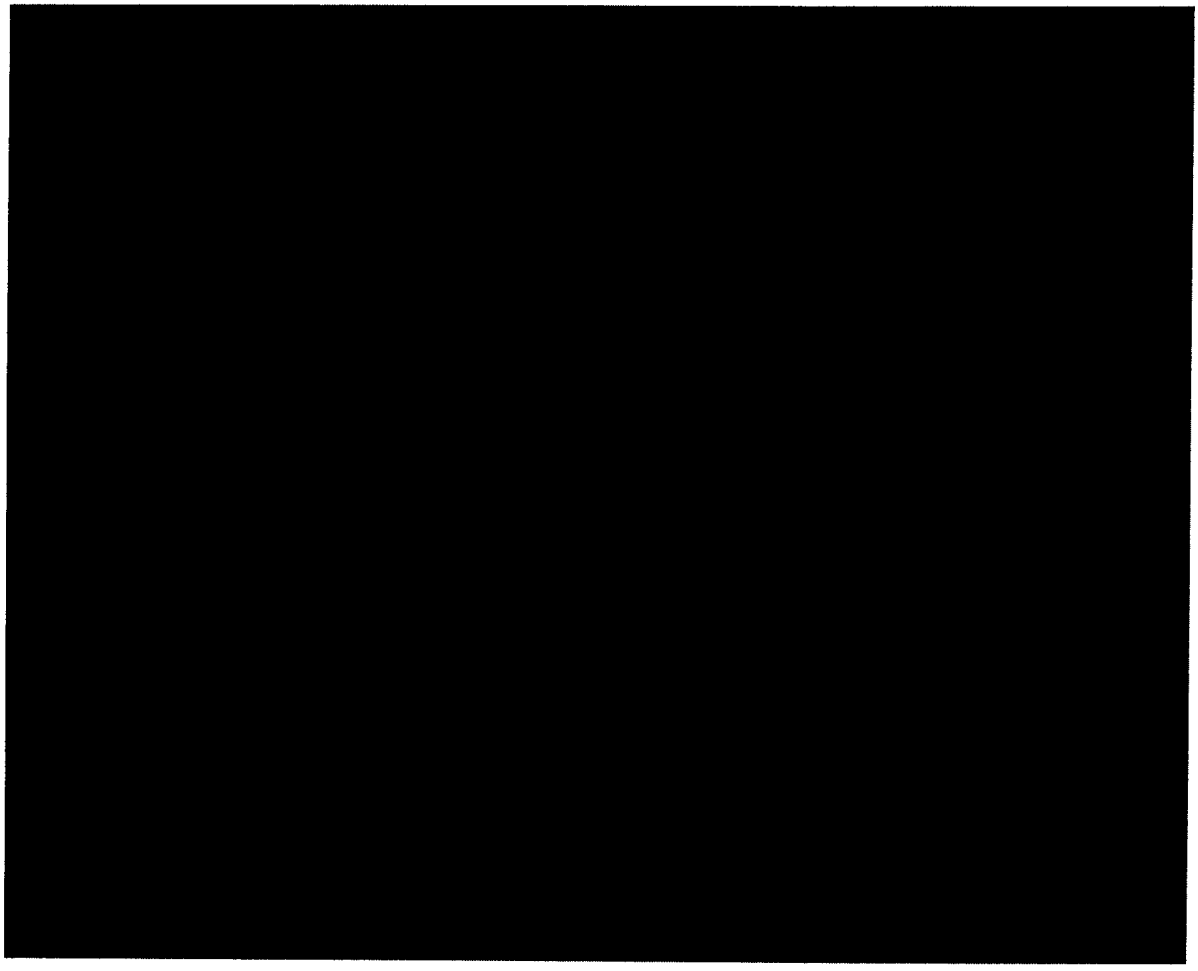
Vol. 1, No. 1
January - March 2005

An official publication of the Information Resources Management Association

INTERNATIONAL JOURNAL OF

Knowledge Management



IDEA GROUP PUBLISHING
Publisher of IT books, journals and cases since 1988
www.idea-group.com

International Journal of Knowledge Management

Jan-March 2005, Vol 1, No. 1

Table of Contents

Continuum of Context Explication: Knowledge Discovery Through Process-Oriented Portals

Stefan Smolnik, University of Paderborn, Germany
Stefan Kremer, University of St. Gallen, Switzerland
Lutz Kolbe, University of St. Gallen, Switzerland

ABSTRACT

For a company to be consistently oriented toward its customers and their processes, it needs to customize its intra-corporate processes and systems. Customer process-oriented portals that integrate companies' systems and provide transparent access to information objects stored in these systems seem to offer a solution. However, one key problem is finding relevant information objects in systems that are not only growing, but also being disseminated. There is the additional challenge of making knowledge available at the right time and the right place. Companies' competitive advantage is rooted in this knowledge advantage as well as in the capability to transform this superior knowledge into market-driven business processes. The research questions addressed in this article are how the value of information objects is affected by the context in which it is considered, and how associated contexts can be uncovered for given situations. We introduce a continuum of context explication comprised of the relationships between data, information objects, and knowledge, and their contexts, according to their degree and ease of context explication. The extremes of the continuum would therefore be data with no context to explicate and knowledge with rich, person-specific context. We conclude that discovering implicit meanings and expressing those meanings explicitly increases the potential value of information objects. In addition, we evaluate the full-text search, attribute-based search, and topic maps as approaches for knowledge discovery through customer process-oriented portals as well as providing patterns that indicate when to apply which approach. Two small case studies of knowledge discovery through such portals are presented. We conclude with suggestions for future research based on our final deductions in respect of the study.

Keywords: *action research; case studies; information search & retrieval; knowledge creation; knowledge discovery; portals*

INTRODUCTION & OVERVIEW

Challenge

The use of information technology has given many organizations access to vast internal and external information repositories. Intranets, content management systems, and enterprise portals have become commonplace, providing employees with opportunities to discover knowledge enshrined in information objects, for example, in electronic documents (Latham, 2001; WebCKS, 1999). Nonetheless, dealing with information and finding the right content are inefficient actions. Davenport, Harris, and Kohli (2001) stated, "Information management must begin by thinking about how people use information..." (p. 63). This way of thinking is a precondition for the feasibility of the use of information.

Although organizations currently have access to different information repositories, the process of knowledge discovery still has major shortcomings, such as:

- *Lack of information.* Finding information objects on a topic is frustrating if users know that they exist but cannot trace them.
- *Overload of information.* Knowledge discovery is time-consuming if too many information objects are found which have no or little relevance.

