although they have not matched the terms of his search query. In this way the personal maps are made available as a kind of recommendation of relevant documents and their semantic contexts (topics) suggested by other users, delivered by the system in a contextualized way that matches the specific need expressed by a user’s search query.

9. Empirical Evaluation in a Comparative Laboratory Study

This chapter presents the results of the empirical evaluation of the proposed approach to using knowledge visualisation for supporting cross-community knowledge exchange based on the prototypical implementation in a concrete interactive system: the Knowledge Explorer. The developed method for eliciting and visualising implicit knowledge structures of individuals and communities and its application to supporting typical tasks of cross-community knowledge exchange mediated through information access have been evaluated through a comparative laboratory study.

9.1. Evaluation Objectives

The results of the formative evaluation have confirmed the overall user acceptance and usefulness of the basic concept of our approach and of its first prototypical implementation. The goal of the summative evaluation is a more specific validation of the adequacy of the developed method and system for supporting typical tasks and requirements of cross-community knowledge exchange defined by the theoretical framework introduced in Chapter 3. By assessing the extent to which these requirements are supported by the developed solution we can asses to which extent our solution is adequate for the stated goal of supporting cross-community knowledge exchange.

The aim of this study is to qualitatively investigate how well the proposed approach supports the identification and understanding of relevant knowledge from unfamiliar communities mediated through access to unfamiliar community information spaces. In particular, the main objectives are:

- to evaluate whether the proposed method for computation and visualisation of implicit knowledge structures of individuals and communities and its application to supporting access to knowledge of unfamiliar communities improves the quality and effectiveness with which tasks relevant for cross-community knowledge exchange can be solved,
- to investigate user acceptance of the Knowledge Explorer tool implementing the developed method into an interactive system for multi-perspective access to community information spaces,
- to collect empirical data about the ways in which people use such a knowledge visualisation system for performing complex information seeking tasks in unfamiliar knowledge domains.

The main hypothesis of the proposed approach is that supporting access to unfamiliar community information spaces through dynamic knowledge maps that visualise personal and shared knowledge structures of members from different communities offers valuable support for cross-community knowledge exchange. This overall hypothesis shall be validated by assessing the following specific aspects:

- the adequacy of the developed knowledge visualisation method for eliciting and representing knowledge structures of individuals and communities (personal and community maps),
- the adequacy of applying this method to support cross-community knowledge access by means of multi-perspective access to community information spaces through personal and community maps.

To this end, the following specific hypotheses are examined:

(H1) The proposed method for eliciting implicit knowledge structures of human users based on their interaction with information is well-suited for eliciting personal knowledge structures of individuals and shared structures of communities.

(H2) The proposed knowledge map visualisation model integrating document and concept maps is well-suited for supporting semantic exploration and navigation in community information spaces.

(H3) Community knowledge maps better support the identification and understanding of relevant knowledge in unfamiliar community information spaces than standard information seeking tools without knowledge visualisation.

(H4) Personal knowledge maps better support the identification and understanding of relevant knowledge in unfamiliar community information spaces than standard information seeking tools without knowledge visualisation.

(H5) Personal knowledge maps better support the identification and understanding of relevant knowledge in unfamiliar community information spaces than shared community maps.

In order to be empirically tested, such hypotheses need to be made operational. This includes defining an experimental setup with measures for evaluating the hypotheses and practicable tasks that can be performed by test persons. In order to formally verify hypotheses H1 and H2 in an absolute manner, the notion of "well-suited" should be defined both in terms of concrete measures of the quality of elicited knowledge structures (Section 9.4.1) as well as in terms of quality thresholds, above which we can affirm such a judgment. However, as the discussion in Section 9.4.1 will show, already establishing appropriate measures for quality of knowledge structures is still an open and difficult research challenge - let alone determining precise quantitative thresholds. For this reason, the hypotheses H1 and H2 cannot be formally operationalized and tested in a way that would allow a definitive formal statement of their rebuttal or acceptance. Rather, they serve as a guide for qualitatively analysing the extent of suitability of developed methods based on a qualitative analysis of results of objective measures and subjective user feedback. In other words, the actual challenge lies in finding a space of possible interpretations of the notion of "suitability" of the developed methods in the given application context and building up an appropriate and practically applicable measurement and interpretation framework (Section 9.4.1)

Hypotheses H3-H5 can be verified by a comparison between a group using the system support proposed by this thesis and a control group using a standard reference system. To this end, criteria for measuring the quality of task solutions created by the users and the perceived adequacy of provided system support must be defined (Section 9.4.2-9.4.3). As basis for achieving this, the next two sections introduce a high-level evaluation framework and the overall structure of the evaluation study.

9.2. Evaluation Framework

The first pillar of our evaluation framework are ill-defined *information seeking tasks* that cannot be solved within the users' own community domain (Fig. 9-1, left) thus motivating cross-community access.

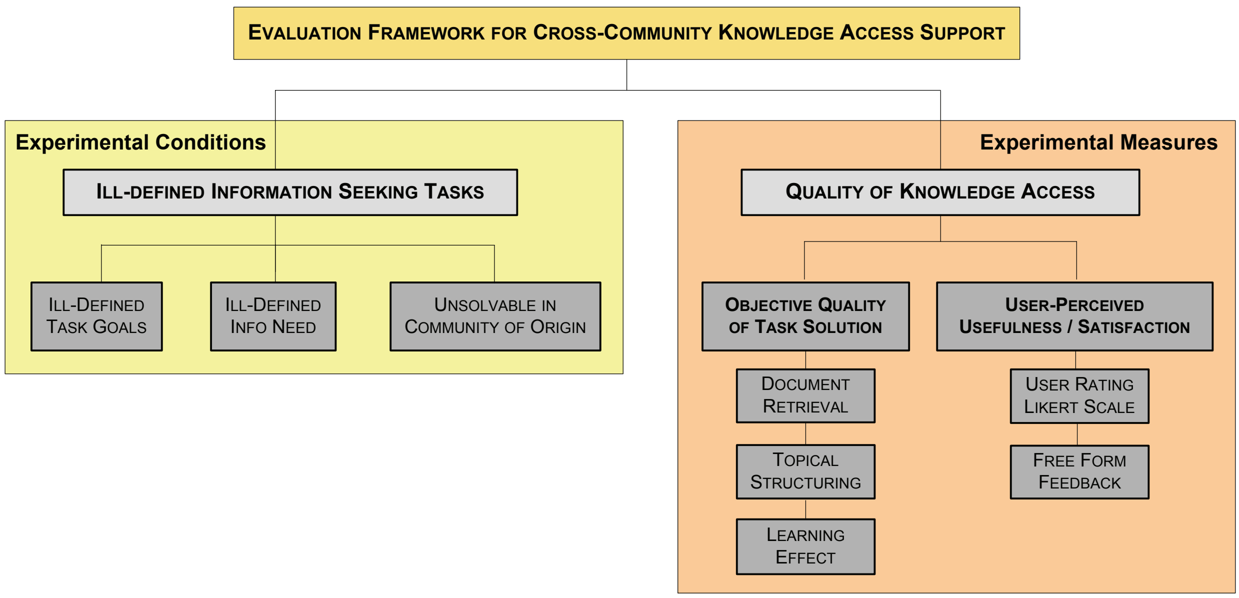


Fig. 9-1 Framework for evaluating cross-community knowledge access support

The basic experiment design includes a group of users from one community accomplishing tasks that require them to access the information space of an unfamiliar community in order to identify and understand knowledge relevant for the task. The second pillar of the evaluation framework is the definition of criteria for measuring the quality of support for identifying and understanding knowledge from unfamiliar community spaces (Fig. 9-1, right). We shall measure the quality of such *knowledge access support* in two ways:

- *objectively*: by the *quality of task solutions* created by the users,
- *subjectively*: by the *user feedback* on *perceived usefulness and satisfaction*.

The theoretical framework of requirements for supporting cross-community knowledge exchange introduced in Chapter 3 provides the basis for defining the criteria for measuring the objective quality of task solutions and the user-perceived usefulness and satisfaction. It identifies typical processes and tasks occurring in cross-community knowledge exchange mediated through information seeking in unfamiliar community information spaces. The corresponding high-level evaluation criteria based on this analysis are depicted in Table 9-1.

Evaluation Criteria 1: Perspective Taking and Perspective Making
• How well does the proposed method support users in gaining insight into implicit knowledge structures of communities different from their own? • To what extent does the proposed method enable users to discover relationships between knowledge from unfamiliar communities and a) knowledge from their own community ? b) needs of a specific task ? • To what extent are users supported in expressing the discovered insights and new knowledge in their own terms?
Evaluation Criteria 2: Sensemaking Tasks in Unfamiliar Community Domains
• How well does the proposed method support users in: • Finding appropriate concepts to formulate an information need • Finding and creating representation schemas organizing information into meaningful structures for a given task • Using representation schemas of others to find and classify information for a task

Table 9-1. High-level evaluation criteria for cross-community knowledge exchange

9.3. Overall Structure of the Evaluation Study

In order to test the hypotheses formulated in Section 9.1 we need to design an appropriate experimental setup consisting of operative measures and practical tasks that can be accomplished by users and that shall provide us with indicators for measuring the high-level criteria listed above.

Hypotheses H1 and H2 require us to create an experimental setup in which the adequacy of the proposed method for the elicitation and visualisation of personal and shared knowledge structures of a community of users can be tested. This is important since the proposed method for supporting cross-community knowledge exchange is based on eliciting and visualising personal and shared structures of community knowledge. Hypotheses H3-H5 require us to create an experimental setup in which the adequacy of the use of personal and shared community knowledge maps for accessing knowledge of unfamiliar communities can be evaluated by comparing it to a standard information seeking system that doesn't use knowledge visualisation.

Accordingly, we adopt a two-stage experiment design as follows. Two different community information spaces representing two different communities (Community C1, Community C2) shall be selected as the basis for a cross-community knowledge access scenario. In the experiment, members of one community shall use the system to generate personal and community knowledge maps that will then be used by members of another community for accessing knowledge from the first community's information space (Fig. 9-2).

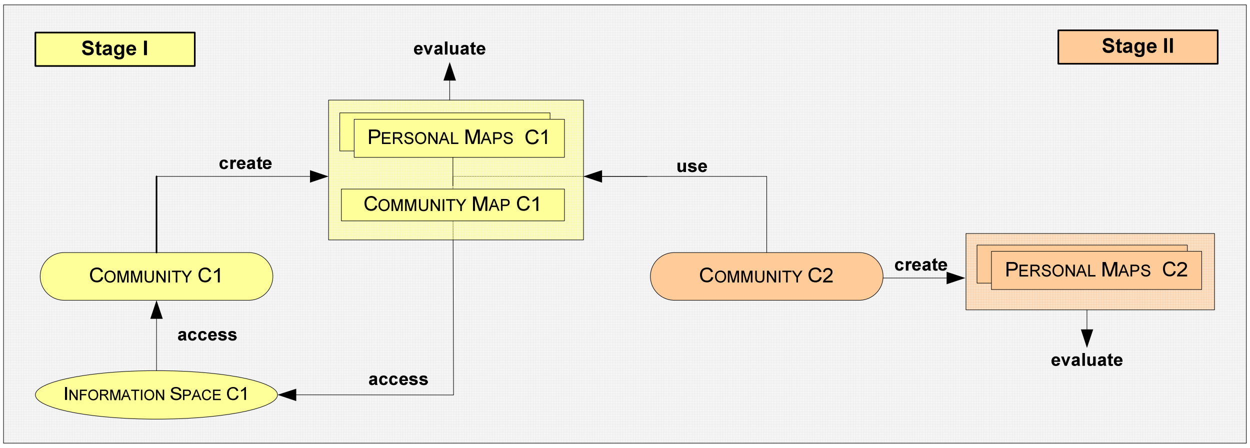


Fig. 9-2. Overall structure of the experimental design.

The evaluation study is thus divided into two phases i.e. two separate experiments corresponding to the two main groups of hypotheses H1-H2 and H3-H5:

Experiment I: Eliciting personal and shared knowledge structures of a community from members' information seeking activities in own community information space.

In this experiment a group of test persons from one community (Community C1) accomplishes a set of information seeking tasks in their own community information space. To this end they use the Knowledge Explorer with system-generated community maps (Fig. 9-3). Based on the personal document maps created as a result of accomplishing the assigned tasks, implicit knowledge structures of individual users and of the entire user group are calculated and visualised by our system.

By introducing an appropriate quality measure and applying it to the resulting personal and shared community knowledge maps created by the system, we can test the hypotheses H1 and H2. The use of the Knowledge Explorer with system-generated maps in this phase further allows us to gain user feedback on the adequacy of the use of system-generated maps as a bootstrapping solution when there are no personal maps and no shared community structure available yet.

Note: The personal and shared community maps of Community C1 created in this experiment are used in Experiment II by members of Community C2 to accomplish tasks requiring them to identify relevant knowledge from the information space of Community C1.

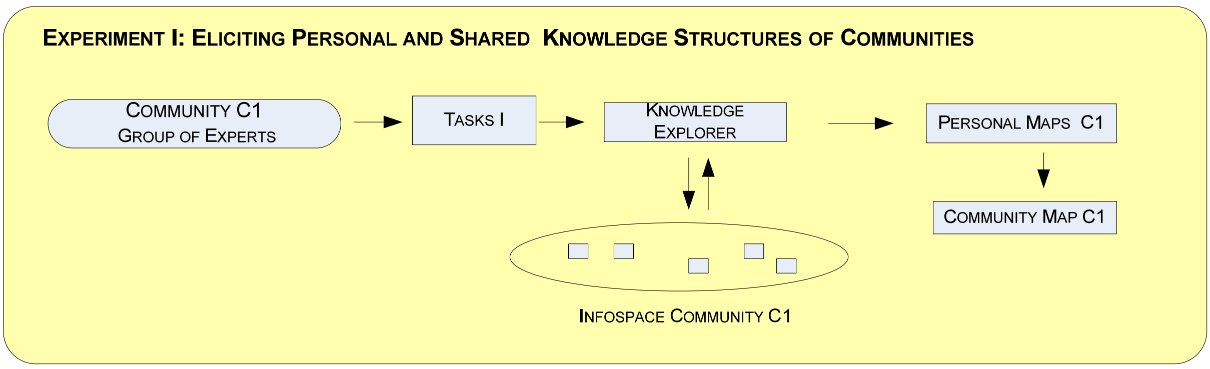


Fig. 9-3. Structure of Experiment I

Experiment II: Using personal and shared knowledge structures of one community by members of another community for accessing knowledge from an unfamiliar community information space.

This experiment is divided in two phases: the initalization phase (Experiment IIa) and the test phase (Experiment IIb).

Experiment IIa: First, personal and shared community knowledge structures of Community C2 are elicited by having test persons from Community C2 accomplish a set of information seeking tasks in their own community information space.

The maps generated in this experiment serve as initialization of own personal and shared structures to be used in connection with structures of the unfamiliar Community C1 in Experiment IIb.

Experiment IIb: Following the initialization phase in own community space, three different groups of test persons from Community C2 accomplish a set of information seeking tasks in the unfamiliar information space of Community C1 (Fig. 9-4). The first group uses the shared community knowledge map created in the previous phase by users from the Community C1. The second group uses personal knowledge maps created in the previous phase by users from Community C1. The third group uses a standard information seeking system without knowledge visualisation to accomplish the same tasks.

By comparing the results of the first and second group to results of the third group we can test the hypotheses H3 and H4. The comparison of the results of the first and second group between each other allows us to test the hypothesis H5.

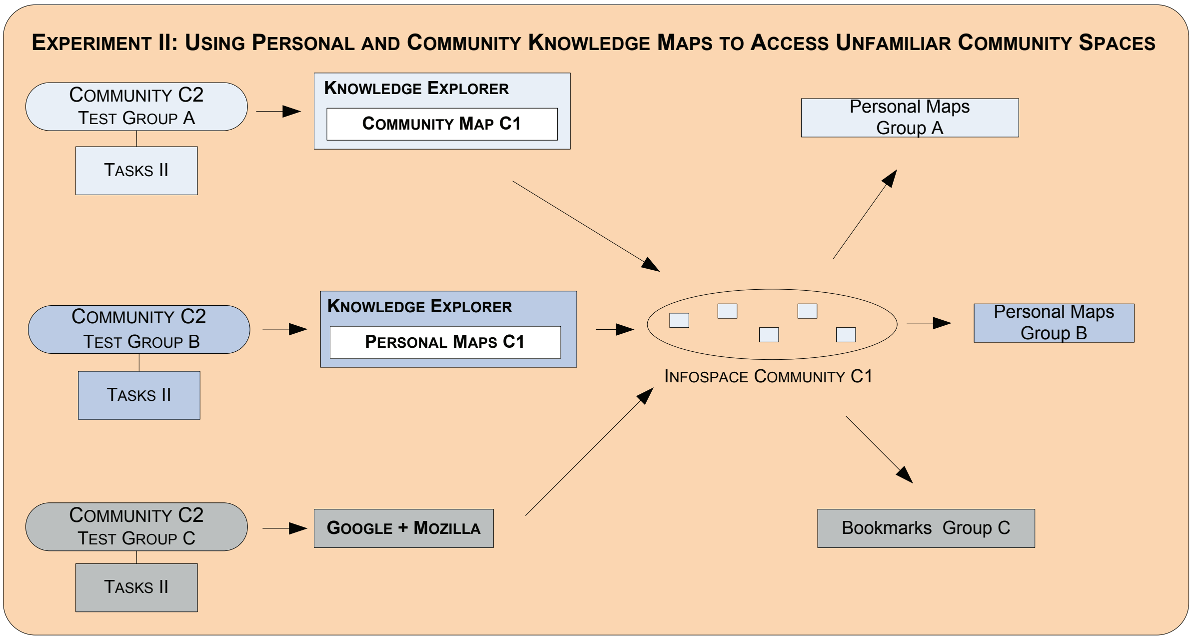


Fig. 9-4. Structure of Experiment II

9.4. Definition of Experimental Measures

As outlined in the previous section, the first experiment aims at evaluating the developed method for eliciting knowledge structures of individuals and communities of users itself. The measures for assessing the quality of knowledge structures elicited by the method presented in Chapter 6 are introduced in Section 9.4.1. In order to evaluate the application of this knowledge visualisation method to supporting cross-community knowledge in the second experiment, the high-level criteria introduced in Table 9-1 are transferred into concrete measures and indicators that can be obtained from information seeking tasks to be accomplished by the users (Sections 9.4.2-9-4.3). A summative discussion of how the defined operational measures implement the high-level criteria from Table 9-1 is given in Section 9.4.4.

9.4.1. Measures for the Quality of Elicited Knowledge Structures

Different approaches to evaluating knowledge structures in the area of ontologies and concept maps exist. Ontology evaluation approaches distinguish between evaluation of structural properties, functional evaluation and application or usability evaluation (Gangemi et al.,2005; Brank et al., 2005). Concept map evaluation methods generally consider content and application validation (Albert & Steiner, 2005).

Structural evaluation metrics of ontology-related methods are often based on graph theory and measure indicators such as breadth, depth, tangledness and fan-otuness (Gangemi et al., 2005). Apart from being rather abstract and thus not providing much insight into content-based quality, in our case they are not applicable for two additional reasons: we do not implement a classical ontological structure (no complex logical relations, no inferencing, see Chapter 6.6) and our method is based on setting the majority of these properties to a predefined value in advance.

Functional evaluation and application evaluation are preferred methods of ontology evaluation in literature. Functional evaluation is often considered with respect to information retrieval use and measured in terms of precision, recall and coverage (Hartmann et al., 2004). The first two assess the classification correctness of the ontological structure whereas coverage assesses the indexing completeness of a document space.

A precondition for application of such measures is the existence of a reference solution based on which the correctness of a classification can be determined. This is also not given and highly impracticable in our case: different personal structures by default represent personal points of view and are equally valid. The whole premise is that there exist different relevant views on the assignment of documents to concepts and characterization of document groups based not only on a personal bias but also on the need of a specific task. This applies to the shared community structure as well. Its very creation is based on the assumption that multiple relationships revealed by different users are inter-connected into one structure.

The methods of content validation of concept maps are also based on comparison with a reference solution (Albert & Steiner, 2005) and are hence not applicable for the same reasons. The supporters of such methods themselves acknowledge the limitation of this procedure and point to the difficulties of establishing methods for assessing “objective” content validity of concept maps (ibid).

The application evaluation both in the area of ontologies and of concept maps is concerned with assessing the suitability of the ontological structure for a specific purpose for which it was built in a given application domain. This involves aspects such as quality of support for achievement of stated goals (e.g. problem solving, acquisition of knowledge in a given domain), perceived usefulness by the users as well as usability issues (Hartman et al., 2004; Gangemi et al., 2005). Application evaluation is also the major focus of our evaluation and is undertaken in a dedicated experiment (Experiment II) for which the Experiment I serves as a preparatory population phase.

Finally, different authors point out that the evaluation of knowledge structures is not only a still widely open research area but that there is inherently no single best approach: a choice of an appropriate approach highly depends on the purpose of the application and the purpose that the evaluation aims to achieve (Brank et al., 2005).

Accordingly, in order to gain some insight into the intrinsic quality of the elicited knowledge structures produced by our method, in this section we establish a set of structural properties that are relevant for our application context. To this end we adapt and extend structural indicators used for the evaluation of probabilistic ontology extraction in (Kunz, 2005). This includes the quality of structural composition (single-world vs. multi-world concepts), concept distribution and quality of role assignments (sub-concept clustering to main concept), document coverage (nr. of indexed documents) and the quality of concepts in terms of intelligibility and descriptive power (Table 9-2).

The structural composition distinguishes between single-word and multi-word concepts, with the latter being assumed to be more expressive and hence of higher quality and usefulness. To describe the descriptive power of a given concept map we distinguish between the following concept classes:

- topical concepts – are concepts that characterize a thematic context in the community domain (e.g. “mixed reality”, “cultural heritage”, “interactive installations”),
- structural concepts – are concepts that do not refer to a single thematic context but describe a structural aspect regardless of a specific theme (e.g. “events”, “projects”, “tools”),
- narrative concepts – are concepts that are not formulated in a way that corresponds to a specific theme or subject area from the community domain but use colloquial language to describe a topic (e.g. “How to work with digital archives”),
- idiosyncratic concepts – are concepts used in a very specific, personal way whose context of meaning is not readily understandable without familiarity with the personal intention of the author (e.g. “fragmentation”, “passages”, “paris”),
- unusable concepts – are concepts whose meaning is either so generic or so unintelligible that they convey no usable meaning as a characterization of a document set.

The higher the use of topical concepts and the lesser the amount of idiosyncratic concepts the higher the descriptive power of a given concept map. Structural concepts contribute to the descriptive power but should not be the dominating ones. The use of narrative concepts in general implies little familiarity with the specific concepts relevant in a given community domain. Thus it indirectly implies a less experienced user and thus a lower quality personal concept map.

On the other hand, narrative concepts tend to exhibit high immediate intelligibility and usefulness for novice users. Idiosyncratic concepts are also expected to be less useful than topical concepts as they are more difficult to be understood by another user. Hence, with respect to descriptive power, higher quality concept maps will have a higher proportion of topical concepts than narrative and structural concepts and contain little idiosyncratic and unusable concepts.

MEASURES FOR THE QUALITY OF ELICITED KNOWLEDGE STRUCTURES	
Structural Composition	Percentage of single-word vs. multi-word concepts
Concept Distribution	Number of concepts • unique concepts • multi-occurrence concepts
Document Coverage	Amount of documents indexed by the concepts • total number of documents • average number of documents per cluster • percentage of coverage in document pool
Role Assignment	Number of main concepts and sub-concepts Quality of sub-concept → main concept relations
Descriptive Power	Percentage of: • topical concepts • structural concepts • narrative concepts • idiosyncratic concepts • unusable concepts

Table 9-2. Measures for the quality of elicited knowledge structures

9.4.2. Measures for the Quality of Knowledge Access

According to the high-level criteria introduced in Table 9-1, to assess the quality of knowledge access mediated by information seeking in an unfamiliar community information space we shall measure three main effects:

- How well could users gain insight into knowledge structures of unfamiliar communities?
- How well could users discover relationships between the knowledge of an unfamiliar community and one's familiar knowledge domain?
- How well could users accomplish typical sensemaking tasks occurring during information seeking in unfamiliar community domains?

Such high-level criteria provide us with the basis for constructing appropriate tasks for our experiment design. They define *what* effects should be measured. But they still do not tell us *how* we can do this in concrete operative terms. In considering how to achieve this, it is important to note that the question *"how well the users were supported in…"* can be answered at two different levels:

- *outcome-based:* in terms of quality of the end result of users efforts in accomplishing the task (quality of task solution)
- *process-based:* in terms of quality of the process through which users have arrived at their solutions (quality of solution process support)

The outcome-based assessment allows us to measure the quality of support based on those effects that can be readily derived from the end-result of users' efforts (e.g. the proportion of relevant documents retrieved with respect to a reference solution). The process-based assessment aims at gaining an understanding of the different factors that have influenced the process through which the users have arrived at a given solution. This allows us to capture and assess the quality of system support based on factors that may not be readily measurable from the end-solution alone.

On one hand, this enables us to better assess the actual quality of different aspects of a given system support. On the other it may enable us to interpret the causes of the positive or negative quality of end-solutions as well as possible discrepancies between the objective solution quality measures and user-reported usefulness and satisfaction.

The process-based solution quality will have an important role in our case since some of the important effects we need to measure can be induced from end-solutions only to a limited extent. This is due to the inherent nature of the problem we are addressing: knowledge construction in information access. The basic premise of our approach is that information access is a process in which users use actively construct new knowledge (the sensemaking paradigm, Chapter 3.4). Thus what we are interested in is not merely the amount of relevant information identified by the users but the level of understanding of knowledge from an unfamiliar community that users develop in the process.

Obviously, there are significant limits to the extent in which this can be measured in a simulated task situation with a significant time limit. More naturally such processes occur in a longer period of time as the users actions are also spread across different time occasions. Measuring knowledge construction effects is very difficult, on two main accounts. On one hand, the knowledge internalized by users will only partly be directly expressed in their information seeking results. On the other hand, the very task model on which our approach and the developed method are based, emphasises the role of complex cognitive tasks - such as concept formation and information classification and structuring - accomplished by the users *during* the information seeking process.

It is difficult to measure the quality of users' end-solution in a way that will give us an indication of such process-based effects we are interested in. At the same time, while better process support should eventually result in a better end-solution of a task, this may not always be the case: e.g. when factors not foreseen in advance have significantly influenced the eventual outcome or when the knowledge construction effects have not been externalized in a measurable way. Hence, finding a way that allows us to link task solutions to knowledge construction effects that have occurred during the process is both very important and a difficult challenge.

In order to deal with that challenge we adopt the following strategy:

- we recognize that objective indicators linking the users' end-solutions to knowledge construction effects which we need to measure can be established only to a limited extent,
- we place an important role on subjective user feedback to perceived usefulness and satisfaction as a way of capturing process-based outcomes that are otherwise difficult to capture.

Against this background we can refine the high-level criteria for which we will need to define operative measures in order to assess the quality of knowledge access, in the following way:

QUALITY OF KNOWLEDGE ACCESS	
QUALITY OF TASK SOLUTION	• the extent to which users have identified relevant knowledge from an unfamiliar community space for a given task, • the extent to which users have identified relevant relationships between knowledge from an unfamiliar community and their familiar community domain, • the extent to which users have expressed the discovered insights through newly created information structures, • the extent to which users have created suitable representation schemas for structuring ill-defined tasks
SUBJECTIVE USER FEEDBACK	• the extent to which users have developed an understanding of concepts describing unfamiliar community knowledge relevant for a given task, • the extent to which users were able to relate knowledge from the unknown community to the familiar community domain, • the extent to which users found the support provided by the system adequate for solving the assigned tasks, • the extent to which users found the support provided by the system adequate for typical sensemaking tasks occurring during the information seeking process.

Table 9-3. High-level measures for measuring quality of knowledge access and user-perceived usefulness of a given support system

In order to operationalize the above criteria into concrete operative measures we need to relate the high-level effects from Table 9-3 to main informational activities occurring during information seeking in unfamiliar community domains.

Based on the analysis presented in Chapter 3.6 (Table 3-10) we can distinguish the following classes of informational activities related to knowledge access in unfamiliar communities:

- identifying appropriate concepts for expressing an information need,
- identifying documents relevant for a given information need,
- creating conceptual structures that organize information in a meaningful way relevant for a given task and information need,
    - o understanding semantic contexts of documents and organizing related documents into meaningful groups,
    - o identifying and understanding unfamiliar concepts relevant for a given task,
- assigning documents to conceptual structures relevant for a given task

- assigning related documents to newly acquired concepts,
- assigning documents from unfamiliar domains to familiar concepts.

According to the adopted sensemaking model of information seeking the result of an information seeking task is a set of documents organized into semantically related groups, characterized by appropriate concepts (representation schemas, Chapter 3.4). This corresponds to bookmark-like structures in standard information seeking support systems and to the personal maps created by the users in our model.

To measure the quality of knowledge access based on assessing the quality of task solutions we define the following three indicators:

- Document retrieval effect,
- Quality of topical structuring,
- Learning effect.

For each of these indicators we define measures that can be applied to users' task solutions. They are depicted in Table 9-4 and introduced in the following sections.

Measure	Definition
Document precision - p_d	$p_d = \frac{\lvert D \cap R \rvert}{\lvert D \rvert}$
Document recall - r_d	$r_d = \frac{\lvert D \cap R \rvert}{\lvert R \rvert}$
Topical precision - p_t	$p_t = \frac{\lvert C \cap Q \rvert}{\lvert C \rvert}$
Topical recall - r_t	$r_t = \frac{\lvert C \cap Q \rvert}{\lvert Q \rvert}$
Cross-community topical precision p_{tx}	$p_{tx} = \frac{\lvert C \cap Q_x \rvert}{\lvert C \rvert}$
Intra-community topical precision p_{tx}	$p_{th} = \frac{\lvert C \cap Q_h \rvert}{\lvert C \rvert}$,
Shared community topical precision p_{tx}	$p_{ts} = \frac{\lvert C \cap Q_s \rvert}{\lvert C \rvert}$
Overall topical coverage – t_c	$t_c = \sum_{i \in U} \lvert C(i) \cap Q \rvert$
Cross-community topical coverage - t_x	$t_x = \sum_{i \in U} \lvert C(i) \cap Q_x \rvert$

Table 9-4. Summary of operational measures for quality of knowledge access based on task solutions.

9.4.2.1. Document Retrieval Effect

The document retrieval effect is described by well-known measures of precision and recall (Baeza-Yates & Ribeiro-Neto, 1999). Precision considers the ratio between the number of relevant documents retrieved by the user and the total number of documents retrieved by the user. Recall describes the ratio between the number of relevant documents retrieved by the user and the total number of relevant documents in the collection (defined by a reference solution).

Document precision and recall: Let D be the set of documents found by a test person for a given task and R the set of relevant documents defined by the reference solution for that task.

User document precision p_d and user document recall r_d are defined as:

$$p_d = \frac{|D \cap R|}{|D|}, \qquad r_d = \frac{|D \cap R|}{|R|}.$$

9.4.2.2. Quality of Topical Structuring

The quality of topical structuring is very difficult to define in our situation. In order to characterize groups of semantically related documents users can proceed in several different ways:

- they can relate documents to familiar concepts from their own community,
- they can relate documents to concepts from the unfamiliar community,
- they can use non-topical characterization concepts.

The question posed here is how to asses which concepts were “good characterizations” for a given task and which not - i.e. how to construct a reference solution against which the quality of topical structuring can be measured. One way of constructing a reference solution is to consider how members of the source community to whom the information space belongs, would solve this task. In this way, only the concepts used by the members of this community would be considered as relevant choices for naming document groups. But this is the correct solution only from the point of view of that community. Members of the test user community could well argue that using the concepts form their own community to characterize the document groups is an appropriate solution as they need to identify documents and topics that are relevant to *them*.

Finally, the users might use a mixture of concepts from own and unfamiliar community and argue that this satisfies best the needs of the given task from their own community perspective. Defining an “optimal” ratio for the use of concepts from own and unfamiliar community is also an impracticable suggestion. Moreover, for an ill-defined need as in our task design, many different but equally valid structuring schemes will exist, depending on a specific focus given to the topic.

In other words, the extent to which the use of concepts from the unfamiliar community or the use of concepts from the own community should indicate a better solution will highly depend on the nature of the task. Accordingly, the reference solution for topical structuring of the task must include the points of view of both communities and the relevant importance of one view with respect to the other.

The quality of topical structuring can then be assessed by considering the extent to which the topical structure of the users’ task solutions contains:

- relevant concepts from the unfamiliar community,
- relevant concepts from the users own community of origin,
- relevant concepts shared by both communities,
- non-topical concepts.

To assess these aspects we apply the measures of precision and recall to concepts that users have chosen as semantic descriptors of document groups in their solution. Since in our tasks the users are required to group documents into relevant subtopics structuring the task information need, we refer to these measures as *topical precision* and *topical recall*. Topical precision signifies then the ration between the number of relevant concepts and the total number of concepts used by a user in a given task solution. According to the above distinction of concept origin, we distinguish between:

- overall topical precision: precision with respect to all concepts in the combined reference solution including both community points of view,
- cross-community topical precision: precision with respect to the reference solution from users community of origin,
- intra-community topical precision: precision with respect to the reference solution from the unfamiliar target community,
- shared topical precision: precision with respect to concepts shared in both communities.

The overall topical precision tells us the proportion of the relevant concepts used by the user regardless of their origin. The topical precision in own community tells us to which extent the user had used familiar concepts to structure the task information (the portion of relevant familiar concepts in the solution). The topical precision with respect to the unfamiliar community tells us to which extent the user had adopted previously unfamiliar concepts (the portion of unfamiliar concepts). In our experimental design a more extensive use of concepts from the unfamiliar community will signify a higher level of understanding of the relevant knowledge from the unfamiliar community space.

In contrast, the related measure of topical recall cannot be readily applied in our case since the reference solution is not *the* unique correct topical structure but merely a set of relevant topics out of which different topical combinations are possible and equally valid. Thus the notion of recall cannot be straightforwardly used as an indicator of task solution quality as the users are not expected to retrieve *all* relevant concepts in order to create a valid structure.

Nonetheless the measure of topical recall does provide us with a useful quality indicator as it tells us how much of the overall topical spectrum available for the task (i.e. its reference solution) was captured by the user. A higher topical recall signifies a greater use of concepts from the reference solution and hence a better structured solution. We must only remember that in our case this is not an absolute measure (i.e. 100% recall doesn't represent the ideal solution) but only a *relative* indicator by which different task solutions can be compared between each other.

Topical precision and topical recall: Let C be the set of topics created by a test person in the solution of a given task and Q the set of all relevant topics defined by the reference solution for that task, with following subsets:

- Q_h - topics from the own community reference solution,
- Q_x - topics from the unfamiliar community reference solution,
- Q_s - topics shared by both community reference solutions.

Overall topical precision p_t and recall r_t are defined as: $p_t = \frac{|C \cap Q|}{|C|}$, $r_t = \frac{|C \cap Q|}{|Q|}$.

Cross-community topical precision p_{tx} is defined as: $p_{tx} = \frac{|C \cap Q_x|}{|C|}$, Intra-community topical precision p_{tx} is defined as: $p_{th} = \frac{|C \cap Q_h|}{|C|}$,

Shared community topical precision p_{tx} is defined as: $p_{ts} = \frac{|C \cap Q_s|}{|C|}$

Given a small number of users (with heterogeneous personal backgrounds and interests) forming the community sample and the limited time available for completing the tasks, it can be expected that the choice of concepts for topical structuring of the solutions will differ greatly between individual users (unless trivial solutions are produced). Accordingly, an important indicator of quality of task solutions is the *cumulative topical coverage.* This describes the total number of topics used in the solutions of all users in a given test group, differentiated by the community origin of concepts used in naming the topics.

Our particular focus of interest here is the cumulative coverage of concepts from the unfamiliar community in the group solutions. This allows us to gain insight into the degree of diversification of cross-community topics identified by the users. For ill-defined information needs where personal user preferences play an important role in determining relevant information, a greater cross-community topical diversification implied by a greater group coverage would imply a better group performance, as more of the relevant unfamiliar concepts were identified. Hence, we focus on cross-community topical coverage.

Cross-community topical coverage: Let C(*i*) be the set of concepts used by the test person *i* in the solution of a given task, U the set of test persons for a given test group and Q the set of all relevant topics defined by the reference solution for that task, with following subsets:

- Q_h - topics from the own community reference solution,
- Q_x - topics from the unfamiliar community reference solution,
- Q_s - topics shared by both community reference solutions.

Overall topical coverage t_c is defined as: $t_c = \sum_{i \in U} |C(i) \cap Q|$.

Cross-community topical coverage t_x is defined as: $t_x = \sum_{i \in U} |C(i) \cap Q_x|$.

9.4.2.3. Learning Effect

The learning effect refers to the extent to which users were able to internalize the insights discovered in accomplishing the tasks into new knowledge. Topical structuring discussed in the previous section reflects the extent to which users have externalized these insights and applied them for structuring information relevant for the task. However, several factors may have influenced the results of this process. For example, the effectiveness in using an unfamiliar system support may limit the effectiveness with which users are able to express the relevant knowledge from the unfamiliar community they have acquired. On the other hand, since the knowledge maps directly visualise conceptual structures of a given community and related documents, it is possible that users could identify relevant concepts and documents without significantly learning about their meaning.

Accordingly, the learning effect indicator complements the measure of the extent to which users were able to identify and understand relevant knowledge from the unfamiliar community. It is measured by a qualitative questionnaire asking users to name concepts and topics describing specific aspects of the information seeking task they had just accomplished. This includes naming relevant concepts from the unfamiliar community and related concepts from own community. The user answers are then assessed by the measures of *topical precision, topical recall* and *cumulative topical coverage* introduced in the previous section. The structure of the questionnaires measuring the learning effect is given in Table 9-5.

QUESTION	MEASURED ABILITY
Q1: Name concepts and topics related to task topic "…" in community "…" *(unfamiliar community)*	Understanding an unfamiliar community concept
Q2: Name concepts exemplifying the meaning of the concept "…" in community "…" *(unfamiliar community)*	Relating a familiar concept to an unfamiliar community domain
Q3: Name concepts related to topics "…." *(own community topic)* and "…." *(unfamiliar community topic)*	Integrating knowledge across communities
Q3. Name related topics in community "…" *(own community)*	Relating an unfamiliar concept to own community domain.

Table 9-5. Structure of questionnaires for measuring the learning effect.

9.4.3. Measures for User-Perceived Usefulness and Satisfaction

Subjective user feedback on perceived usefulness and satisfaction is gathered by means of a qualitative questionnaire submitted to the users after all tasks have been completed. To this end, the questionnaire combines questions requiring users to provide a rating on 5-point Likert-scale and open questions with free form answers. The first part of this questionnaire is concerned with the task-adequacy of system support. Its structure is depicted in Table 9-6. The second part of the questionnaire is concerned with the overall usability of the system and general user satisfaction (Table 9-7).

QUESTION SUBJECT	TYPE OF ANSWER
the degree to which the given tasks are realistic tasks in practice	Likert-scale (very unrealistic – very realistic)
the overall adequacy of the system for accomplishing the tasks	Likert-scale (very poor – very good)
the functions perceived as useful by the users	free form
the extent of use of individual system functions in accomplishing the tasks,	Likert-scale per function (very little - very much)
the perceived usefulness of individual functions for accomplishing the tasks	Likert-scale per function (of very little use – very useful)

Table 9-6. Structure of the questionnaire on system task-adequacy for Knowledge Explorer test groups.

QUESTION SUBJECT	TYPE OF ANSWER
the ease of use of the system	Likert-scale (very easy – very difficult)
the ease/difficulty of learning to use the system	Likert-scale (very easy – very difficult)
the extent to which solving the tasks with the system was strenuous	Likert-scale + free form (very easy – very strenuous)
the existence of some particularly appealing aspects of the system	free form
the existence of some particularly practical aspects of the system	free form
the existence of some displeasing aspects of the system	free form
suggestions for improvement of the system	free form

Table 9-7. Structure of the questionnaire on usability and user satisfaction for Knowledge Explorer groups.

This questionnaire aims to elicit feedback on functionality aspects related specifically to the proposed Knowledge Explorer system. As such it isn't suitable for the test group using a standard reference system, since a number of questions don't apply. In order to assess the critical aspects of the standard reference system in comparison to the proposed solution, we need to identify main problems that users experience in using the reference system and than compare them to the support provided by our solution. To measure this, a separate questionnaire is constructed for the test group using the standard reference system. Its structure is depicted in Table 9-8.

QUESTION SUBJECT	TYPE OF ANSWER
the degree to which the given tasks are realistic tasks in practice	Likert-scale (very unrealistic – very realistic)
the overall adequacy of the system for accomplishing the tasks	Likert-scale (very poor – very good)
the adequacy of the system for accomplishing the tasks	Likert-scale (very poor – very good)
difficulties experienced in accomplishing the tasks	free form,
what functions were useful for accomplishing the tasks	free form
what functions would have been desirable to have	free form

Table 9-8. Structure of the questionnaire on usability and user satisfaction for the reference test group.

The feedback on the comparison of the two systems after the exposure of the reference group to the Knowledge Explorer is gathered by an additional questionnaire, depicted in Table 9-9.

QUESTION SUBJECT	TYPE OF ANSWER
the comparison of the adequacy of the Knowledge Explorer and the standard reference system for accomplishing the tasks	Likert-scale (much worse – much better)
the functions of the Knowledge Explorer that have been useful for accomplishing the tasks	free form
the existence of some particularly practical aspects of the system	free form
the existence of some displeasing aspects of the system	free form
suggestions for improvement of the system	free form

Table 9-9. Structure of the questionnaire for direct system comparison by the reference test group.

9.4.4. Linking Operational Measures to High-Level Evaluation Criteria

The defined measures for the quality of task solutions operationalize the high-level measures from Table 9-3. The first indicator of the extent to which users were able to identify relevant knowledge from the unfamiliar community space is the extent to which they were able to identify relevant documents for a given task. This is measured by the standard *document precision and recall.*

This however is a very limited measure for our purpose, since selection of documents from a list of retrieved documents does not necessarily reflect the understanding of document contents, as it can be performed by simple term matching by the users: selecting documents containing the terms from the task definition as relevant ones. Moreover, document selection tells us little of the extent to which the users were able to understand the knowledge context in which the documents make sense within the unfamiliar community.

A more direct indicator of a user's understanding of unfamiliar knowledge is reflected in the concepts used to structure the information relevant for the task. A general indicator is to consider the overall number of subtopics created by the users. This provides an immediate insight of the extent to which users engaged in structuring the information need of a given task, regardless of the structuring criteria they have chosen. It accounts for foreseeable differences in sensemaking behaviour (Qu & Furnas, 2005) that may result in valid structures based on different criteria than the reference solutions that assume topical structuring by the target community concepts (assessed by the topical precision and recall measures).

The extent to which users have used concepts from the unfamiliar community to characterize the subtopics in their solutions is an important measure of the extent to which they succeeded in identifying and appropriating knowledge from the unfamiliar community. This is measured by the *topical precision and recall* and even more specifically by *cross-community topical precision.* The overall topical precision and recall reflect the overall extent of identified knowledge as they consider all topics regardless of the origin of concepts with which they were characterized (unfamiliar community / user's community).

The use of concepts from the unfamiliar community given by the *cross-community topical precision* is the best direct measure of the extent to which users have gained insight into the unfamiliar community knowledge structure. The associated *cross-community topical coverage* complements this by showing the amount of concepts transferred by a whole user group. The greater this amount, the greater the richness of the knowledge transfer that occurred at the group level. Finally, the overall extent to which users have explicitly expressed the discovered insights through newly created information structures is measured by *topical recall* that indicates the overall degree of structuring of users' task solutions.

The extent to which users were able to identify relationships between unfamiliar knowledge and concepts used in their familiar community domain can be measured by *document precision* and *recall* in specifically designed tasks focusing on the identification of relevant documents for a specific concept.

Some of the operative measures can also be used as indicators for the suitability of support for typical sensemaking tasks. The effectiveness with which users were able to find appropriate concepts for formulating their information need is implicitly reflected in extent to which they were able to identify relevant concepts from the unfamiliar community domain.

Since all tasks are undertaken within a strict time limit, *cross-community topical precision* is a good indicator of the effectiveness with which unfamiliar concepts were identified and used by the users in order to produce good task solutions. In the same way, the *topical recall* of solutions produced in a limited timeframe, tells us about the effectiveness with which users were able to create suitable representation schemas for a given task. The higher the topical recall and the higher cross-community topical precision, the higher the effectiveness of finding and creating appropriate representation schemas.

