ABSTRACT

MODELING USAGE OF AN ONLINE RESEARCH COMMUNITY

by
Yi Zhang

Although online communities have been thought of as a new way for collaboration across geographic boundaries in the scientific world, they have a problem attracting people to keep visiting. The main purpose of this study is to understand how people behave in such communities, and to build and evaluate tools to stimulate engagement in a research community. These tools were designed based on a research framework of factors that influence online participation and relationship development.

There are two main objectives for people to join an online community, information sharing and interpersonal relationship development, such as friends or colleagues. The tools designed in this study are to serve both information sharing and interpersonal relationship development needs. The awareness tool is designed to increase the sense of a community and increase the degree of social presence of members in the community. The recommender system is designed to help provide higher quality and personalized information to community members. It also helps to match community members into subgroups based on their interests.

The designed tools were implemented in a field site – the Asynchronous Learning Networks (ALN) Research community. A longitudinal field study was used to evaluate the effectiveness of the designed tools. This research explored people's behavior inside a research community by analyzing web server logs. The results show that although there are not many interactions in the community space, the WebCenter has been visited

extensively by its members. There are over 2,000 hits per day on average and over 5,000 article accesses during the observation period. This research also provided a framework to identify factors that affect people's engagement in an online community.

The research framework was tested using the PLS modeling method with online survey responses. The results show that perceived usefulness performs a very significant role in members' intention to continue using the system and their perceived preliminary networking. The results also show that the quality of the content of the system is a strong indicator for both perceived usefulness of the community space and perceived ease of use of the community system. Perceived ease of use did not show a strong correlation with intention to continue use which was consistent with other studies of Technology Acceptance Model (TAM). For the ALN research community, this online community helps its members to broaden their contacts, improve the quality and quantity of their research, and increase the dissemination of knowledge among community members.

MODELING USAGE OF AN ONLINE RESEARCH COMMUNITY

by
Yi Zhang

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Information Systems

Department of Information Systems

August 2004

APPROVAL PAGE

MODELING USAGE OF AN ONLINE RESEARCH COMMUNITY

Yi Zhang

Dr. Starr Roxanne Hiltz, Dissertation Advisor — Date
Distinguished Professor of Information Systems, NJIT

Dr. Marilyn Tremaine, Committee Member — Date
Professor of Information Systems, NJIT

Dr. Il Im, Committee Member — Date
Assistant Professor of Information Systems, NJIT

Dr. Yi-fang Wu, Committee Member — Date
Assistant Professor of Information Systems, NJIT

Dr. Loren Terveen, Committee Member — Date
Associate Professor, Computer Science & Engineering Department,
University of Minnesota, Twin Cities

BIOGRAPHICAL SKETCH

Author: Yi Zhang

Degree: Doctor of Philosophy

Date: August 2004

Undergraduate and Graduate Education:

- Doctor of Philosophy in Information Systems,
  New Jersey Institute of Technology, Newark, New Jersey, USA, 2004

- Master of Science in Information,
  New Jersey Institute of Technology, Newark, New Jersey, USA, 2001

- Bachelor of Science in Electrical Engineering,
  University of Science and Technology, Heifei, China, 1995

Major: Information Systems

Presentations and Publications:

Cathy Dwyer, Yi Zhang, and Starr Roxanne Hiltz, "Using Web Analytics to Measure the Activity in a Research-Oriented Online Community," Proceedings of the Tenth American Conference on Information Systems, New York, 2004.

Yi Zhang and Starr Roxanne Hiltz, "Factors that Influence Online Relationship Development in a Knowledge Sharing Community," Proceedings of the Ninth Americas Conference on Information Systems, Tampa, 2003.

Yi Zhang and Il Im, "Recommender Systems: A Framework and Research Issues," Proceedings of the Eighth Americas Conference on Information Systems, Dallas, 2002.

Yi Zhang and Starr Roxanne Hiltz, "Enhancing the WebCenter for ALN Research," Presented at the Eighth Sloan-C International Conference, Orlando, November 2002.

Starr Roxanne Hiltz, Yi Zhang, and Murray Turoff, "Studies of Effectiveness of Learning Networks," In John Bourne and Janet C. Moore (Eds) *Elements of Quality Online Education*, Volume 3 in the Sloan-C Series, pp. 15-41, 2001.

Yi Zhang, Razvan S. Bot and Starr Roxanne Hiltz, "Building a Virtual Community for ALN", Proceedings of the Seventh Americas Conference on Information Systems, Boston, 2001.

Starr Roxanne Hiltz, Yi Zhang, and Razvan S. Bot, "Building a WebCenter for the ALN Research Community," Presented at the Seventh Sloan-C International Conference on Online Learning, Orlando, 2001.

This dissertation is dedicated to my parents and my sister.

Their endless love and support gave me strength to sustain.

ACKNOWLEDGMENT

I would like to express my enduring gratitude to Dr. Starr Roxanne Hiltz, who not only served as my research advisor and was an invaluable resource, but who sustained my efforts through with her tireless guidance and support. My deepest appreciation is also expressed to my other committee members, Dr. Marilyn Tremaine, Dr. Il Im, Dr. Yi-fang Wu, and Dr. Loren Terveen, for their valuable insight and guidance. I consider myself very fortunate indeed that the entire committee made great professional and personal investments in my work.

Many other faculty members, fellow Ph.D. students, lab assistants in the Information System Department, and co-workers also contributed to this research. My special thanks go to Dr. Murray Turoff, Dr. David Mendonca and Dr. Quentin Jones for their valuable suggestions. I extend deepest appreciation to Razvan Bot and Saad Tariq, who assisted me in developing the underlying technology for the ALNResearch WebCenter.

Finally, I am deeply grateful to my friends, Morgan Benton, Peishih Chang, Hee-kyung Cho, Hyojoo Han, Eunhee Kim, Kevin Walsh, and Dezhi Wu, for their support and encouragement.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF TABLES
(Continued)

LIST OF TABLES
(Continued)

LIST OF TABLES
(Continued)

LIST OF TABLES
(Continued)

LIST OF TABLES
(Continued)

LIST OF FIGURES

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

CHAPTER 1

INTRODUCTION

The chapter starts with a brief description of the research motivation and goals followed by an introduction to online research community and findings from previous studies of online research community.

1.1 Motivation

The appearance of online communities has expanded our life from a local and physical world to a global and virtual world. Online communities have touched all different aspects of our lives including professional endeavors. For researchers, this new phenomenon has become a new way to gather knowledge and exchange ideas. However, web-based research communities have been claimed to be very inactive (Brazelton, 2003). Research has shown that it is erroneous to expect people to participate automatically in online community systems after they are built (Andrew, 2002). Lurkers are reported to make up over 90% of online communities (Katz, 1998; Mason, 1999), yet little is known about what lurkers do in web-based communities. This research aimed to get a better understanding of people's behavior in a web-based online research community; to identify factors that cause peoples' engagement in the community; and to design and develop software applications that may better serve to this type of online community.

1.2 Goals and Approach

Corresponding to the above motivations, there were three main goals for this study. The first goal was to better understand people's behavior in an online research community: including what do people do in the community, what do they expect, what are the things they care about most in the community? The second goal was to design and develop

system features that may improve the community cohesiveness and motivate members' engagement in the community. Although some people may think that technology can solve all problems, it is not true in most real world cases. Even with "perfect" technology, an online community may still not attract people to join. The third goal of this study was to develop and test a theoretical framework which identifies factors that may influence individuals' engagement in an online community. A longitudinal study was conducted with a particular online research community in order to achieve the above research goals. Both web log data and online survey responses were analyzed to test hypotheses and answer research questions.

1.3 Introduction to Online Research Communities

In the 21st century, the Internet is not only an information sharing place, more importantly, it is a social technology that connects people together regardless of time and location limitations. Because of the communication revolution, groups of people from different corners of the world who share similar interests can exchange information and interact with each other via computer networks. These groups, which reside on the computer networks, are referred as online communities or virtual communities. With the expansion of the Internet, various kinds of online communities have evolved. Accompanied by the growth of these online communities, our everyday life and work activities have also changed. This research is particularly interested in how online communities impact researchers in achieving their professional goals.

Scientists are not just independent individuals who work in laboratories all the time, dedicated to curiosity, pursuing their own particular interest in the study of natural phenomena, as often imagined by people. Communication is the essence of science;

research shows that in some scientific fields, over 65% of publications are jointly authored (Over, 1982). Different research groups all have annual meetings and journals to help exchange and transmit ideas or knowledge among group members. Cole and Cole (1973) describe the importance of communication in science as follows:

> *"Scientific advance is dependent on the efficient communication of ideas. The communications system then is the nervous system of science; the system that receives and transmits stimuli to its various parts" (Cole & Cole 1973, p16)*

Communication is observed in the whole process of scientific information generation. Both formal and informal communication are crucial to the research community (Garvey, 1979; Meadows, 1974). In the informal world, scientists exchange their ideas within the scientific community. The product of their exchange is new information, and is the start of new research ideas. In conducting research on these new ideas, scientists always need to communicate with their colleagues or other scientists in the same area to solve the problems they meet during the research. Research done by Kraut and his colleagues showed that informal communication is an important mechanism to help achieve both the production goals and the social goals of groups (Kraut et al. 1990). Finally, scientific information must be disseminated in order to become scientific knowledge.

> *"The well-trained, creative scientist realizes that the information he contributes cannot be fully appreciated unless he communicates his scientific thoughts and ideas in a manner such that further interactive communication with other scientific information (past, present, future) is feasible. That is, new scientific information*

must be embedded in the contemporary grammar of its subject matter." (Garvey 1979, p14)

In traditional ways, communications between scientists were usually through journal articles, annual conferences in the field or within the local area. It was difficult to communicate with other scientists who have the same research interests but are located in a remote place. The degree of communication and collaboration was very limited because of the spatial and time restriction. With the information superhighway, computer networks bring new possibilities to communicate and share knowledge with people throughout the world, and also disseminate new knowledge in a much faster way than traditional methods. The potential of online publication is also influencing the formal publication system (Peek and Newby, 1996). Academic mailing lists and newsgroups also have been adopted as tools for informal communication between researchers. However, there is not much research evidence on how web-based online communities help researchers with their work. This research aimed to examine the impacts of a web-based online community for researchers or on Asynchronous Learning Networks, a form of e-learning.

With the advantage of eliminating time and space limitations, online scientific communities are expected to have a very promising position for improving generation and dissemination of scientific knowledge (Weil, 1998; Kling, 1996). However, reality is not as simple as expected, some web-based research communities have been claimed to be very inactive (Brazelton, 2003). However, there is no research which takes a closer look at what was happening in this type of online community. Although there are few postings in these communities, will this be sufficient to show that such a community

space is not useful to members? Not much research has been done in checking the silent population of these communities. This research aimed to examine lurkers behavior in the communities space and find out what aspects of a community space are perceived to be useful by members.

Previous studies on online research communities have provided us some optimistic results. Hiltz (1984) studied scientific research communities using a computer conferencing system as a communication medium. She found that by using the computer conferencing system, scientists broadened their contacts with other scientists in the same domain and increased their communication across disciplines; they obtained a better understanding of others' work; they got more clarification of theoretical and methodological controversies. In terms of productivity, the results showed that the more time spent online, the more likely users reported increases in quality and quantity of work they accomplished. Star & Ruhleder (1996) studied the Worm Community System which was designed for a geographically dispersed community of geneticists. The community consists of about 1,400 scientists distributed around the world in some 120 laboratories. The system provides a knowledge base for its members to share their research data and reports, and also provides opportunity for them to communicate with each other (see also (Schatz 1992)). The results show that collaborations among scientists may take place across disciplinary or geographic boundaries:

"{Are you collaborating with anyone?} I'm collaborating with people in the worm and non-worm community. Mostly immunologists in the non-worm community, people interested in the immune system. In the worm community, I'm collaborating with [a person in another state], on [a particular gene]." (interview transcript from Star 1996, p123)

Research on academic discussion groups shows that researchers build up weak contacts that make their research more visible and that make them more aware of other researchers' work (Matzat, 2002). All of the above studies were focused on people's behavior in online community. My research will focus on exploring new technology for online research community systems.

Currently most commonly used community systems are mail-lists, Usenet newsgroups, or discussion forums. However, there are questions of their abilities to support knowledge sharing and transfer (Lee et al., 2003). All the studies I mentioned above were based on mail-list or discussion forums. There is not much research on web-site based online community systems, and these systems are more knowledge management systems which focus on knowledge gathering, integration and dissemination than communication systems.

1.4 Chapters Ahead

Knowing more about what happens inside an online research community is very important to the design of online research community systems. However, currently there is not much research in this direction. This research used web usage mining technology to help understand people's behavior inside a knowledge-based research community where the majority of members never post in the community. This research also presented a research framework to identify factors that affect members' engagement in an online community. Awareness tools and a recommender system were designed and developed to increase the cohesiveness of the community and better serve members' social needs and information needs. Literature reviews of online communities and corresponding software applications are presented in Chapter 2. Chapter 3 reviews literature on recommender

systems, which was selected as a solution to solve information overload problem in an online community space. Chapter 4 presents a theoretical research framework that identifies various factors that contribute to or inhibit members' engagement in an online community based on literatures of different areas. Chapter 5 describes design of the software tools developed in this research to serve individuals' social and information needs. Chapter 6 presents research hypotheses and questions based on the framework. Chapter 7 introduces research methodologies that will be used in this study. Chapter 8 shows the results of behavior data of the ALNResearch community. Chapters 9 through 11 focus on data analysis and results for testing the research framework. Chapter 12 summarizes all the findings in this research, discusses contributions and limitations of this research, and gives directions for future research.

CHAPTER 2

ONLINE COMMUNITIES

In this chapter, the background of online community, its history, definition, categories, applications and people's needs in the online community are introduced.

2.1 History

The Internet was invented to provide a rapid, secure exchange of data between federal government offices in times of crisis. In 1969 DARPA (Defense Advanced Research Projects Agency) began the first contract for ARPANET, which later became the Internet. As inventors of the Internet, J.C.R. Licklider and Bob Taylor envisioned online communities even thirty years ago. In one of their articles, they envisioned that this new form of community would bring people more satisfaction because:

> "The people with whom one interacts most strongly will be selected more by commonality of interests and goals than by accidents of proximity. Secondly, communications will be more effective and productive, and therefore more enjoyable. Third, much communication and interaction will be with programs and programmed models, which will be (a) highly responsive, (b) supplementary to one's own capabilities, rather than competitive, and (c) capable of representing progressively more complex ideas without necessarily displaying all the levels of their structure at the same time-and which will therefore be the levels of their structure and rewarding. And, fourth, there will be plenty of opportunity for everyone to find his calling, for the whole world of information, with all its fields and disciplines, will be open to him – with programs ready to guide him or to help him explore." (Licklider & Taylor 1968, pg20) .

Most of his predictions have become true with today's Internet. The Internet not only provides a new way to communicate and exchange information, but also serves as a kind of virtual settlement for a new kinds of communities – the online community (Jones 1997). The Internet fosters online communities by creating new forms of online interaction, providing meeting spaces for people with shared interests that overcome the limitations of space and time (Baym 1995; Sproull & Kiesler 1991; Wellman 2001). The new cyber relationships help individuals to extend their personal social network. A social network is a set of people (or organizations or other social entities) connected by a set of socially-meaningful relationships (Wellman et al. 1996. pp.5).

> *"The relative lack of social presence online fosters relationships with network members who have more diverse social characteristics than is normally encountered in person. It also gives participants more control over the timing and content of their self-disclosures (Walther 1995). This allows relationships to develop on the basis of shared interests rather than be stunted at the onset by differences in social status (Hiltz & Turoff 1993)"*

Despite the lack of social presence, the advantages of no physical limitation and no time limitation provide a larger space for individuals to meet different people online and provide more chances to find people who share the same interests. Moreover, people can be in multiple communities at the same time online without interfering with each other. With all these advantages, it now can be understood why people choose to join different kinds of online communities.

2.2 Definition

Although "online community" or "virtual community" is a very familiar term to people, it is difficult to give an accurate definition to the term. Different people in different fields have different understandings of the term. Hill in 1995 stated:

> "The term community means a group of people who share characteristics and interact. The term virtual means 'in essence or effect only' Thus, by virtual community it means "a group of people who share characteristics and interact in essence or effect only" (Hill et al. 1995, pp.194).

Hagel and Armstrong define a virtual community as a computer-mediated space where there is an integration of content and communication with an emphasis on member-generated content (Hagel & Armstrong 1997). A more popular definition was given by Howard Rheingold:

> "... virtual communities are culture aggregations that emerge when enough people bump into each other often enough in cyberspace. A virtual community is a group of people who may or may not meet one another face-to-face, and who exchange words and ideas through the mediation of computer bulletin boards and networks" (Rheingold 1993, Introduction).

In Jenny Preece's book "Online Communities", she gives the definition more from a development perspective. "An online community consists of:

- *People*, who interact socially as they strive to satisfy their own needs or perform special roles, such as leading or moderating.
- *A shared purpose,* such as an interest, need, information exchange, or service that provides a reason for the community.
- *Policies,* in the form of tacit assumptions, rituals, protocols, rules and laws that guide people's interactions.
- *Computer systems,* to support and mediate social interaction and facilitate a sense of togetherness." (Preece 2000, p.10)

2.3 Community Type

From the purpose of the community, online communities can be categorized into the following types (from http://www.fullcirc.com/community/communitytypes.htm):

- Collaborative workgroups
- Family groups
- Social space
- Role playing
- Illness support groups
- Space for primarily face-to-face groups
- Ethnic groups
- Professional groups
- Software support
- Intellectual discussion groups
- Special interest groups
- Creative groups sharing techniques and work

Different types of community will have very different objectives, content, different system requirements, different activities, and different policies. This research will be focused on research communities (e.g. www.isworld.org, www.chiplace.org), which serve people who are interested in a certain scientific research area.

This appearance of online community has enabled new approaches for global research partnerships, and virtual project teams. In the modern scientific environment, it is becoming relatively rare for a solitary person to come up with a revolutionary new concept and see the idea come to fruition. The changing nature of research has mandated

team-based projects (Tilley & Smith 1996). Virtual professional communities are communities for special research or professional interests. They combine people with similar research interests together, provide opportunities for collaboration among them. However, those communities for researchers seem to be less active than other kinds of online communities (e.g. entertainment communities). How to help online research community members realize this great opportunity and support them in using the system is potentially very important for both research communities' growth and improving productivity for individual researchers.

2.4 Needs in Online Communities

Information seeking is no doubt one of the preliminary reasons to drive people to come to the online communities. Computer networks are the fastest way to spread information. As Cowan proposed:

"Communities are repositories of large amounts of heterogeneous information that need to be searched, read, explored, acted upon, updated, and that offer opportunities for collaboration and other forms of two-way communication... The ability to browse, search, and acquire data is essential for the establishment of community network" (Cowan et al. 1998, pp.1)

Now, people can almost get any information they want from the Internet. On the other hand, the information exploding on the Internet also brings the problem of information overload. Information overload problem and some solutions for it will be discussed in the next chapter.

Besides information, the other most important reason to drive people online is other people online. As William Whyte said: "What attracts people most, in sum, is other people" (Whyte 1988). One very important factor to draw people into online communities is that it provides a chance to meet with other people. People come to the

virtual space to seek social acceptance and develop social relationships as well as look for information:

> "People on the net should be thought of not only as solitary information processors but also as social beings. People are not only looking for information; they are also looking for affiliation, support and affirmation." (Sproull and Faraj 1997, quoted in Donath thesis, p11)

Thus, one of the goals of an online community is to create atmosphere and chances to help community members develop relationships through the community space. For this purpose, computer systems that communities settled in should provide ways to help members recognize other people in the community and provide a friendly, warm virtual environment to motivate community members to start to communicate with each other.

For an online research community, in order to establish collaborations among community members, members of the community need to maintain awareness of one another on different levels. People need awareness information in order to find opportunities for collaborations and people to collaborate with (Kraut et al., 1990). However, the cues and mechanisms used to maintain awareness of others in face-to-face situations are difficult to provide in virtual environments. Awareness is always a very important issue in Computer Supported Collaborative Work (CSCW). In the literature, awareness has been achieved through:

1. Audio/Video support: such as the Media Space (Bly et al., 1993) which provided a video view of remote collaborator's activities in the office; and Portholes (Dourish and Bly, 1992), which reduced the bandwidth requirement by using occasional video snapshots instead of live video.