One key to successfully minimizing these deficits is by controlling the semantics (i.e., the meaning of terms), making explicated context available, and methodically classifying information objects utilized in business environments (Dale, 2001; Felber & Budin, 1989). Different technological approaches — based on different degrees of context explication, that is, discovering implicit meanings and expressing those meanings explicitly — have been proposed to address the lack of relevant and overload of remotely related information problems in knowledge discovery. Examples of such approaches are search and classification engines. Nevertheless, there are hardly any criteria available with which to support organizations' choice of an appropriate solution. Consequently, we present a comprehensive overview of several approaches, their underlying principles, advantages and constraints. To fulfill specific organizations' needs, criteria are also provided in respect of the degree of context explication required.

Objective & Research Approach

The overall objective of this article is to propose a continuum of context explication comprised of the relationships between information objects and their contexts to foster knowledge discovery. We will demonstrate that the continuum allows organizations to make deductions with regard to the appropriate approach with which to stimulate knowledge discovery through portals.

A review of the knowledge management literature provided a comprehensive overview of the topic and related work. It furthermore revealed the gap between research on the context of an information object and that on its explication. Desk and action research, for example, prototypical implementations of our conceptual approaches, led to logically deduced concepts (see Checkland & Holwell, 1998), while the case research allowed the deduction and validation of these concepts. In terms of our research questions, the latter was

particularly suitable since the research and theory are still in the early stages of formulation (Benbasat, Goldstein & Mead, 1987). Consequently, the research and descriptive processes were also influenced by the results from workshops conducted and projects undertaken with our corporate partners during the action research (Gummesson, 2000; Whyte, Greenwood & Lazes, 1991). We are currently testing and expanding the findings with other partners as well.

Structure of the Article

The subsequent section deals with related work in the area of knowledge discovery and portals. It also defines the most important and relevant terms for an understanding of the research field.

In the third section we describe the challenges facing knowledge discovery. Thereafter we introduce three major approaches for discovering knowledge through portals by providing a chronological outline of the different development phases. We describe the three approaches — the full-text search, attribute-based search, and topic maps — according to their characteristics, advantages, and restrictions in respect of context explication.

Knowing the three approaches' capabilities and constraints, and based on given prerequisites, we then propose a continuum of context explication, providing criteria and advice for choosing an appropriate solution.

In the penultimate section we provide two examples of how the continuum was successfully applied in a normal work situation. This was done at our institutions where, based on different prerequisites, we chose and implemented different solutions for knowledge discovery. Finally we draw a conclusion and propose directions for further research.

RELATED WORK & DEFINITIONS OF TERMS

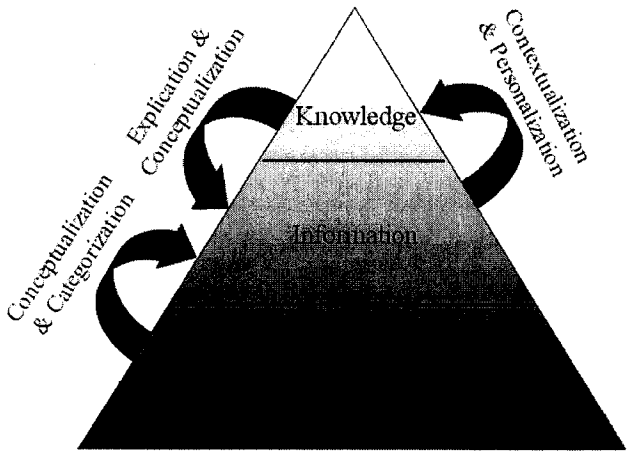
In this section we introduce the theoretical background and define the most relevant terms. We identify related work and explain how they differ from our approach.

Knowledge & Context

Within the literature there are many definitions of knowledge (e.g., Biggam, 2001; Davenport & Prusak, 1998; Lai & Chu, 2000; Murray, 1996; Nonaka & Takeuchi, 1995; Polanyi, 1966; Sveiby, 1997; Table 1; some of these references also provide detailed discussions on the differentiation of the terms data, information, and knowledge as well as discussing knowledge types and their classification; Figure 1).

In our view, knowledge comprises both information and person-specific aspects such as experiences, values, and insights. An important characteristic of knowledge, which simultaneously distinguishes it from information, is its strong affinity to activities (see Davenport & Prusak, 1998). Individuals act and react according to their experiences and intrinsic attitudes. Knowledge, on the other hand, is much more than transformed information and therefore cannot be represented in the form of information objects or data. Polanyi (1966) developed a concept of implicit knowledge that he described as follows: "We can know more than we can tell" (p. 4). We concur with Polanyi's basic concept that the implicit and the explicit dimensions of knowledge are complementary (Polanyi, 1966) — all knowledge contains both dimensions. Pure explicit or im-

Figure 1: Data, Information, and Knowledge (Klemke, 2000)



plicit knowledge, or the conversion of one into the other, is thus impossible.

Many of these definitions of knowledge have context as an important common aspect. One form of transition from information to knowledge is contextualization. Dey and Abowd (1999) define context as "any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves" (p. 3). Similarly, Sowa (2000) describes context in its nonlinguistic meaning as "situation, environment, domain, setting, background, or milieu that includes some entity, subject, or topic of interest" (p. 275).

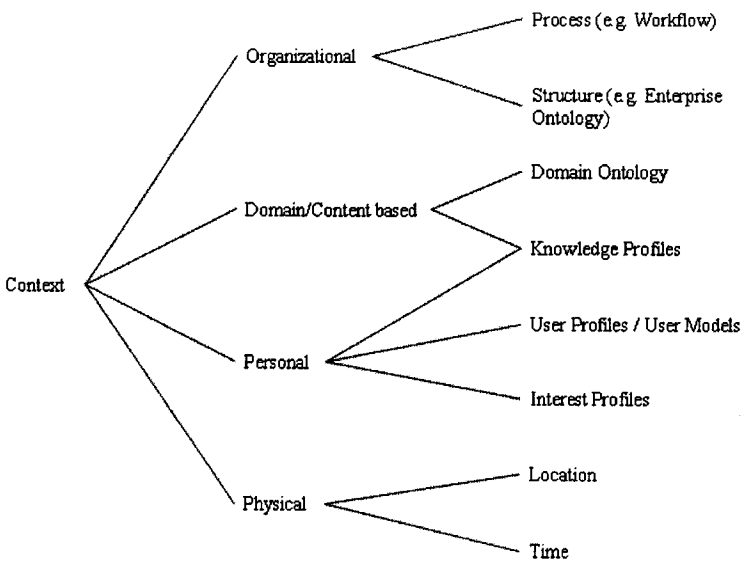
Klemke (2000) describes the differentiation of context types by means of a level-based approach. The first level identifies the following context dimensions: organizational, domain-/content-based, personal, and physical. These dimensions are specified in more detail on the second level; for example, the organizational dimension is subdivided into a process and a structure component. In spite of the common assumption that context only consists of implicit information, these definitions allow context to be either explicit or implicit. In this article we reveal that the explication of information objects' implicit context, that is, the discovery of implicit meanings and expressing those meanings explicitly, supports the creation of new knowledge. We moreover describe different approaches with which to achieve this.