As discussed in Section 9.4.2, the quality of knowledge access is reflected only partially in the task solutions. Accordingly, the quality of knowledge access can be assessed by the described objective measures only to a limited extent. The introduction of the learning effect indicator is one way of addressing this. Eliciting direct user responses to task-related questions requiring them to actively engage and express the knowledge acquired in solving the task, is a way of measuring otherwise inaccessible extent of knowledge construction occurred during the information seeking process.

Assessing the extent to which the users are able to understand unfamiliar community concepts, to relate familiar concepts to an unfamiliar community domain, to relate unfamiliar concepts to own community and to integrate knowledge across communities directly reflects high-level criteria defined in our evaluation framework.

Another issue raized in Section 9.4.2 is the need of eliciting user feedback on the quality of support for complex cognitive tasks accomplished by the users *during* the information seeking process. One way of eliciting user feedback on process-based support is by asking them to rate the perceived usefulness of specific functionalities aimed at addressing specific sensemaking tasks (Chapter 3.4), such as:

- finding concepts for expressing an information need,
- finding and creating representation schemas for a ask,
- finding and using representation schemas of others for identifying and classifying information for a task.

This needs to be translated into appropriate system functionalities aimed at supporting such tasks. In our solution this includes in particular the following functionalities:

- visualisation of unfamiliar concepts in the concept maps related to a search query (Chapter 8.6.3.3),
- navigation in concept maps for finding relevant concepts and topics for the given task (Chapter 8.6.3.4),
- navigation in concept maps for identifying documents related to a given topic (Chapter 8.6.3.4).

For other functionalities that are more difficult to relate to specific task aspects or cover different aspects at the same time it would be potentially misleading to restrain them to a specific purpose. Rather we can elicit feedback on perceived overall usefulness and gain additional user explanation of their kinds of use through open-end questions and informal interviews. Such functionalities include:

- visualisation of the document clusters in the document map of the unfamiliar community space (Chapter 8.6.2, 8.6.3.1),
- visualisation of the search results in the document map (Chapter 8.6.3.2).

Accordingly, the two different types of functionalities are incorporated into the questions aimed at eliciting user feedback on the usefulness of specific system functionalities in the questionnaire described in the previous section.

9.5. Basic Elements of the Test Setting

The relevance of the experiment results is critically dependent on a realistic test scenario's that is as close as possible to situations found in practice. To ensure that, we proceed as follows.

9.5.1. Experimental Simulation Method

As a basic framework for creating an appropriate test setting and task designs we adopt the method of "simulated work task situation" (Borlund & Ingwersen, 1997) also known as the "experimental simulation method" (McGrath, 1995). In this method, test users are presented with scenarios similar to what they could meet in a real-world situation. The tasks to be accomplished are then defined in the context of such a concrete scenario. Such contextualisation has proven helpful for ensuring that participants invest as much effort in fulfilling the tasks as they would do in real life (Borlund, 2000).

In our approach we are examining effects of specific system support on knowledge access mediated by access to community information spaces that are represented as unstructured document collections. A common approach to evaluating information access is based on experiments with artificially constructed document collections that have no reference to users' professional backgrounds. In our case this is not a viable approach.

In order to create a realistic setting in which our hypotheses can be tested we need to select example community information spaces and test user groups that can serve as realistic samples of community members. The selected community information spaces need to represent communities that are sufficiently different between each other while also having some intersections in their domains of interest; they have to be readily accessible over the Internet and we must have practicable ways of recruiting test users that can provide a realistic sample of community members.

9.5.2. Creation of Test Collections for Simulated Community Spaces

Accordingly, for the purpose of our experiment the following two community information spaces have been chosen:

- the information space of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) - as a representative example of a HCI community information space concerned with the development and evaluation of interactive systems and techniques,
- the information space of the netzspannung.org community of media art, design and technology – as a representative example of a media art community space concerned with artistic and cultural uses of interactive media and technology.

The test collections have been compiled from selected portions of the two information spaces. For the HCI community specific archives of proceedings of CHI conferences have been chosen and merged into a test information space. For the media art and design community information space, the open information pool netzkollektor and portions of editorial contributions of netzspannung.org have been chosen.

9.5.3. Selection of Test Persons

The test group for Experiment I was composed out of 7 participants including media art curators, media artists and interface designers. The test persons were recruited from the alumni network of the Media Arts Research Studies Dept. of the Fraunhofer Institute for Media Communications. Six of them were registered members of netzspannung.org. All users were experts with several years of professional experience in the fields of media art, design and media technology. They actively participated in the netzspannung.org community by following the discourse on the corresponding Internet platform, exploring the contents of its information space and providing own contributions. In a qualitative sense, they were a very realistic sample of typical members of netzspannung.org.

The HCI test user groups have been composed out of researchers and graduate students in the fields related to the two communities. For the HCI test groups the participants were recruited among research assistants, Master and PhD students at the University of Duisburg-Essen, Dept. of Computer Science. Master students came from the Master's in Computer Science and from the Master's in Applied Communication and Media Sciences. Doctoral students and research assistants came from the Chair for Interactive Systems and from the Chair for Cooperative Systems.

The researcher assistants and PhD students in interactive and cooperative systems qualified naturally as realistically representative members of the HCI community at large. To ensure that student participants qualified as suitable test persons they were required to have completed or have been attending the courses Interactive Systems and Usability Engineering and to have expressed an interest in specific HCI topics that served as the basis for compiling test collections. In this way, 18 qualified participants, divided in 3 test groups of 6 participants each, could be recruited for the experiment.

9.6. Experiment I: Eliciting Knowledge Structures of Individuals and Communities

In this experiment a group of test persons from the community netzspannung.org accomplishes a set of information seeking tasks in the test collection compiled from the netzspannung.org information space. Based on the personal document maps created as a result of accomplishing the assigned tasks, knowledge structures of individual users (personal concept maps) and of the entire user group (community concept map) are generated by the system (Chapter 6.4, 6.5).

The quality of the generated knowledge structures is evaluated by examining structural properties of the created concept maps (e.g. amount of indexed documents, connectedness of concept clusters) and by a qualitative assessment of concept quality (Section 9.6.5). In addition, the task-adequacy for intra-community information seeking as well as overall usability of the specific system configuration are assessed by means of subjective user feedback gathered by qualitative questionnaires (9.6.6, 9.6.7).

9.6.1. Task Design: Intra-Community Information Seeking

The main requirement on the tasks for this experiment is that they need to provide realistic intra-community information seeking tasks that stimulate user engagement and allow their personal knowledge to be naturally reflected in the task solutions. The information structuring activity is one important aspect of user actions as it provides the main source from which personal and community knowledge structures are extracted. Our approach relies on information structuring occurring naturally during information access (Chapter 6.1).

Accordingly, the tasks must not explicitly require users to structure an area or domain but need to simulate situations which users would encounter in practice - they need to elicit structuring actions as part of natural information seeking behaviour of the users. Furthermore, in order to represent different points of view typically present in a community in real-world conditions, the individual tasks need to reflect personal interests of different community members (test persons).

The task topics must not be too narrowly defined: they should leave some leeway for personal interpretation by the users, so that the results can reflect a specific personal point of view. On the other hand, the task topics also need to exhibit some degree of intersection so that there is a realistic chance that relationships between concepts used by different users actually exist and a shared community structure can be constructed.

Finally, a reasonable amount of personal maps and concepts created by the users needs to be available so that sensible structures can be extracted. Due to the simulated nature of the experiment and time limits we will work with a relatively small sample of user maps. But this is actually an appropriate condition, since an important aspect of the proposed method is that it should function in an acceptable way also in the bootstrapping phase when little user input is available.

In order to meet these requirements we adopt the following approach. The test group is composed of 6 users and each user is presented with two information seeking tasks, with the task topics being especially tailored to his/her personal interests. One kind of tasks is based on broad topics so that the personal view of the user on the topic can be expressed and reflected in the concepts s/he decides to use for characterizing document groups. The other kind of tasks is based on a more specifically defined topic.

The task structure is:

- **Task 1**: well-defined task with a familiar but ill-defined (broad) information need, requiring users to identify relevant information for solving the task (documents and topics) in their own community information space.
- **Task 2:** well-defined task with a familiar and well-defined information need, requiring users to identify documents of personal interest in their own community information space and organize them into thematic clusters.

The topics are selected and assigned to individual test persons in such a way that:

a) they correspond to the professional background and personal interests of a given user,

b) there are non-obvious intersections between some of the topics.

Task 1: Exhibition „Communicating Culture with Digital Media“
You are preparing an exhibition on the topic of „Communicating Culture with Digital Media“. Find projects and topics relevant for your exhibition in the information space of the media art community netzspannung.org. Organize the results of your search in a way that will be helpful for preparing your exhibition.

Table 9-10. Example of task description for Experiment I.

In order to create a meaningful test situation for the users the tasks are accompanied with practically relevant scenarios, adjusted to the professional background of the user (curators, artists, designers). They include work contexts such as "preparing an exhibition", "writing an article" and "developing a project idea". An example of a task description is given in Table 9-10. A list of all task topics and corresponding simulated work contexts is presented in Table 9-11.

Topic	Context
knowledge spaces	preparing an exhibition
linking real and virtual spaces	writing an article
digital cities	preparing an exhibition
interactive environments	writing an article
interactive concepts for museums	writing an article
digital archives	writing an article
communicating culture with digital media	preparing an exhibition
artistic production with digital media	writing an article
interactive performance	developing a project idea
media spaces	preparing a workshop
knowledge interfaces	developing a project idea
interactive installations	writing an article
cultural production with digital media	writing an article
digital media in public space	developing a project idea

Table 9-11. Task topics and simulated work contexts for Experiment I.

9.6.2. System Configuration

As this experiment is situated in the bootstrapping phase, in which no personal maps of individual community members exist yet, the Knowledge Explorer system presented in Chapter 8 has been configured in such a way that only those functionalities that are applicable in this phase were active.

This configuration provides only a system-generated map Document Map and system-generated Concept Map based on the text-analysis of the documents in the test collection of a given community space. It includes search query and visualisation of search results as well as basic map visualisation and means of interaction: browsing, semantic zoom, visualisation of related concepts and concept map navigation, display of document abstracts on the map and document details in the browser window (Chapter 8.6.3)

The results of users' information seeking activities can be organized and saved in personal maps containing named clusters of related documents (Chapter 8.6.3.5). The back-end system provides only the search query engine and basic map management functionalities (retrieving system-generated maps, saving personal maps). The corresponding Knowledge Explorer interface is depicted in Fig. 9-5.

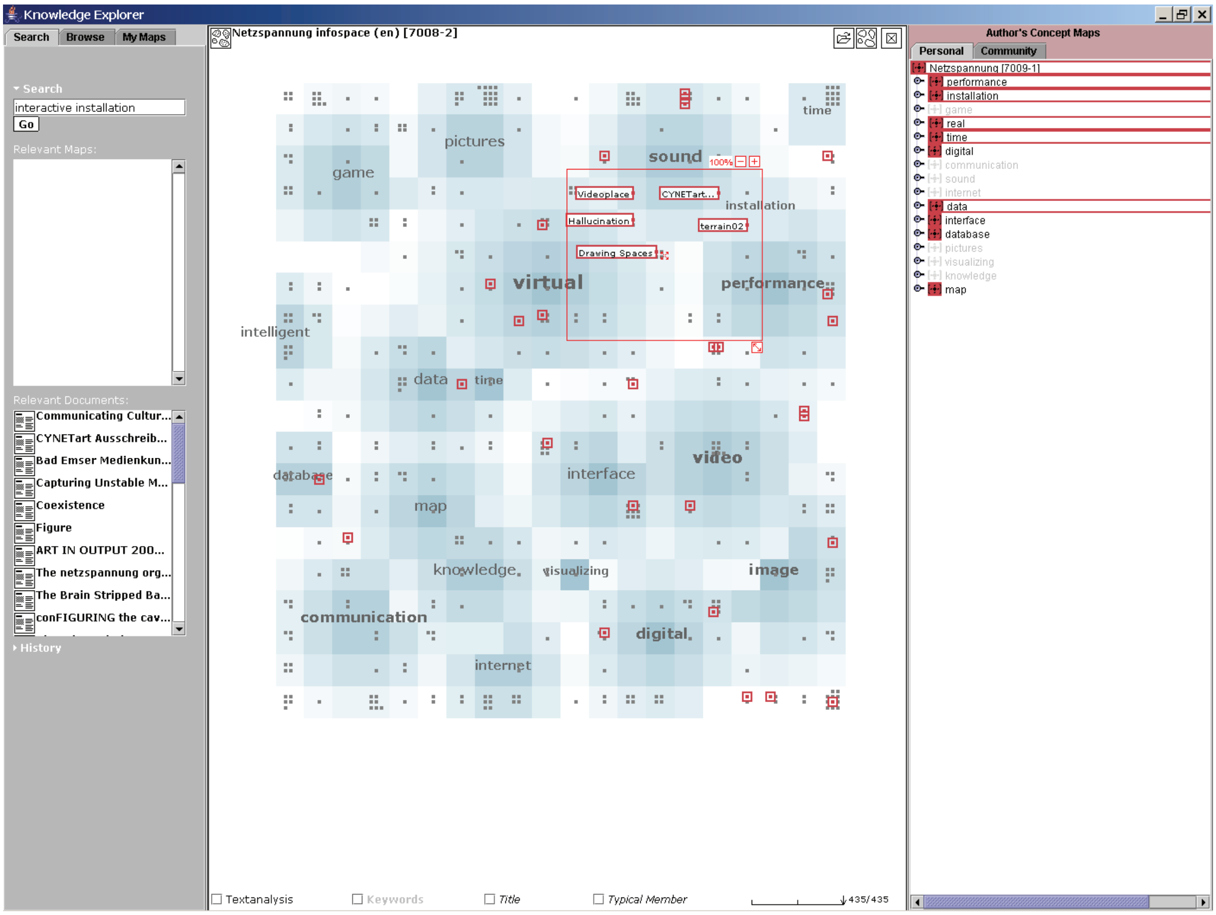


Fig. 9-5.Screenshot of the Knowledge Explorer interface used in Experiment I

9.6.3. Course of the Experiment

The experiment was held in November 2005 at the Fraunhofer Institute for Media Communication. A total of 7 participants took part in the experiment, divided in several sessions over two weeks (November 14 - November 29). In each session one or more users completed the experiment[66].

The sessions were accompanied by one supervisor who controlled the course of the experiment and answered questions regarding the general understanding of tasks. Two additional technical assistants observed the technical performance of the system and recorded informal feedback given by the participants during the course of solving the tasks. s

Total duration of each session was approx. 3h. This includes a short introduction and presentation of the system (ca. 25min), a try-out of the system by the test users (10min), performing the actual tasks (2h) and compiling the questionnaire (ca. 15min). The exact course of a session is given in Table 9-12. The following data was collected during the experiment:

- the task solutions (recorded by the system),
- the questionnaires compiled by the test-subjects,
- informal feedback of the test subjects gathered during the experiment,
- informal feedback from post-experiment group discussion.

[66] Due to participants' professional commitments and geographical distribution, the appointments had to be made on an individual basis.

COURSE OF AN EXPERIMENT session
• Introduction (5 min): – Explaining the purpose & course of the session – Brief outline of the tasks to be solved • Compiling the pre-testing questionnaire (5 min) • System demonstration (15 min) • Free exercise for trying out the system (10 min)
• Performing Task 1 (1h)
• Short break & refreshment with informal feedback (10 min)
• Performing Task 2 (1h) • Compiling the task-adequacy and usability questionnaire (15 min)
• Group discussion (15 min)
Time total: 3h 15min

Table 9-12. Course of an experimental session for Experiment I.

9.6.4. Characterization of the Test Subjects

At the start of each test session socio-demographic data on the test subjects were collected by means of a pre-testing questionnaire. The data collected include general information such as gender, professional background and topics of interests. This has been complemented with a more specific information such as technical skills in working with computers and the familiarity with background knowledge relevant for community domains treated by the experiment design.

The general characterization of the test subjects shows a strong dominance of female participants (71% female, 29% male participants). Regarding technical skills, 5 test subjects declared themselves as "advanced users" of personal computers and 2 as "professional users". Thus, all test users had a high-degree of technical skills. The most important data collected includes the capabilities of test subjects regarding background knowledge that may strongly influence the experiment results. To this end the users have been asked to rate their capabilities on following indicators:

- the degree of familiarity with the community domain of expertise to which users have been assigned (media art)
- the degree of familiarity with the information space of netzspannung.org.

Since the test groups simulate a sample of the community of media art and design (represented by netzspannung.org), a high level of familiarity with that field and a fair level of familiarity with the information space of netzspannung.org is required for a realistic setting. The collected data confirms the suitability of the selected test subjects with respect to these requirements. As can be seen in Fig. 9-6 all test subjects exhibit at least a moderate level of familiarity with the field of media art with the large majority being very well (71%) or well familiar with it (14%). The level of familiarity with netzspannung.org is similarly good though somewhat differently distributed: 70% of test subjects are very well (43%) or well (29%) familiar with it while the remainder of the users knows it moderately well.

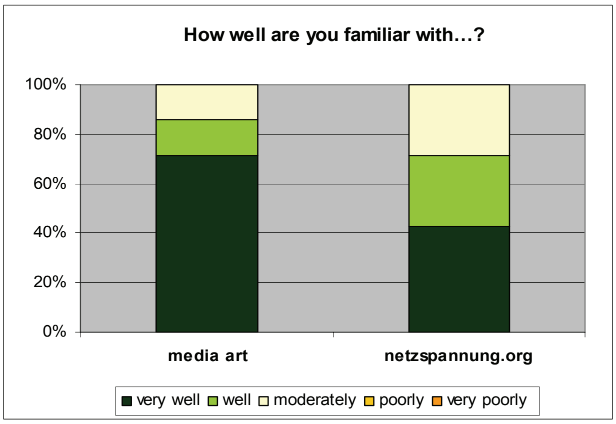


Fig. 9-6. Distribution of individual test subjects capability levels on background knowledge in Experiment I.

9.6.5. Quality of Elicited Knowledge Structures

9.6.5.1. Personal Knowledge Structures

Out of 14 personal document maps produced by the 7 test persons in accomplishing the two assigned tasks (one map per task), 13 personal document maps available for the evaluation (one map was not saved properly by a user). Table 9-13summarizes the main properties of the created personal maps.

Document Coverage	
avg. doc per map	29,31
median	29
lower quartile	18
upper quartile	37
avg. doc per cluster	3,40
total documents	381
unique documents	202
total coverage	46%
avg. coverage per map	13%

Concept Distribution	
avg. user concepts per map	8,62
median	9
lower quartile	7
upper quartile	10
total user concepts	112
unique user concepts	92
multi-occurence concepts	15

Table 9-13. Document coverage and concept distribution of personal document maps in Experiment I.

The data show that already with a small group of users (seven) working on a small number of tasks (two) in limited time (2h) a large number of concepts was created (92 unique concepts). On average, each personal map indexed 13% of documents from the test collection (435 documents total) with a total of 46% of documents in the test collection contained in all maps.

The documents were structured into 9 clusters on average (median) characterized by user-defined concepts with an average of 3,4 documents per concept (cluster). The inter-quartile range is relatively small (7-10 concepts) showing that the majority of users exhibited very similar structuring behaviour with respect to map granularity.

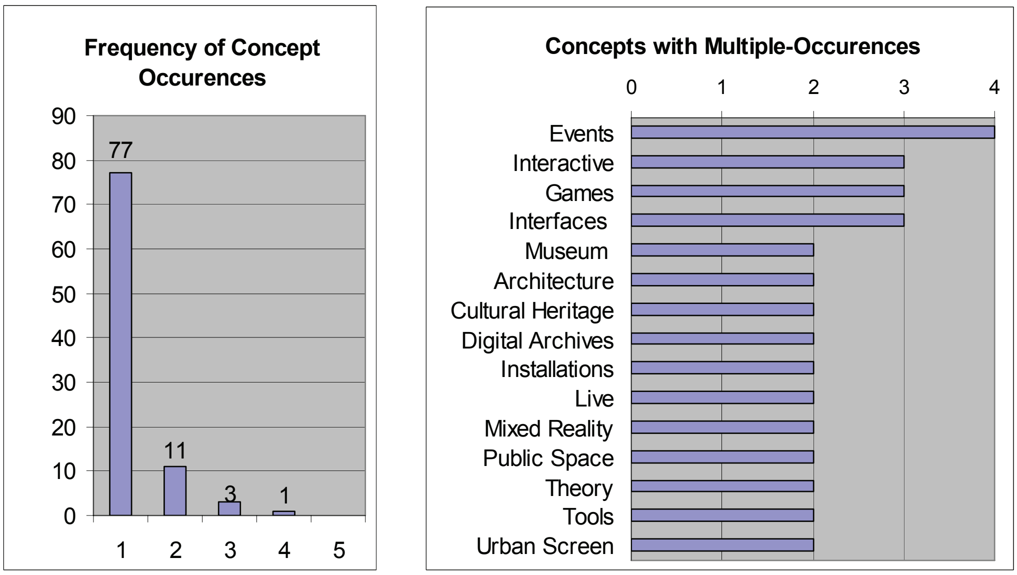


Fig. 9-7. Distribution of concepts in personal document maps in Experiment I.

The use of concepts differs very much between the users. A total of 112 out of which 92 unique concepts were used. This shows very little concept overlap and demonstrates that personal maps clearly represent very specific personal points of view.

At the same time, the significant topical intersection between the tasks is reflected in the large amount of documents indexed by more than one concept (179 multi-facetted documents i.e. 88% of unique documents indexed by the users). This provides further evidence that individual users largely used different concepts to characterize the same thematic aspects.

The distribution of concepts given in Fig. 9-7 shows that few concepts occurred more than once (16%) and rarely more than twice (4%). With regard to concept quality, the user concepts are very expressive: they include both broader themes such as “Interfaces” as well as specific topics such as “Cultural Heritage”, “Mixed Reality” or “Urban Screens”.

Based on the personal document maps of each user a corresponding personal concept map has been generated, according to the method presented in Chapter 6.4. According to the knowledge map model introduced in Chapter 6.1 the combination of the personal document maps and a personal concept map serves as a means of externalizing personal knowledge structures of individual users and making them useful for others. Two examples of personal Concept Maps are given in Fig. 9-8.

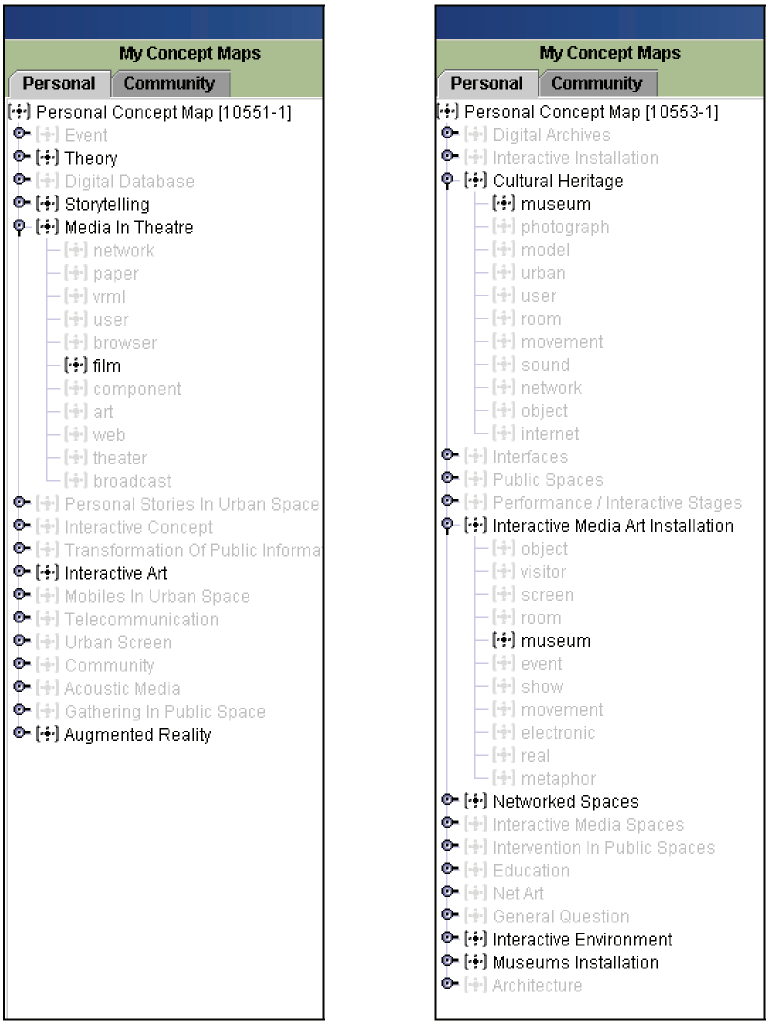


Fig. 9-8. Example of two personal Concept Maps generated in Experiment I.

The quality of personal concept maps can be characterized with respect to the number of concepts, the descriptive power and the structural composition of concepts. The higher the number of concepts, the better granularity with which personal knowledge of a given user is visualised and made usable for navigation in a document space.

The average number of main concepts per personal concept map is 15,3 with a median of 16 and an inter-quartile range 12-17,5. To each main concept a set of most related 10 terms from the text-analysis is assigned (Chapter 6.1).

The average number of documents indexed by personal concept maps is 54,4 with 28,9 unique documents per map (6,6% of the test collection). The average number of multi-facetted documents is 25,6. Such data suggest that approximately 50% of documents is indexed by more than one concept which signifies good interconnectedness within the maps.

The results of the analysis of the personal concept maps with respect to descriptive power and structural composition are given in Table 9-14. and Fig. 9-9. On average, the personal concept maps exhibit a high relative proportion of topical concepts (63%) with some structural concepts (19%) amounting jointly to 82% of concepts. There is occasional appearance of idiosyncratic (11%) and narrative concepts (7%) with marginal occurrence of unusable concepts (2%).

Characterization of User Concepts					
Descriptive power	**avg. nr.**	**%**	**median**	**LQ**	**UQ**
topical concepts	9,6	63%	7,0	5,5	13,5
structural concepts	2,9	19%	3,0	2,0	3,0
narrative concepts	1,0	7%	0,0	0,0	1,5
idiosyncractic concepts	1,7	11%	1,0	0,0	2,0
unusable concepts	0,3	2%	0,0	0,0	0,5
Structure					
single-word concepts	7,1	47%	5,0	3,5	11,5
multi-word concepts	8,1	53%	16,0	12,0	17,5

Table 9-14. Characterization of personal concept maps by descriptive power and structural composition.

There is a relatively large fluctuation in the number of topical concepts between individual maps suggested by a large inter-quartile range (5,5-13,5). The maps contain a comparable average amount of single-word and multi-word concepts, with some preference for the latter (47% vs. 53%). The fluctuation of multi-word concepts is also relatively high (inter-quartile range of 12-17,5).

Such results suggest a high topical quality of personal concept maps with significant differences between individual users, reflecting their personal structuring behaviour. The analysis of individual maps confirms this: some concept maps are highly topical with a majority of multi-word concepts, whereas others exhibit a more balanced relation between topical and structural concepts with a preference for single-word concepts.

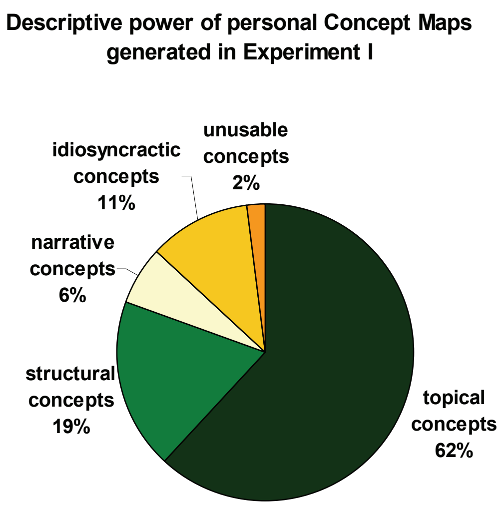


Fig. 9-9. Descriptive power of personal Concept Maps generated in Experiment I.

9.6.5.2. Community Knowledge Structures

The generation of the community concept map from the personal maps created by the users (Chapter 6.5) was configured with a two level depth[67] containing 20 top level concepts and 10 concepts at the sub-concept level. The number of concepts on each level was established with respect to the manageability of the user interface based on experiences reported in (Davies et al., 2003) and on user feedback from the formative evaluation.

The population of the sub-concept level was divided to equal parts between user-defined sub-concepts and terms from the text analysis (5 each). The number of cross-concept relations has been set to 5 as well. Basic structural properties of the community concept map are given in Table 9-15.

Concept Distribution	nr	%
main concepts	20	9%
sub-concepts	200	91%
total concepts	220	100%
total user concepts	120	55%
total terms	100	45%
unique user concepts	**59**	**27%**
unique terms	**40**	**18%**
Concept Coverage	**0,6**	**64%**
Cross-Connectedness	**nr**	**%**
multi-role concepts	18	33%
multi-role frequency	2,2	-

Table 9-15. Basic structural properties of the community concept map generated in Experiment I.

The map contains 59 unique user concepts (incl. both main and sub-concepts) which amounts to 27% of the total number of user-defined concepts. This points to a very high level of inter-connectedness within the map. Similarly, there are 39 unique terms from the text-analysis which amounts to 18% of the total number of terms. This suggests that the structural differentiating power of user-defined concepts is notably higher than of the terms from text-analysis.

Another important aspect is the concept coverage describing the proportion of the overall available user concept set from all personal maps, covered by the unique user concepts contained in the community concept map. Concept coverage of 64% suggests a good degree of coverage of the overall user concept population[68].

High cross-connectedness is already suggested by the 25% of unique concepts. This is further confirmed by the number of multi-role concepts (18 or 80% of main concepts and 33% of unique user concepts) describing how many main concepts also appear as sub-concepts in other concept clusters.

Simultaneously, a low average occurrence frequency ensures that cross-connectedness is balanced in favour of differentiating power. Such data suggests that the generated community concept map displays a good coverage of the concepts from users' personal maps and offers a well-connected as well as well-differentiating conceptual structure.

[67] With respect to common depth measure for hierarchically organized ontological structures this would correspond to a three level depth, since there is an implicit root concept in our case which corresponds to the user or community, whose knowledge structure is represented. But it makes no sense to count this as a level in our case.

[68] This degree can be increased to some extent by increasing the proportion of user concepts to be included at the sub-concept level and decreasing the proportion of text-analysis terms.

Characterization of User Concepts				
Descriptive Power	**nr unique**	**% unique**	**nr total**	**% total**
topical concepts	37	63%	83	69%
structural concepts	8	14%	16	13%
narrative concepts	6	10%	10	8%
idiosyncractic concepts	6	10%	7	6%
unusable concepts	2	3%	2	2%
Structure				
single-word concepts	21	36%	43	36%
multi-word concepts	38	64%	77	64%

Table 9-16. Characterization of descriptive power of the community concept map generated in Experiment I.

The characterization of descriptive power of the community concept map is given in Table 9-16 and visualised in Fig. 9-10. The majority of concepts is topical, with comparable distribution between unique (63%) and total concepts (69%). There is a small portion of structural concepts (14% unique, 13% total), an even lesser part of narrative (10% unique, 8% total) and idiosyncratic concepts (10% unique, 6% total) and only a few unusable concepts (3% unique, 2% total). Regarding concept structure there is a significant majority of multi-word concepts (64%) which increases the expressiveness of the concepts.

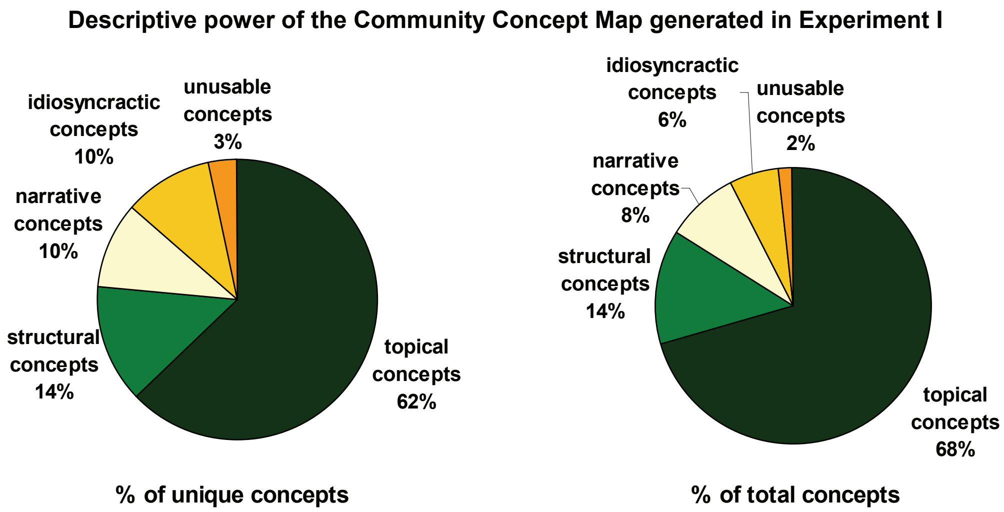


Fig. 9-10. Descriptive power of the Community Concept Map generated in Experiment I.

The distribution of user concept occurrences in the map (Fig. 9-11, left) confirms the high degree of inter-connectedness implied by the previously discussed indicators. Half of the concepts occurs more than once: 20% of concepts occurs twice and another 20% three or four times. The set of concepts with multiple occurrences is shown in Fig. 9-12.

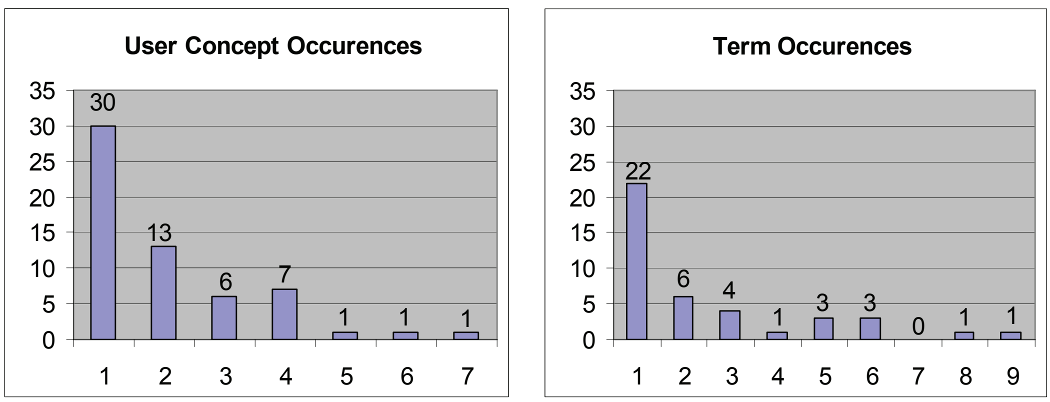


Fig. 9-11. Distribution of user concepts and text-analysis terms in the community Concept Map.

The three most occurring concepts occur more than four times each and reflect important topics from the community domain: “knowledge spaces” (7 occurrences), “mixed reality” (6 occurrences), “interactive installations” (5 occurrences). While two of them are also in the top three concepts with respect to the number of indexed documents (“interactive installations”, “mixed reality”), the third doesn’t occur at all among the main concepts sorted by document coverage (Fig. 9-13). A high number of occurrences of such a concept implies a high degree of similarity of documents indexed by this concept with the documents indexed by other main concepts. This suggests a good ability of the proposed method to discover implicit relationships between different personal concepts used by different users.

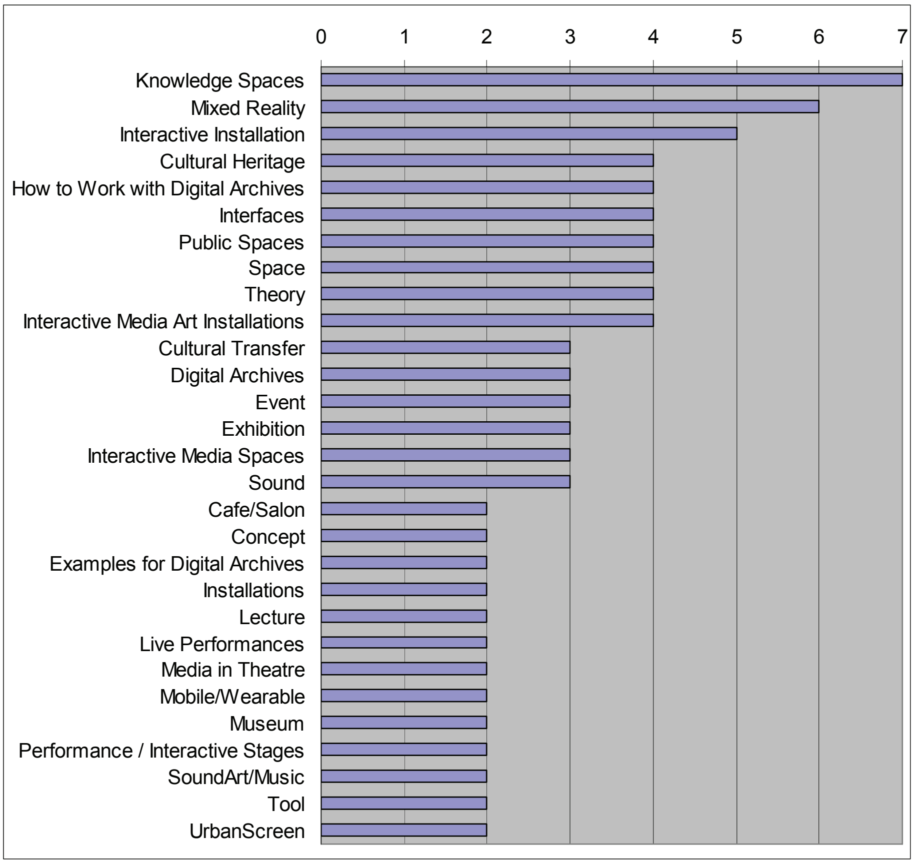


Fig. 9-12. User concepts with multiple occurrences in the generated community Concept Map.

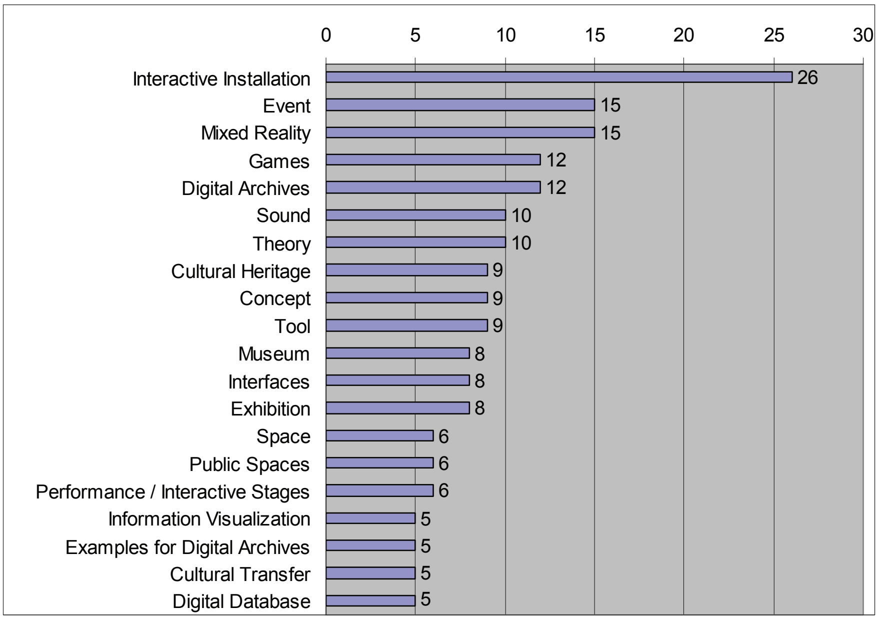


Fig. 9-13. Main concepts in the community Concept Map with respect to the number of indexed documents.

MAIN CONCEPT	SUB-CONCEPTS	
USER CONCEPTS FROM PERSONAL MAPS	USER CONCEPTS FROM PERSONAL MAPS	TERMS FROM TEXT-ANALYSIS
Interactive Installation	Space, Networked Spaces, Public Spaces, Interactive Media Art Installation, Interactive Media Spaces	room, visitor, event, sound, person
Event	Cultural Transfer, Exhibition, Important Exhibition, Theory Archives, How to Work with Digital Archives	network, art, electronic, field, concept,
Mixed Reality	Theory, Interactive Installation, Cafe/Salon, Cultural Heritage, Virtual Reality	concept, feedback, user, model, room,
Games	Citygames, Storytelling, Environment/immersion, Augmented Reality, Interactive Installation	player, film, art, sound, graph,
Digital Archives	Knowledge Spaces, Exhibition, Tools for Archives, Examples for Digital Archives, Paris	field, internet, art, content, sound,
Sound	Cultural Heritage, Sound Art/Music, Interactive Installation, Installations, Mixed Reality	sound, visitor, object, show, model,
Theory	Mixed Reality, Information, What to do with Digital Archives, Mobile/Wearable, Public Spaces	paper, art, environment, network, concept,
Cultural Heritage	Sound, Architecture, Mixed Reality, Interfaces, Cultural Transfer	museum, photograph, model, urban, room,
Concept	Interfaces, How to Work with Digital Archives, Tool, Mobile/wearable, Performance/Interactive Stages	sound, human, model, network, channel,
Tool	Interactive Media Spaces, Technology, Media in Theatre, Sound Art/Music, Concept	network, graph, art, sound, visitor,
Museum	Interactive Media Art Installations, Exhibition, Installations, Museum Installations, Interactive Environments	museum, visitor, networld, network, internet,
Interfaces	Forms of Interface and Interaction, Cultural Heritage, General Questions, Technical, Video	movement, user, touch, person, control,
Exhibition	Important Exhibition, Event, Digital Archives, Society, Interactive Art	art, internet, field, museum, public,
Space	Interactive Installation, Knowledge Spaces, Cafe/Salon, Examples for Digital Archives, Interactive Media Spaces	visitor, network, sound, track, event,
Public Spaces	Interactive Installation, Urban Screen, Theory, Digital Art Project	event, urban, person, internet, screen,
Performance/Interactive Stages	Live Performances, Concept, Media in Theatre, Museum, Mixed Reality	sound, actor, dancer, field, event,
Information Visualisation	Knowledge Spaces, Public Space, Urban Screen, Event, Timebased	art, model, field, content, domain,
Examples for Digital Archives	Knowledge Spaces, Mixed, Space, Digital Archives	sound, art, event, actor, learn,
Cultural Transfer	Event, How to Work With Digital Archives, Digital Database, Cultural Heritage	field, internet, electronic, content, mail,
Digital Database	Lecture, Digital Archives, How To Work With Digital Archives, Cultural Transfer, Theory Technology	internet, public, mind, concept, object

Table 9-17. Composition of the community Concept Map generated in Experiment I.

The map also exhibits a high quality of discovered relationships between different user concepts, as can be observed in Table 9-17, depicting the complete composition of the community concept map. The assignment of sub-concepts to main concepts shows a high correctness and rich extent of semantic relationships.

For example, the sub-concepts of the main concept "Mixed Reality" describe a number of important aspects of this topic in the community domain:

- "Interactive Installations" are the main kinds of interactive systems realizing mixed reality concepts in netzspannung.org
- "Cultural Heritage" is an application field for which different projects can be found in the test collection,
- "Theory" corresponds to the importance of theoretical considerations of the topic which are well-represented in the community,
- "Café/Salon" is an idiosyncratic but correct concept reflecting referring to projects dealing with mixed reality in public spaces that is an important community topic,
- "Virtual Reality" establishes a link to a related topic.

Similar quality can be observed for other concepts as well. Sub-concepts of "Interactive Installation" describe specific aspects of the topic (all installations are situated in "Space"), refer to more specific characterization of interactive installations ("Interactive Media Art Installation") and related topice ("Interactive Media Spaces") or describe special application contexts ("Networked Spaces", "Public Spaces") in which such installations occur. The concept "Performance/Interactive Stages" that is not intuitively understandable for non-community members becomes easily comprehensible by looking at its sub-concepts: "Live Performances" suggest real-time action, "Media in Theatre" and "Museum" imply fields of application and "Mixed Reality" a related technology. In a similar way, the most typical approaches and meaning of the notion of "Cultural Transfer" within the community are exemplified by its sub-concepts referring it to the field of digital archives and databases, applications in cultural heritage and events as common ways through which cultural transfer is supported.

Overall, the presented analysis suggests that the proposed method is well-suited for generating implicit community knowledge structures based on eliciting personal points of view of individual community members. The resulting community concept map exhibits high topical quality and a high-degree of inter-concept relationships and good quality of discovered relationships between concepts used by individual community members. This shows that the proposed method achieves good results in connecting implicitly expressed personal structures of individual users into a shared structure.

