Peer-to-Peer Information Workspaces in Infotop

Ronald Maier [1], Johannes Sametinger [2]

[1] Martin-Luther-University Halle-Wittenberg, D-06108 Halle, Germany
maier@wiwi.uni-halle.de
[2] Johannes Kepler University Linz, A-4040 Linz, Austria
sametinger@acm.org

Abstract. Knowledge workers collaborate in teams, networks and communities in order to accomplish knowledge processes. They have to be supported with adequate organizational as well as information and communication technological (ICT) infrastructures. From an ICT perspective, requirements have changed when compared to more traditional (office) work due to the considerable higher complexity of data, the focus on communication across the boundaries of corporate ICT infrastructures and the mobility of knowledge workers. This requires the systematic handling of context and substantially extended functionality for collaboration in the knowledge workers' personal workspaces.

In this paper, we outline typical knowledge processes and discuss ICT support for the personal management of information, of web content, of collaboration and of knowledge. We present Infotop, a tool that supports the creation and management of shared-context information workspaces and organizes knowledge resources in a peer-to-peer (p2p) architecture. We show how Infotop can be used to support typical knowledge work processes and discuss its dimensions, its user interface, its shared context workspaces, its architecture, and some thoughts on a prototype implementation currently under development.

1 Introduction

During the last decade, knowledge-intensity of products, services as well as organizational processes has substantially increased. Knowledge workers collaborate in teams, networks and communities and have to be supported with an adequate organizational as well as an information and communication technological infrastructure. A knowledge management system (KMS) promises enhanced support for knowledge work through an integrated combination of information and communication technologies (ICT).

So far, KMS in organizations provide technocratic and document-focused basic support for organization-wide information sharing. Even though corporate KMS are already advanced in many organizations, they rarely offer support for the design and management of knowledge workers' personal workspaces as well as advanced functions for knowledge sharing and collaboration. The need for ICT support of knowledge workers is rarely considered when corporate KMS solutions are developed. Consequently, actual KMSs often fail to convince users of their advantages. Goals of this paper are:

- to describe requirements for ICT support of knowledge work, especially personal information, collaboration and knowledge management, with the help of typical knowledge work processes,
- to review the current technological support of knowledge work, and
- to present Infotop, a tool that helps to overcome the shortcomings of current ICT.

Infotop is a personal workspace designed to help knowledge workers (1) to organize their personal information and knowledge resources and (2) to share context and collaborate on the basis of peer-to-peer (p2p) information workspaces.

In Section 2 we will discuss knowledge work, compare it to traditional work, introduce concepts of knowledge management and describe typical knowledge work processes. In Section 3 we will review traditional ICT and recent proposals for tools in the areas of personal information management, web content, collaboration and knowledge management (KM). In Section 4 we will present the main concepts needed for the creation and management of peer-to-peer information workspaces in Infotop and discuss how these are used to support the knowledge work processes identified in Section 2. Finally, we conclude the paper and give an outlook to directions of future work in Section 5.

2 Knowledge work

Knowledge represents the key concept to explain the increasing velocity of the transformation of social life in general and the way businesses and social institutions work in particular [4, 9, 19]. Employees' roles and their relationships to organizations have changed dramatically as knowledge workers have replaced industrial workers as the largest group of the work force. 60% of US organizations think that between 60% and 100% of their employees are knowledge workers [7]. Knowledge workers are well educated, creative and self-motivated people engaged in joint, complex problem-solving processes. Knowledge workers have to be supported with an organizational and ICT infrastructure in which knowledge work can be handled more effectively and efficiently. In the following, we will first contrast more traditional (office) work to knowledge work. We will then elaborate on a set of knowledge work processes for which ICT support is required.

2.1 Traditional work versus knowledge work

Knowledge work can be characterized by a high degree of variety and exceptions and requires a high level of skill and expertise [36]. Knowledge work consists of a number of specific practices, e.g., generating new knowledge, interpreting and representing it, producing and reproducing knowledge or, in Schultze's ([35], 50f) terms, practices of informing, such as expressing or extracting knowledge and experiences, monitoring, translating and networking. Knowledge work typically is also characterized by attributes such as mobility, flexibility, teamwork and the use of intellectual abilities and specialized knowledge rather than physical abilities ([35], 43).The increasing specialization means that knowledge workers have to work together in various kinds of groups and teams which differ in their social structure and interactions. Virtual teams, expert networks, best practice groups and communities complement traditional organizational forms such as work groups or project teams and aid collaboration between knowledge workers within and increasingly across organizations. When compared to traditional work, knowledge work can be characterized by stronger communication needs, weakly structured and less foreseeable processes, increased mobility of work spaces and the need for semi-structured data, e.g., hypertext documents, messaging and learning objects, experiences or skill directories.

From an ICT perspective, the main changes in the requirements occur due to the considerably higher complexity of data and the focus on organization-wide and inter-organizational communication and mobility of personally responsible knowledge workers. This requires the systematic handling of context and extended functionality for collaboration in the knowledge workers' personal workspaces (see Section 4.2). From an organizational perspective, process-orientation has been proposed to help organize knowledge work, especially the recent additions to business process management aimed at weakly structured knowledge-intensive processes which are characteristic for knowledge work, see e.g., [6, 12, 23]. Table 1 compares the traditional IS-related workspace of an office employee with the modern workspace of a knowledge worker and shows the changed requirements for the design of an ICT workspace for knowledge workers.

Table 1: Traditional Work versus Knowledge Work

	traditional work	knowledge work
orientation	data-oriented	communication-oriented
structure	highly structured, deterministic processes (pre-structured workflow)	weakly structured, less foreseeable processes (ad-hoc workflows)
data types	structured data (tables, quantitative data, "hard facts")	semi-structured data (links, hypertext documents, container, office workflows, "soft data")
data storage	(relational) data base management system	content-oriented "knowledge bases", experience data bases, newsgroups, mail folders etc.
data handling	coordination of accesses, integrity, con-	replication, information sharing, distribution

	trol of redundancy	of message objects, search and retrieval, valuation
boundaries	organization-internal focus	focus across organizational boundaries, alliances, co-opetition, (virtual) networks
organizational design	central organizational design	decentral organizational design
group structure	work group, department	project team, network, community
role	one job position per person	multiple roles per person
workspace	fixed workspace	mobile office (virtual office), multiple workspaces
IS design	top-down design of IS	bottom-up and top-down design of IS (prototyping, component-based, evolutionary software design)
equipment	personal desktop computer	laptop, personal digital assistant, mobile phone

2.2 Knowledge Management

Success of an organization is dependent on taking these changed requirements into account. An organization has to create an effective environment for knowledge generation and application and depends on the knowledge and talent it can recruit, develop and retain in order to provide value innovation [19]. Consequently, organizations need concepts and instruments that help them to establish such an environment. KM promises guidance in this matter and therefore has recently received increasing attention from a variety of fields and disciplines, see e.g. [21].

KM can be defined as the management function responsible for regular selection, implementation and evaluation of goal-oriented knowledge strategies that aim at improving an organization's way of handling knowledge internal and external to the organization in order to improve organizational performance. The implementation of knowledge strategies comprises all person-oriented, organizational and technological instruments suitable to dynamically optimize the organization-wide level of collective competencies, education and ability to learn [21]. KM initiatives can be described with the help of four levels of intervention:

- strategy: KM strategy and goals
- organization: roles, tasks and organizational culture
- contents and systems: KMS architecture, contents and functions
- economics: evaluation areas and evaluation categories

Moreover, recently process orientation has been viewed as a good starting point for formulating knowledge strategies and for redesigning the organizational and the ICT environment for KM (Maier 2002a). These levels of intervention can play a crucial role in the design of a knowledge environment for knowledge workers.