Klemke (2000) recommends a holistic understanding of context by means of several dimensions (cf. Figure 2) and the implementation of an integrated architecture to trace and maintain context models. In addition, the literature regards contexts as having different characteristics and uses different approaches to model these contexts; for example, workflow process context is modeled by workflow management systems (Wargitsch & Habermann, 1998), while organizational structures are modeled by enterprise ontologies. Conversely, we focus directly on information objects and

Table 1: Overview of Selected Definitions of the Term "Knowledge"

Author	Definition
Davenport & Prusak, 1998	"Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers." (p. 5)
Nonaka & Takeuchi, 1995	"First, knowledge, unlike information, is about beliefs and commitment. ... we consider knowledge as a dynamic human process of justifying personal belief toward the 'truth'." (p. 58)
Alavi & Leidner, 1999 Lai & Chu, 2000	Knowledge is created and organized by the very flow of information, anchored by the commitment and beliefs of its holders. Information becomes knowledge when it is processed in the mind of an individual and knowledge becomes information when it is articulated or communicated to others in the form of text, computer output, speech or written words, and so forth.
Murray, 1996	"Knowledge solves a problem; it produces competence leading to effective action." "Making the tacit explicit often includes the following activities: ... Identifying terminology that is clearly understood and using language that is appropriate for the culture and context."
Sveiby, 1997	Knowledge is the capacity to act within context.
Polanyi, 1966	Tacit knowledge is personal, context-specific, difficult to express in verbal, symbolic, and written form, and therefore hard to formalize and communicate.
Biggam, 2001	"- It must be true. - The perceiver must believe this to be the case - The perceiver must be in a position to know this to be the case" (p. 3)

Figure 2: Context Typology (Klemke, 2000)



their contexts and provide approaches with which to discover, explicate, and use these contexts in different situations to increase the information objects' potential value and to stimulate knowledge discovery.

Eventually, we believe that all documents are information objects. Users are able to create knowledge by processing and understanding them but the information objects do not comprise knowledge. However, we recognize that many KM researchers differentiate between information and knowledge object documents; that is, they allow that documents with context can be knowledge objects. This article considers all documents are information objects and requests that readers accept this viewpoint throughout the rest of the article.

Knowledge Discovery as an Important Knowledge Management Activity

Many knowledge management activities, methods, or modules have been discussed within the literature. Lai and Chu (2000) suggest an integrated knowledge management framework that comprises the following activities: initiation, generation, modeling, repository, distribution and transfer, use, and retrospect. Davenport and Prusak (1998) differentiate between specifying a requirement, capturing, distributing, and using knowledge. Probst, Raub, and Romhardt (1999) present a pragmatic approach to the organization-wide management of knowledge. This approach comprises six core processes and two pragmatic modules: the identification, acquisition, development, distribution, use, and preservation of knowledge as well as the objectives and performance measurement of knowledge. More or less similar classifications of knowledge management activities are also offered in Nonaka and Takeuchi (1995), Arthur Andersen (1996), and Alavi (1997).

All these approaches have a method for the identification or use of knowledge, either implicitly or explicitly, in common. Unused knowledge that is generally to be found within organizations can be uncovered with appropriate methods and thereafter utilized. Knowledge discovery methods additionally foster knowledge transparency in organizations as well as supporting users to find relevant information objects. They are a necessary precondition for the core processes of knowledge identification and knowledge use (Probst, Raub & Romhardt, 1999). They not only improve the organizational use of existing individual and common knowledge, but also contribute to the knowledge generation process, that is, the development or collecting of new knowledge (Güldenberg, 1996).

Portals

Portals have been discussed as an integration concept for user access to personalized information and applications since 1998 (Bristow et al., 2001). Although there are many different descriptions of portals (e.g., Davydov, 2001; Dias, 2001; Kalakota & Robinson, 2001; Röhricht & Schlögel, 2001; Schwarz, 2000), we focus on them as Web-based, personalized, and integrated access systems to internal and external applications and information repositories.

Portals support knowledge-oriented processes by providing users with a graphical front-end integration of back-end systems that comprises, amongst others, integration, personalization, and administrative services. Knowledge discovery methods through portals are mainly supported by navigation and search mechanisms (Fleisch

& Österle, 2001; Puschmann, 2003). The role of search mechanisms is especially significant in these methods, as the following section shows.

KNOWLEDGE DISCOVERY THROUGH PORTALS

As stated earlier, search and retrieval play a vital role in the concept of portals, but knowledge discovery through portals faces special challenges (see Baeza-Yates & Schäuble, 2002; Raghavan, 2002):

- *Heterogeneous structures and formats.* Information objects are stored in multiple, roughly structured formats classified differently and presented in different languages. Portal users therefore need a standardized view of all the available information objects.
- *Distributed and redundant information.* Organizations have information objects residing partly redundantly in a variety of sources, for example, e-mail, content management, and file server systems. Knowledge discovery processes have to offer mechanisms that connect these repositories to the portal to provide users with a consolidated view.
- *Protected content.* The role of each individual portal user dictates which information objects that individual is able to access. In the process of knowledge discovery, navigation entries and search results have to be filtered to display only the information objects accessible to the user; that is, secure access has to be provided.

Addressing these challenges is fundamental to supporting knowledge discovery methods through portals (Andrews, 2003). The following sections illustrate three major approaches with which to achieve this objective by providing a chronological outline of different development phases.

Full-Text Search

The classic full-text search has been an established retrieval approach since the early 1990s (Rappoport, 2002). A search engine is an information technology component of a portal that acts as a central instance between the user's information need and the available information objects that are stored in one or more repositories. Users transform their information need into a search query and enter it in a search field provided within a portal. In order to respond to the user queries, a search engine indexes each information object, representing it as a set of weighted words. The search engine compares the entered terms with the previously indexed information objects and provides the users with a result list.

The benefits for users are

- *Speed:* searching the content of multiple repositories with a single query is faster then searching each application individually with separated queries.
- *Ease of use:* currently, the full-text search is well known and most users have some experience of this (Gordon & Pathak, 1999).
- *No pre-structuring required:* since the relevant terms are indexed automatically, no human-driven intervention is necessary.