9.6.6. User Assessment of System Task-Adequacy

This section presents the results of users' assessment of the task-adequacy of the Knowledge Explorer gathered by means of a qualitative questionnaire. All test persons have perceived the tasks to be realistic (43%) to very realistic (57%) as can be observed in Fig. 9-14.

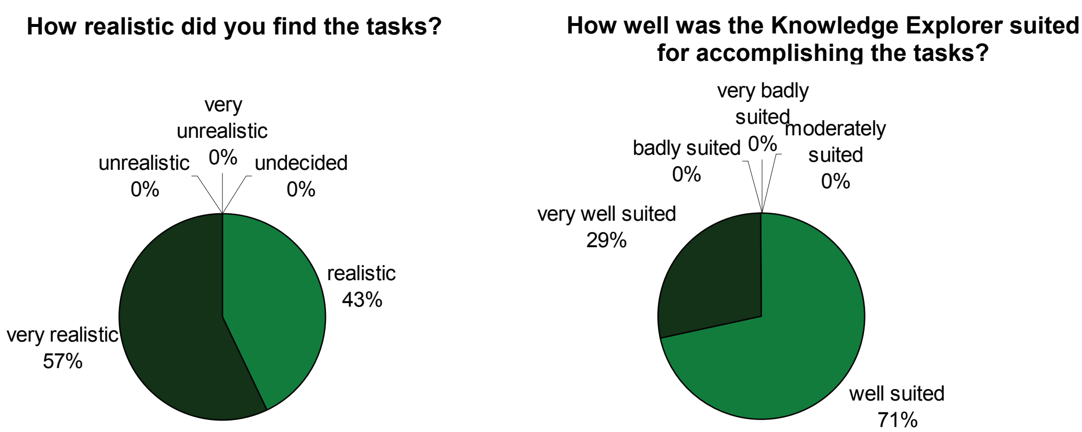


Fig. 9-14. User-perceived degree of realism of the assigned tasks (left) and of the task- adequacy of Knowledge Explorer in Experiment I

The overall suitability of the Knowledge Explorer for accomplishing the assigned tasks has also been rated very high: 71% of users considered it well suited for accomplishing the tasks and 29% very well suited (Fig. 9-14). No users reported a negative judgment. Such results confirm the adequacy of the application of the proposed approach to supporting intra-community information seeking in the bootstrapping phase and its prototypical realization in the Knowledge Explorer. However, the extent of positive judgment is an interesting result. Alongside with free form user comments in the usability questionnaire and informal feedback, it suggests that even in intra-community knowledge access there is a need for tools that support the discovery of unexpected relationships diverging from familiar community views and structures.

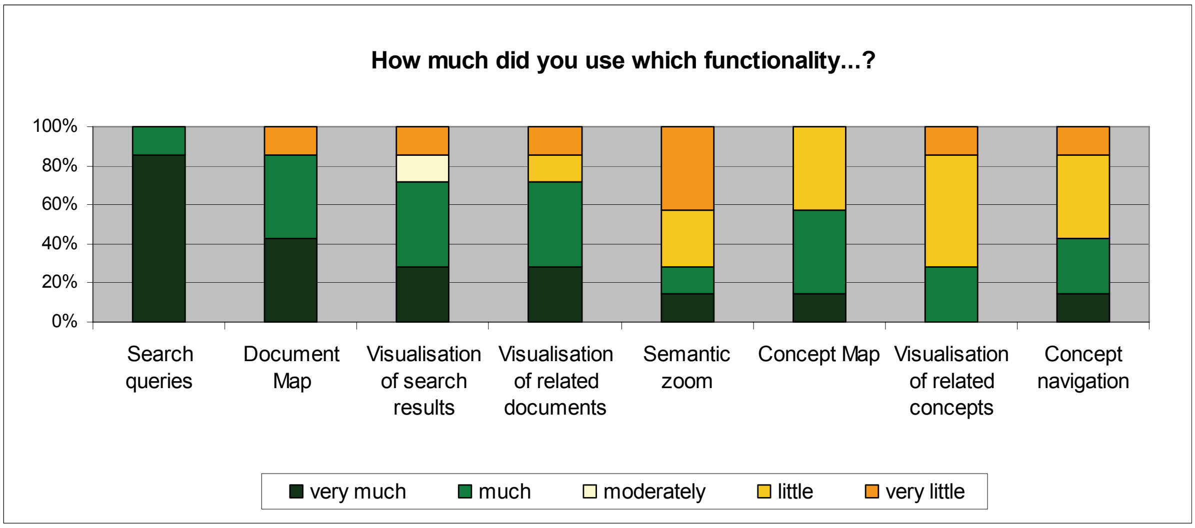


Fig. 9-15. User-reported extent of use of individual Knowledge Explorer functionalities in Experiment I.

The extent of use of individual functionalities and their usefulness as perceived by the users is depicted in Fig. 9-15 and Fig. 9-16. The results show that most extensively used were the search queries (86% "very much" use) and the Document Map (86% "much" or "very much" use). The visualisation of search results and related documents were also used to a great extent (72% of "much" or "very much" use). The Concept Map was also significantly used by roughly half of the users (57% of "much" or "very much" use) with concept navigation used more than the visualisation of related concepts. The semantic zoom was used the least with 72% of users reporting "little" or "very little" use (the user feedback in the overall usability questionnaire shows that the reason were usability problems experienced by the users).

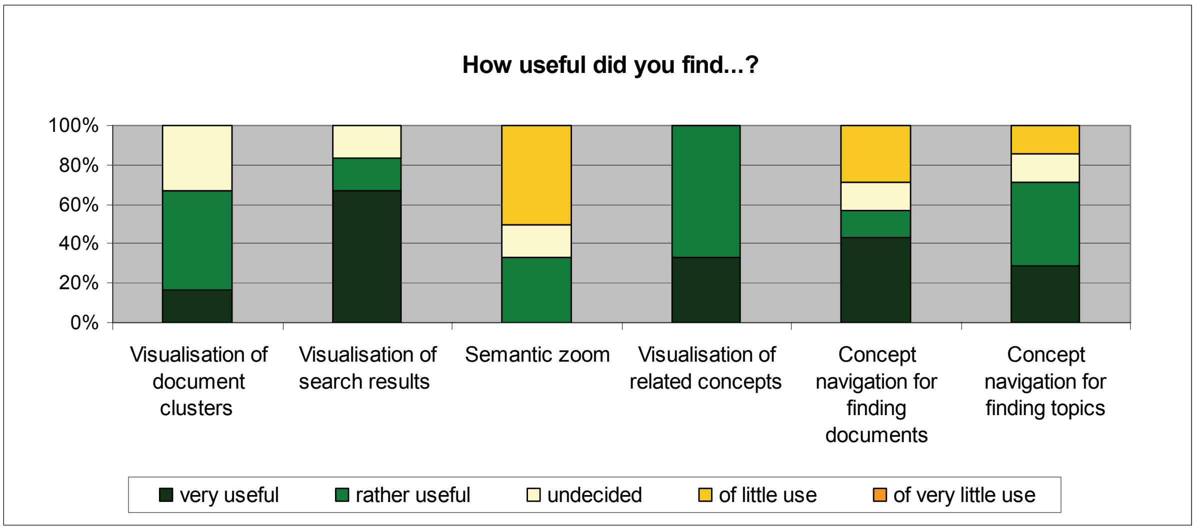


Fig. 9-16. User-reported usefulness of individual Knowledge Explorer functionalities in Experiment I.

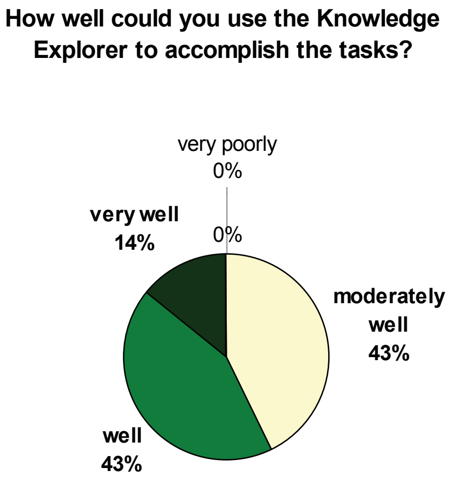
The high extent of use of search queries and the Document Map visualisation was expected, since users are familiar with the domain of the task topics and the community vocabulary but the task topics were broadly defined. This lends itself to a combination of goal-directed search queries and explorative access of topical clusters. The same kind of reasoning also explains the relatively low extent of use of Concept Map functionalities.

But surprisingly, user reported usefulness of individual Knowledge Explorer functionalities (Fig. 9-16.) emphasises a high usefulness of Concept Map functionalities as well. The two best rated functionalities are the visualisation of search results (57% ratings “very useful” and 14% ratings “rather useful”) and the visualisation of concept related to results of a search query or a set of manually selected documents (86% of users giving it a useful rating: 57% “rather useful” , 29% “very useful”). Concept navigation for finding both documents and topics was also rated well: 72% rather useful or very useful for the latter and 57% for the former.

Such results are surprising in a positive sense. These Concept Map functionalities were designed with the aim of supporting the discovery of appropriate concepts for expressing an information need in an unfamiliar domain. Accordingly, the users in general didn’t tend to use them much (Fig. 9-15). But as it turns out, these functionalities do have a special purpose in the intra-community use context as well. Users repeatedly remarked that seeing specific concepts visualised as a result of their search query or document selection made them aware of concepts they would not use themselves or expected to be related to a given context. This provoked their interest and curiosity into discovering the reason for this relationship which led them to selecting those concepts and inspecting the results in more detail. Accordingly, they report to have discovered interesting projects and relationships between topics they were not aware of before, though being familiar with the field and with similar projects in the community space. Thus they rated this Concept Map functionality highly useful because it led them to discovering new and previously unavailable points of view.

9.6.7. Overall Usability Assessment

Fig. 9-17 displays the results of the questionnaire concerning the overall usability of the Knowledge Explorer. The majority of users was able to use the Knowledge Explorer well for accomplishing the tasks (57% well or very well, Fig. 9-17 left), found it easy to use (86% easy or very easy, Fig. 9-17 right) and very easy to learn using it (86% very easy, Fig. 9-18 left).



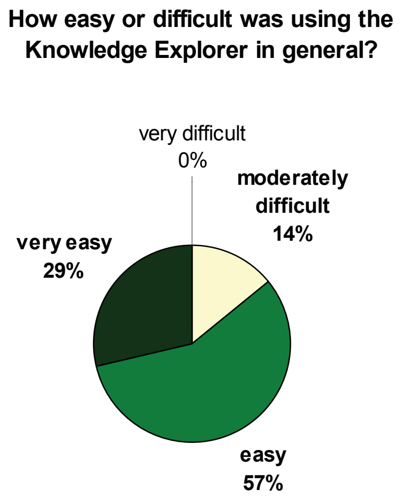


Fig. 9-17. User assessment of the ease of use of the Knowledge Explorer in Experiment I.

Working on the tasks was perceived as strenuous by the majority of users (62% moderately strenuous or strenuous, Fig. 9-18 right) whereby most users remarked that the strenuousness stemmed from the cognitive demand of the tasks and the length of the experiments, rather than from using the system itself.

Especially rewarding was unsolicited user feedback to the system experience. After the experiment, a great majority of users self-initiatively remarked that they found the tool to be very interesting and that they would have liked to try it out some more. Several users stated that the tool was “fun to use” and inquired about the possibility to use the tool for their professional work and research.

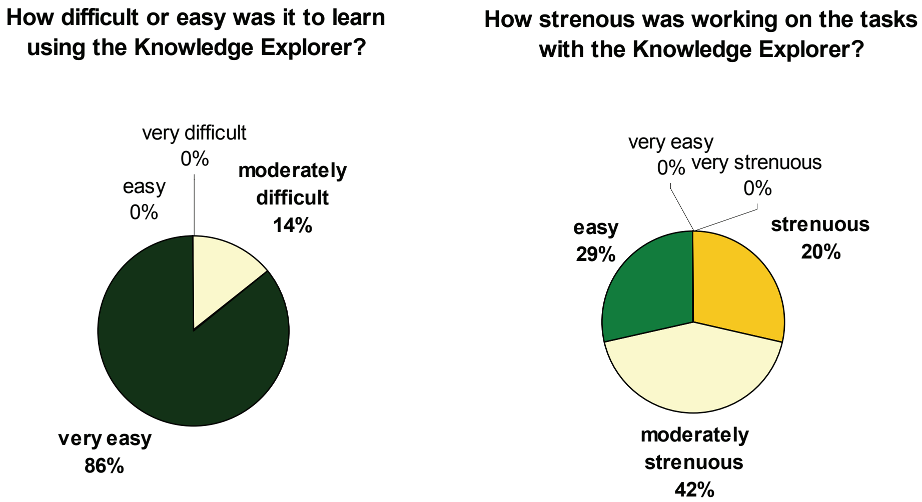


Fig. 9-18. User assessment of the ease of learning to use the Knowledge Explorer and the streneousness of accomplishing the tasks in Experiment I.

9.6.8. Conclusions

The results of Experiment I discussed in the previous sections suggest that there is significant supportive evidence for accepting the hypothesis (H1) stating that: *the proposed method for eliciting implicit knowledge structures of human users based on their interaction with information is well-suited for eliciting personal knowledge structures of individuals and shared structures of communities* (Section 9.1).

The analysis of elicited personal knowledge structures (Section 9.6.5.1) has shown that already a small group of users working on only two tasks within limited time produced a large number of unique concepts, indexing almost half of the test collection (46% out of 435 documents). Both the personal document maps and personal concept maps exhibit significant differences in the use of concepts between the users (low concept overlap, Fig. 9-7) despite significant topical intersection between the tasks (definition of task topics, large number of multi-facetted documents). This demonstrates that the elicited knowledge structures of individual users clearly represent very specific personal points of view on related thematic aspects. The qualitative analysis of the concepts created by the users showed a high expressiveness and good descriptive power characterized by the majority of topical concepts, balanced presence of structural concepts and low occurrence of idiosyncratic expressions.

The analysis of the community concept map has shown that already from a small number of personal document maps an appropriate shared structure inter-connecting personal points of view could be created. The resulting conceptual structure is characterized by high descriptive power (large proportion of topical concepts, low occurrence of idiosyncratic concepts, Fig. 9-10), a high degree of interconnectedness and good coverage of the concept population used by the members of the community test sample (Table 9-15). The subject areas and topics of the community domain contained in the test collection are well reflected in the discovered relationships between different user concepts. In particular, the assignment of sub-concepts to main concepts shows a rich extent of semantic relationships (Table 9-17). Such results confirm the suitability of the proposed method for uncovering implicit community knowledge structures based on eliciting personal points of view of individual community members.

Overall, the results of Experiment I suggest that we can accept the hypothesis (H1) at a qualitative level. In addition, the positive user assessment of Knowledge Explorer task adequacy (Section 9.6.6) confirms the suitability of the proposed bootstrapping solution and use of the Knowledge Explorer for intra-community information access. In particular, the results of the usefulness assessment for individual functionalities confirm the suitability of the proposed knowledge map model based on a document map - concept map combination (Fig. 9-16). This provides supportive evidence for accepting hypothesis (H2) at a qualitative level. Finally, the results of the overall usability assessment (Section 9.6.7) show a high degree of user acceptance of the Knowledge Explorer system, with particularly high ease of use (Fig. 9-17) and ease of learning to use it (Fig. 9-18).

9.7. Experiment II: Using Personal and Community Knowledge Maps for Accessing Community Spaces

The principal aim of the evaluation study is to investigate how well the proposed approach based on collaborative knowledge visualisation supports the identification and understanding of relevant knowledge from an unfamiliar community. Such knowledge access is thereby mediated through access to the unfamiliar community information space (Chapter 2.8). To this end, in this second part of the study the prototypical implementation of the proposed approach in the form of an interactive system is compared to a standard information seeking system which does not offer knowledge visualisation. This both reflects the common situation found in practice (Chapter 2) and serves as the basis for a comparative evaluation of the contribution of the own approach, based on the hypotheses H3-H5 (Section 9.1).

9.7.1. Experiment Design

As outlined in Section 9.3 this experiment is divided into two phases i.e. sub-experiments: the initialization phase and the test phase:

- **Initialization phase (Experiment IIa)**: First, the test users perform three different tasks in succession within their own community information space (CHI information space). This phase serves for generating the necessary personal and community structures that will be used alongside with knowledge maps of the unfamiliar community (netzspannung.org) in the subsequent test phase. No hypotheses are tested here.
- **Test phase (Experiment IIb)**: In the actual experiment, the test users perform four different tasks in succession within the unfamiliar community information space of netzspannung.org. The hypotheses H3-H5 are tested here.

The assignment of participants to test groups and the assignment of systems to test groups in the initialization phase is the same as in the actual test phase. Only the configuration of the Knowledge Explorer differs in the two phases. In the initialization phase both Knowledge Explorer test groups use the same, bootstrapping configuration of the Knowledge Explorer system, employing system-generated maps.

In this way, the necessity of using the Knowledge Explorer system in the initialization phase doesn't influence the results of the actual experiment in the subsequent test phase. Furthermore, as the tasks completed in the initialization phase are performed on a different information space than the one in the test phase, no carry-over effects exist between the two experiments[69].

The main purpose of the tasks in the initialization phase is to simulate intra-community information seeking tasks from which some personal and community knowledge structures can be extracted by the system. The general structure of the tasks for the initialization phase is as follows:

- **Task 1.1 / Task 1.**2: present users with a familiar specific topic and requires them to identify relevant documents in their own community information space.
- **Task 1.3:** requires users to identify topics and documents of personal interest in their own community information space and organize the results in an appropriate topical structure.

[69] With respect to the knowledge of the unfamiliar community domain.

Experiment IIa	Between-Subjects Design		
	Task 1.1	Task 1.2	Task 1.3
Group A	System-generated maps	System-generated maps	System-generated maps
Group B	System-generated maps	System-generated maps	System-generated maps
Group C	Google + Mozilla	Google + Mozilla	Google + Mozilla

Table 9-18. Test design for the initialization phase of Experiment II (IIa)

The actual experiment in this part of the study (Experiment IIb) is based on a between-subjects design (Table 9-19). Three test user groups perform four different tasks in succession with an assigned system configuration. Each test group is composed out of six participants from the simulated HCI community sample (Section 9.5.3). A different system configuration is assigned to each group. Two groups are assigned different configurations of the proposed knowledge visualisation system (Knowledge Explorer): the configuration with community maps and the configuration with personal maps, respectively. The third group uses a standard information seeking reference system.

Experiment IIb	Between-Subjects Design			
	Task 2.1	Task 2.2	Task 2.3	Task 2.4
Group A	Community Maps	Community Maps	Community Maps	Community Maps
Group B	Personal Maps	Personal Maps	Personal Maps	Personal Maps
Group C	Google + Mozilla	Google + Mozilla	Google + Mozilla	Google + Mozilla

Table 9-19 Basic between-subjects test design for the test phase of Experiment II (IIb)

Accordingly, the basic design of the actual test phase (Experiment IIb) is a standard between-subjects design. A cross-over design where each test group is exposed to all system configurations would provide the advantage of being able to directly measure user-perceived system adequacy by eliciting user rating of system preference. This is not feasible in our case due to significant carry-over effect (acquired knowledge of the unfamiliar community domain) that would occur after the first system exposure.

The advantage of the chosen between-subjects design is thus that we can be confident that no carry-over effects have influenced the results. The disadvantage is the lack of explicit user rating of comparative system preference between the alternative system configurations. This has been partially compensated by extending the basic design with an additional exposure of the third group to the Knowledge Explorer system (Table 9-20).

After completing their tasks with the standard reference system, the control group repeats two tasks with both configurations of the Knowledge Explorer system, one after another. Due to obvious carry-over effects, the objective solution quality is of no relevance in this case. Instead, the aim of this exposure is to elicit direct user feedback on comparative system preference between Knowledge Explorer and the standard reference system.

The rationale behind this choice is the assumption that in accomplishing the tasks the control group using the standard system will have experienced and reported the very problems that were the motivation for the design of our solution in the Knowledge Explorer. Having used the Knowledge Explorer for the same kind of task after the standard system, the users can provide some indication as to whether they perceived the Knowledge Explorer as helpful for alleviating the problems they experienced when using the standard system. Although, this additional exposure will have to be very limited in time, so as not to exceed the overall acceptable duration of the test, it should provide us with valuable insights on the direct user comparison of task adequacy of our solution with respect to the standard system.

Experiment IIb	Between-Subjects Design			
	Task 2.1	Task 2.2	Task 2.3	Task 2.4
Group A	Community Maps	Community Maps	Community Maps	Community Maps
Group B	Personal Maps	Personal Maps	Personal Maps	Personal Maps
Group C	Google + Mozilla	Google + Mozilla	Google + Mozilla	Google + Mozilla
	Partial Cross-Over Extension			
Group C	Community Maps / Personal Maps	Community Maps / Personal Maps		

Table 9-20. Extended test design for Experiment IIb with partial cross-over for eliciting subjective user feedback on comparative system preference

A critical factor of an appropriate test design is the selection of tasks to be accomplished by the users. Designing appropriate tasks requires them to be relevant in practice and that an objective reference solution exists against which the quality of task solutions can be assessed. To ensure the practical relevance of the tasks they have been based on the theoretical analysis and requirements framework presented in Chapter 3. The general task structure follows the guidelines of the evaluation framework introduced in Section 9.2:

- **Task 2.1:** presents users with an ill-defined task and an ill-defined information need in the unfamiliar community knowledge domain. Users must identify relevant documents from the unfamiliar community information space and organize them into appropriate topical groups.
- **Task 2.2:** presents users with a familiar concept from their own community and requires them to identify relevant documents in the unfamiliar community information space. The presented concept is not used in the unfamiliar community.
- **Task 2.3:** presents users with a familiar concept from their own community and requires them to identify relevant documents in the unfamiliar community information space. The presented concept is also used in the unfamiliar community in ways that differ from the usage in the users own community.
- **Task 2.4:** presents users with an ill-defined task and a well-defined information need in the unfamiliar community knowledge domain. Users must identify relevant documents from the unfamiliar community information space and organize them into appropriate topical groups.

Following the simulated work task method, the tasks have been accompanied by practically relevant scenarios in order to create a meaningful test situation for the users and ensure their motivation and understanding of the purpose of their actions. A detailed description of the tasks is given in Section 9.7.3.

9.7.2. System Configurations

The prototypical implementation of the proposed knowledge visualisation approach is provided by the Knowledge Explorer, an interactive system enabling multi-perspective access to community information spaces through personal and shared community knowledge maps (described in Chapter 8). According to our hypotheses H3-H5 we shall compare the two main aspects of our solution (community maps and personal maps) to a standard reference system as well as between each other. As a reference system against which the adequacy of the functions provided by the Knowledge Explorer shall be tested, we take a combination of the Google Desktop Search and the Mozilla Browser as interface.

This combination provides us with a well-known and widely accepted information seeking system against which the added value of our approach can be compared. In order to ensure that both systems provide the same search and information organization functionalities, the Google Desktop Search engine has been integrated into the Knowledge Explorer query interface. The creation of personal maps in the Knowledge Explorer has been provided in a tree-based view much like the organization of bookmarks in Mozilla.

The main difference between the two systems is then the absence of the functionalities related to the visualisation and application of personal and community knowledge maps in the reference system. This allows us to isolate the effects of the main elements of our solution whose added value and adequacy for typical information access tasks occurring in cross-community knowledge exchange we want to evaluate. In other words, we are able to investigate the difference between access to community space without visualisation of personal and community knowledge structures – as it normally occurs in practice (Chapter 2) – and the expected benefits of our proposed knowledge visualisation solution.

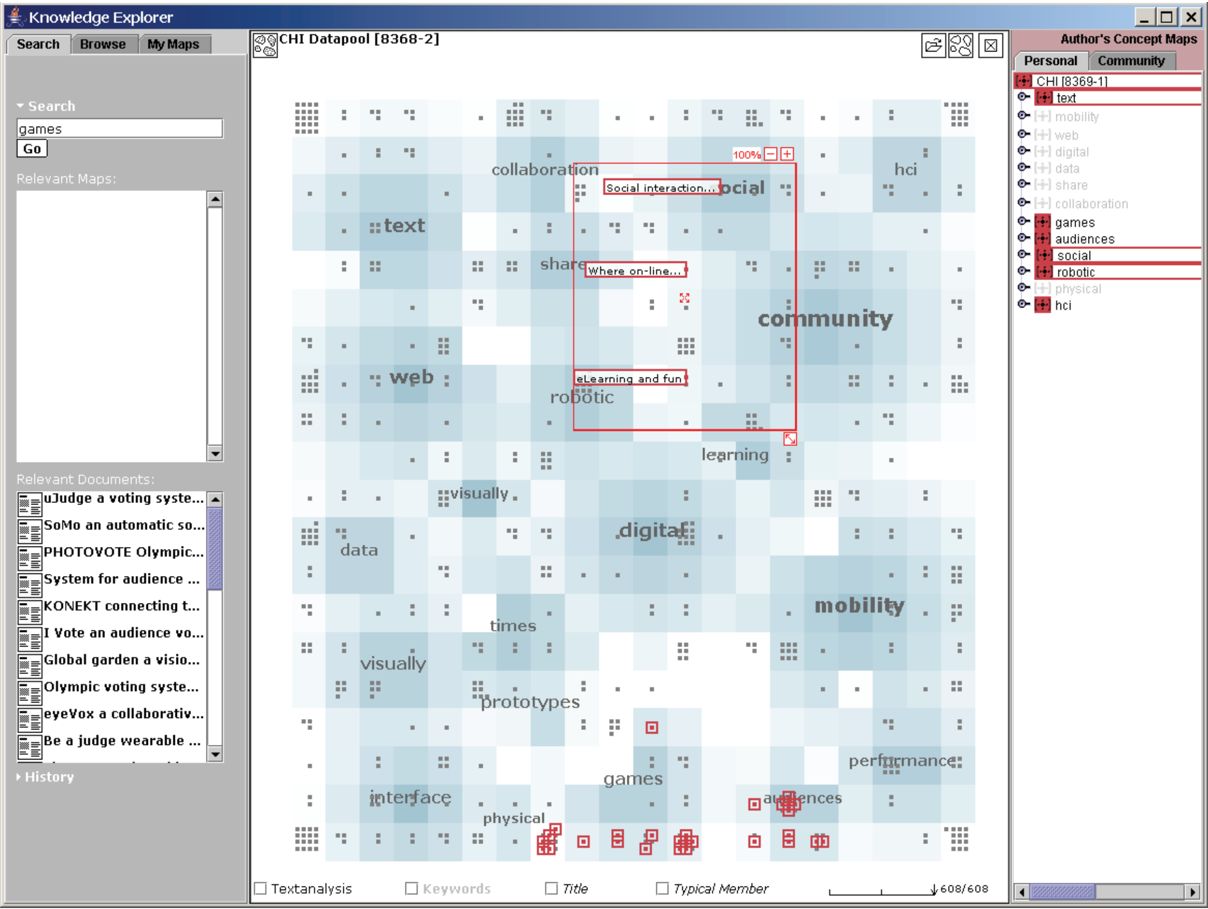


Fig. 9-19 Screenshot of system-generated map configuration of the Knowledge Explorer interface used in the initialization phase (Experiment IIa) by Group A and Group B. Provides only a Document Map and Concept Map based on text-analysis of documents in the community information space, with search visualisation and basic map interaction.

In order to test the individual hypothesis, the Knowledge Explorer system presented in Chapter 8 has been separated into three different configurations. The first configuration provides only a system-generated map Document Map and system-generated Concept Map based on the text-analysis of the documents in the test collection of a given community space. It includes search query and visualisation of search results as well as basic map visualisation and means of interaction: browsing, semantic zoom, visualisation of related concepts and concept map navigation, display of document abstracts on the map and document details in the browser window (Chapter 8.6.3)

The results of users' information seeking activities can be organized and saved in personal maps containing named clusters of related documents (Chapter 8.6.3.5). The back-end system provides only the search query engine and basic map management functionalities (retrieving system-generated maps, saving personal maps). This configuration is used in the intialization phase (Experiment IIa) by both Group A and Group B. The corresponding Knowledge Explorer interface is depicted in Fig. 9-19.

The community map configuration provides a multi-view visualisation of the knowledge map of the unknown community (Fig. 9-20, right) and of one's own familiar community knowledge structure (Fig. 9-20, left). The first consists of a Document Map presenting the semantic structure of the unfamiliar community information space and a Concept Map presenting the structure of the shared vocabulary used by members of the unfamiliar community, extracted from their personal maps (Chapter 6.5) created in Experiment I. The Concept Map of the familiar community presents main concepts used by users belonging to both Knowledge Explorer test groups representing the same community. It is generated from their personal maps created in the initialization phase (Experiment IIa).

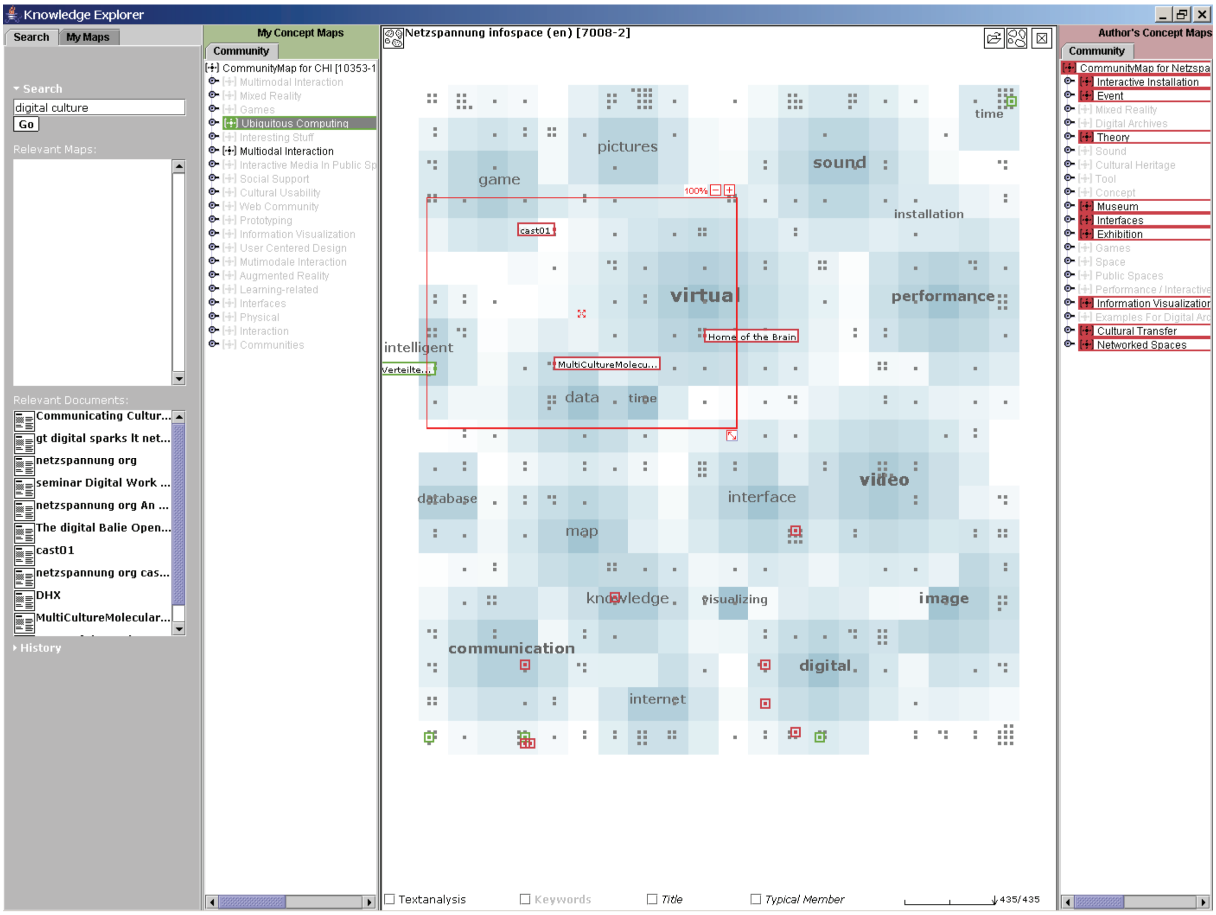


Fig. 9-20. Screenshot of the Community Map configuration of the Knowledge Explorer interface, used in the test phase (Experiment IIb) by Group A. Provides a multi-view visualisation of Document and Concept Map of the unfamiliar community map (right) and of the Concept Map of the familiar community (left). Interaction includes search visualisation with basic map interaction and multi-perspective navigation.

The interface includes all main map visualisation and interaction possibilities described in Chapter 8.6.3: search query with visualisation of search results, map browsing, semantic zoom, visualisation of related concepts and concept map navigation, display of document abstracts on the map and document details in the browser window.

It also allows multi-perspective navigation by simultaneous selection of concepts in the two different community Concept Maps (with colour coding) as described in Chapters 7.1.5 and 8.6.3.4. The results of users' information seeking activities can be organized and saved in personal maps containing named clusters of related documents (Chapter 8.6.3.5).

The back-end system provides only the search query engine and basic map management functionalities (retrieving community maps, saving personal maps). This configuration is used in the test phase (Experiment IIb) by Group A. The corresponding Knowledge Explorer interface is depicted in Fig. 9-20.

The personal map configuration differs from the community map configuration in two main respects. First, instead of visualising shared community maps it allows the user to identify personal maps of individual users relevant for a given information task in the unfamiliar community space. To this end, the basic search functionality is coupled with the matchmaking service that determines relevant personal maps based on the given query (Chapter 7.1.3, 7.2.2.2).

Accordingly, a ranked list of relevant personal map created the user from the unfamiliar community is presented and the first map from the list is displayed. Thus, instead of shared community maps this configuration provides a multi-view visualisation of a personal Document Map and Concept Map of a specific user from the unfamiliar community (Fig. 9-21, two right-most frames) alongside with one's own personal Concept Map (Fig. 9-21, second frame from the left).

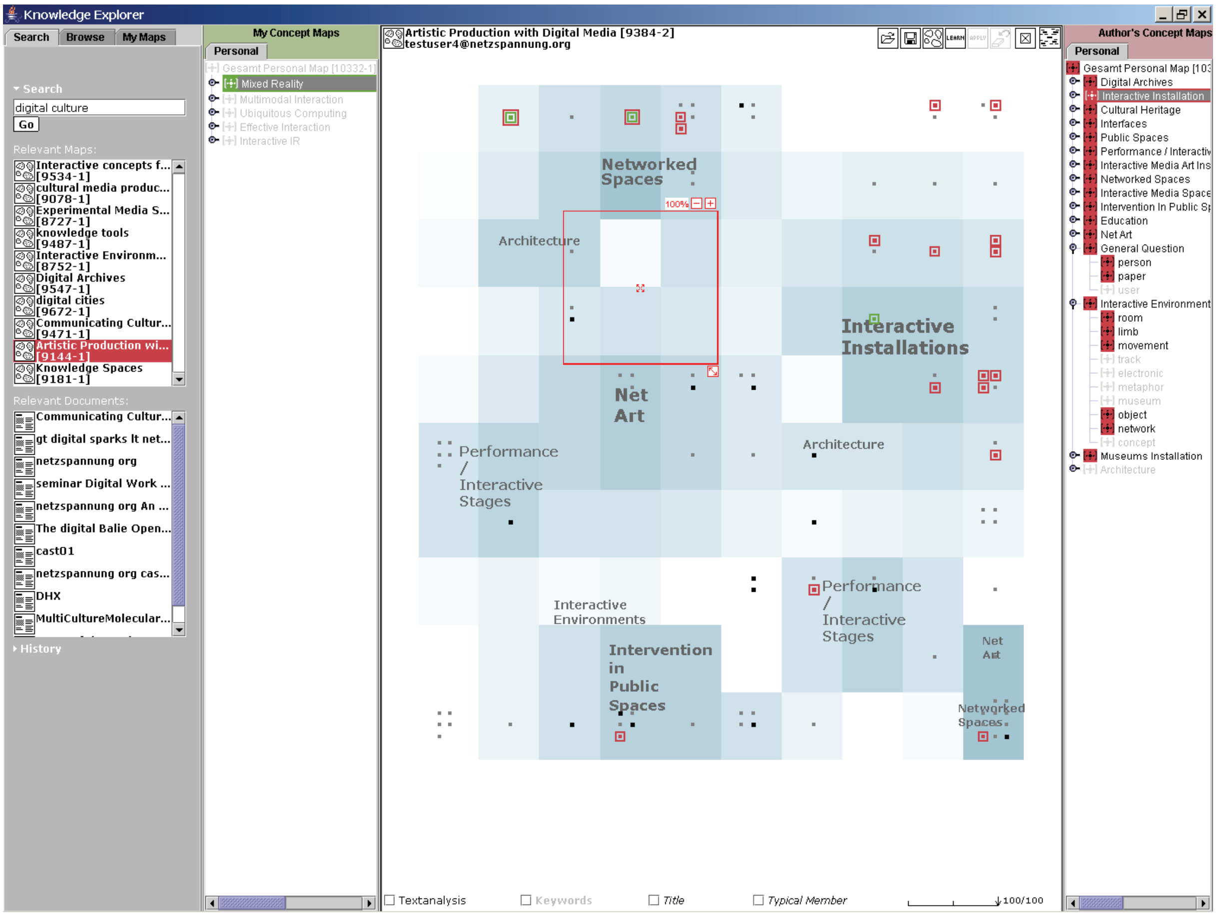


Fig. 9-21. Screenshot of the Personal Map configuration of the Knowledge Explorer interface, used in the test phase (Experiment IIb) by Group B. Provides automatic identification and multi-view visualisation of personal maps relevant for a given search query: personal Document and Concept Map of a given user from the unfamiliar community map (right) with one's own personal Concept Map (left). In addition to standard map interaction (search query and visualisation, semantic zoom, multi-perspective navigation) the user can select individual personal maps to be displayed at will.

The personal maps from the unfamiliar community are selected from the set of maps created by users from the netzspannung.org test group in Experiment I (Section 9.6). The personal concept map of the user currently using the system is extracted from this personal maps created in the intialization phase (Experiment IIa) by the method described in Chapter 6.4.

The interface includes the same map visualisation and interaction possibilities as the community map configuration: search query with visualisation of search results, map browsing, semantic zoom, visualisation of related concepts and concept map navigation, display of document abstracts on the map and document details in the browser window (Chapter 8.6.3). It also allows multi-perspective navigation by simultaneous selection of concepts in the two different personal Concept Maps (with colour coding) as described in Chapters 7.1.5 and 8.6.3.4. In addition, the user can browse the list of personal maps relevant for a given search query and select individual maps to be displayed at his own will.

As in all other system configurations, the results of users' information seeking activities can be organized and saved in personal maps containing named clusters of related documents (Chapter 8.6.3.5). In addition to the search query engine and basic map management functionalities (retrieving community maps, saving personal maps) the back-end system here provides also the matchmaking service for identifying maps relevant for a given query (Chapter 7.1.3, 7.2.2.2). This configuration is used in the test phase (Experiment IIb) by Group B. The corresponding Knowledge Explorer interface is depicted in Fig. 9-21.

The interface of the standard information seeking reference system provided by a combination of Google Desktop Search and Mozilla Browser is depicted in Fig. 9-22. Using the Google Desktop Search allows us to constrain the search index only to the selected document collections representing a given community information space in individual experiments.

This system configuration provides standard search query capabilities based on the Google search engine and uses a common web browser for list-based visualisation of search results. The main functionalities of the Mozilla browser used for the experiments include the display of document details in multiple tabs and bookmarks sidebar for organizing search results into named groups of related documents.

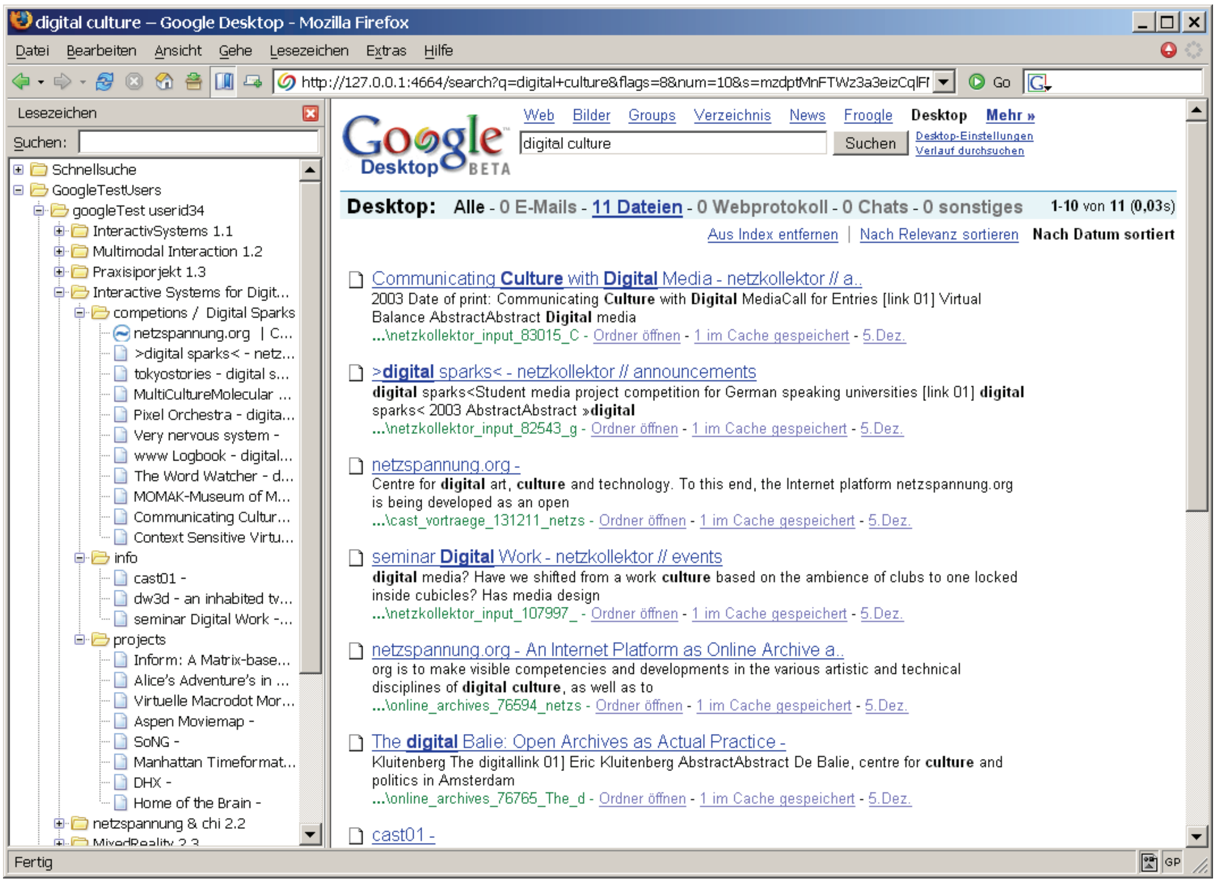


Fig. 9-22. Screenshot of the reference system Google Desktop Search with Mozilla Browser, used by Group C in both initialization and test phase (Experiment IIa, Experiment IIb). It provides standard search functionalities with bookmarks for organizing search results.

9.7.3. Task Design

The general task structure for this experiment has been introduced in Section 9.7.1. In this section we present the details of individual tasks through which the defined hypotheses shall be experimentally investigated. As described in Section 9.7.1, the test design is divided in two main sets of tasks corresponding to the two phases of the experiment: initialization phase and the test phase. Each of these task groups has a particular character stemming from the purpose of the corresponding phase. Following the simulated work task method, each task is preceded with a scenario that aims at creating a realistic setting to provide a broader context for guiding user actions and the motivation for their efforts in accomplishing the tasks. Accordingly, the two tasks sets corresponding to the two phases of the experiment are defined as follows.

9.7.3.1. Task 1: Intra-Community Information Seeking

The tasks in the initialization phase have a twofold purpose. On one hand, users need to create a set of personal maps from their own community information space, so that their personal concept maps and the concept map of the HCI community test sample can be generated. But at the same time, since the users involve students familiar with the HCI topics but not working with them on a daily basis, the initialization phase also provides a way for the users to refresh their memory of specific topics that will be involved in the tasks of the test phase. Furthermore, for the test groups using the Knowledge Explorer this is also the phase in which the subjects can gain some experience in using the system to solve realistic tasks.

Defining a realistic task scenario

Following the simulated work task method, first an overall task with a practically relevant scenario is formulated in order to create a meaningful situation for the users. Then specific subtasks are given that need to be carried out by the users within the overall scenario.

