A Four-Layer Model for IT Support of Knowledge Management

Matteo Bonifacio, Thomas Franz, Steffen Staab

Abstract: Strategies for knowledge management as well as strategies for IT-systems supporting knowledge management have often been conceived to serve one particular knowledge management paradigm. We argue that in practice there is not one paradigm that suits all needs of all organizations. Considering knowledge management as a measure to further communication in an organization we find that different ways of organizing communication exist - some of which are not well supported by current IT-systems for knowledge management. In particular, we investigate two dimensions of communication, first, the vocabulary or knowledge structures used to organize, i.e. annotate, store and retrieve, the knowledge using means such as Semantic Web technology, second, the way of organizing access to the knowledge in a more or less distributed manner involving technologies like distributed databases, agent-based or peer-to-peer based systems. Based on these dimensions we overview existing knowledge sharing systems and elaborate on novel knowledge sharing models and tools. Thus, we derive a four layer model for IT support of knowledge management, where only one layer for centralized knowledge sharing is well understood, where the layers of individual and decentralized knowledge management are currently intensively researched and where the forth layer of evolutionary knowledge management is just appearing on current research agendas.

Keywords: IT Knowledge Mangement, Semantic Web, Knowledge Sharing, P2P

1 The Organization is a Network

The core objective of knowledge management is the creation of value for an organization by establishing a principle way of handling knowledge in the organizations. Over the last two decades a large number of organizational measures have been proposed or rediscovered to meet this objective. Organizational means for this purpose range from creating coffee corners, over discovering and nourishing communities of interest, to establishing new comprehensive, organizational learning procedures. The essence of all these organizational experiments and changes was that the knowledge fabric of the organization is not depicted by the organizational chart and it is not found in the established business processes, but rather it is constituted by the *network of people* communicating their knowledge along and beyond established organizational domains and processes. Organizational measures towards knowledge management harvest added value from such networks.

Intriguingly though, IT support for knowledge management has hardly kept pace with the richness of knowledge communication established in such human networks. Either, it has just been used as is, like e-mail, in order to facilitate human communication without considering special needs for knowledge management. Knowledge management support here would have considered knowledge life cycle support: creation, organization, sharing, use, and evaluation of knowledge. Or, it included knowledge life cycle support, but was mostly bound to *one* dominating model of knowledge communication: *centralized sharing*.

The centralized sharing model as depicted in Figure 1 is quite simple. Knowledge is evaluated by the individual. It is then contributed to a central information repository, e.g. a knowledge database. It is organized into categories and can be retrieved by a designated audience. Typically, there are curators that take care of initial filling, that organize community assessments and maintain the basic structures. The centralized model of sharing knowledge is wide-spread, it is powerful – but it covers the wide variety of knowledge management needs in human networks only to a very limited extent.

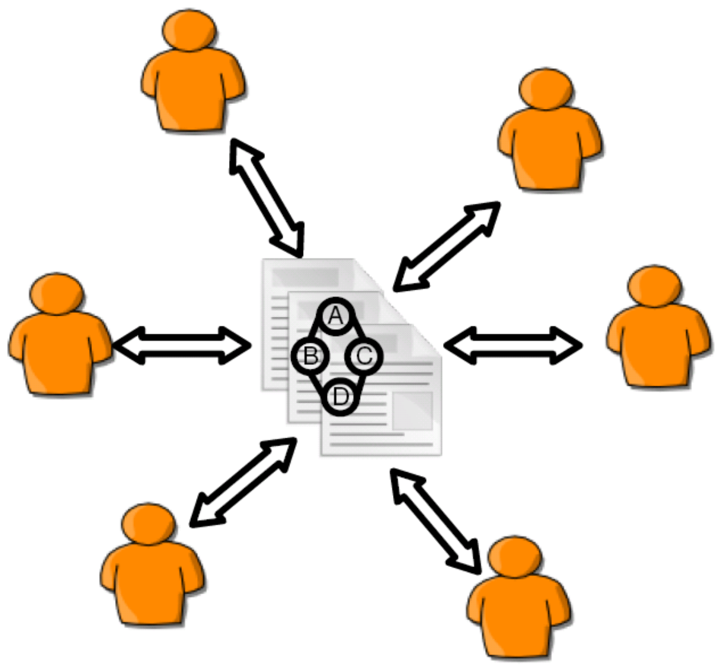

Figure 1: Centralized Sharing of Knowledge

In this chapter we investigate the IT needs of organizations with explicit knowledge management from a human network perspective. Furthermore, we assume that IT structures work best if they do not get into conflict with organizational structures, but if they reflect as closely as possible the underlying organizational needs and practices. For this purpose, we start with a consideration of various organizational practices and relate them to knowledge management efforts in Section 2. This perspectives let us resolve the one model of centralized sharing of knowledge into a layered model with varying degrees of knowledge sharing as discussed in more detail in Section 3. In Section 4, we use the layered and more fine-grained model to survey the space of established, new and forthcoming IT solutions and technologies for knowledge management[1].

2 The Organization Evolves

The very notion of organization and its related field of inquiry, emerged when, after the so called scientific management experience, researchers and practitioners felt the need to overcome the one best way approach. According to this view, the administration of an organization could be done, as much as a laboratory experiment, according to scientific principles. Moreover, management was seen as the function that centralized organizational knowledge; by means of scientific principles, managers could organize work optimally since they could possess all the needed knowledge. Individuals' knowledge and experience was very important but in a particular sense: through the systematic measurement and observation of working practices, managers could collect best practices and spread them to the overall organization. Employees were asked to actively engage in such process since those productivity gains obtained through knowledge reuse would become a benefit for all the company's stake holders. Knowledge was a core asset that was systematically codified and disseminated by means of methods and procedures.

The organizational debate rose when scholars started questioning this monolithic model. The debate produced a wide range of positions, options and alternatives that will not be analyzed here. For the purpose of this chapter, we want to underline some common points that can provide a useful framework to formulate our argument on KM. First of all, although from different perspectives, researchers agreed that the main goal of an organization science should have been the one of studying differentiation rather than conformity: namely, the fact that different models of organizations exists and, further, that within the same organization different parts or units may work according to different models. Second, the various models proposed hoovered along a continuum that went from the formal, centralized, and specialized model proposed by the scientific management (the so called mechanic model), to more decentralized, unstructured and despecialized models (the so called organic models). Third, choices along this continuum should be judged in terms of appropriateness, rather than in terms of ideological opposition. That is, each model is appropriate under particular conditions. Fourth, the models express different solutions in gathering and processing knowledge about the world, whatever these terms (knowledge and environment) mean. From this perspective, while the mechanic orientation seems to be more suitable when the "outside" is more stable (and knowledge more standardizable), the organic one seems to be more appropriate the more the outside is uncertain (and the more knowledge becomes contextual). As a corollary, when the environment changes, the organization (or part of it) may evolve accordingly.