2. Virtual reality environment: such as MUDs (Multi-user domain). As users navigate through regions in the virtual world, they will encounter others who are in the same region and they are able to interact with each other via text descriptions and commands (Curtis, 1992). More recent graphical virtual environments added visual representations of places and people, thus making it even easier to notice and start interaction with others (Morningstar and Farmer, 1991; Jeffrey and McGrath, 2000).
3. Chat-like communication tools: for example, an Instant Messaging (IM) system can provide some awareness information such as who is online now, who is busy, who is away, etc. The BABBLE project provides more awareness information about other people's activity while in the chat space (Erickson et al., 1999; Bradner et al., 1999).

However, most of the tools above are for synchronous communication. There are not many awareness tools designed for asynchronous communication. For an online research community, in order to help members interact with each other, first of all, the system should help people to recognize other members in the community. Kollock's principle of building online community also emphasizes the need for information about the people and their behavior (Kollock, 1999). Systems that provide perceptually-based social cues which afford awareness and accountability are called "Socially Translucent Systems" (Erickson et al., 1999). In online communities, because of the lack of social cues, the degree of social presence tends to be low. Thus, a system with "Socially Translucent" features should help users to be aware of other users in the system and encourage relationship development among them.

2.5 Applications that Support Online Communities

There are different online applications that provide support for hosting online communities. These applications include:

- Email-list forums (Grier and Campbell, 2000; Thomas, 1997)
- Computer conferencing/Bulletin board systems (Hiltz, 1984; Hiltz and Turoff, 1978; Kuehn, 1993; Rheingold, 1993)
- Usenet news-groups (Kollock and Smith, 1994; Smith, 1992)
- Chat-systems such as IRC (Reid, 1991; Stewart et al., 1999; Vincent, 1992)
- Instant messaging systems, such as MSN, Yahoo Messenger (Handel and Herbsleb, 2002; Isaaces et al., 2002; Tang et al., 2000)
- MUDs and MOOs (Ackerman and Muramatsu, 1998; Ciskowski and Benedikt, 1995; Reid, 1994; Turkle, 1994)
- Webblogs (Conhaim, 2002; Gruhl et al., 2004; Kumar et al., 2003; Grimes, 2003)
- Websites (Chandler, 1998; Donath, 1996; Miller & Mather, 1998; Wynn & Katz, 1997)

In this section, some of the most used systems for supporting online communities will be discussed in more detail.

2.5.1 Listserv

Listserv is an automatic mailing list server developed by Eric Thomas, a graduate student in Paris, for BITNET in 1986. This application helped BITNET grow rapidly during that period (Grier and Campbell, 2000). Listserv is one of the early types of systems that supported online community when the Internet did not exist. Listservs or mailing lists

have been around since the Internet's beginnings as a way of facilitating one-to-many and many-to-many communications using the most common application of the Internet – email (Thomas, 1997). Because of the long history of the system, there are thousands of these email discussion lists on different kinds of topics on the Internet.

When e-mail is addressed to a listserv, it is automatically broadcast to everyone on the list. The function of a listserv is similar to a newsgroup or forum, except that it transfers messages through email and is only available to individuals who are on the list. Members of a listserv interact only with the list server and not other users (http://www.webopedia.com/TERM/L/Listserv.html). This is the most simple and straightforward system for supporting an online community.

As defined by the manual, Listserv is a system that makes it possible to create, manage and control electronic "mailing lists" on your corporate network or on the Internet. The listserv maintains a database for the archive of all posted topics. People can ask for some special one, or they can request a list. When a user sends mail to the Listserv, it will automatically distribute the mail to subscribed users. In addition to being a mail server, Listserv also serves as a file server. It collects files and makes them available for users to collect.

A listserv can be either public or private. A public list is a totally open list, it allows anyone to join and leave, search for archives, etc. This kind of Listserv often can attract many people to subscribe. A private list is a list that has some access control. People have to apply to the owner of the list to get access. Some Listservs use human moderators to decide whether your message can be accepted or be rejected.

No matter whether the listserv is public or private, it can be moderated or unmoderated. In moderated listservs, messages are reviewed by a central authority figure before being passed on to the subscribers. This moderation is often used to reduce spam in messages and filter out uninformative comments like advertisements. However, this moderation requires human effort, and lists with higher activity will cost more effort. Unmoderated lists pass everything that they receive to all subscribers, and do not need any human intervention.

Listservs often support two modes of delivery, either a real-time delivery or a digest delivery. Real-time delivery distributes messages to subscribers as soon as they reach the server. This can result in a number of messages per day for each subscriber, which can even be hundreds if the activity level of the list is very high. The alternative and more commonly used one, digest delivery, summarizes all messages sent to the server over a period of time. When subscribers request digest delivery, the server will send only one message per day (or per week) to the subscriber, which contains all messages through that period. Currently, list servers also offer a Web-based delivery, which lets subscribers, or even non-subscribers read all messages from the list through a browser interface. Web-based listservs also provide functions like navigate list archives, search inside archives. Listservs can also connect to a Usenet news server, allowing Usenet readers to access the messages in the list.

(see http://serverwatch.internet.com/listservers_2.html)

A list server can host multiple mailing lists at the same time, and they may be monitored by different individuals. A typical list server usually has three levels of management:

- Server Administrator: has highest priority and accessibility to configure the list server as a whole.
- List server administrator: has control of list server software, such as the ability to create or delete a mailing list and assign list owners.
- List Owner: Manages a particular mailing list. (http://serverwatch.internet.com/listservers_3.html)

There has been some research done recently on improving listservs. Some of it aims to improve the delivery speed (Nakamura, 1999); some to find a better approach for managing large mailing lists (Chalup et al., 1998). There are also efforts on retrieving information from archives of mailing lists (Smith et al., 1998; Saito et al., 1998). There are also attempts to make mail list service suited to mobile devices, like the cell phone (Sakata and Kurashima, 2001).

Listserv as a group communication tool is best used to exchange information or share information among group members as it sends messages to all the members no matter whether or not they are really interested in the topic (Xu, 1998). The communication is sequential and usually unstructured. Messages on different topics can follow each other and become intertwined. E-mail and listservs are "push" technologies, which means that the sender of the message determines who receives it, not the receivers (Dennis and Wheeler, 1997). This can bring problems with information overload because it is not easy for users to filter or summarize messages (Hiltz and Turoff, 1981).

Most of the communities online today have a Listserv. Even a very loose community like a university or commercial site will use it as a way to connect to their members. Universities will use listservs to distribute school news and events; commercial

sites will use them to broadcast their new products. Figure 2.1 is an example of Listserv service: NIH (National Institute Health) – Listserv.

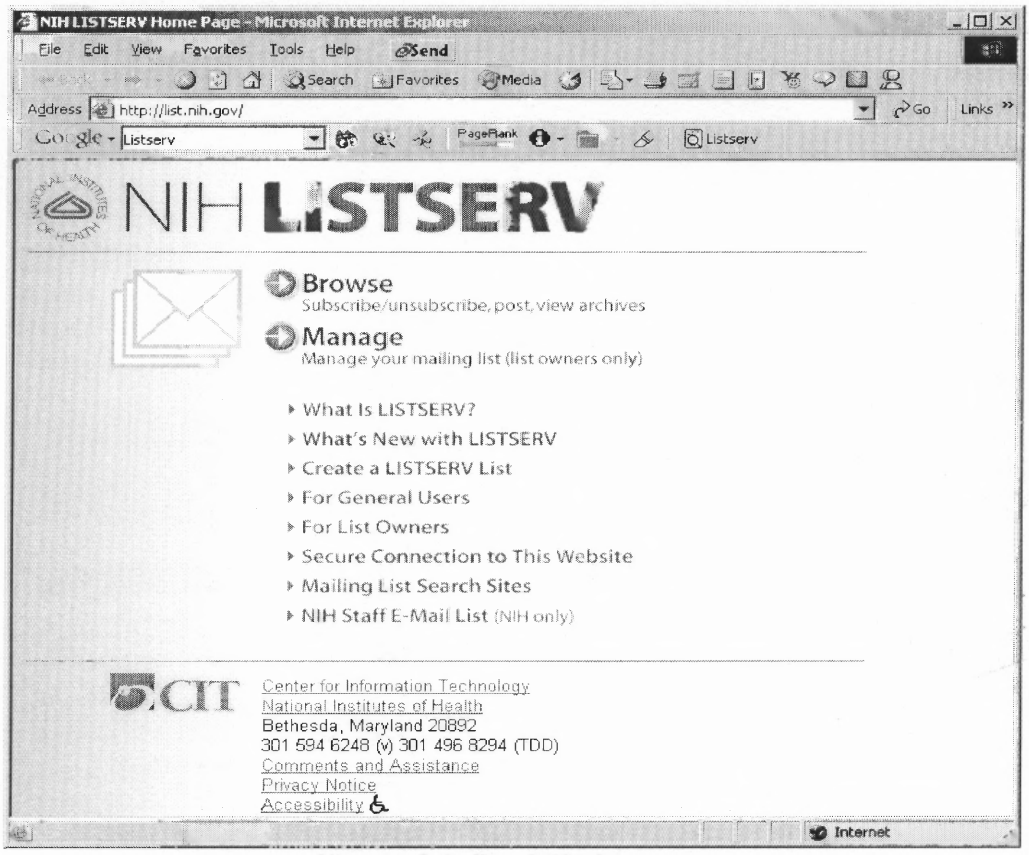

Figure 2.1 The NIH Listserv

Tables 2.1 and 2.2 at the end of this chapter summarize strong and weak points of listservs and also compare listservs with other applications for online communities.

2.5.2 UseNet News

"Usenet is a world-wide distributed discussion system. It consists of a set of "newsgroups" with names that are classified hierarchically by subject (http://www.faqs.org/faqs/usenet/what-is/part1/). Messages are posted to these newsgroups using appropriate software. Usenet is hosted on Internet servers. Usenet encompasses government agencies, large universities, high schools, businesses of all sizes, home computers of all descriptions, etc. Usenet is a tool for a set of people who

exchange articles tagged with one or more universally-recognized labels, called "newsgroups". Usenet can also have moderators to check the messages and decide whether to accept or reject a posted message. The 10 broad categories are:

- Alt (alternative newsgroups)
- Comp (computer newsgroup)
- News (Usenet newsgroups)
- Sci (science newsgroups)
- Talk (Usenet talk newsgroups)
- Misc (miscellaneous newsgroups)
- Biz (Usenet business newsgroups)
- Rec (recreation newsgroups)
- Soc (social issues newsgroups)
- All (a catch all group)

When you choose to subscribe to certain groups, you will find thousands of people are there exchanging information on the subject. Again, some of the newsgroups are very active, they have thousands of messages per week, and some of them are not that active, subscribers there will also be fewer.

Users get all messages from a newsgroup via a newsgroup server, this server can be hosted at any place on the Internet. The server computer should have enough capacity to host all news messages and also the network speed should be high enough to host the site. The administrator of the site can control his own site. When building a virtual community using Usenet, you can assign your members to one category. However, there

are trade-offs between ease of creation and speed of propagation. Launching a new group in the well-known categories can take a very long time.

Usenet newsgroups add some structure to messages. All the messages are organized in threads so that the responses to a message are always linked to the original message. With this structure, readers can follow the related discussions. Newsgroups are "pull" technologies, which means that users (receivers) can select which messages to read. By adding some structure to the message, it helped to reduce the problem of information overload. Also because newsgroups only save one copy of the message at the server, it reduces the network traffic and storage. However, there are still problems like security, because it is more difficult to restrict access to newsgroups than listservs (Dennis and Wheeler, 1997). Figure 2.2 shows how it looks like when reading news from netscape.

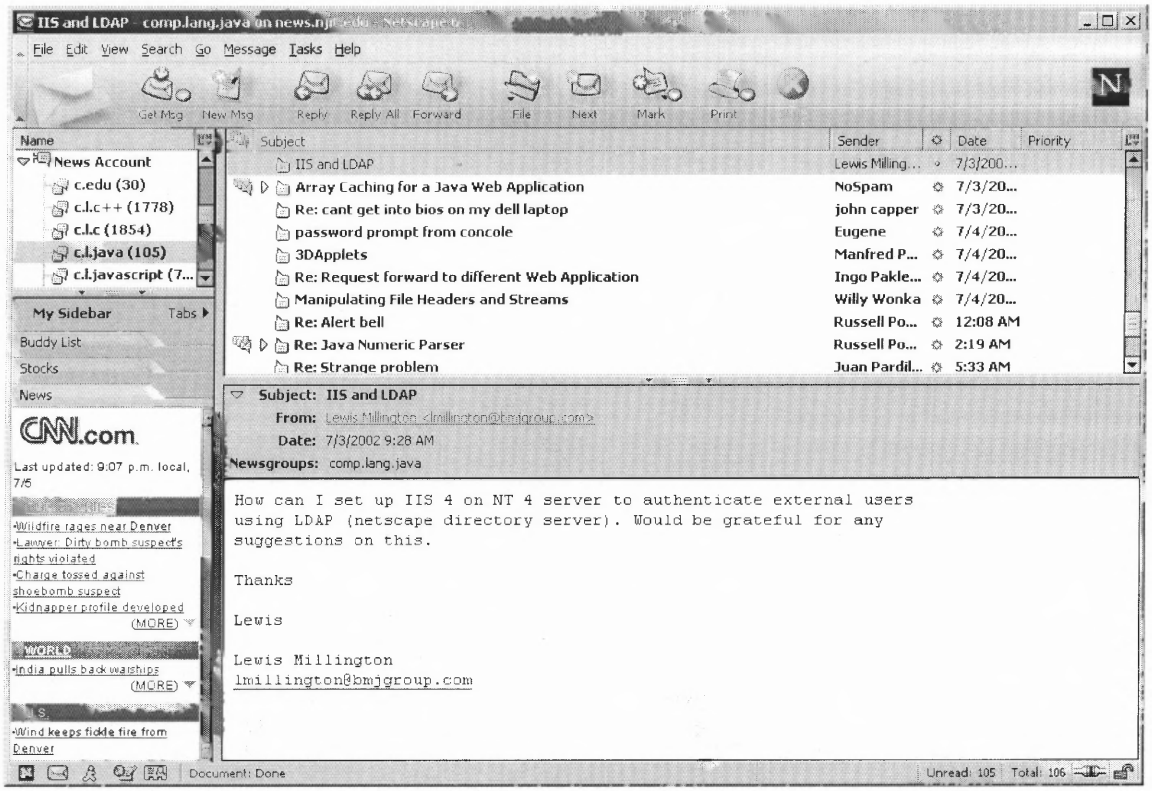

Figure 2.2 Netscape Interface for News Reading

Table 2.1 and 2.2 summarize strong and weak points of newsgroups and also compare newsgroups with other applications for online communities.

2.5.3 Chats

IRC(Internet Relay Chat) was developed in Finland in 1988 by Jarkko Oikarinen. By mid-1999, there were more than 45,520 IRC channels on 27 IRC networks.

Unlike Usenet and Email, Chats are usually synchronous text systems. For example, in IRC, using IRC client (a computer program), you can exchange text messages instantly with people all over the world. There are different channels that allow users to become involved in different conversations. On large networks, such as Efnet, more than 2,000 channels can exist. Some of the more popular chat clients are mIRC, Pirch, and Virc for Windows and Homer or Ircle for Macs. But IRC is not something that you should just jump into without knowing how it works. You could feel very lost if you do not know some of the basic commands before you enter IRC, and you might not even know how to get out of IRC. So, you'd better read some documentation (readme files, faqs, etc) before connecting to an IRC server.

Because of the effort people need to spend when using IRC, there are alternatives now everywhere on the Internet. For instance, people can access chat rooms and start chatting with other people via a browser interface, which does not need extra effort. The figure below is a chat room in MSN. It shows participants who are in the chat room now. You can talk with them in the public area, while you also can talk to a particular person by sending him or her a private message. Figure 2.3 shows a chat room on MSN.

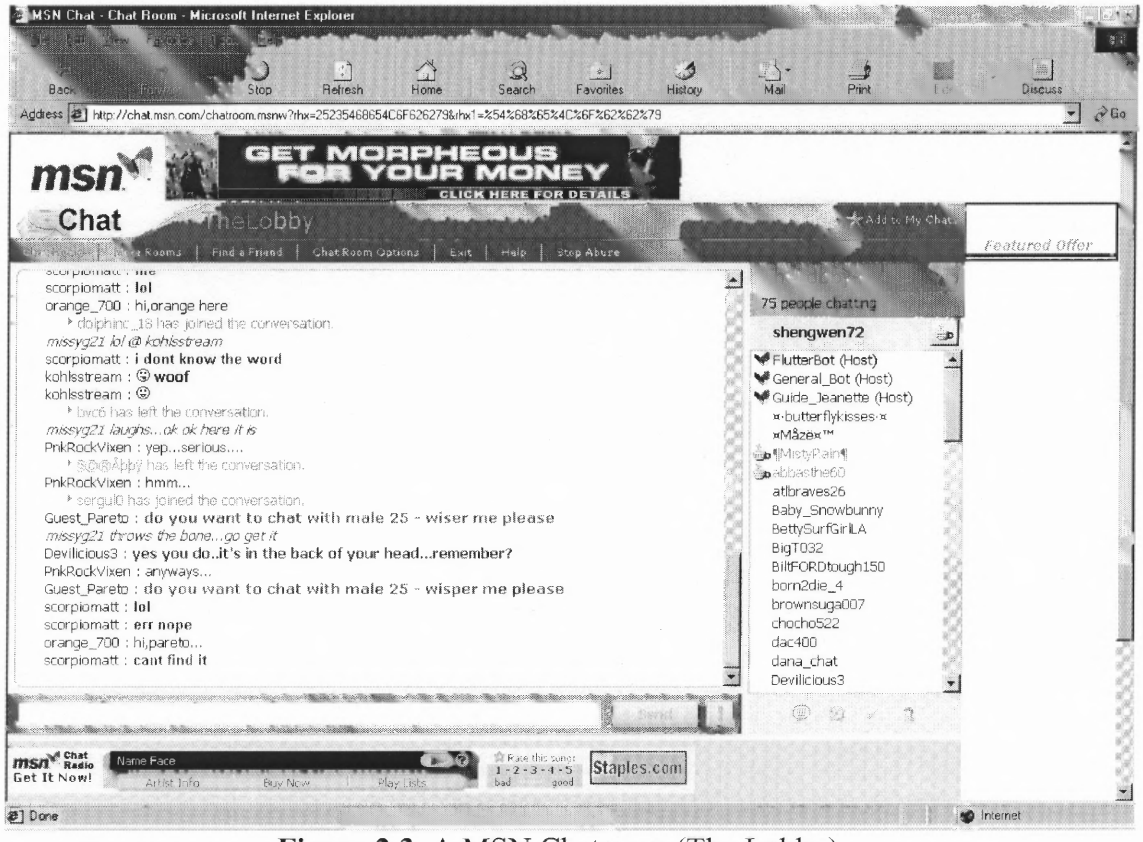

Figure 2.3 A MSN Chatroom (The Lobby)

Compared to mailing lists and Usenet newsgroups, chat room applications are "consummatory communication" which is more suited for people to share enjoyment, casual conversation among friends. However, Chat applications are not only used to socialize, to get to know new people, but also used a lot in the Online education field (Erickson et al., 2001; Cohen et al., 2000; Irvine et al., 1999) and distance collaboration work (Graveline et al., 2000). In distance education, because students do not need to come to the institution to take courses, they take courses online, discuss in a conference system, and send email to other students or teachers, while chat may be the most convenient tool that they can use for synchronous communication. There is also lots of effort on how to visualize chat logs to make them more useful (Borner and Lin, 2001; Donath et al., 1999).

Currently, a variety of kinds of chat software exists, some of it traditional text-based chat, others provide beautiful 2D or 3D chatting environments, which makes it more interesting and adventurous. Chat software with graphics and roles make it look more like a real community, as if you were stepping in to a club or coming across someone on the street, and it is more fascinating. However, this kind of 3D graphical application requires a high performance computer and a connection with higher speed. In order to let all kinds of users use the application, the software usually provides an alternative text-based version for those who do not have high speed connectivity or who have low computer performance. Figure 2.4 shows the homepage of "Activeworld" - a 3D graphical virtual environment application. This virtual environment has a very good graphical effect that user can enjoy. However, if you want to experience the virtual environment, you need to download the client application and install it on your own machine. Also because of the graphical effect, every page of the Activeworld is presented by a huge amount of data, it requires high transfer speed, otherwise it will take too long for each page to load. Figure 2.5 shows the virtual reality chat room from the Activeworld.