Since the expressing of an information need in a single query has a strong impact on the quality of the search results,

the main restrictions of the full-text search emanate from semantic issues:

- "Wrong" or too many results. Receiving search results referring to information objects with no or little relevance is time consuming (Cathro, 1997). In this context, the challenge is for users to anticipate the correct terminology, that is, to match the authors' terminology.
- Spelling. A user's query should be orthographically correct.

As shown, the full-text search is dependent on the information object's content, because its context is exclusively contained within the information object itself. The authors do not provide explication during the information object's creation, nor does a system later do so.

Attribute-Based Search

To overcome the restrictions of the full-text search, the attribute-based search was developed in the mid 1990s (Cathro, 1997). This approach is based on a context explication model that stores the context of an information object as metadata, that is, as data about the data (Berners-Lee, 1997). The metadata are stored with the information object itself and can be viewed and retrieved by users and applications. Common metadata attributes that are associated with information objects include the author's name, the date of publication, the source of publication and so forth. The attribute-based search during knowledge discovery through portals would therefore permit structured queries on the context explicated in information objects' metadata (McGovern, 2001). Currently there are several metadata standards, for example the Dublin Core Metadata Element Set, which proposes 15 fields or attributes according to which a document can be described (Baeza-Yates & Ribeiro-Neto, 1999; Dublincore, 2003).

The major benefits of the attribute-based search are:

- *Reduced result set.* Compared to the full-text search, users retrieve relevant information objects more swiftly.
- *Controlled vocabulary.* Users can choose standardized terms from drop-down lists.
- *Personalization capabilities.* Search queries can be automatically enriched with personalized information (user attributes, e.g., roles, language, and organizational unit).

But there are also certain constraints:

- *Maintenance of controlled vocabulary.* Although this approach is less time consuming when users want to find relevant information objects, human intervention is required at the time of creation to provide them with appropriate context attributes.
- *Metadata are stored with the information object itself.* Since terms could change over time, reclassification may be necessary. Alternatively, reclassification could be avoided by mapping old terms to new ones separately, for example, with a customized thesaurus.

From this it is clear that because the information objects contain content and explicated context, both maintained by the their author at the time of creation or during maintenance changes, the attribute-based search is actually based on context explication.

Topic Maps

The initial idea behind topic maps — which date back to the early 1990s — arose from the need to model intelligent electronic indexes (of books), tables of contents, glossaries, thesauri, or cross references in order to merge them automatically. During many years of discussion and evolutionary development cycles, the topic map model has developed into something far more powerful that is no longer restricted to simply modeling indexes. The ISO standard ISO/IEC 13250 Topic Maps, adopted in 1999, defines a model and architecture for the semantic structuring of link networks. Topic maps establish an associative network between subjects, which represent information objects, and provide navigation paradigms that allow them to be searched. By applying topic maps to large sets of heterogeneous information repositories, reusable and structured semantic link networks are created on a level above those resources (Rath & Pepper, 1999). The key concepts of topic maps are topics, which represent real-world subjects, occurrences of topics, and relationships between topics (topic associations). In addition, the topic map standard provides the extended concepts of scope, public subject, and facets. For a comprehensive introduction and reference, refer to Rath and Pepper (1999) and ISO/IEC 13250 (2002).

Topic associations describe the relationships between topics. They are completely independent of the information object itself and therefore represent the essential added value of the topic map. The addition of topic associations to the concept of topics enables topic maps to model information networks.

Topic maps organize information repositories in a new knowledge space by relating them to topics, and structurally associating these. They furthermore enable heterogeneous sets of information repositories to be used in an integrated way by interrelating them by means of a unifying conceptual framework. Another characteristic of topic maps is that they are well suited to represent ontologies. Consequently, they facilitate the description of a shared common understanding, for example, about the kinds of objects and relationships that are being discussed (Wrightson, 2001).

The link mechanism between topics and topic occurrences provides a means with which to "bridge the gap" between knowledge representation and information management fields (Pepper, 1999).

Since the human brain always remembers memorized issues in a specific context (Goldfarb & Prescod, 2000), association is the basic way of thinking. Topic maps support this way of thinking by pointing to related themes when a user searches for a specific theme.

To summarize, topic maps have the following benefits:

- *Creation of knowledge structures.* Applying topic maps to information repositories generates knowledge structures. They form structured semantic link networks above large sets of information repositories.
- *Creation of meta-layers.* Transparent access to information objects is provided by searching and navigating knowledge structures, that is, a meta-layer above the information objects. Modifications of the meta-context do not affect the information objects or their descriptors. Searching in topic maps can be compared to searching in knowledge structures.

- *Discovery of new context.* Added value is achieved by the creation of new knowledge through the discovery of new contexts.
- *Support of human thinking.* Topic associations support the basic way of thinking by interrelating themes.

Some basic constraints are:

- *Effort required for topic map creation and maintenance.* Intense human effort is needed to define, create, and maintain topic maps. Persons who manage topic maps need expertise in both topic map concepts and paradigms as well as in the specific domain to which the topic map applies.
- *New search paradigm.* Users have to learn to use the topic map search concepts and to adopt the associative way of thinking, while they are familiar with the full-text or attribute-based search concepts and their flat result sets.

As has been described, topic maps provide strong concepts and paradigms with which to discover and explicate information objects' contexts, relieving authors and users of the need to provide metadata or descriptions. The explicated context does not form part of the information object and can even be used without it. However, specifically skilled persons are required to support the process of context explication. Concepts for the organizational and process integration of such "knowledge workers" are introduced in detail in Smolnik and Nastansky (2002). In general, they need expertise in managing topic maps as well as in the specific domain to which the topic map applies.

THE CONTINUUM OF CONTEXT EXPLICATION

As pointed out in our motivation, context has been recognized as being an important aspect to consider when looking at the meaning of information with respect to knowledge discovery and knowledge creation. In the previous section, we presented three approaches with which to find information objects and with which to recognize, represent, and use contextual information through portals. Even though these approaches have supporting users to find relevant information objects in common, they focus on contextual information in different ways and with varying intensity.

The introduction of the continuum of context explication was one of the major results of our research. This continuum focuses on data, information objects, and knowledge as basic subjects of portals, as well as on their varying embodied degree of context explication. It furthermore describes approaches with which to find and use information objects and contextual information (Table 2). We define five approaches, each with a differing degree of context and explication ease: three approaches relate to information objects and the chronological development of search methods, with the other two forming a logical extension in the transition of data into information and information into knowledge. We additionally provide criteria and advice for choosing an appropriate solution, based on given prerequisites.