[1]Being aware of the different notions of knowledge management, we would like to clarify that the approaches desribed in this chapter represent codification-based knowledge management systems that enable the exchange and personalization of information.

Interestingly, the recent KM debates have had a similar origin but didn't follow a similar evolution. In fact, on the one hand it started assuming a one best way approach that focused on the centralized codification, standardization and reuse of knowledge. On the other hand, when alternative models have been proposed, these where intended to function more in terms of contraposition than in terms of appropriateness. As an example, the communities of practice approach (that is quite consistent to the organic approach) has been proposed as a view on organizational knowledge dynamics which is alternative to the centralized one (similar to the mechanic). In this chapter we propose that KM, instead of argueing for or against a particular approach, should follow the more fruitful route that was originally taken by the organizational debate. That is, to view the various models of KM as responses which are to be judged as appropriate to different conditions and, as a corollary, to view KM models as part of an evolution that occurs within and across different organizations. We will refer to this approach as to evolutionary KM.

In particular, we propose that the typical contraposition between decentralized KM approaches, that privilege organic issues such as informality and lack of structure, and centralized ones, that privilege mechanic issues such as formality and standardization, should be re-interpreted in terms of appropriateness. From this perspective, *Local knowledge* is needed to generate appropriate solutions to heterogeneous and changing environments. *Centralized knowledge* is needed to share these solutions among different contexts. Both levels should be preserved; without locality, conformity might surmount. Without centrality, each wheel might be reinvented. A process is needed in order not just to preserve the two levels, but rather to connect them as phases of an evolutionary process: what is locally valuable knowledge should be consolidated centrally when a new domain stabilizes to sustain economies of reuse; what is centralized should be re-localized when what was known comes into crisis and locally re-elaborated. KM technologies should be consistent with these challenges and need to bridge between centralized and localized technologies and must allow for evolution of the organization.

3 KM System Classification by Degrees of Networking: Requirements and Use Cases

The model for centralized sharing of knowledge constitutes only one point in the space of possible solutions for knowledge management. In particular in professional working environments, it focuses on one kind of communication that comes with a considerable set of assumptions on the characteristics of the communication:

- Communication is mostly *unidirectional*: more people read than write.
- Shared knowledge is *mature* knowledge: immature knowledge generated on the fly is mostly not dealt with. Shared knowledge is expected to be evaluated locally before it is published and further examined by dedicated

authorities, e.g. moderators, or editors.

- Sharing of knowledge is an *active* act: the sharing of knowledge requires that one actively writes and publishes a memo, a web page or a data entry.
- Knowledge management happens in a *semi-open* domain: access is not restricted to one individual, but is extended to a group such as specified by corresponding access rights.
- Knowledge management happens in a *semi-closed* domain: the sharing of knowledge is exactly confined to the group with formal access rights. The *vocabulary* used is *shared* knowledge by the audience.

All these assumptions may be valid, but they may be questioned, too. Seeing knowledge management through the glasses of knowledge management as communication in human networks, an extended model will topicalize the individual node in the network as well as the topological structure of the network and the way it evolves. A corresponding picture of knowledge sharing may then distinguish (at least) four layers of knowledge sharing as depicted in Figure 2 and explained in the following.

3.1 A Layered Model of Sharing Knowledge

The layering of our model shown in Figure 2 should not be interpreted in its strictest sense, i.e. in a sense that each lower layer is a premise for the next layer. However, it is common that one has to manage individual knowledge first, before it comes to sharing of knowledge. Second, centralized sharing is a special case for third, decentralized sharing (and hence, we argue, precedes it). Fourth, the evolutionary model presumes that there is some minimal model of (de)centralized sharing in order that evolution may have a signficant basis and may be useful.

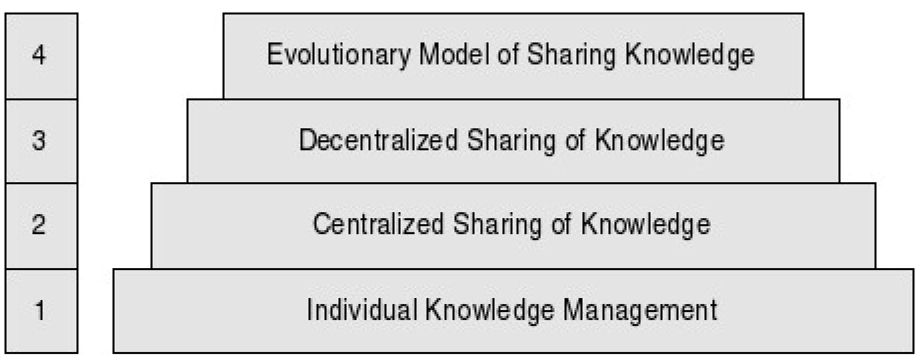

Figure 2: A Consolidated, Layered Model of Sharing Knowledge

3.1.1 Individual Knowledge Management

At the lowest layer, people have found over the last years that the *individual* needs more powerful IT systems to share knowledge *with himself or herself* in the sense that the individual organizes his information for later retrieval by himself or herself. There are plenty of digital knowledge creation and communication systems that the individual uses: e-mail, instant messaging, group chats, document repositories, portals, disk drives, bookmarks, databases (from local address books to large ERP databases), web pages, text documents, spreadsheets, PDF documents, images, videos, calendars, SMS, voice-over-IP, fax, etc. While large companies have worked hard to digitize their major business processes starting with scanning all incoming papers in order to avoid media ruptures (e.g. for managing car insurance claims), such smoothness is often not found when it concerns more ad-hoc work of less regularity and when it concerns the individual desktop leading to redundancy of data (e.g. telephone numbers on different media, such as mobile phone book vs PC address book), low quality of data (e.g. outdated phone numbers) and the difficulty or impossibility to retrieve the data easily (e.g. to search for a phone number by “Smith” with one search in multiple sources). This led to a recent increase of research and development in the areas of Personal Information Management (PIM) and desktop search[2] (cf. Sect. 4.1.1 for further explanations).

A typical **use case** for individual knowledge management lies in organizing and retrieving project communications from ones desktop. Such project communication may involve mails, appointments or documents. There are partial solutions for integrating some of these concerns, e.g. by Microsoft’s Outlook for organizing mail and calendar or by Google™ Desktop Search for searching mail and desktop, but these need to further develop to fully cover individual knowledge management needs (as we will also see further in Section 4).

While individual knowledge management benefits the individual, the organization wants to integrate processes of local knowledge management, e.g. knowledge creation, into organizational procedures: First, to exploit the benefits of individual knowledge management efforts, and second, to ensure appropriate integration by some corporate process instead of ``accidental’’ sharing. Hence, means such as apprenticeship have been invented long ago and knowledge management as an organizational topic (and also coined organizational learning) has required and will always require the next layer.