2.4 Knowledge work processes

Knowledge workers work together in knowledge-intensive business processes and in especially designed service processes, also called knowledge processes. The latter represent a (portion of a) knowledge life cycle consisting of the activities create (or externalize), value, organize and refine, store, distribute, search, apply as well as feedback/improve knowledge [1, 27, 41]. The life cycle steps can be combined to patterns of typical knowledge processes that are initiated and handled ad-hoc by knowledge workers. A typical knowledge process might handle knowledge externalized in a business process according to the life cycle steps value, organize & refine, store and distribute until it is re-applied in a different business process. Further examples of typical knowledge processes are (a) the acquisition of knowledge from outside the organization or (b) managing communities-of-interest or networks of knowledge workers. The following list provides examples for personal knowledge work processes corresponding to these knowledge processes.

- *Externalization process*
  is used e.g., to create and prepare a paper or a presentation for a conference co-authored by geographically dispersed knowledge workers who share a portion of their knowledge context, i.e., electronic knowledge sources. Examples of knowledge sources are documents, web sites, data bases, expert contacts.
- *Submission process*
  comprises the sub-processes submit, value, organize & refine as well as store knowledge. The process is triggered by an individual or a group of knowledge worker(s), evaluated by (members of) a community, e.g., a program committee, reviewed, refined and linked to other knowledge elements by a subject matter specialist. These value-added knowledge elements are occasionally repackaged for specific target groups. Finally, the target audience is granted physical and intellectual access to the submitted knowledge elements.
- *Distribution process*
  uses interest profiles specific to individual knowledge workers, so that, new knowledge elements as well as links to events, learning offerings, meetings or expert advice can be distributed according to these profiles. More generally, this process handles all distribution of knowledge to knowledge workers in geographically dispersed locations.
- *Search process*
  identifies and connects several steps of an individual or joint search for knowledge elements and/or expert advice by a group of connected knowledge workers. Crucial steps in a joint search process include the definition of search locations, the combination and weighing of personal preferences, and the amalgamation of individual search results.
- *Application process*
  integrates knowledge into the operative work processes and ICT environment of the knowledge worker, e.g., a scientist who uses the shared knowledge workspace to conduct research and to improve teaching.
- *Feedback and improvement process*
  comprises activities concerning the follow-up on feedback that one has gained through the comments on knowledge elements. Scientists' sources of feedback include conferences, newsgroup discussions, email on certain topics, etc.
- *Acquisition process*
  defines and handles the exploitation of external knowledge sources. Access to knowledge sources that have to be paid for is organized centrally and has to be integrated into the information workspace.
- *Community or network management process*
  supports the identification, foundation of and participation in communities-of-interest. For example, a scientist wishes to value, organize and integrate the communities' knowledge resources into her personal knowledge workspace.

3 Technological Support

There are many tools that support knowledge work. In this section we will present representative approaches in the categories of personal management of information, of web content, of collaboration, and of knowledge.

3.1 Personal Information Management

The desktop is the primary metaphor being used as interface on our computers. Thus, it not only manages personal information stored in files and folders, but it also serves as the main access point to personal information management systems, e.g., calendars, address books. A metaphor is one thing conceived as representing another. Using metaphors takes advantage of people's knowledge about them. For example, people in offices have been used to store paper documents in file folders. It makes sense to these people to store computer documents in folders on the computer, i.e., in containers that look and behave like folders. The desktop is the primary metaphor being used as interface on our computers. The desktop metaphor was introduced when computers were quite different to today's machines, see [15]. While computers, users and the environment have changed, interfaces and the basic handling of data have stayed the same, see Table 2. Today, the situation is quite different with professional users in addition to novices, a wide range of applications including web applications, rich resources, permanent network connections, and comprehensive communication features. The desktop has become an unmanageable mess [38]. Countless files are stored on increasingly more capacious storage drives. This has resulted in big hierarchies of folders that make it hard to retrieve information. The “find file” function sometimes helps in averting the desktop’s limitations. This function can be based on the files’ names, or even more helpful, on the files’ textual contents. Today, we not only have text documents, but images and multimedia files. Text-based searches do not help much in finding such files.

Table 2: Comparison of early and today’s personal computers.

	early personal computers	today’s personal computers
metaphor	desktop	
content handling	applications	
interaction	WIMP (windows, icons, menus, pointing device)	
structure	hierarchy	hierarchy + network
administration of collections of contents	files, folders	files, folders + applications with complex GUIs (browser, e-mail, explorer, etc.)
users	novice users (experts used command lines)	novice, intermediate and expert users
user interaction	weak i.e., keyboard, mouse, small b/w screen	strong i.e., keyboard, mouse, large color screen, speakers, microphone, video camera
software	small range primarily office applications	wide range including web applications
resources	poor i.e., 128KB RAM, 400KB hard disk, floppy disk drive (170KB)	rich i.e., 512 MB RAM, 80GB disk, zip, CD-ROM, DVD (hundreds of MB, GB)
connectivity	stand-alone	permanent, fast network connections + mobile devices with synchronization and replication needs

Alternatives to the desktop have been proposed to overcome the hierarchical file structure. Some of these alternatives try to completely replace the desktop, while others are placed on top of the desktop. Personal information management systems aim at organizing and visualizing the increasing volume of information that we have to handle. This should help to reduce knowledge workers’ information overload leading to more effective decisions and knowledge-related activities [13].

Desktop alternatives. An example for an alternative to the desktop metaphor is Lifestreams [14]. Lifestreams uses a time-ordered stream of documents rather than conventional files and folders. Incoming information is organized, located, summarized and monitored by stream filters and software agents. Each document that is created is also stored in a lifestream which functions as a diary. Work in progress or the latest e-mail messages are in front of the stream. The tail of the stream contains documents from the past. Moving beyond the present, the stream may also contain documents, e.g., reminders and calendar items. Similar time-ordered lists can be found in some applications, e.g., an e-mail list sorted by date or a web history list. [14]. Incoming information is organized, summarized and monitored by stream filters and software agents.

Another time-centric approach called ‚Time-Machine Computing’ has been proposed in [32]. Four key features characterize time-machine computing, i.e., lifelong archival of information history, chronological navigation over archived information (time-traveling), visualizing time in different ways, and inter-application communication of time (time-casting). A first desktop environment based on these concepts is called Timescape.

Presto is another approach to overcome hierarchical filing structures as the basis for organizing, storing and retrieving documents [8]. Presto is a prototype document management system that provides rich interaction with documents through meaningful, user-level document attributes. A uniform document model is provided for arbitrary kinds of documents to which attributes with arbitrary names and values can be attached. In Presto, documents may have multiple names, or no name at all. They may also appear in many places, or in none. The "pile" metaphor is being used for document appearance [8].

Desktop add-ons. PersonalBrain is a tool for managing information by visually organizing resources according to whatever scheme makes sense to the user. [37]. SixDegrees helps in managing the relationship network of personal information. It concentrates on messages, files and people, rather than on a strict file system hierarchy [5]. SemioTagger is a categorization and indexing engine that (semi-)automatically organizes online sources and documents of different data formats into meaningful categories. SemioMap then uses the detailed document profiles to create multi-layered concept maps [11].

3.2 Personal Web Management

In the World-Wide-Web we are drowning in data, but starved of information. Web tools provide the functionality of bringing some order into the multitude of web pages being visited by today's Internet users.

Web Browsers. The favorites or bookmarks mechanism of web browsers is a simple remedy for the administration of many web sites that are visited repeatedly. Spatial memory for document management is used in Data Mountain [33]. It allows users to place documents at arbitrary positions on an inclined plane in a three-dimensional desktop virtual environment.

Enhanced Web Support. TopicShop helps users to evaluate and to organize collections of web sites [2]. It provides support for finding web pages relevant to someone's interests. The Resource Description Framework (RDF) is part of the Semantic Web Activity. It is intended for representing meta-data about web resources, e.g., title, author, creation date of web pages [40]. This meta-data will allow software like web browsers to provide enhanced features for the organization and retrieval of information on the Internet.

