Peer-to-peer knowledge management

Anjana Susarla, De Liu and Andrew B. Whinston

In *Handbook on Knowledge Management*. C. W. Holsapple (eds). Berlin, Springer. Volume 2: 129-140. 2003.

1. Introduction: The potential of P2P in knowledge management initiatives

The literature in strategic management has long recognized the role of organizational knowledge as a source of competitive advantage. The knowledge within an organization confers innovational capabilities and can help the organization deal with change in the environment. However, organizational knowledge is created through a complex interaction of routines, organizational culture and individual endeavors and can be difficult to locate or transmit (ref). It is a formidable task for the organization to isolate the knowledge from a particular person or a process and transmit it to others who need the knowledge. Promoting efficient exchange of knowledge and enabling collaboration among individuals in an organization to foster a spirit of sharing knowledge is the key challenge for knowledge management in organizations. Peer-to-peer networking offers an innovative way to manage organizational knowledge since peer-to-peer networks provides a flexible way of connecting people while offering a robust and communication rich information protocol.

Several organizations have attempted to build centralized knowledge management systems, which given the effort necessary to codify knowledge and create mechanisms for knowledge transfer, can be a Herculean task. With such conventional KMS efforts, the knowledge discovery process can only elicit information about encoded knowledge as opposed to being a tool aiding the discovery of tacit knowledge. A peer-to-peer knowledge exchange environment, on the other hand, allows for dynamic knowledge transfer and can also enable interaction among knowledge owners, allowing knowledge owners to be proactive in disseminating their new knowledge to others in the organization. A peer-to-peer knowledge management system can mediate a decentralized collaboration of various individuals in the organization, complementing the search technology with a 'human' intelligence component. Further, the dynamic nature of knowledge exchange facilitates the dissemination of up-to-date and as yet unrecorded knowledge through the organization. Peer to peer knowledge management systems enable richer communication between participants and can contribute to the propagation of hidden and implicit knowledge, which needs a great amount of organizational effort to be transferred through a centralized system of knowledge management. P2P technology can ensure search is up to date and fosters an organization's ability to tap vital organizational intelligence. P2P technology also does away with the limitations posed by storing data at a central server, creating a strategic point of failure. Current client server based knowledge management initiatives typically provide little privacy for their users, and P2P networks can address the privacy concerns of users. (See Table 1)

The term Peer-to-peer computing gained currency after the widely publicized Napster case. Napster, the best-known example of a P2P network, was set up as a network of PCs with a central server acting as a clearinghouse. Users could download napster client software and install it to make selected folders on their PCs available for sharing, as well as searching for and downloading files from other nodes on the network. The central server acted as a lookup service that would direct a user to a particular node where the requested file was available. Thus, the lookup function was a client-server interaction, followed by the file transfer that was peer-to-peer. Though attacked initially as a medium for piracy, industry has now woken up to exploring possible positive uses of P2P for information sharing, and electronic commerce in general. Most of the current applications of client server technology use networking as an ancillary feature to aid computing at, and sharing from, centralized nodes. Peer-to-Peer Computing may be thought of as isolating the networking aspect and focusing on it as the mainstay of the business model. The wide acceptance of personal computers in personal and corporate computing, coupled with the ubiquity and rapid acceptance of Internet technology implies that there is a vast untapped demand for applications based on peer-to-peer technology (Parameswaran et. al. 2001). One of the most promising applications of the peer-to-peer technology is in the area of knowledge management systems.

2. Role of P2P in creating organizational knowledge

The process of creation of organizational knowledge is complex and no technology can completely address the nuances of creating organizational knowledge. KMS technologies, however, can help in effective codification of knowledge, provide the right incentives for people to create and share knowledge (Ba et. al. 2001), and offer a superior infrastructure for knowledge transfer and retrieval.

Right now, most knowledge management initiatives try to create a searchable database of stored knowledge by exhaustively collecting knowledge from individuals within the organization. The knowledge that is collected should also include knowledge embedded in organizational processes that is harder to encode. Once the knowledge is created, the organization needs to implement a mechanism for dissemination of knowledge. This could include efficient algorithms to retrieve data from the knowledge management system. However the technological initiatives do not really solve the fundamental problem of knowledge management systems. The success of a KMS system depends on how well knowledge that is specific to individuals can be routed to other individuals and can be disseminated into the collective organizational memory. In short, the success of a KMS system depends on how the technology can enable the creation of a community of knowledge practitioners.