Data Approach

Data are meaningless symbols without content and context that have no context to explicate. Depending on the data

quantity and the relevant domain, there are several methods with which to transform data into information objects, or even into domain-specific knowledge. For example, in the knowledge discovery in databases and in the data mining research domains, the identification of patterns in large structured data sets results in the non-trivial extraction of implicit, previously unknown, and potentially useful knowledge (Fayyad, Piatetsky-Shapiro & Smyth, 1996). Processed and conceptualized data, such as documents created by authors, are defined as information.

The data approach is appropriate for the following situation:

- No or little interaction with users, authors, or knowledge workers
- Large structured data sets
- Possible automatic data generation or collection

Information Approach

Even a simple information object contains some kind of content, for example text, an audio annotation, or a spreadsheet. Although the information object may provide no explicit context like descriptors or other contextual information, it inherently contains context. The context is, however, interwoven with the content and difficult to conceptualize, which means that the methods implemented to find requested information objects have to rely on the content and cannot access contextual information. An exemplary method is the full-text search as previously described. Normal full-text search engines use information objects' indexed contents to respond to a query and do not access contextual information at all. No effort is therefore made to explicate context, as neither the authors nor the users provide or use explicit contextual information.

The information approach is appropriate for the following situations:

- Many users who have little or no experience with enhanced searching approaches, or who are unwilling to use them
- Authors who have no experience with describing their information objects
- Numerous unstructured information objects

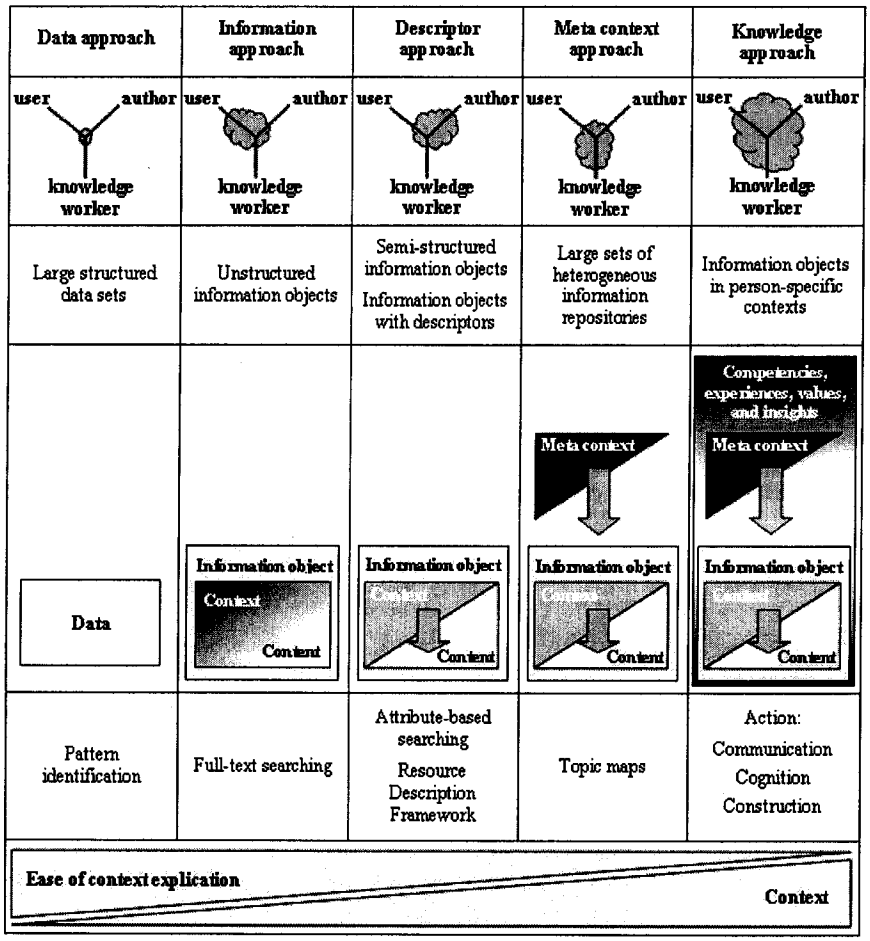
Descriptor Approach

Information objects are often enriched with metadata; that is, they contain content and explicit contextual information. Examples are Microsoft Word documents, Adobe PDF documents, or semi-structured documents in a groupware-based office environment.

In contrast to the information approach, the information objects do not only contain implicit contextual information but also explicit contextual information. As previously explained, a standard for formulating contextual information is the Dublin Core Metadata Element Set that proposes specific attribute classes for the description of an information object. Another concept for structuring and providing metadata is the Resource Description Framework (RDF), which is resource-oriented. Its main objective is the description of resources and their relationships to other resources, with the description mostly residing in the resource.

In contrast to the information approach, some effort is necessary to enrich an information object with explicit contextual information. Authors have to provide this information at the time of creation. In

Table 2: Continuum of Context Explication



addition, software systems try to maintain some of the contextual information.

The advantage of the attribute-based search as a retrieval method for information objects is dependent on the quality of the provided explicit contextual information (see the introductory section on the attribute-based search). If the metadata are wrong, misleading, or incomplete, the attribute-based search will provide insufficient result sets — if not, the attribute-based search provides more accurate results, which will fit to the users' context to some degree.

The descriptor approach is suitable for the following situations:

- The authors are both trained in and skilled at describing their information objects
- The information objects contain descriptors
- There is a large quantity of semi-structured information objects

Meta-Context Approach

When extending the descriptor approach, information objects are not only

described by metadata that reside in the information object, but also by subjects, concepts, or themes that form contextual information in a meta-layer above the information objects. This contextual information is not necessarily explicitly stored within the information object.

Topic maps provide strong paradigms with which to discover, maintain, navigate, and visualize this meta-context and thus explicate the context of an information object (see the introductory section on topic maps).

Semantic relationships between information objects are expressed by associating topics. This semantic network links the explicated contextual information of different information objects and discovers new contexts. The discovery of these new contexts supports users in creating new knowledge when they associate known information objects in a new way with other information objects. To achieve these benefits from explicated and new contexts, substantial effort has to be invested to define, create, and maintain a topic map. This effort is disproportionally higher than the definition of metadata in the descriptor approach. In the latter case, authors or software systems explicitly provide contextual information. Authors know what they publish and can easily describe their information objects. In the meta-context approach, knowledge workers are needed to provide and maintain a topic map.

The benefit for users depends on the quality of the knowledge workers' work. If the meta-contexts layer covers the entire domain of interest and contains rich and numerous topic associations, users will be able to explore the search domain easily. They will thus be enabled to discover new contexts and to leverage and enhance their knowledge.