3.1.2 Centralized Sharing of Knowledge

The validity of this model *for the appropriate set of use cases* is not jeopardized by the need for other models. The centralized sharing of knowledge is facilitated by plenty of established software systems, server-based and typically organized around documents, sometimes organized around folder structures, taxonomies or similar kind of metadata. Folder structures, vocabularies, taxonomies, ontologies etc. denote differently sophisticated and complex means to define

[2] E.g. Copernic™ or Google™ desktop search.

knowledge and organize information. Throughout this chapter, we use the more general notion of *knowledge structure (KS)* to summarize any such means.

A typical **use case** for this model is demonstrated through the creation and maintenance of best practice knowledge bases. Here, business practices are reconsidered along the core business process (e.g. at project touchdowns), distilled for interesting experiences and then fed into a centralized best-practices knowledge base. Such a knowledge base is exactly characterized by the dominance of asynchronous write-read, mature knowledge, active sharing, and semi-open and semi-closed domains as elaborated above.

This approach encounters its **limits** when it has to seek a compromising position, where the compromise runs counter to the structures in which the human communication network works most productively. For instance, when researchers develop their own new way of structuring knowledge, it may be counterproductive or at least of little help to make them shape their thoughts into predefined templates or vocabularies – such as required by the model of centralized sharing of knowledge. Thus, at this point the view of the individual and the organizational compromise get into conflict and require a next layer of knowledge sharing.

3.1.3 Decentralized Sharing of Knowledge

The next layer is about the decentralized sharing of knowledge as indicated by Figure 3. This model circumvents a knowledge server bottleneck. Instead it allows for sharing all the knowledge (one wants to share) that one has on one's own desktop. It allows for a peer-to-peer communication structure that reflects individual expertise and knowledge needs and standard organizational processes. It requires from the system that it allows for distributed authorization and authentication of information access and for distributed search and knowledge organization. Typical system infrastructures are (but need not be exclusively) peer-to-peer systems or agent infrastructures.

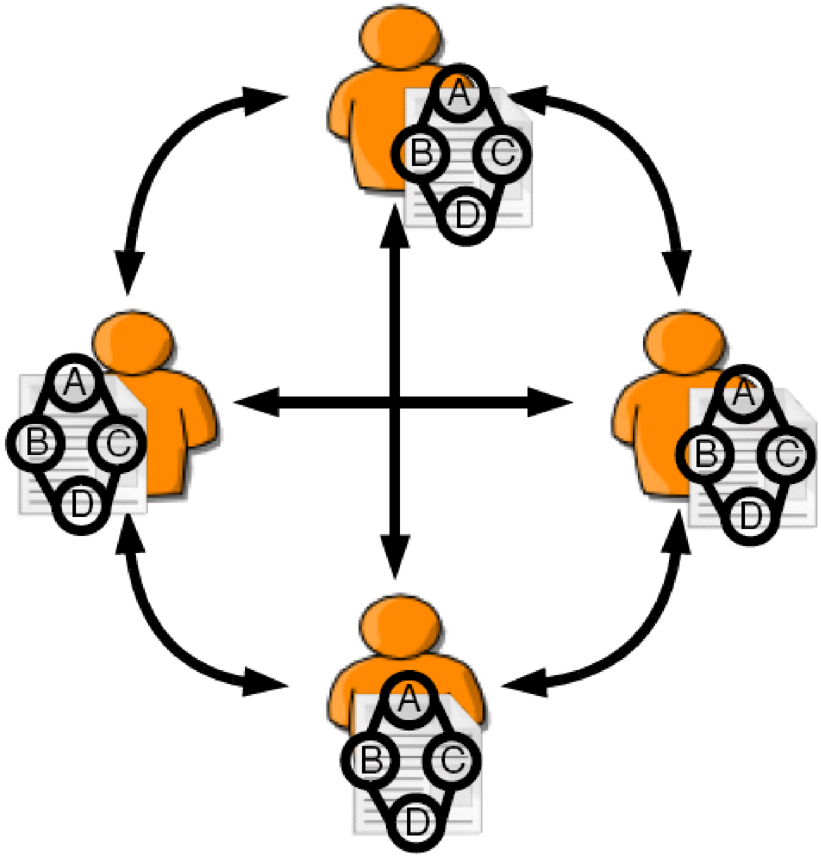

Figure 3: Decentralized Sharing of Knowledge

A typical **use case** for decentralized sharing arises during project work spanning several organizations. For a centralized system in such a setting an important question would be who controls the server, pays the associated costs and after project touchdown has access to the knowledge. In a peer-to-peer organization new project members are just added to the peer-to-peer network giving them rights to provide and search knowledge. Each project member benefits from the advantages of individual knowledge management as she personally controls and organizes the shared knowledge. In comparison to centralized sharing environments, decentralized sharing allows for a desktop-like user experience where the burden of up- and downloading information, common in centralized systems, is not applicable.

This approach requires the location of knowledge and it still expects that the various knowledge structures of different participants are either sufficiently similar or that they can be translated back and forth (a problem that the centralized sharing of knowledge never has, because it ``allows‘' only one way of structuring knowledge). It hits its **limits** when vocabularies and the meaning of a vocabulary change too fast – which often happens when also the social and the communication structures change very quickly. Then, the need for a new layer of sharing knowledge arises that reconciles between the centralized and the decentralized model of sharing knowledge:

3.1.4 Evolutionary Model of Sharing Knowledge

The evolutionary model of sharing knowledge draws from communication structures, knowledge structures and content in order to propose revisions to the communication in the network and revisions to the structures and vocabularies used for organizing and communicating knowledge (cf. Figure 4). Thus, it needs to reconcile between the individual views of layer 1, pre-existing standards such as established for layer 2 and it needs to use the communication structures such as established for layer 3. In doing so, the evolutionary model of sharing knowledge does not abolish the need for layers 1 to 3. Rather it is to be used when more traditional means (if layer 3 can be at all be categorized into this phrase) fail to provide the appropriate solution.