Figure 2.4 The Entry Page for Activeworld.com

Once you enter the activeworld, the whole screen contains four parts. The left window shows choices of different worlds to visit and also shows where you are. The choices for environments are: Yellowstone National Park, the planet of Mars, a university, shopping mall, etc. Using cursor keys, participants can move their avatars through the environments, usually it is accompanied with sound effects. The university environment can be used for online learning, while shopping malls are more like an e-commerce site where you can buy things. Figure 2.5 shows the gateway of Activeworld university.

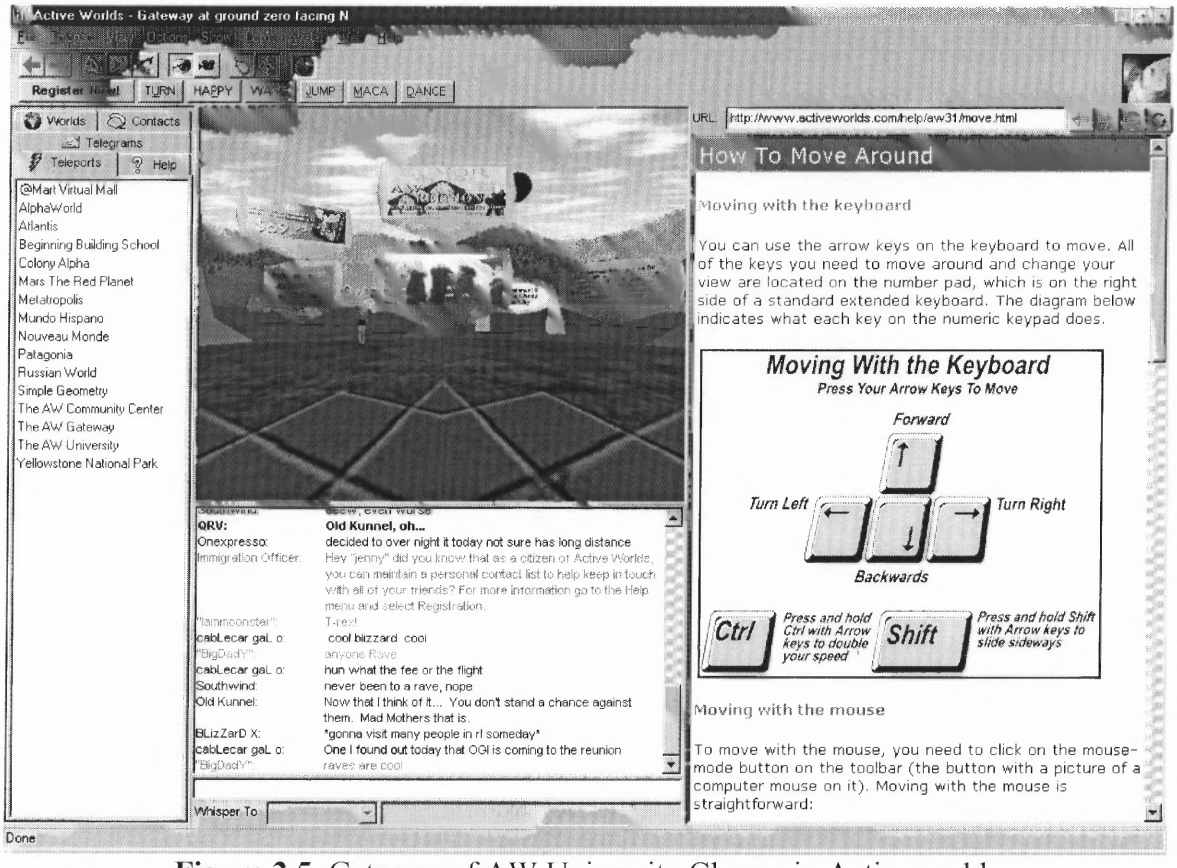

Figure 2.5 Gateway of AW University Classes in Activeworld

Table 2.1 and 2.2 summarize strong and weak points of chats and also compare chats with other applications for online communities.

2.5.4 Instant Messaging

Instant messaging systems such as MSN, yahoo messenger, ICQ, AIS, etc. allow groups of users to stay in contact while doing other things at the same time. IM (Instant messaging) is a very simple technology with complex implications. The application usually is very small, its basic function is to help users identify who in their contact list is online at the moment and also communicate with them using text, audio or even video. It also provides different statuses that user can choose, including "I am available", "I am busy", "Be right Back", "I am not at my desk", etc. Users can also customize the message of status. Also, there are times when, although you are online, you do not want to talk to

people; then you can choose to be "invisible" to others, or you can block certain people in your list so that they won't be aware that you are online at the moment. Instant messenger does not only support two people "talking", it also supports a "conference" feature where you can invite more friends to join and talk together. The application has rather complex implications because it gives people enough capability and flexibility to show their identity and awareness of others.

The IM (Instant Massaging) products fall into two broad categories. Public products let anyone download a client and establish an account with a central, Internet-based service. This kind of product includes Yahoo Messenger, MSN Messenger, and AOL Instant Messenger; these products are widely used now on the Internet. When the provider is also hosting chat room services, IM applications are often connected with chat room applications; users can use the IM interface to send a private message to other people in the chat room, and also add them in their contact list. The second category is private products that you can set up on your own IM server for users on your local network; this kind of application is always used in an organization's intranet to let employees contact each other.

IM is not only used for people to talk with other people online for fun, for social reasons, but it also presents its power in the work place. This is a convenient and undisturbing way to contact colleagues (Herbsleb et al., 2002; Bradner et al., 1999; Churchill and Bly, 1999). Because IM applications are usually very light and convenient, they are among the first applications to try on pervasive computing devices, like cell phone or handheld devices. There are already palm or pocket pc versions of IM

applications. Text messaging services on cell phones are also very common now (Cherry, 2002).

Although IM applications provide a very nice way for people to connect to other people online, and get to know new people all over the world. They also bring danger at the same time. When you use instant messenger online, your personal information and your location can be easily detected online, and this will give hackers information and also virus as a way to attach the computer. IM is also used by people for online violence as there is no way for people to monitor the behavior in instant messenger. The following graph shows the MSN instant messenger, it shows your contact list in the middle, using different colors to show whether the person is online or not, the left bottom shows your current status. It also has functions to let people call other people using the Internet or having voice chat or video chat. Figure 2.6 shows the interface of MSN messenger. Table 2.1 and 2.2 summarize both strong and weak points of IMs, and also compare IMs with other applications for online communities.

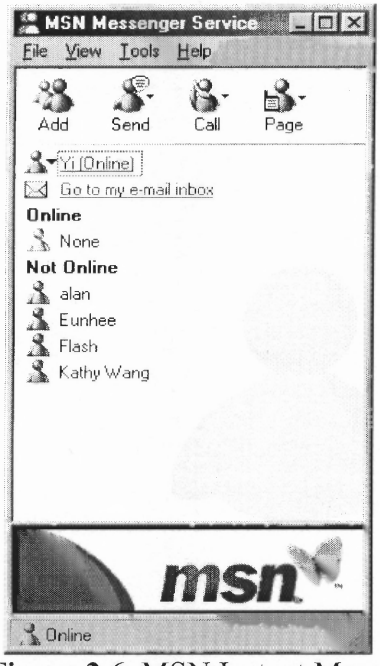

Figure 2.6 MSN Instant Messenger

2.5.5 Computer Conferencing Systems (Bulletin Board Systems)

The first computer conferencing system was created in 1970, and was used by tens of organizations and a few thousand people by the late 1970s (Hiltz and Turoff, 1978). EIES and "The Well" are among the earliest computer conferencing systems (Hiltz and Turoff, 1978; Rheingold, 1993). It is also one of the earliest types of systems that support online community. Conferencing systems provide people with a space to discuss together, each topic will be arranged as a thread. Bulletin boards, which have evolved to be indistinguishable from computer conferencing, are a very commonly used system for virtual community as they provide a very convenient way of discussing among community members.

Conferencing systems also have private and public conferences. The public ones always allow guest participation and do not control the entry and leaving of the community. However, usually "guest" users are only lurkers, they can read and browse the board, but they cannot participate, they need to join and get their own id to post a message. Private BBS is a more tight community, the board administrator controls who can be in and who cannot. Members all have their own identity to participate in the discussions that are going on in the board. The content of the board is also different for public boards and private boards. Public boards usually have a wider range of discussions, more topics, while the private ones are more specific and have fewer topics. Figure 4.7 shows a BBS hosted on the Internet for a learning environment.

Conferencing systems have been used as the main media in online education or distance collaboration work (Middleton, 1999; Rice, 1983).

In recent years, there have been lots of efforts on the multimedia conferencing systems, like video conferencing systems or audio conferencing systems (Onishi et al., 2001;Wang, Y. et al., 2000; Wang, C. et al., 2000). There are also real-time 3D environments for conferencing systems (Dai et al., 2002). Figure 2.7 shows a sample conference system – webboard system. Table 2.1 and Table 2.2 summarize strong and weak points of discussion forums and also compare forums with other applications for online communities.

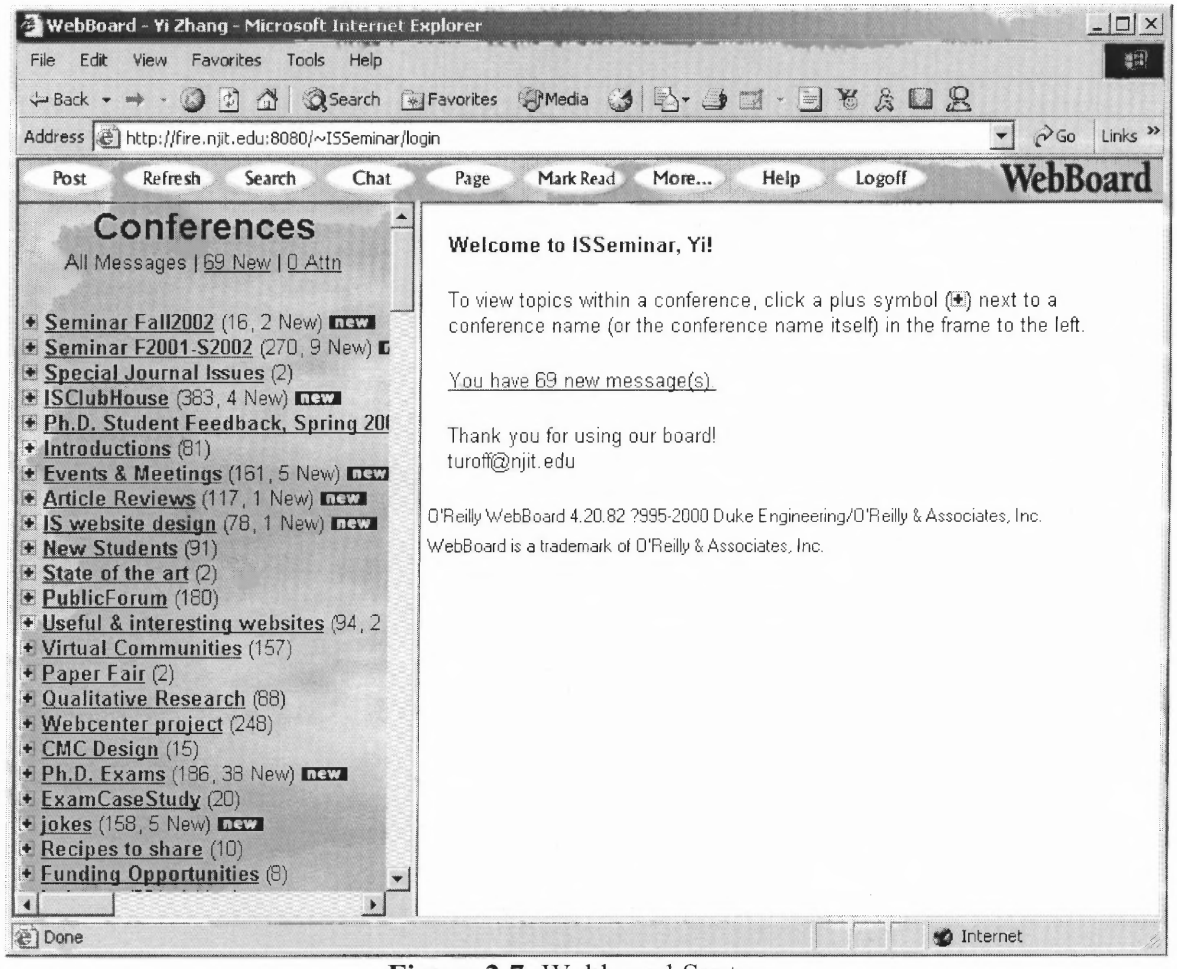

Figure 2.7 Webboard System

2.5.6 MUDs

A MUD (Multiple User Dimension, Multiple User Dungeon, or Multiple User Dialogue) is a computer program which users can log into and explore. Each user takes control of a computerized persona/avatar/incarnation/character. You can walk around, chat with other characters, explore dangerous monster-infested areas, solve puzzles, and even create your very own rooms, descriptions and items. A MUD system is a Client/Server infrastructure. The MUD server manipulates the database of objects in the virtual world, is programmable in some languages that allow one to extend the set of objects, and accepts network connections from clients. The client's primary task is to send and receive I/O between the server and other users. The MUD server exists on one machine on the network, while the client is run on the user's local machine. MUD has very good features for group communication, such as: MUDs interact in real-time; MUDs are multi-user capable; MUDs are extensible; MUDs hold all the exchanged messages as history that can be retrieved later.

As a MUD application provides a metaphor of real life, every user in it has a specific role; this makes it more like a real world. Today, there are perhaps a dozen popular types of MUDs available on the Internet. They vary in many details, such as their embedded programming language, but all of them have the capacity to allow multiple users to interact within some shared context. MUD servers are usually running on a UNIX machine, its client can run on any system. It only needs telnet to connect to the server. Usually, the beginners need to see a help file before using the system as there are a lot of unfamiliar commands; you also can get help from other people who are in the

"world". However, as it is a game based environment, learning some new command is not hard for people to accept.

The MUD world is just like the real physical world. It has its own policies. If you behave badly, you will get a very bad reputation, and eventually receive the "NO FUN" stamp of disapproval.

2.5.7 MOOs

Multiple-user, Object-Oriented environments (MOOs) are very similar to MUDs, they are synchronous online multiuser spaces. A MOO is a virtual place created on the Internet. A MOO is also a Client/ Server infrastructure. The MOO server is hosted on an Internet server and is accessed through a text relaying client program, such as telnet. MOOs are written in a language that is a cross between C++ and LISP, called MOOCode. MOOCode is a form of OOP (Object Oriented Programming). Users of MOOs can get what is called a Programming Bit (basically a level of access to the core computer) that allows them to program in MOOCode. Although a MOO is a synchronous space, it is unlike chat; besides real time dialogue, users in a MOO can move around the virtual space and can build rooms and create objects.

Because of its game based nature, MOOs have been developed to be used in an educational environment. Students can use their creativity to work in the virtual space with their classmates, it gives them more fun and attracts them to involve them more in the discussions. For example, AquaMOOSE 3D is a research project developed to create an interactive learning environment for high school aged students. It is geared for ages 15-18, although all ages are welcome. AquaMOOSE 3D is designed to allow users to

explore and experiment with 3D mathematics. Figure 2.8 shows an example of a MOO environment.

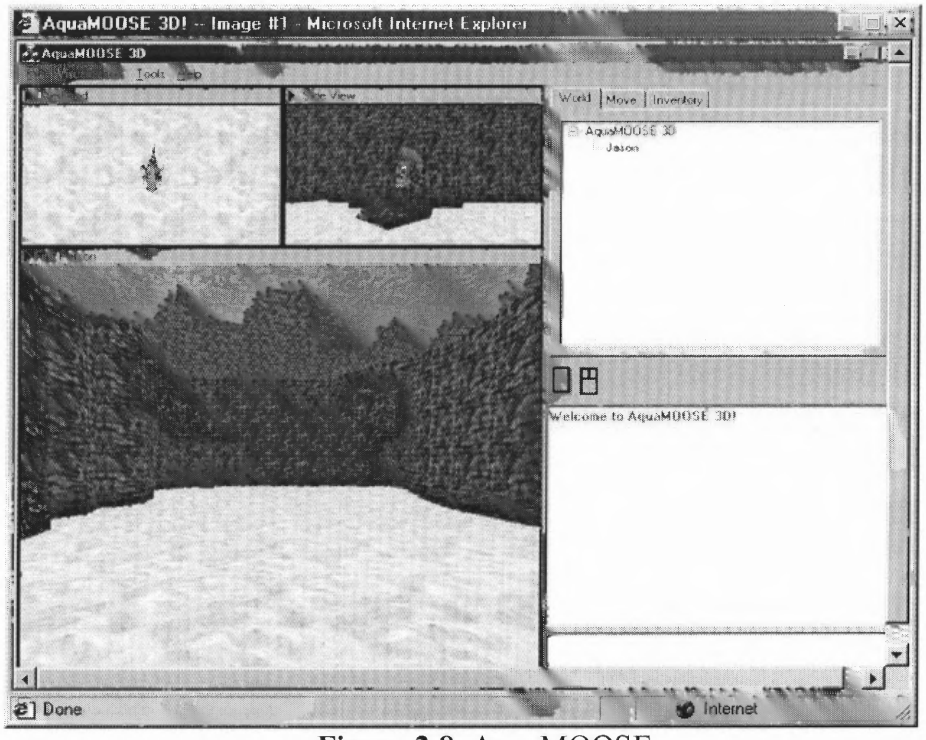

Figure 2.8 AquaMOOSE

Table 2.1 and 2.2 summarize strong and weak points of MUD (MOO)s and also compare them with other applications for online communities.

2.5.8 Blogs

A blog is a web page made up of usually short, frequently updated posts that are arranged chronologically – like a what's new page or a journal and it also allows users to discuss a news item. The contents and purpose of blogs can be very different, depending on the people who host them. Many blogs are personal, used to share their thoughts or certain news on the topic a person is interested in. Bloggers are also excellent team/department/company/family communication tools. They help small groups communicate in a way that is simpler and easier to follow than email or discussion forums. It can be used on an intranet to let team members post related links, files, quotes,

etc. A blog helps keep every team member in the loop, promote a cohesive group culture. (http://www.blogger.com/about.pyra)

A blog can be operated with a browser interface, so it is very simple, does not need any extra effort for users to use. The first page of a blog changes automatically as the new items have been received, so it always keeps the newest news on the top. It also provides categories and a search function for people to get what they are interested in.

One of the most active blogs is www.slashdot.org, it is a blog about science and technology, particularly open source. It has millions of members all over the world reading and posting every day. A news item posted in the morning will get 200 comments in the afternoon. It also has a good and flexible strategy to organize all the messages, users can choose their preferred format to display all the messages, and also each message has been scored 1-5 according to its contents. Because of its large member population, Slashdot has many volunteer members as monitors to filter the messages and also to give a score to the message. Everyday, there are millions of postings, and it has almost all news and events and new books that are related to science and technology. It is amazing how such a huge traffic can be organized beautifully. Figure 2.9 shows the first page of Slashdot, the content of menus and top icons also changes according to the news content in the first page.