The meta-context approach is suitable for the following situation:

- Knowledge workers who are familiar with both topic map concepts and the domain of interest
- Manageable domains of interest
- Existing taxonomies of the domains of interest
- Users experienced in searching and navigating topic maps
- Large sets of heterogeneous information repositories

Knowledge Approach

So far, we have only focused on the human factor in very specific perspectives such as authors defining the metadata of information objects, or knowledge workers developing topic maps. The human factor plays a decisive role in the conversion of information into knowledge. We subsequently differentiate two facets of the human factor.

Firstly, there are the competencies, experiences, values, and insights that form a rich, person-specific context. This context is a feature of knowledge's implicit dimension and can hardly be explicated (Polanyi, 1966). Within this context, a highly individual and subjective meaning is assigned to an information object. Secondly, there is the users' active involvement that is a necessary precondition to convert information into knowledge. This active involvement comprises actions like communication, construction, or, more intrinsically, cognition. If users experience an "I see!" event, that is, an epiphanic event, as a result of some action, knowledge is created.

Characteristics of the knowledge approach are:

- Competencies, experiences, values, and insights
- Information objects in person-specific contexts
- Creation of knowledge through human actions, for example, cognition of information objects.

SMALL CASES & LESSONS LEARNED

In this section we present two small cases derived from prototypical implementations at our institutes. They illustrate the benefits and constraints of the previously discussed approaches presented with respect to the discovery of information objects and thus the stimulation of knowledge creation.

The first case meets the criteria of both the information and the descriptor approaches in the context explication continuum introduced in the previous section. The second case is an example of a solution addressed by the meta-context approach. Motivated by the participatory action research theory (Whyte, Greenwood & Lazes, 1991), our selection of these two cases was based on their significance and the available information in order to achieve an appropriate reliability and validity (see Yin, 1994).

Combination of Full-Text & Attribute-Based Searches at the IWI

Within the Institute of Information Management (IWI) at the University of St. Gallen, we have several departments, each with two or more competence centers. Project managers lead these competence centers and are responsible for achieving their objectives. Each competence center produces many information objects, for example, lecture materials, presentations, and publications. These materials are stored in different systems, for example, file server, groupware-based office environments, or Web content management systems.

From a terminological point of view, all information objects have one thing in common when contextualizing the content: they all deal with specific topics, for example, knowledge management, enterprise application integration, business networking and so forth. Since most of the information objects are semi-structured and the maintenance of the metadata is manageable, the continuum of context explication led us to a hybrid approach. In order to reduce both the maintenance effort required to achieve the controlled terminology of an attribute-based approach and the risk of a misspelled full-text search, we chose a combination of the two.

Within an internal project we proclaimed "topic" as the most important descriptor in contextualizing the content of an information object for storage and eventual retrieval. Relevant topics had been locally collected from all the competence centers and stored in a single parameter database. As far as a specific topic, for example, portal, is concerned, the following contextual information is embodied in our definition framework: *Item* (preferred term for topic), *Assigned to* (responsible competence center), *Status* (draft, active, or frozen), *Synonyms* (similar terms or different languages) and *Description* (description of the term). A document history supports the traceability of modifications (Kremer, Kolbe & Brenner, 2003). These topics are subsequently used to classify information objects, for example, within our team databases or literature and publication applications.

The following challenges had motivated us to conduct the previously described project and to implement a combination of the full-text search and the attribute-based search:

- *Availability and access.* Users inside and outside the IWI should be able to search and to access IWI's information objects in an effective and transparent way. They had not been provided with any navigation and search mechanisms.
- *Consistent and controlled terminology.* Authors should be supported by a consistent and controlled terminology. Authors had defined information objects' metadata without following any organizational rules or standards, or they had not used metadata at all. This had led to an uncontrolled and not utilizable terminology and thus to no reasonable classification of information objects.

Currently, there are about 350 topics overall, owned by 30 competence centers. Approximately 11,000 documents have been classified according to the introduced topics for eventual retrieval through intranet and extranet portals. Having used this approach for almost a year, we have been able to derive the following success factors from our observations to solve the challenges:

- *Simplicity.* A "lean" context explication framework with only a few dimensions reduces the workload of information object classification, which increases user acceptance. Authors classify their information objects during the creation phase according to the introduced terminology. Thus, no subsequent reclassification and editorial work is needed.
- *Mixture of centralization and decentralization.* A few simple, centralized "rules" for topic definitions are helpful, for example, naming conventions. Decentralized, responsible team members make the detailed decisions regarding terms, thus reducing coordination overheads.

On the one hand, these success factors ensure the maintenance of the terminology with a small effort. On the other hand, users benefit from the manageable and consistent terminology during their search.

Topic Maps at the GCC

As pointed out in Smolnik and Nastansky (2002), groupware-based office systems provide an excellent environment for organizational knowledge management. Within the Groupware Competence Center (GCC) of the University of Paderborn, the GCC K-Pool (GCC Knowledge Pool) is used in almost every facet of operative work. It is a groupware-based repository for several kinds of information objects, which chiefly maintains information on books, conferences, links, media objects, contributions, articles, and software. The different information objects are enriched with numerous descriptors: categories are used to set information objects in different contexts, keywords describe the information objects in detail, and publishing information provides further explanation.

Even though there were many semantic relationships between the information in these databases, it was scarcely possible to navigate between them or to identify knowledge structures. The capabilities to access information objects are restricted to a basic full-text search and

navigating through context-sensitive views and categories. Full-text indexes are insufficient when searching for information, while structures, like document types or taxonomies, are sometimes too confining to qualify or categorize information objects (Biezunski & Newcomb, 2001). Furthermore, the usage and the scope of these techniques are limited to a single database.

Using the generic approach of applying topic maps to groupware-based organizational memories as described in Smolnik and Nastansky (2002), we applied the search and navigation concepts discussed in the introductory section of topic maps to the GCC K-Pool. We exhaustively defined a topic map template that comprises topic types as well as association types and describes the skeletal structure of the topic map. Typical topic types are *author, title, publisher,* or *place;* whereas typical association types are *author writes title, publisher publishes title,* or *publisher is located in place.* Software agents automatically create and maintain the topic map that is applied to the GCC K-Pool.

The GCC K-Pool topic map facilitates the creation of knowledge structures and meta- layers, the discovery of new contexts, and supports the users' cognitive capabilities. Users are furthermore able to search and navigate the GCC K-Pool topic map in several ways. A text-oriented Web browser interface also provides intuitive access. Additionally, users can explore the GCC K-Pool topic map by using two visualization tools: The K-Viewer, a two-dimensional approach with auto-layout capabilities for restructuring the topic map visualization, and the Sky Surfer, a three-dimensional approach with extensive navigation and search functions. These different topic map visualization approaches are described in detail in Smolnik, Nastansky, and Knieps (2003).