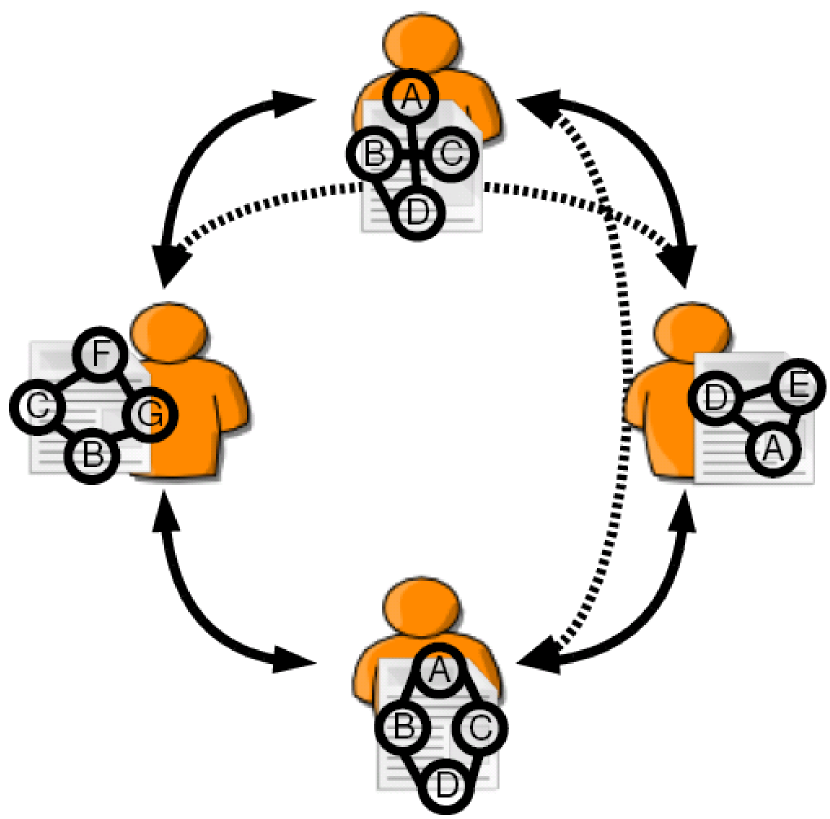

Figure 4: Evolutionary Model of Knowledge Sharing

A typical **use case** for evolutionary sharing exists with scientific research in computer science. Scientific research can be characterized by the need to share a lot of knowledge (e.g. about latest research papers), by the need to cope with large and quickly evolving vocabularies and by the need to share and cooperate far beyond organizational boundaries based on mutual trust and previously successful cooperations. Thereby the communication network structure is in constant flux, because the interests of people move in different directions and the need to find and use new methods and tools is ubiquitous. Not accidentally, the need for IT tools such as Grid environments (Foster, Kesselman, 2004), which are about sharing data and computational resources and building virtual organizations, has been pushed by academic institutions.

Therefore there is the need to manage trust by the individual rather than to have it managed by an overarching organization. The evolutionary model requires complex computations at each peer as vocabularies must be reconciled, learned, and mapped. Correctness and completeness of match-making are not guaranteed and thus may reduce the quality of the shared data.

3.2 Dimensions for Knowledge Sharing in Human Networks

We have now provided a set of criteria that layout the space of IT solutions for knowledge management based on different organizational models: *local, centralized, decentralized,* and *evolutionary*.

To better capture the differences between various systems and tools we need to distinguish at a more fine-grained level *what* is actually local, centralized, decentralized or evolving. When communicating knowledge via IT systems we see two main dimensions that vary along these values:

First, the knowledge structures (including the vocabulary) used to organize knowledge may be available (only) *locally*. There may have been a kind of (simple) standardization effort to agree on a common, *centralized* knowledge structure, there may be a plurality of such knowledge structures that are *decentral* and are mapped to each other, or there may be a continuously *evolving* set of knowledge structures the overview of which needs to be automatically changed and adapted all the time.

Second, the ways to actually access knowledge may vary. Knowledge may only be accessible if it is *local*. Or, it may be accessible if it lies on a *central* server. Or it may be accessible in a well-defined *decentralized* mode. Finally, it may be available even if one needs to bridge several trusted relationships in an *evolving* network. Thus, we have a classification scheme for two dimensions with four different possible attribute values each as summarized in Table 1.

Table 1: Dimensions of Knowledge Sharing Systems

	Knowledge Structure	Access
Local	The KS is individual and proprietary.	No information sharing takes place, no access.
Centralized	The KS is maintained and employed at a single location for all.	Global access to a single repository. Static access control.
Decentralized	The KS is fixed, but distributed, thus shared by all instances that employ it.	Access to multiple distributed repositories, statically controlled at each repository.
Evolutionary	The KS is maintained individually and shared between distributed nodes. Local vocabulary changes are communicated to be processed or converted to ensure interchangeability.	Access to multiple distributed repositories. Automatic access control based on trust between agents/users. Automatic mapping between different knowledge structures.

The reader may note that the dichotomy between these dimensions has also been explored in related work, e.g. in “Semantic Web and Peer-to-Peer” (Staab, Stuckenschmidt, 2005), where the Semantic Web accounts for ways to organize knowledge structures and peer-to-peer for ways to organize knowledge access. The editors there, however, have focused on the two layers of centralized and decentralized systems. Semantic Web (SW) research deals with methods to build a World Wide Web containing machine interpretable information. The development of the Resource Description Framework (RDF) that provides a machine interpretable, semi-structured language to record, interlink, and distribute metadata is one outcome of SW research. RDF has been further extended towards ontologies that can be, amongst others, expressed by the Web Ontology Language (OWL) which adds knowledge representation and reasoning to RDF based on Description Logics (DL). The advances in knowledge representation, exchange, and reasoning as building blocks for the SW also leverage IT-support for KM in general. Several of the systems described in the following section are based on and contribute to that work.

4 Evolving KM Systems and Technologies

In this section, we contrast actual KM as reflected by current KM systems with novel KM approaches. We discuss KM systems with respect to the organizational model to provide a classification for these systems and to pinpoint the current status of KM. The latter defines the starting point from which we progress towards the evolutionary KM approach that reflects human networking and dwell on recent research and technologies that foster progress towards this model.

4.1 Knowledge Management Today

4.1.1 Local

Within the last years, KM entered the local desktop in the form of personal information management (PIM) applications that are aiming at better user support for common tasks such as communication, time management, and data management. PIM systems combine and integrate multiple knowledge sources to improve availability, management, and reuse of information. Most PIM systems combine an address manager, an e-mail client, and a scheduler so that e.g. an appointment can be seamlessly e-mailed, or an entry in the address book can be used to denote a participant of a meeting or the recipient of an e-mail. Such functionality is provided by software like the Mozilla™ Suite, Microsoft™ Outlook, and KDE's Kontact.

Improved management of information stored on the local file system has also been addressed lately. Enhanced search tools such as Google's desktop search, Gnome's search tool Beagle, and Copernic support the user in finding and locating information stored on the file system. These systems are often superior to common search features provided by the operating system as they are able to search the content of various proprietary file formats so that a search may include e-mails, bookmarks, contacts, images and several text document types (PDF, Microsoft Word etc.).