Task 1: Developing a project on the topic "Interactive Systems"
Scenario: You have to develop a project in the field of "Interactive Systems". In order to develop ideas for the project, you want to gain an overview of the research field "Interactive Systems".
Task: Search for relevant information in the information space of the CHI (Computer Human Interaction) community regarding possible topics for your project and save the results of your research in a personal map.

Table 9-21. Scenario description for the initialization phase of Experiment II[70].

The scenario is selected in such a way that it has a high degree of realism both for researchers and students comprising the test groups. For students, the notion of project is additionally described as "practical semester project" (orig. "Praxisprojekt" in German). For doctoral students and researchers, it is described as a "research project" or "phd project".

Structuring the task into practicable subtasks appropriate for our experiment design

Letting users perform such broadly defined task would be influenced very much by user's individual interest, producing in turn potentially very diverging personal knowledge structures and having very different levels of familiarity with specific topics. In order to gain some comparable results in the next phase, we must ensure that there is a comparable level of users' familiarity with the main topics addressed in the test phase. To this end, we define more specific subtasks to guide user actions as follows:

- two subtasks presenting users with a familiar specific topic and requiring them to identify relevant documents in their own community information space (Task 1.1, Task 1.2).
- one subtask requiring users to identify topics and documents of personal interest in their own community information space and organize the results in a topical structure (Task 1.3).

[70] Originally, a much longer description of the scenario has been created with the assumption that it would help create a more engaging setting for the users. But the experience from Experiment I and pre-test with a pilot user has shown that longer scenario descriptions tend to distract users and cause confusion as to the actual task focus, rather than help them.

Task 1.1. Find documents dealing with the topic „Mixed Reality“. What related topics exist? Organize the results of your search in thematically named document clusters in a personal map.
Task 1.2. Find documents dealing with the topic „Multimodal Interaction“. What related topics exist? Organize the results of your search in thematically named document clusters in a personal map.
Task 1.3. Find other documents and topics that may be interesting for your project. Organize the results of your search in a personal map.

Table 9-22. Definition of tasks for the intialization phase of Experiment II.

Tasks 1.1 and 1.2 ensure that all users have at least a minimal common denominator of familiarity with topics from own community that will play a role in the next experiment phase. Moreover, they ensure that the personal knowledge structures extracted by the system will have at least two topics in common for all users and that they will appear in the community structure as well. This provides the basis for testing the suitability of specific system functionalities that take advantage of users personal and shared community concept maps in the next phase of the experiment.

Task 1.3 allows that personal interests and particular background of individual test users can come to bear on the eventual solution and thus ensures the necessary degree of diversity of personal views, needed for a realistic community structure. The topics for Tasks 1.1 and 1.2 have been chosen in accordance with Tasks 2.1-2.4 and such that they correspond to topics treated in the courses "Interactive Systems" and "Usability Engineering" taken by the test persons during their studies. In this way the required background knowledge of the participants could be ensured.

Compiling a suitable test collection

The test collection simulating the CHI community space for Tasks 1.1-1.3 has been compiled from the proceedings of the ACM SIGCHI conferences CHI 05, CHI 04 and CHI 03 including both full papers and short papers and poster proceedings. The proceedings are available online from the ACM Digital Library.

The collection contains 608 documents and has been chose such that it contains a good proportion of HCI topics treated in the courses "Interactive Systems" and "Usability Engineering" taken by test persons. In particular, it has been verified that the collection contains the necessary amount of documents dealing with the topics defined by the subtasks and a practicable diversity of typical HCI topics that can be explored for Task 1.3.

Constructing reference solutions

Objective reference solutions for Tasks 1.1 and 1.2 have been constructed by asking one HCI researcher to examine abstracts of all documents in the collection and assign them to the appropriate topics. For Task 1.3 no reference solution could be compiled as the choice of relevant documents obviously depends on personal interests and background of individual test persons.

9.7.3.2. Task 2: Cross-Community Information Seeking

The purpose of Task 2 is to create a setting for testing the main hypotheses regarding the user of personal and community maps to support knowledge access in unfamiliar community spaces. This is the test phase of Experiment II.

Defining a realistic task scenario

The overall formulation of Task 2 provides a practically relevant scenario aimed at creating a meaningful situation for the test persons, in the following way:

<u>Task 2:</u> **Report on the topic "Interactive systems for digital culture"** **Scenario:** You have to write a report on the topic "Interactive systems for digital culture". In order to find relevant information for your report you investigate the information space of the netzspannung.org community for media art and design.

Table 9-23. Scenario description for the test phase of Experiment II.

The scenario is selected in such a way that it has a high degree of realism both for researchers and students comprising the test groups. The notion of report has been referred to "homework assignment" (orig. "Referat" in German) and for researchers and doctoral students it has been additionally described as "research report" or "state of the art report".

The task topic is formulated as a combination of familiar concepts from users' own community ("interactive systems") and of unfamiliar concepts used in the unknown community netzspannung.org ("digital culture"). Furthermore, the latter concepts is ill-defined even within its own community of origin, as different meanings and interpretations exist, without an explicit definition of the term in any of the documents contained in the community space. An understanding of the meaning of the concept in the netzspannung.org community is commonly acquired by following the community discourse which has not been available to the test persons. In this way, our scenario simulates the typical problem situation of cross-community knowledge access as described in Chapter 2.8.

Structuring the task into practicable subtasks appropriate for our experiment design

In order to assess different aspects of task-adequacy described in Section 9.4.5. the task has been structured into separate subtasks, As previously outlined (Section 9.7.1), this experiment phase is comprised of following subtasks:

- two ill-defined information seeking tasks in the unfamiliar community domain, requiring users to identify relevant documents and concepts for:
  - a broad ill-defined information need in the unfamiliar community domain (Task 2.1) and
  - a specific information need in the unfamiliar community domain (Task 2.4)
- two information classification tasks requiring users to retrieve relevant documents from an unfamiliar community domain related to a specific concept from own community, which:
  - is not used in the unfamiliar community (Task 2.2) and
  - is used in the unfamiliar community but in different ways (Task 2.3)

The exact definitions of individual tasks are given in Table 9-24.

Task 2.1	Find relevant documents for the topic "Interactive systems for digital culture" and group them into subtopics. Organize your search results in a personal map. <u>Note:</u> You can use the following questions for guidance in solving the task: 1. What is understood in the netzspannung.org community under the concept "Digital Culture"? 2. What topics in netzspannung.org are related to the topic "Interactive Systems"? 3. What kinds of interactive systems are distinguished in netzspannung.org? 4. What topics and kinds of interactive systems can be related to the topic "Digital Culture"?
Task 2.2	Find documents in netzspannung.org related to the HCI topic "Multimodal Interaction". Organize documents in thematic subgroups in a personal map.
Task 2.3	Find documents in netzspannung.org related to the HCI topic "Mixed Reality". Organize documents in thematic subgroups in a personal map.
Task 2.4	As a focus of your report now take the topic "Interactive Systems in Public Space". Find relevant documents and group them into subtopics in a personal map.

Table 9-24. Definition of tasks for the test phase of Experiment II.

Accordingly, the design of tasks 2.1 and 2.4 focuses on the extent to which the users are able to gain an understanding of knowledge from an unfamiliar community by identifying relevant documents and concepts from that community and by using those concepts to characterize groups of related documents. The extent to which users are able to establish relationships between unfamiliar knowledge and familiar concepts is then examined separately in tasks 2.2 and 2.3, designed explicitly for that purpose.

Compiling a suitable test collection

The test collection for this experiment has been compiled by selecting all English entries from the real community space of netzspannung.org[71]. No other restrictions have been applied. Since the information space of netzspannung is largely created collaboratively by the community members themselves (extended by some editorial sections), the used test collection provided a very realistic test setting. The specific topics for the tasks have then been chosen based on the analysis of the available topics and corresponding distribution of documents in the collection. In particular, it has been verified that the selected topics have an intersection with the HCI community field.

Constructing reference solutions

The reference solution for Task 2.1 has been constructed based on the personal maps created by members of the netzspannung.org community in Experiment I. The results of this experiment can be considered as a process of collaborative indexing of documents in the netzspannung.org community space, since documents are assigned to concepts describing their semantic context by different community members.

Relevant topics for the themes “digital culture” and “interactive systems”, respectively, have been selected from the clusters created by the netzspannung users, in a discussion with several experienced experts from the netzspannung.org community. From the resulting set of documents those related to one but not to the other concept have been filtered out. In this way, the set of relevant concepts and the associated set of documents have been compiled. Documents that have not been indexed by any map have been additionally examined based on their contents.

As discussed in Section 9.4.2.2, the reference solution for topical structuring must include the points of view of both communities and the relevant importance of one view with respect to the other. Accordingly, we proceed as follows. We define a set of all relevant topics and documents from the viewpoint of the unfamiliar community in whose information space the tasks are accomplished. To this end, the expert members of the netzspannung community are consulted in the way described in Section 9.7.3.1. We then consult a HCI expert well familiar with the netzspannung information space to identify related topics in the HCI domain and expand the reference solution accordingly.

For our experimental design a more extensive use of concepts from the unfamiliar community will signify a higher level of understanding of the relevant knowledge from the unfamiliar community space. The ill-defined information need is located largely in the unfamiliar community space and can be appropriately described only by using the concepts from that community. The establishment of relationships to familiar concepts can occur only after this understanding has been developed and there very few concepts from own community will lend themselves to structuring the relevant information.

The resulting set of documents, concepts and document-concept relationships reflects an “objective solution” from the netzspannung community perspective, as it reflects the shared vocabulary used by the community with respect to the subject of the task. Objective reference solutions for Tasks 2.2 and 2.3 have been constructed by asking one HCI researcher to examine abstracts of all documents in the test collection and choose the relevant documents from the HCI point of view. The reference solution for Task 2.4 has been compiled in the same way as for Task 2.1.

[71] The original information space of netzspannung.org is bilingual containing both documents in English and in German.

9.7.4. Course of the Experiment

The experiment was held in December 2005 in the Usability Lab at the Institute for Computer Science and Interactive Systems of the University of Duisburg-Essen. A total of 17 participants took part in the experiment, divided in three sessions over two days (December 7 and December 8). In each session one test group completed the experiment. The planned size of 6 participants for each test group could be realized for two test groups using the Knowledge Explorer system. There was one drop-out in the reference system test group, which was accordingly composed out of 5 test subjects.

The sessions were accompanied by one supervisor who controlled the course of the experiment, answered questions regarding the general understanding of tasks and recorded informal feedback given by the participants during the course of solving the tasks. The total duration of each session was ca. 3h. This includes a short introduction and presentation of the system (ca. 25min), a try-out of the system by the test users (10min), performing the actual tasks (2h) and compiling the questionnaires (ca. 30min). The exact course of a session is given in Table 9-25.

The following data was collected during the experiment:

- the task solutions (recorded by the system),
- the questionnaires compiled by the test-subjects,
- informal feedback of the test subjects (e.g. adopted information seeking strategies).
- informal feedback from post-experiment group discussion.

COURSE OF AN EXPERIMENTAL SESSION
• Introduction (10 min): – Explaining the purpose & course of the session – Brief outline of the tasks to be solved • Compiling the pre-testing questionnaire (5 min) • Demonstration of the system configuration for the initialization phase (15 min) • Free exercise for trying out the system (10 min)
Initialization phase: Experiment IIa • Performing Task 1 (25min): – Task 1.1 (10 min) – Task 1.2 (10 min) – Task 1.3 (15 min)
• Short break & refreshment (10 min) • Demonstration of the system configuration for the test phase (10 min)
Test phase: Experiment IIb • Performing Task 2 (1h 25min): – Task 2.1 (40 min) • Compiling questionnaire 1 - knowledge construction effects (10 min) – Task 2.2 (10 min) – Task 2.3 (10 min) – Task 2.4 (15 min) • Compiling questionnaire 2 - knowledge construction effects (10 min)
• Compiling questionnaire 3 - user-perceived usefulness & satisfaction (15 min) • Free feedback & group discussion (15 min)
Time total: 3h 20min

Table 9-25. Course of an experimental session for Experiment II

9.7.5. Characterization of the Test Groups

At the start of the experiment socio-demographic data on the test subjects comprising the test groups were collected by means of a pre-testing questionnaire. The data collected include general information such as gender, professional background and topics of interests. This has been complemented with a more specific information such as technical skills in working with computers and the familiarity with background knowledge relevant for community domains treated by the experiment design.

Out of 18 persons recruited for the experiment, there was 1 drop out (in group C). From the remaining 17 users who participated in the experiment, 1 user didn't possess the minimal necessary background knowledge in the field of interactive systems for accomplishing the tasks (in group A). Accordingly the final test user population contained 16 participants: 5 in group A, 6 in group B and 5 in group C.

The general characterization of the test subjects shows a strong tendency of dominance of male to female participants in the overall test population (81% male, 19% female participants) and an unequal gender distribution across groups: Group A - 100% male participants, Group B – 50% male and 50% female participants and Group C – 80% male and 20% female participants. Such gender polarisation could not be circumvented for two reasons.

First, it reflects the general distribution of students and researchers at the Dept. of Computer Science and Dept. of Cooperative Systems. And second, in order to avoid possible bias in experiment results based on large differences in background knowledge between the groups it was important to maintain an even distribution of users between groups with respect to their degree of familiarity with the field "Interactive Systems". As a result, there was no feasible composition of the groups that would satisfy both the criteria of the equal distribution of background capabilities and gender.

Regarding technical skills, 16 test subjects declared themselves as "professional users" of personal computers and 1 as "advanced user". Thus, all test users had a high-degree of technical skills which is expected for a HCI community sample. The most important data collected includes the capabilities of test subjects regarding background knowledge that may strongly influence the experiment results. To this end the users have been asked to rate their capabilities on following indicators:

- the degree of familiarity with the community domain to which users have been assigned (HCI)
- the degree of familiarity with the supposedly unfamiliar domain (media art and design)
- the degree of familiarity with the information space of netzspannung.org.

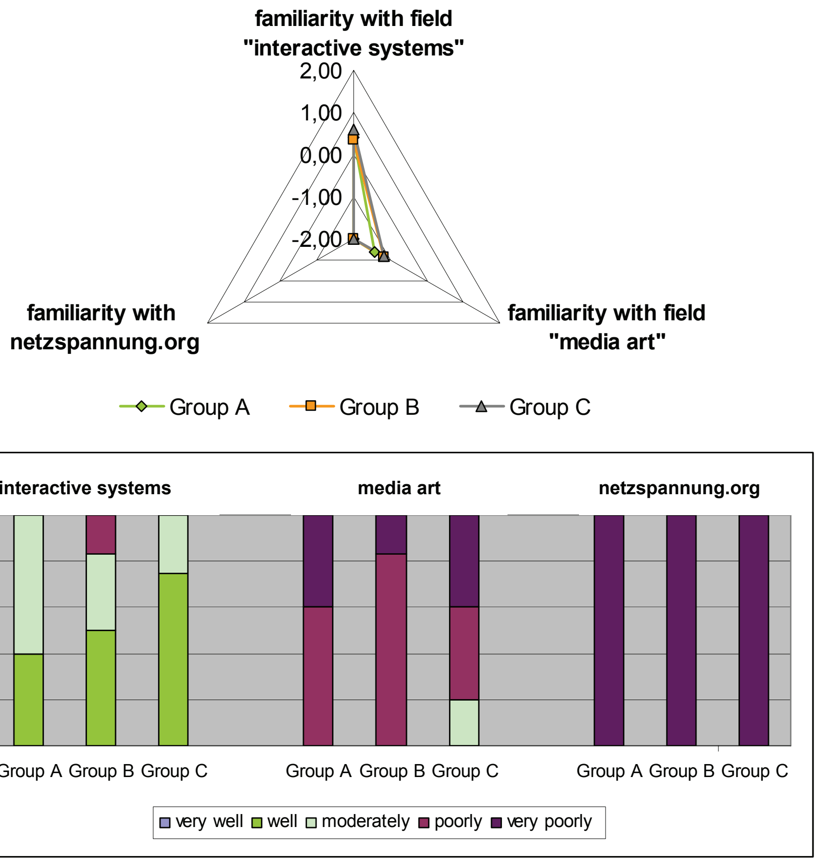


Fig. 9-23. Graphical characterization of test groups regarding background knowledge

Since the test groups simulate a sample of HCI community members a fair level of familiarity with the field of interactive systems is required for a realistic setting. On the contrary, a low level of familiarity with the field of media art (preferably none) is required. Finally, since the main tasks include access to an unfamiliar community space it is essential that the test subjects are not at all familiar with the test collection of netzspannung.org.

The test groups were composed in such a way that the degree of familiarity of individual users with the field “Interactive Systems” was approximately equally distributed in all groups. The results of the socio-demographic questionnaire confirm the suitability of the selected test subjects and their distribution in individual test groups. The results for each capability and test group are displayed graphically in Fig. 9-23. The histograms of individual groups show intra-group tendencies. The network diagram of averaged results presents the group capabilities in relation to each other.

The collected data show that the subjects in all test groups exhibit a good to at least moderate level of familiarity with the field of interactive systems, a low to very low level of familiarity with the field of media art and no familiarity with the information space of netzspannung.org. This confirms the suitability of the test subject population and their group distribution for our experimental design.

9.7.6. Task Results: Quality of Knowledge Access

To present and interpret the task results we use methods of descriptive statistics and exploratory data analysis (Polasek, 1994). The results of each task quality measure are described by the mean and median values followed by the standard deviation, the lower (LQ) and upper quartile (UQ) and the inter-quartile range of individual user solutions.

The median has been chosen as a more robust measure of central tendency the than mean, as it is less influenced by extreme values in the small sample size (Polasek, 1994). The inter-quartile range encloses 50% of all data and provides a good measure of the data spread. In this way the central tendency and the distribution of results for each test group can be assessed.

Given small samples sizes we also do not apply the statistical significance tests as any inductive statistics would be misleading. We focus on descriptive statistics and display the results from the tables graphically in form of boxplot-and-whisker diagrams which allow easy comparison of results between individual groups.

For drawing the boxplots we take the median and the inter-quartile ranges extended with the minimum and maximum values for each data series. Such boxplots allow us to quickly identify the central tendency (median), the spread of data (interquartile range) and distribution skew of each data series.

This enables us to visually compare the results of different test user samples between each other in order to determine whether we can safely assume that the observed difference is not due to chance as well as to determine which group exhibited better results. In particular, the absence of overlap between the inter-quartile ranges and the direction of the distribution skew allow us to establish whether the difference between results of two different user groups is “significant” in the sense that it is not due to chance (for the observed samples).

The degree of difference between median values can then be used to assess to degree to which one group performed better than the other. Accordingly, we use the term “significant difference” as an explanative characterization (ibid., p. 4) of the observed difference between the different test groups rather than expression of statistical generalization[72].

[72] The use of the notion of “significance” in this context must not be confused with its canonical meaning in inductive statistics where it is used to indicate the generalization validity of results beyond the test sample.

9.7.6.1. Results of Task 2.1: Ill-defined Task with Ill-defined Information Need in Unfamiliar Community Domain

Table 9-26 to Table 9-28 present the results of the task 2.1 with respect to the measures of quality of task solution introduced in Section 9.4.2. The document retrieval effect is given in Table 9-26, the quality of topical structuring in Table 9-27 and the measurement of learning effect in Table 9-28. For each of these three indicators the results of individual measures used for assessing the corresponding quality of task solution are presented. For most important measures, the best median and mean are displayed in boldface. This shows which group obtained the best result by a given measure. The corresponding cells are shaded with different levels of grey: the best value with the lightest grey, the worst with the darkest. Equal values are shaded with equal grey levels (this principle is also applied to all remaining tables).

Document retrieval effect

	Document retrieval effect											
	Group A (Community Maps)				Group B (Personal Maps)				Group C (Google + Mozilla)			
Measure	**median**	μ	σ	**IUQ**	**median**	μ	σ	**IUQ**	**median**	μ	σ	**IUQ**
document precision	**0,94**	0,94	0,04	0,07	**1,00**	0,98	0,03	0,04	**0,81**	0,82	0,07	0,06
document recall	**0,10**	0,12	0,04	0,01	**0,09**	0,12	0,06	0,02	**0,09**	0,10	0,02	0,02
Total Doc = 435, Ref. Solution Doc = 145												

Table 9-26. Descriptive statistics for document retrieval effect in task 2.1

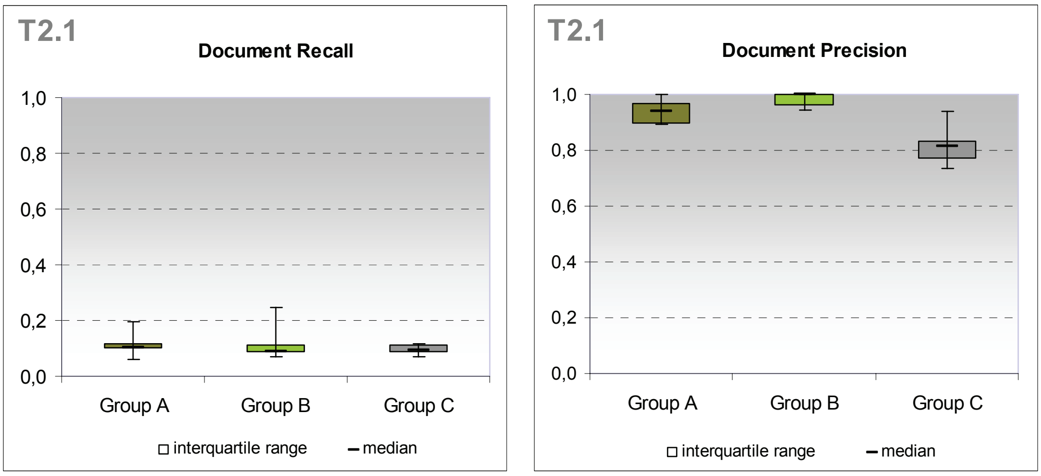


Fig. 9-24. Boxplot of document retrieval effect for task 2.1

The boxplot in Fig. 9-24 (left) suggests that for document recall no significant differences can be observed between different groups. With respect to document precision (Fig. 9-24, right) there are significant differences in result distributions between the groups A and B on one hand and the group C on the other. Both group A and group B have a significantly higher median than group C and their inter-quartile ranges do not intersect with the range of group C.

The maximum range of group C falls within the inter-quartile range of group A and is barely below the minimum of group B (see also Table 9-26). At the same time, the distribution of group C is skewed towards the lower quartile. Such distribution indicates that the document precision results of group C differ significantly from both group A and B. This difference allows us to asses the quality of results between the groups A and B on one hand and group C on the other, since given same recall, higher precision signifies a better result.

There is also a significant difference between the distribution of document precision results between group A and B though less than for group C. Group A has both a higher median and a higher mean than group A (1.00 vs. 0.94 and 0.98 vs. 0.94 respectively) and the inter-quartile ranges are non-overlapping but with with very close limits (UQ of group A and LQ of group B are almost coinciding). There seems to be a tendency towards the higher median for group B as the graph suggests a concentration of the data distribution in the opposite direction of group A. Group B shows the most homogeneous spread of all groups while the distribution of group A is more dispersed and skewed towards the lower quartile.

Thus, we can observe that while there are no significant differences with respect to document recall (Fig. 9-24, left) both groups using knowledge maps (group A, group B) achieved significantly higher precision than the group using the standard information seeking system without knowledge visualisation (group C).

This implies that with respect to document retrieval effect both groups using knowledge maps (Group A, B) achieved better results than the group using a standard information seeking system without knowledge visualisation (Group C).

With respect to the comparison of document retrieval effect between Group A and group B the results suggest a possible tendency toward better performance for group B given by higher precision at same recall and homogeneous results with smaller spread than group A.

Quality of topical structuring

The results of measures of the quality of topical structuring are given in Table 9-27 and Fig. 9-25 to Fig. 9-27. A first indicator of the quality of topical structuring is the number of topics identified and used by the test persons to structure relevant information for the task.

The small boxplot in Fig. 9-25 (upper right) depicts the distribution of the number of relevant topics across groups. It indicates that users from group B created significantly more topics from users in group C (the interquartile range being 3-5 vs. 0-3). The results for group A also display a significantly higher median than group C (4 vs. 2) but the inter-quartile range is partially overlapping.

However, the difference between the medians and the highly skewed distribution of group C toward the lower quartile (equalling zero) suggests a significant difference with respect to the number of relevant topics between group A and group C as well.

With respect to the overall number of topics group A and group B exhibit a similar behaviour, having the same median and upper quartile. A significantly smaller spread (higher lower quartile) of group B points to more homogeneous data and a tendency to a higher number of topics.

	Topical structuring effect											
	Group A (Community Maps)				Group B (Personal Maps)				Group C (Google + Mozilla)			
Measure	median	μ	σ	IUQ	median	μ	σ	IUQ	median	μ	σ	IUQ
nr. topics	4,00	4,20	1,17	2,00	5,00	5,17	1,95	3,50	3,00	2,20	1,94	3,00
nr. relevant topics	4,00	3,80	1,60	3,00	**4,00**	**4,50**	1,80	2,00	2,00	1,60	1,36	3,00
cross-community topics	2,00	2,20	1,47	1,00	**3,00**	**3,17**	1,07	1,50	1,00	0,80	0,75	1,00
intra-community topics	**1,00**	**1,40**	0,49	1,00	1,00	0,67	0,47	0,75	0,00	0,40	0,49	1,00
shared topics	0,00	0,60	0,80	1,00	0,00	0,33	0,47	0,75	0,00	0,00	0,00	0,00
non-topical concepts	0,00	0,00	0,00	0,00	0,00	0,50	0,76	0,75	0,00	0,40	0,80	0,00
topical recall	**0,10**	**0,10**	0,04	0,08	**0,10**	**0,11**	0,05	0,05	0,05	0,04	0,03	0,08
topical precision	**1,00**	**0,87**	0,16	0,33	0,92	0,88	0,12	0,23	0,40	0,48	0,45	1,00
cross-community	0,50	0,51	0,23	0,33	**0,67**	**0,67**	0,21	0,09	0,20	0,24	0,25	0,33
intra-community	**0,33**	**0,35**	0,11	0,07	0,13	0,13	0,11	0,13	0,00	0,11	0,14	0,20
shared	0,00	0,13	0,17	0,25	0,00	0,05	0,07	0,09	0,00	0,00	0,00	0,00
cumulative topical coverage												
overall	0,49				**0,57**				0,09			
cross-community	0,35				**0,62**				0,08			

Table 9-27. Descriptive statistics for the quality of topical structuring in task 2.1.

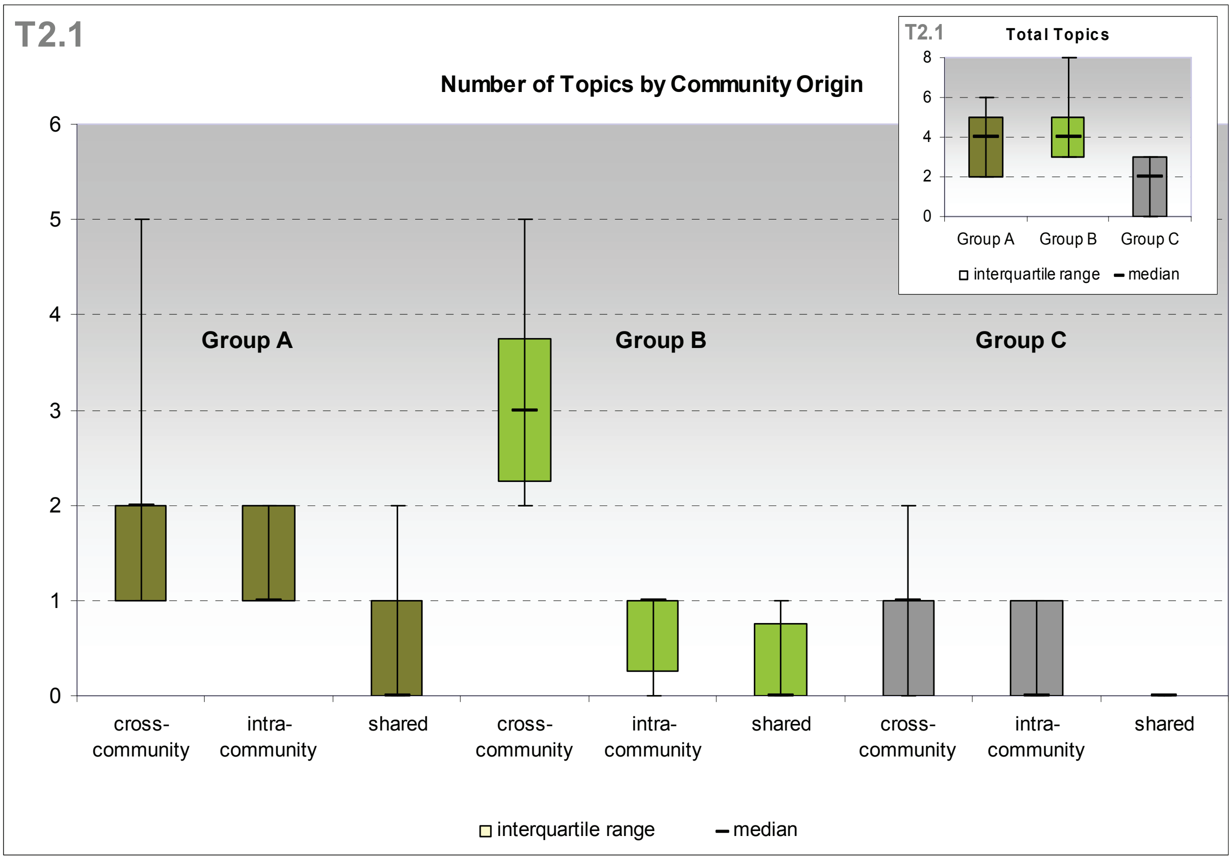


Fig. 9-25. Boxplot of the distribution of topics by community of origin per test group for task 2.1.

More significant differences in the quality of topical structuring between the three groups can be observed with respect to the community of origin of the topics as indicated by the main boxplot in Fig. 9-25.

Group B displays by far the highest amount of cross-community topics used in the task solution, the median is higher both from group C and group A (3 vs. 2 vs. 1), the inter-quartile range is non-overlapping and well above both other groups. The use of intra-community topics is very low in both groups ranging between no topics at all and 1 topic at most: in group A there is a tendency of at least 1 intra-community topic per user (median=1, distribution skewed towards upper quartile) with only rare use of shared topics (median=0, UQ=0.75) .

Group C exhibits poor quality of structuring in general: it shows only occasional occurrence of both cross-community topics (median=1, distribution skewed towards no topics at all) and intra-community topics (median=0, UQ=1) with no occurrence of shared topics. Group A exhibits a tendency of slightly higher use of cross-community topics (median=2) than intra-community topics (median=1) within the group, but this is relativized by equal inter-quartile ranges of the two distribution.

Overall, the boxplot in Fig. 9-25. suggests significant differences in structuring behaviour between groups with respect to topic origin. Group C shows the highest incidence of cross-community topics, group B the highest incidence of intra-community topics while group A shows the poorest structuring behaviour for both cross-community and intra-community topics. Shared topics were used little by all groups.

The measures of topical recall and topical precision further confirm the above findings. As expected, topical recall is low in general (Fig. 9-26, left). One reason for the low recall is the limited timeframe for accomplishing a task in an unfamiliar domain. The other is the relatively high number of correct reference topics (N=35) and the nature of the reference solution for this task: as discussed in Section 9.4.2.2, in this case the reference solution is not *the* unique correct topical structure but merely a set of relevant topics out of which different topical combinations are possible and equally valid.

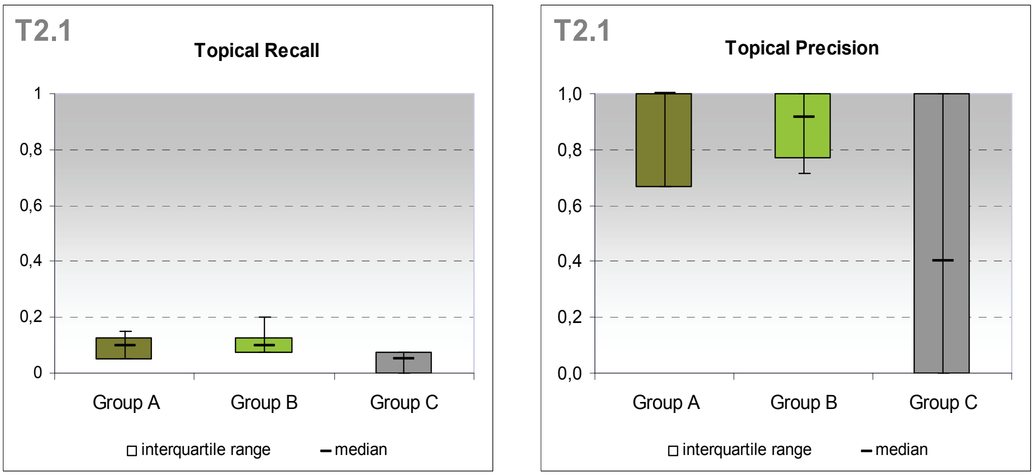


Fig. 9-26. Boxplot of topical recall and topical precision in task 2.1.

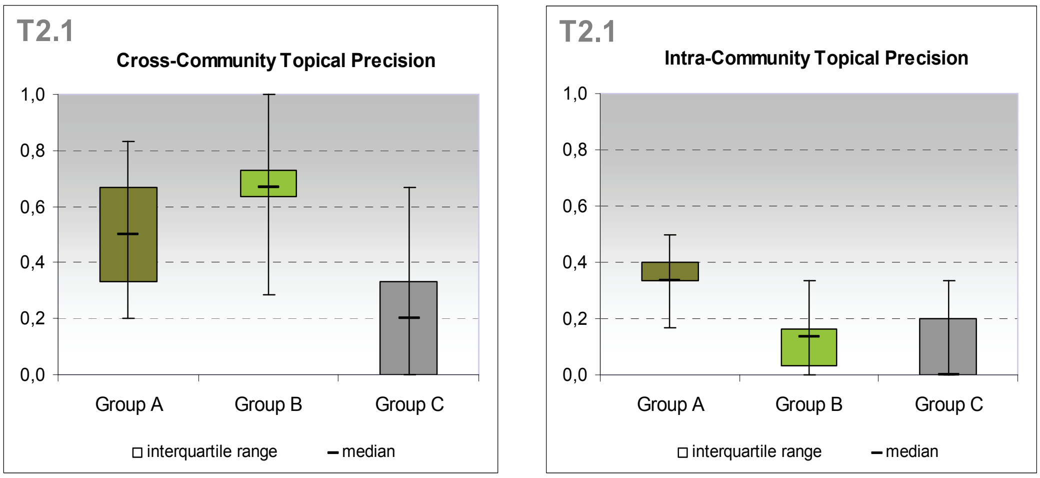


Fig. 9-27. Boxplot of cross-community and intra-community topical precision in task 2.1.

Hence, the measure of topical recall cannot be considered primarily with respect to its absolute value but as a comparative measure between the groups themselves. The boxplot in Fig. 9-26 (left) indicates significant difference in topical recall between group B and group C (higher median of group B and non-overlapping inter-quartile range) and a tendency of higher recall for group B than group C (higher median, partially overlapping inter-quartile range but distribution of group C skewed towards zero).

With respect to topical precision (Fig. 9-26, right) both group A and group B display significantly different behaviour than group C. The values for group A and B are grouped in the upper quarter of the precision scale (0.67-1.00 and 0.77-1.00, respectively) while results of group C display a completely random distribution over the whole scale. The latter is due to the absence of any topical structuring for several test users in group C (Fig. 9-25). In a small sample such as ours, this significantly influences precision which ranges from 0 to 1 with a median roughly in the middle.

It also confirms our observation of significant difficulties in identifying relevant topics experienced by test persons from group C. In contrast, group A and group B exhibit roughly similar behaviour with largely overlapping ranges a slightly higher median for group A (1.00 vs. 0.92, Table 9-27). The spread of group B is significantly smaller suggesting more homogeneous results. Thus, no significant difference with respect to overall topical precision can be identified between group A and group B.

With respect to the origin of topics, both group A and group B exhibit significantly different distribution of cross-community topical precision than group C, with higher median values, as shown in Fig. 9-27 (left).

The boxplot also suggests a significant difference between group A and group B: though inter-quartile ranges are slightly overlapping, group B exhibits a higher median and a much more homogenous spread, skewed towards the upper quartile and farther away from group A.

With respect to intra-community topical precision group A displays significantly different and better result than group B and group C (Fig. 9-27, right), whereas distributions of group B and group C show no significant differences. In our case topical precision reflects the proportion of topics from a particular community of origin with respect to all topics used in task solutions of a given group. Thus, higher cross-community topical precision indicates a greater extent of the (correct) use of cross-community topics with respect to intra-community and shared topics, which is the most important criteria for the quality of topical structuring in task 2.1 (Section 9.4.2.2, 9.7.3.2).

Finally, the cumulative topical coverage is significantly higher by both groups A and B than group C, with group B exhibiting both the highest overall coverage and the highest coverage of cross-community topics. The absolute level of cross-community coverage is also high with 57% overall and 62% of cross-community topics covered by all individual solutions with respect to the reference solution. In contrast, group C exhibit only 9% overall and 8% cross-community topical coverage. This indicates that users without knowledge map support not only found fewer cross-community topics individually, but also that their solutions tended to be more homogeneous with respect to topical structures created.

The high coverage in case of knowledge map supported groups indicates that users were able not only to identify relevant cross-community topics for the task, but also to select those more appropriate from a personal point of view. Given a relatively large number of topics in the reference solution due to the highly ill-defined nature of the task, this suggests a very positive result.

Overall, we can observe that there are significant differences between groups A and B, on one hand and group C on the other hand, with respect to both topical recall and topical precision. Both groups using knowledge maps (group A, group B) achieved significantly higher overall topical recall and precision than the group using the standard information seeking system without knowledge visualisation (group C).

Furthermore, the results of topical distribution by community of origin (Fig. 9-25), the analysis of corresponding components of topical precision (Fig. 9-27) and cumulative topical coverage demonstrate a significantly higher use of cross-community topics by group B than by group A, with both groups also performing better than group C.

This implies that with respect to the quality of topical structuring both groups using knowledge maps (Group A, B) achieved better results than the group using a standard information seeking system without knowledge visualisation (Group C). Furthermore, group B (personal maps) achieved significantly better results than group A (community maps) with respect to extent of identifying and using cross-community topics for structuring the task solution.

Learning effect

Table 9-28 summarizes the learning effect indicators based on the results of the post-task questionnaire (Section 9.4.2.3). Boxplots in Fig. 9-28 - Fig. 9-31 visualise the distribution of results of the topical recall and topical precision measure for each question.

Fig. 9-28 depicts the topical recall and topical precision for answers to the question Q1 concerning the acquired level of understanding of the unfamiliar concept from the unfamiliar community ("digital culture"). The graph clearly shows significant difference between group A and group B with respect to group C for topical recall and even more for topical precision.

The medians of both group A and group B are higher than for group C and the interquartile ranges are non-overlapping in both cases. Furthemore the distribution of results for group C is skewed toward the lower quartile in such a way that 50% of results lie below the median.

The difference in topical recall between group A and group B is not so clear: even though group B exhibits a higher median, the interquartile ranges overlap. In contrast, the topical distribution results differ to a great extent.

This suggests that while topical recall for Q1 may be comparable for group A and B, there is a much better performance with respect to topical precision for group B. Topical recall corresponds to the proportion of correctly named topics with respect to the reference solution for this question. Topical precision provides insight into the proportion of correctly named topics within all topics named by the users.

The above analysis thus implies that both group A and group B acquired a significantly better understanding of the unfamiliar concept from the unfamiliar community than group C. In addition, there is a noticeable tendency for group B to have acquired better understanding than group A, due to greatly better topical precision at comparable recall.

		Learning effect											
		Group A (Community Maps)				**Group B** (Personal Maps)				**Group C** (Google + Mozilla)			
		median	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**
Q1: digital culture in netzspannung	**tr**	0,12	0,14	0,05	0,04	**0,19**	0,18	0,07	0,11	0,12	0,10	0,04	0,04
	tp	0,75	0,74	0,16	0,13	**1,00**	0,98	0,05	0,00	0,50	0,45	0,17	0,35
Q2: interactive systems in netzspannung	**tr**	0,07	0,09	0,10	0,07	**0,25**	0,21	0,14	0,23	0,21	0,20	0,07	0,00
	tp	0,25	0,41	0,42	0,80	**0,61**	0,58	0,32	0,36	0,50	0,55	0,13	0,24
Q3: interactive systems for digital culture	**tr**	0,10	0,10	0,06	0,00	**0,30**	0,25	0,19	0,30	0,10	0,08	0,07	0,10
	tp	**1,00**	0,80	0,40	0,00	0,90	0,63	0,45	0,80	0,50	0,30	0,24	0,50
Q4: related CHI topics	**tr**	0,14	0,10	0,06	0,07	0,14	0,17	0,15	0,21	**0,21**	0,14	0,09	0,14
	tp	**1,00**	0,70	0,40	0,50	0,75	0,58	0,45	0,88	0,50	0,55	0,33	0,25

Q1 - understanding an unfamiliar community concept
Q2 - relating familar concept to unfamiliar community domain
Q3 - integrating knowledge accross communities
Q4 - relating unfamilar concept to familiar community domain
tr = topical recall, tp = topical precision

Table 9-28. Descriptive statistics for the learning effect in task 2.1

The results for question Q2 concerning the acquired level of understanding of relationships between the familiar concept ("interactive systems") and the unfamiliar community domain ("netzspannung.org") are visualised in Fig. 9-29.

Group C exhibits significantly better performance for this question than for the previuous one: topical recall is comparable with group B (similar medians and inter-quartile range) and significantly better than group A. Topical precision is also comparable to group B: the median of group C is notably lower but the interquartile ranges partially overlap in the same direction. The inter-quartile range of topical precision of group A is spread almost across the whole scale which suggests highly variable results between individual users.

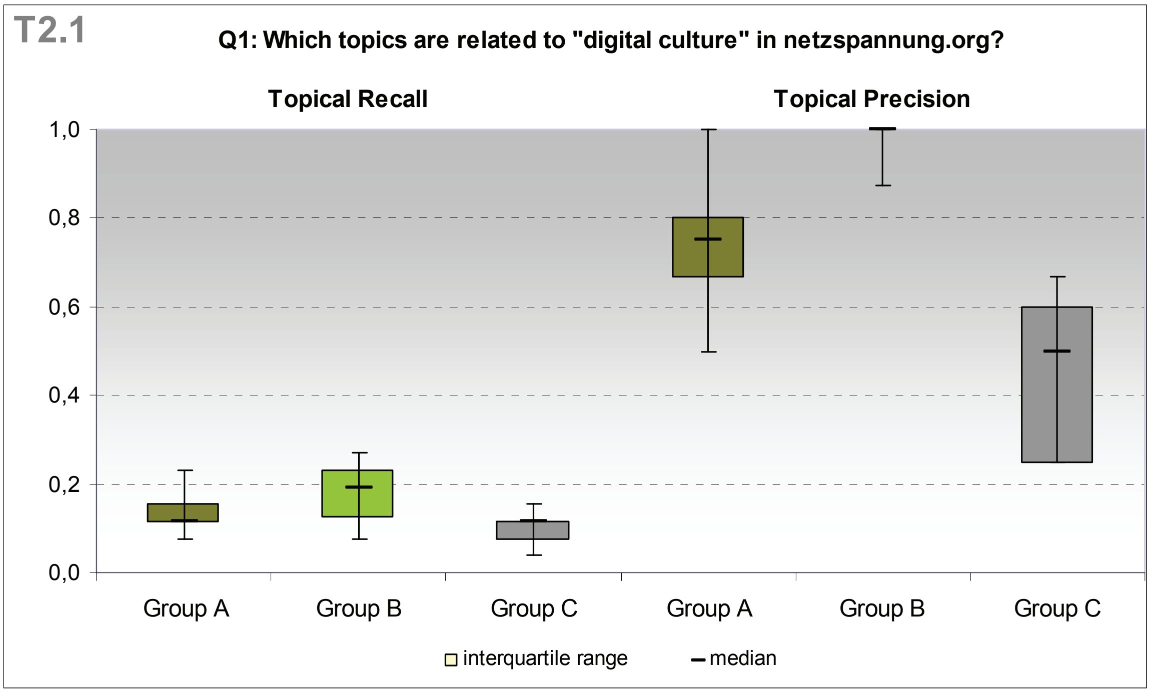


Fig. 9-28. Boxplot of topical recall and topical precision concerning the acquired level of understanding of the unfamiliar concept ("digital culture") from the unfamiliar community ("netzspannung.org") in task 2.1.