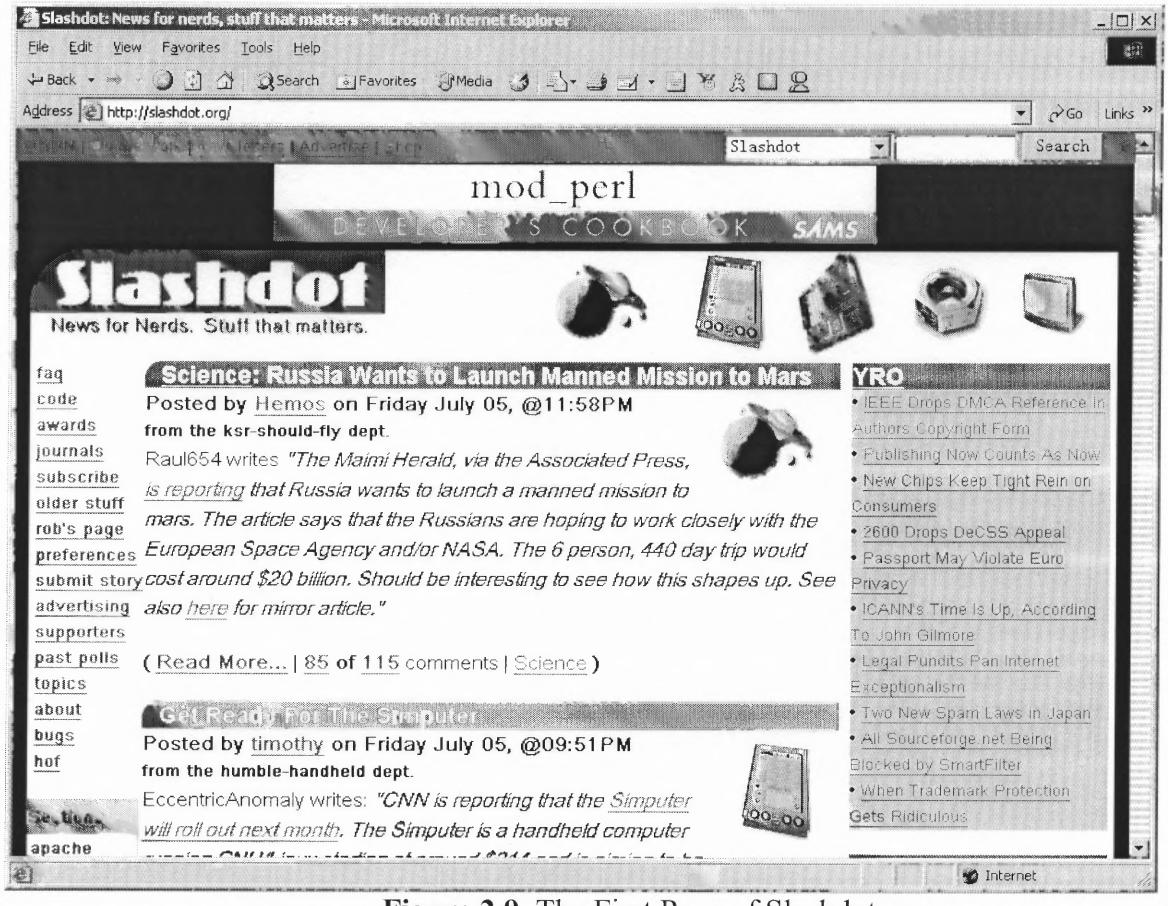

Figure 2.9 The First Page of Slashdot

Table 2.1 and 2.2 summarize strong and weak points of bloggers and also compare bloggers with other applications for online communities.

2.6 Summary

With the appearance of online communities, on the one hand, users are trying to adapt the technologies that enable them to participate in cyberspace. On the other hand, more computer applications are made to fit this new phenomenon. As this is a quite new area, researchers and developers are trying to use knowledge from existing fields like sociology, psychology, and information systems to study such communities, and to seek better ways to serve and foster online communities to thrive, which means there are sufficient people and enough activity to make it attractive and worthwhile (Markus, 1987,

1990; Morris and Ogan, 1996). As mentioned in the review of virtual community (Lee et al., 2003), there is a particular need for research on technology development for online communities.

This chapter also introduces most popular software applications for hosting an online community. Table 2.1 summaries the characteristics of different applications mentioned in this chapter. Table 2.2 summarizes strong and weak points of different applications.

In an online research community, information need is the preliminary requirement for such a community. In the next chapter, the focus will be on applications to help online communities with their information needs.

Table 2.1 Comparison of Different Systems for Online Communities

	Listserv	Newsgroup	Chats (IM)	MUDs (MOOs)	BBS (conferencing)	Blogger
Communication mode	Asynchronous	Asynchronous	synchronous	synchronous	asynchronous	asynchronous
Presentation mode	Push	Pull	Pull	Pull	Pull	Pull
Community size	large	large	small	small to medium	medium	large
Identity	Email address	Email address	Users choose their id	Users choose their own id	Usually requires it, also can post anonymous	Use id and email address
Social awareness support	Only shows person who sends out messages. Doesn't show who are the members and members' behavior	Only shows who sends messages. Doesn't show who are the members and members' behavior	Shows who are currently in the chatting space. For IM, it will also show different status of the person (e.g. away or busy). Doesn't show members' behavior	Shows who is around in the same region. Doesn't show all users and members' behavior	Shows who posts messages in the conference. Usually shows membership list and last on date	Shows who provides the news and also shows person who did discussion on a certain message. Only shows the behavior for the member who current logged in, doesn't show other members' behavior
Information awareness support	Information is gathered from members and reaches to members whenever someone contributes a message or combines all messages in time period. Most listservs provide archives for members to search for information they wanted. However, there is no categorizing, integration, or synthesis of the information.	Similar to listserv	No storage for interaction between individuals	Individual can save their sessions in MUDs. However, there is no knowledge base for the whole community	All messages posted in the bulletin board are categorized by thread. Besides thread, there is no categorization or organization. This form of knowledge base is more appropriate for problem solving.	Has broad category and archive for past messages. However, there is no other organization of the information.

Table 2.2 Strong and Weak Points of Different Applications

	Listserv	Newsgroup	Chats (IM)
Strong Points	➢ Simple and straightforward ➢ Easy to learn, easy to use, the only thing the user needs to know is email software ➢ High accessibility, people only need to access email to receive all the messages	➢ Easy to use. Members only need news reader software to access the service ➢ The thread structure of the message make it easy for readers to follow the discussion ➢ Can serve a large number of people	➢ Real time system, participants can get very quick response. ➢ Provide e-emoticons to help people express their feeling while talking ➢ Graphic environment makes it feel more like a real world meeting ➢ Very good environment to start a casual conversation
Weak Points	➢ Because all the messages are distributed to a group of people, sometimes, two people's private argument will go to everyone's mailbox, which is not wanted by the owner ➢ As people can usually post anything to the list, it is less controlled ➢ Needs human moderator to judge the messages, which is a large amount of work for the moderator ➢ Cannot provide real time communication ➢ Flaming and Spamming often occur	➢ As newsgroups often contain a large amount of information, it always takes a long time to download messages. ➢ It is very difficult to restrict access to newsgroups ➢ It will take a long time to create an active group ➢ Not all newsgroup servers are updated simultaneously. There might be some inconsistencies in message structures among different newsgroup servers (Dennis and Wheeler, 1997). ➢ Need an expert to run the server or to access someone else's server.	➢ Graphics Chat is not fit for a low speed connection ➢ Chat with audio features also needs a high speed connection ➢ Very hard to control the flaming and spamming as everything is going on real time ➢ In a graphic environment, when the number of participants increases to a certain point, it will make the entire screen look crowded and hard to search

Table 2.2 Strong and Weak Points of Different Applications (Continued)

	MUD (MOO)s	Discussion Forums	Blogger
Strong Points	➢ Easy to access ➢ Can do all activities in realistic world, with more adventurous experience ➢ Fun, interesting and exciting ➢ Can also be used as a communication method	➢ Easy to post a message ➢ Easy to set different access privilege level for member. ➢ Easy to create private forums for a closed group	➢ Very easy to use, user only needs to know a web browser to get access to blog ➢ Dynamic content, every time a user visits the blog, it will have a new look with new items on the top ➢ Can support large member population without lost organization
Weak Points	➢ Users need to learn commands and programming to be an expert ➢ Most systems are text-based ➢ When many people are in the room, conversation can be hard and easily get confused ➢ MUD can be a big distraction, when it's constantly running somewhere on your screen, it can interrupt your concentration	➢ Hard to find related information when number of postings exceeds certain level ➢ There might be a lot of duplicated messages as the structure intentionally organized.	➢ The administrator will have a tough job to maintain the blog, especially when member size becomes large ➢ Does not support synchronous discussion ➢ When blogger gets bigger, it is hard to structure and present items

CHAPTER 3

RECOMMENDER SYSTEMS

For an online research community, it is very important to fulfill members' information needs. Community members need a very well structured knowledge base to help their research. In this chapter, recommender systems will be introduced as a way to structure a knowledge base and effectively find the most valuable information based on individual preferences.

3.1 Information Overload Problem

Berghel claimed that cyberspace is the only term that will come to be associated with the 21st century (Berghel, 1997). As we step into this digital age, together with the capability of getting more information than ever before, people also have to face the information overload problem as it becomes more and more serious. Now, we do not get information just from newspapers, colleagues, family members and friends, but also largely from the Internet. The Internet grows so fast that it has already become the largest information source ever. It is estimated that there were over two billion Web pages, and thousands of newsgroups and forums by the year 2002, on the Internet - covering virtually every topic imaginable (http://www.jisc.ac.uk/techwatch/reports/ikm1.htm).

Defined traditionally by Sheridan and Ferrell, information overload means "information presented at a rate too fast for a person to process" (Sheridan and Ferrell, 1974). However, when describing information overload in the context of Computer Mediated Communication (CMC), the definition needs to add some new content because the term refers first to the delivery, if there is more communications access to individuals than they can easily respond to (e.g. emails, newsgroup messages), and second to what

might be termed information entropy, where incoming messages are not sufficiently organized by topic or content to be easily recognized as important or as part of the history of communication on a given topic (e.g. the result returned from a search engine)" (Hiltz and Turoff, 1985, p682).

Because of the problem of information overload on the Internet, people expend great effort to filter out useful information through this digital plethora. Various filtering technologies are used to help people find useful and related information. Malone and his colleagues propose three types of information filtering activities: cognitive, economic and social (Malone et al., 1987). Cognitive or content filtering techniques filter documents based on their content. For example, people may want to look for a certain category of information. Economic filtering techniques filter information based on estimated search cost and benefits of producing and reading the item. For example, Malone argues that mass mailings have a low production cost per address and should therefore be given lower priority. Social filtering techniques filter information based on individual judgments of quality communicated through personal relationships. For example, you might want to see a movie because a friend of yours suggests you watch it.

3.2 Recommender systems

Among all the filtering technologies that have been developed, recommender systems have gained more attention recently. Even e-commerce sites have started to use such systems to provide suggestions to their online customers (e.g. www.amazon.com, www.Cdnow.com). Recommender systems are information systems that are designed to provide recommendations about goods or services to some specified population of people (Resnick and Varian, 1997). In applications of recommendation systems, there are two

main filtering technologies mentioned by Malone that are involved: one is content filtering, also called information filtering (IF); the other one is social filtering, also called collaborative filtering (CF).

IF systems are very similar to Information Retrieval (IR) systems in that they both help to find relevant information and filter out the undesirable items. However, they also have fundamental differences which make IF a distinct research field. Traditional IR systems have been designed to retrieve relatively short-term information needs quickly. These systems usually deal with a static set of documents and short-term information needs presented as keyword search or queries (Salton and McGill, 1983). In contrast, IF systems help people to fulfill their long-term information needs. IF systems usually deal with dynamic document sets (Belkin and Croft, 1992; Good et al., 1999; Mostafa et al., 1997). The characteristics of the information filtering systems match some of the needs that people want from the Internet quite well, i.e. users' interests are long-term and the Internet is a dynamic document set as it changes all the time. IF systems gained popularity mainly as a part of various information services based on the Internet (Edwards et al., 1996). Information filtering focuses on tasks involving classifying streams of new content into categories. Information filtering (IF) systems require a user profile that presents items based on the user's long-term information needs or preferences. Early IF systems usually used the user profile that purely depends on the user's explicit inputs. More advanced IF systems maintain dynamic user profiles which update themselves based on information learned from user behavior. The Vector Space Model (VSM) and Inference net model are usually used in IF systems to cluster items in the data set. The advantage of IF techniques is that they can work well without other users'

involvement. However, depending only on term matches without understanding the content sometimes gives low quality recommendations. Because of the dynamic dataset and the changing of user profiles, it is always hard and expensive to get accurate predictions.

Effectiveness is usually emphasized on evaluation of IF systems. The best-known measures for effectiveness are precision and recall. Precision is the proportion of a retrieved set of documents that is actually relevant. Recall is the proportion of all relevant documents that are actually retrieved.

Because of the problems with IF systems, Collaborative/Social filtering (CF) technologies were introduced in the last decade. They attempt to solve some of the problems with content-based systems (Herlocker et al., 1999). Instead of considering items' similarity, Collaborative filtering (CF) systems build a database of user opinions on available items. This kind of system employs users' ratings on items to predict users' preferences and also calculates similarities between users to find users whose interests are similar and create "neighbor groups" depending on the similarities. It then generates recommendations to a specific user based on the ratings that are given by other users in his/her neighbor group. Ratings are usually collected by asking users to fill in their evaluations explicitly in the earlier CF systems.

More recently, a number of systems have begun to use implicit ratings, such as history file, browsing behavior or suggestions from Usenet groups. This kind of system infers user preferences from user's behavior rather than requiring the user to explicitly rate the item. CF systems derive recommendations based on evaluations of other users who share similar interests with the particular user. They fulfill a computerized process of

'Word of Mouth' (Konstan et al., 1997; Shardanand and Maes, 1995). For example, a CF based system would recommend a book to a user because other users who have similar interests rated the book highly. As CF systems are based on other users' opinions about the item, they provide a degree of quality of the item based on humans' judgments, not on the item's attributes. Because of this characteristic of CF systems, they are generally perceived to be more useful than IF based systems (Herlocker et al., 1999; Resnick and Varian, 1997). The algorithms that are used in CF systems to generate like-minded people include correlation, vector similarity, Bayesian network model, etc. Detailed discussions about these algorithms can be found in (Breese et al., 1998; Herlocker et al., 1999; Ungar and Foster, 1998).

One key advantage of CF systems is that they help users to find items of interest that are unexpected by the user because these are recommendations from like-minded people, not based on keyword matching. At the same time, CF's dependency on human ratings can also be a significant drawback. There are two major problems related to collaborative filtering methods: *sparsity* and *first-rater* effects. The sparsity problem is that for a CF system to work well, several users must evaluate each item. The first-rater problem refers to the fact that new items cannot be recommended until some users take the time to evaluate them (Good et al., 1999).

3.3 A System Framework

Figure 3.1 shows a framework that presents all components of a recommender system and how they work together to form recommendations (Zhang and Im, 2002)

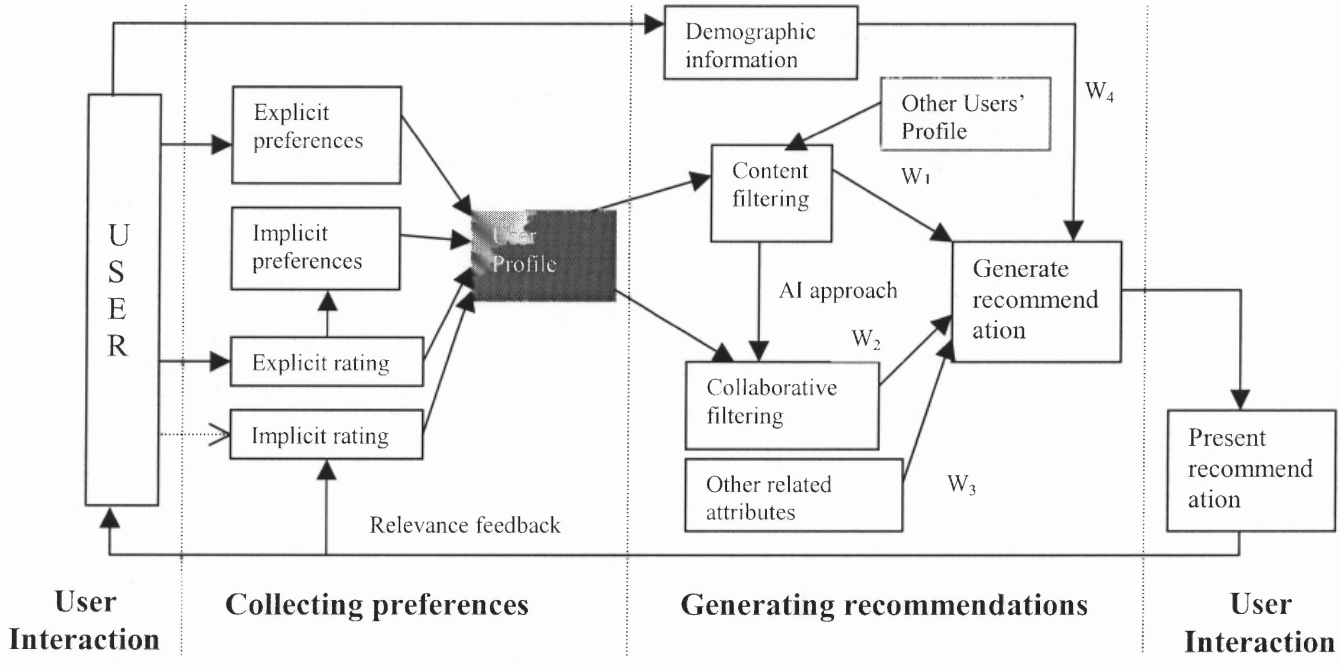

Figure 3.1 System Framework for A Recommender System

From the above framework, the whole system can be divided into three parts: user interaction, collecting preferences and generating recommendations.

User input is the direct and initial source for a system to know a user's preferences. All sources of preference need direct or indirect user input. Implicit preference is the information that indirectly represents user's preferences, such as user's browsing behavior or user's purchase history in the system.

From the user input and user behavior, four types of information are extracted – explicit preference, implicit preference, explicit rating, and implicit rating. These sources are either combined or directly generate the user's preference. The reaction of the user towards the recommended items serves as relevance feedback, which can be used to better understand the user's preference.

Based on the preferences collected, content filtering will generate a list of recommendations based on matching of user's preference and items' content. Collaborative filtering will first find like-minded neighbors for each user by calculating the similarities between the ratings provided by users. After finding neighbors of a user, collaborative filtering systems will generate a list of recommendations based on those neighbors' ratings. Besides recommendations provided by content filtering and collaborative filtering method, other attributes, such as demographic information, and expert judgment, can also contribute to generate recommendations. The recommendations mentioned above can be used separately as the final recommendations or they can be integrated into a mathematical model to generate the final recommendations for the user.

After recommendations are generated, they will need to be presented to users. Presentation strategy (including what to present, how to present, and when to present), will influence the users' perceptions and satisfaction toward the system.

In the following sub-section, the three parts of the framework are discussed separately. Pervious studies and examples of different systems are reviewed which will be the basis for the discussions on research issues in the next section.

3.3.1 User Interaction

The user interaction part includes the user's input and recommendation presentation. For systems that use explicit preference and explicit rating, user input for preference or ratings are vital to the system, since they serve as the source for recommendation generation. This is a very tedious part because users who want to use the system usually do not want to spend time and effort entering their interests or ratings on the items they

know. The amount of effort for signing up and entering ratings will increase the resistance to using the system. The accuracy of the final recommendations will also be affected if users do not provide their information or do not provide real information. Users' input of ratings also depends on the rating scale, whether it is clear and distinctive. The study by Swearingen and Sinha (Swearingen and Sinha, 2001) showed that the users' impatience seemed to have less to do with the absolute number of ratings and more to do with how the information was displayed (e.g. information about item being rated, rating scales for input items).

The other part of user interaction is the presentation of generated recommendations. Systems can display the recommendations with only estimated ratings, like Movielens, GroupLens, or they can display ratings with additional information. Swearingen & Sinha's (2001) study showed that users perceived higher usefulness of the system when providing descriptions about recommended items. Herlocker et al. (2000) compared display with or without the explanations on how recommendations are generated. Their study's result demonstrated that most users valued the explanations and would like to see the explanation features in their Automatic Collaborative Filtering (ACF) system (86% of survey respondents). The Knowledge Pump system was designed as a recommender system for organizations. Each recommended item consists of a link to the item, the predicated score for the user, a list of names of the users who reviewed it and links to their comments (Glance et al., 1998). The information about other users who reviewed the item allows users to track who is active and knowledgeable in their community as well as in return providing them a way to build their own reputations.

There have been several other systems that tried to combine additional information with recommendations. The Tapestry system provides annotation together with the messages to show recommendations (Goldberg et al., 1992). The Pointer system contains hypertext links to the source documents as well as contextual information to help recipients determine the interests and relevance of the documents prior to accessing them (Maltz and Ehrlich, 1995). Displaying social navigation information together with recommended items, it will help people to be aware of other people in the same space, and also helps users to follow traces from other users (e.g. EFOL system, CoWeb system, Footprints system, etc)

There are not many studies of user interaction with recommender systems. This area is promising and has a lot of important research issues to be explored in the future, which will be discussed later in the research issues section.