Increased popularity of enhanced search tools indicates that the common organizational model of folder hierarchies meets its borders with the increasing amount and heterogeneity of data stored on computers today. Folder hierarchies are simple and resemble real world file storage thus being easy to grasp by novice users. A shortcoming of folder hierarchies is that content can only be classified following a single view disregarding that content usually has multiple contexts. As a simplistic example, consider some text documents that need to be stored in a folder structure: One could distinguish the documents by the type of the document such as *Contract, Invitation, Invoice* etc., or by the person who created the document, or by the time it was created. Either classification provides different context and is differently suitable depending on the task that needs to be established with these documents.
Improved file systems as well as tools that address the shortcomings of conventional file systems are already available or close to release: Microsoft's

new file system WinFS™ [3], the Gnome Storage[4] project, Apple's Spotlight™ [5], and the database file system project (DBFS[6]) are samples of such novel approaches. The conceptual e-mail manager (CEM) (Cole, Stumme, 2000) addresses the same shortcoming in the specific context of e-mailing, enabling multiple categorizations for e-mails and offering different views onto the so classified e-mails.
A model for content classification that is less restrictive than the taxonomic model employed by folder hierarchies, however more complex to maintain, is known as tagging and recently gained popularity in centralized information sharing systems[7] but is also applicable locally.

Further local KM systems are Haystack (Huynh et al., 2003) and Gnowsis (Sauermann, 2005), which are research prototypes aiming at more advanced management features. The Haystack system combines several knowledge sources and applications in a single PIM system and supports the user in relating arbitrary information from any such knowledge source to feature enhanced management and utilization that goes beyond the features of the popular market players mentioned before. Amongst others, Haystack manages bookmarks, e-mails, contacts, appointments, and local files and includes a web browser, e-mail client, scheduler, and contact manager. Any information managed through Haystack is accompanied by meta data that can be extended and modified by the user, e.g. a local file can be related to an e-mail or a contact so that the additional context can be exploited when searching and browsing for information. The Gnowsis system follows a different approach, as it does not provide any application specific user interface. Instead, Gnowsis provides a single data store to integrate arbitrary information and offers an interface that can be used by regular desktop applications to contribute their data. Similar as with Haystack, all information managed with Gnowsis can be related to each other to enable enhanced information management features.

In all systems mentioned so far, knowledge is defined utilizing individual knowledge structures such as individual hierarchies on the file system, in e-mail clients, or by user tags, while no other than local access is offered. Such desktop applications commonly offer high usability and simplify ad-hoc knowledge creation due to high personalization of the knowledge structures and knowledge processes. It can also be noticed that the formerly separated retrieval tasks such as *searching* and *browsing* merge with the rise of novel data management tools as presented.

---

[3] http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnwinfs/html/winfs03112004.asp
[4] http://www.gnome.org/~seth/storage/
[5] http://www.apple.com/macosx/features/spotlight/
[6] http://dbfs.sourceforge.net/
[7] http://flickr.com, http://del.icio.us

4.1.2 Centralized KM Systems

In contrast to systems implementing a local model that omits networked information access, systems that store knowledge in a single centralized repository are the most popular ones at the time of this writing. The centralized model has two significant advantages:

1. Centralization permits knowledge sharing.
2. Centralization simplifies access control.

Most widely employed are systems that centralize both, knowledge structures and information access. Centralized knowledge structures conform to predefined standards that are agreed on by the knowledge community and ensure that knowledge can be easily shared within that community. A multitude of standard groupware systems including document management systems (DMS), and content management systems (CMS) follow this organizational model. Nearly all such systems provide a browser-based user interface that offers entry, manipulation, and retrieval options for the content provided by the system. Groupware solutions usually offer similar features as PIM systems do for the local desktop, while DMSs often resemble a local file system characterized by a folder hierarchy where each folder contains documents. Compared to DMS, CMS offer more flexibility to present information as they do not only provide a centralized storage but also a presentation system.
Centralized systems that utilize tagging for information organization represent a hybrid approach between centralized access and evolutionary vocabularies. Flickr™ is such a system where users can share and organize images using taggings.

Another example of a partly centralized KM application are semantic e-mail processes (McDowell et al., 2004). An example that illustrates semantic e-mail processes is the organization of a meeting: In a common scenario, the initiator of a meeting first needs to invite potential participants to the meeting and afterwards has to take appropriate actions upon incoming replies that may confirm or dismiss the proposed meeting date. Semantic e-mail processes reduce the effort of the meeting initiator who usually needs to act on positive (a guest confirms the invitation) and negative (e.g. a guest rejects the proposed meeting date) replies by automating decision processes such as taking an appropriate action on incoming replies to the meeting proposal. Decisions can be automatically taken based on a prior definition of rules and a semantic specification of the particular task, e.g. a meeting organization. Semantic e-mail processes rely on e-mail servers which represent a centralized IT infrastructure. The knowledge structure utilized for the definition of a semantic e-mail process, however, is not shared as it is only locally available by the creator who defined the process.

4.1.3 Decentralized Samples

Peer-to-Peer (P2P) systems became popular through file sharing applications such as Napster™ [8], or Kazaa™ [9] and represent the next group of KM systems we identify. The most significant difference to the organizational models mentioned before is the decentralization of information access. Information is not available at a single centralized repository, but is kept by multiple so called peers each of which provides a portion of the information available in the network which consists of all participating peers.

The Bibster system (Schnizler et al., 2005) is a P2P system that provides a decentralized repository for bibliography data established by multiple local repositories. While Bibster allows each user to maintain his own bibliography database, bibliographic information maintained by other users is also available through a single user interface. If a user searches for a particular bibliography entry, the local database as well as the network of connected Bibster clients is searched to retrieve the entry. For the organization of entries, Bibster employs an additional taxonomy of research topics that can be used to classify bibliography data by associating bibliography entries to one or multiple entries in that taxonomy. The taxonomy and the file format for bibliographical data represent the knowledge structures used in Bibster. As the whole network agrees on these knowledge structures, the organizational model for the knowledge structures is centralized. Compared to systems that centralize access, Bibster reflects a common local organizational model although being networked and decentralized.

Xarop (Tempich et al., 2004) is an example for a completely decentralized knowledge sharing system. In contrast to Bibster where a fixed taxonomy is used to organize information, Xarop allows to organiz information with individually defined taxonomies. Thus, the knowledge structures describing the shared knowledge are not centralized, however, they follow a globally agreed meta model (the SWAP common model) that defines typical concepts for shared knowledge sources such as the concept of a file, folder et cetera. Based on this shared meta model, Xarop implements means to map individual taxonomies to enable the retrieval of information that is not available locally and thus not organized by the known local taxonomy. A similar example is the Kex P2P system (Bonifacio et al., 2002) that supports document search within the network utilizing multiple individual contexts at each peer. A context is a topic-hierarchy used to classify local content and follows a commonly agreed schema that is exploited to implement context matching and utilize context information to represent perspectives onto the shared knowledge. Both, Xarop and Kex, allow for defining individual knowledge structures within the boundaries of a given meta model thus providing higher personalization than centralized models.