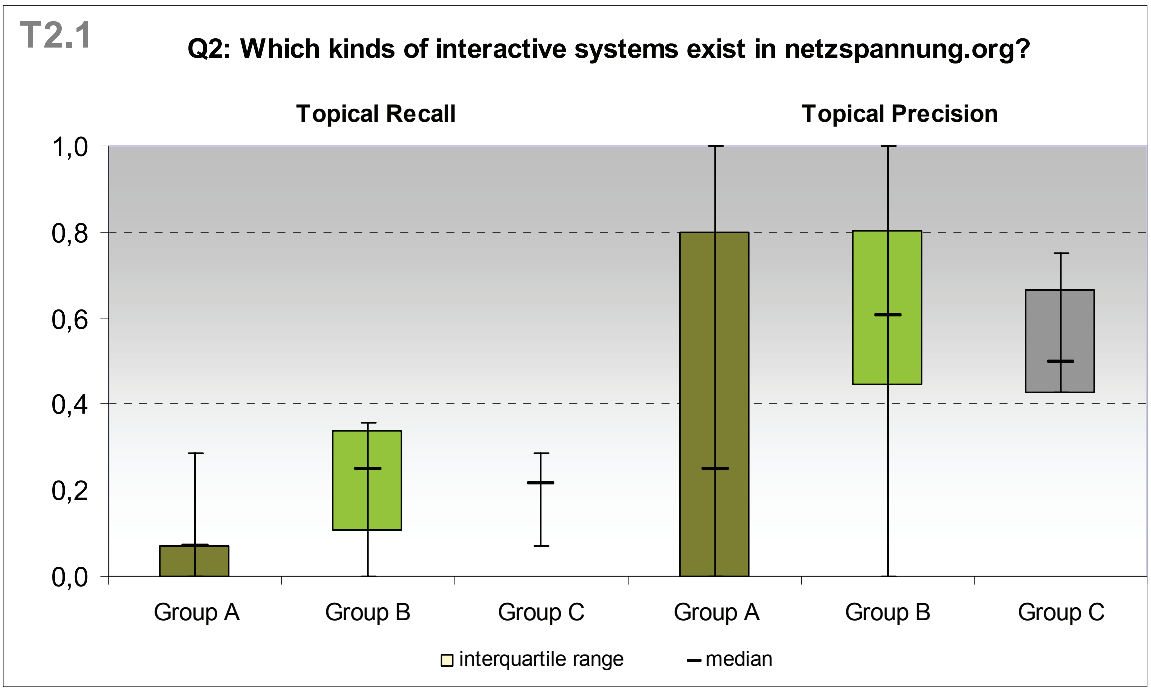


Fig. 9-29. Boxplot of topical recall and topical precision concerning the acquired level of understanding of relationships between the familiar concept ("interactive systems") and the unfamiliar community domain ("netzspannung.org") in task 2.1.

The boxplot in Fig. 9-30 graphically presents the results for question Q3 concerning the acquired level of knowledge integration between unfamiliar and familiar concepts ("digital culture" and "interactive systems") in the unfamiliar community domain ("netzspannung.org").

The results of group A and group B again show significant differences to group C for both topical recall and topical precision. Group B exhibits by far the highest median for topical recall (0.30 vs. 0.10 vs. 0.10) while group A is characterized by the best topical precision: it exhibits the highest median (1.00 vs. 0.90

vs. 0.50) and the most homogeneous distribution of all groups. Although group B exhibits somewhat lower precision with a significant spread, the notable difference in recall suggests better performance of group B than group A for this question.

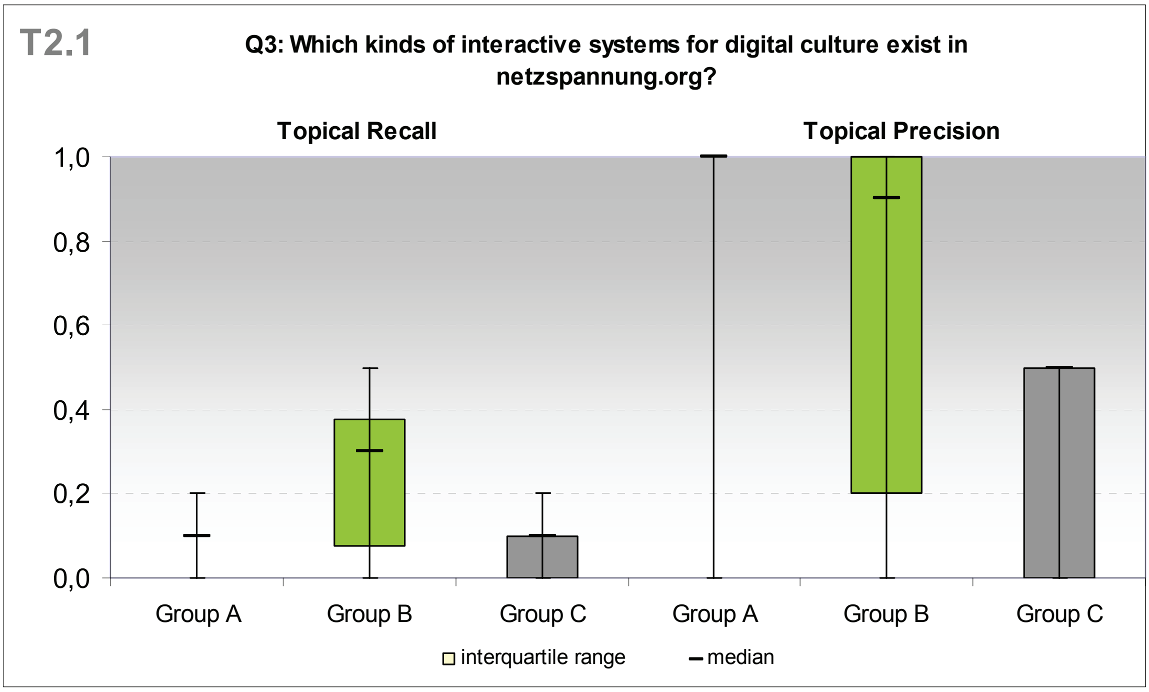


Fig. 9-30. Boxplot of topical recall and topical precision concerning the acquired level of knowledge integration between unfamiliar and familiar concepts ("digital culture" and "interactive systems") in an unfamiliar community domain ("netzspannung.org").

Overall, the results for Q3 suggest a significant difference between all three groups with respect to knowledge integration between communities: group A and group B performing significantly better than group C and group B exhibiting a tendency of better performance than group A.

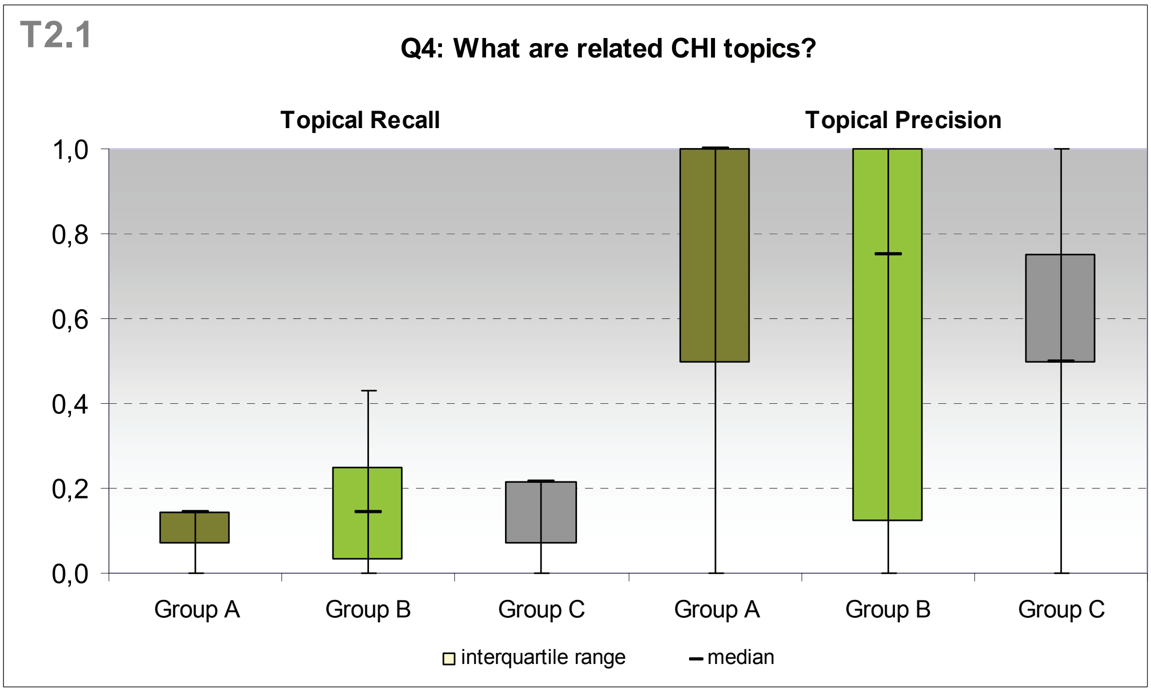


Fig. 9-31. Boxplot of topical recall and topical precision concerning the acquired level of understanding of relationships between the task in the unfamiliar community ("interactive systems for digital culture") and the own community domain ("HCI").

The results for question Q4 concerning the acquired level of understanding of relationships between the task in the unfamiliar community domain ("interactive systems for digital culture") and the own community domain ("chi") are visualised in Fig. 9-31. With respect to topical recall no significant differences between the groups can be observed. The median is slightly higher for group C but the inter-quartile ranges of all three groups largely overlap.

With respect to precision group A exhibits the highest median, second next is group B while group C exhibits lowest precision. The precision spread is most homogeneous for group C and the most variable for group B. Overall the results of this question suggest a slight tendency of better performance of group A than group B and group C.

A qualitative analysis of user answers to this question explains such results. Several users from group B that provided a low quality or no answer added an explanatory comment stating they hadn't been considering this question during the task, since focused on the task goal. On the contrary, the topical structuring of group C exhibited in task solutions shows the use of known intra-community concepts only (Fig. 9-25).

This points to an important observation: while group B focused on understanding the unfamiliar domain, group C retreated to looking for familiar concepts in order to make sense of the unfamiliar context. Similar conclusions can be drawn for the performance of group C. The better scoring on this question is in concordance with the higher use of intra-community than cross-community concepts for structuring their task solutions (Fig. 9-25). The analysis of the qualitative questionnaires in the next section sheds some more light on such important indicators of the relationship between group results, the structuring behaviour and sensemaking patterns exhibited by different groups.

9.7.6.2. Results of Tasks 2.2 and 2.3: Relating Knowledge from Unfamiliar Community to Familiar Concepts

As discussed in Section 9.7.3.2 task 2.2 and 2.3 are concerned with evaluating the extent to which users are able to establish relationships between unfamiliar knowledge and familiar concepts. To this end users have been required to identify relevant documents in the unfamiliar community information space related to a specific topic from their own community. In task 2.2. the presented topic is based on a concept not used by the other community while in 2.3 the presented concept is used in the unfamiliar community but in different ways than in its community of origin. Accordingly, the quality of task solution is measured only with respect to the document retrieval effect.

	Document retrieval effect											
	Group A (Community Maps)				Group B (Personal Maps)				Group C (Google + Mozilla)			
Measure	**median**	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**
document precision	0,67	0,78	0,18	0,33	**0,83**	0,81	0,19	0,33	0,50	0,51	0,08	0,05
document recall	0,08	0,17	0,13	0,08	**0,19**	0,21	0,07	0,07	0,19	0,20	0,06	0,04
Total Doc = 435, Ref. Solution Doc = 24												

Table 9-29. Descriptive statistics for the document retrieval effect in task 2.2.

The boxplot of document recall (Fig. 9-32, left) suggests that there are small but significant differences between group B and group C with respect to group A in task 2.2. Both medians of group B and group C are higher than the median of group A and the respective inter-quartile ranges do not overlap with the inter-quartile range of group A.

On the other hand, no significant differences can be concluded between group B and group C who have similar medians and overlapping inter-quartile ranges. Relatively low overall recall can be attributed to a very limited time (10 min) with respect to a large number of documents in the reference solution (N=24) in an unfamiliar domain.

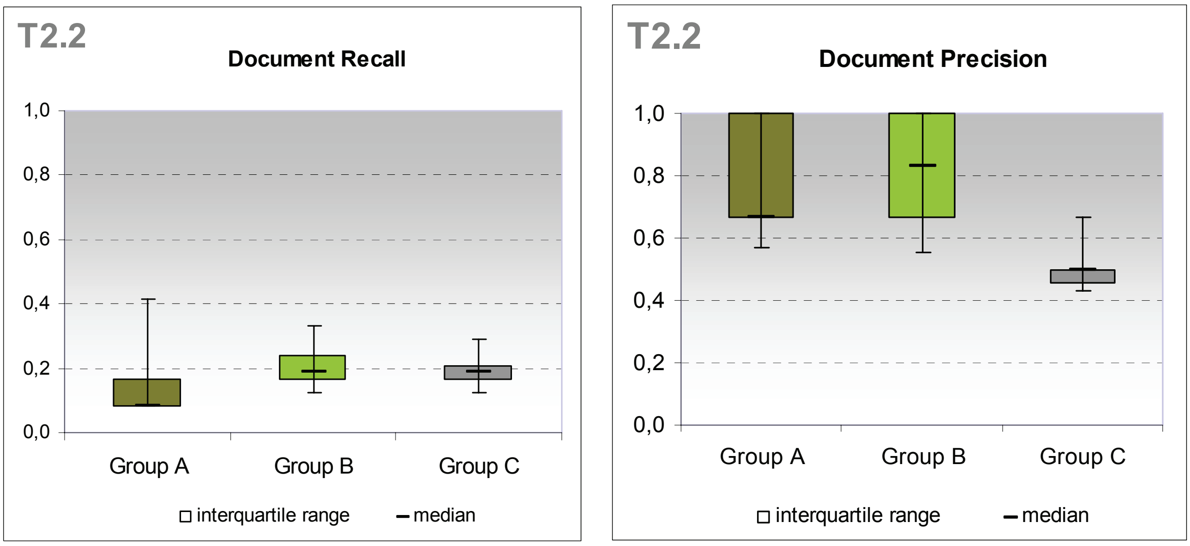


Fig. 9-32. Boxplot of document retrieval effect for task 2.2.

With respect to document precision group A and group B exhibit significantly better results than group C. Their medians are much higher (0.83 and 0.67 vs. 0.50), the inter-quartile ranges are non-overlapping and significantly higher positioned than that of group C. The distributions of document precision results of group A and group B are comparable. Group B has a significantly higher median (0.83 vs. 0.67) but the inter-quartile ranges are coinciding, with group B being skewed towards to upper end.

The relatively small overall differences in recall between all groups and the significantly better precision of group B and group A suggests an overall better performance of group B and group A compared to group C. Due to a somewhat better recall at a comparable level of precision group B exhibits a tendency to better performance than group A. Overall, such results suggest that users from group B and group A were able to discover relationships between unfamiliar knowledge and a familiar concept not used in the target community better than group C. While no such statement can be established between group B and group A, there is a likely tendency of better results for group B.

	Document retrieval effect											
	Group A (Community Maps)				Group B (Personal Maps)				Group C (Google + Mozilla)			
Measure	**median**	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**
document precision	0,80	0,78	0,16	0,28	**0,96**	0,92	0,09	0,13	0,67	0,64	0,06	0,07
document recall	0,23	0,20	0,06	0,09	**0,33**	0,30	0,07	0,05	0,26	0,31	0,12	0,14
Total Doc = 435, Ref. Solution Doc = 35												

Table 9-30. Descriptive statistics for the document retrieval effect in task 2.3.

The results of document retrieval effect for task 2.3. are given in Table 9-30 with a graphical overview in Fig. 9-33. For document recall group B exhibits the highest median with the smallest spread that suggest best performance (Fig. 9-33, left). There are significant differences between results of group B and group A given by a much lower median of group A (0.23 vs. 0.33) and relatively distant inter-quartile ranges. While group C exhibits only a slightly higher median than group A (0.26 vs. 0.23) the inter-quartile ranges are non-overlapping and the distributions are skewed in opposite directions. This suggests better document recall of group C than group A.

The difference between group B and group C is not so obvious: although group C has a notably lower median (0.26 vs. 0.33) the inter-quartile range of group B is subsumed by the inter-quartile range of group C. Group B exhibits a very small spread which gives high expression validity of the median, whereas the results of group C are spread much wider. Thus, while a significant difference cannot be concluded such results do suggest a tendency of better performance for group B than group A.

With respect to document precision, group B performed significantly better than group C: the median of group B is much higher (0.96 vs. 0.67) and the lower quartile of group B is significantly higher than the upper quartile of group C (Fig. 9-33, right). The difference between group B and group A is much smaller. Though the median of group B is much better than the median of group A (0.96 vs. 0.80) the inter-quartile ranges are somewhat overlapping (LQ_B=0.87, UQ_A=0.91) However, since the distribution of group A is skewed towards the lower quartile we can observe a tendency of better performance of group B.

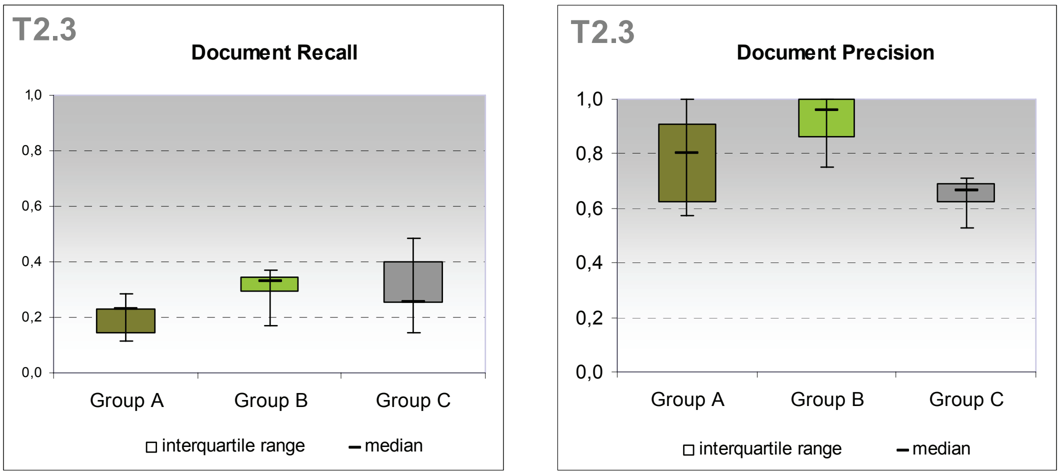


Fig. 9-33. Boxplot of document retrieval effect for task 2.3.

Finally, group A exhibits a much higher median than group C (0.80 vs. 067). The inter-quartile ranges are somewhat overlapping with both distributions skewed towards the respective lower quartiles. As the spread of group C is very little and contained in the neighbourhood of the lower quartile of group A such results imply a tendency of better performance for group A.

Overall, such results suggest that users from group B were able to discover relationships between unfamiliar knowledge and a familiar concept also used in the target community better than group A (better recall, tendency of better precision) and group C (better precision, tendency of better recall). Since group C exhibits better recall than group A but the latter shows a tendency of better precision we cannot conclude a significant overall difference in the discovery of relationships between group A and group C.

9.7.6.3. Results of Task 2.4: Ill-defined task with well-defined need

In task 2.4 users they were required to identify relevant documents and organize them into appropriate topical groups, just like in task 2.1. The only difference is in the nature of the task which is still ill-defined but with a more specific, narrower information need. Accordingly, the quality of task solutions is measured based on all three indicators just as in task 2.1: the document retrieval effect, the quality of topical structuring and the learning effect.

Document retrieval effect

	Document retrieval effect											
	Group A (Community Maps)				Group B (Personal Maps)				Group C (Google + Mozilla)			
Measure	median	μ	σ	IUQ	median	μ	σ	IUQ	median	μ	σ	IUQ
document precision	0,90	0,91	0,09	0,10	1,00	0,95	0,10	0,00	0,75	0,73	0,02	0,04
document recall	0,09	0,12	0,07	0,04	0,11	0,11	0,04	0,01	0,10	0,10	0,04	0,06
Total Doc = 435, Ref. Solution Doc = 145												

Table 9-31. Descriptive statistics for document retrieval effect in task 2.4

The boxplot of document recall (Fig. 9-34, left) for task 2.4 shows no significant differences between results of different groups. The medians are almost identical (0.10 vs. 0.11 vs. 0.9) and the inter-quartile ranges are almost completely overlapping. The only notable difference is a very homogeneous distribution for group B that has an extremely small spread and the highest median (0.11).

The document precision measure (Fig. 9-34, right) shows a more differentiated picture. Group B achieves significantly better precision than both group C and group A, with the highest median (1.00 vs. 0.90 vs. 0.75) and almost no spread. Group A achieves much better precision than group C demonstrated by a significant difference both in the median (0.90 vs. 0.75) and in the respective inter-quartile ranges.

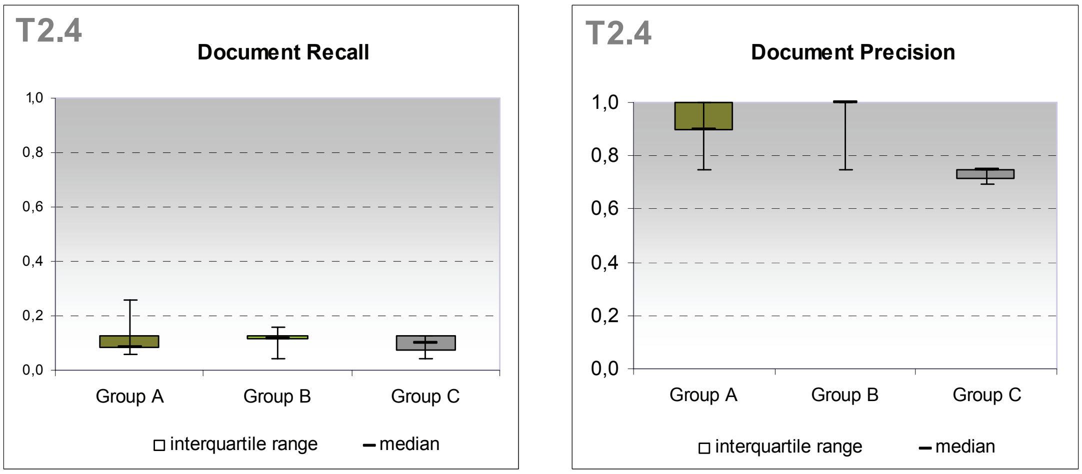


Fig. 9-34. Boxplot of document retrieval effect for task 2.4: document recall and document precision.

Given comparable recall but much better precision suggests much better document retrieval effect for group A and group B using the knowledge maps than group C using standard information seeking system without knowledge visualisation. Furthermore, due to better precision the document retrieval effect is better for group B using personal maps than for group A using community maps.

Quality of topical structuring

	Topical structuring effect											
	Group A (Community Maps)				Group B (Personal Maps)				Group C (Google + Mozilla)			
Measure	median	μ	σ	IUQ	median	μ	σ	IUQ	median	μ	σ	IUQ
nr. topics	3,00	2,20	1,47	2,00	2,00	2,40	1,36	3,00	0,00	0,00	0,00	0,00
nr. relevant topics	2,00	1,60	1,02	1,00	**2,00**	**2,40**	1,36	3,00	0,00	0,00	0,00	0,00
cross-community topics	2,00	1,40	0,80	1,00	**2,00**	**2,00**	0,89	2,00	0,00	0,00	0,00	0,00
intra-community topics	**0,50**	**0,50**	0,50	1,00	0,00	0,20	0,40	0,00	0,00	0,00	0,00	0,00
shared topics	0,00	0,00	0,00	0,00	0,00	0,20	0,40	0,00	0,00	0,00	0,00	0,00
non-topical concepts	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
topical recall	**0,07**	**0,06**	0,04	0,03	**0,07**	**0,08**	0,05	0,10	0,00	0,00	0,00	0,00
topical precision	**0,67**	**0,63**	0,37	0,50	1,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00
cross-community	0,67	0,57	0,33	0,17	**1,00**	**0,90**	0,12	0,25	0,00	0,00	0,00	0,00
intra-community	**0,00**	**0,13**	0,16	0,33	0,00	0,05	0,10	0,00	0,00	0,00	0,00	0,00
shared	0,00	0,00	0,00	0,00	0,00	0,05	0,10	0,00	0,00	0,00	0,00	0,00
cumulative topical coverage												
overall	0,35				**0,42**				0,00			
cross-community	0,41				**0,53**				0,00			

Table 9-32. Descriptive statistics for the quality of topical structuring in task 2.4.

With respect to the number of relevant topics used to structure relevant information for task 2.4 group C exhibits by far the worst results, as no topics what so ever have been created by the test persons in that group (Fig. 9-35).

This suggests that users without knowledge map support have had extreme difficulties in identifying topics relevant for an ill-defined task in an unfamiliar community domain within an very limited period of time (10min) (this is also confirmed by results of the qualitative questionnaires, discussed in Sections 9.7.7, 9.7.8).

The results of group A and group B are characterized by the same median and overlapping inter-quartile ranges but with a differently skewed distribution. Whereas the distribution of results for group A is skewed towards the lower quartile (median equals the upper quartile), the one of group B is oriented towards the upper quartile, which also significantly exceeds the upper quartile of group A (UQ_A=4, UQ_B=2). This suggests a tendency of somewhat better structuring by cross-community topics of group B.

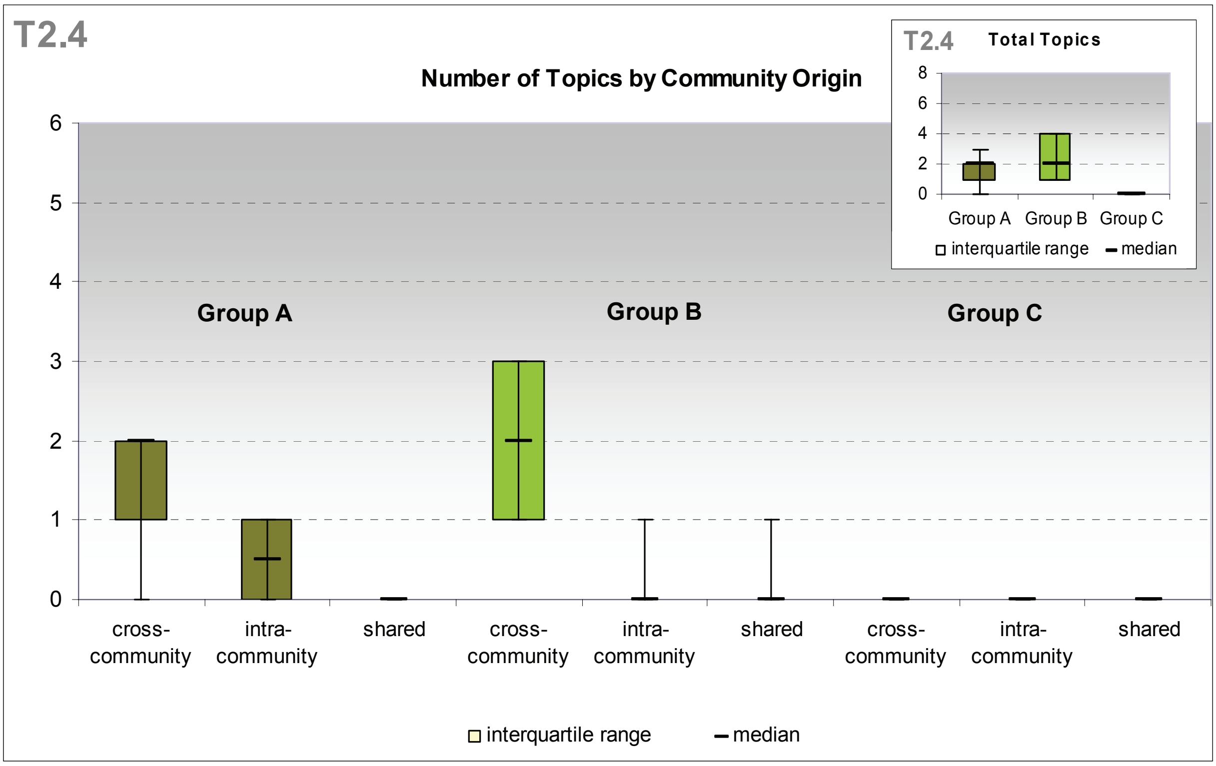


Fig. 9-35. Boxplot of the distribution of topics by community of origin per test group in task 2.4.

The measures of topical recall and topical precision further support such conclusions. While both groups have the same median of topical recall the distribution of group B is skewed towards higher values than that of group B, whose inter-quartile range is contained completely below the median level (Fig. 9-36). Group B further exhibits much better topical precision with a much higher median (1.00 vs. 0.67) and virtually no spread.

As can be seen in Fig. 9-37 (left) this is largely the consequence of the significantly higher precision for cross-community topics. This indicates a greater extent of the (correct) use of cross-community topics and thereby better performance by the most important criteria for the quality of topical structuring in this task.

The cumulative topical recall confirms these results. Both overall and cross-community topical coverage are significantly higher in group B than in group A (Table 9-32). The absolute level is also relatively high, with 42% of overall and 53% of cross-community topical coverage for group B (confirming the tendency of more extensive use of cross-community than intra-community concepts for the task solutions).

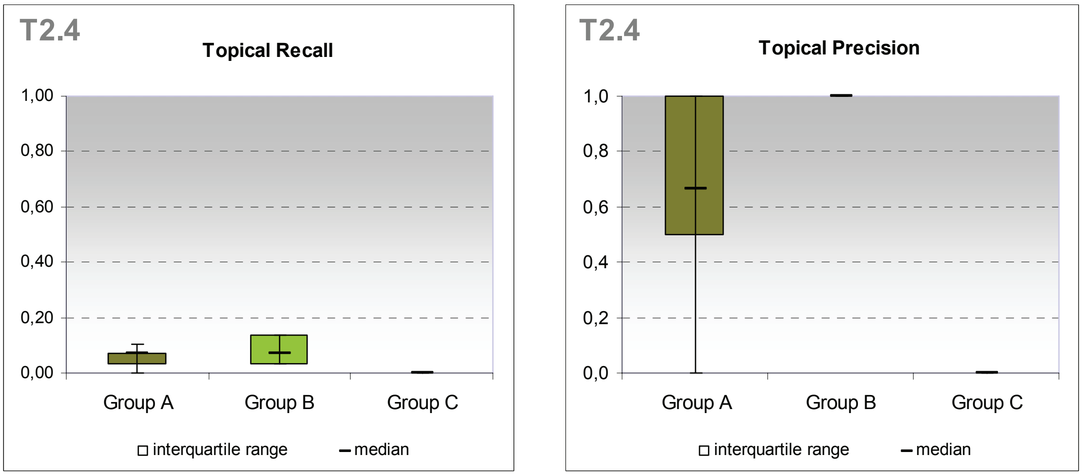


Fig. 9-36. Boxplot of topical recall and topical precision in task 2.4.

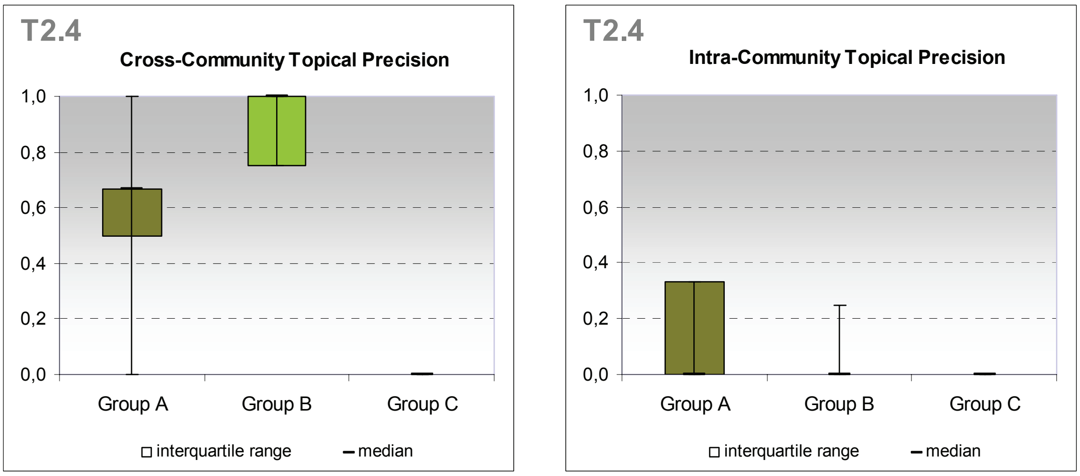


Fig. 9-37. Boxplot of cross-community and intra-community topical precision in task 2.4.

This indicates a high diversification of concepts used in individual maps, which reflects high levels of personal judgment in matching the understanding of the task information need against concepts from unfamiliar community maps.

This in turn suggests both high levels of sensemaking and the suitability of the proposed collaborative knowledge visualisation method for the discovery of relationships between different domains: a higher range of relevant topics identified by the users results in a higher range of created relationships between unfamiliar concepts and their own community knowledge, available for later use by others.

Overall, we can observe that both groups using knowledge maps (group A, group B) achieved much better results with respect to topical structuring than the group using the standard information seeking system without knowledge visualisation (group C), which produced no topics at all.

Furthermore, based on a tendency of better topical recall, significantly higher cross-community topical precision and cumulative topical coverage group B achieved better quality of topical structuring than group A.

Learning effect

Table 9-33 presents the learning effect indicators based on the results of the post-task questionnaire for task 2.4 (Section 9.4.2.3). Boxplots in Fig. 9-38-Fig. 9-40 give a graphical overview of the distribution of topical recall and topical precision measures for each question.

		Learning effect											
		Group A (Community Maps)				**Group B** (Personal Maps)				**Group C** (Google + Mozilla)			
		median	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**	**median**	**μ**	**σ**	**IUQ**
Q1: digital culture in netzspannung	**tr**	0,12	0,12	0,04	0,00	**0,18**	**0,18**	0,08	0,12	0,12	0,10	0,04	0,04
	tp	0,67	0,65	0,20	0,17	**1,00**	**1,00**	0,00	0,00	0,50	0,45	0,17	0,35
Q2: interactive systems in netzspannung	**tr**	0,10	0,12	0,07	0,10	0,20	0,20	0,06	0,00	**0,21**	**0,20**	0,07	0,00
	tp	0,50	0,50	0,33	0,33	**1,00**	**0,93**	0,13	0,00	0,50	0,55	0,13	0,24
Q3: interactive systems for digital culture	**tr**	0,13	0,15	0,09	0,13	**0,38**	**0,33**	0,19	0,25	0,10	0,08	0,07	0,10
	tp	0,33	0,40	0,39	0,67	**0,50**	**0,50**	0,45	1,00	0,50	0,30	0,24	0,50
Q4: related CHI topics	**tr**	0,14	0,17	0,11	0,14	0,14	0,17	0,11	0,14	**0,21**	0,14	0,09	0,14
	tp	0,50	0,50	0,45	1,00	0,50	0,60	0,37	0,50	**0,50**	0,55	0,33	0,25

Q1 - understanding an unfamiliar community concept
Q2 - relating familar concept to unfamiliar community domain
Q3 - integrating knowledge accross communities
Q4 - relating unfamilar concept to familiar community domain
tr = topical recall, tp = topical precision

Table 9-33. Descriptive statistiscs for the learning effect in task 2.4.

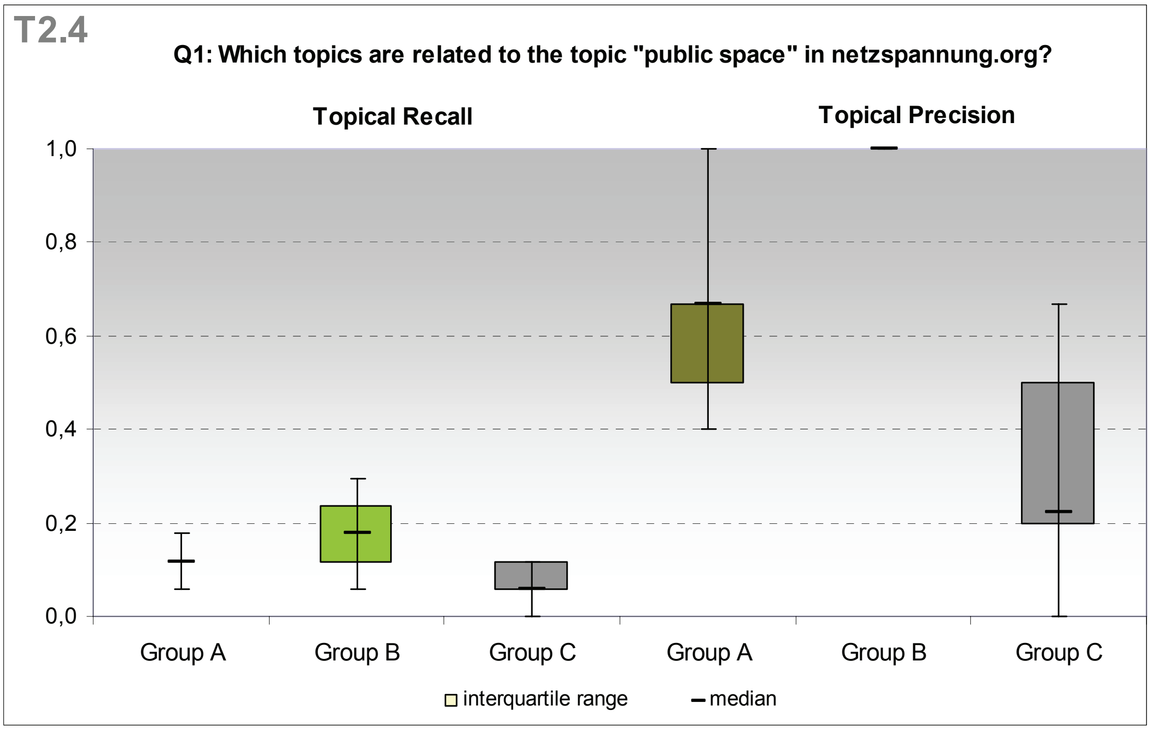


Fig. 9-38. Boxplot of topical recall and topical precision concerning the acquired level of understanding of the unfamiliar concept (“public space”) from the unfamiliar community (“netzspannung.org”) in task 2.4.

Fig. 9-38 depicts the topical recall and topical precision for answers to the question Q1 concerning the acquired level of understanding of the unfamiliar concept (“public space”) from the unfamiliar community (“netzspannung.org”).

It shows a significant difference between the results of group A and group B with respect to group C, by both measures. The medians of both group A and group B are higher than for group C and the inter-quartile ranges are non-overlapping in both cases. The largest difference is between the medians of group B and group C both for recall (0.18 vs. 0.12) and for precision (1.00 vs. 0.50).

Group A displays the most homogeneous results with virtually no spread for recall, whereas group B displays the same properties for precision. Given the highest recall and the absolutely highest median for precision with no spread (1.0) group B achieved the best results, followed by group A and group C.

The topical recall corresponds to the proportion of correctly named topics with respect to the reference solution for this question. Topical precision provides insight into the proportion of correctly named topics within all topics named by the users.

The above analysis thus implies that both group A and group B acquired a significantly better understanding of the unfamiliar concept from the unfamiliar community than group C. In addition, the results suggest that group B acquired a significantly better understanding than group A, due to greatly better topical precision and significantly better topical recall.

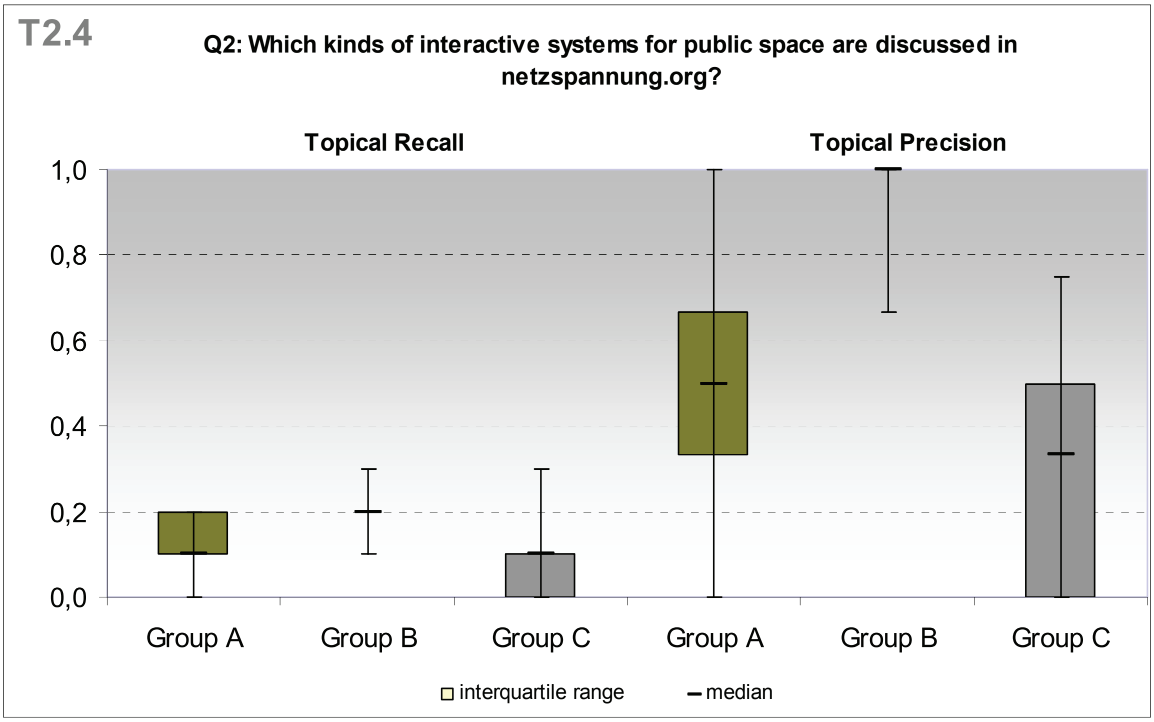


Fig. 9-39. Boxplot of topical recall and topical precision concerning the acquired level of knowledge integration between unfamiliar and familiar concepts (“public space” and “interactive systems”) in an unfamiliar community domain (“netzspannung.org”) in task 2.4.

The results for question Q2 concerning the acquired level of knowledge integration between unfamiliar and familiar concepts (“public space” and “interactive systems”) in an unfamiliar community domain (“netzspannung.org”) are visualised in Fig. 9-39.

The boxplot indicates significant differences between different grops. Group A and group B exhibit better results for both topical recall and topical precision than group C: their medians are higher, the inter-quartile ranges non-overlapping for recall and only partiall for precision (between group A and C), with distributions skewed in opposite directions. Group B exhibits by the highest median for recall and by far highest precision, indicating best results between all groups.

Such results suggest better ability of knowledge integration for both groups using knowledge map support (groups A, B) than the group without knowledge visualisation (group C). Furthermore, the group using personal maps (group B) performed better than the group using community maps (group A).

Fig. 9-40. shows a boxplot of topical recall and topical precision for question Q3 concerning the acquired level of understanding of relationships between the task in the unfamiliar community ("interactive systems for public space") and the own community domain ("HCI"). Group A and group B exhibit very similar behaviour for both measures.

With respect to topical recall their medians and inter-quartile ranges are identical, while for topical precision equal medians are differentiated only by a greater spread of group A. Significantly better results are achieved by group C: its median is significantly higher than those of group A and group B for both topical recall and precision. In addition, there is almost no spread for recall and the inter-quartile range of precision is comparable to that of group B.

Such results suggests that the best understanding of relationships to own community domain was exhibited by group C, while group A and B performed less well and in a comparable manner with respect to each other.

