KNOWLEDGE MANAGEMENT AND PEER-TO-PEER COLLABORATION TECHNOLOGY: KEY ISSUES AND RESEARCH CHALLENGES

Saurabh Gupta, The University of Georgia
Robert P. Bostrom, The University of Georgia

ABSTRACT

The importance of knowledge management (KM) has been developed in both academic and practitioner literatures. Organizations are currently searching for best practices and technologies for implementing KM. This paper looks at a new technology, Peer-to-Peer (P2P) computing, and explores how it fits in the knowledge management paradigm. The paper explores the kinds of P2P applications that exist, identifying those most relevant to KM. The paper then focuses on collaboration applications, comparing P2P tools with client server tools by mapping tool usefulness in terms of KM process and at various levels of collaboration. The paper concludes with the identification of key research issues concerning the use of P2P applications for KM.

INTRODUCTION

Organizations are implementing knowledge management practices and technologies with the intention of increasing their effectiveness, efficiency and competitiveness. The importance of knowledge, and how it can be used to create competitive advantage, has been examined extensively in both MIS and strategic management literature (Beaman & Zeithaml, 1989; Liebeskind, 1996). This literature clearly shows the importance of contextualizing and applying information and knowledge. Based on Von Krogh (1998), we can look at the definition of knowledge management and summarize its importance as: Knowledge management refers to identifying and leveraging the collective knowledge in an organization to help the organization compete.

The process of knowledge management (KM) is focused around human interaction via collaboration (Pearlson, 2001). Thus, collaboration is an integral part of KM. Currently, the client server (C/S) architecture dominates the thinking in the knowledge management world at all levels of collaboration. We feel this thinking may be limiting, thus creating missed opportunities. Peer-to-Peer (P2P) architecture offers some interesting benefits and challenges to KM. For example, getting people to share knowledge through a server has proved to be a challenging task (Kini, 2002). P2P environments remove this hurdle since information can be kept on person’s desktop, and still be shared.

Research is needed to explore KM in P2P environments. One of the reasons causes for the lack of research is this area is the lack of understanding of P2P environments within Management Information Systems. The purpose of this paper is to define P2P, show how it can be applied to KM and identify key research areas.

The paper starts with review of P2P networks, identifying those most relevant to KM. Next, the paper compares P2P and C/S architectures and tools in terms of their support for KM process and for different levels of collaboration. The paper concludes with the identification of key research issues concerning the use of P2P applications for KM.

PEER-TO-PEER NETWORKS

At the most basic level, P2P is simply one computer node talking directly to another without any intermediary (Kini, 2002). Thus, P2P is a type of network in which each workstation has equivalent responsibilities. This
differs from C/S architecture in which one computer (server) is dedicated to serving the others (clients). Communication in this scenario exists only between nodes and server, and no direct communication exists between nodes. A more formal definition is provided by Barkai (2001 pg 13) “Peer-to-peer computing is a network-based computing model for applications where computers share resources via direct exchanges between the participating computers”. The above discussion highlights two elements that are fundamental to P2P computing - the sharing of resources and direct communication between users. These two together are common thread in describing P2P applications (Barkai, 2001).

A typology for P2P environments has been developed by Fattah (2002). He divides uses of P2P architecture into four applications: User collaboration, application interaction, resource utilization and supercomputing. The first two are grouped under active applications and the later two under passive applications. Active applications are ones in which users or systems do things with P2P to accomplish a task. For example, instant messaging applications, such as ICQ or MSN Messenger, and file-sharing systems, such as Napster, would be active applications.

Passive applications are one in which idle resources are roped in uses es other than their primary function. Resource utilization applications allow combining of resources to produce large database out of the documents stored on computers scattered around the world. McAfee Corporation has produced software that helps manage the distribution of software among users using this idea. Supercomputing applications harness the computing power of PCs on the network or on the internet to aggregate their power and produce virtual supercomputers (Fattah, 2002). SETI@home is a prime example of such an application.

In a knowledge management environment, where emphasis is on using collaboration to generate results, active technologies are more applicable to the process of knowledge management. Current P2P collaboration applications are combining the instant messaging capabilities with the file-sharing capabilities. Companies such as Groove Networks, Endeavors Technology, and Ikimbo sell systems of this type. Other P2P software interaction systems let users of different applications tie those applications together to pass along relevant information, while maintaining ownership of the documents. Applications from companies like Oculus Technologies and Nextpage are using these ideas for databases and design software to accelerate development efforts and improve outcomes (Fattah, 2002).