3.3.2 Collecting Preferences

Collecting user preferences is one very important method for generating accurate recommendations. User preferences determine both matching of the items in content filtering and matching of similar user groups in collaborative filtering.

There are many studies on how to collect preference information. Early systems, such as GroupLens and Fab (Balabanovic and Shoham, 1997) use explicit ratings for preferences; ReferralWeb (Kautz et al., 1997), PHOAKS (Terveen et al., 1997) and Siteseer (Rucker and Polano, 1997) use mining technologies to get preference information from public data sources such as Usenet postings or existing bookmark folders (Resnick and Varian, 1997). The RAAP (Research Assistant Agent Project) system asks users to select their research area when they register. This information is

used as the initial user profile to match the items in the database to give recommendations. This user profile is modified each time the user rejects, accepts or reviews the recommended items. User preference (profile) changes with user behavior to match their interest more accurately (Delgado and Ishi, 1999).

A system called GroupMark is totally based on implicit information to collect user preferences; it does not need users' direct input. The GroupMark system is a system to recommend bookmarks to users. They use users' existing bookmarks as the interpretation of their preferences to give recommendations (Pemberton et al., 2000). Another study by Ahmad and his colleagues built user profiles by collecting user access patterns (Ahmad and Ahmad, 1999). They built an autonomous agent to learn users' preferences by analyzing their accessing pattern to web pages. There are also systems that generate preferences based on the user's personal history (Terveen et al., 2002). In the MOVIES2GO system, voting theory was used to help multiple individuals with conflicting preferences arrive at an acceptable compromise by collecting preferences in multiple dimensions (Mukjerjee et al., 2001).

3.3.3 Generating Recommendations

After getting user preferences, these preferences are sent to content filtering or collaborative filtering systems as the input for recommendation generation. As users' perceived usefulness of the system correlates most highly with % of good and % of useful recommendations (Swearingen and Sinha, 2001), it is very important to choose algorithms to generate accurate recommendations.

Content based IF systems calculate similarities between a user's preferences and document content. Then they generate recommendations based on these similarities.

Information filtering methods are very close to information retrieval technology in calculating similarity between user profile and the data pool. Algorithms often used include vector space model and inference net model. The vector space model puts both user profiles and documents as vectors in a multi-dimensional space and calculates the similarity between vectors (Salton and McGill, 1983; Tak and Garcia-Molina, 1999). The Inference net model uses a probabilistic model based on Probability Based Ranking Principle (Robertson, 1977) to calculate similarities (Belkin and Croft, 1992).

Collaborative filtering systems first generate a neighbor group for a particular user by calculating similarities of users based on their ratings. Then they generate recommendations based on the rating of the neighbor group. There are different ways to calculate the similarity in collaborative filtering methods, and past studies have explored different algorithms and compared their results (Breese et al., 1998; Herlocker et al., 1999; Shardanand and Maes, 1995). Some researchers are working on the mathematical model for generating recommendations. There are Baysian networks approaches (Breese et al., 1998); dimensionality reduction (Goldberg et al., 2000; Sarwar, 2000); clustering techniques (Ungar and Foster, 1998b) and horting technique (this new algorithm is graph-theoretic, based on the twin new concept of horting and predictability) (Aggarwal et al., 1999); a hybrid memory- and model-based approach (Pennock, 2000), factor analysis with concern of privacy (Canny, 2002). Researchers in this field are trying to explore new mathematical models to calculate similarities in order to generate more accurate recommendations. As both content filtering and collaborative filtering have their own drawbacks when used alone, recent applications are trying to combine these two technologies. Among the systems mentioned above, RAAP, PHOAKS, and GroupMark

all combine content filtering and collaborative filtering, and Referral Web is a system combining social networks and collaborative filtering (Kautz et al., 1997). With the development of AI techniques, agent approaches and machine learning are now being broadly used in recommender systems.

The GroupLens project implemented agents to help overcome the problems in collaborative filtering. They built several filter bots based on the content of the Usenet messages and combined the results given by filter bots and collaborative filtering to generate final recommendations (Sarwar, 1998). Researchers are also trying to apply other theories to help generate recommendations. For example, Decision Theory has been tried in the DIVA project. DIVA represents user preferences using pairwise comparisons among items rather than ratings (Nguyen and Haddawy, 1998).

Besides user preferences, researchers are trying to take into account other attributes that might influence recommendation results. A previous study has shown that the accuracy of collaborative filtering systems is affected by domain, user characteristics, and purpose of use of the users (Im and Hars, 2001). Attributes of recommended items and the relationship between person and items have been used to help improve the effectiveness and efficiency of recommender systems (Schafer et al., 2001). Ansari and others built a Bayesian preference model that allows statistical integration of five types of information useful for making recommendations: a person's expressed preferences, preferences of other consumers, expert evaluations, item characteristics, and individual characteristics (Ansari et al., 2000). This model performed well in generating recommendations.

3.4 Summary

There are a lot of research issues and unsolved questions for the recommender system based on the three parts explained above. In this research, a recommender system is chosen to help fulfill individual information needs for community members, and also increase community information awareness. A recommender system will be built to provide better knowledge transfer among community members and also help them find their possible research partners. The detailed design will be discussed in the system design chapter. The next chapter will be focusing on the theoretical context of this research.

CHAPTER 4

THEORETICAL CONCEPTS

Online engagement is the key factor for judging whether an online research community is successful or not. Here, engagement refers to both social engagement and also information seeking activities.

As mentioned in the previous section, one reason for people to join the online community is to meet other people, to seek help, support, friendship, love, etc. In other words, they are driven to develop social relationships with other people inside the community. Interpersonal communication can also create structures inside the community (Weick, 1979). Thus, it is very critical for an online community to help its members establish interpersonal relationships. This is so crucial that it directly affects the satisfaction of its members, and satisfaction is a potent predictor of continued participation in the online community; from the system aspect, it is also a crucial factor for the continued use of the system (Kiesler et al., 1985; Rockart and DeLong, 1988).

Relating is one very important part of our daily lives, it is defined as:

> " One-to-one, interactive process, involving first recognition, then a developing understanding of the other person, self-disclosure, accompanying arousal and affect, and the carrying forward of this to future encounters." (Whitworth, 1997, p88)

As Millar and Rogers state: "*Functionally, the communication process is largely a negotiation process whereby persons reciprocally define their relationships and themselves*" (Millar and Rogers, 1976, p88). Relational communication consists of the messages and message dimensions people use to define or redefine relationships (Millar

and Rogers, 1976; Parks, 1977). Relational messages are transmitted through verbal or nonverbal expressions that indicate how people regard each other, regard their relationship or regard themselves within the context of the relationship (Burgoon and Hale, 1984; Burgoon and Saine, 1978). Walther extended this definition by adding linguistic, and textual manipulations which are more from a CMC perspective (Walther, 1992). The dimensions of relational messages are inherited from interpersonal exchange. Burgoon and Hale proposed seven dimensions for relational communication – *similarity/depth, receptivity/trust, immediacy/affection, dominance, equality, composure, and formality* (Burgoon and Hale, 1987).

On the other hand, for a research community, knowledge seeking is one of the primary goals that members want to achieve. The advert of the Internet provides the best opportunity to share information ever, it provides faster and more information than any other media. Anyone can put information online, and everyone can seek any information online. However, this advantage also causes the biggest problem for information seeking online – information overload. The Internet as an information repository doesn't have its own structure for organizing all the information. Although search engines may help people to filter out some unrelated information, it still returns thousands of related documents for many keyword searches.

One of the purposes for online community is to gather and share all information that relates to the community's general interests. Thus, how well the information is presented, how well the information has been organized, and the quality of the information will be very important factors for an online community to retain its members.

In this section, literature from computer-mediated communication (CMC), research on behavior in online communities, and literature from individual technology acceptance, have been reviewed from the media aspect, community aspect and individual aspect to form a theoretical framework identifying factors that influence online engagement.

4.1 Media Characteristics

There have been two themes of research on the effectiveness of communication media. One is the media-dependent perspective, which focuses on the fact that different media differ mechanically in their bandwidth, the communication cues, and feedback speed. The other perspective is the social construction perspective, which focuses on how social factors influence users' perceptions of the media and media choices (Yoo and Alavi, 2001).

A debated issue in CMC literature is whether CMC is a lean medium that is not appropriate for a complicated communication task, like forming relationships. Some research argues that people communicating online perceive less communication context and interpret less meaning from the conversation. "*Without nonverbal tools, a sender cannot easily alter the mood of a message, communicate a sense of individuality, or exercise dominance or charisma... Communicators feel a greater sense of anonymity and detect less individuality in others*" (Kiesler, 1986, p48). The media with a small number of cues and communication channels tend to "depersonalize" the communication interaction (Rice, 1984, Siegel et al., 1996, Sproull and Kiesler, 1986). Culnan and Markus (1987) also suggest that the mechanical characteristics of the media, like

bandwidth and number of communication cues will alter interpersonal variables. Studies with this theme belong to media-dependent perspective (Yoo and Alavi, 2001).

In other studies, people were found to establish relationships via Internet Relay Chat (IRC) (Reid, 1991), commercial online services (Van Gelder, 1985), virtual reality systems (Reid, 1995), and Usenet newsgroups (Parks and Floyd, 1996) Korzenny (1978) proposed that communication through electronic media creates a feeling of greater propinquity with others, regardless of their actual geographic dispersion. Such "electronic propinquity" might be expected to foster friendships, as traditional propinquity is known to do (see also Walther, 1992). All these studies are classified in the social construction perspective, which is inconsistent with the media-dependent perspective. There have been a number of studies that support this perspective (e.g. Chidambaram, 1996, Carlson and Zmud, 1999, Walther, 1995, McGrath et al., 1993). Chidambaram found that the perceived social presence of group support systems increased as group members spent more time with their communication partners, in spite of the initial text-based technology constraints. Carlson and Zmud found that higher familiarity among group members would expand the perceived richness of electronic mail.

In the following, two theories will be discussed in more detail from the above two perspectives which are applicable for online community research. They are Social Presence Theory (media-dependent perspective) and Social Information Processing Theory (Social Construction Perspective).

4.1.1 Social Presence Theory

Social Presence (Short, Williams and Christie, 1976) theory argues that different media foster different levels of perceived intimacy and immediacy. Social presence is the degree

to which a medium is perceived as conveying the presence of communicating participants. Face-to-face communication provides the highest level of social presence and asynchronous writing provides the lowest level of social presence (Rice, 1993; Short et al., 1976). Short and his colleagues hypothesized that: "communications media vary in their degree of Social Presence, and that these variations are important in determining the way individuals interact" (Short et al., p65). In this sense, social presence is a term to describe the ability of a communications medium to give a group member the feeling of the presence of other members. Getting to know someone, maintaining friendly relations/staying in touch, exchanging information, exchange opinions, and generating ideas are all listed in a set of activities that are likely to be affected by differences in a medium's social presence (Rice, 1993). Social presence depends on verbal, and nonverbal cues and communication context. It "varies between different media, ...affects the nature of interaction... interacts with the purpose of the interaction to influence the medium chosen by the individual who wishes to communicate" (Short et al., 1976, p65). It is more difficult to carry out relational communication in an online environment than face to face situation because of the lack of non-verbal cues.

Social presence, or the salience of another person in communication, is said to depend on the number of channels or cues available within a medium; the fewer the channels, the less attention paid by the user to the presence of other participants (Walther, 1994). Because there is a sparcity of verbal and nonverbal channels in CMC, the social presence of CMC should to be lower than other media, like face-to-face communication; and messages are assumed to be more impersonal. Social presence theory has been used to explain task orientation and impersonality in CMC research (Culnan and Markus, 1987;

Hiltz et al., 1986; Kiesler et al., 1984; Rice, 1984; Siegel et al., 1986; Williams and Rice, 1983).

The fundamental difference between the real world and the virtual world is the lack of physical body. Consequently, other social cues that are related to body are lost, such as individual recognition, gestures, facial emotions, and tone. Understanding how these social cues help in communication will help to assess the impact, both positive and negative, in cyberspace with their absence. In some cases, this absence is beneficial as it allows other modes of communication, while in others; the communication seems to be handicapped. In order to increase the proximity in CMC environment, people tried to invent different tools to overcome the lack of social cues. In 1980, Carey reported electronic paralanguage, which is the use of intentional misspelling (like "weeeeeeellll"), lexical surrogates for vocal cues (Carey, 1980). There are also symbols which represent emotions, like :) represents a smiley face; recently this has been a standard to be recognized in browser or word applications; when you type :), it will automatically change to ☺. Emotion icons and paralanguage have been widely used in Chatrooms, instant messages, emails. With more advanced computer technology and higher bandwidth, it is also possible to reconstruct those social cues functions in other forms online, such as emotion icon, photos, audio support and video support.

4.1.2 Social Information Processing Theory

Social Information Processing Theory was proposed by (Walther, 1992), This theory asserts that:

> "communicators using any medium experience similar needs for uncertainty reduction and affinity, and to meet these needs CMC users will adapt their

> linguistic and textual behaviors to the solicitation and presentation of socially revealing, relational behavior" (Walther, 1994, p465)

According to this theory, the critical difference between FtF and CMC is rate, not capacity. It states that social identity and relational cues can be transmitted by plain text, but this occurs at a slower rate than would occur in a richer channel such as voice, or face-to-face meeting. Thus, CMC relations should require more time for social information to get through (summarization from Withworth, 1997, p88). Walther predicted that:

> "over time, computer mediation should have very limited effects on relational communication, as users process the social information exchanges via CMC." (Walther, 1992, p80)

Walther also used social information processing theory to explain the contradictory findings in CMC literature. The early experimental studies found CMC to be more impersonal and task oriented (Connolly et al., 1990; Dubrovsky et al., 1991; Hiltz, 1975; Hiltz et al., 1978; Hiltz et al., 1986; Hiltz et al., 1989; Kiesler et al., 1984; Kiesler et al., 1985; Rice, 1984; Siegel et al., 1986; Sproull and Kiesler, 1986), while field studies produced more supportive or more varied communication among CMC interactants (Feldman, 1987; Ord, 1989; Weedman, 1991) Among these studies, there are case studies that found that CMC supports colleagueship formation (Johansen et al., 1988); it provides ongoing social and professional support (Beals, 1991). People were found to establish friendships or even romances in Internet Relay Chat (IRC) (Reid, 1995); in commercial online services (Van Gelder, 1985); in virtual reality systems (Reid, 1995); and also in asynchronous Usenet newsgroups (Parks and Floyd, 1996). A large amount of supporting information exchanges were found in varieties of online social

support communities using bulletin board systems or email systems (Egdorf and Rahoi, 1994; Preece and Ghozati, 1998; Rheingold, 1993).

Field experiments showed that CMC task orientation is affected by factors like task complexity, task interdependence, environmental uncertainty, the need for communication across locations and organizational culture (Hiemstra, 1982; Steinfield, 1986). All the results from field studies suggest that the lack of nonverbal cues is not a satisfactory predictor of the interpersonal nature of CMC social interaction (Walther, 1994).

According to Walther, time frame is one of the differences between experimental studies and field studies. In lab experiment studies, CMC and FTF groups had very limited time to finish their task, ranging from 20 to 90 minutes, while in field studies, the time interval for interaction is not limited. Social information processing theory predicts that the exchange of social information in CMC may be slower than FtF but will eventually get to the same point, which explains why with a limited time period, CMC tends to be more impersonal and task oriented, while as time goes by, it shows more social exchange.

Imposing equal time limits on FtF and CMC groups may have an unequal effect. It may interrupt slower moving CMC interaction before sufficient messages are exchanged for interpersonal effects to occur; for FtF participants, as they process social information at a faster rate via multiple channels, the interpersonal effects won't be hindered. Other studies, not from the social information processing perspective, also found that CMC may change relationally over time (Hesse et al., 1988; Johansen et al., 1988; Kerr and Hiltz, 1982; Rice and Love, 1987). Walther and his colleagues did a

meta-analysis with over 350 articles and book chapters to test the social information processing theory. The findings of the study showed that the degree of socioemotional communication in CMC is greater where interaction was not limited and the ratio of CMC to FtF on socially-oriented communication is more positive when interaction is unrestricted. However, the results did not support that when interaction time was not restricted, the extent of antisocial communication in CMC is less than when it is restricted; also there is no difference between CMC and FtF on antisocial communication regardless of time (Walther et al., 1994).

Social Information Processing theory tells us when studying social activities in a CMC environment, a short-term experiment (1-3 hours) is seldom as effective as a longitudinal study.

4.2 Community Characteristics

In our research, community refers to a social group connected by computer networks (Wellman, 1997). Community size and communication activities are very important factors for attracting people to join the community and sustain their membership with the community.

4.2.1 Critical Mass Theory

In order for a community to survive, it needs a certain number of active members to keep it attractive and sustained. If there are too many participants, the community will become chaotic and people will tend to leave; if there are too few people, there won't be sufficient activities to maintain people's motivation to stay in the community (Preece, 2000). A community thrives only if there are sufficient people and enough activity to

make it attractive and worthwhile (Markus, 1987, 1990; Morris and Ogan, 1996; Rice, 1994).

The number of people an online community needs to attract its critical mass (Preece, 2000). In 1985, Oliver and his colleagues used the critical mass theory to explain the extent and effectiveness of group actions in creating a public good. They define critical mass as a subgroup of a community that is motivated enough to make significant contributions toward the establishment of a public good (Oliver et al., 1985). Critical mass theory originally was used to explain the diffusion of interactive media, such as phone, electronic mail, etc (Markus, 1987). She argues that all interactive media share two characteristics:

- Universal access: the ability to reach all members in a community through the media
- Interdependence: the use of the media involves reciprocal interdependence, in which later users are influenced by earlier users as well as vice versa.

Because of these characteristics, interactive media are extremely vulnerable to start-up problems and discontinuity (Markus, 1987). From an individual point of view, when thinking about whether or not to adopt a new medium, it is very likely that a person won't use it unless there are already a sizeable number of his or her communication partners using it. Critical Mass Theory can also be used to explain the diffusion of new technology. For every new technology (such as video conferencing, mobile devices), among its potential users, it is necessary that there be initial users who feel the system has sufficient worth, or who can afford the startup costs. These initial users then decrease the costs and increase the value of adoption for later users (Markus, 1990). Researchers have

referred to enough initial users to stimulate rapid later adoption by others and sustain interaction as "critical mass" (Bair and Mancuso, 1985; Culnan and Bair, 1983; Hiltz, 1984; Hiltz and Turoff, 1978; Rogers, 1986; Rice, 1982, 1990; Uhlig et al., 1979).

Granovetter (1973) proposed a similar approach - threshold models to explain such phenomena as rioting and diffusion of innovation. In threshold models, an individual's decision on whether or not to follow the innovation depends on the distribution of individual thresholds, which is defined as the number of other people who must be doing the activity before a given individual joins in. There must be enough active members who sign on regularly to generate the motivations for all members to sign on frequently and to contribute to the system. The number of the active members is also referred to as "effective size" (Hiltz 1984; Hiltz and Turoff, 1978).

Palme (1981) also noticed that the effective size of a group is very important, but what this effective size is varies with the type of task, characteristics of technology, etc. The critical mass for different online communities also varies. It depends on both characteristics of the community and the system that supports the community. In group CMC, Turoff and Hiltz predicted that "computer conferencing might enable large, diverse groups to tap the resources of all the members without losing the ability to communicate freely or experiencing other negative effects of large group size." (Hiltz and Turoff, 1978). Among the applications that support online community, blogger, Usenet, and listservs are likely to hold large size communities; Conferencing systems tend to hold middle sized communities; while Chats, MUDs, and MOOs are more likely to hold small sized communities. The larger the community, the more need for tools and structures.

Although critical mass is a good term to help understand activities in online communities, it is very difficult to measure. There is some research on mass activities in Usenet newsgroups (Jones, 2000; Smith and Kollock, 1999; Whittaker et al., 1997) and listserv communities (Nonnecke, 2000), but still this is a field that needs to be explored. The Internet is a public network, it is open to all users, and it has the potential for universal access. This is also true with the online communities, there is no special rule for forcing people to stick with the community, and also there is no other rule to exclude users. The number of people in a community is very hard to control. It might often go from initial members (small group of people) -> evolving -> reach critical mass -> overwhelming -> members' leaving -> then go either balanced or die. For developers of the community, they might need to consider the community's characteristics to choose the appropriate system that is able to support the critical mass. They will also use different techniques to stimulate community activities to attract more people to participate and reach the critical mass.