Peer-to-Peer clients in the domain of file-sharing have proven the abilities of decentralized systems. Such systems often ``feel'' local, i.e. providing similar

[8] http://www.napster.com/
[9] http://www.kazaa.com

ease of use as systems following a local organizational model, and combine this degree of personalization with networking.

Table 2: Knowledge Management Systems

		Access			
		Local	Centralized	Decentralized	Evolutionary
Knowledge Structure	Local	MS Outlook, Mozilla Suite, KDE Kontact, Google Desktop Search, Beagle, Copernic, WinFS, Spotlight, Haystack, Gnowsis	Semantic E-mail		
	Centralized		Document Management Systems (e.g. BSCW), Content Management Systems	Bibster, Distributed Databases (e.g. Lotus Notes)	
	Decentralized			Xarop, Kex	
	Evolutionary		Flickr, Delicious, Bibsonomy		*Tagora Project*

4.2 Shortcomings of current KM Systems

All of the systems introduced before address some challenges of KM such as personalized knowledge management, efficiency of knowledge work or knowledge sharing, and distribution of knowledge. The overview reveals that a generally accepted KM model that fulfills the multitude of requirements imposed on KM does not exist, and explains the near continuous appearance of new systems that tackle various issues in KM. Table 2 visualizes the classification of the presented systems given by the overview. Note that the the entry for fully evolutionary approaches, Tagora, points to an integrated project of the European Union under which a system with evolutionary access models and knowledge structures is going to be developed. No systems that follow this model are currently available as more research is needed to enable such systems. In the following, we enumerate shortcomings of available KM systems that result from the organizational model employed by these systems:

Omitting Knowledge Sharing: Local PIM systems often improve time management and communication by integrating data and applications that are related to these tasks. Integration of further knowledge sources such as the documents stored on the local filesystem, and additional communication channels (instant messaging, blogging) is partly addressed by research projects such as Haystack or Gnowsis. The value and potentials of knowledge sharing, however, are completely disregarded by these approaches and denote their main shortcoming.

Single Sourced: Centralized systems lack the integration of multiple knowledge sources even more than localized ones: They are completely disconnected from knowledge sources other than the single centralized one disregarding additional available knowledge and the relations existing between the information stored at the central repository and available at various additional sources. Knowledge of the type ``write once, read never'' as well as process knowledge that could be captured locally can hardly be shared with centralized models even though it may be valuable for others.

Lacking Personalization: Centralized knowledge structures prevent users from contributing knowledge as it requires them to think in different structures than they are used to, or even in structures they reject, increasing their effort of contribution while decreasing motivation, believe in and utilization of the KM system.

Imposing Security Issues: Although a decentralized model supports the combination of multiple knowledge sources while enabling networked knowledge sharing, the decentralized systems we introduced before are usually bound to one specific knowledge source. In comparison to centralized models, access control and knowledge sharing are costly tasks as locally defined knowledge cannot be shared as easily and access control needs to be implemented and coordinated at each peer.

4.3 Towards Evolutionary KM

Standardization of knowledge structures in traditional knowledge sharing environments collides with the individual needs of the user to organize knowledge according to his individual knowledge structures. Knowledge sharing should be seamlessly integrated without requiring users to invest significant additional effort into converting their work to a representation that is different from or even contradicts their mental models. Minimizing user effort can certainly increase acceptance and quality of the KM environment but imposes several challenges dealing with the generation, sharing, and utilization of knowledge. The following paragraphs elaborate on such challenges and demands towards evolutionary KM:

Integration of Multiple Knowledge Sources: To maximize the amount of knowledge that can benefit the community, any available knowledge source should possibly contribute, e.g. file system hierarchies provide classifications of the content stored in folder hierarchies, replies and responses in e-mail threads can include process knowledge and exchange frequency can indicate user-to-user relations, bookmarks can indicate a user's expertise in certain fields, and group calendars contain knowledge about the presence and location of persons. Furthermore, the Web represents a vast knowledge source which can be used as indicated by the Flink project (Mika, 2005) explained in Section 4d. While new knowledge is not necessarily beneficial as it may impose contradictions and overload, the integration process needs to include the assessment of the quality of the integrated knowledge and the manageability of the resulting enhanced knowledge base.

Support for the Individual: Knowledge work should not confine the knowledge worker in how she accomplishes her tasks. To encourage the knowledge worker, increase ease of use, user acceptance and support for knowledge work, everyone should be able to define knowledge as desired by utilizing individual knowledge structures. Most of the systems introduced before lack the ability of providing this level of personalization in combination with knowledge sharing support. For example, DMS often provide a folder hierarchy to organize containing documents. The hierarchy, however, is shared by all users that share documents with the system. Different users prefer different categorizations for their data depending on which tasks they accomplish with that data, or based on their personal preferences and mental models.

Interconnecting Information: Human memory heavily relies on mnemonics that tie various information together. Today's systems often focus on only one knowledge source and prohibit defining and exploiting relations between information existing in other places. Consider a user A that received a text document from user B by email, saved that document on his harddisk, and edited it. At a later time, A might want to further edit the file, or send it back to B but has forgotton where the file is stored. However, A knows who sent the file, but the association between B, the file, and its location is often lost. Any content, whether available in local sources or on the network should possibly be interconnected to support human memory as best as possible to improve reuse, efficiency of search, retrieval, and perception of knowledge.

Dynamic Access Control: Access control is an important requirement in KM to secure knowledge of businesses, and to ensure privacy for the individual. Static access control becomes unsuitable when arbitrary knowledge sources are employed, when participants cannot be centrally managed and should have different access rights for different information. Real world information access is based on trust between people where people implicitly infer trust between two persons when they transfer information originating from or being about one person to another person. The evolutionary KM model requires similar features to automate access control on each peer.

The evolutionary KM model defines a new KM scenario that deals with the dynamicity of vocabularies, users, information, and access. Evolutionary KM systems improve usability allowing ad-hoc knowledge creation, providing high personalization while reducing overhead, participating in a networked knowledge management environment by propagating or providing local knowledge and being capable of searching, importing, and combining knowledge available locally and on the network. The next section explains research approaches and results that tackle the challenges raised by the evolutionary model.

4.4 Research and Technologies

In the following, we discuss various aspects of evolutionary knowledge management and mention several tools as listed in Table 3 that represent samples of state-of-the-art research and technologies dealing with several challenges of this KM model.