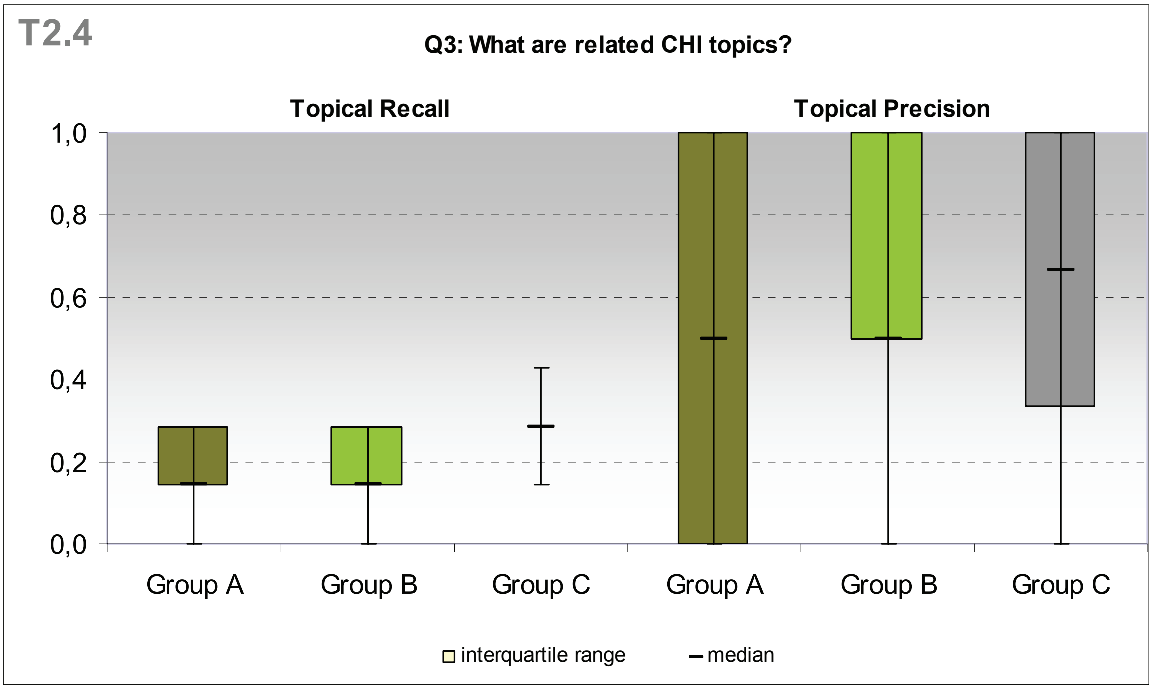


Fig. 9-40 Boxplot of topical recall and topical precision concerning the acquired level of understanding of relationships between the task in the unfamiliar community ("interactive systems for public space") and the own community domain ("HCI")

9.7.7. User Assessment of System Task-Adequacy

This section presents the results of the qualitative questionnaires on user-perceived task adequacy of the proposed knowledge map system.

First the results of the test groups using only the Knowledge Explorer are presented (Fig. 9-41 to Fig. 9-45), followed by the results of the comparative assessment of the Knowledge Explorer to the standard reference system (Google+Mozilla, Fig. 9-47).

9.7.7.1. Knowledge Map Support System (Knowledge Explorer)

In both groups using the Knowledge Explorer approximately two thirds of the users perceived the assigned tasks as realistic to very realistic and no users expressed a negative judgment (Fig. 9-41). Such positive assessment confirms the practical relevance of the task design.

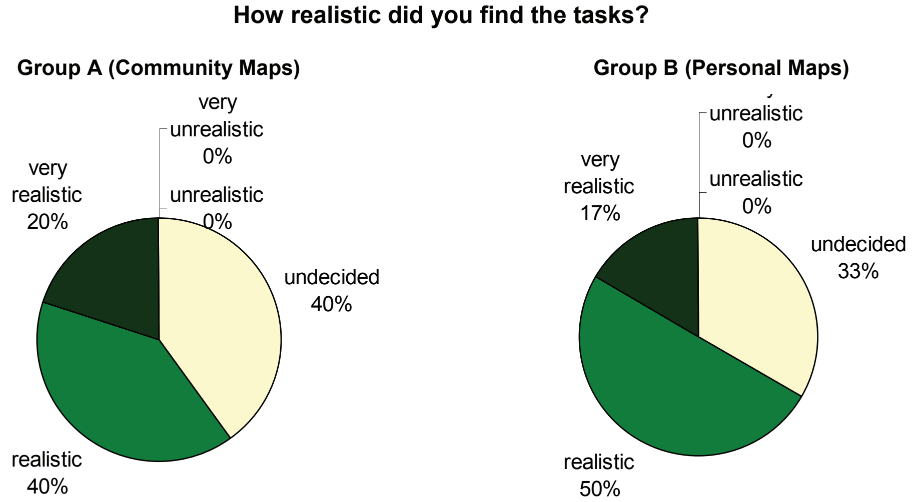


Fig. 9-41. User-perceived degree of realism of the tasks in Knowledge Explorer test groups (Experiment II).

Similar responses can be observed with respect to the assessment of overall suitability of the Knowledge Explorer for accomplishing the assigned tasks. In the group using the Community Maps configuration 60% of test persons considered the Knowledge Explorer to be well-suited and 40% as moderately suited for the assigned tasks (Fig. 9-42, left). The assessment of overall suitability by the group using the Personal Maps configuration is even somewhat higher with 67% of well-suited and 33% of moderately suited judgments (Fig. 9-42, right). Such highly positive assessment of the overall system suitability in both configurations provides supporting evidence for the validity of the proposed approach and its prototypical implementation in the Knowledge Explorer.

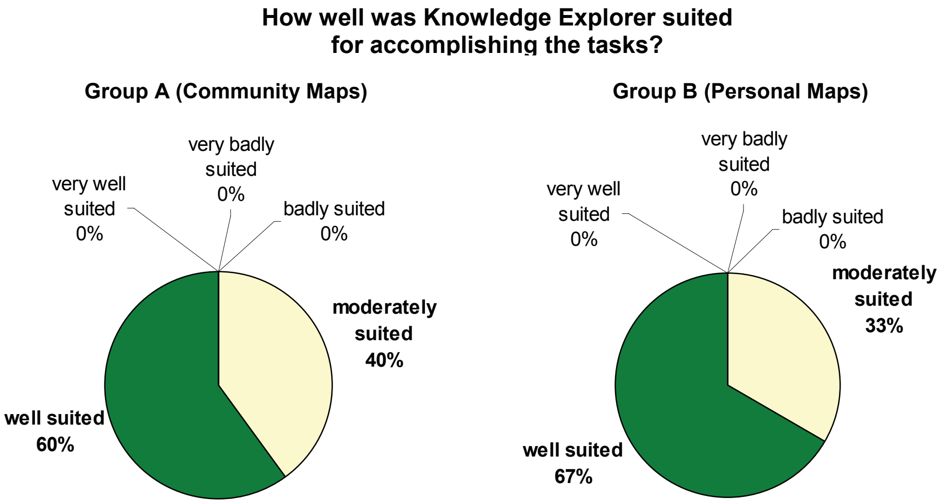


Fig. 9-42. User-perceived adequacy of the Knowledge Explorer for accomplishing the tasks in Experiment II.

Fig. 9-43 and Fig. 9-44 present the reported extent of use of individual functionalities and their usefulness as perceived by the users. The results show that both groups extensively used the Document Map (100% of users reported “much” or “very much” use), as well as the search queries and the visualisation of search results and related documents. The search queries were thereby used with a lesser intensity by users of personal maps (67% “much” or “very much” use) than by users of community maps (80% “much” or “very much” use).

On the other hand, users of personal maps reported greater use of the visualisation of search results (50% “much” and 50% “very much”) than users of community maps (60% “much” and only 20% “very much”). The visualisation of related documents has been used a lot by approximately the same number of users in both groups (50% “much” and 33% “very much” in group B, 40% “much” and 40% “very much” use in group A). But while only 17% of personal map users reported a moderate use of that functionality, 20% of community map users reported very little use.

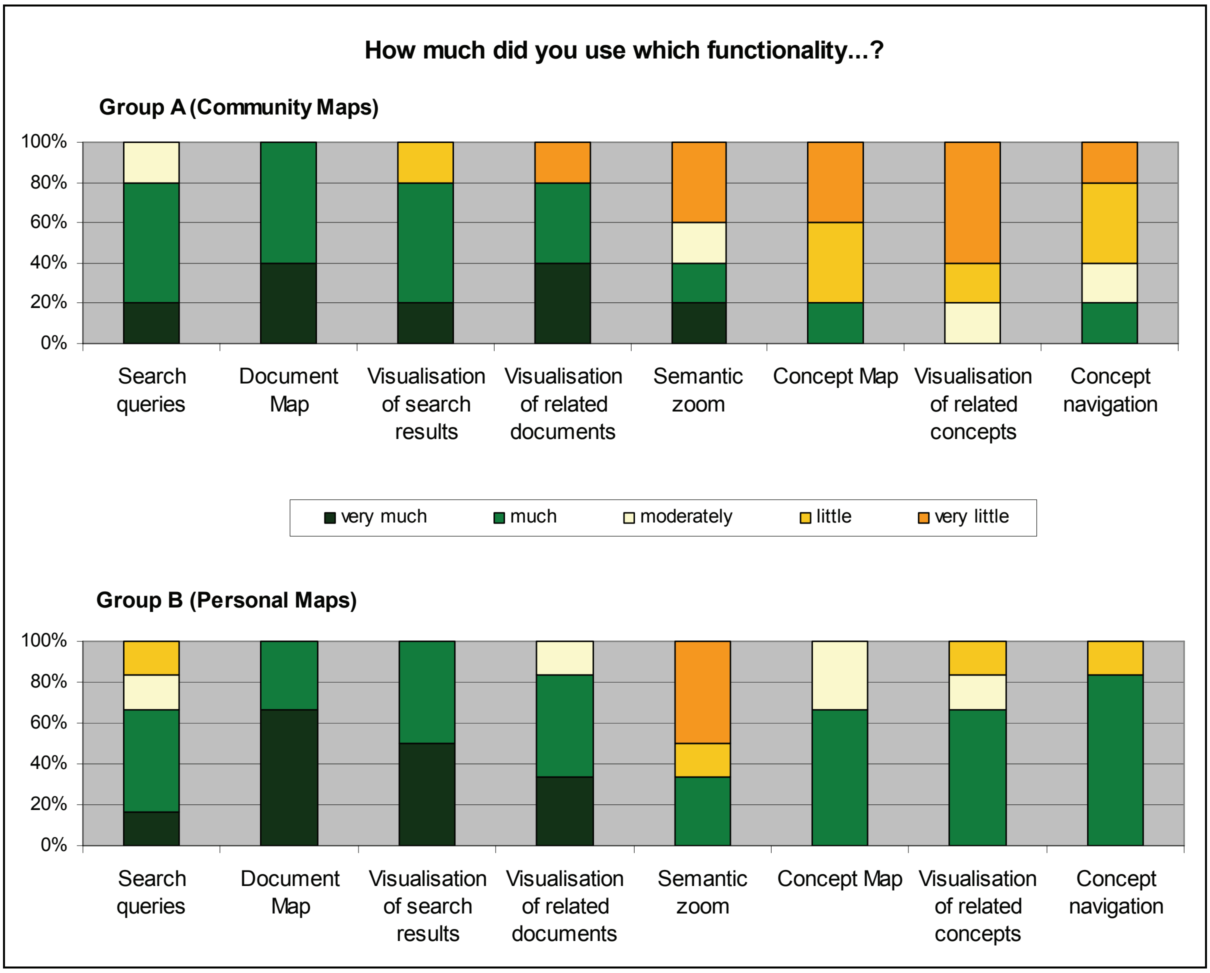


Fig. 9-43. User-reported extent of use of individual Knowledge Explorer functionalities in Experiment II.

This tendency of more extensive use of special visualisation functionalities by users of personal maps (group B) than by users of community maps (group A) is most clearly observable in the extent of use of Concept Map functionalities. Overall use of the Concept Map was low in group A (20% “much”, 80% “little” or “very little”) while it was extensively used in group B (63 % “much” and 17% “moderate” use). In particular, in group A the visualisation of related concepts has almost not been used at all (highest rating is “moderate” use, reported by 20% of users) while concept navigation has been used more but still only to a limited extent (20% “much” and 20% “moderate” use). On the contrary, group B used extensively both concept navigation (83% “much” use) and the visualisation of related concepts (67% “much”, 20% “moderate” use).

This difference in the usage of the Concept Map functionalities is the main difference in the usage of the system reported by the two different groups. As we will see later this coincides with the difference of reported information seeking strategies and with the difference in satisfaction and usability feedback between the users of community maps and users of personal maps. The perceived usefulness of individual functionalities reported by the users (Fig. 9-44) coincides with the described frequency of use. Both groups reported a high usefulness of Document Map functionalities whereas users of personal maps reported much higher usefulness for the Concept Map functionalities than users of community maps. The visualisation of document clusters and the visualisation of search results were rated equally well in group A, where both functionalities have been perceived as "rather useful" by 80% of the test persons. Users of personal maps rated the usefulness of the visualisation of document clusters roughly the same (83% "rather useful"), whereas the visualisation of search results was rated even better: all 100% of users found it "rather useful" (50%) or "very useful" (50%).

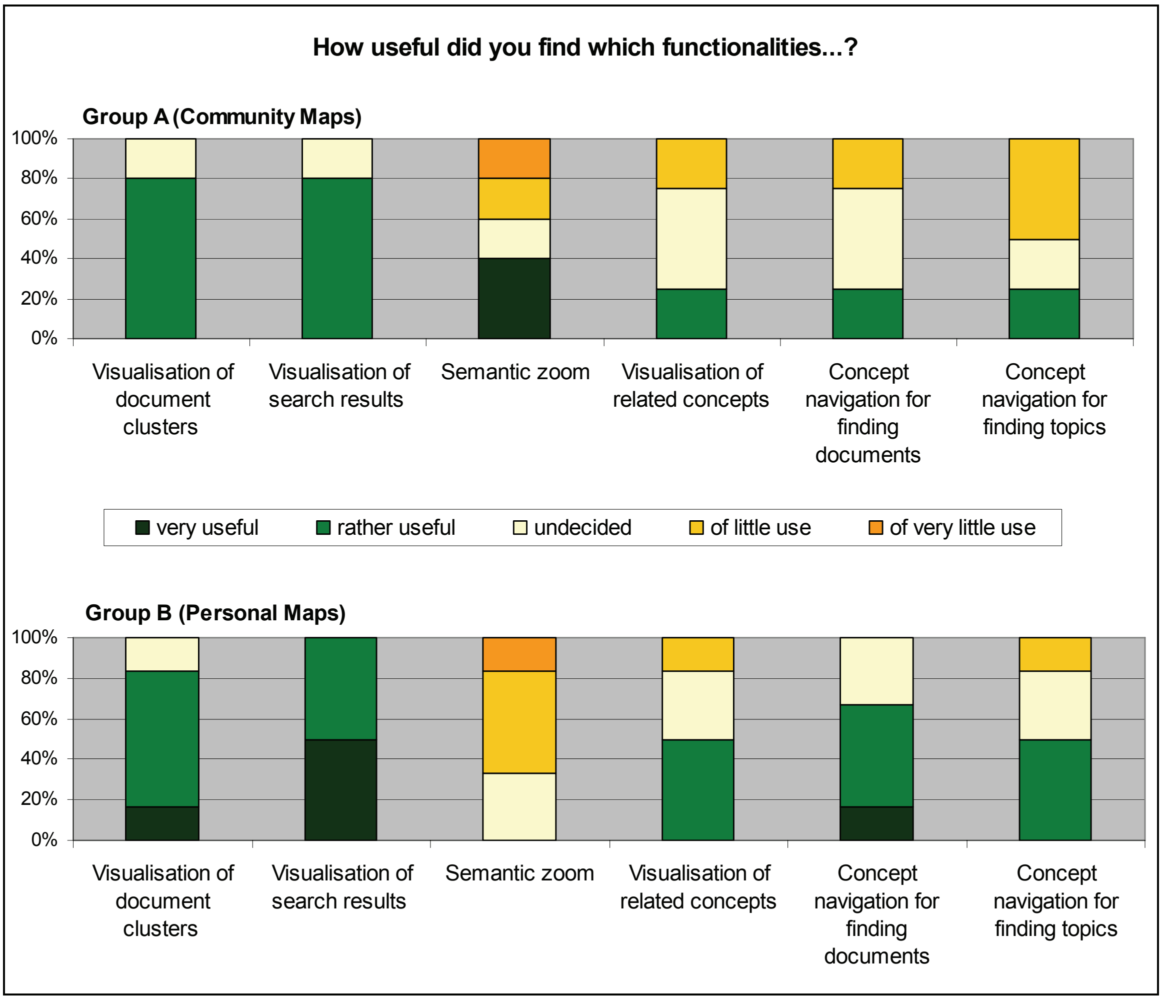


Fig. 9-44. User-perceived usefulness of individual Knowledge Explorer functionalities in Experiment II.

The visualisation of related concepts in the Concept Map was judged more useful by users of the personal maps (50% "rather useful" vs. 17% "of little use") than by users of the community maps (only 25% "rather useful" vs. 25% "of little use"[73]). Concept navigation was also perceived to be of little use by users of community maps: both for finding documents (25% "rather useful", 25% "of little use") and for finding topics (25% "rather useful", 50% "of little use") relevant for a task. Quite contrary, users of personal maps report concept navigation as a very useful function: 67% reported it "rather useful" or "very useful" for finding documents and 50% rated it "rather useful" for finding topics relevant for a task.

With respect to the Concept Map origin, a minority of users of community maps perceived community Concept Maps useful: both unfamiliar and own community Concept Maps were judged as "rather useful" by only 25% of users (Fig. 9-45, top). In contrast, personal map users perceived both Concept Maps very useful: 67% reported the own personal Concept Map to be "rather useful" while the personal Concept Map of unfamiliar community members was judged "rather useful" by 50% but also "very useful" by additional 17% of users (Fig. 9-45, bottom).

[73] Responses on the usefulness of individual Concept Map functionalities in Group A (Community Maps) are based on data from 4 users only: 1 user provided no answers with a remark he couldn't comment due to not having had used the Concept Map at all.

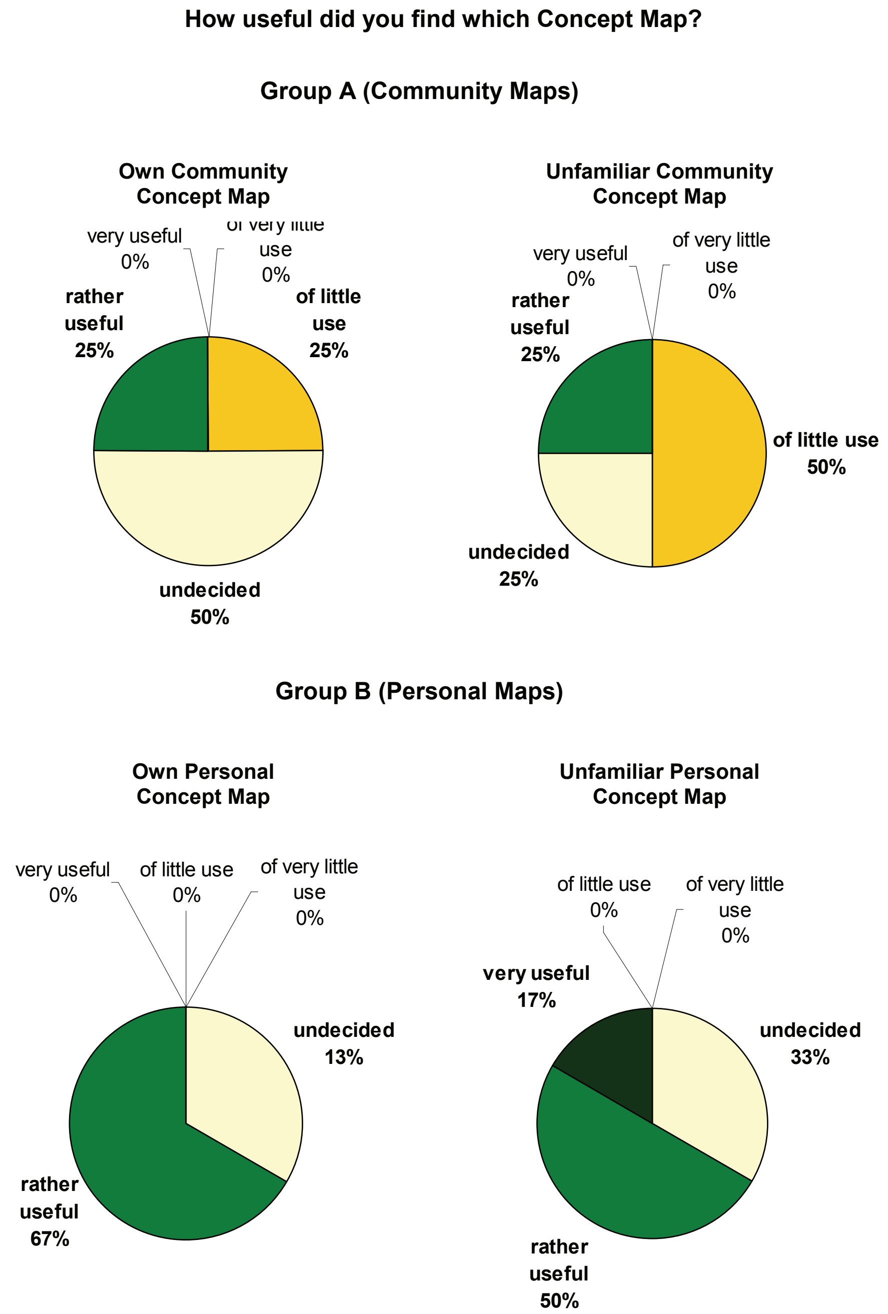


Fig. 9-45. User-perceived usefulness of different kinds of Concept Maps.

The semantic zoom wasn't perceived as useful as expected. The results exhibit relatively low levels of usage and of perceived usefulness in both groups, with somewhat better results for community map users. The community map users (group A) exhibit two opposing tendencies of the extent of use: while 40% reported "much" or "very much" use of the semantic zoom 40% also reported "little" use. The judgment of its usefulness is similarly divided: while 40% found the zoom "very useful", 40% judged it "of little" or "very little use".

The results of personal map users show a much clearer picture. For them the semantic zoom provided no significant support: 50% reported "very little" and 17% "little" use with exactly the same judgment of perceived usefulness. The results of the personal map group are understandable, since the personal maps present relatively small portions of the entire community space, with a human-defined semantic structure reflecting a specific information need and personal knowledge of the map author. In such a situation, overview is relatively easy to gain and navigation by points of interest is the main means of finding relevant information. This is also in accord with reported high levels of usage and usefulness of the Concept Map visualisation and navigation functionalities.

The divided feedback on the usage and usefulness of the semantic zoom by the community map group is rather surprising since the community document maps are much larger covering several hundreds of documents (435 and 608 in our case). Furthermore, the results of formative evaluation (Chapter 8.5) showed quite a different picture with great levels of user acceptance of the semantic zoom. The explanation of the polarized feedback of community map users is found in user comments in the usability part of the questionnaire and informal feedback (Section 9.7.8). Both of those point to some usability problems with the display of overlapping titles in high density areas, experienced by the users who didn't take advantage of the semantic filter to reduce the number of simultaneously displayed titles.

The described results of reported usage and perceived usefulness indicate an important difference in the use of the system between the users of community maps and the users of personal maps. While both groups extensively used the Document Map functionalities and reported high levels of usefulness, the personal map exhibits a much higher extent of usage and perceived usefulness of Concept Map functionalities, such as visualisation of concepts related to a search query and concept navigation for finding documents and topics relevant for a task.

As discussed in Chapter 8.6.3 these functionalities provide important support for typical sensemaking tasks. Eliciting user feedback on their usefulness was the basis of measuring the process-based quality of support for knowledge construction occurring during information seeking activities (Section 9.4.4). Accordingly, the significant difference in the usage and usefulness of Concept Map suggests that personal maps provide better support for typical sensemaking tasks occurring during information access than community maps.

On the other hand, the extensive use and greatly positive feedback in both groups on all other functionalities of the system (especially related to the Document Map) confirms the suitability of the proposed approach and the overall design of the Knowledge Explorer tool. In particular, the individual differences in the extent of use of different functionalities reported by the users suggest that such a combination of different functions is a good way of accommodating the differences in personal information seeking strategies. The next section describing the informal feedback of the test persons provides us with more insight into this important issue.

9.7.7.2. Comparison with the Standard Reference System (Google+Mozilla)

The degree to which the given tasks were perceived to be realistic in group using the standard reference system (Google+Mozilla) is comparable to that of the groups using the Knowledge Explorer: 60% of users considered the assigned tasks as realistic and no user expressed a negative judgment (Fig. 9-46). Such positive results for all three test groups indicate a high practical relevance of the assigned tasks and confirm the adequacy of the task design.

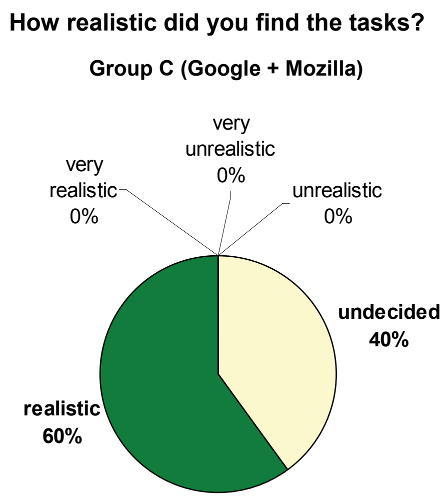


Fig. 9-46. User-perceived degree of realism of the assigned tasks in the reference system test group.

As described in Section 9.7.1, after completing the tasks by using the standard reference system, this test group repeated Task 2.1 (20min) and Task 2.2 (10min) by using the Knowledge Explorer. Due to obvious carry-over effects from the previous work on the same tasks, assessing the quality of these task solutions was not feasible. The goal was to obtain a direct user comparison of the task-adequacy of the Knowledge Explorer with respect to the standard reference system. The results in Fig. 9-47 indicate that users of the standard information seeking system (Google + Mozilla) didn't find it very adequate for accomplishing the assigned tasks. Only 40% of users reported that the system was "well suited" for accomplishing the assigned tasks, 40% judged it only "moderately well suited" whereas 20% considered it "poorly suited" for solving the tasks[74]. Such assessment is in line with the problems encountered in solving the tasks that were reported by the users.

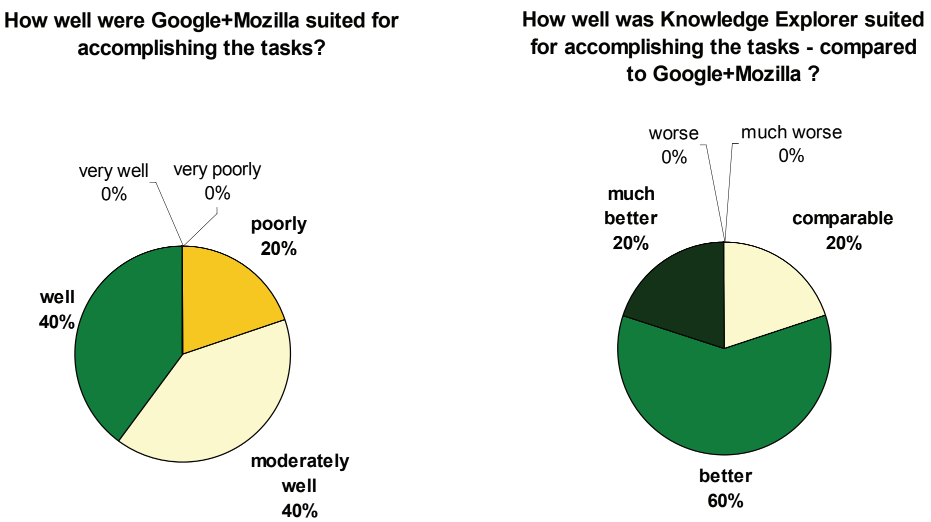


Fig. 9-47. User-perceived task-adequacy of the reference system and comparison to the Knowledge Explorer.

All users in this group complained about the difficulty of finding appropriate concepts to describe the information need of the task and of identifying topics for organizing document groups. As a critical problem they mentioned the ill-defined nature of the topic of task 2.1. The notion of "digital culture" was not strictly defined even in its community of origin: there were many different ways in which the concept of "digital culture was used by different community members. This empirically confirms the gravity of the problem of gaining insight into implicit community knowledge in cross-community knowledge access (Chapter 3). The difficulty of getting an overview of topics in the community space and their possible relationships to the topic of the task was also perceived as a big obstacle to the successful completion of the task. These problems explain the low level of structuring of task results for group C (Section 9.7.6).

The lack of support for addressing such problems by the reference system is reflected in the highly favourable assessment of the task-adequacy of the Knowledge Explorer compared to the Google and Mozilla system. Fig. 9-47 shows that 80% of users judged the Knowledge Explorer to be "better" (60%) or "much better" (20%) suited for accomplishing the assigned tasks.

9.7.8. Personal Information Seeking Strategies of the Test Persons

In addition to eliciting user feedback through qualitative questionnaires, informal feedback was gathered through questions during the accomplishments of tasks and in the group discussion following each session. To this end the users have been asked to briefly describe how they went about accomplishing the task and the problems experienced in using the system. The user feedback reveals that different users

[74] The exposure of group C to Knowledge Explorer (30min) was much shorter than that of group A and B (aprox. 2h). Hence, while the users were able to compare the impression of usefulness with respect to the experience of using the reference system, a differentiated assessment of individual functions as in groups A and B could not be expected. Accordingly, the questionnaire for group C contained regarded only the overall assessment of task-adequacy and comparison between the two systems.

employed different strategies in solving the task based on their usual personal information seeking behaviour. All users reported an initial phase in which they rather randomly tried out different exploration and navigation possibilities offered by the system, before adopting a more consistent strategy which they perceived the most fruitful and easy to use. The next paragraphs summarize the most frequently occurring patterns of system use reported by the users. Although the users tended to describe their personal strategies in working with the system quite precisely, each user exhibited a slightly different behaviour and a number of users combined different strategies. Accordingly, the findings summarized in the following paragraphs should be taken with due caution as an overall qualitative interpretation of user's informal feedback.

Several users of the Knowledge Explore Community Map configuration reported a systematic Google-like approach: issuing search queries and inspecting document from the search result set visualised on the map, one by one. After all documents from a search result were inspected, they turned to examining other documents that were located in the spatial proximity of the search results on the map. The visualisation and navigation by concepts in the Concept Map was only occasionally used by such users (e.g. when they couldn't find any more relevant documents following the described procedure). Interestingly, such users were located largely in the community map group and gave low ratings of the usage and usefulness of Concept Map functionalities. We can establish this cause relationship for two users with certainty since they explicitly commented on the reasons of their low Concept Map usage in the free form part of the questionnaire.

Another group of users exhibited a variation of this approach by starting with a search query and inspecting documents from the search result set, but with more use of specific Document Map functionalities. Rather than inspecting individual documents going through the set visualised on the Document Map, these users went through the search result list selecting documents one by one. As a result the selected document would be displayed on the map alongside with a set of potentially related documents (not necessarily appearing in the results of the search query).

The users would then first inspect all of the related documents before proceeding to the next document in the search result list. Only after exhausting all documents in the search result list, their attention would turn to the Concept Map. Such users also reported higher levels of use of the Document Map than the Concept Map. There were more of such observations in the group using community maps than in the group using personal maps.

A third pattern of use shared by different users (across both Knowledge Explorer groups) has been a berry-picking like method (Bates, 1989): after issuing a search query the users would examine some of the documents from the search result set visualised on the Document Map. In doing so they would take notice of related concepts visualised on the Concept Map and if some drew their attention they would continue their seeking by navigating through concepts on the Concept Map.

In this way, the clarification of the ill-defined information need evolved with the discovery of interesting concepts presented in the map. The use of the Concept Map would thereby increase as the number of documents retrieved so far increased and as more attention was given to structuring documents into topical groups and identifying relevant topics. Only two users in the community map group followed this strategy, whereas different users in the personal map group reported such behaviour.

In addition, several users in the personal map group adopted a rather systematic strategy of interrelated use of Concept Map and Document Map functionalities. After issuing a search query they would examine a few of the documents visualised in the search results that seemed most relevant. They would then examine the concepts visualised in the Concept Map that were related to individual documents in order to identify which concepts from the unfamiliar community were related to the topic of the task. They then selected promising concepts in order to find additional documents. When this led to little additional results, they continued by reformulating the search query by combining the most promising concepts with another.

All users of the standard reference system (Google+Mozilla) reported the same strategy based on iterative formulation of search queries. Their main problem was the difficulty in finding appropriate concepts to describe the information need of the task in the unfamiliar community space. Accordingly, the users attempted to formulate different search queries by combining concepts contained in the task description (e.g. “interactive” + “culture” = “interactive culture”). As this produced only limited results and the combinations were quickly exhausted, the users spent more time reading full-texts of the documents in order to identify potential topics. Thus they exhibited a classical iterative search cycle of formulating the query, analysing the results, reformulating the query accordingly and so forth. In contrast to the groups using the Knowledge Explorer they were confronted with the problem of identifying a set of relevant topics that they could use in formulating appropriate search queries, without any topical indication other than the contents of documents from an unfamiliar community domain.

The feedback in the group discussion highlighted two general attitudes of the two different groups of Knowledge Explorer users that explain the results of the task-adequacy questionnaires. The majority of users of the community maps referred to the relative complexity of the community Concept Maps as an important reason for lesser extent of use and perceived usefulness. Furthermore, two users in this group voiced strong personal preference on Google-like information interfaces over visual displays. Nonetheless, all users (including the latter) expressed great satisfaction and appeal of the Document Map.

Two users asserted that the Concept Map was useful for identifying relevant concepts and topics in order to formulate and structure the (ill-defined) information need of the task. All users expressed that they would have expected a great help of the Concept Map for orientation in the unfamiliar community domain, but that this was diminished by the large number of different and potentially relevant topics contained in the community Concept Maps. In contrast, the navigation by the Concept Map of the own community was perceived more helpful since it was simpler and contained familiar topics. The general group attitude of was that for retrieving documents they preferred explorative access in the Document Map whereas the community Concept Maps provided better support for topical structuring.

The users of the personal maps painted quite a different picture. They described personal maps as giving a good immediate overview of topics and documents relevant for a task and being easy to use. Since in the initialization phase they had used a community map configuration (with system-generated maps) they perceived the personal map configuration as much less complex. This group found the personal Concept Maps quite useful and liked that they contained a small number of topics with great discriminatory power.

This group also exhibited a general preference for the visualisation mode of the Document Map. Some users suggested that Concept Map visualisation be more tightly integrated with Document Map into one single information display. On the other hand, the explicit separation of one’s own personal Concept Map from the Concept Map of a member of an unfamiliar community was found helpful. A number of users reported a high usefulness of the comparison between documents related to own personal concepts and documents related to concepts from the unfamiliar community member. Both groups perceived the overall duration of the experiment as too long given the significant cognitive effort of the tasks and appreciated the two breaks between the tasks. This suggests that the reported relative strenuousness is likely not due to the interface but to the cognitive demands of the tasks themselves. A number of users also explicitly explained their answers in this way.

9.7.9. Overall Usability Assessment

The results of the usability questionnaire (Section 9.4.3, Table 9-7) are presented in Fig. 9-48. The results indicate that there is no significant difference between the two groups with respect to perceived ease of use and learning effort. In both groups the large majority of users reported that they were able to use the system well in accomplishing the tasks (60% well and 40% moderately well in group A; 50% well and 33% moderately well in group B). The majority of users also reported that the system was easy to use (60% group A, 67% group B) and that it was easy to learn using it (80% group A, 66% group B).

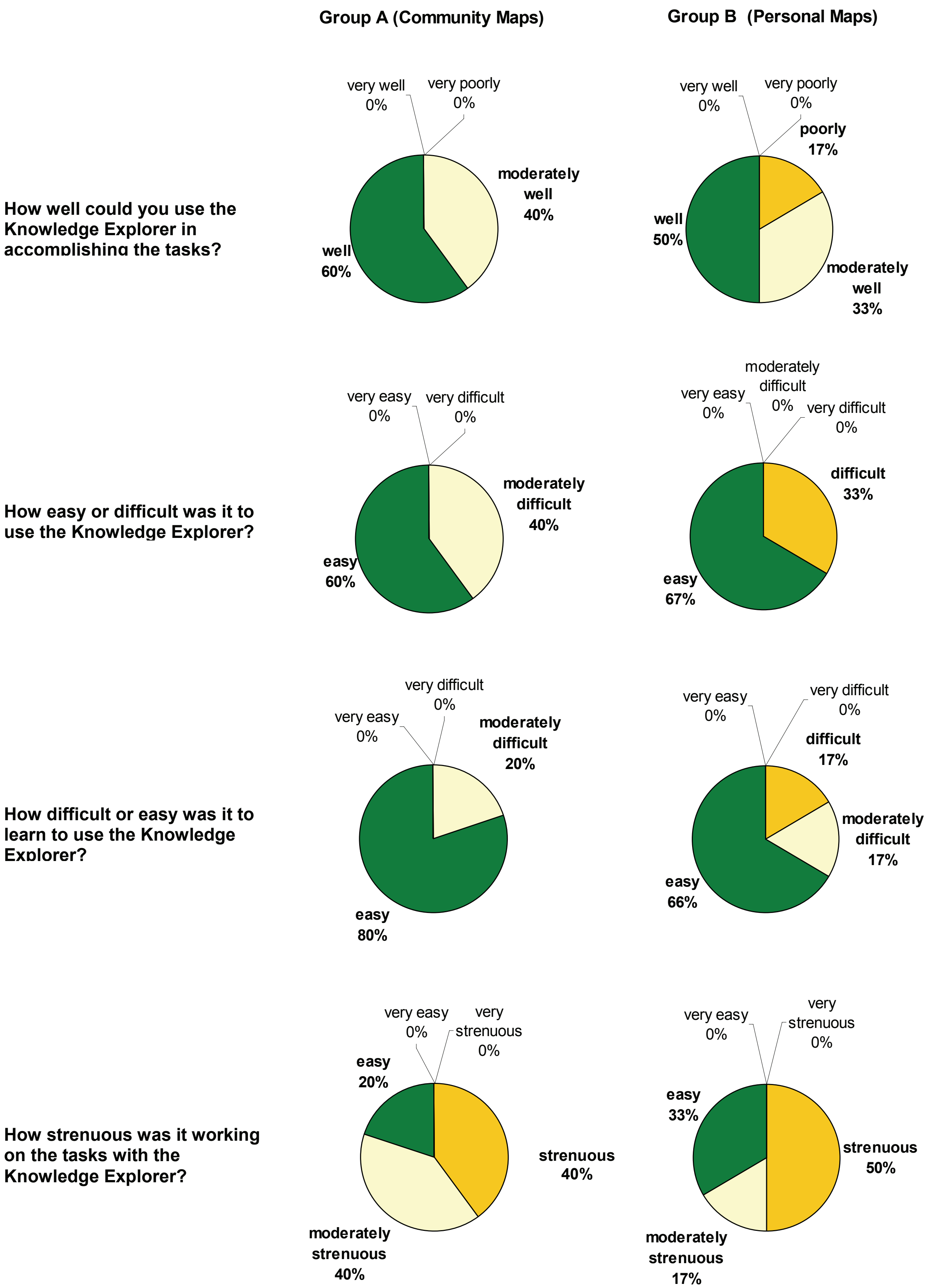


Fig. 9-48. Results of the questionnaire concerned with overall usability of the Knowledge Explorer.

The percentage of users reporting some difficulties in using the system is higher in the personal map group than in the community group. Whereas in the community map group no user judged the system to be difficult to use or learn, in the personal map group 17% of users reported they were only "poorly" able to use the Knowledge Explorer for accomplishing the tasks, 33% that it was difficult to use and 17% that it was difficult to learn using it.

The majority of users in both groups reported that working on the tasks with the system was to some extent strenuous (20% strenuous and 40% moderately strenuous in group A; 50% strenuous and 17% moderately strenuous in group B). The latter responses assume a completely different meaning when the free form comments of the users are concerned. Not so much using the system was perceived to be strenuous, but the tasks themselves were strenuous due to their cognitive complexity and above all the length of the experiment. Although two breaks have been incorporated in the course of the experiment (Table 9-25) the total length of 3h 20min (including filling out questionnaires) was perceived as very demanding by all users. The total time of the reference system test group (group C) was 30min longer due to the additional usage of the Knowledge Explorer for the comparative assessment (Section 9.7.7.2).

Question	Freq.
…something you particularly liked about the Knowledge Explorer?	
document map	****
visualisation of document clusters	***
visualisation of relationships between own and unfamiliar community concepts	***
visualisation of related concepts	**
concept map	**
contextualisation of search results	**
visualisation of related documents	**
personal maps	*
direct access to abstracts and full-texts	*
semantic zoom	*
spatial organisation of the workarea	*
…something very practical?	
document map	****
drag & drop	***
visualisation of document clusters	***
visualisation of search results	**
visualisation of similar documents	*
personal concept map	*
concept navigation	*
semantic zoom	*
direct access to abstracts and full-texts	*
…something you disliked?	
semantic zoom didn't work well	****
overlapping document titles	***
bugs	***
switching between search panel and personal maps	**
…what would you improve?	
better zoom mechanism: overlapping titles shoud spread faster	****
limit zoom to selected area. only focus, no overview	****
building subclusters in personal maps	***
connecting related items with lines	**

Table 9-34. Free-form answers to questions on usability and user satisfaction with the Knowledge Explorer.

In addition to the presented specific usability aspects, users were also asked about what they particularly liked or disliked about the Knowledge Explorer in general, what they found especially practical as well as about any suggestions for improvement. Whereas the detailed usability questionnaire was presented only to the groups A and B, the free form questions were submitted also to group C after their additional exposure to the Knowledge Explorer[75].

The user answers to the free-form questions from all three test groups are summarized in Table 9-34 along with an indication of approximate frequency of their occurrence (*** = 4 or more times, *** = 3 times, ** = 2 times, * = 1 time). Since the answers were freely formed they needed to be interpreted and clustered. The resulting summary is thus not exact but provides us with a qualitative impression of user feedback. We can observe that the document map was by far the most liked and practical functionality and that the users liked and found practical different visualisation functions such as the visualisation of search results and related documents. Several users also liked and found practical the visualisation of related concepts in the Concept Map and the visualisation of relationships between own and unfamiliar community concepts. The main aspect with which users weren't satisfied were overlapping titles of closely positioned documents and a zooming mechanism effecting to late title spread-out.

In addition to the free-form questionnaire answers, just like in the Experiment I, users tended to provide unsolicited feedback after the experiment. A number of users self-initiatively reported the use of the tool as an interesting experience and suggested possible fields of applications (e.g. "researchers workbench", teaching and learning support etc.). Several users declared that the tool was "fun to use", that they discovered a lot of interesting projects in themes they don't commonly encounter and inquired about the possibilities of using the tool as a research tool for their own personal work.

9.7.10. Summative Interpretation and Conclusions

The analysis of results presented in previous sections shows that the groups using knowledge map support (groups A, B) demonstrated overall better performance in accomplishing the assigned tasks than the group using a standard information seeking system without knowledge visualisation (group C). This is supported in particular by significantly better results with respect to topical structuring quality, the learning effect and task-adequacy assessment. Furthermore, in most cases users of the Personal Map configuration of the Knowledge Explorer (group B) achieved better results than users of the Knowledge Explorer Community Map configuration (group A).

This suggests supportive evidence for the hypothesis (H3) stating better suitability of the community map configuration than the standard reference system without knowledge visualisation, the hypothesis (H4) stating better suitability of the personal map configuration than the standard reference system and for the hypothesis (H5) stating better suitability of the personal map configuration than the community map configuration (Section 9.1). The next sections summarize the interpretation of the experiment results with respect to these hypotheses based on the discussion of the quality of task solutions, user perceived task adequacy of the proposed solution and informal user feedback presented in previous sections.

9.7.10.1. Quality of Task Solutions

Document retrieval effect

Regarding document retrieval effect, users of the knowledge map system (Knowledge Explorer, groups A and B) tended to achieve comparable document recall but a significantly higher precision than groups using the alternative information seeking system without knowledge visualisation (Google + Mozilla). In addition, the group using the Personal Map configuration of the Knowledge Explorer consistently achieved better results than the group using the Community Map configuration (Fig. 9-24, Fig. 9-32, Fig.

[75] Due to the short exposure to the Knowledge Explorer of group C, compared to groups A and B submitting the full usability questionnaire to that group was not appropriate. Only the free-form questions were presented to users from group C.

9-33, Fig. 9-34). Accordingly, for overall document retrieval effect we can observe a tendency of better performance for both groups with knowledge map support, since given at least comparable recall they display significantly higher precision. While comparable recall signifies comparable amount of retrieved relevant documents, higher precision is a first indicator of a higher level of acquired understanding of the unfamiliar community domain by the users of the Knowledge Explorer system. This supports our hypotheses stating better suitability of community maps (H3) and personal maps (H4) than a standard information seeking system withouth knowledge visualisation.