With a very small number of people, a community cannot really be established. So, does this mean "the more the better"? Let's discuss it from common ground theory from the communication literature.

Common ground is defined as the mutual knowledge, beliefs and suppositions shared by individuals (Clark, 1992). It provides the basis for interpersonal communication (Clark and Brennan, 1991). In other words, common ground is the things people know about what is known by the person they are talking to. Common ground comes from the common culture and social history, public history of the interaction, current public state of the interaction, and so on. It accumulates as the interaction

continues (Clark and Brennan, 1991). When there are too many people in a community, it is harder to establish the common ground that supports the interaction. Whittaker et al.'s study also shows that it is easier to establish common ground in communities with small size and familiarity of the users (Whittaker et al., 1998).

4.2.2 A Resource-Based Model for Sustainable Social Structure

In order to make online communities thrive, technology itself is not enough. Information technology can only provide infrastructures to support social activities (Hagel and Armstrong, 1998, Hof et al., 1997), it cannot guarantee that individuals will willing to join and participate. A community systems' ability to support social activities is strongly influenced by its ability to support establish sustainable online social structures. Social structures are sustainable when the benefit that they provide can overweigh the cost for an individual to join the social structure (Moreland and Levine, 1982). Butler (2001) proposed a resource-based model to explain how community size and communication activity may influence building sustainable social structures. Members contribute their time, energy, and other resources to visit the community; it must provide them with higher benefits than the resources they spend to attract and retain them. These benefits include valuable information, an audience who can read their postings, etc. His model is shown in Figure 4.1.

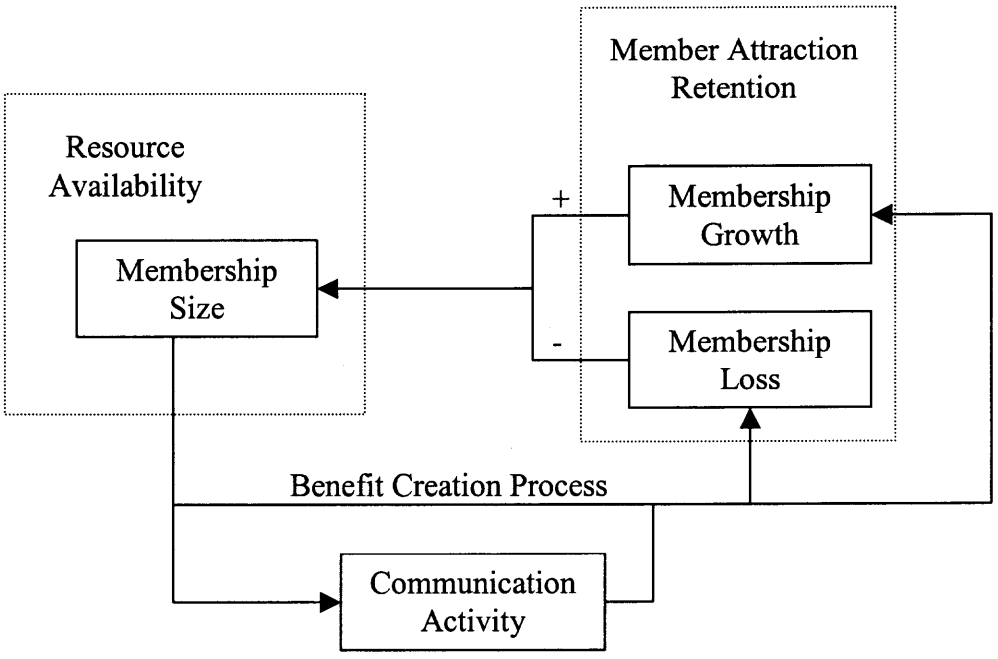

Figure 4.1 A Resource-Based Model of Sustainable Social structures (Butler 2001)

In this model, resource availability is measured by the size of a structure's membership as members are the sources of information, time, energy, and social and emotional support. However, having just sources of resources available is not enough to generate benefit. Some channels to transfer from resources to benefit are needed, and this transformation is made by communication activities. In larger groups, it is more likely that someone in the group has the needed knowledge, or has the ability to provide support. Thus, larger groups will be more able to provide valuable benefits to members as a result of possible communication activities, and hence be sustainable over time (Haveman, 1993, Rafaeli and LaRose, 1993). However, when member size becomes too large, it generates obstacles to converting all resources into valuable benefits. In larger face-to-face groups, individuals have few opportunities to participate and less time to talk. Online

larger groups are more likely to have information overload problems and also it is also harder to form personal relationships as there are too many members.

Jones and Rafaeli (2000)'s model of virtual public user populations shows that increase in communication activities will be less than increase in user population because of the negative influences that size cause such as free-riding, social loafing and information overload problems. Studies of Usenet newsgroups and email lists show a significant relationship between community size and member attitude variables and also a link between size and communication activity (Jones and Rafaeli, 2000, 2002; Butler 2001, Schoberth, Preece and Heinzl, 2003).

4.2.3 Other Community Characteristics

Research also shows that usability of the community system, the volume and quality of information (this will include communication activities), and the timeliness of the community content are factors that affect members' engagement in the community (Preece, 2000, Nonnecke and Preece, 2000). As people in an online community share similar interests, one obvious purpose for them to come to the community is to seek for valuable information and exchange information on their special interests. In a research community, information is more important than interaction. Timely information is also very important as it shows the current trend and also meets the users' immediate needs. If the community cannot provide both good quality and volume of information, it will drive its members away.

4.3 Individual Characteristics

Another component that contributes to users' engagement online is their individual characteristics. People come to an online community with different needs, different

constraints and different personalities. All these factors will influence their behavior online.

4.3.1 Technology Acceptance Model

In order to understand why individuals choose to join and stay in an online community, Theories and models in the technology acceptance literature are reviewed to understand individual factors. From synthesis of the literature, Venkatesh and others found eight theories and models that are related to individual acceptance (Venkatesh et al., 2003). They are Reasoned Action (TRA) (Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1975), Technology Acceptance Model (TAM) and TAM2 (Davis, 1989; Venkatesh and Davis, 2000), Motivational Model (MM) (Vallerand, 1997; Davis et al., 1992), Theory of Planned Behavior (TPB) (Taylor and Todd, 1995), Model of PC Utilization (MPCU) (Thompson et al., 1991; Plouffe et al., 2001), Innovation Diffusion Theory (IDT) (Rogers, 1995; Tornatzky and Klein, 1982; Moore and Benbaset, 1991, 1996; Agarwal and Prasad, 1997, 1998 Karahanna et al., 1999), and Social Cognitive Theory (SCT) (Bandura, 1986; Compeau and Higgins, 1995; Compeau et al., 1999). Among them, the Technology Acceptance Model (TAM) (Davis, 1989, Davis et al., 1989) is the most widely applied model. TAM was developed from the social psychology Theory of Reasoned Action (TRA) (Ajzen and Fishbein 1980, Fishbein and Ajzen 1975). It theorizes that an individual's behavioral intention to continue use a system is determined by two beliefs: perceived usefulness and perceived ease of use. Perceived usefulness is defined as the extent to which a person believes that using the system will enhance his or her job performance. Perceived ease of use is the extent to which a person believes that using the system will be free of effort. The original TAM is shown in Figure 4.2.

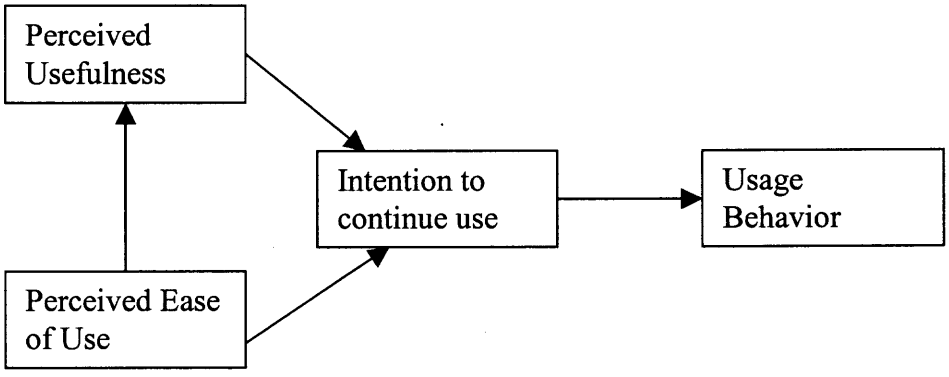

Figure 4.2 Technology Acceptance Model (Davis et al. 1989)

TAM is a very robust model in understanding why users intend to use information technology. Many empirical studies have found that TAM consistently explains about 40% of the variance in usage intentions and behavior (Adams et al., 1992; Chin and Gopal, 1993; Chin and Todd, 1995; Davis, 1993; Davis and Venkatesh, 1996; Szajna, 1994; Taylor and Todd, 1995; Venkatesh, 1999; Venkatesh and Davis, 1996; Venkatesh and Morris, 2000). However, the initial TAM did not include the determinants of both perceived ease of use and perceived usefulness. Later on, Vankatesh and Davis did extend the initial TAM and add both determinants of perceived ease of use (Venkatesh and Davis, 1996; Venkatesh, 2000) and perceived usefulness (Venkatesh and Davis, 2000). The Figure 4.3 is a complete framework which includes TAM, TAM2, and anchoring and adjustment for perceived ease of use.

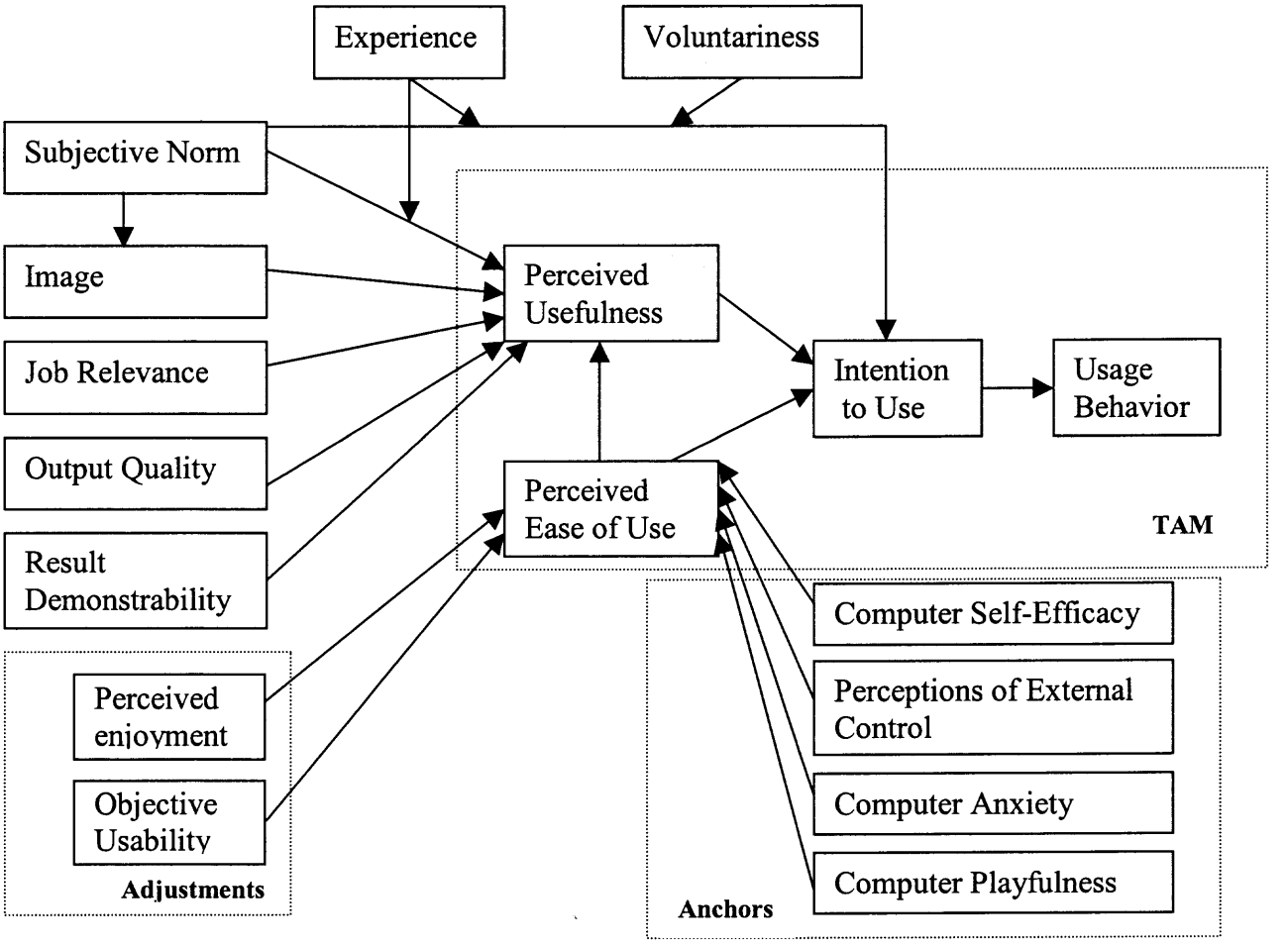

Figure 4.3 Extended TAM Model

4.3.2 Other Individual Factors

Mason (1999) reported that lurkers make up the majority of members in online groups. Lurkers are members who rarely post, but are known to read the community's content regularly. Although the activities of lurkers are a legitimate form of participation, and being a lurker is a very important transition mechanism for a novice to lean about novel topics (Whittaker et al., 1998), more people are expected to actively participate from the point of view of community development. Studies on lurkers will help us to understand what factors hinder people's active participation online.

Nonnecke and Preece (2001) did a study to explore reasons why lurkers lurk, using structured interviews. Their results show that the most mentioned reasons for lurking include an insistence upon being anonymous, a concern about privacy and safety, work-related constraints, shyness about public posting, and time restrictions. Utz' (2000) longitudinal study of the factors that influence development of friendships in MUD environments found a significant correlation between skepticism and making friends, and a moderate correlation between sociability and making friends. In other studies, researchers also found that people's expectations will influence their behavior online (Hiltz, 1984; Utz, 2000). For all the above issues can be summarized into one factor – "attitude toward the online community". Attitude toward behavior is one construct from TRA, it is "an individual's positive or negative feelings (evaluative affect) about performing the target behavior" (Fishbein and Ajzen, 1975, p. 216). Although these factors are individual characteristics that cannot be directly affected by technologies, the atmosphere of the community space, information that the community space provides, and the communication activity in the community may change them.

4.4 Research Framework

After combining all the factors discussed from media aspects, community aspects and individual aspects, A research framework is presented in Figure 4.4 to predict member's engagement with the online research community and their preliminary networking behavior.

From the framework, in one direction, Expectation relevancy, output quality, result demonstrability, members' attitudes, sense of community, and perceived ease of use are determinants for perceived usefulness of the community system; objective

usability, perceived enjoyment, perception of external control are determinants of perceived ease of use. Perceived usefulness and perceived ease of use in turn will change members' intention to continue use the system. This intention, the designed technology and other individual factors (e.g. gender, status, time constraint) determines members' engagement inside the community place. In the other direction, in the feedback loop, members' engagement will change community activities and volume and quality of information in the community knowledge base, which are benefit creation process. Thus, this process will influence job relevance, output quality and result demonstrability. During this process, the designed technology will influence the final results of members' perceptions on job relevance, output quality and result demonstrability.

Detailed explanations about each factor will be presented in the research questions and hypotheses chapter.

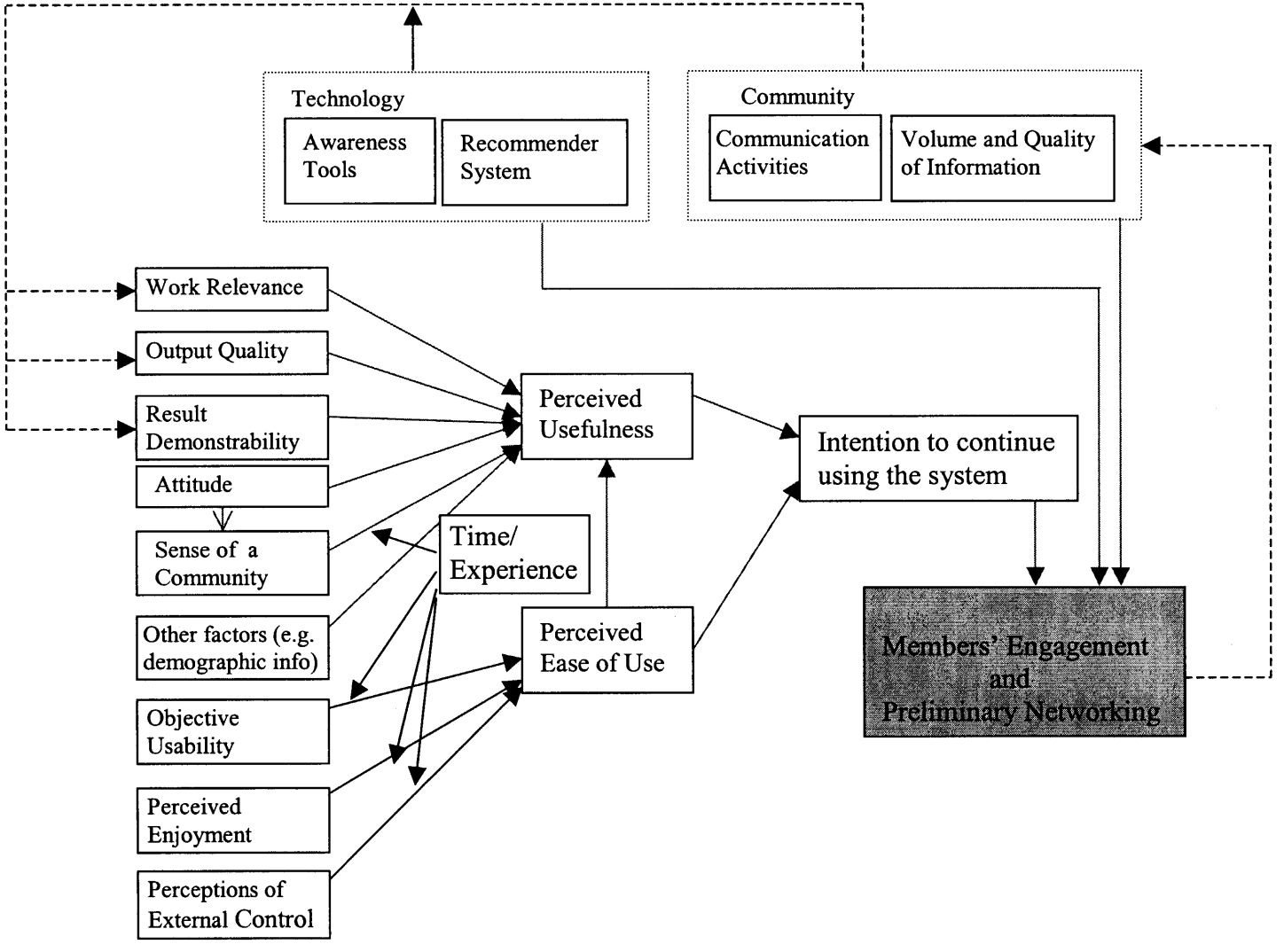

Figure 4.4 Research Framework

In the above framework, arrows in red are factors that will be measured in this study. This research presents two applications for the community system: one is awareness tools which are based on "Socially Translucent System Design"; the other is a recommender system which is based on providing valuable information on a personal basis.

4.5 Summary

Before going into more research detail, factors that were planned to be explored in this study that will contribute to people's engagement online are summarized in Table 4.1.