PEER-TO-PEER AND KNOWLEDGE MANAGEMENT

Knowledge Management Process

Knowledge management represents a set of activities for management of knowledge. At a minimum, one considers four basic processes of creating, storing/ retrieving, transferring/ distribution, and applying knowledge, see Figure 1 (Alavi & Leidner, 2001).

Knowledge Creation: Organizational Knowledge creation involves developing new content or replacing existing content within organization’s tacit and explicit knowledge (Pentland, 1995).

Knowledge Storing & Retrieving: Empirical studies have shown that while organizations create knowledge and learn, they lose track of the acquired knowledge (Argote, Bechman, & Epple, 1990; Darr, Argote, & Epple, 1998). Thus, storage, organization, and retrieval of organizational knowledge, also referred to as organizational memory (Stein & Zwass, 1995; Walsh & Ungson, 1991), constitute an important aspect of effective organizational knowledge management.
Knowledge Transfer/Distribution

An important process in the knowledge management process is that of knowledge transfer. Transfer occurs at various levels: transfer of knowledge between individuals, from individuals to explicit sources, from individuals to groups, between groups, across groups, and from group to organization (Alvi & Leidner, 2001).

Knowledge Application

An important aspect of the knowledge-based theory of the firm is that the source of competitive advantage resides in the application of the knowledge rather than in the knowledge itself. (Alavi & Leidner, 2001).

Adapted from Alavi and Leidner (2001), Figure 1 shows the process of knowledge management and what IT enables within each activity. The entire process of knowledge management is self-feeding where knowledge application leads to further knowledge creation and storage, thus a repeat of the cycle.

**KNOWLEDGE MANAGEMENT USING P2P COLLABORATION TOOLS**

We use the model developed by Schrage (1997) to help explain the benefits and limitations of P2P compared to C/S and discuss how this model is implemented in one of the state-of-the-art P2P tools, Groove. An integral part of the knowledge management process is collaboration. According to Schrage’s model, see Figure 2, collaboration builds on conversation or pure communication. Conversations are serial and ephemeral in nature. In most conversations, people take turns in exchanging information, not sharing it. In most conversations, the absence of memory means a useful phrase or expression can be distorted or lost. Thus, most conversations are like a transaction model.

The collaboration model (Figure 2) looks at shared spaces as a new dimension to conversation, a dimension embracing symbolic representation, manipulation and memory. Participants can communicate with one another directly through the medium of shared space. Changing the conversation can lead to change in shared space and visa versa. Symbols, ideas, processes, sketches, music numbers, and words can be put in the shared space to be expanded, organized, altered, merged, clarified and otherwise manipulated to build new meanings. The model emphasizes the need for collaboration beyond pure communication for increased effectiveness and efficiency by providing two essential elements—Memory and the ability to access memory while having a conversation.

Groove, an application offered by Groove networks demonstrates the implementation of this model (www.Groove.net). At the heart of the application is shared space. A Groove user creates a shared space and invites other people into it. From the moment the invitation is accepted, Groove keeps install copies of the space on the participants computer and keeps the copies synchronized via the internet or network. When any member makes any change to the space, the change is sent to all copies for update (Edwards, 2002). Groove offers both synchronous and asynchronous tools for communication. Synchronous tools allow for real-time audio and text-
chat conversations and simultaneous access to shared space such as co-editing of documents; while asynchronous tools allow for utilization of the shared space. Groove also offers a rich set of tools that can be used in the shared spaces in order to customize the functionality of each space. It also offers application development platform to develop additional tools. Finally, it keeps in memory all the collaboration taking place though the shared space thereby enhancing the group’s ability to interact with each other.

‘People create knowledge and they store it on their desktops,’ says Andrew Mahon, director of strategic marketing for Groove networks. ‘It would be nice if they would publish it to a central location, the way client/server knowledge management tools are supposed to work, but people are lazy. P2P solutions allow people to look around and find the knowledge’ (Kini, 2002).

This quote highlights differences between P2P and client/server models. These differences are easier to see when P2P and C/S models are mapped to the KM process, see Table 1. The primary differences between the C/S tools and P2P tools are in the process of creation to storage and transferring of knowledge. With C/S technology, the transfer of knowledge involves the creator’s and the receiver’s action in terms of transferring and fetching data to/from the server respectively. With P2P technology, the data transfer process is taken care by the technology, once the transfer is approved by the collaborators. Acceptance can occur at various levels, e.g. at file level as in the case of Napster or at a shared space level as in the case of Groove. The creator can simply put documents on his desktop and the receiver can pick it up from his desktop. P2P combines the storage and transferring processes in the KM process, rather than going through the additional layer of the server. This P2P approach has the potential for significantly improving the process of knowledge creation and transfer (Kini, 2002).