Automating the Creation of Individual Ontologies: Information management is commonly alleviated by a categorization of the information that provides a context and can be exploited in tasks such as information retrieval. A simple form of an ontology, a taxonomy, is often employed for this purpose, e.g. file system hierarchies, mail folders, and catalogs. Ontology learning denotes the generation of a knowledge structure (ontology) based on the analysis of some content. Tools like Text-To-Onto (Maedche, Staab, 2004) analyze text documents to automate the extraction of contained concepts and concept relations that can be used to build an ontology that structures the information represented by such documents. Ontological descriptions can then be exploited to improve knowledge retrieval and to combine knowledge. Ontology learning plays an important role for knowledge preparation helping to overcome cold start problems, i.e. to ease the transition from existing, however, hardly described content to richer descriptions that can be exploited in exchange and reuse.

Knowledge Sharing with Individual Ontologies: Centralized systems are based on the standardization of knowledge representation in order to ensure shareability and consistency. In the case of distributed and evolutionary KM, sharing knowledge with others is challenged by the deviation between individual knowledge structures upon which the formalizations of the knowledge of each participant are based. In order to establish interoperability between KM systems under such circumstances, individual knowledge structures such as illustrated in Figure 5 have to be matched. Multiple research works deal with related tasks referred to as ontology matching, semantic coordination, ontology mapping, ontology merging, and ontology alignment. Given that two individual ontologies have been successfully aligned, the knowledge defined with any of the ontologies can be transferred. Quick Ontology Mapping (QOM) (Ehrig, Staab, 2005), CtxMatch (Bouquet et al., 2003), and PROMPT (Noy, Musen, 2000) all are samples of algorithms and applications dealing with this challenge.

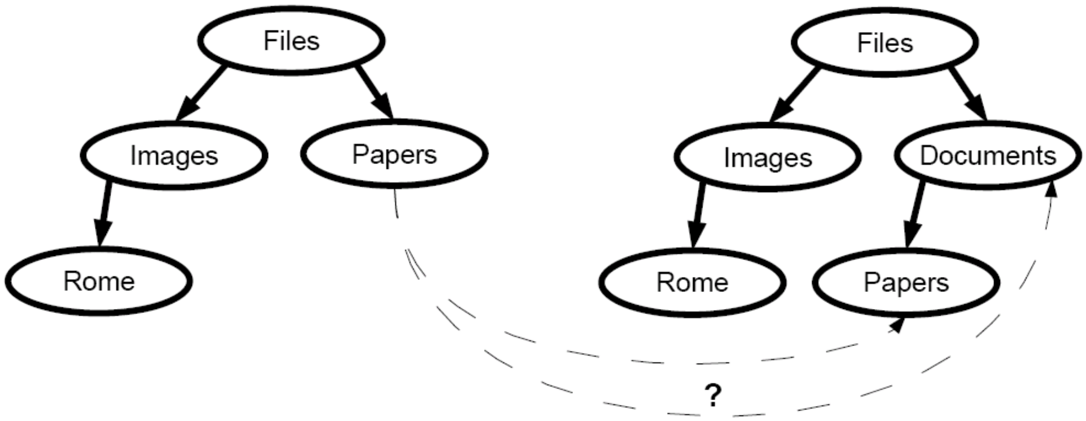

Figure 5: Ontology Matching

Combination of Knowledge: Given that multiple knowledge sources are combined by appropriate means, e.g. ontology alignment tools as mentioned before, the knowledge contained in multiple sources needs to be combined in order to derive new knowledge and to sufficiently present it to the user.
An example of knowledge combination is given by Flink (Mika, 2005), a web portal that aggregates information about researchers from email lists, web pages, publication archives, and user profiles. The conglomerated information, e.g. researcher A cooperated with B to publish a paper P at conference C, is exploited by social network analysis to compute new knowledge. For instance, visualizations of the social network of a researcher are offered, and rankings of the importance of a researcher are given.

Supporting Dynamic Individual Knowledge Structures: The evolutionary KM model respects individual knowledge structures that can be constantly modified by the user to suit his needs. Knowledge sharing under such circumstances is partly addressed by the semantics aware messenger (SAM) (Franz, Staab, 2005) that allows two communication partners to classify messages utilizing user-defined tags while chatting. The taggings benefit the retrieval of messages, e.g. they allow to retrieve messages from conversations with different persons that all deal with the same topic. SAM utilizes the taggings of multiple users so that a user can browse messages based on the views (represented by the taggings) of others.

Trust to Control Access: Similarly to the dynamics of knowledge structures, dynamics of access control are imposed by distributed and evolutionary KM models: Such a networked KM infrastructure may have an unrestricted number of participants as no centralized membership and access control exists. Consequently, the network consisting of multiple local KM systems, needs to define who has access to which knowledge, and how much credibility is given to knowledge from a certain participant. Golbeck et al. developed an ontology for defining trust in general but also more fine grained options to define trust bound to certain topics as one often trusts a particular person concerning one issue but may distrust the same person on other issues. Given such trust information, a ``trust network'' can be established that defines how much one should trust another. Further research towards automating such trust ratings to cope with dynamic access control is required to establish evolutionary KM in the dimension of information access.

Table 3: Methods and Tools

Knowledge Structure \ Access	Local	Centralized	Decentralized	Evolutionary
Local	Categorization not applied because of multitude of tools			Trust-based spam filters
Centralized				Trust-based recommendations (Ziegler, Golbeck) Trust Networks on the Semantic Web (Golbeck, Parsia, Hendler)
Decentralized	Prompt (Noy, Musen)	Glue (Doan et al) CtxMatch (Bouquet et al) Qom (Ehrig, Staab)		
Evolutionary	Ontology learning/ TextToOnto (Mädche, Staab)	Flink (Mika)	Sam (Franz, Staab)	Full vision of the dynamic Semantic Web

Automating access control and alignment of knowledge defined by individual knowledge structures are the main challenges in dealing with the dynamics imposed by evolutionary KM.

While we here aimed to formulate our vision from a user's point of view, the Semantic Web crowd as a technical community recently proposed the Networked/Social Semantic Desktop which exhibits several similarities to a individual, but networked KM system as rendered here. The Networked Semantic Desktop (Decker, Frank, 2005) is a computer desktop that provides many of the features required for local, centralized, decentralized and evolutionary knowledge organisation combining social networking, P2P infrastructure, and local knowledge management to establish a novel working environment that naturally supports collaboration, knowledge exchange, and access control (by computing trust in people and credibility of information) in a networked environment.

5 Conclusion

We developed a model for distinguishing different IT based KM approaches and applied this model to state-of-the-art systems. Based on this model, we analyzed challenges, benefits, and shortcomings of common KM systems and envisioned novel evolutionary KM approaches. Different research contributing to such novel KM environments has been skeched subsequently.