3.3 Personal Collaboration Management

Personal collaboration management includes activities that are needed to communicate, coordinate and cooperate.

Workgroup Computing. A large number of ICT have been proposed to support work groups called workgroup computing, groupware or computer supported cooperative work. Groupware can be classified into communication, coordination and cooperation systems or along the two dimensions space (same/different location) and time (synchronous/asynchronous). Examples of groupware applications are [e.g., 10, 31, 39]: co-authoring systems, electronic discussion groups, electronic meeting systems, group calendars, group (decision) support systems, shared screen systems, teleconferencing systems or workflow management systems. These systems primarily support communication, coordination and decision making in groups as well as the joint handling of objects. A groupware platform combines many of these functions and provides general support for collecting, organizing and sharing information within (distributed) collectives of people. The best known groupware platforms are Lotus Notes and Microsoft Exchange [20, 26].

Information Agents. Information agents help to manage the explosive growth of information that we are experiencing. They perform the role of managing, manipulating or collating information from many distributed sources [28]. Software agents differ from traditional software systems with respect to their autonomy, ability to communicate and cooperate, mobility, reactive and proactive behavior, reasoning, and adaptive behavior [3]. Information agents can be used for many purposes, e.g., to scan email messages, to group and automatically update user-specific messages and information items, to search, integrate, evaluate and visualize information from many sources, to intelligently handle information subscriptions, and to identify and network experts or generally knowledge seekers and providers [21].

3.4 Personal Knowledge Management

There are a number of basic ICT that together form a corporate infrastructure for knowledge management. The situation as found in many organizations is that there is an advanced ICT infrastructure in place. This is regularly a solution based on a set of Internet technologies or based on a groupware platform like Lotus Notes. Consequently, knowledge workers are increasingly supported by advanced ICT systems. The ever-increasing pace of innovation in the field of ICT support for organizations has provided numerous technologies ready to be applied in organizations to support these approaches. Examples for information and communication technologies that are related to KM include Intranet infrastructures, document and content management systems, workflow management systems, business intelligence tools, visualization tools, groupware and e-learning systems.

Knowledge Management Systems. More recently, KMS have been proposed as an ICT platform that combines and integrates many, if not all of these ICT, i.e., a number of functions for the contextualized handling of knowledge in organizations [21]. Examples are the KMS platforms Opentext Livelink or Hyperwave Information Server [17, 29]. A KMS promises significantly enhanced functionality through an integrated combination of a substantial portion of the information and communication systems from a KM perspective. It should not be seen as a voluminous centralized database, but rather as large networked collections of contextualized data and documents linked to directories of people, roles and skills. A KMS provides intelligence to analyze these documents, links, employees' interests and behavior, offers support for personalized access to the knowledge base as well as advanced functions for knowledge sharing and collaboration. Figure 1 gives an overview of an ideal KMS architecture (see [21]).

The knowledge worker accesses the organization's KMS with the help of a variety of access services (I), that translate and transform the contents and communication to and from the KMS to heterogeneous applications and appliances. Main aim of the personalization services (II) is to provide a more effective access to the large amounts of knowledge elements and thus to avoid information overload. On the one hand, subject matter specialists or managers of knowledge processes can organize a portion of the KMS contents and services for specific roles or develop role-oriented push services, e.g., with a role-specific knowledge portal. On the other hand, both, the portal and the services can be personalized with the help of e.g., interest profiles, personal category nets and personalizable portals.

The core knowledge processes—search and retrieval, publication, collaboration and learning—are supported by knowledge services (III). These are key components of the KMS architecture and provide intelligent functions for (1) publication, the joint authoring, publication, structuring and contextualization of knowledge elements supported by workflows, (2) discovery of knowledge elements and experts with the help of search, mining, visualization, mapping and navigation tools, (3) collaboration between knowledge providers and seekers with the help of location and awareness management tools, community homespaces and experience management tools and (4) learning supported by tools managing courses, tutoring, learning paths and examinations. Knowledge services work on the basis of integration services, e.g., a knowledge repository which handles the organization's meta-knowledge describing knowledge elements that come from a variety of sources with the help of meta-data, e.g., person, time, topic, location, process, type. A taxonomy or a knowledge structure helps to meaningfully organize and link the knowledge elements and is used to analyze the semantics of the organizational knowledge base. These layers are based on infrastructure services (V), an Intranet infrastructure which provides basic functionality for messaging, teleconferencing, data (file server) and web content management as well as extract, transformation and loading tools which can be viewed in analogy to a data warehousing architecture. Inspection services support viewing of documents without the corresponding application, e.g., a text document without the text processing software that created the document. The data and knowledge sources (VI) give some examples of the wide variety of electronic sources which have to be integrated into a KMS.

Current KMS implementations are rather technocratic and document-centered infrastructure-oriented solutions. Thus, they are limited to the lower levels of Figure 1. Personalization services, easy access as well as integration into daily work practices are missing. Knowledge workers access knowledge services with the help of a number of isolated applications, such as a web browser, the desktop, a mailing system, office systems and a Groupware client.

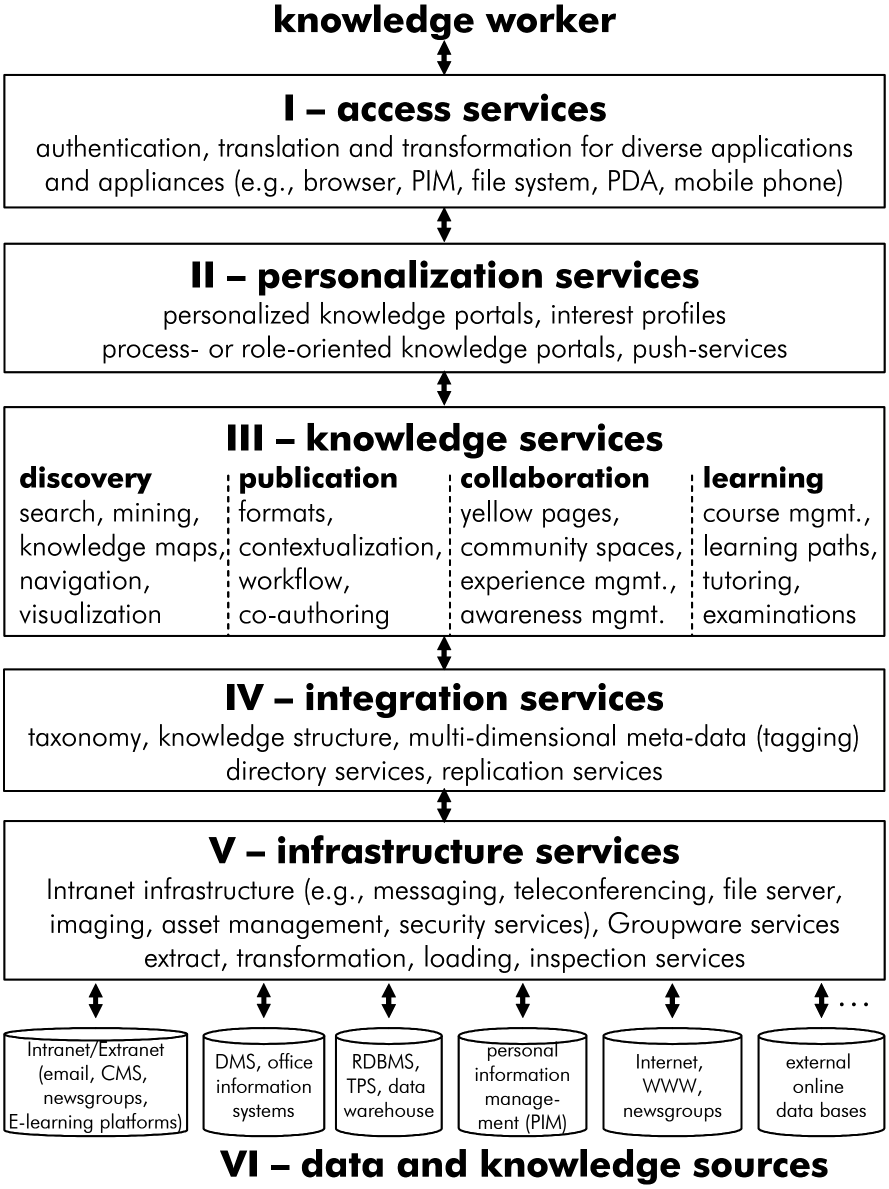

Figure 1: KMS architecture, for details see [21]