Peer-to-peer technology is a decentralized mechanism of file transfer and relies on group norms for its survival. Instead of creating a vast repository of knowledge, P2P offers a way of managing distributed repositories of knowledge, where the repositories can be individuals in the organization or processes in the organization. By allowing for transfer of heterogeneous document formats and with its highly scalable architecture, a P2P network can create a dynamic method of transfer of knowledge as opposed to a static model of knowledge transfer. Search processes in a P2P setting can be more flexible than in a server based KMS, with different systems interacting with each other regardless of access mode. With a minimum interoperability between the Corba / DCOM/ Java/ XML standards, there could be a wealth of information that is missing in the traditional KMS setting. Knowledge management theorists (Nonaka and Takeuchi, 1995) posit that knowledge is created through a continuous and dynamic interpersonal interaction between tacit and explicit knowledge; P2P embodies the spirit of dynamic interaction and is therefore an ideal knowledge management vehicle. Successful P2P initiatives such as Freenet have demonstrated that the individuals who provide content on a P2P network can monitor the access and usage of the resources they contribute. The peer-to-peer technology can therefore be structured to provide selective incentives to some members for enhancing organizational knowledge and further reinforce the reward system by excluding access to the non-cooperators. Thus P2P transfer of knowledge can solve the incentive problems plaguing knowledge transfer systems by ensuring that the right knowledge is created and distributed. Additionally P2P can address the issues of recombination of knowledge that conventional KMS cannot deal with. For example if the users in the manufacturing department have access to product designs from marketing, they would be able to combine the new product ideas with their own knowledge about design. With a client server based KMS, there is a problem of who does the updating on the server, without protecting the privacy of the individuals who are recombining the knowledge. The users in manufacturing could be apprehensive of too many individuals viewing their plans, and would like to control access to their knowledge. P2P systems such as Freenet can solve this problem by providing for inserted files to be placed selectively on nodes already possessing files with similar keys.

----

fig 2. Inserted here

----

3. The free rider issue in P2P

In economics literature knowledge is treated as a public good []. According to this view, in an organization, once the knowledge is created it belongs to all the individuals in an organization. This raises the classic free rider problem, i.e. all the individuals in the organization can benefit from the knowledge in the organization but no one has the incentive to contribute to the creation of knowledge. In a knowledge intensive organization such as a consulting company or a software development organization, knowledge transfer within an organization has a public good aspect to it. This problem is compounded by the inter-related and interdependent nature of information flow. Free riding in a P2P network, as in the case of other public goods scenarios, could lead to a tragedy of the commons. The preponderance of free riders could lead to less and less individuals sharing knowledge, and eventually the knowledge management system itself breaking down due to lack of adequate knowledge being shared.

A significant technological constraint in a P2P network is that each participant node that does share content can only handle a limited number of connections. This aspect is all the more relevant since these nodes are typically not dedicated servers with built-in redundancy in storage and connections which would be the case in a client-server scenario. So, as the ratio of free riders to contributing nodes goes up, the pressure on the contributing nodes increases, and this could lead to a breakdown. Thus, excessive free riding could lead to collapse of the system not merely by way of the economic reasons of erosion of the value gained by participating in the system, but by failure of the technological infrastructure as well. This problem could be compounded in an organization that encourages individual performance, which actually discourages people from exerting effort that benefits others rather than improve their own performance. We propose that the way to ensure a smoothly running P2P network is to implement cost sharing mechanisms. For example if a user does not contribute to content provision, he can at least lend cycles of CPU units. An internal currency system can effectively privatize what is a public good. Also there needs to be a consensus on what is the value of the content to be provided, say by means of an auction among the P2P users. Mojonation (www.mojonation.com) is one of the P2P engines that is implementing mechanisms by which free loaders cannot consume more than they contribute to the system; they must purchase more resources from others. This in turn should serve as a means of excluding some people who will not voluntarily contribute to the P2P network.

4. Aligning the incentives of individuals in creating and disseminating knowledge

The rise of Peer-to-peer technology parallels the growing trend in organizations towards a leaner and more autonomous structure that is flexible in dealing with change. Peer-to-peer technology highlights the role of contribution and participation by individuals in an organization, which is the thrust of enterprise knowledge management systems. Organizational support for the peer to peer technology addresses the issues of providing the right incentives and mitigating free riding of participants, which poses a challenge for an essentially unregulated technology such as a peer to peer system. Coping with free riding and ensuring efficient knowledge dissemination is the idea that lies at the heart of our proposed incentive system. Any node in the P2P network may 'steal' quality links or information from other nodes by replicating it and offer such information as part of bundled products. There may be no way to differentiate the original sources of knowledge, and even if there were, the knowledge buyer may not care about the source as long as he receives quality information. Some form of internal regulation within the P2P network may be required to monitor such activity.