Table 4.1 Factors that Influence Members' Engagement in An Online Research Community

Factors	Specification in this study	Related system features
Intention to Continue Using the System	Whether the member decides to stay and continue using the community system	No
Time / Experience	The change of member's engagement and preliminary networking online should be observed over a long time	No
Communication Activities	Number of topics in community space and number of postings and people involved in the topics.	Knowledge base, discussion topics
Volume and Quality of information	Whether the community can provide valuable and timely information for every individual user	Recommender system
Perceived Usefulness	The degree to which a member believes that the community system would provide valuable benefits	Recommender system
Perceived Ease of Use	The degree to which a member believes that using the community system would be free of effort	No
Attitude	An individual's feelings about performing the target behavior inside the online community	No
Expectation relevancy	The degree to which members believe that the community can meet their expectations	No
Output Quality	How well the community system supports individuals to perform their tasks	Recommender system
Result Demonstrability	The tangibility of the results of using the community system	No

Table 4.1 Factors that Influence Members' Engagement in An Online Research Community (Continued)

Factors	Specification in this study	Related system features
Objective Usability	Allows for a comparison of systems based on the actual level of effort required to complete specific tasks	Whole system
Perceived Enjoyment	The degree to which a member enjoys to using the system	No
Perception of External Control	The required knowledge, resources, and access for the community system	No
Other Individual Factors	Gender, status, whether a member has enough time to check community activities and be involved	No

CHAPTER 5

SYSTEM DESIGN

Based on the theoretical framework in the previous section, two specific system tools are designed, developed and studied in this research. They are awareness tools and a recommender system.

5.1 Social Awareness Tools

Research in organizations has shown that physical proximity is associated with several personal perceptions. Physical proximity strongly correlates with opportunity for informal communication to happen, so does one's familiarity with his/her colleagues work, as well as liking them (Kraut et al., 1990). In an online world, where people are not physically together, it is very important for computer technology to provide a way to establish the equivalent of physical proximity. However, the cues and mechanisms used to maintain awareness of others in face-to-face situations are difficult to provide in virtual environments. Awareness is always a very important issue in Computer Supported Collaborative Work (CSCW). In the literature, awareness has been achieved through:

- Audio/Video support: such as the Media Space (Bly et al., 1993) which provided a video view of remote collaborator's activities in the office; and Portholes (Dourish and Bly, 1992), which reduced the bandwidth requirement by using occasional video snapshots instead of live video.
- Virtual reality environment: such as MUDs (Multi-user domain). As users navigate through regions in the virtual world, they will encounter others who are in the same region and they are able to interact with each other via text descriptions and commands (Curtis, 1992). More recent graphical virtual

environments added visual representations of places and people, thus making it even easier to notice and start interaction with others (Morningstar and Farmer, 1992; Jeffrey and McGrath, 2000).

- Chat-like communication tools: for example, an Instant Messaging (IM) system can provide some awareness information such as who is online now, who is busy, who is away, etc. The BABBLE project provides more awareness information about other people's activity while in the chat space (Erickson et al., 1999; Bradner et al., 1999).

However, most of the tools above are for synchronous communication. There are not many awareness tools designed for asynchronous communication. For an online research community, in order to help members interact with each other, first of all, the system should help people to recognize other members in the community. Kollock's principle of building online community also emphasizes the need for information about the people and their behavior (Kollock, 1999). Systems that provide perceptually-based social cues which afford awareness and accountability are called "Socially Translucent Systems" (Erickson et al., 1999). In online communities, because of the lack of social cues, the degree of social presence tends to be low. Thus, a system with "Socially Translucent" features should help users to be aware of other users in the system and encourage relationship development among them. For this purpose, the following features are designed.

- Member directory: contains all the profiles of members of the community. A profiles includes a member's picture, contact information, research interests, and member's participation in the community. Photographs have been shown to increase

affectionality among people in the CMC environment (Walther, 2001). They help members get to know each other and help reduce initial uncertainty about others. Figure 5.1 shows the screenshot of this feature

Figure 5.1 Screenshot of People Directory

- Provide visibility/awareness of people and their activities. The visibility of others in the community will foster familiarity among members, help build self-esteem and responsibility (Erickson et al., 1999, Constant et al., 1997, Girgensohn and Lee, 2002) and also help to increase aliveness of the community (Dieberger, 1997). Whenever a

person posts in the community, her ID (login name) and email will be linked together with the postings, and the id is linked to the individual's profile, which will also show all the postings that the member contributes to the community. Figure 5.2 shows the screenshot of this feature.

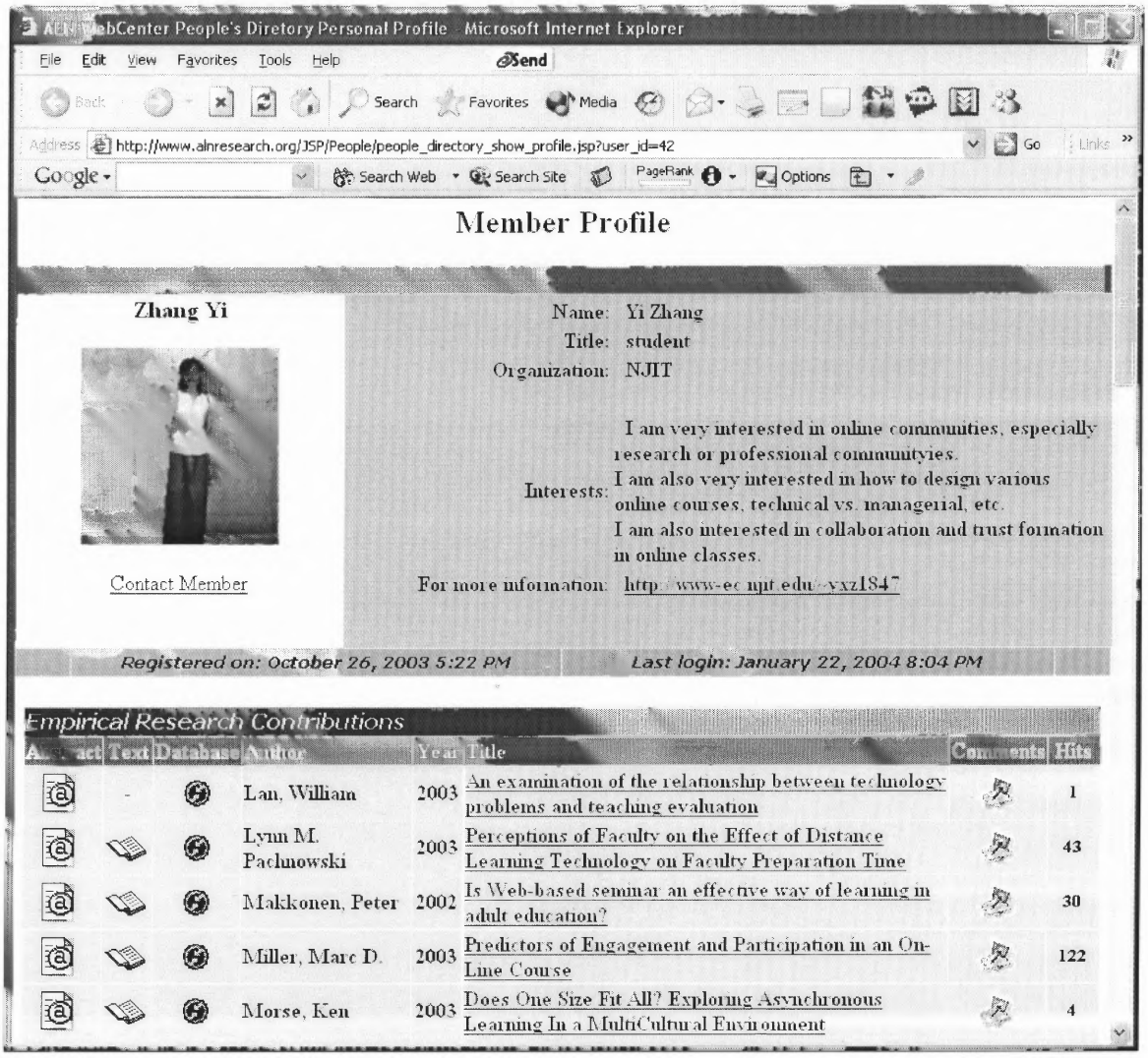

Figure 5.2 Screenshot of A Member's Profile

- In the community forum, each message is displayed together with its author, a link to the author's profile and summary of the author's activity in the forum. Figure 5.3 shows the screenshot of this feature.

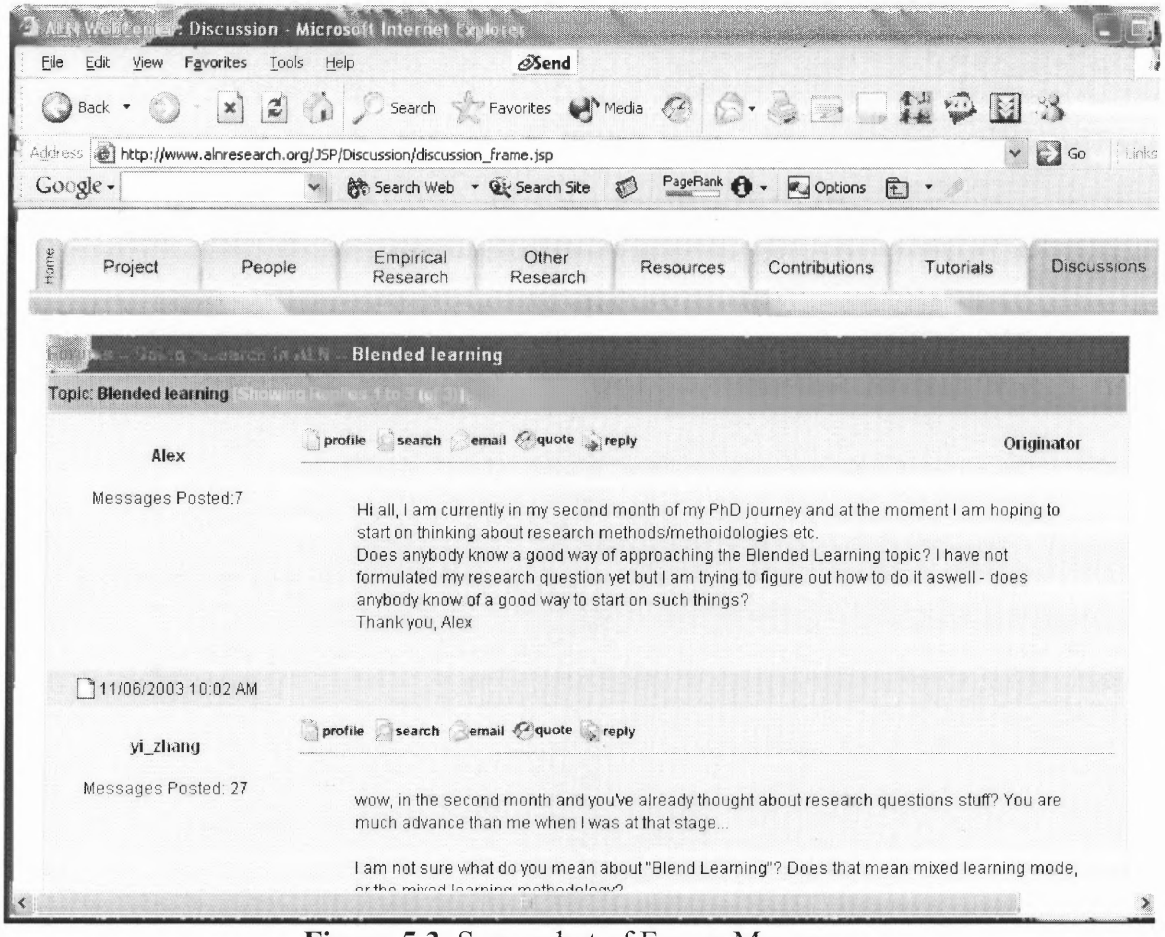

Figure 5.3 Screenshot of Forum Messages

5.2 Recommender System

In order to reduce possible information overload problems and provide personalized guidance, a recommender system which is based on both content filtering and collaborative filtering technologies will be implemented to help members access information that is most useful to them on the site. Recommender systems can be viewed as intelligent agents that can suggest artifacts of interest using stored information (e.g. user preferences, performance data, artifact characteristics, and cost) of a given domain of artifacts (e.g. books, music, movies, internet resources, or research papers) (Ramakrishnan, 1997; Resnick and Varian, 1997; Terveen et al., 1997; Im and Hars, 2001). Content filtering and collaborative filtering are the most commonly used

technologies for developing recommender systems. Content filtering is based on personal preference and collaborative filtering is based on other like-minded people. The design for the recommender system is shown in Figure 5.4.

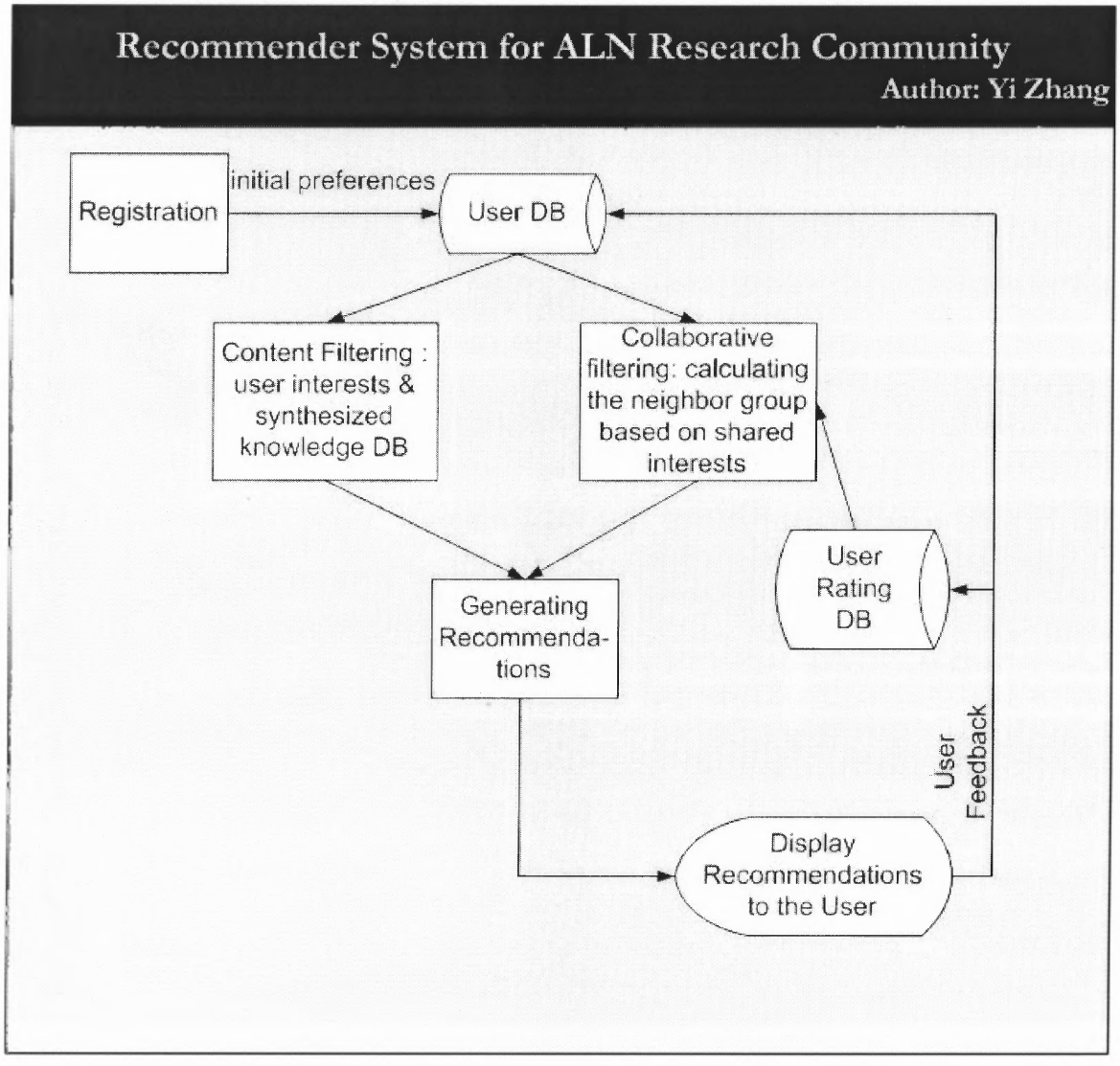

Figure5.4 Recommender System Design

User input is first collected at the time of registration. Users are required to register with the WebCenter to access its resources. In the registration process they are encouraged to provide their interests, in order to obtain personalized content. These

initially specified interests serve as the basis of initial content filtering and recommendations. Users are also encouraged to regularly update their interests.

Users are further encouraged to provide ratings for articles they read based on the "Value of Article for ALN Researchers". The rating is on a 5-point Likert scale. (Figure 5.5) This can serve as the basis of collaborative filtering discussed earlier.

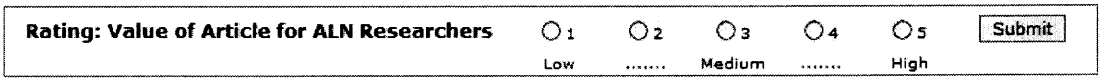

Figure 5.5 Screenshot for Rating on Articles

Because of the hesitation by users to explicitly rate articles and the resulting data sparseness problem, it was also decided to update ratings based on user behavior. These implicit ratings are determined on the basis of a visit to an article (Rating of 2 – on a 5-point scale), repeat visit (Rating of 4 – on a 5 point scale) and printing of an article (Rating of 3, on a 5-point scale).

Separate recommendations are presented based on content and collaborative filtering. These are presented to the user when they first log into the system. In addition to presenting these recommendations, a scheduler is developed which runs a routing query incorporating the user profile, the correlations (similarities) and the recommended articles which are newly added. The results are periodically to be e-mailed to the users. This is intended to keep the members updated on new content based on their interests or system-perceived interest based on the filtering technologies and which can keep them updated on the continuously updated knowledge resource. In addition to the content and collaborative filtering, articles with most visits, a history of short-term user searches and all newly added articles are also presented to the user in categorized lists. (Figure 5.6, 5.7, 5.8)

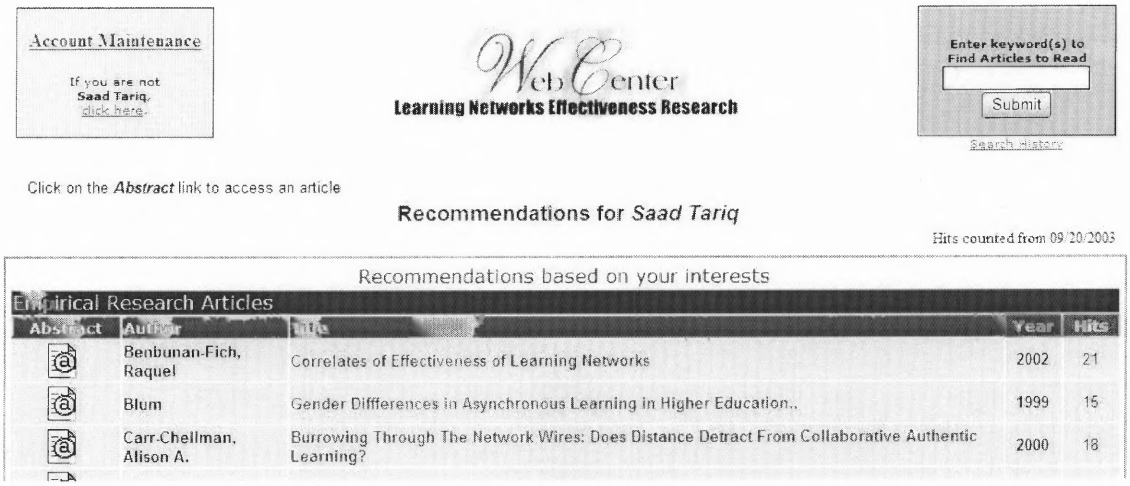

Figure 5.6 Screenshot of Recommendation Items from Content Filtering

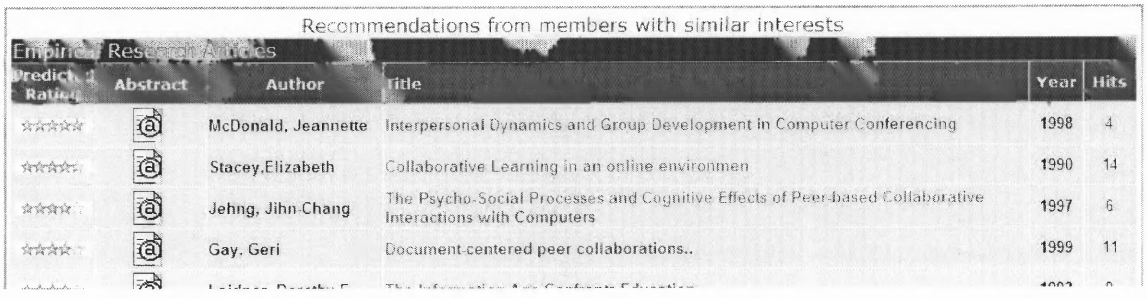

Figure 5.7 Screenshot of Recommendation Items from Collaboration Filtering

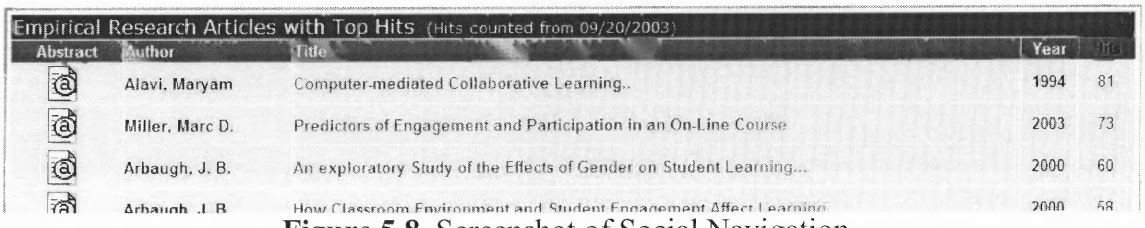

Figure 5.8 Screenshot of Social Navigation

5.3 Summery

The ALNResearch WebCenter is designed as a knowledge-based community space. It is a place to consolidate and synthesize existing knowledge in Asynchronous Learning Networks (ALN) research area, to increase the quality, quantity, and dissemination of results of research on the effectiveness of ALN. The community space is also designed to

help this research community become more cohesive. Awareness tools have been designed and developed to create proximity among members. A recommender system was designed and developed to help better serve personal information needs. The basic structure of the community system besides the awareness tools and recommender system includes:

- Empirical studies: contains contributed empirical articles in ALN research, each article includes abstract, database entry (synthesis of the article), and comments from readers. Some articles also contain a full-text link with the permission from the publisher.
- Other studies: contains contributed theses, books, qualitative research, technology related research, and other research that relates to ALN.
- Resources: contains contributed research instruments, online resources related to learning theories, online journals, evaluations, methods, and others.
- Tutorials: contains web-based tutorials of measuring learning effectiveness and stream videos on lectures of basic research methodology.
- Discussions: contains different discussion forums of ALN.
- Project: contains proposal of the ALNResearch WebCenter project and people who are involved in the project.