The efficient and effective use of a KMS, which is smoothly integrated with an organization's (knowledge-intensive) business processes, requires a systematic redesign of its architecture. Knowledge workers need a personalized knowledge environment that allows for an integrated multi-perspective view on knowledge services, on collections of contents and networked fellow knowledge workers as part of the organizational knowledge base. From the concept of a process-oriented strategic KM initiative we can derive some of the perspectives required, see [21]:

- organization: structure, business, work and knowledge processes (and projects), people and roles, groups, teams and communities,
- contents: types of contents / media, topics, knowledge structures, ontologies,
- systems: formats, integration of document, messaging, personal information, office, Groupware and KMS applications.

Corporate ICT infrastructures are quite advanced in many organizations, for empirical results see e.g., [21]. However, they rarely offer support for the design and management of the knowledge workers' personal workspaces as well as advanced functions for knowledge sharing and collaboration using these workspaces.

4 Infotop

Infotop is a metaphor for a shared-context information workspace. The term Infotop stresses in analogy to desktop to be "on top of the information" and thus covers the dynamic aspect, i.e., the flow of knowledge, in contrast to desktop. In this section, we will present six dimensions for the categorization and visualization of knowledge, shared context of collaborating users, the support of knowledge work processes, the proposed peer-to-peer (p2p) architecture, and some thoughts about a possible implementation.

4.1 Dimensions

The dimensions time (when?), topic (what?), location (where?), person (who?), process (why?) and type (how?) have been identified as being essential for effective categorization, visualization and navigation of collections of contents [24], see Table 3. We identify these dimensions to be essential for effective categorization, visualization and navigation of collections of contents. In analogy to OLAP (online analytical processing) techniques, these dimensions can be used for slicing, dicing, drilling down, rolling up, and ranging operations on contents of a personal knowledge environment. Several hierarchies of any of these dimensions can be used for display in addition to well-known visualization techniques like icons, thumbnails or lists, see Figure 2. Business intelligence software allows users to quickly analyze data that has been transformed into a subject-oriented, multidimensional data warehouse [18]. OLAP tools are used to perform trend analysis and statistics on e.g., sales and financial information in an interactive question-answer way

Table 3: Dimensions

Question	Dimension	Explanation
When ?	Time	any representations with a timed order
What?	Topic	any topics a user is interested in
Where?	Location	any geographic location like a city or country; local vs. lan vs. web
Who ?	Person	any person, physical or not, e.g., a company, an organizational unit
Why?	Process	any project or process, e.g., a conference, a paper writing process, an administrative task with many steps
How?	Type	any type of document, e.g., text document, Word document, audio, video

In Figure 2 information about sets of documents is displayed using dimensions and the pile metaphor [25]. Additionally, the numbers of documents are indicated for each displayed category. Views may be restricted to contents with specific attributes in any of these dimensions, e.g., contents of a specific process or of a specific age. Infotop's one-dimensional views are shown in [24], however, without considering knowledge work processes and shared context. Facts, i.e. the information on sets of contents represented in each cell, could be e.g., the number of elements as represented in Figure 2, the amount of data, e.g., the number of pages or Mbytes used, the number of contributions or of questions answered of knowledge providers, an aggregate valuation of elements, e.g., the number of accesses to elements, a measure of the skill levels of knowledge providers in a domain, or, in finer granularity, any other meta-information that is stored along with elements, e.g., the titles of documents, or a comparative measure, e.g., the proximity of competencies between a number of potential knowledge providers in a certain domain.

Time is one of the most crucial attributes of documents, e.g., time of creation, time of last modification, time of last read only access. Typically, only appointments are displayed in calendars, rather than e-mail messages, text documents and other forms of documents, e.g., comments, yellow stickers. It is also useful to display a selection of documents, e.g., we may want to see all documents of a project displayed in the calendar, or all documents of a person, i.e., all e-mail messages from and to that person, all files exchanged with that person, all web documents about that person that we have visited, etc. The meta-data for time and the other dimensions can easily be extracted from the context that comes with a content element or the activities that are performed on such an element, e.g., in the case of an e-mail message we can derive sender, receiver (person, location), date (time), subject (topic, process) and type of attached file (type).

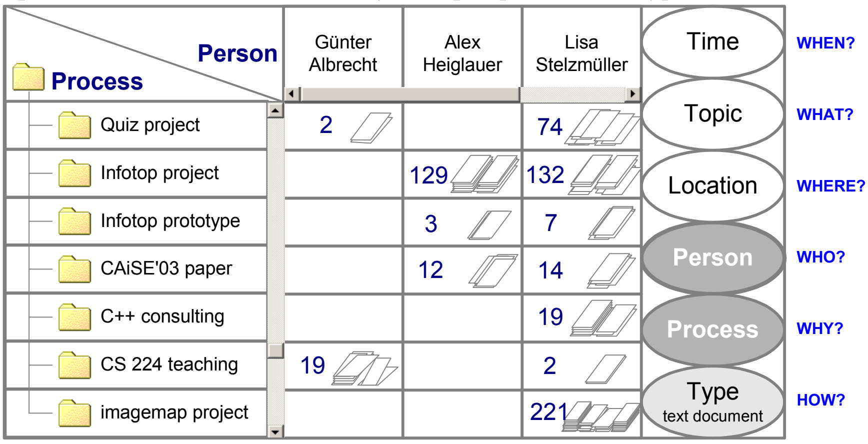

Figure 2. The two dimensions "process" and "person" are selected, the display is limited to text documents.

4.2 Shared Context

Users have information on their private computers and can also access public resources, typically on the Internet. Additionally, servers on local area networks provide extra information that is not accessible to the public, but to a restricted number of users only. We imagine a private, a protected and a public workspace for users. Private workspaces contain information that is stored locally on our computers and accessible only for the owner of the private workspace. Public workspaces include information that is published via the Internet and accessible by an undefined group of users. Protected workspaces lie somewhere in between. They contain information that is not accessible for everyone, but for whoever the owner grants explicit access, e.g., digital libraries. Private, protected and public workspaces of an individual can be placed on her computer, see user 3 in Figure 3. Additionally, user 3 shares parts of other users' workspaces. The dashed line and the gray boxes indicate her shared-context information workspace, i.e., a virtual workspace that includes her private, protected and public workspace as well as all public and parts of protected workspaces of other users. It is important to note that a user's protected workspace is not open to the public, but rather allows restricted access only to those individuals that she wishes. Thus, access privileges of the protected workspace have to be configurable in a flexible manner. Typically, public workspaces grant permission to read only, whereas protected workspaces may be open to write.

For a shared-context information workspace, we imagine to have private, protected and public workspaces institutionalized on all workplaces. Additionally, any information in these workspaces has to have meta-information attached, according to the six dimensions mentioned above, such that powerful query mechanisms can be supported. Assignment to e.g., topics is crucial for workspaces. This allows us to have several virtual workspaces for different topics of interest, i.e., several dashed lines in Figure 3. Virtual workspaces can overlap, because workspaces and sets of documents can be assigned to more than one topic.