The GCC team consists of highly skilled researchers familiar with the topic map concepts and with expertise in the Center's everyday subjects. Users and authors are supported by a slightly distinctive taxonomy. These preconditions meet the criteria that are required for the meta-context approach of the context explication continuum. An excellent environment for the deployment of topic maps has therefore been created. We have used this approach for several months now and have observed the following main results:

- *Understanding of work contexts.* Users understand better how their work subjects are related when interrelating themes or information objects are utilized. They are able to explore the domain of interest in an intuitive way, and thus they are able to retrieve relevant and related information objects. The result of both observations is that users' creation of knowledge is stimulated.
- *Low maintenance.* Once configured and created, the topic map is updated automatically. Software agents insert new information objects, topics, and topic associations and delete outdated ones. A low effort of maintenance is important for the acceptance of such an IT system in an organization or its sub-units.

CONCLUSIONS

As shown, information objects' context and context explication play an important role in the area of knowledge discovery and portals. As there are several possible approaches, the real task for knowledge discovery begins with the selection of the appropriate solution for context explication. Consequently, we have illustrated three approaches — full-text search, at-

tribute-based search, and topic maps — each of which has been described according to its characteristics, benefits, and constraints as far as context explication is concerned, and aligned in our context explication continuum. Successful application of the specified criteria has been illustrated by the two implementations at our institutes.

FUTURE AREAS OF RESEARCH

To enrich our proposed model of context explication, we see at least four areas of future research. Firstly, we will further evaluate the distinctness of situations in terms of applying the continuum's elements by adapting GCC's topic map framework to IWI's content. Even if the preconditions are different, this might lead to insights into the degree of exchangeability of the different approaches. Secondly, we have to determine whether patterns can be found which will facilitate transition, for example, from topic maps to the full-text search or the full-text search to the attribute-based search. Thirdly, we would like to extend our continuum with indications regarding the implicit dimension of knowledge by including the explication of skills and skill management. Fourthly, we will validate and expand our findings with other external partners. We will particularly focus on industries other than academia to generalize the validity of the continuum. In addition, while we have focused on customer process-oriented portals, we will evaluate the application of the continuum to portals that are designed for other purposes. We therefore envision that knowledge discovery through context explication will provide a comprehensive framework with which to support knowledge management processes productively.

REFERENCES

Alavi, M. (1997). *KPMG Peat Marwick U.S.: One giant Brain.* Report Nr. 9-397-108. Boston, MA: Harvard Business School.

Alavi, M., & Leidner, D. (1999). Knowledge Management Systems: Emerging Views and Practices From the Field. *Proceedings of the 32nd Hawaii International Conference on System Sciences.* IEEE Computer Society Press, Los Alamitos, CA; Washington; Brussels; Tokyo.

Andersen, A. (1996). *The Knowledge Management Assessment Tool: External Benchmarking Version.* Chicago, IL.

Andrews, W. (2003). *Visionaries invade the 2003 search engine magic quadrant.* Report. Gartner, Inc., Stamford, CT.

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval.* New York: Addison Wesley.

Baeza-Yates, R., & Schäuble, P. (2002). Retrieving Information: A Discipline with a Tradition. *UPGRADE (The European Online Magazine for the IT-Professional), III*(3), pp. 3-4.

Benbasat, I., Goldstein, D.K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly, 11*(3), 369-386.

Berners-Lee, T. (1997). *Web Architecture: Metadata.* Retrieved May 7, 2004 from: http://www.w3.org/DesignIssues/Metadata

Biezunski, M., & Newcomb, S.R. (2001). Topic maps and XTM – A manager's overview. *Conference Proceedings of the Knowledge Technologies 2001.*

Biggam, J. (2001). Defining knowledge: An epistemological foundation for knowl-

edge management. *Proceedings of the 34th Hawaii International Conference on System Sciences.* Los Alamitos, CA: IEEE Computer Society Press.

Bristow, P., Dickinson, C., Duke, S., Henry, S., & Makey, P. (2001). *Enterprise Portals: Business Application and Tech.* East Yorkshire: Butler Group.

Cathro, W. (1997). Metadata: An overview. Retrieved May 7, 2004 from: *http://www.nla.gov.au/nla/staffpaper/cathro3.html*

Checkland, P., & Holwell, S. (1998). Action Research: Its Nature and Validity. *Systemic Practice and Action Research, 11*(1), 9-21.

Dale, A. (2001). Designing Taxonomies at Unilever. *Knowledge Management Review, 3*(6), 30-34.

Davenport, T.H., Harris, J.G., & Kohli, A.K. (2001). How Do They Know Their Customers So Well? *MIT Sloan Management Review, 42*(2), 63-73.

Davenport, T.H., & Prusak, L. (1998). *Working Knowledge: How Organizations Manage What They Know.* Boston, MA: Harvard Business School Press.

Davydov, M.M. (2001). *Corporate Portals and e-Business Integration.* New York: McGraw-Hill.

Dey, A.K., & Abowd, G.D. (1999). Towards a better understanding of context and context-awareness. *GVU Technical Report GITGVU -99-22.* College of Computing, Georgia Institute of Technology, Atlanta, GA.

Dias, C. (2001). Corporate Portals: A Literature Review of a New Concept in Information Management. *International Journal of Information Management, 21,* 269-287.

Dublincore. (2003). *Dublin Core Metadata Initiative.* Retrieved May 7, 2004 from: http://dublincore.org

Fayyad, U.M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery: An Overview. In U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth & R. Uthurusamy (Eds.) (1996), *Advances in Knowledge Discovery and Data Mining* (pp. 1-36). AAAI Press (American Association for Artificial Intelligence), Menlo Park, CA.

Felber, H., & Budin, G. (1989). *Terminologie in Theorie und Praxis.* Gunter Narr Verlag, Tübingen.

Fleisch, E., & Österle, H. (2001). Das tor zur IT-welt: Thesen zum erfolgreichen portaleinsatz. *Computerwoche Extra, 2,* S. 28-31.

Goldfarb, C.F., & Prescod, P. (2000). *XML Handbook – Anwendungen, Produkte, Technologien.* Verlag, München: Addison-Wesley.

Gordon, M., & Pathak, P. (1999). Finding Information on the World Wide Web – The Retrieval Effectiveness of Search Engines. *Information Processing and Management, 35,* 141-180.

Güldenberg, S. (1996). *Wissens management und Wissens controlling in Lernenden Organisationen: Ein Systemth eoretischer Ansatz.* Deutscher Universitäts-Verlag GmbH, Wiesbaden.