The regulation of knowledge creation and propagation is easier in the case of a centralized network. The decentralized nature, ease and speed of growth and open access policies of the P2P networks may render this quite an insurmountable difficulty. The challenge for the software designer is to think of ways to regulate the formation of organizational P2P networks for knowledge transfer while ensuring that creators of knowledge in the P2P network get compensated for their efforts (Groves, 1973).

Emerging uses of P2P technologies have tried to implement a basic set of communication protocols that let computer nodes find each other, organize into peer groups, exchange messages within groups, and manage the nodes forming the group. One such concept is implemented in Project Jxta of Sun Microsystems where "peer monitoring," is used to manage the nodes in a group. This feature can detect if one node is using too much bandwidth or capacity and then pass an alert to a management console so an administrator can reconfigure the faulty node. However, if P2P has to be used for knowledge management initiatives, these technological rules should be combined with other rules restricting participation through access rights for knowledge management. The P2P network needs to have restrictions in place that ensure the right individuals get to know the appropriate knowledge. The network should supply participation incentives so that there is no one who might free ride on the efforts of others in the network, and creators of knowledge get compensated for their effort. Strong alignment within the organization, i.e. ensuring that individuals in the organization work for a common set of goals is a prerequisite for an organization seeking efficient knowledge management (Nonaka and Takeuchi, 1995). Cohesive social groups where norms can be imposed on public behavior ensure that knowledge is created and distributed in an equitable manner.

To be effective, therefore, the P2P KMS should enable a spirit of collaboration and teamwork within the participants in the organization. The decentralized nature of the P2P network, which causes the free rider problem, can also make P2P an ideal medium for individuals to collaborate without a central enforcing authority. The agents within the different teams have to cooperate to achieve the desired knowledge management goals of the group. Research has shown (Holmstrom, 82) that these types of self sustaining cooperative mechanisms can be far more effective than one where each individual needs to be monitored separately by a central authority. Since the success of KMS initiatives depends on team works, the organization should promote the formation of decentralized teams that can motivate each other and work for cooperation enhancing incentive schemes such as those that reward group performance (Che and Yoo, 2001). The economic rationale of P2P for knowledge management initiatives lies in the ability of P2P as a communications infrastructure fostering collaboration between agents in an organization.

5. Technological infrastructure: Content sensitive addressing to improve the resource identification

P2P technology can support distributed searches through the network. An individual looking for some particular knowledge can query the P2P network for the information. Further, there could be restricted access to certain types of documents to a select set of groups or individuals in the organization. Although searchers can still search for abstract info, they may have to ask for grant to access the complete document. Since access depends on the user who owns the knowledge, a document owner would be able to monitor the demand for her knowledge. However this system needs a well-defined system of resource ownership within the network. When the ownership of a knowledge management resource itself can be blurred, it is difficult to implement a system of knowledge transfer that needs explicit property rights.

A peer-to-peer KMS can also allow knowledge seekers to initialize the knowledge request to the network. The network can have intelligent agents to direct the requests to the appropriate knowledge experts who can send their responses to those seeking the knowledge. A peer to peer KMS can facilitate the creation of a virtual space where knowledge providers and users can interact intensively, thus translating into reality the vision of dynamic knowledge interaction envisaged by KM theorists.

In an organization where different individuals could use heterogeneous formats to encode knowledge which then gets distributed over P2P networks, the challenge for the KMS is to set up a knowledge exchange environment which facilitates knowledge communication between multiple devices and document formats. The user does not need to know the exact node on which a particular item of knowledge resides, and not worry about where recent updates to knowledge have been made. Additionally the P2P KMS technological architecture should address security and reliability issues as well as provide a protocol to verify access rights of different users. One such initiative is TRIAD architecture developed by the distributed computing group at Stanford University. The group's *Name-Based Routing Protocol* (NBRP) is a dynamic mechanism for updating the routing information in relay nodes. NBRP distributes name suffix reachability among (and within) address realms. Routing using names helps to ensure the availability of those names, even in the absence of multiple address realms.