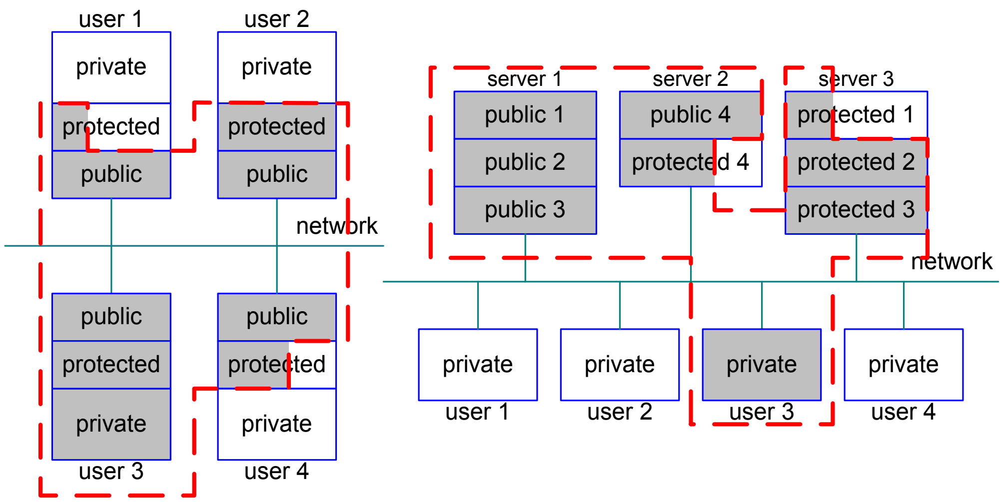

Figure 3. Four users have their private, protected and public workspace on their individual machines (left side), or distributed on server machines (right side).

In Section 4.1 we have used multiple dimensions to organize and visualize information of an individual, isolated workspace. In this section we have depicted a shared-context information workspace where users share parts of their workspaces. Organizing and visualizing this shared-context information workspace for each individual remains a challenging task. We argue that the multi-dimensional workspace can be used with minor modifications in a shared context.

Figure 4 shows how the dimensions of Infotop can be used to define shared-context workspaces and, thus, to distinguish private from protected information. Users 1, 2 and 3 all have access to their personal data store that is visualized by the data base symbol. The data store can contain text documents, personal information management documents, e.g., addresses, bookmarks, calendar with appointments, to-do-lists, hypertext documents, messaging objects, such as e-mails, contributions to newsgroups, multimedia elements, etc. Infotop provides access to the entire personal data store using the six dimensions. In Figure 4, the two dimensions process and topic are used to define shared-context knowledge workspaces with user 2. User 2 grants access to user 1 for all data in her data store that are assigned to "Infotop" and "Seminar DotNet Programming" in the process dimension and all data assigned to "KMS architecture" in the topic dimension whereas the "EBRP project" and the topic "software engineering" are not accessible to user 1. User 1 grants access to user 2 for all data in his data store that are assigned to "Infotop" and "Seminar DotNet Programming" in the process dimension and all data assigned to "KMS modeling" and "KMS architecture" in the topic dimension whereas the "KnowCom" process and the topic "KMS success" are not accessible for user 2. Consequently, workspace management is easily accomplished in a flexible manner by assigning instances of each of the six dimensions to (groups of) users.

The six dimensions are helpful, no matter whether the information is private or shared. They have been introduced to get rid of the rigid file hierarchy. The shared context should conceal network structures and stress the logical boundaries among knowledge elements. However, explicit consideration of workspaces and thus a seventh dimension may be necessary to visualize social networks and promote the sharing of context.

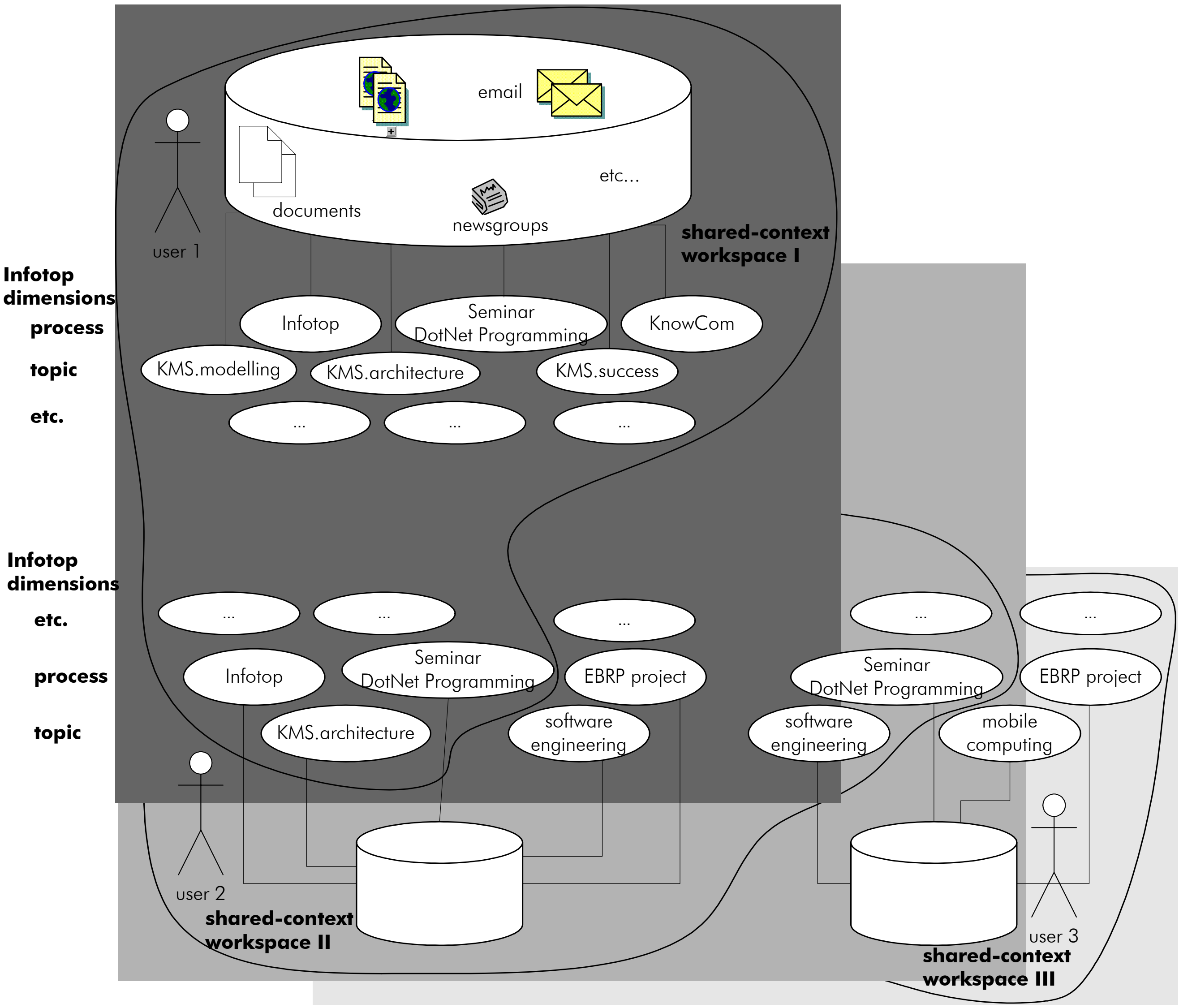

Figure 4. Shared-context workspaces in a peer-to-peer knowledge management system.

4.3 Knowledge Work Processes

In Section 2.4 we have outlined several knowledge work processes that are important for collaborating knowledge workers. Subsequently, we will outline how these processes can be supported by Infotop. A user externalizes, distributes, submits, acquires, searches, applies information in her shared-context information workspace.

– *Externalization process*

  Externalization of information is done with regular applications, e.g., a word-processing software, or (co-) authoring tools. This process results in documents that typically are at first stored in the private workspace. It is important to have meta-information attached to these documents. This is not sufficiently supported by today's applications. Infotop provides rich contextualization of documents using the six dimensions.