The benefits of P2P collaboration tools can be examined by looking at the utility of a shared space. Schrage (1997) cites four characteristics of an effective shared space:

- Highly malleable and manipulability: These applications should be easy to tinker with, edit, or alter them. Groove allows this by allowing seamless access to the resources because they exist at the users desktop, usage of any software that the user is comfortable with, and easy alternation of the shared document. This is unlike the case in a C/S architecture where the document to be manipulated has to be downloaded, changed and then uploaded again to be shared.

- Work both in real time and asynchronous mode: Most C/S architectures work in asynchronous mode. P2P tools, like Groove, not only work in an asynchronous mode but also in real time.

- Are divorced of time and distance: Successful shared spaces create the aura of co-presence; they make collaborators feel they’re together, even if they’re not. This is primarily done with access to thoughts and data captured by technology. Since centrality is the core of the C/S architecture, access to information is restricted by the access and availability of the server. The core of P2P architecture is data being distributed at the peer level. Thus, access to data is available at all times removing the bottleneck of a server in terms of the need to upload and download the data. With Groove, one can take the laptop anywhere and still have access to all the data without having to be connected to the network when needing the document.

- Have Interoperability (Barkai, 2001): A C/S architecture usually restricts the technological parameters like applications, protocols, etc. that can be used to create/retrieve and apply knowledge. Since these are universal
to all knowledge workers across the organization, they are usually not task specific. In case of a P2P tool, the
data is created and stored at a client level. This means that the collaborators have the freedom to use any kind
of applications useful for getting a successful collaboration. For example, Groove applications can store any
kind of data, and can communicate over most of the protocols. Increasingly, P2P applications are getting
increased interoperability across operating platforms.

Table 2 summarizes these benefits. The benefits have been categorized into two levels—technological and
social.

In contrast to these advantages, the following are the three major weaknesses that the current P2P architecture
has compared to the C/S architecture:

1. Authentication: A centralized environment has greater control over the creation and disbursal of
information. Thus, information in such an environment is perceived to be more valid. P2P architectures
do not have such controls. Information can be generated and shared without controls, and thus, there is a
greater chance of contamination.

2. Searching/transferring: An important part the knowledge management process is the ability to find the
stored knowledge at the right time. This involves having effective search capabilities. (Paul, 1991). In a
C/S environment, because of information being centralized and in a standard format, the search and transfer
process can be easily undertaken. In a P2P environment, the data is spread over a wide network, over
different computers, in potential different formats making searching more difficult. The second problem
is the ability to transfer. P2P systems presently assume participation in the shared space before being able
to search for information in that space. This presumption is not necessarily valid in organizations.

<table>
<thead>
<tr>
<th>Process Technology</th>
<th>Creation</th>
<th>Storage/retrieving</th>
<th>Transferring</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server Centric</td>
<td>On Client (Creator)</td>
<td>On Server</td>
<td>On Server</td>
<td>On Client (Receiver)</td>
</tr>
<tr>
<td>Peer-to-Peer</td>
<td>On Client (Creator)</td>
<td>Once initiated, the system transfers data between clients</td>
<td>Receiver fetches data from server</td>
<td>On Client (Receiver)</td>
</tr>
</tbody>
</table>

**TABLE 2**

**STRENGTHS OF P2P COLLABORATION TOOLS IN KM PROCESS**

<table>
<thead>
<tr>
<th>Technology/Process</th>
<th>Creation</th>
<th>Storage/retrieving</th>
<th>Transferring</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>Interoperability</td>
<td>Remains &quot;Single source&quot; bottleneck</td>
<td>Interoperability</td>
<td>Remote</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Load balancing over the net</td>
<td></td>
<td>maintenance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distributed services between clients reducing need for costly data centers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>Ease of creating multiple shared spaces</td>
<td>Bypass centralized control—direct interaction between people</td>
<td>Highly malleable and manipulability</td>
<td>Synchronous and asynchronous modes of working</td>
</tr>
<tr>
<td></td>
<td>Highly malleable and manipulability</td>
<td>Divorced of time and distance</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Divorced of time and distance</td>
<td>Synchronous and asynchronous modes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Technological: Technological issues are in terms of reliability of connectivity, security and privacy, fault-tolerance and availability, performance and bandwidth, scalability, self-management of systems, interoperability, complexity of the infrastructure, etc. (Barkai, 2001). Again, a network administrator has more control over these in a centralized IS environment than a decentralized IS environment resulting in greater control over the above mentioned issues.