Quality of topical structuring

This assumption is further supported by the extreme difference in the quality of topical structuring between both groups using the Knowledge Explorer and the group using the strandard reference system (Google+Mozilla). The number and quality of topics identified and applied for structuring documents relevant for the task by the users of the standard reference system is very low. In task 2.1 Google users created 1.6 relevant topics on average (median=2) in contrast to 3.80 (median = 4) and 4.5 (median = 4) for the Knowledge Explorer users of community and personal map configurations, respectively (Fig. 9-25). The extent of the use of cross-community topics in the Google group is even lower (average = 0.80, median = 1) and several users created no topics at all. This problem can be observed even more extremely in task 2.4 where no single user of the Google group has undertaken any structuring at all (Fig. 9-35). In spite of a well-defined information need (narrower focus of task 2.1) and gained familiarity with the information space of netzspannung.org, the much shorter time than in task 2.1 (15 vs. 40min) made it even more difficult to identify relevant topics without knowledge visualisation support.

The difficulty of finding appropriate topics in the group using the standard reference system without knowledge visualisation is confirmed by the user feedback in the questionnaires. All users from this group heavily complained about the difficulties in understanding the vocabulary of the unfamiliar community (netzspannung.org) and identifying relevant topics for the task (Section 9.7.7.2). This issue has also been brought up in the informal feedback session following the conclusion of the experiment. Such feedback suggests an explanation for the good performance of the standard reference group with respect to document retrieval measure, especially in task 2.1 and 2.2. Given the difficulties in identifying appropriate topics in the unfamiliar domain, at some point the Google group users focused their attention more on retrieving relevant documents and less on finding topics. This is a possible explanation why Knowledge Explorer groups produced only comparable recall, since their time was spent between retrieving documents and producing significant topical structures.

The results of the group using the reference system are comparable and sometimes better with respect to tasks requiring the identification of documents from the unfamiliar community related to very specific familiar concepts from own community. This is suggested by better document recall performance than the group using the Community Map configuration of the Knowledge Explorer in tasks 2.2 and 2.3 (Fig. 9-32 left, Fig. 9-33 left). On the other hand, the better recall of the standard reference group in task 2.2 and 2.3 is sacrificed by significantly lower precision (Fig. 9-32 right, Fig. 9-33 right). This actually suggests a lower level of comprehension of the document contents and their semantic context.

The richness of the knowledge transfer that occurred at the group level in both Knowledge Explorer groups compared to the Google group is also indicated by a much higher cumulative cross-community topical coverage (35%-62% compared to 0-8% in tasks 2.1 and 2.4). Overall the results of topical structuring measure provide supportive evidence for confirming the hypotheses stating better suitability of community maps (H3) and personal maps (H4) than a standard information seeking system without knowledge visualisation.

The main difference in topical structuring between the two Knowledge Explorer groups lies in the higher use of previously unfamiliar concepts from the netzspannung.org community by users of personal maps (group B) whereas the users of community maps (group A) largely used familiar HCI concepts in organizing retrieved documents (Fig. 9-25, Fig. 9-35). This is in accord with greater usage and user satisfaction with Concept Maps in the Personal Map group compared to the Community Map group (Fig. 9-43 to Fig. 9-45). This correspondence suggests that the usage of Concept Maps from an unfamiliar community helped in understanding the relationship between retrieved documents and unfamiliar concepts. Furthermore, it supports the assumption that the availability of personal concept structures of individual community members is a better means for supporting this than the sole use of community maps. This additionally supports both the hypothesis (H5) as well as the proposed approach to eliciting personal knowledge maps and their application to supporting for information access in unfamiliar community spaces (Chapter 6.4, Chapter 7.1.3, 7.1.5).

Learning effect

The results of the learning effect questionnaire provide a similar picture. The "learning effect as reflected in the number and quality of topics found reported in the questionnaire is very high for both Knowledge Explorer groups and very low for the Google group in both tasks in which the learning effect has been assessed (task 2.1, task 2.4).

In both tasks the group using the standard reference system performed much worse on questions related to the acquired level of understanding of the unfamiliar concepts from the netzspannung community (Fig. 9-28, Fig. 9-38) and on questions assessing the level of cross-community knowledge integration (Fig. 9-30, Fig. 9-39). The Google group was able to match the performance of Knowledge Explorer groups on the learning effect questions related to the acquired level of understanding the relationships of the knowledge from the unfamiliar community to the own domain (Fig. 9-31, Fig. 9-40). There, it exhibited a tendency of better performance from the group using the Community Map configuration and comparable performance to the group using the Personal Map configuration of the Knowledge Explorer.

The qualitative analysis of the answers explain why: since the users in this group were hardly able to acquire concepts from the unfamiliar community, their information seeking patterns focused on identifying those concepts that they could recognize from their existing background knowledge and own community. Accordingly, they matched familiar concepts that also appear in the unfamiliar community space (such as "mixed reality", "augmented reality" and "multimodal interaction"). On the contrary, since users in both Knowledge Explorer groups were able to acquire a much better understanding of previously unfamiliar concepts from the unfamiliar community space they had more difficulties in recalling related concepts from own community. This explanation is strongly supported by explicit user remarks (such as "didn't focus on this aspect").

Thus, with respect to the learning effect we can observe much better performance of both Knowledge Explorer with respect to identifying and understanding relevant knowledge from the unfamiliar community space both for highly ill-defined and for specific information needs (task 2.1 vs. task 2.4). The performance of the group using personal maps (group B) is here also consistently better than the group using community maps (group A). Users of personal maps exhibit significantly results than users of community maps for an ill-defined information need, whereas the differences become smaller but still notable in the case of a more specific, well-defined information need. Such results provide supportive evidence for the hypotheses H3, H4 and H5 with respect to the learning effect measure.

9.7.10.2. Task-Adequacy and Sensemaking Behaviour

The majority of users considered the Knowledge Explorer to be well-suited for accomplishing the tasks requiring the identification and understanding of relevant knowledge in an unfamiliar community space (Fig. 9-42). Explicit comparison by the users of the standard reference group shows a clear preference and higher task-adequacy of the knowledge map support (Knowledge Explorer) compared to the reference information seeking system without knowledge visualisation (Google+Mozilla). Less than half of the users of the standard reference system (Google + Mozilla) considered this system to be well-suited for solving the assigned tasks (Fig. 9-47). In a direct comparison of the two systems, the majority of the same users perceived the Knowledge Explorer to be more suitable for accomplishing the assigned tasks than the standard reference system without knowledge visualisation (80% of users from group C, Fig. 9-47).

The results of user reported extent of use and perceived usefulness of individual Knowledge Explorer functionalities show good acceptance by the users (Fig. 9-43, Fig. 9-44). They provide supporting evidence for the suitability of the proposed knowledge map model based on a Document Map and Concept Map combination and its application to supporting information access in unfamiliar community domains (hypothesis H2).

More specifically, the results of Knowledge Explorer groups indicate important differences between the users of community maps and users of personal maps. While the both groups exhibit a high extent of use and usefulness of the Document Map functionalities, users of personal map reported a much higher usage and usefulness of Concept Map functionalities (such as visualisation of related concepts and concept-based navigation). Since these functionalities provide important support for typical sensemaking tasks (Chapter 8.6.3, Section 9.4.4) such user feedback suggests that personal maps provide somewhat better sensemaking support in unfamiliar community spaces than community maps.

This is partly relativized by informal user feedback, which suggests the existence strong personal preference of a search strategy based on goal-directed information retrieval rather than explorative access for several users in the group using community maps (Section 9.7.8). This implies that the low extent of use and perceived usefulness of the Concept Map in the group using community maps could be a result of personal preference of a specific information seeking strategy rather than system functionality. The same users reported a perception of higher complexity of using the community concept maps with in addition to the document map, due to a wide range of concepts contained in it.

In contrast, those users that did use the community Concept Maps more extensively both reported a greater degree of usefulness and exhibited a preference for explorative access and navigation. These users reported dominant usage of concepts created by human users and little use of the terms from the text-analysis. This suggests that even for initial community concept maps based on a small sample of user activities the proposed method of eliciting user-defined concepts produces feasible results, without the need to be complemented by terms from the text-analysis. This is an important result as it shows the robustness of the method and a possible improvement based on a user-centred evaluation of the application validity of produced knowledge structures.

At the same time, users of personal Concept Maps describe personal Concept Maps as simple to use, with limited number of concepts offering quick topical overview with high relevance for a given search query. Thus, while personal preference specific search strategies did play an important role in the subjective assessment of community Concept Map functionalities, there is also a differentiation in the use of personal and community concept maps based on their structural properties.

On the other hand, the extensive use and greatly positive feedback in both groups on all other functionalities of the system (especially related to the Document Map) confirms the suitability of the proposed approach and the overall design of the Knowledge Explorer tool. In particular, the individual differences in the extent of use of different functionalities reported by the users suggest that a combination of Document Map and Concept Map functionality is a good way of accommodating the differences in personal information seeking strategies. This suggests supportive evidence for hypothesis (H2) assuming the suitability of such a model.

Finally, the lack of adequate support of standard information seeking system without knowledge visualisation for cross-community knowledge access is further confirmed by the results of the group using the standard reference system. Not only is the reported task-adequacy much lower than the assessment of the Knowledge Explorer, but the user feedback also points to serious problems experienced in working on the tasks. All users of the standard reference system reported great difficulties in finding appropriate concepts to describe the information need of the task and in identifying appropriate topics for organizing documents into groups. This illustrates the perceived lack of appropriate support for typical sensemaking tasks and their critical importance in accessing unfamiliar community spaces.

Such problems were resolved by users of the Knowledge Explorer in adopting information seeking strategies that took advantage of the provided sensemaking support functionalities. In particular, this is visible in combination of goal-directed search, explorative browsing and concept-based navigation enabled by functionalities such as visualisation of search results on the Document Map, visualisation of related concepts on the Concept Map and concept-based navigation for finding relevant topics to structure retrieved documents and task need (Section 9.7.8). Users taking advantage of those functionalities reported an information cycle corresponding to a typical sensemaking learning loop. Most users exhibiting such behaviour were located in the group using personal maps. The better overall performance of this group (and in particular of topical structuring) suggests that best sensemaking support was provided by the Personal Map configuration of the Knowledge Explorer.

9.7.10.3. Overall Usability Assessment

The results of the overall usability assessment show good acceptance of the prototypical realization of the Knowledge Explorer. The majority of users in both groups using the Knowledge Explorer reported the system as easy to use (60%-67%, Fig. 9-48) and easy to learn (66%-80%, Fig. 9-48). The majority of users reported that working on the tasks with the system was to some extent strenuous (Fig. 9-48) while attributing this primarily to the cognitive complexity of the tasks and the length of the experiment, rather than to using the system.

With respect to individual functionalities, some usability problems were reported regarding the use of the semantic zoom. The users complained about too slow spreading out of overlapping titles. In contrast, free-form user feedback highlighted the document map as the most liked and practical functionality followed by different visualisation functions such as the visualisation of search results and related documents. Other functionalities emphasized as practical by the users include the visualisation of related concepts in the Concept Map and the visualisation of relationships between own and unfamiliar community concepts. Such results confirm the overall usability of the proposed solution and its prototypical implementation in the Knowledge Explorer, with largely positive user feedback especially with respect to ease of use and suitability for accomplishing the assigned tasks.

10. Conclusions and Outlook

In this thesis we have developed a method and an interactive system for supporting cross-community knowledge exchange based on collaborative elicitation and visualisation of personal and shared community knowledge perspectives and their application to supporting information seeking in unfamiliar community spaces. The proposed approach has been prototypically implemented in a concrete system and a visual information interface enabling multi-perspective access to community information spaces, the Knowledge Explorer. The evaluation in a comparative laboratory study confirmed the adequacy of the developed solutions with respect to specific requirements of the cross-community problem and demonstrated significantly better quality of knowledge access than with a standard information seeking reference system without knowledge visualisation.

The challenge of understanding and supporting knowledge exchange across the boundaries of individual communities has been emphasized as an important practical problem in knowledge management and an open research issue. But while different studies analysed the problems involved with knowledge sharing between communities and proposed theoretical frameworks to address them, the development of concrete systems and tools incorporating these findings has received little attention. Furthermore, none of the existing solutions for knowledge sharing has been theoretically grounded nor empirically evaluated with respect to the specific requirements of the cross-community application context. Accordingly, the work presented in this thesis contributed to extending existing research on several main accounts:

- it has addressed an important but previously little investigated problem of supporting cross-community knowledge exchange from the HCI perspective and shown how this enriches and extends the existing approaches from knowledge management and computer-supported cooperative work,
- it has proposed an approach to supporting cross-community knowledge exchange based on using knowledge visualisation to support information access for solving ill-defined tasks in unfamiliar community domains,
- it has realized this approach by developing a novel method for eliciting and visualising implicit knowledge structures of individuals and communities in form of dynamic knowledge maps that make the elicited knowledge usable for supporting information access,
- it has shown how this method can be applied to support cross-community knowledge exchange by realizing an application framework, system and interface supporting typical tasks and needs of information access in unfamiliar community spaces,
- it has empirically evaluated the developed solutions in a comparative laboratory study which confirmed the adequacy of the solutions for the specific requirements of the cross-community application context.

10.1. Summary of Results and Contributions

The next sections summarize the achieved results and contributions in more detail.

10.1.1. Theoretical Analysis and Requirements Definition

In this work we have shown how knowledge visualisation can support cross-community knowledge exchange by facilitating information access in unfamiliar community spaces. Based on an extensive analysis of existing work regarding processes of collaborative creation and exchange of knowledge within communities and findings of studies of knowledge sharing across community boundaries, we have mapped out the main barriers for cross-community knowledge exchange and the limits of existing intra-community solutions for overcoming them (Chapter 2). This analysis points out the lack of a shared context of implicit understanding between different communities as the main problem of cross-community knowledge exchange. Different thought worlds and interpretive perspectives of different communities make it difficult for knowledge to cross community boundaries as information loses its original context and the recipients and authors do not share common vocabularies and implicit knowledge.

A set of requirements that need to be met in order to overcome these barriers has been developed by integrating an existing theoretical framework for cross-community knowledge exchange (perspective making – perspective taking) with practical requirements of typical knowledge management processes and relating them to sensemaking tasks occurring during information access in unfamiliar domains (Chapter 3). This synthesis suggests that visualising implicit knowledge structures of communities and supporting the discovery of relationships between them in a way that makes them usable as tools for information access in unfamiliar domains, could provide valuable support for cross-community discovery and sharing of knowledge. This has been the main hypothesis investigated in this thesis.

10.1.2. Method for Eliciting and Visualising Implicit Knowledge Structures of Individuals and Communities

Accordingly, we have developed a novel method for eliciting and visualising implicit knowledge structures of individuals and communities of users in form of dynamic knowledge maps that make them usable for semantic exploration and navigation of community information spaces (Chapter 6). Three main aspects distinguish this solution from other approaches. First, it addresses the problem of shared knowledge structures of communities being highly implicit, rarely explicitly expressed by communities themselves and composed of many different but interrelated points of view. The developed method enables the construction of both personal and community knowledge maps reflecting personal points of view and relationships between them. Instead of trying to connect structures of different communities into a single, unified structure, this supports the co-existence of different local perspectives and the discovery of relationships between them.

Second, the developed technique elicits personal knowledge structures in a way that is unobtrusively embedded into a user’s primary task. It follows the sensemaking paradigm that views the process of information seeking as a rich source of knowledge-intensive activities reflecting user’s personal knowledge (e.g. finding concepts for expressing an information need, creating representation schemas for organising information). The creation of personal maps is realized as an intrinsic part of users’ common activities of information access and requires no additional effort.

Finally, the resulting maps are not merely static structures but function as dynamic semantic templates allowing dynamic classification of information or structuring of unfamiliar community spaces from a specific point of view. This makes them usable both for the map author as well as for other users, which eliminates the cold-start and free-rider problems of traditional collaborative filtering and community rating approaches. Moreover, rather than complex ontology formalisms, the maps are modelled as light-weight conceptual structures. They are realized as interactive visual artefacts that can be shared and interactively manipulated by the users in order to generate and explore different views of unfamiliar information collections. This supports the inherently explorative nature of learning about unfamiliar knowledge.

The visualisation model combines Document Maps visualising main topics, document clusters and relationships between knowledge reflected in document collections comprising community spaces with Concept Maps visualising personal and shared conceptual structures of community members. The technical realization integrates methods for generation and visualisation of semantic overviews of unstructured information spaces based on self-organizing maps with extraction of word categories from texts and methods for collaborative indexing and personalised classification based on user-defined templates.

Such a combination of methods for generation of semantic overviews of document space with techniques involving the knowledge of human users enables the elicitation of semantic structures unavailable in other approaches. It allows the elicitation and visualisation of knowledge structures that incorporate personal views of individual members and visualise shared community vocabularies. Since they are based on the members' actual use of concepts in organizing information the resulting structures will evolve with the dynamics of personal use and the development of community knowledge. The application of visualisation techniques for representing similarity structures and relationships in complex information spaces makes the extracted structures usable for contextualized access to information from different points of view.

10.1.3. Application Framework and System for Cross-Community Knowledge Exchange

To demonstrate different ways in which such maps can be used for implicit and explicit exchange of knowledge between members of different communities, we have developed an application framework defining a set of typical use cases and supporting services for cross-community knowledge exchange. This is accompanied by a system architecture implementing the services of the framework, giving special attention to performance capabilities required for an interactive system (Chapter 7).

The framework relates the developed knowledge map method to specific tasks relevant for cross-community knowledge exchange identified by the requirements model. It demonstrates how visualising personal and shared knowledge structures of members of unfamiliar communities can support knowledge construction during access to the respective community information spaces. This includes the use of community maps for semantic exploration of unfamiliar community spaces, the use of personal maps for personalised classification and filtering of unfamiliar information as well as multi-perspective navigation for discovering how unfamiliar information is related to one's own knowledge, and vice versa.

A special aspect is a contextualized recommendation service that supports the sharing of personal knowledge by automatically matching personal maps of members from different communities to the need of a specific information task. In this way, ill-defined tasks requiring knowledge from an unfamiliar domain can be supported by providing relevant information in a user-specific context: both with respect to the needs of a given task as well as within its original context reflecting personal knowledge of the map author (documents contextualized within personal concept structures).

10.1.4. Visual Interface for Multi-Perspective Access to Community Information Spaces

The functionalities of the developed method and application framework are made available to the users by means of an interactive knowledge map tool and visual information interface for multi-perspective access to community information spaces, the Knowledge Explorer (Chapter 8). The Knowledge Explorer incorporates a multi-view visualisation model allowing simultaneous visualisation of different personal and community knowledge structures and their use for structuring, exploring and navigating community information spaces from different points of view. It presents a specific blend of capabilities unavailable in existing knowledge visualisation solutions, information workspaces and sensemaking-support interfaces (Chapter 4).

While other multi-perspective interfaces focus on specialized analysis tasks (Chapter 4) the Knowledge Explorer is intended as a powerful and yet simple to use visual information interface for normal users. It provides the functionalities of the underlying application framework (Chapter 7) in a way that naturally accompanies users' everyday activities of information access without distracting them from their primary task. Its design and implementation have been informed by insights from well-known information access task models and by findings of studies on knowledge construction during information seeking in unfamiliar domains. In this way the vast amount of experience on developing visual information interfaces could be productively applied, while considering the specific requirements of the cross-community application domain.

10.1.5. Results of the Empirical Evaluation

The empirical evaluation undertaken in a comparative laboratory study confirmed the adequacy of the proposed solution to supporting cross-community knowledge exchange. It has shown that the developed method is well-suited for eliciting and visualising implicit knowledge structures of individuals and communities in a way which makes them usable for facilitating information access. The results also provided supportive evidence for confirming the main hypothesis that constructing and visualising such personal and shared community knowledge maps can provide valuable support for cross-community discovery and sharing of knowledge.

More specifically, we were able to demonstrate that users using developed knowledge map support were able to accomplish ill-defined tasks requiring the identification and understanding of knowledge from an unfamiliar community information space, better than users of a standard information seeking system without knowledge visualisation. The evaluation results also demonstrated a better overall performance of users of personal maps than users of community maps. The latter confirms the importance of the method allowing the elicitation, visualisation and use of personal maps, which reflect both personal knowledge of individual community members and relate it to the information need of a specific task .

The subjective user feedback also demonstrated a positive assessment of the overall usability and perceived task-adequacy of the prototypical implementation of the Knowledge Explorer. Most users considered the Knowledge Explorer to be easy to use, easy to learn and well-suited for accomplishing the assigned tasks. In particular, the results of a direct subjective comparison with standard information seeking reference system show that users perceived the Knowledge Explorer to be better suited for accomplishing tasks occurring during cross-community knowledge access.

Finally, the developed evaluation framework with operative measures for assessing the quality of knowledge access in the cross-community application context is also a contribution of its own. It establishes a theoretically grounded and practically feasible method for evaluating the suitability and effectiveness of technological support for cross-community knowledge exchange mediated by information access in unfamiliar community spaces. In doing so it adapts and extends well known evaluation metrics used in information access (e.g. information retrieval precision and recall) with measures capturing specific aspects of the quality of knowledge access (e.g. quality of topical structuring, learning effect) and relates them to the cross-community context. This provides a basis that can be used and further refined by other researchers in order to develop and evaluate new solutions addressing this important but difficult and still little investigated problem.

10.1.6. Dissemination and Publication of Results

The results of this work have been presented and published on different occasions including international conferences, workshops, books and journals (see list of bibliographic references). Among others, this includes the best paper award at the 4th IEEE Workshop on Knowledge Media Networking (Novak et al., 2002) and the invited contribution to the Springer book “Knowledge and Information Visualisation – Searching for Synergies” (Novak & Wurst, 2005a) representing a first attempt at bridging the gap between different scientific communities addressing knowledge visualisation from inter-related but distinct perspectives. The basic ideas of the approach, the knowledge map method and the realization of the application framework were partially developed within the project “AWAKE – Networked Awareness for Knowledge Discovery” (BMBF, 2001-2003) bringing together two universities and two Fraunhofer Institutes in a research consortium led by the author. The results of this project were nominated for the cooperative research award of the Foundation for German Science (“Stiffterverband für die Deutsche Wissenschaft”, 2003).

The research experiences and the results of the thesis have also been transferred into studies for management consulting companies and international organizations such as the eCultureNetwork consortium, the UNESCO Digiarts Commission, the European Roadmap for Knowledge Management and eLearning of the Fraunhofer Society (Encarnação et al., 2003) and most recently into a new research project on knowledge tools and virtual services for interdisciplinary cooperation (i-WIZARDS; BMBF Research Programme "e-Science and Networked Knowledge Management", 2005).

10.2. Possibilities for Improvement and Further Research

10.2.1. Technical Optimization

Due to its prototypical character, the server-side implementation of the application framework leaves some room for improvement, largely with respect to optimization of the map processing and response times. One optimization possibility regards the format of data exchange between modules and map storage. Early on in the design process XML has been chosen as a universal exchange scheme between different modules of the system. While this simplified system development and communication between both different modules and different developers contributing to the implementation of the system, it incurred additional processing effort for building and parsing XML documents. For large document pools the XML representation of maps quickly bloats and the DOM trees become quite large. This results both in increased processing and transmission time between the server modules and the client.

In some cases, where response times are critical the use of XML has already been substituted by more compact data formats (e.g. the datapool service delivering document abstracts). Additional performance improvement in all server-side modules as well as in the client could be achieved by substituting the current XML-based map representation format with a more efficient representation. In particular, the MapManager service implementing an XML-based database storage for the maps could be replaced by a native relational database which would significantly speed up map retrieval and querying (e.g. by the client and by the recommender service).

This would also result in an improved performance of the personalised classification service which is a novel solution producing personalised document maps based on a Kohonen SOM in near real-time (Chapter 7.2.1.3). Another improvement here could also be achieved by implementing direct access by the service to the objects of the SOM produced by the SOM_PAK module, instead of passing results through intermediary files which is the default way of using the SOM_PAK package. In this way, it is quite likely that the current solution based on a combination of direct response and caching to satisfy the performance requirements of an interactive system could be replaced with direct real-time response.

Another interesting approach for scaling the system would be to consider a distributed peer-to-peer organization instead of the current client-server scheme. Instead of storing all maps only in a centralized database, the system could be organized as a collection of active user clients each containing its own map collection storage and indexing as well as local instances of the personalised classification and recommendation services. Both personal maps created by the user as well as maps fetched from other users or community maps from the central repository would be stored in the local collection and thus be immediately available for next use. In a similar way, the personalised classification service would process local maps faster while the recommender service could first check for local matches and then query other instanced for additional results.

A major conceptual benefit would be that users could apply the system also to structure and explore personal collections and not only shared community spaces. This is in accord with common patterns of use e.g in scientific communities or commercial organizations where individual users commonly build-up own personal collections often containing relevant portions of the shared community space but also portions collected personally from other sources that are not necessarily shared in the community space. Approaches such as (Mierswa & Wurst, 2005) have also shown possibilities for distributed processing of personal taxonomies for discovering topical relationships and performing classification tasks based on collaborative knowledge of a group of users operating in a peer-to-peer environment.

10.2.2. Visualisation and Interface

With respect to the Knowledge Explorer interface a better algorithm for label placement in order to avoid overlapping titles in high-density map areas should be implemented. User feedback indicated this as one of the few points of critique. Related to this is also the focus+context zoom functionality that could be improved by introducing a dynamic scaling factor, regulating the magnification level in focus and demagnification spread with distance based on the number of documents contained in the zoom focus. More extensive availability of tool tips and a help functionality are also straightforward improvements required by the users. The visualisation of search results and navigation selections could also be extended by including the display of phrases in which the search matches were found, which could be display in an additional window offering more detailed information on a given document set.

A more conceptually interesting issue would be to consider a closer integration of the concept map and document map visualisation. The evaluation has shown some user preference for the spatial visualisation model of the document map and several users have actively asked about the possibility to have the exploration and navigation capabilities integrated into one visual element. This is also inline with results of studies on the use of graphical overviews for access to text collections (Hearst, 1999) which suggest that users like to see more meta-information integrated into map displays than is commonly available.

In our case, one possibility may be to consider the use of overlaying techniques allowing the document map and concept map to be displayed as different layers of the same visual surface. An experimental approach to the use of overlaying in combination with semantic zooming has been presented in (Lieberman, 1997). Another interesting direction, could be to display only those portions of the concept map relevant to the semantic context of the current user focus on the document map (or vice versa). Overall, developing techniques for such visual integration of complex semantic structures remains a difficult challenge due to the inherent trade-off between the level of information semantics desired by the users and the required visual simplicity respecting the limits of cognitive strain.

Finally, the capabilities for user-defined information structuring in personal maps could be extended to allow multiple hierarchical levels. Though the current prototypical solution allowing essentially two-level (top map level and one-level folders) was adequate for assessing the viability of the approach and developed method focusing on explorative access, several users asked for the possibility to create sub-folders at different hierarchical levels. The ontological structure of the maps can be easily extended to allow this by simply nesting additional sub-cluster nodes within clusters. This can also be straightforwardly replicated in the tree-folder visualisation in which personal maps are created.

The main problems to be solved are the extension of the personalised classification service to hierarchical classification and the visual representation of multi-level cluster hierarchies in the 2D visualisation of the Document Map. The existing realization of the personalised classification based on kNN-classification (k-nearest neighbour) can be readily applied by simply considering all clusters as equally valid categories regardless of the hierarchical level. Documents classified to a given sub-cluster will be naturally linked to parent clusters due to the ontological structure of the map model (nested clusters are connected by parent-child relations). The visual complexity of displaying multi-level clusters in the 2D Document Map could be solved by displaying multiple hierarchical levels in the folder-tree visualisation while retaining only the first level cluster in the 2D Document Map visualisation. Information on sub-hierarchies could be provided interactively on user-demand (e.g. by sub-cluster zoom).

10.2.3. Method Application

The developed method for knowledge visualisation and exchange through personalised and community knowledge maps has been applied to community spaces represented as textual collections. An interesting extension would be the application to multimedia collections, including images, audio and video objects. This is interesting on one hand in the commercial application contexts where information and knowledge from working processes are increasingly documented and exchanged in multimedia form. On the other hand, due to the widespread availability of general use multimedia devices (e.g. digital cameras, mp3 sticks and camera phones) the capturing and sharing of multimedia data has become widespread also in more general communities of interest in leisure and entertainment (e.g. blogs, photo and music sharing).

Since the developed knowledge map method for eliciting personal structures and connecting them to each other relies primarily on context-similarity information (e.g. objects frequently co-occuring in user-defined groups, Chapter 6.5.1) it can be readily applied to multimedia data as well. The main modification to be made includes the substitution of text-analysis for encoding object similarities in the bootstrapping phase with a method relying on multimedia meta-data provided by human descriptions or by feature extraction from media content. Some experiments in this direction have been undertaken in the Nemoz project (Network Media Organizer)[76], which uses collaborative similarity measures based on personal taxonomies of different users in order to classify, filter and recommend music files based on different points of view.

10.2.4. Longitudinal Evaluation

However appealing the further investigation of different technical aspects both at method, visualisation and interface level may be, probably the most interesting challenge for further research is the need for longitudinal evaluation in day-to-day real-world use.

The laboratory evaluation undertaken in this work has provided valuable insights into the possible contributions of using knowledge visualisation to support cross-community knowledge exchange. It is the first evaluation of a sophisticated knowledge visualisation system in this specific application context (to the authors' best knowledge) and as such offers valuable points of departure for further work in the field. However, the relevance of these findings is limited by the method of the laboratory setting. This concerns in particular the possibilities of measuring the quality of cross-community knowledge access on aspects such as learning effect or cross-community concept transfer. In order to better investigate such issues a longitudinal evaluation analysing the application of the proposed approach or similar techniques in day-to-day use in real-world contexts is required.

In order to fruitfully perform such an evaluation the method and the system should have reached a certain level of maturity, having been sufficiently examined in previous usability and task-adequacy studies. The evaluation framework should have been tested and demonstrated its applicability to uncover the desired effects and the nature of possible test settings and classes of tasks relevant in practice should have been identified. The methods, the system and the evaluation framework developed in this work have aimed at achieving these goals. It is our hope that in doing so we have provided a fruitful ground for further work on this challenging and promising research topic.

[76] http://nemoz.sourceforge.net/web/Nemoz_Project

References

Ackermann, F., Eden, C., Cropper, S.: Getting Started with Cognitive Mapping, Management Science, University of Strathclyde, 1993-1996
http://www.banxia.com/depaper.html

Aha, D., Kibler D., and Albe, M: Instance based learning algorithms. Machine Learning, 6:37--66, 1991.

Ahlberg, M.: Varieties of Concept Mapping. In Concept Maps: Theory, Methodology, Technology Proc. of the First Int. Conf. on Concept Mapping, A. J. Cañas, J. D. Novak, F. M. González (Eds.), Pamplona, Spain 2004

Albert, D., Steiner, C.M.: Empirical Validation of Concept Maps: Preliminary Methodological Considerations. In Proc. of Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@ICALT'05), Kaohsiung, Taiwan, July 2005
http://www.win.tue.nl/SW-EL/2005/swel05-icalt05/final/W3-5.pdf

Allee, V.: Knowledge Networks and Communities of Practice. In OD Practitioner, Journal of the Organization Development Network, Vol. 32, No. 4, 2000
http://www.odnetwork.org/odponline/vol32n4/knowledgenets.html

Alpert, S.R., Grueneberg, K., Concept Mapping with Multimedia on the Web. Journal of Educational Multimedia and Hypermedia, 9(4), 313-330, 2001

Amento, B., Terveen, L. G., Hill, W.Experiments in Social Data Mining: The TopicShop System, ACM Transactions on Computer-Human Interaction, Vol. 10, No. 1, pp. 54–85, March 2003.

Argyris, C., Schön D.A.: Organizational Learning II, Addison-Wesley, Reading, 1996

Argyris, C.; Schön D.A.: Organizational Learning, Addison-Wesley, Reading, 1978

Arias,, E.G., Fischer, G.: Boundary Objects: Their Role in Articulating the Task at Hand and Making Information Relevant to It, International ICSC Symposium on Interactive & Collaborative Computing (ICC'2000), University of Wollongong, Australia, ICSC Academic Press, Wetaskiwin, Canada, 2000

Baeza-Yates, R., Ribeiro-Neto, B., (Eds.) Modern Information Retrieval, Chapter 10, Addison Wesley, 1999.

Baldonado, M., Woodruff, A., Kuchinskiy, A.: Guideliness for using multiple views in information visualisation. Proc. of ACM Conference on Advanced Visual Interfaces (AVI'00), 2000.

Baldonado, M.Q.W, and Winograd, T., SenseMaker: An Information-Exploration Interface Supporting the Contextual Evolution of a User's Interests, Proc. CHI 97, ACM Press, 1997

Bates, M.J.: The design of browsing and berrypicking techniques for the online search interface. Online Review, 13(5):407.424, 1989

Becks, A., Seeling, C.: SWAPit: A Multiple Views Paradigm for Exploring Associations of Texts and Structured Data. Proceedings of AVI 2004 Int. Working Conf. on Advanced Visual Interfaces, Gallipoli, Italy, May 2004.

Becks, A.: Visual Knowledge Management with Adaptable Document Maps, Doctoral Dissertation, RWTH Aachen /GMD Research Series, GMD – German National Research Center for Information Technology, 2001, ISBN 3-88457-398-5

Belkin, N.J., Stein, A., Thiel, U.: Cases, Scripts and Information Seeking Strategies. On the Design of Interactive Information Retrieval Systems. Expert Systems with Applications, Vol. 9(3):pp. 379--395, 1995.

Bentley, R., Appelt, W., Busbach. U., Hinrichs, E., Kerr, D., Sikkel, S., Trevor, J., Woetzel, G.: BasicSupport for Cooperative Work on the World Wide Web. In: International Journal of Human-Computer Studies, 46, Special issue on Innovative Applications of the World Wide Web, 1997

Berger P.L, Luckmann T.: The Social Construction of Reality: a Treatise in the Sociology of Knowledge. Doubleday, 1966

Bly, S., Harrison, S.R., Irwin, S.: Media spaces: Bringing people together in a video, audio, and computing environment. Communications of the ACM, vol. 36, no. 1, pp. 28–47, 1993

Bobrow, D.G., Whalen, J.: Community Knowledge Sharing in Practice: The Eureka Story, In: Reflections, Journal of the Society for Organizational Learning, Vol. 4, No. 2, 2002

Boissevain, J.: Friends of Friends. Basil Blackwell, 1974

Boland J.R., Tenkasi R.V.: “Perspective Making and Perspective Taking in Communi-ties of Knowing”, Organization Science, 6,4 (July-August) 1995

Bonifacio M., Bouquet P., Cuel, R.:The role of classication(s) in distributed knowledge management, Proceedings of 6th International Conference on Knowledge-Based Intelligent Information Engineering Systems & Allied Technologies (KES'2002), Special Session on Classification.

Bonifacio, M., Bouquet, P., Lanzardo, A.: A distributed intelligence paradigm for knowledge management. In Proc. of AAAI Spring Symposium: Bringing Knowledge to Business Processes, pages 69–76. AAAI Press, March 2000.

Bonifacio, M., Bouquet, P., and Traverso, P.: Enabling distributed knowledge management. Managerial and technological implications. Novatica and Informatik/Informatique III(1), 2002

Booker, A., Condliff, M., Greaves, M., Holt, F., Kao, A., Pierce, D., Poteet, S., Jason Wu, Y.: Visualising Text Data Sets. IEEE Computing in Science & Engineering, pp. 26-35, July/August, 1999

Borghoff, U. M., Pareschi, R. (Eds.): Information Technology for Knowledge Management, Springer Verlag, Berlin, 1998.

Borghoff, U.M., Schlichter, J. H.: Computer-Supported Cooperative Work. Springer, 2000.

Borlund, P., Ingwersen, P.: The development of a method for the evaluation of interactive information retrieval systems. Journal of Documentation, 53(3):225.250, 1997

Borlund, P.: Experimental components for the evaluation of interactive information retrieval systems. Journal of Documentation, 56(1):71.90, 2000

Bowker, G.C., Star, S. L..:Sorting Things Out: Classification and its Consequences. Cambridge: MIT Press, 1999

Bradshaw, J.M., Ford, K.M., Adams-Webber, J.R., Boose, J.H.: Beyond the repertory grid: new approaches to constructivist knowledge acquisition tool development. International Journal of Intelligent Systems 8(2) 287-33, 1993

Brank, J. et al.:A survey of ontology evaluation techniques. In Proc. of Conference on Data Mining and Data Warehouses (SiKDD 2005), Ljubljana, Slovenia, October 2005
http://kt.ijs.si/dunja/sikdd2005/Papers/BrankEvaluationSiKDD2005.pdf

Brown J. S., Duguid P.:The Social Life of Information. Boston, MA, Harvard Business School Press, 2000

Brown, J.S., Duguid, P.: Knowledge and Organization: A Social-Practice Perspective. Organization Science. Vol 12 (2) pp198 – 213, 2001

Brown, J.S., Duguid, P.: Organizing Knowledge. California Management Review, Vol. 40 Issue 3 (spring), pp. 90-111, 1998

Brown, J.S., Duguid, P.: Organizational Learning and Communities of Practice: Toward a Unified View of Working, Learning, and Innovation, in Organization Science, Vol.2, No. 1, pp. 40-57, February, 1991

Bruckmann, A., Resnick, M.: Virtual Professional Community: Results from the MediaMOO project, Proceedings of 3CyberConf, The Third International Conference on Cyberspace, May 1993.

Bruillard, E., Baron, G.L.: Computer-Based Concept Mapping: a Review of a Cognitive Tool for Students. In D. Benzie, D. Passey (Eds.), Proceedings of Conference on Educational Uses of Information and Communication Technologies (ICEUT 2000) (pp. 331-338). Beijing, Publishing House of Electronics Industry (PHEI), 2000

Bruner, J., Acts of Meaning, 1990.

Bryson., J.M. et al.: Using the Oval Mapping Process to Identify Strategic Issues and Formulate Effective Strategies. In: Bryson J (ed) Strategic Planning for Public and Nonprofit Organisations. San Francisco, Jossey-Bass, pp 257 – 275, 1995

Bullinger, H., Baumann, T., Fröschle, N., Mack, O., Trunzer, T., Waltert, J.: Business Communities. Galileo Business, Galileo Press, Bonn, Germany, January 2002.

Burkhard, R.A.: Towards a Framework and a Model for Knowledge Visualisation: Synergies between Information and Knowledge Visualisation, In: Knowledge and Information Visualisation – Searching for Synergies, Tergan, S.-O., Keller, T. (Eds.), Springer, Heidelberg, 2005

Cañas, A. J., Ford, K. M., Novak, J. D., Hayes, P., Reichherzer, T. R., Suri, N.: Online Concept Maps: Enhancing collaborative learning by using technology with concept maps. The Science Teacher, 68(2), pp. 49-51, 2001.

Cañas, A.J. et al.: Concept Maps: Integrating Knowledge and Information Visualisation. In Knowledge and Information Visualisation: Searching for Synergies,Tergan, S.O., Keller, T. (Eds), Springer Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2005

Cañas, A.J., Ford, K., Coffey, J.: Concept Maps as a hypermedia navigational tool. Paper presented at the seventh Florida Artificial Intelligence Research Symposium, Pensacola (FL), USA, 1994

Cañas, A.J., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T., et al.: CmapTools: A knowledge modeling and sharing environment. In A.J. Cañas, J.D. Novak, & F.M. González (Eds.), Concept maps: Theory, methodology, technology. Proceedings of the 1st international conference on concept mapping (pp. 125-133). Pamplona, Spain: Universidad Pública, de Navarra, 2004

Cañas, A.J., Leake, D.B., Wilson, D.C.: Managing, mapping, and manipulating conceptual knowledge. In Proceedings of the AAAI-99 Workshop on Exploring Synergies of Knowledge Management and Case-Based Reasoning, pp. 10-14, Menlo Park, AAAI Press, 1999.

Card, S. K., Robertson, G. C., Mackinlay, J. D.:The information visualiser, an information workspace. In Proceedings of CHI'91 (New Orleans, April), ACM, New York, 1991.

Card, S.K., Mackinlay, J.D., Shneiderman, B.: Information Visualisation. In S.K. Card, J.D. Mackinlay, B. Shneiderman: Readings in Information Visualisation: Using Vision to Think, Morgan Kaufmann Publishers, San Francisco, 1999a

Card, S.K., Mackinlay, J.D., Shneiderman, B.: Readings in Information Visualisation: Using Vision to Think, Morgan Kaufmann Publishers, 1999b

Carotenuto, L., Etienne, W., Fontaine, M., Friedman, J. Newberg, H., Muller, M., Simpson, M., Slusher, J., Stevenson, K., Communityspace: Towards flexible support for voluntary knowledge communities. In Proc. Workshop on Workspace Models for Collaboration, London, April, 1999

Carvalho, M., Hewett, R., Cañas, A.J.: Enhancing web searches from concept-map based knowledge models. In Proceedings of the SCI. Fifth Multi-Conference on Systems, Cybernetics and Informatics. Orlando, FL, 2001

Catarci, T. et al.: An Ontology Based Visual Tool for Query Formulation Support. In: Proc. of On The Move to Meaningful Internet Systems (OTM 03). Springer-Verlag Heidelberg, October 2003.

Chalmers, M., Rodden, K., Brodbeck, D.: The Order of Things: Activity-Centred Information Access. In Proc. WWW7, Brisbane, April 1998.

Chen, C. & Carr, L.: Visualising the evolution of a subject domain: A case study. Proceedings of IEEE Visualisation 99. San Francisco, California, USA. pp. 449-452, October, 1999.

Chen, C., Yu.Y: Empirical studies of information visualisation: a meta-analysis, International Journal of Human-Computer Studies, Vol. 53, No. 5, pp. 851-866, 2000

Chen, C.: Information Visualisation and Virtual Environments. Springer, 1999

Chen, C.: Structuring and visualising the World-Wide Web with Generalized Similarity Analysis. Proceedings of the 8th ACM Conference on Hypertext (Hypertext '97). Southampton, UK, April, 1997

Chen, C.: Visualisation of Knowledge Structures, In: Handbook of Software Engineering and Knowledge Engineering, 2002

Chen, C.:Visualising semantic spaces and author co-citation networks in digital libraries, Information Processing and Management 35, 401-420, 1999a

Chiu, C-H., Huang, C-C., Chang, W-T., Liang, T-H., Interaction Processes in Network Supported Collaborative Concept Mapping, In Proc. of ACM ITiCSE '99, Cracow, Poland, 1999

Christophides, V. et al.: Community Webs (C-Webs): Technological Assessment and System Architecture, Technical Report, EU-IST C-Web project deliverable, 2000

Cody, W. F., Kreulen, J. T., Krishna, V., Spangler, W. S.: The integration of business intelligence and knowledge management, IBM SYSTEMS JOURNAL, VOL 41, NO 4, 2002

Coffey, J. W., Hoffman, R. R., Cañas, A. J., Ford, K. M.: A concept-map based knowledge modeling approach to expert knowledge sharing. In M. Boumedine (Ed.), Proceedings of IKS 2002 - the IASTED international conference on information and knowledge sharing (pp. 212-217). Calgary, Canada: Acta Press, 2002

Cohen, D., Prusak, L.: In Good Company: How Social Capital Makes Organizations Work, Harvard Business School Press, Boston, MA, 2001

Cohendet, P., Creplet, F., Dupouët, O.: Communities of Practice and Epistemic Communities: A Renewed Approach of Organisational Learning within the Firm, Proc. of the project "Nouvelle Economie du logiciel", 2001 http://www.marsouin.org/IMG/pdf/dupouet.pdf

Conklin, J., Begeman, M.: gIBIS: A hypertext tool for exploratory policy discussion, Proc. of CSCW 98, Portland (OR), USA, 1998

Cook, S., Brown, J.S.: Bridging Epistemologies: The Generative Dance Between Organisational Knowledge and Organisational Knowing. Organisation Science 10, 4, pp. 381-400, 1999

Cousins et al.: The Digital Library Integrated Task Environment (DLITE). In Proc. of 2nd ACM International Conference on Digital Libraries, pp. 142-151, Philadelphia (PA), USA, 1997

Cox, R.: Representation, construction, externalized cognition and individual differences. Learning and Instruction, 9, pp. 343-363, 1999

Crossley, J.D., McGrath, A., Rejman-Greene, M.: The Knowledge Garden, In Proc. of the 2nd International Conf. on Practical Aspects of Knowledge Management (PAKM99), 1999.

Cutting, D. R. et al.: Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections. In: Proceedings of the Fifteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 318-329, Kopenhagen, 1992.

Dansereau, D.F.: Node-Link Mapping Principles for Visualising Knowledge and Information, In Knowledge and Information Visualisation: Searching for Synergies,Tergan, S.O., Keller, T. (Eds), Springer Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2005

Davenport, T.H., Prusak, L., Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, Boston, MA, 1998.