- *Submission process*
  In the simplest case, submission means publication of a new knowledge element and its distribution towards a topic-oriented network, i.e., in a protected or public workspace. Versioning of information and the support of workflows is required for the submission process.
- *Distribution process*
  The distribution process involves moving or copying information from one's private to one's protected or public workspace. It is useful to have this process combined with information subscription and some sort of notification, especially in the protected workspace.
- *Search process*
  Searching is done primarily based on meta-information in one's workspace consisting of one's private, accessible protected and public workspaces. Protected and public workspaces have to be prioritized according to topics, e.g., workspaces of research groups have to be considered only when the search process is aimed towards the research topics of these groups. Findings in protected workspaces are typically more relevant than findings in public workspaces.
- *Application process*
  The application process involves any usage of information that has been retrieved from an arbitrary source, i.e., from protected and/or public workspaces.
- *Feedback and improvement process*
  Responses or reflections to information in an arbitrary workspace can improve the quality of information. Feedback includes communication to information holders, i.e., workspace owners, citations, etc.
- *Acquisition process*
  The acquisition of information includes the extension of the search domain to include new workspaces, the location of information in any of the accessible workspaces and copying this information or a link to it into one's individual workspace.
- *Community or network management process*
  Communities share their interest in certain topics. It is necessary to have topic directories in public work spaces, where users can register and obtain permission to participate in protected workspaces that are assigned to these topics. The consideration of new topics results in new dashed lines, see. Figure 3. The acquisition of information is supported by the extension of one's workspace by including additional protected workspaces.

Figure 5 depicts these knowledge work processes involving the entire shared-context information workspace of a user. The solid ellipse in Figure 5 depicts the user's individual workspace, while the dotted ellipse depicts the user's shared-context information workspace.

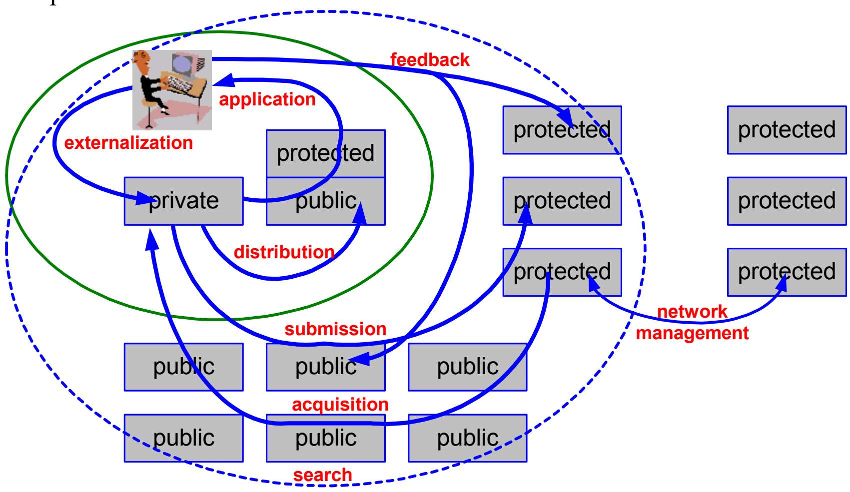

Figure 5. Knowledge work processes and workspaces.

Figure 6 shows how the knowledge work processes discussed above can be supported in a setting with a number of knowledge workers collaborating in a number of (overlapping) knowledge communities each using Infotop (see also [22]). In Figure 6, three communities are visualized. Communities correspond to shared-context workspaces in which a number of knowledge workers participate. In the terms of Figure 4, externalization of knowledge requires documentation of a knowledge element, organization according to the six dimensions and moving it into Infotop's knowledge base. Submission simply means that access privileges are granted to members of a community for instances of one or more Infotop dimensions. The search domain used in a search process consists of all locatable peers that have granted access to their knowledge base. Priority is given to those peers that participate in the same community the topic of which most closely matches the search term.

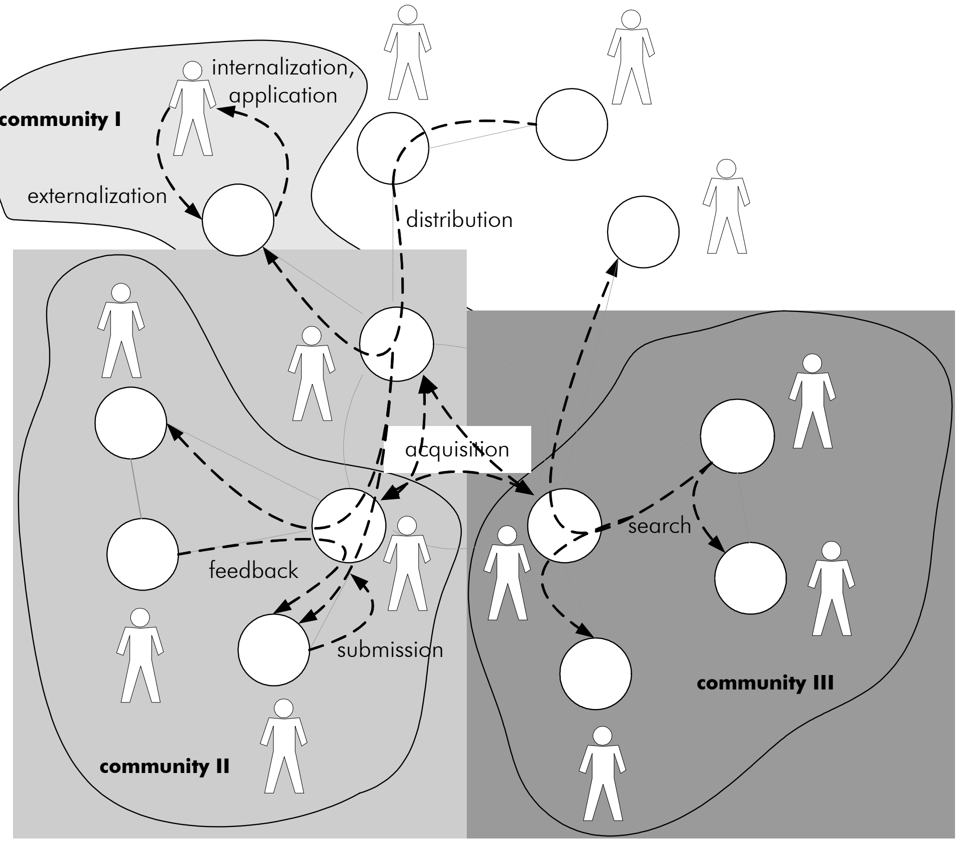

Figure 6. Knowledge work processes in a peer-to-peer knowledge management system.

4.4 Peer-to-peer architecture

Architectures in general play an important role in MIS as blueprints or reference models for corresponding implementations of information systems. The term architecture as used in MIS origins in the scientific discipline architecture and is used in a variety of ways: e.g., application architecture, system architecture, information system architecture and especially software architecture. There are basically three main sources for architectures describing the structure of KMS:

Theory-driven architectures. The first group of KMS architectures is the result of theoretic investigations which represent a theory-driven decomposition of an organizational knowledge base or organizational memory and derive ideal components of a corresponding ICT system. Zack classifies KM tools and systems into one of the following two segments: KMS with an integrative versus an interactive architecture [42]. This classification corresponds to the two main directions of KM research, human orientation and technology orientation, and the distinction of KM strategies into a personalization and a codification strategy [16].

Vendor-specific architectures. Secondly, vendors of KMS publish white papers in which they describe their perspective on knowledge management and place their tools in a knowledge management architecture that regularly pays attention to the ICT infrastructure already available in the organizations.

Market-driven architectures. A third group of authors applies a more pragmatic approach and empirically distills the most important components of an organizational knowledge management environment which is integrated with more traditional data and document management systems as well as communication systems. The authors mostly rely on the offers of (a number of) vendors of standard software tools, platforms and systems to support KM or analyze the individual KM environments of organizations that are regarded as KM pioneers and develop their own KMS solutions. These architectures are mostly layer models. The number, naming and inclusion criteria of the layers differ from author to author.