P2P tools, such as Groove, are addressing 1 and 3 to some extent through integrating P2P with some C/S capabilities. Number 2 remains an area in need of research.

LEVELS OF COLLABORATION

A good knowledge management system enhances collaboration within and among businesses (Kini, 2002). Collaboration can be viewed at five levels- (i) Individual (ii) Project/ Team (iii) Community of interest/practice (iv) Organizational (v) Across enterprises (Barkai, 2001; Bostrom, Kadlec, & Thomas, 2002; Fattah, 2002), see Table 3. Each of these levels serves a different purpose in the organization. The most basic form of communication within an organization occurs at an individual level. The value generated at this level is the knowledge generated during the interaction. There are various applications at this level, which have been broadly categorized as instant messengers (MSN, AIM, Yahoo!, ICQ being the prime examples). Though these started out as an internet application, corporations are recognizing the need at this level and implementing them over their intranets.

Organizational teams are formed to achieve a project goal. Technology helps in achievement of these goals faster through the use of interactive work spaces. This is the level where P2P tools have had the most impact (Merriden, 2001). For example, Groove implements this concept by helping the users to share plans, schedules, documents and discussions in a shared project space. Departments or more broadly referred to as communities of interest within an organization, essentially share and manage varied amount of information on a particular area.

Over a period of time organizations capture a lot of knowledge in the form of documents, rules, policies, practices, insights, etc; but most of this stored at an individual or team levels. Organizations need tools to help them store and transfer this knowledge in a way that it can be easily retrieve and shared across the organization. P2P tools like Nextpage and PC docs help catalog documents on individual computers in an organization so that they can be shared across the organization. The C/S response in this area is an organizational or enterprise portal.

Inter-organizational collaboration within the supply/value chain has become a major focus of organizations. Although organizations could collaborate at any of the levels outlined (individual, cross-organizational team, etc.) using tools outlined in Table 3, specific B2B P2P applications have just started to emerge. For example, Liquiditynet, a tool for the financial security industry, builds a P2P network that lets institutions do the large block trades outside the traditional market mechanisms (Fattah, 2002).

Although there is no formal research comparing P2P tools versus C/S tools at the various levels, the practitioner press indicates that P2P applications are having their biggest impact and success at individual and team levels, where C/S based tools dominate at other levels. Initial review suggests that both C/S as well as P2P technologies have their advantages and are more suited to different levels. Thus, it seems that both C/S and P2P can co-exist, but more research is needed to understand how each can be exploited. Also, as combinations of these two technologies emerge, the need for research becomes more acute.

RESEARCH ISSUES

Much of the interest surrounding P2P computing is based not so much on what has been demonstrated so far, but on its potential (Barkai, 2001). P2P computing facilitates an entirely new way of working within and among businesses because it will allow spontaneous, secure communication between small groups that can be created
TABLE 3
LEVELS OF COLLABORATION

<table>
<thead>
<tr>
<th>Collaboration level</th>
<th>Type of content</th>
<th>Duration of context</th>
<th>Business Value</th>
<th>P2P Technology/products</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>Presentation</td>
<td>Minutes, hours</td>
<td>Interaction Impact</td>
<td>Instant Messaging Voice over IP Application Sharing Messenger</td>
</tr>
<tr>
<td></td>
<td>Meeting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dialogue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project/Team</td>
<td>Plans</td>
<td>Days, weeks, months</td>
<td>Accelerate Project Work</td>
<td>Interactive Work spaces Groove OpenCola Oculus CO</td>
</tr>
<tr>
<td></td>
<td>Schedules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Working Drafts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discussions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community of interest/practice/department</td>
<td>Relevant content to Community Discussions References Norms &amp; Ground Rules</td>
<td>Days, weeks, months, years</td>
<td>Manage and share information and knowledge</td>
<td>Interactive work spaces Groove/Sharepoint</td>
</tr>
<tr>
<td>Organizational</td>
<td>Workflow</td>
<td>Weeks, months, years</td>
<td>Manage Information</td>
<td>Document Sharing Indexing Groove/Sharepoint Nexpage PC Docs</td>
</tr>
<tr>
<td></td>
<td>Applications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Policies &amp; Practices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Templates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Documents &amp; reports</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analyses</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-Organization (Defined users)</td>
<td>Marketing</td>
<td>Weeks, months, years</td>
<td>E-commerce</td>
<td>Product inventory Catalog Trading Firstpeer Liquidlaynet</td>
</tr>
<tr>
<td></td>
<td>Advertising</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Retailing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trading</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

on the fly (Kini, 2002). This is why research in this area is critical. We know of no studies that have or are looking at P2P applications in the KM area. In this section, we want to outline some key research issues. These issues are broadly classified into structural, level and content.