Davidson, G.S., Hendrickson, B., K.Johnson, D., Meyers, C.E., Wylie, B.N. Knowledge mining with VxInsight: Discovery through interaction. Journal of Intelligent Information Systems, 11 (3), pp. 259-285, 1998

Davidson, M.L.: Multidimensional Scaling, John Wiley & Sons, 1983

Davidson, R., Harel, D.: Drawing graphs nicely using simulated annealing. ACM Transactions on Graphics, 15(4):301-31, 1996

Davies, J. et al: Ontoshare – a knowledge management environment for virtual communities of practice. In Proc. of K-CAP'03, Sanibel Island, Florida, USA, October 2003

Deerwester, S., Dumais, S. T., Landauer, T. K., Furnas, G. W., & Harshman, R. A. (1990). Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science, 41(6), 391-407.

DeMichelis, G.: Cooperation and Knowledge Creation, in Nonaka.I, Nishiguchi, T. (eds), Knowledge Emergence: social, technical and evolutionary dimensions of knowledge creation, New York, Oxford University Press, 2001

Dervin, B.: From the Mind's Eye of the User: the Sense-Making Qualitative-Quantitative Methodology. In Glazier, J.D. and Powell, R.R. (Eds.) Qualitative Research in Information Management, 61-84, 1992

Dillon, A., McKnight, C., Richardson, J.: Navigating in hypertext: A critical review of the concept. In: EAD Diaper (ed.), Human-Computer Interaction – INTERACT 90, Elsevier (Oxford, UK), 1990

Ding, Y., Foo, S.: Ontology research and development. Part 2 – a review of ontology mapping and evolving Journal of Information Science, 28(5): pp. 375 – 388, October 2002

Donath, J. , Karahalios, K. , Viegas, F.: Visualising Conversation. In: Proc. Hawaii International Conf. on System Sciences (HICSS-32), 1999

Dougherty, D.: Interpretative barriers to successful product innovation in large firms. Organization Science, 3(2), 1992

Dourish, P., Lamping, J., Rodden, T.: Building bridges: Customisation and mutual intelligibility in shared category management. In Proceedings of the ACM Conference on Supporting Group Work (GROUP '99), November 1999.

Dourish, P.: Extending awareness beyond synchronous collaboration. In T. Brinck and S.E. McDaniel (eds.): CHI'97 Workshop on Awareness in Collaborative Systems, Atlanta, Georgia, 22–27 March 1997.

Eades, P.: A heuristic for graph drawing. Congressus Numerantium, 42: 149-60, 1984

Eden C., Ackermann F.: Making Strategy: The Journey of Strategic Management. London, Sage Publications, 1998.

Eden, C. Ackermann, F., A Mapping Framework for Strategy Making, In: Mapping Strategic Knowledge, Huff, A.S.,, Jenkins, M. (eds.), Sage Publications, 2002

Encarnação, J.L., John, M., Bendig, T., Boris, G. (ed.): iKnowS - Research Roadmap for a European e-Learning and Knowledge Management Infrastructure, Fraunhofer ICT Group, June 2003, ISBN: 3-8167-6318-9

Eppler, M.J.: Making Knowledge Visible Through Intranet Knowledge Maps: Concepts, Elements, Cases, In Proc. of 34th Hawaii International Conference on System Sciences (HICSS-34), Maui, Hawaii, January 2001

Eppler, Martin., Burkhard, Remo.: Knowledge Visualisation: Towards a New Discipline and its Fields of Applications. Working Paper, Universität Lugano (USI), Universität St. Gallen (HSG) , 2004 http://www.wirtschaftsinformatik.de/wi_arbeitsberichte.php?op=anzeigearbeitsbericht&abid=142

Erickson, T., Kellogg, W.A.: Knowledge Communities: Online Environments for Supporting Knowledge Management and its Social Context. In Beyond Knowledge Management: Sharing Expertise. (eds. Ackerman, Mark, Volkmar Pipek, and Volker Wulf). Cambridge, MA, MIT Press, 2001.

Erickson, T., Smith, D.N., Kellogg, W.A., Laff, M.R., Richards, J.T., Bradner, E.: Socially translucent systems: Social proxies, persistent conversation, and the design of Babble. CHI 99 Conference Proceedings: Human Factors in Computing System, 72-79, ACM Press, 1999.

Faloutsos, C., Lin, K.I.: FastMap: A Fast Algorithm for Indexing, Data-Mining and Visualisation of Traditional Multimedia Datasets. In Proc. of ACM SIGMOD 95, ACM Press, 1995

Fensel, D. et al.: OIL in a Nutshell. In: Proceedings of the 12th European Workshop on Knowledge Acquisition, Modeling and Management. Springer-Verlag, 2000, ISBN 3. 540.41119.4, 1.16

Fensel, D. et al.: Ontobroker: The Very High Idea. In Proc. of the 11th International Florida Artificial Intelligence Research Society Conference. AAAI Press, 1998

Fernandez, M., Gomez-Perez, A., Juristo, N. Methontology: From ontological art toward ontological engineering. In Spring Symposium Series, Stanford University. Stanford, CA., 1997.

Fillenbaum, S., Rapoport, A.: Structures in the subjective Lexicon.New York: Academic Press, 1971

Fischer, F., Mandl, H.: Facilitating the construction of shared knowledge with graphical representation tools in face-to-face and computer-mediated scenarios. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), Proceedings of Euro-CSCL 2001 (pp. 230-236). Maastricht: McLuhan Institute, 2001

Fischer, G.:External and shareable artifacts as opportunities for social creativity in communities of interest, in J. S. Gero and M. L. Maher (eds), Proc. Computational and Cognitive Models of Creative Design, Key Centre of Design Computing and Cognition, University of Sydney, Heron Island, Australia, December 2001.

Fleck, L.:Genesis and Development of a Scientific Fact, Trenn, T.J. and Merton, R.K. (Eds), translated by Bradley, F. and Trenn, T.J., The University of Chicago Press, Chicago, IL, 1979

Foner, L.N.: Political Artifacts and Personal Privacy: The Yenta Multi-Agent Distributed Matchmaking System. PhD Thesis, Massachussets Institute of Technology, June 1999.

Fowler, R. H., Fowler, W. A. L., Wilson, B. A.: Integrating Query Thesaurus and Documents through a common visual Representation. In: Proc. of ACM SIGIR '91, ACM Press, 1991

Fruchterman, T.M.J., Reingold, E.M.: Graph drawing by force-directed placement. Software Practice and Experience, 21:1129-64, 1991.

Fuchs-Kittowski, F., Stahn, P., Walter, R.: Wissensmanagement und E-Collaboration - Ein Framework für Communities, Teams und Netze zur Unterstützung kooperativer Wissensarbeit?, In Proc. KnowTech 2003 - 5. Conference on Knowledge Engineering & Management, München, October 2003.

Fuhr, D., Fuchs-Kittowski, F.: Against Hierarchy and Chaos: Knowledge Coproduction in Nets of Experts, J.UCS: Journal of Universal Computer Science, Vol. 10, No. 3, March 2004

Fuhr, N., Klas, C.P., Schaefer, A., Mutschke, P.: Daffodil: An integrated desktop for supporting high-level search activities in federated digital libraries. In Research and Advanced Technology for Digital Libraries. 6th European Conference, ECDL 2002, pp. 597-612, Springer, 2002

Furnas, G. W.: Generalized fisheye views. In Proc. of CHI '86. ACM Press, New York, 1986.

G. A. Hillery. Definitions of community: Areas of agreement. Rural Sociology, 20:111–123, 1955.

Gaines, B., Shaw, M.L.G: Supporting the Creativity Cycle through Visual Languages, in Proc. of AAAI Spring Symposium: AI and Creativity, Menlo Park, California, 1993

Gaines, B.R., Shaw, M.L.G.: Basing knowledge acquisition tools in personal construct psychology. Knowledge Engineering Review 8(1) 49-85, 1993a

Gaines, B.R., Shaw, M.L.G.:WebGrid: Knowledge Modeling and Inference through the World Wide Web, In Proceedings of KAW'96, 10th Knowledge Acquisition Workshop Banff, Alberta, Canada, November 1996 http://ksi.cpsc.ucalgary.ca/KAW/KAW96/gaines/KMD.html.

Gaines, B.R.,M.L.G. Shaw. Using knowledge acquisition and representation tools to support scientific communities. Proc. of the AAAI-94, 12th National Conference on Artificial Intelligence, Menlo Park, California, USA, AAAI Press/MIT Press, 1994

Gaines, B.R., Shaw, M.L.: Knowledge Acquisition Tools based on Personal Construct Psychology, in Special Issue on “Automated Knowledge Acquisition Tools” of the Knowledge Engineering Review, 1992

Gangemi, A. et al.: A theoretical framework for ontology evaluation and validation. In Prof. of 2nd Italian Semantic Web Workshop, University of Trento, Trento, Italy, December 2005
http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-166/9.pdf

Garnsey, E., Wright, S.: Technical innovation and organizational opportunity." International Journal of Technology Management, 5(3): 267-291, 1990

Garton, L., Haythornthwaite, C., Wellman, B.: Studying Online Social Networks, Journal of Computer Mediated Communication, 3(1), June 1997
http://jcmc.huji.ac.il/vol3/issue1/garton.html

Genesereth, M., Fikes, R.: Knowledge Interchange Format Reference Manual - Version 3.0., CSD Tech-Report Logic 92-1, Stanford University, 1992 .

Glance, N., Arregui, D., Dardenne, M.: Knowledge pump: Supporting the flow and use of knowledge, in U. Borghoff, R. Pareschi (eds.), Information Technology for Knowledge Management, Chapter 3, Springer-Verlag, 1998.

Gloor, P.: Collaborative Knowledge Networks, eJETA electronic Journal on E-Business Technologies and Applications, Vol 1, No 3, 2002.
http://www.ejeta.org

Gongla, P., Rizzuto, C.R.: Evolving communities of practice: IBM Global Services experience, IBM Systems Journal, Vol. 40, No. 4, 2001

Granovetter, M. (1973). The strength of weak ties. American Journal of Sociology, 78, 1360-1380.

Gräther, W., Prinz, W.: The Social Web Cockpit - Support for Virtual Communities. In Proc. ACM Conference on Supporting Group Work - GROUP 01, Boulder, Colourado, USA, 2001

Grey, D.: Knowledge Mapping :A Practical Overview, SWS e-Journal Talking Culture, Talking Knowledge, Smith Weaver Smith Inc, March, 1999

Gruber, T.R.: A translation approach to portable ontology specifications. Knowledge Acquisition, 5 (2):199–220, 1993.

Gruninger, M., Lee, J.: Ontology applications and design. Communications of the ACM, 45(2), p. 39 – 41, 2002

Guarino, N.: Formal Ontology and Information Systems. In: Guarino, N. (Ed.): Proceedings of the 1st International Conference on Formal Ontologies in Information Systems, FOIS'98. Trento, Italy: IOS Press, June 1998

Guarino, N.: Formal Ontology, Conceptual Analysis and Knowledge Representation. In: International Journal of Human-Computer Studies, 43, Nr. 5-6, 1995.

Hartmann, J. et al.:Methods for ontology evaluation. EU-IST Knowledge Web Project, Deliverable D1.2.3, 2004 http://www.aifb.uni-karlsruhe.de/Publikationen/showPublikation?publ_id=1021

Hearst, M. A.: User interfaces and visualisation. In R. Baeza-Yates, B. Ribeiro-Neto, (Eds.) Modern Information Retrieval, Chapter 10, pp. 257-323. Addison Wesley, 1999.

Hearst, M., Karadi, C.:Cat-a-cone: An interactive interface for specifying searches and viewing retrieval results using a large category hierarchy. In Proc. of 20th ACM/SIGIR Conf., Philadelphia, PA, July 1997.

Hearst, M.: Tile Bars: Visualisation of term distribution information in full text information access. In Proc. of CHI 95, ACM Press, 1995

Hearst, M.: T User Interfaces & Visualisation for Information Access, Tutorial at ACM SIGIR 2000 http://bailando.sims.berkeley.edu/talks/sigir-viz-tutorial-2000.ppt

Heflin, J., Hendler, J.: Searching the Web with SHOE. In: Artificial Intelligence for Web Search. In Proc. of the AAAI Workshop WS-00-01. AAAI Press, 2000

Heinrich, L., Roithmayr, F.:Wirtschaftsinformatik-Lexikon (6th ed). Munich and Vienna, Oldenbourg Verlag, 1998

Henry Lieberman, A Multi-Scale, Multi-Layer, Translucent Virtual Space, IEEE International Conference on Information Visualisation, London, September 1997

Herlocker, K., Konstan, J., and Reidl, J.: Explaining Collaborative Filtering Recommendations, in Proc. of CSCW 2000, Philadelphia, PA, USA, December 2000

Herman, I., Melancon, G., Marshall, M.: Graph visualisation and navigation in information visualisation: a survey, In IEEE Transactions on Visualisation and Computer Graphics 6(1), pp. 24-43, 2000

Hetzler, B., Harris, W. M., Havre, S., Whitney, P.: Visualising the full spectrum of document relationships. Proceedings of the 5th International ISKO Conference: Structures and Relations in the Knowledge Organization Lille, 1998a

Hetzler, B., Whitney, P., Martucci, L., Thomas, J.:. Multi-faceted insight through interoperable visual information analysis paradigms. Proceedings of IEEE Information Visualisation '98, 1998b

Hill, W. C., Terveen, L. G.: Using frequency-of-mention in public conversations for social filtering. In Proceedings of CSCW'96, pp. 106–112, Boston, November, 1996.

Hodge, G.: Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files, CLIR Pub91, April 2000. http://www.clir.org/pubs/reports/pub91/pub91.pdf

Honkela, T., Kaski, S., Krista, L., Kohonen, T.: Newsgroup Exploration with WEBSOM Method and Browsing Interface. Technical Report A32, Helsinki University of Technology, Laboratory of Computer and Information Science, 1996

Honkela, T., Kaski, S., Lagus, K., Kohonen, T.:WEBSOM – Self-organizing maps of document collections. In Proc. of WSOM'97, Workshop on Self-Organizing Maps, Espoo, Finland, pp. 310-315, 1997

Honkela, T., Self-organizing maps in natural language processing, PhD thesis, Helsinki University of Technology, 1997.

Hoxmeier, J.A., DiCesare, C., 2000, System response time and user satisfaction: an experimental study of browser-based applications. Proceedings of the Americas Conference on Information Systems, Long Beach, California: Association for Information Systems, August 2000.

Huff, A. S., Jenkins, M. (Eds.):Mapping strategic knowledge. London: Sage, 2002

Improving Performance in Small and Medium Sized Enterprises Through Collaborative Innovation Networks, International Research Group Manifesto, University of Cologne, MIT - Massachusetts Institute of Technology, Helsinki University of Technology et al.
http://www.wim.uni-koeln.de/Internationale_Forsc.447.0.html

Ishida, T.: Community Computing. John Wiley and Sons, 1998

Joachims, T.: Text categorization with Support Vector Machines: Learning with many relevant features. In Proc. of 10th European Conference on Machine Learning (ECML'98), 1998.

Jonassen, D.H., Beissner, K.,Yacci, M. (Eds.).: Structural knowledge. Techniques for representing, conveying, and acquiring structural knowledge. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993

Jonassen, D.H., Reeves, T.C., Hong, N., Harvey, D., & Peters, K.: Concept mapping as cognitive learning and assessment tools. Journal of Interactive Learning Research, 8(3/4), pp. 289-308, 1997

Jones, S.: Graphical Query Specification and Dynamic Result Previews for a Digital Library. In Proc. of ACM UIST'98, ACM Press, 1998.

Kautz, H., Selman, B., Shah, M.: Referral Web - Combining Social Networks and Collaborative Filtering, Communications of the ACM, Vol. 40, No. 3, pp. 63—65, 1997

Keller, T., Tergan, S.O.: Visualising Knowledge and Information: An Introduction. In Tergan, S.O., Keller, T. (Eds.): Knowledge and Information Visualisation – Searching for Synergies, Springer LNCS 3426, Springer, 2005

Kelly, G.A., The Psychology of Personal Constructs. New York: Norton, 1955.

Kleinberg, J. M.: Authoritative sources in a hyperlinked environment. In Proceedings of ACM-SIAM Symposium on Discrete Algorithms, San Francisco, January, 1998.

Koch, J.H. Unterstützung der Formierung und Analyse von virtuellen Communities. PhD Thesis, Faculty of Computer Science, Technical University Munich, July 2002.

Koch, M.: Community-Unterstützungssysteme - Architektur und Interoperabilität, Technical University Munich, Faculty of Computer Science, December 2003

Köhntopp, M.: Identitätsmanagement. In: Bäumler, H. (Hrsg.) , Breinlinger, A., Schrader, H. (Eds.): Datenschutz von A-Z, Luchterhand, Neuwied, 2000

Kohonen, T., Hynninen, J., Kangas, J., Laaksonen, J.: SOM_PAK: The Self-Organizing Map Program Package. Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 1996.

Kohonen, T., Kaski S., et al.: Self Organization of a Massive Document Collection, IEEE Transactions on Neural Networks, Vol. 11, No. 3, May 2000.

Kohonen, T.: Self-Organizing Maps. Springer Verlag, Berlin, 2nd edition, 1995.

Kohonen. T.: The Self-Organizing Map. Proceedings of the IEEE, 78(9):1464-1480, 1990.

Kommers, P., Lanzing, J.: Student´s concept mapping for hypermedia design. Navigation through the world wide web (WWW) space and self-assessment. Journal of Interactive Learning Research 8(3/4), pp. 421-455, 1997

Kriewel, S., Klas, C.P., Schaefer, A., Fuhr, N.:Daffodil - Strategic Support for User-Oriented Access to Heterogeneous Digital Libraries, D-Lib Magazine, Vol. 10, No. 6, 2004

Kunz, C.: Ein integrierter Ansatz zur wissensbasierten Informationsrecherche, Doctoral Dissertation, Institute for Human Factors and Technology Management, University of Stuttgart, December 2005

Lacher, M.: Supporting the Exchange of Knowledge in Communities of Interest via Document Catalog Mediation, PhD Thesis, Dept. Computer Science, Technical University Munich,June 2003

Lamping, J., Rao, R., Pirolli, P.: A focus+content technique based on hyperboic geometry for viewing large hierarchies. In Proc. of CHI'95, Denver, May 1995.

Lave, J., and Wenger, E.: Situated Learning: Legitimate Peripheral Participation. New York: Cambridge University Press, 1991.

Lazar, J., Preece, J.: Classification Schema for Online Communities. Proceedings of the 1998 Association for Information Systems, Americas Conference, pp. 84-86, 1998

Le Grand, B., Soto, M.: Information management - Topic maps visualisation. In: Proc. of XML Europe 2000.

Lee, Y.J.: Facilitating Web Search with Visualisation and Data Mining Techniques. In Knowledge and Information Visualisation: Searching for Synergies,Tergan, S.O., Keller, T. (Eds), Springer Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2005

Lehner, F.: Organisational Memory - Konzepte und Systeme für das organisatorische Lernen und das Wissensmanagement. C. Hanser Verlag, München, 2000

Lenat, D., Guah, R.: Building large Knowledge-based Systems. Representation and Inference in the Cyc Project. Menlo Park: Addison-Wesley, 1990

Lesser, E., Storck, J.: Communities of practice and organizational performance. In IBM Systems Journal, Vol. 40, No. 4: Knowledge Management, 2001

Leuf, B., Cunningham, W. The Wiki Way: Quick Collaboration on the Web, Addison-Wesley, Boston, 2001

Leuski, A., Allan, J.: Lighthouse: Showing the way to relevant information. In Proc. of IEEE Symposium on Information Visualisation (InfoVis'00), 2000

Lin, X., Soergel, D., and Marchionini, G.: A self-organizing semantic map for information retrieval, in Proc. of 14th ACM SIGIR Conf., 1991

Lin, X.: Map Displays for Information Retrieval. Journal of the American Society for Information Science, 48:40-54, 1997

Lin, X.: Searching and Browsing on Map Displays. In Proc. of ASIS '95, Annual Conf. of the American Society for Information Science, Chicago, October, 1995

Lin, X., Qin, J.: Building a Topic Map Repository, In Proc. Knowledge Technologies Conference, Seattle (WA), USA, 2002
http://www.knowledgetechnologies.net/proceedings/02/tuesday.asp

Lokuge, I., Gilbert, S., Richards, W.: Structuring Information With Mental Models: A Tour of Boston, In. Proc. of ACM CHI ’96, ACM Press, 1996

Mack, R., Ravin, Y., Byrd, R.J.: Knowledge portals and the emerging digital knowledge workplace, IBM Systems Journal, Issue on Knowledge Management , Vol. 40, No. 4, 2001 http://www.research.ibm.com/journal/sj40-4.html

Mackinlay, J. D., Rao, R., Card, S. K.: An organic user interface for searching citation links. In Proceedings of CHI’95 (Denver, May), ACM, New York, 1995.

Madhavan, J., P. A. Bernstein, P. Domingos, and A. Y. Halevy. Representing and reasoning about mappings between domain models. In Proc. of the 18th National Conference on Artificial Intelligence and 14th Conference on Innovative Applications of Artificial Intelligence (AAAI/IAAI 2002), August 2002.

Maedche, A., Staab, S: Mining Ontologies from Text. In: R.Dieng/Corby, O (Ed.): Proc. of 12th International Conf. on Knowledge Engineering and Knowledge Management, EKAW-2000. Juan-les-Pins, France: Springer, October 2000

March J. G.: Exploration and exploitation in organizational learning, Organization Science, vol. 2, pp.71-87, 1991

Mark, G., Gonzalez, V., Sarini, M., Simone, C.: Reconciling Different Perspectives: An Experiment on Technology Support, In. Proc. of COOP02, IOS Press, 2002

Markus, M.L.:Toward a Theory of Knowledge Reuse: Types of Knowlwdge Reuse Situations and Factors in Reuse Success " Journal of Management Information Systems, 18 (1), pp 57-93, 2001.

Marshall, C., Shipman, F., Coombs, J.: VIKI: Spatial hypertext supporting emergent structure. In Proc. of ACM ECHT ’94 (Edinburgh, September). ACM, New York, 1994.

Marwick, A.D.: Knowledge Management Technology, IBM Systems Journal, Vol. 40, No. 4, 2001

McCrickard, S. D., Kehoe, C. M. Visualising Search Results using SQWID, In Proc. of 6th International World Wide Web Conference (WWW6), April 1997.

McGill, M.et al.:An evaluation of factors affecting document ranking by IR Systems,” Project Report. Syracuse University, School of Information Studies, Syracuse (NY), USA, 1979

McGrath, J.E.: Methodology matters: Doing research in the behavioural and social sciences, In Baecker, R. M., Grudin, J., Buxton, W. A. S. and Greenberg, S. (ed), Readings in Human-Computer Interaction: Toward the Year 2000, 2nd ed., Morgan Kaufmann, S.Francisco, 1995

McGuinness, D.L., Harmelen, F. van: OWL Web Ontology Language Overview. W3C, March 2003. http://www.w3.org/TR/owl-features/

Merali, Y., Davies, J.: Knowledge Capture and Utilization in Virtual Communities, In Proc. of K-CAP’01, Victoria, British Columbia, Canada, October 2001

Michelis, G. de, Cremers, A. B. (Eds.): Services and Communities in the Information Society. Lecture at the 4th Services Symposium of the German Ministry for Education and Research, Bonn, October, 2002

Mierswa, I,. Wurst, M.: Efficient Case Based Feature Construction for Heterogeneous Learning Tasks. In J. Gama et al. (ed.), Proc. of the European Conference on Machine Learning (ECML 2005), pp. 641-648, Springer, 2005.

Munro, A. J, Höök, K., Benyon, D. (Eds.): Social Navigation of Information Space. Springer-Verlag, New York, 1999.

Munzner, T.: H3: Laying Out Large Directed graphs in 3D Hyperbolic Space. In Proceedings of InfoViz’97, IEEE Symposium on Information Visualisation Phoenix, AZ, IEEE Computer Society Press, pp. 2–10, October 1997.

Neumann, A., Gräber, W., Tergan, S.O.: ParIS – Visualising Ideas and Information in a Resource-Based Learning Scenario. In Knowledge and Information Visualisation: Searching for Synergies,Tergan, S.O., Keller, T. (Eds), Springer Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2005

Nielsen, J,: The need for speed. Jakob Nielsen’s Alertbox for March, 1997 . http://www.useit.com/alertbox/9703a.html

Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company, Oxford University Press, London 1995.

Nonnecke, B., Preece, J.: Lurker demographics: Counting the silent, In Proc. of CHI 2000, The Hague, Amsterdam, ACM Press, 2000

Novak, J., Cuel, R., Sarini, M., Wurst, M.: A tool for supporting knowledge creation and exchange in knowledge intensive organisations, Proc. of I-KNOW 04, 4th International Conference on Knowledge Management, Graz, Springer & Co., July 2004b

Novak, J., Kunz, C., Wurst, M.: Supporting Innovation by Facilitating Knowledge Exchange in Heterogeneous Expert Communities based on Visualisation and Utilization of Tacit Knowledge, Proc. of CINet 2004, 5th Intl. Conference on Continuous Innovation, Sydney, Sept., 2004c

Novak, J., Wurst, M.: Knowledge Explorer: An Interactive Tool for Supporting Cross-Community Discovery and Sharing of Knowledge, Proc. of HCI International 2005, Lawrence Erlbaum Publishers, 2005b

Novak, J., Wurst, M., Schneider, M., Fleischmann, M., Strauss, W.: Discovering, visualising and sharing knowledge through personalised learning knowledge maps – an agent-based approach, Proc. of AAAI Symposium on Agent-Mediated Knowledge Management (AMKM03), Springer Lecture Notes on AI, Springer, 2003a

Novak, J., Wurst, M.: Supporting Knowledge Creation and Sharing in Communities based on Mapping Implicit Knowledge, Journal of Universal Computer Science, Special Issue on Communities of Practice, Springer & Co., March 2004a

Novak, J.D.: Concept mapping: a useful tool for science education. Journal of Research in Science Teaching, 27, pp. 937-949, 1990

Novak, J.D., Gowin, D.B.: Learning how to learn, New York, NY, Cambridge University Press, 1984

Novak, J, Fleischmann, M., Strauss, W., Wurst, M., Morik, K., Ziegler, J., Kunz, C.: Augmenting the knowledge bandwidth and connecting heterogeneous expert communities through uncovering tacit knowledge in Proceedings of IEEE 4th International Workshop on Knowledge Media Networking, IEEE Computer Society Press, CRL, Kyoto, Japan, July 2002.

Novak, J, Wurst, M., Kunz, C: Entdeckung und Nutzbarmachung von stillem Wissen in heterogenen Experten-Gemeinschaften, i-com Zeitschrift für interaktive und kooperative Medien, Heft 3/2003, Oldenbourg Verlag, München, 2003c

Novak, J, Wurst, M., Kunz, C, Morik, K., Ziegler, J., Fleischmann, M., Strauss, W.: Verbindung heterogener Experten-Communities durch die Entdeckung, Visualisierung und Nutzbarmachung von stillem Wissen – das AWAKE Projekt, in Proc. Mensch & Computer 2003, Stuttgart, Sept. 2003b

Novak, J, Wurst, M: Collaborative Knowledge Visualisation for Cross-Community Learning, in Tergan, S. and Keller, T. (eds), Knowledge Visualisation and Information Visualisation – Searching for Synergies, Springer Lecture Notes in Computer Science, 2005a

Orr, J. E. Talking About Machines: An Ethnography of a Modern Job. Ithaca, NY: Cornell Univ. Press, 1996

Paal, S., Kammüller, R., Freisleben, B.: Crossware: Integration Middleware for Cross-Platform Internet Application Environments. Journal on Integrated Computer-Aided Engineering. Special Issue on Autonomous Computing. Vol. 13, Nr. 1, pp. 41-62, IOS Press, 2006

Page L., Brin, S., Motwani, R., Winograd, T.: The PageRank citation ranking: bringing order to the Web. Stanford Digital Libraries Working Paper, 2002

Pallant, A., Timmer, P. McRae, S.: Cognitive Mapping as a Tool for Requirements Capture, In: Robertson, S.A. (Ed.), Contemporary Ergonomics. (pp. 495 –500). Taylor & Francis, London, 1996.

Papert, S.: Introduction: Constructionist Learning. Idit Harel (ed.). Cambridge, MA: MIT Media Lab, 1990.

Park, J. Hunting, S.: XML Topic Maps: Creating and Using Topic Maps for the Web. Addison-Wesley, 2002

Pepper, S., The TAO of Topic Maps, 2000, http://www.ontopia.net/topicmaps/materials/tao.html

Pirolli, P., Card, S.: Information Foraging in Information Access Environments . Proc. of CHI 95, May 1995.

Pirolli, P., Pitkow, J., Rao, R.: Silk from a sow's ear: Extracting usable structures from the Web. In Proceedings of CHI'96, Vancouver BC, April, 1996.

Polanyi, M., Personal Knowledge, Routledge, London, 1958.

Polanyi, M., Tacit Dimension, Peter Smith Pub., 1983.

Polasek, W.: Explorative Daten-Analyse: Einführung in die deskriptive Statistik, 2nd edition, Springer, Berlin, 1994

Pratt, W., Hearst, M. A., Fagan, L. M.: A knowledge-based approach to organizing retrieval documents. In Proc.s of the 16th National Conference on AI (AAAI '99), Orlando, Florida, July 1999.

Preece, J.: Empathic Communities: Reaching out Across the Web, Interactions, 1998

Preece, J.: Online Communities: Designing Usability, Supporting Sociability. John Wiley, Chichester, U.K., 2000.

Probst, G., Raub, S., Romhardt, K.: Wissen managen. Wie Unternehmen ihre wertvollste Ressource optimal nutzen. Gabler/FAZ, Frankfurt am Main, Wiesbaden, 1997

Prusak L.: Knowledge in Organizations. Butterworth-Heinemann, Oxford, 1997

Purser, R., Pasmore, W: Organizing for learning. In Woodman, R and Pasmore, W. (Eds.), Research in Organizational Change and Development (pp.37-114), Vol. 6, Greenwich, Conn, JAI Press, 1992

Qu, Y., A Sensemaking-Supporting Information Gathering System, Proc. of CHI 2003, Extended Abstracts and Applications, ACM Press, 2003

Qu, Y., Furnas, G.: Sources of Structure in Sensemaking. Proc. of ACM CHI 2005, Extended Abstracts and Applications, Portland, Oregon, USA, April 2005

Resnick, P. et al, GroupLens: An Open Architecture for Collaborative Filtering of Netnews, Proc. of ACM 1994 Conf. on CSCW, Chapel Hill, N. Carolina, 1994.

Resnick, P., Varian, H.R.:Recommender systems, Communications of the ACM, Vol. 40, No. 3, 1997.

Reynolds, D., Shabajee, P., Cayzer, S., Steer, D.: Semantic Portals Demonstrator-Lessons Learnt. SWAD-Europe deliverable 12.1.7, 2004
http://www.w3.org/2001/sw/Europe/reports/demo_2_report/

Rheingold, H.:Virtual Communities:Homesteading on the Electronic Frontier, 1993.

Risch, J., May, R., Dowson, S., Thomas, J.A.: Virtual Environment for Multimedia Intelligence Data Analysis. IEEE Computer Graphics and Applications, pp. 33-41, November 1996

Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, C., Van Dantzich, M.: Data Mountain: Using spatial memory for document management. In Proc. of UIST'98 (San Francisco, November), ACM, New York, 1998

Robertson, G., Mackinlay, J., Card, S.: Animated 3d visualisations of hierarchical information. In Proc. of CHI'91,pp. 189-194. ACM Press, April 1991.

Rosch, E.: Principles of Categorization. In Cognition and Categorization. L. Erlbaum Associates, 1978.

Rosson, M.B., Carroll, J.M.: Usability Engineering, Morgan Kaufmann, 2001

Rughase, O.G.: Linking Content to Process: How Mental Models of the Customer Enhance the Creative Strategy Process, In: Mapping Strategic Knowledge, Huff, A.S.,, Jenkins, M. (eds.), Sage Publications, 2002

Russell, D. M., Stefik, M. J., Pirolli, P., Card, S. K.: The Cost Structure of Sensemaking. Proc. of ACM INTERCHI'93, pp. 269-276, 1993

Russell, D.M., Slaney, M., Qu, Y, Houston, M.: A Cost Structure Analysis of Manual and Computer-supported Sensemaking Behavior, International Conference on Intelligence Analysis, 2005

Sack, W.: Conversation Map: A Content-Based Usenet Newsgroup Browser, In Proc. of ACM Conf. on Intelligent User Interfaces (IUI'00), New Orleans (LA), USA, 2000

Salton, G., Allan, J., Buckley, C.: Automatic structuring and retrieval of large text files. Communications of the ACM, 37(2), 97-108, 1994

Sarkar, M., Brown, M. H. Graphical Fisheye Views. Communications of ACM 37, 12, 1994.

Sarkar, M., Brown, M. H.: Graphical fisheye views of graphs. In Proc. of CHI'92. ACM, New York, 1992

Schlichter, J., Koch, M., Xu, C.: Awareness - the common link between groupware and community support systems. In T. Ishida, editor, Community Computing and Support Systems, Lecture Notes in Computer Science 1519, pp. 77–93. Springer Verlag, 1998.

Schmidt, K.: The Problem with 'Awareness': Introductory Remarks on 'Awareness in CSCW', Journal of Computer Supported Cooperative Work, Vol. 11, pp. 285–298, 2002.

Schubert, P.: Virtuelle Transaktionsgemeinschaften im Electronic Commerce. Josef Eul Verlag, Lohmar, 2000

Schvaneveldt, R.W., Durso, F.T., Dearholt, D.W.: Network Structures in Proximity Data. In: The Psychology of Learning and Motivation, G. Bower, ed., Academic Press, San Diego (CA), USA,1989.

Seeling, C., Becks, A.: Analysing Associations of Textual and Relational Data with a Multiple Views System Proceedings of the 2nd International Conference on Coordinated and Multiple Views in Exploratory Visualisation (CMV 2004), IEEE Computer Society Press, London, England, July 2004

Senge P.: The Fifth Discipline: the art and practice of the learning organization. Doubleday Currency, New York., 1991

Shardanand, U., Maes, P.: Social information filtering: algorithms for automating "word of mouth", In Proc. of CHI 95, ACM Press, 1995.

Sharp, J., Communities of Practice: A Review of the Literature, Working Paper, 1997 http://www.tfriend.com/cop-lit.htm

Shaw, M. L. G., Gaines, B.R.: WebGrid: Knowledge Elicitation and Modeling on the Web, In Proc. of World Conference on the WWW and Internet, 1996

Shaw, M. L. G., Gaines, B.R.:Modeling the Social Practices of Users in Internet Communities, In Proc. of 7-th International Conf.on User Modeling (UM'99), Banff, Canada, June 1999 http://www.cs.usask.ca/UM99/Proc/gaines.pdf

Shaw, M. L. G.: Interactive elicitation and exchange of knowledge. International Journal of Personal Construct Psychology, 2, pp. 215-238, 1989.

Shaw, M. L., Gaines, B. R.: WebGrid II: Developing Hierarchical Knowledge Structures from Flat Grids. In Proc. of 11th Workshop on Knowledge Acqusition, Modelling and Management, Banff, Alberta, Canada, April 1998 http://ksi.cpsc.ucalgary.ca/KAW/KAW98/shaw/

Shaw, M.L.G., Gaines, B.R. Supporting personal networking through computer networking. Proceedings of ACM CHI'91, pp.437-438. New York, ACM Publications, 1991.

Shaw, M.L.G.: On Becoming a Personal Scientist: Interactive Computer Elicitation of Personal Models of the World, London, Academic Press, 1980

Shen, R., Richardson, R., Fox, E. (n.d.).: Concept maps as visual interfaces to digital libraries: summarization, collaboration, and automatic generation. In Proc. of Workshop on Information Visualisation Interfaces for Retrieval and Analysis, Joint Conference on Digital Libraries, 2003
http://vw.indiana.edu/ivira03/shen-et-al.pdf

Shipman, F., Marshall, C., And Lemere, M.: Beyond location: Hypertext workspaces and non-linear views. In Proc. of ACM Hypertext'99, ACM, New York, 1999.

Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualisation. Technical Report 96-66, Institute for Systems Research, University of Maryland, 1996

Simunic, K., Novak, J.: Combining Visualisation and Interactive Clustering for Exploring Large Document Pools, In Proc. of IASTED Conf. on Visualisation, Imaging, and Image Processing (VIIP 2004), Marbella, Spain, 2004

Smith, M.: Invisible Crowds in Cyberspace: Measuring and Mapping the USENET, Communities in Cyberspace, Routledge, 1999.

Soergel, D.: 2000 Taxonomy of Knowledge Organization Sources/Systems , 2000 http://nkos.slis.kent.edu/KOS_taxonomy.htm

Star, S. L., Griesemer, J. R.: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, 1989.

Star, S. L.:Cooperation Without Consensus in Scientific Problem Solving: Dynamics of Closure in Open Systems. In: CSCW: Cooperation or Conflict, S. Easterbrook, (ed.). London, Springer, 1993

Star, S.L.: The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In Readings in Distributed Artificial Intelligence, ed. M.Huhn, L.Gasser, pp. 37-54. Menlo Park, CA, Morgan Kaufman, 1989

Studer, R. et al.: Ontologies and the Configuration of Problem-solving Methods. In: Gains, B.R., Musen, M.A. (Eds.): Proceedings of the 10th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop. Banff, Canada, 1996.

Studying Online Social Networks. Journal of Computer Mediated Communication, 3(1), June 1997 http://jcmc.huji.ac.il/vol3/issue1/garton.html

Swan, J., Knowledge Management in Action: Integrating Knowledge Accross Communities, In Proc. of the 34th Annual Hawaii International Conference on System Sciences (HICSS-34), Maui, Hawaii, January 2001

Swan, J., Scarbrough, H., Robertson, M.: The construction of 'communities of practice in the management of innovation. Management Learning, 33, 4, pp. 477-496, 2003

Swarbrick, A.: Visualising Knowledge Networks: Leveraging the Potential of Work Based Communities, Technical Report, Department of Computer Science at the University of York, 2002 http://www.knowledgeboard.com/download/206/visualising_knowledge_networks.pdf

Szczepanska, A.M.: Searching for the Virtual Community. In Proceedings of 24th Information Systems Research Seminar in Scandinavia, Norway, 2001.
http://www.viktoria.se/results/result_files/152.pdf.

Tergan, S.O., Keller, T. (Eds.): Knowledge and Information Visualisation – Searching for Synergies, Springer LNCS 3426, Springer Verlag, Berlin-Heidelberg, 2005

Tergan, S.O.: Digital Concept Maps for Managing Knowledge and Information. In Tergan, S.O., Keller, T. (Eds.): Knowledge and Information Visualisation – Searching for Synergies, Springer LNCS 3426, Springer Verlag, Berlin-Heidelberg, 2005

Terveen, L. G., Hill, W.: Beyond recommender systems: Helping people help each other. In HCI In The New Millennium, J. Carroll, Eds. Addison-Wesley, Reading, Mass, 2001

Tönnies, F.: Gemeinschaft und Gesellschaft, Grundbegriffe der reinen Soziologie. Wissenschaftliche Buchgesellschaft, Darmstadt, 1991 (originally published in 1887)

Totland, T.: Enterprise modelling as a means to support human sense-making and communication in organizations", NTNU PhD Thesis 1997, ISBN 82-471-0133-5.
http://www.idi.ntnu.no/grupper/su/publ/html/totland/thesis.htm

Tulsa, O.K., Eletronic Statistics Textbook. StatSoft Inc., 1997 http://www.statsoft.com/textbook/stathome.html

Turkle, S.: Life on the Screen: Identity in the Age of the Internet, Simon and Schuster, New York, 1995.

Ultsch, A.: Self-organizing neural networks for visualisation and classification. In: Opitz, O., Lausen, B., Klar, R.(eds.): Information and Classification, Springer-Verlag, Berlin, 1993

Uschold, M., Gruninger",M.: Ontologies: principles, methods, and applications, Knowledge Engineering Review, Vol. 11, No. 2, pp. 93-155, 1996

Uschold, M., Jasper,R.: A Framework for Understanding and Classifying Ontology Applications, in Proc.of the IJCAI99 Workshop on Ontologie and Problem-Solving Methods(KRR5), Stockholm, Sweden, August 1999. http://sunsite.informatik.rwthaachen.de/Publications/CEUR-WS/Vol-18/

Vesanto, J.: SOM-Based Data Visualisation Methods. Intelligent Data Analysis, Vol. 3, No. 2, Elsevier Science, April 1999

Vliet, W. van , Burgers, J.: Communities in transition: From the industrial to the postindustrial era. In: Altman, I., Wandersman, A. (Eds.): Neighborhood and Community Environments, Plenum Press, New York, 1987

Voss, A., Nakata, K., Juhnke, M.: Concept Indexing, In Proc. of GROUP 99, Phoenix, USA, ACM Press, 1999

Vygotsky, L., Thought and Language. Cambridge, MA: The MIT Press, 1986 (originally published in 1962).

Wehrend, S., Lewis, C.: A problem-oriented classification of visualisation techniques. In Proc. of IEEE Visualisation '90, IEEE Computer Society Press, 1990

Weick E.K. The social psychology of organizing. McGraw-Hill, Inc., 1979

Wellman, B. , Gulia, M.: Net Surfers dont ride alone: Virtual Communities as Communities. In: Kollock, P., Smith, M. (Eds.): Communities and Cyberspace, Routledge, 1999

Wellman, B., Carrington, P.J., Hall, A.: 'Networks as Personal Communities'. In B. Wellman, S.D. Berkowitz, S.D. (ed.), Social Structures: A Network Approach. Cambridge: Cambridge University Press, 1988

Wellman, B.: An Electronic Group is Virtually a Social Network, in Culture of the Internet, S. Kiesler (ed.), pp. 179-208, Lawrence Erlbaum Associates, Mahwah, NJ, 1997,

Wellman, B.: Virtual Community: Introducing a New SIGGROUP Focus Aream, ACM SIGGROUP Bulletin, Vol 19, No. 1, 1998a

Wellman, M.P.: Inference in Cognitive Maps, In: Mathematics and Computers in Simulation, 1994.

Wellman., B.: A computer network is a social network. SIGGROUP Bulletin, 19(3):41–45, Dec. 1998b

Wenger, E., Snyder, W. M.: Communities of practice: The organizational frontier. Harvard Business Review, pp. 139–145, Vol. Jan-Feb 2000.

Wenger, E.: Communities of Practice: Learning, Meaning and Identity. Cambridge Univ. Press, 1998

Wenger,E., McDermott, R., Snyder, W.: Cultivating communities of practice. A guide to managing knowledge, Harvard Business School Press, 2002

Wever, P.: Entwicklung eines Moduls zur Unterstützung adaptiver Visualisierung von Wissenslandkarten, Bachelor's Thesis, University of Applied Sciences Bonn-Rhein-Sieg, February 2005

Wexelblat, A., Maes, P.: Footprints: History-rich tools for information foraging. In Proceedings of CHI'99. Pittsburgh, May, 1999.

Whittaker, S., Isaacs, E., O'Day, V.: Widening the Net, SIGCHI Bulletin (29:3), pp. 27-30, 1997

Wiig, K.M.:Knowledge Management: Where did it come from and where will it go, Expert Systems with Applications, Vol. 13, No. 1, pp. 1-14, September, 1997.

Wittenschlaeger, T.A., Fiedler, J.A..: Current Practices in Perceptual Mapping, Sawtooth Software Conference, 1997

Wittgenstein, L.: Philosophical Investigations, New York, MacMillan, 1974 (originally published in 1953).

Wurst, M., Novak, J., Schneider, M.: Integrating different Machine Learning methods to support search in cross-domain information sources - the project AWAKE, In Proc. of the Workshop of the German Society for Computer Science Machine Learning Special Interest Group (FGML-2002), Hannover, Germany, October 2002

Wurst, M., Novak, J.: Knowledge Sharing im Heterogeneous Expert Communities based on Personal Taxonomies. In Proc. of ECAI Workshop on Agent Mediated Knowledge Management, Valencia, Spain, 2004.

Wurst, M.: Evaluating Knowledge Management in Heterogeneous Domains by Agent-based Simulation. In Proc. of the Workshop on Agent Mediated Knowledge Management (AMKM'05), 2005.

Xiong, R., Donath, J.: PeopleGarden: Creating Data Portraits for Users, In Proc. of ACM Conference on Intelligent User Interfaces (UIST-99), 1999

Yang, Y., Liu, X.: A re-examination of text categorization methods. In SIGIR-99, Berkley (CA), USA, ACM Press, 1999.

Zhang, J. The nature of external representations in problem solving. Cognitive Science, 21(2), p. 179-217, 1997

Zhang, J., Norman, D.A.: Representations in distributed cognitive tasks. Cognitive Science,18, pp.87-122, 199