Recently, there are several attempts of KM researchers to profit from the promised benefits of a peer-to-peer metaphor for the design of an information sharing and especially for a knowledge management system, e.g., [30, 35]. The peer-to-peer metaphor promises to resolve some of the shortcomings of centralized knowledge management systems, e.g.,

- to reduce the substantial costs of the design, implementation and maintenance of centralized knowledge management suites, in terms of hardware, standard software as well as the often underestimated costs of designing, structuring and organizing a centralized knowledge server,
- to reduce the barriers of individual knowledge workers to actively participate and share in the benefits of a knowledge management system,
- to overcome the limitations of a knowledge management system that (almost) exclusively focuses on organization-internal knowledge whereas many knowledge processes cross organizational boundaries,
- to include individual messaging objects (emails, instant messaging objects) into the knowledge workspace that are rarely supported by centralized knowledge management system and, moreover,
- to seamlessly integrate the shared knowledge workspace with an individual knowledge worker’s personal knowledge workspace.

However, there are a number of organizational and technological issues that still have to be resolved before a p2p knowledge management infrastructure can be fully deployed in an organization. Examples are [35]:

- participation issue: there have to be incentives to actively participate in the p2p network in order to foster information sharing and avoid the free rider issue,
- trust issue: security and reliability of the p2p infrastructure have to be guaranteed if the system should be used as the sole, personal knowledge workspace of knowledge workers,
- coordination issue: structuring and quality management of the knowledge contained in a p2p network have to be supported in order to avoid information overload.

Infotop addresses the coordination and the trust issues. The participation issue is out of the scope of a technological solution. It should be no more of a problem than in centralized KMS within organizational boundaries. In p2p knowledge networks that cross organizational boundaries, (professional) communities along with personal contacts, contracts, shared goals and interests might act as a kind of social infrastructure that induces social regulations and also trust into the p2p network.

Figure 7 shows that the architecture of one of the peers includes the same layers as the centralized knowledge management architecture described in Section 3.4, but lacks a centralized knowledge structure, taxonomy and repository (see also [22] for a more detailed account of peer-to-peer KMS architectures). Personal data and knowledge sources are extracted, transformed and loaded into an integrated Infotop knowledge base. Inspection services support the access of documents without the applications that were used to create the documents. The integrated knowledge base comprises a private, protected and public area. A personal knowledge cache is used to optimize network traffic when shortly accessing the same knowledge elements multiple times. Due to the fact that knowledge workers might still at some time prefer to work offline, this knowledge base has an offline cache keeping those knowledge elements that are often needed on the local storage medium preferred by the knowledge worker. Just as in the centralized case, knowledge and access services build upon this integrated knowledge repository. The main difference is that the knowledge repository now is spread across a number of collaborating peers that have granted access to parts of their knowledge repositories.

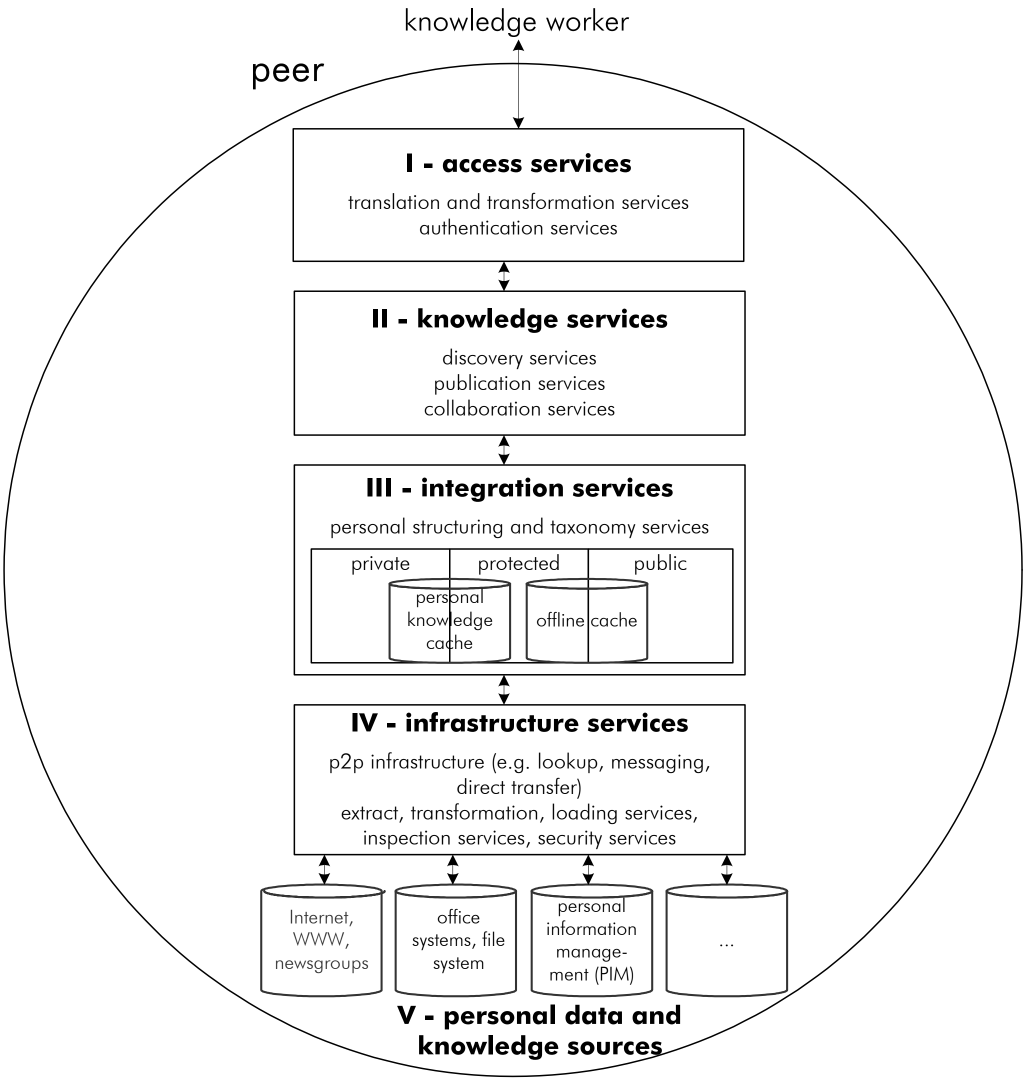

Figure 7. Architecture of a single peer.

4.5 Implementation

We imagine an implementation of a shared-context information workspace based on a combination of web service, data base, peer-to-peer and configuration management technology. Web service and peer-to-peer-technology can be used to seamlessly integrate other users' shared workspaces into one's own workspace in a platform-independent way. A data base is required in order to manage the meta-information created by Infotop. Configuration management and version control is needed to avoid versioning conflicts and to allow coordinated and cooperative work in the shared context. Also, Infotop has to exchange meta-information with other applications, e.g., messaging, office management and a search engine. The presentation of the workspace has to be modeled according to the six dimensions identified in Section 4.1. We are currently implementing a prototype for the creation of shared-context information workspaces.

For efficient document retrieval and for grouping of documents, meta-data has to be associated with documents according to the six dimensions described above. This can become a nuisance to the user, because she may not want to manually categorize each incoming and outgoing e-mail message, or each web page that she has visited. Therefore, we need an automated, or at least a semi-automated approach for this task. We imagine different attributes that should be defined for each document, e.g., title, author, date, event, location, person, process. Each attribute of a document has an undefined or a defined value, e.g., location= Dublin, date= 9/25/2002. We define one or more value sets, which we can easily switch when working on different processes. The meta-data can easily be extracted from the contextual information that comes with a document or the activities that are performed on a document, e.g. in the case of an e-mail message we can derive sender, receiver (person, location), date (time), subject (topic, process) and type of attached file (type).