Networking and seamless local organization are two sides of the same coin. A knowledge habitat must provide means for evolutionary growth of smoothly organizing and accessing knowledge.

The next step now we believe must come from the organizational side. This wealth of new technologies will not easily be absorbed by current management strategies that focus on a lot of control over knowledge flows while decentralized and evolutionary knowledge management systems will radically shift knowledge organization towards the edges. Currently, companies often refuse peer-to-peer sharing, because they fear that it gives too much power to the individual and too little control to the company, but we conjecture that the ones that first learn to harvest the benefits of such edge computing systems will outperform their competitors' knowledge creation values.

Acknowledgements. This work has been partially supported by the integrated projects X-Media (EU IST FP6-26978) and Tagora (EU IST FP6-034721).

6 References

[1] M. Bonifacio, P. Bouquet, G. Mameli, M. Nori, "KEx: a Peer-to-Peer Solution for Distributed Knowledge Management", in Proceedings of the Fourth International Conference on Practical Aspects of Knowledge Management (PAKM02), Vienna, Austria, 2-3 December 2002. Published in Lecture Notes in AI, Vo. 2569, pp.490-500, Springer Verlag, Heidelberg 2002.

[2] M. Bonifacio, P. Camussone, C. Zini, "Managing the KM Trade-Off: Knowledge: Centralization versus Distribution". Journal of Universal Computer Science (J.UCS), Vol.10, No.3, 2004, Special Issue: (Virtual) Communities of Practice within Modern Organizations.

[3] P. Busetta, P. Bouquet, G. Adami, M. Bonifacio, F. Palmieri, "K-Trek: A Peer-to-Peer Approach To Distribute Knowledge In Large Environments", Second International Workshop on Agents and Peer-to-Peer Computing at AAMAS03, Melbourne, Australia, July 2003, Springer. Published in the Lecture Notes in Computer Science series, no. 2872.

[4] V. Dignum, F. Dignum: The Knowledge Market: Agent-Mediated Knowledge Sharing. CEEMAS 2003: 168-179.

[5] A. Doan, J. Madhavan, P. Domingos, and A. Y. Halevy. Ontology Matching: A Machine Learning Approach. In Handbook On Ontologies , pages 385 – 404. 2004.

[6] P. Bouquet, L. Serafini, and S. Zanobini. Semantic Coordination: A New Approach and an Application. In International Semantic Web Conference , pages 130 – 145, 2003.

[7] M. Ehrig, S. Staab, QOM, Semantic Web and Peer-to-Peer, 2005

[8] I. Foster, C. Kesselmann. The Grid: Blueprint for a New Computing Infra-structure, Morgan Kaufmann, 2004.

[9] L. van Elst, V. Dignum, A. Abecker: Towards Agent-Mediated Knowledge Management. AMKM 2003: 1-30.

[10] T. Franz, S. Staab, SAM: Semantics Aware Instant Messaging for the Networked Semantic Desktop. In Proceedings of the Workshop on the Semantic Desktop at the ISWC, 2005.

[11] I. Foster, A. Tamnitchi: On Death, Taxes, and the Convergence Peer-to-Peer and Grid Computing, http://iptps03.cs.berkeley.edu/final-papers/death_taxes.pdf

[12] P. Haase, J. Broekstra, M. Ehrig, M. Menken, P. Mika, M. Plechawski, P. Pyszlak, B. Schnizler, R. Siebes, S. Staab, C. Tempich. In: Proceedings of the 3rd International Semantic Web Conference (ISWC2004), November, 2004, Hiroshima, Japan. LNCS, Springer, 2004.

[13] A. Maedche, S. Staab. Ontology Learning. In Handbook On Ontologies , pages 173 – 190. Springer 2004.

[14] N. F. Noy and M. A. Musen. PROMPT: Algorithm and Tool For Automated Ontology Merging and Alignment. In Aaai/iaai , pages 450 – 455, 2000.

[15] S. Pinto, S. Staab, C. Tempich. DILIGENT: Towards a fine-grained methodology for Distributed Loosely-controllled and evolvInG Engineering of oNTologies. In Proc of ECAI-2004, Valencia, August 2004.

[16] B. Schnizler et al. Bibster, Semantic Web and Peer-to-Peer, Springer, 2005

[17] S. Staab, L. Ding, P. Domingos, T. Finin, J. Golbeck, A. Joshi, P. Mika, A. Nowak, R. Vallacher. Social Networks Applied, IEEE Intelligent Systems, Trends & Controversies, 20(2), Mar/Apr 2005.

[18] S. Staab, H. Stuckenschmidt (eds.), *Semantic Web and Peer-to-Peer*, Springer, 2005.

[19] C. Tempich, M. Ehrig, C. Fluit, P. Haase, E. L. Marti, M. Plechawski, S. Staab. XAROP - A Midterm Report in Introducing Decentralized Semantics-based Knowledge Application. In PAKM-2004.

[20] E. Tsui: Technologies for Personal and Peer-to-Peer Knowledge Management, CSC Report, http://www.csc.com/aboutus/lef/mds67_off/uploads/P2P_KM.pdf

[21] C.-N. Ziegler, J. Golbeck. Investigating Correlations of Trust and Interest Similarity - Do Birds of a Feather Really Flock Together? *submitted to the Journal of Artificial Intelligence Research*, 2005.

[22] A. Mädche, ~, N. Stojanovic, R. Studer, Y. Sure. SEmantic portAL - The SEAL approach. In D. Fensel, J. Hendler, H. Lieberman, W. Wahlster (eds.) *Creating the Semantic Web*. MIT Press, Cambridge, MA, 2003.

[23]BSCW; http://bscw.fit.fraunhofer.de/

[24]D. Huynh, D. R. Karger, D. Quan, and V. Sinha. Haystack: a Platform For Creating, Organizing and Visualizing Semistructured Information. In Intelligent User Interfaces, page 323, 2003.

[25]L. Sauermann. The Gnowsis Semantic Desktop For Information Integration. In WM 2005: Professional Knowledge Management , pages 39–42, 2005.

[26]R. Cole and G. Stumme. CEM - A Conceptual E-mail Manager. In Iccs , pages 438 – 452, 2000.

[27]L. McDowell, O. Etzioni, and A. Y. Halevy. Specifying Semantic E-mail Processes. In WWW Workshop On Application Design, Development and Implementation Issues in the Semantic Web , 2004.

[28]S. Decker, M. Frank. Social semantic desktop. http://www.deri.at/publications/techpapers/documents/DERI-TR-2004-05-02.pdf (visited september 2005)

[29]P. Mika. Flink: Semantic Web Technology for the Extraction and Analysis of Social Networks. Journal of Web Semantics, Vol. 3, 2005