Within the organization there could be different teams exchanging knowledge, and the P2P system can ensure that only members within a team can exchange ‘private’ knowledge whereas ‘public’ knowledge can be transferred across teams. Similarly when there are overlapping P2P groups, the P2P infrastructure can define authorized principals for each group and place the group name on an appropriate access control list, as envisaged by the distributed security infrastructure developed by MIT’s distributed computing group. Thus within each group there is authorization for the users as to who can access certain data or perform certain operations. The principal of each group can establish his own local name space and use identity certificates to establish the credentials of other members. Therefore whenever a user queries the group for a pice of knowledge, the security protocol first verifies if the user node from where the query originates has the right to access the information. The access protocol should also reconfirm the identities of users from time to time.

Access control lists can be set up for interlinked groups where each person is linked to a ‘principal’. The certification list denotes who can get access to certain data or perform certain operations. For example

Marketing –Advertising Division
Research and Development Division –Research scientist
Marketing –New Product Development Group
Virtual team – Customer relationship management -members
Enterprise security group
Product development group-Emerging Markets –Brazil

Group membership assertions can play the roles of credentials, licenses, or ordinary paper certificates. The distributed architecture of the P2P system should be able to provide a user access to multiple groups. In some of these groups the member can have private key access while he could have public key access, i.e. access to general and commonly available information. An organization such as ACM can create a group that can be referred to as ACM's members. A statement signed by ACM might say that a given principal (i.e. a particular public key, not a name for that key) is a member of the group members. The ``meaning'' of this statement is entirely up to ACM, since it is ACM's prerogative to define its groups any way it pleases. The definition of the group members, and the assertion that some principal is a member of it, are things that ACM may export.

The P2P KMS system can address the issue of access rights of users by providing participation rules that are enforced when a member of the organization connects to a node in the P2P network. An individual in the network is assigned a high-level name, similar to the Domain Name system (DNS). The names to address mappings are replicated across nodes. An individual in the organization initiates a knowledge request, and the P2P software translates the request to the specific node that has the product. Addresses are thus grouped based on the contents of respective nodes in the organization. Thus, addressing is taken a level further in the semantic hierarchy, where users specify not physical locations but a content identifier. The identifier corresponds to a peer-to-peer collection of nodes that store this type of content, and the process of identifying the individual node that stores the requested item, and initiating the download, is transparent to the user. Thus, classifying content into specialized groupings distributed over P2P networks can facilitate a more refined information repository and 'truer' uniform resource identifiers.

fig 3.

6. Some potentially deleterious problems of P2P technology

The quality of knowledge can be a serious problem to the survival of P2P management systems. Knowledge is autonomously maintained in P2P, therefore there is no dedicated expert who can look at the centralized knowledge database and weed out unimportant content. The decentralized nature of the network also makes it hard to implement consistent methods of organization of information. P2P networks are vulnerable to overload of data. Besides the problems in ensuring secure delivery of data and the free riding issue pose threats to the viability of P2P as a knowledge management mechanism. Indeed a serious cause for concern in P2P networks is that arbitrary organization and distribution of content can seriously undermine the quality of information search (Parameswaran, et. al. 2001). The degree of efficiency of a P2P search will necessarily depend on how well organized the network is, and its policies on classifying content and building directories, and on the search software itself. The effectiveness and timeliness offered by P2P searches may be undermined if the organization of content is chaotic, and spurious content and duplicates proliferate in the network. Three factors in particular often cause knowledge markets to operate inefficiently in organizations:

Incompleteness of information in a decentralized mechanism poses another problem in the acceptance of P2P technology for knowledge management systems. In the absence of a centralized repository of knowledge, the knowledge user does not have a roadmap to search for knowledge in the appropriate manner. Content based addressing is one way of mitigating the problem; adopting consistent knowledge management practices is another. Organizations trying to implement an internal knowledge market face the problem that a decentralized system may not adopt consistent criteria for pricing of knowledge.

The localization of knowledge poses another problem to using peer-to-peer technology. Individuals in an organization may tend to approach those whom they know for knowledge rather than invest effort into an uncertain process of trying to discover who in the company may have the right knowledge. Theorists of bounded rationality argue that individuals have a tendency towards 'satisfying', i.e. settle for knowledge or information that is "good enough" for their purpose rather than go to considerable efforts to acquire the best possible knowledge. High search cost for optimal knowledge is probably the biggest constraint towards implementing an efficient knowledge market within a firm. The totally decentralized nature of a P2P setting may of course make it vulnerable to breaches of security. The community aspect of P2P computing may call for social mechanisms to enforce regulation.