Gummesson, E. (2000). *Qualitative Methods in Management Research.* London: Sage Publications.

ISO, ISO/IEC 13250. (2002). *International organization for standardization.* Geneva. Retrieved May 7, 2004 from: http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=38068&ICS1=35&ICS2=240&ICS3=30

Kalakota, R., & Robinson, M. (2001). *M-business.* New York: McGraw-Hill.

Klemke, R. (2000). Context framework – An open approach to enhance organisational memory systems with

context modelling techniques. *Proc. of the Third International Conference on Practical Aspects of Knowledge Management* (PAKM 2000).

Kremer, S., Kolbe, L.M., & Brenner, W. (2003). Do You Know Your Terms? A Procedure Model for Terminology Management. *Proc. of the European Conference of Infor. Systems* (ECIS).

Lai, H., & Chu, T.-H. (2000). Knowledge Management: A Review of Theoretical Frameworks and Industrial Cases. *Proceedings of the 33rd Hawaii International Conference on System Sciences*. IEEE Computer Society Press, Los Alamitos, CA.

Latham, L. (2001). Web Content Management and Portals: Who's Doing What? *Research Note SPA-13-9670.* Gartner, Inc., Stamford, CT.

McGovern, G. (2001). *Why Metadata is Important.* Retrieved May 7, 2004 from: http://www.gerrymcgovern.com/nt/2001/nt_2001_10_01_metadata.htm

Murray, P.C. (1996). Information, Knowledge, and Document Management Technology. *Knowledge Management Briefs, 1*(2).

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation.* New York: Oxford University Press.

Pepper, S. (1999). Navigating Haystacks and Discovering Needles. *Markup Languages: Theory and Practice, 1*(4).

Polanyi, M. (1966). *The Tacit Dimension.* Gloucester: Routledge & Kegan Paul.

Probst, G.J.B., Raub, S., & Romhardt, K. (1999). *Wissen Managen – Wie Unternehmen ihre Wertvollste Ressource Optimal Nutzen.* Gabler, Wiesbaden.

Puschmann, T. (2003). *Collaboration Portale – Architektur, Integration, Umsetzung und Beispiele.* Dissertation. Universität St. Gallen, Difo-Druck, Bamberg.

Raghavan, P. (2002). Information Retrieval for Enterprise Content. *UPGRADE (The European Online Magazine for the IT-Professional), III*(3), 5-8.

Rappoport, A. (2002). *Web Search Engines: Users Surprising the Experts.* Retrieved May 7, 2004 from: http://www.searchtools.com/slides/baychi2002/index.html

Rath, H.H., & Pepper, S. (1999). Topic Maps: Introduction and Allegro. *Conference Proceedings of the Markup Technologies 99.*

Röhricht, J., & Schlögel, C. (2001). *cBusiness – Erfolgreiche Internetstrategien Durch Collaborative Business.* München: Addison-Wesley.

Schwarz, J. (2000). Mass Customization von Prozessen Durch Unternehmensportale. *Information Management & Consulting, 15*(2), S. 40-45.

Smolnik, S., & Nastansky, L. (2002). K-Discovery: Using Topic Maps to Identify Distributed Knowledge Structures in Groupware-Based Organizational Memories. *Proceedings of the 35th Hawaii International Conference on System Sciences* (p. 10). Los Alamitos, CA: IEEE Computer Society Press.

Smolnik, S., Nastansky, L., & Knieps, T. (2003). Mental Representations and Visualization Processes in Organizational Memories. *Proc. of the Seventh International Conference on Information Visualization (IV03) – International Symposium on Knowledge Domain Visualization* (IV03-KDViz). Washington: IEEE Computer Society Press.

Sowa, J.F. (2000). *Knowledge Representation: Logical, Philosophical, and*

Computational Foundations. Brooks/ Cole, Pacific Grove.

Sveiby, K.E. (1997). *The New Organizational Wealth: Managing & Measuring Knowledge-Based Assets.* San Francisco, CA: Berrett-Koehler.

Wargitsch, C., & Habermann, F. (1998). *IMPACT: Workflow-Management-System als Instrument zur Koordinierten Prozessverbesserung: Anforderungen.* Universität Erlangen-Nürnberg Bereich Wirtschaftsinformatik, Nürnberg.

Web/CKS. (1999). *Web Portals: Present and Future Positioning in the European Marketplace.* Report.

Whyte, W.F., Greenwood, D.J., & Lazes, P. (1991). Participatory Action Research: Through Practice to Science in Social Research. In W.F. Whyte (Ed.), *Participatory Action Research* (pp. 19-55). Newbury Park, CA: Sage Publications.

Wrightson, A. (2001). *Topic Maps and Knowledge Representation.* Ontopia AS.

Yin, R. (1994). *Case Study Research. Designs and Methods.* London: SAGE Publications.

Stefan Smolnik received a master's degree in computer science from the University of Paderborn, Germany (1998). He is working as a research assistant and lecturer at the Department of Business Computing of the University of Paderborn in the fields of knowledge management and collaborative computing. Mr. Smolnik is an international recognized specialist in the research domain of topic maps and has published several articles on the topics of knowledge and process management, and information visualization. He is a member of the groups such as: the German Computer Society (GI) and the Special Interest Group on Knowledge Management (SIG KM). He is chairman of the DNUG working group "Knowledge Management".

Stefan Kremer works as a consultant with The Information Management Group (IMG AG). He received a PhD in information management from the University of St. Gallen, Switzerland (2004), and a joint master's in electrical engineering and economics (Diplom-Wirtschaftsingenieur) from the University of Paderborn, Germany (2000). He worked as a research assistant and lecturer at the Department of Information Management of the University of St. Gallen in the fields of knowledge management, process analysis and design. Mr. Kremer has several years of experience in numerous industrial knowledge management, portal and search engine projects and has published several articles on these topics. He is a member of the German Computer Society (GI) and the German Notes User Group (DNUG). Since 1999, Mr. Kremer is a certified Lotus professional (CLP) and certified Lotus instructor (CLI).

After having worked as financial consultant, Dr. Lutz M. Kolbe studied information management at Brunswick Technical University, Germany, where he received a master's degree. He went on working on his dissertation at Freiberg Technical University, Germany, and the University of Rhode Island, USA. He received his PhD in 1997. After that he worked at Deutsche Bank in Frankfurt and New York where he became managing director in 2001. Since July 2002, he has led the Competence Center Customer Knowledge Performance (CC CKP) and teaches at AACSB-accredited University of St. Gallen, Switzerland. His research interests are customer relationship management, security management as well as advanced technologies in the residential environment.