1. Structure issues

   a. Organizations can be viewed as knowledge processing units. Technology and organizational structure provide two mechanisms to increased the knowledge processing capacity of the firm (Daft & Lengel, 1986; Galbraith, 1974). Technologies such as P2P provide the organizations with options of technology enabled structures where people can collaborate and do business, transcending the artificial boundaries of organization and geography based on shared values and shared interests (Barkai, 2001; Werbach, 2002). It provides opportunities for decentralization and since decentralization inherently breaks down boundaries, it raises opportunities with organizational design. Thus, research is needed to explore the organizational design implications of this new technology and how it can be best used to maximize the knowledge processing capacity of the firm.
b. P2P technologies are also offering new ways to do business. Napster and Liquiditynet are two examples of alternative ways of doing business. Napster saved $600 million in storage costs and over $6 million per month in bandwidth expense (Werbach, 2002). P2P offers options for new business models. Formal research is needed to understand these technology enabled business models.

c. P2P technology provides knowledge workers to access knowledge and work even when they are mobile and not connected to a network. This is because knowledge is proactively pushed to their mobile devices (laptops, PDAs, cell phones, etc) whenever they are connected to the network. Once the documents are pushed to them, they have access to them without having to connect to the network. The idea is to push intelligence/knowledge to the edge. Mobile P2P applications may be one of the key future growth areas for P2P tools. Formal research is needed to analyze the impact on performance and productivity of the ability to actively push documents and being able to access them when the knowledge worker needs them especially from mobile devices.

2. Level issues

a. Table 3 explains the various levels of collaboration. Although P2P applications exist at all levels, but P2P is among the many options that exist for collaboration. The emphasis is on increasing the information processing capacity of the organization (Daft & Lengel, 1986). Thus, P2P applications might not be the best tools for suited at all levels of collaboration. Ray Ozzie, founder of Groove, imagined Groove as a tool for groups of two to 25 people. Any more would probably result in too many voices for coherent conversation (Werbach, 2002). Practitioner literature shows significant application of P2P products at individual and team levels and C/S products lead in other areas of collaboration. However, we need formal research that compares P2P versus C/S tools at each level of collaboration.

b. Exploring tools that combine C/S and P2P architectures or exploring the integration of P2P tools with C/S tools is a rich area of research. Kini (2002) outlines a set of models ranging from P2P to a combination of P2P and C/S. Different models could fit at different levels of information processing needs. Research is needed to find the fit between collaboration needs and different technologies like P2P, C/S or a combination of both.

c. Knowledge captured at various levels of collaboration needs to flow across the levels to provide for a richer knowledge management process. For example, team learning nuggets need to find their way to organizational level so other teams can benefit from the learning. Research is needed to find out how best knowledge can flow across levels and how different combinations of technologies can facilitate this flow. For example, Groove (P2P, team level) recently announced a linkage with Microsoft Sharepoint portal software (C/S, organizational level).

3. Content issues: Since P2P applications force decentralization, there is a loss of control over the content. Research issues include validity of content generated, trust between remote users, protection from undesirable content and ad-hoc sharing of information (Barkai, 2001). Content security and user authentication also represent major challenges in content generation and sharing (Barkai, 2001; Fattah, 2002). In addition, the paper raised a number of issues around transfer of knowledge such as finding the right content. Research addressing any of these issues would be beneficial.

As you can see from the above list, KM using P2P applications is a rich research area. We encourage fellow researchers to explore these opportunities! Of course, good research needs good theories to guide it. Due to space constraints, we could not address applicable theoretical frameworks. However, good theories exist to guide the research we have outlined. For example, the structural research issues could be addressed with such theories looking at organizational agility (as Sambamurthy, Bharadwaj & Grover, 2002), and theories looking at structural designs (Daft & Lengel, 1986). [Samba's and Daft and Engle?]. There are a number of useful frameworks, from organizational behavior and Group Support Systems areas, to address team level research.
issues (for example see Hiltz paper Fjermestad & Hiltz, 1998). Research on the appropriateness of P2P technology at various levels, based on the specific characteristics of tasks at that level, can be addressed through theories such as task technology fits (Goodhue & Thompson, 1995). Research that addressed cross level issues, such as team and organizational, would need to integrate theoretical perspectives applicable to each level.

**BIBLIOGRAPHY**


*Proceedings of the 6th Annual Conference of the Southern Association for Information Systems* 183