In the knowledge market…

- The right seller is often *hard to locate*,
- And can be *hard to reach* even if we know her location.
- It's also difficult if not impossible to *judge the quality of knowledge* beforehand.
- The value of knowledge and the likelihood of *eventual payment* are both uncertain.

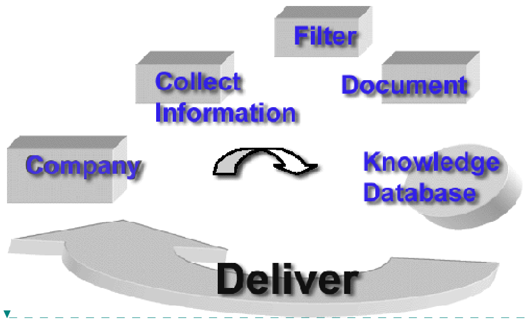

fig 1. centralized knowledge management processes

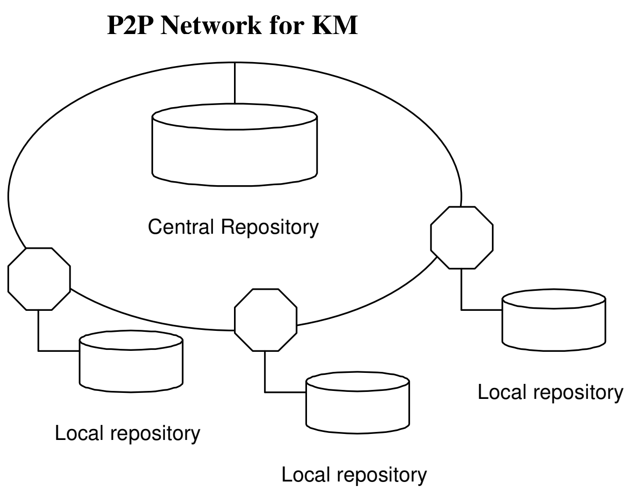

fig.2 Conceptual framework of peer-to-peer network for KM

Deleted: ***9. Role of P2P in creating organizational knowledge¶***
The process of creation of organizational knowledge is complex and no technology can completely address the nuances of creating organizational knowledge. KMS technologies, however, can help in effective codification of knowledge, provide the right incentives for people to create and share knowledge, and offer a superior infrastructure for knowledge transfer and retrieval. ¶
¶
Peer-to-peer KMS will provides a comprehensive knowledge transfer mechanism, which defines the classification and codification rules of knowledge. Individuals are equipped th tools to create knowledge embedded documents, memos, emails, illustrations etc, which are then wrapped up (XML) with all kinds of information (key words, authors, abstract, classification etc). Those knowledge medias are then registered in the system and tracked automatically. System will execute security control and access controls on those knowledge documents.¶
¶

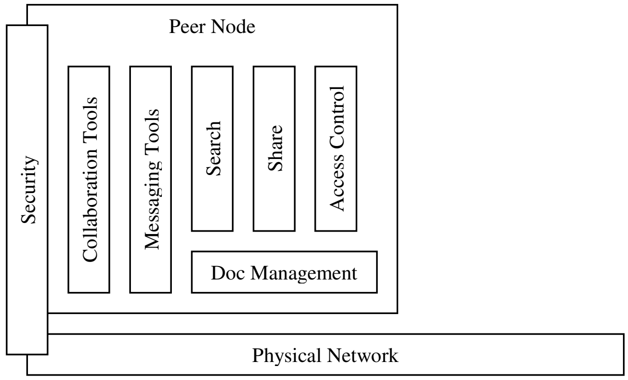

fig 3. Modules of Peer-to-peer Knowledge Management

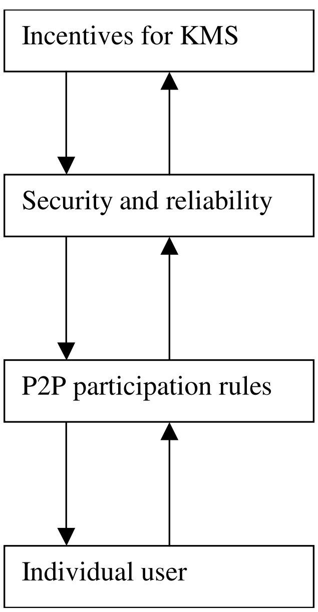

Fig 4. Organizational and technological layers facilitating P2P KMS