5 Conclusion

We have discussed the differences between traditional work and knowledge work and have outlined typical knowledge work processes in which knowledge workers collaborate. ICT support has been found insufficient for the personal management of information, of web content, of collaboration and of knowledge. We have proposed peer-to-peer information workspaes and discussed six dimensions in a shared-context information workspace that support typical knowledge work processes. Infotop comprises the six dimensions time, topic, location, person, process and type. They were derived as essential perspectives on collections of contents of organizational knowledge bases. Furthermore, Infotop establishes a flexible way to share parts of the organizational knowledge base with other knowledge workers and avoids the shortcomings of centralized KMS due to the individual management of access privileges.

KMS are typically restricted to one organization's boundaries. A significant portion of knowledge work processes crosses these boundaries and thus can only be supported on the level of a personal knowledge workspace. We imagine Infotop as the main access point both for personal knowledge management and for ad-hoc collaboration in a shared context. It is important to include multiple ways to visualize the structure of elements in the dimensions, such as hierarchies, networks (knowledge maps) and geographical information systems in order to meet individual visualization needs. Another promising direction for future research is how to integrate personal KM techniques, e.g., portfolios, visualization of individual knowledge workers' knowledge status, learning and networking needs, with corporate KM instruments, e.g., content management, yellow pages, communities, project staffing or competence development programs. We see Infotop's role as an enabler and catalyst to spark usage of corporate KMS solutions and start a positive, reinforcing cycle of more and more active, motivated participants handling knowledge in organizations.

References

1. Alavi, M., Leidner, D. E. "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues". Management Information Systems Quarterly - MISQ, 25 (1) (2001) 107–136
2. Amento, B., Terveen, L., Hill, W., Hix, D.: TopicShop. Proc. 13th annual ACM symposium on User interface software and technology (2000)
3. Brenner, W., Zarnekow, R., Wittig, H.: Intelligent Software Agents. Foundations and Applications. Berlin et al. (1998)
4. Brown, J.S., Duguid, P.: Organizing Knowledge. California Management Review 40(3) (1998) 90–111
5. Creo Inc. http://www.creo.com/
6. Davenport, T., Jarvenpaa, S., Beers, M.: Improving Knowledge Work Processes. Sloan Management Review, 37(4) (1996) 53–65
7. Delphi Consulting Group: Delphi On Knowledge Management. Research & Perspectives on Today's Knowledge Landscape, Boston MA (1997)
8. Dourish, P., Edwards, W.K., LaMarca, A., Salisbury, M.: Presto: an experimental architecture for fluid interactive document spaces. ACM Transactions on Computer-Human Interaction (TOCHI) 6(2) (1999)
9. Drucker, P.: The Age of Social Transformation. The Atlantic Monthly 274(5) (1994) 53–80
10. Ellis, C. A., Gibbs, S. J., Rein, G. L.: Groupware. Some Issues and Experiences. C ACM. 34(1) (1991) 38–58.
11. Entrieva. Semio Map, SemioTagger, http://www.entrieva.com/
12. Eppler, M., Seifried, P., Röpnack, A.: Improving Knowledge-Intensive Processes through an Enterprise Knowledge Medium. Proc. 1999 ACM Conference on Managing Organizational Knowledge for Strategic Advantage, New Orleans (USA) (1999)

13. Eppler, M., Mengis, J.: The Concept of Information Overload. A Review from Organization Science, Marketing, Accounting, MIS, and related Disciplines. MCM Research Paper No. 28, University of St. Gallen, Media and Communications Management, St. Gallen (CH) (2002)
14. Fertig, S., Freeman, E., Gelernter, D.: Lifestreams: An Alternative to the Desktop Metaphor. Proc. CHI '96 conference companion on Human factors in computing systems: common ground, Vancouver, BC, Canada (1996)
15. Gentner, D., Nielson, J.: The Anti-Mac Interface. Comm. of the ACM, 39(8) (1996)
16. Hansen, M. T., Nohria, N., Tierney, T.: What's Your Strategy for Managing Knowledge?, in: Harvard Business Review, Vol. 77, No. 3-4, 1999, 106-116
17. Hyperwave Corporation. http://www.hyperwave.com/
18. Inmon, W. H.: "Building the Data Warehouse". New York, NY, 1992.
19. Kim, C., Mauborgne, R.: Strategy, Value Innovation, and the Knowledge Economy, Sloan Management Review 40(3) (1999) 41–54
20. Lotus, an IBM Corporation. http://www.lotus.com/
21. Maier, R.: Knowledge Management Systems. Information and Communication Technologies for Knowledge Management. Berlin et al. (2002)
22. Maier, R., Hädrich, T.: Knowledge Management Systems. Definition and Architectures, internal report, Martin-Luther-University Halle-Wittenberg, Dept. of MIS, Information Systems Leadership, Halle (Germany) 2003
23. Maier, R., Remus, U.: Defining process-oriented knowledge management strategies. Knowledge and Process Management 9(2) (2002) 103–118
24. Maier, R., Sametinger, J.: Infotop – An Information and Communication Infrastructure for Knowledge Work, ECKM2002, Proc. 3rd European Conference on Knowledge Management, Trinity College Dublin, Ireland (2002)
25. Mander, R., Salomon, G., Wong, Y.Y.: A 'Pile' Metaphor for Supporting Casual Organization of Information. Proc. ACM Conf. on Human Factors in Computing Systems (CHI '92, Monterey, CA, May 3–7), P. Bauersfeld, J. Bennett, and G. Lynch, Eds. ACM Press, New York, NY (1992) 627–634
26. Microsoft Corporation. http://www.microsoft.com/
27. Nissen, M., Kamel, M., Sengupta, K. "*Integrated Analysis and Design of Knowledge Systems and Processes". Information Resources Management Journal. 13 (1) (2000). 24-43
28. Nwana, H. S. Software Agents: An Overview. Knowledge Engineering Review (1996)
29. Opentext Corporation. http://www.opentext.com/
30. Parameswaran, M., Susarla, A., Whinston, A. B.: P2P Networking: An Information Sharing Alternative. IEEE Computer, 34 (7) (2001), 1-8
31. Petrovic, O.: Workgroup Computing. Heidelberg (1993)
32. Rekimoto, J.: Time-machine computing: a time-centric approach for the information environment. Proc. 12th annual ACM symposium on User interface software and technology, Asheville, North Carolina (1999)
33. Robertson, G., Czerwinski, M., Larson, K., Robbins, D.C., Thiel, D., van Dantzich, M.: Data mountain. Proc. 11th annual ACM symposium on User interface software and technology (1998)
34. Schultze, U.: On Knowledge Work, in: Holsapple, C. W. (2003) (ed.): Handbook on Knowledge Management, Vol. 1, Berlin et al. 2003, 43-58
35. Susarla, A., Liu, D., Whinston, A. B.: Peer-to-Peer Enterprise Knowledge Management, in: Holsapple, C. W. (ed.): Handbook on Knowledge Management 2. Knowledge Directions, Berlin et al. (2003), 129-139
36. Sveiby, K.-E.: The New Organizational Wealth. Managing and Measuring Knowledge-Based Assets, San Francisco (1997)
37. TheBrain Technologies Corporation. http://www.thebrain.com/
38. Claire Tristram. "The Next Computer Interface". Technology Review, December 2001. http://www.techreview.com/articles/tristram1201.asp
39. Valacich, J.S., Dennis, A.R., Nunamaker, J.F.: Electronic Support for Meetings: The GroupSystems Concept. International Journal of Man-Machine Studies. 34 (1991) 261–282
40. W3C: Semantic Web Activity. http://www.w3.org/2001/sw/
41. Wiig, K. M.: Knowledge Management Foundations. Thinking about Thinking. How People and Organizations Create, Represent, and Use Knowledge. Arlington (TX, USA) (1993)
42. Zack, M. H.: Managing Codified Knowledge, in: Sloan Management Review, Vol. 40, No. 4, Summer 1999, 45-58