CHAPTER 6

RESEARCH QUESTIONS AND HYPOTHESES

In order to explore the impact of various factors on members' engagement in online research communities over time, a longitudinal field study was conducted on an online research community. The virtual space of this community is a website, which is a pull-based technology. The website is built to be a community knowledge management system. The main purpose of the community is to improve the quality, quantity, and dissemination of the knowledge.

6.1 Study Design

A longitudinal field study is proposed to explore how the technologies, community, and other individual factors influence members' engagement in an online research community. Ideally, the complete model presented in the previous section should be examined. However, due to the nature of the study and the nature of the online community, a very extensive questionnaire cannot be conducted in this situation. One also cannot ask the members in the community to fill out several questionnaires at different times. Thus, factors that relate to perception of change over time will not be examined in this research. Also for the same reason, in this study, the changes of members' perceptions before and after implementing the new technologies is not measured. Instead, the real system usage before and after implementing the new technology is used to measure the effectiveness of the proposed technology. The differences in outcomes before and after implementing the added tools and also the changes of the outcomes over time will be studied. The simplified model that is measured in this study can be found in Figure 6-1.

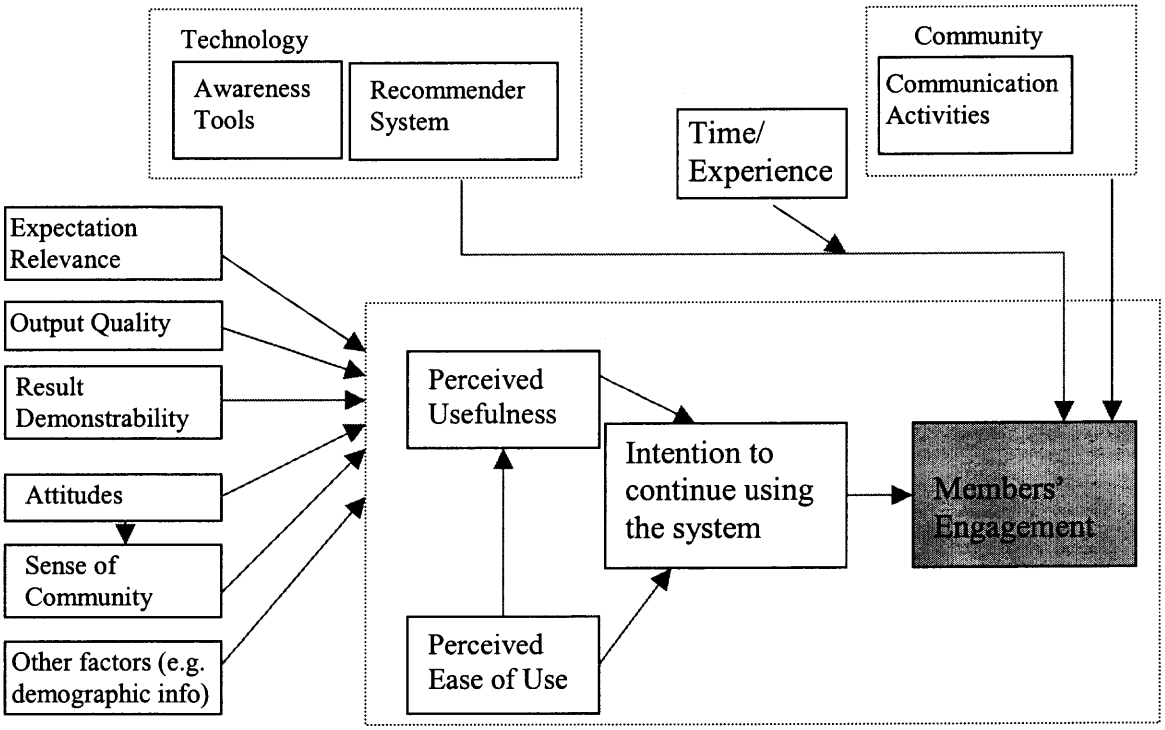

Figure 6.1 Simplified Research Model

6.2 Research Questions and Hypotheses

This research is an exploratory study of how the different factors influence members' behavior in an online research community. The hypothesized interrelationships between different factors and how they change the outcome – **members' engagement** - are presented in the research framework. Because this is an exploratory study, research questions will be presented for relationships that do not have much literature support and hypotheses will be presented for those that do have literature support.

6.2.1 Awareness Tools

Awareness of other people's existence and their behavior is very important for starting a relationship. Designed to compensate for less social cues in CMC, the awareness tools in the community system can help community members learn not only the basic information about other members (e.g. Who are they? Where have they come from?

What are their research interests? What do they look like?), but also learn about other members' activities in the community (e.g. When did they become a member? When was their last visit? What did they post in the community?). Making people and their activities visible to others will potentially reduce the initial uncertainty among members, thus encouraging networking among them. It will also foster feelings of self-esteem, reputation, responsibility and accountability as people become aware that others are "looking" at them (Erickson et al., 1999, Constant et al., 1997, Girgensohn & Lee, 2002).

RQ1: Will the support of awareness help to motivate social engagement in the community?

RQ1a*:* Will members use the awareness tools to get to know other people?

RQ1b: Will members start networking with other members with the information they obtain from the awareness tools?

6.2.2 Recommender System

One common purpose of people who join a research community is to gain knowledge in the specific research area. Thus, information need in research communities is higher than some other types of communities, such as social space, emotional support, etc. Valuable information is one of the main reasons that attract people to come to these online communities. There are two difficulties in providing valuable information: 1) as the volume of information is growing day by day, it will exceed the limit that users can absorb and cause information overload; 2) one person's treasure might be another's trash. It is impossible to make a set of information to be valuable to every member. In this research, a recommender system has been designed to solve the above two problems, an intelligent agent for picking up valuable information based on individual preferences

(Ramakrishnan, 1997; Resnick and Varian, 1997; Terveen et al., 1997; Pemberton et al., 2000; Im and Hars, 2001). With the support of a recommender system, the community place is expected to better serve individual information needs.

RQ2: Will the implementation of a recommender system serve members' information seeking demand better?

RQ2a: Will there be more information seeking activities after implementing the recommender system than before its implementation?

RQ2b: Will members read the items that the recommender system provided for them?

6.2.3 Communication Activities

From the resource-based social structure model (Butler, 2001), community size and communication activities have a dynamic effect on members' engagement. At the community growing stage, larger size and more community activity will attract more people to join and participate. However, when community size becomes too big or demands excessive parallel communication activities, these factors will cause problems such as information overload, free-riding and social loafing (Rafaeli and LaRose, 1993, Jones, Ravid and Rafaeli, 2002). As the field study community is not a large community, the effect of community size on members' engagement is not expected in this research. Instead, this research will focus on the effect of community activities.

RQ3: How do communication activities influence members' engagement in the community?

RQ3a: Will more communication activities attract more people to join the community?

RQ3b: Will more communication activities let members to engage more in the community?

6.2.4 Expectation Relevancy

Expectation relevancy here is defined as the degree to which the community place is relevant to an individual's expectations. This construct is extended from the "job relevance" construct in the TAM2 model, which is defined as an individual's perception regarding the degree to which the target system is applicable to an individual's work (Venkatesh and Davis, 2000). Job relevance is one of the key components for users to judge whether they will use a technology or not. Here, in this context, expectation relevancy refers to an individual's perception about the degree to which the community place is applicable to the goal an individual wants to achieve. Similar to job relevancy, the Expectation relevancy in this model is very important for people to judge whether they want to belong to a community place or not. Kieras and Polson (1985) and Polson (1987) claimed that users have distinct knowledge about their job situation. This knowledge in turn will influence individuals' judgment on what tasks a certain system performs. Other studies have linked user acceptance to variables similar to job relevance, such as job-determined importance (Leonard-Barton and Deschamps, 1988), involvement (Hartwick and Barki, 1994), task-technology fit (Goodhue, 1995), and cognitive fit (Vessey, 1991). In the study of TAM2, job relevance and output quality work together as an interaction effect for determining perceived usefulness (Venkatesh and Davis, 2000). Here, Expectation relevancy serves as a cognitive judgment that directly affects users' perceptions of the usefulness of the community place, and also influences their intention to stay in the community place. As perceived ease of use is an individual's judgment

regarding community system usability, expectation relevancy should not have any significant influence on it. The following hypotheses were proposed related to expectation relevancy.

H1a. Expectation relevancy will have a significant positive influence on perceived usefulness of a community place.

H1b. Expectation relevancy will NOT have a significant influence on perceived ease of use of a community system.

H1c. Expectation relevancy will have a significant positive influence on intention to continue using a community place.

H1d. Expectation relevancy will have a significant positive influence on real usage of a community place.

Furthermore, in order to understand expectation relevancy, understanding members' expectations from the community is also important. Outcome expectation is a construct from social cognitive theory. It has been divided into performance expectation and personal expectation in IS studies (Compeau and Higgins, 1995). Performance expectation is the performance-related consequences of the behavior; it is directly related to work-related outcomes (Compeau and Higgins, 1995). Personal expectation is the personal consequences of the behavior. It deals with individual esteem and sense of accomplishment (Compeau and Higgins, 1995). For this study, the following general research question was asked:

RQ4: What are individual's expectations when they joined the community place?

6.2.5 Output Quality

Relevancy of members' expectations is not sufficient for people to decide whether they will continue using a community place. People will also take into account how well the content of the community place meets their expectations, which is referred to as perception of output quality. Image theory (Beach and Mitchell, 1987, 1990, 1998) distinguishes two stages of decision making. First, decisions are consist of adopting or rejecting potential candidates from future consideration. Second, a profitability test is used to compare acceptable options with one another to determine the final choice. If expectation relevancy is used in the first stage to eliminate some alternative choice from further consideration, judgments on output quality, in contrast, serve in the second stage for a profitability test in which, given a choice set of multiple relevant spaces, one would be inclined to choose a space that delivers the highest output quality. In studies of technology acceptance, output quality refers to how well the system performs the tasks that the user expects to fulfill (Venkatesh and Davis, 2000). Studies have found that there is a relationship between output quality and perceived usefulness (Davis et al., 1992; Venkatesh and Davis, 2000). In this research context, output quality is expected to not only affect people's perception of usefulness of the community place, but also their decision of staying in the community place. This in turn will influence the real usage of the community system. Again, output quality is not expected to have any influence on perceived ease of use of the community place as the latter is more concentrated on the interface of the system, not the content of the place.

H2a. Output quality will have a significant positive influence on perceived usefulness of a community place.

H2b. Output quality will NOT have a significant influence on perceived ease of use of a community system.

H2c. Output quality will have a significant positive influence on intention to continue using the community place.

H2d. Output quality will have a significant positive influence on real usage of a community place.

6.2.6 Result Demonstrability

Even effective systems can fail to gain user acceptance if people have difficulty recognizing accomplishment in their job performance specifically due to their use of the system. Result demonstrability is defined as the "tangibility of the results of using the innovation" (Moore and Benbasat, 1991). According to TAM2, result demonstrability has a strong positive effect on perceived usefulness (Venkatesh and Davis, 2000). The Agarwal and Prasad (1997) study also found a significant correlation between usage intentions and result demonstrability. This relationship is also consistent with the job characteristics model, which emphasizes knowledge of the actual results of work activities as a key psychological state underlying work motivation (Hackman and Oldham, 1976, Loher et al., 1985). In this research context, result demonstrability was predicted to have influence on perceived usefulness, intentions for future usage and the real usage of the community place. As with expectation relevancy and output quality, result demonstrability should not have a direct influence on perceived ease of use of the community system.

H3a. Result demonstrability will have a significant positive influence on perceived usefulness of a community place.

H3b. Result demonstrability will NOT have a significant influence on perceived ease of use of a community system.

H3c. Result demonstrability will have a significant positive influence on intention to continue using the community place.

H3d. Result demonstrability will have a significant positive influence on real usage of a community place.

6.2.7 Attitude

Attitude toward behavior in Theory of Reasoned Action (TRA) is defined as "an individual's positive or negative feelings (evaluative affect) about performing the target behavior" (Fishbein and Ajzen, 1975, p216). When applying this construct for our community engagement research model, attitude refers to an individual's positive or negative feelings about joining the online community and performing different activities in the community. People who regard online communities only as a place to get information are not likely to start a conversation in the community; people who are skeptical about the online environment will be unlikely to interact with other people (Utz, 2001). Because of their attitude, people's perception of usefulness of the community place and their real usage of the community place might be different. People who join the community for just information or people who are skeptical about online environments will not care much about whether the community place is also a social space. Thus, there will be less social engagement activities carried out by these members, and lower degree of perceived preliminary networking activities. Their activities will be mainly focused in information seeking. The attitude towards the online community will also influence the perceived usefulness of awareness tools.

H4a: People who think the community place is just for information will have less social engagement activities than people who do not agree with this.

H4b: People's attitude towards online community will have a significant positive influence on perceived usefulness of the community place.

H4c: People's attitude towards online community will have a significant positive influence on continued use of the community place.

H4d: People who think the community place is just for information will have a lower level of perceived sense of community.

H4e: People who think the community place is just for getting information will use the awareness tools less than others.

H4f: People who think the community place is just for getting information will have a lower perceived usefulness of awareness tools.

H4g: People's attitude towards online community will have a significant positive influence on their perceived preliminary networking.

6.2.8 Sense of Community

Sense of community has drawn great interest in community psychology research. McMillan in his article about sense of community proposed:

> "I view Sense of Community as a *spirit* of belonging together, a feeling that there is an authority structure that can be *trusted*, and awareness that *trade*, and mutual benefit come from being together, and a spirit that comes from shared experience that are preserved as *art*." (McMillan, 1996, p315)

People join an online research community not only for getting information on a certain research area, but also for the need to feel belonging to the group of the same research area. Thus, how well the online space makes members feel they are belonging to a real group will also influence their decision to stay or leave the community place, which is referred to as sense of community. Different from a traditional community where people inside a community always meet together face-to-face, people in an online community cannot see each other, and they might visit the online space at different times. The sense of community shows how cohesive a community is. It is a construct that applies at both the individual level and the community level (Hill, 1996). The awareness tools designed in this study are used to increase the sense of community. However, because of the limitations of the field study, it is impossible to ask the same group of subjects to answer two questionnaires (before and after implementing the awareness system). Thus, the influence of technology on sense of community cannot be directly measured. Only the relationships between sense of community and people's perception about the awareness tools and their social engagement activities can be measured. .

H5a. Perceived sense of community will have a significant positive influence on people's perception of the awareness tools.

H5b. Perceived sense of community will have a significant positive influence on member's social engagement activities.

H5c. Perceived sense of community will have a significant positive influence on perceived usefulness of the community place.

H5d. Perceived sense of community will have a significant positive influence on perceived intention to continue using the community place.

H5e. Perceived sense of community will have a significant positive influence on perceived Preliminary networking.

6.2.9 Other Factors

Other factors here refer to gender, role, Internet connection speed, and time spent online. Besides the factors that discussed above which influence the perceived usefulness of the community place, other factors such as demographic information and time constraints, and the accessibility of the Internet may also influence users' real usage of the community place (Nonnecke and Preece, 2001). Gender, age, and experience have also been used as moderators of social influence on behavior intention in the Unified TAM model (Venkatesh et al., 2003). In this research model, how these other factors influence users' perceptions and their real usage of the community place were examined.

RQ5a. Will there be any gender differences in perception and real usage of a community place?

RQ5b. Will person's role affect perception and the real usage of the community place?

RQ5c. Will the connection speed affect members' perception of a community place and their real usage of the community place?

RQ5d: Will average online time have any effect on members' perception and real usage of the community place?

6.2.10 Basic TAM Model Constructs

The Original TAM model has been adapted in our research model to explain the process of how different factors influence the users' engagement in the community (real system usage). It contains three constructs: perceived usefulness, perceived ease of use and

intention to continue use. In this study, instead of studying initial intention to adopt the community place, intention to continue using the community place was studied. Perceived usefulness is defined as the extent to which a person believes that using the system will improve his or her task performance. Perceived ease of use is defined as the extent to which a person believes that using the system will be free of effort (Venkatesh and Davis, 2000). According to TAM, the perceived usefulness and perceived ease of use will greatly influence people's intention to continue to use the system. In other studies, perceived ease of use becomes an insignificant factor after the initial adoption of the system (Agarwal and Prasad 1997, 1998; Tompson et al. 1991, 1994; Karahanna et al., 1999; Venkatesh et al. 2003). In this study, as the dependent variable is intention to continue use, perceived ease of use was not expected to have a great influence on it. Also different from the original context of TAM, the relationships between different constructs in TAM were expected to predict people's acceptance of an online community place.

H6a: Perceived usefulness will have a significant positive influence on intention to continue using the system.

H6b: Perceived ease of use will NOT have a significant influence on intention to continue using the system.

H6c: Perceived ease of use will have a significant positive influence on perceived usefulness.

H6d: Intention to continue using the system will have a significant positive influence on real system usage.

H6e: Perceived Usefulness will have a significant positive influence on preliminary networking

H6f: Perceived Usefulness will have a significant positive influence on real usage of the community system.

6.2.11 Time

After implementing the new system features, people need time to get familiar with the system. Similarly, when forming a new online community, people also need time to be familiar with the community place and be familiar with other members. Members' experience in the community place will change their view of the community, and in turn change their usage of the system. Changes of the community use are expected over time.

RQ6: How will the usage of the community system change over time?

6.3 Summary

This research aims to discover behavior of people in an online research community, and to find out factors that influence peoples' adoption to and sustainability in an online research community. The proposed research model presents technology influences (social awareness tools, recommender system), community influence (community activities), and individual factors influence (expectation relevancy, output quality, result demonstrability, attitude, sense of community, perceived ease of use, perceived usefulness) on the social and information engagement activities of an online research community.

The next chapter will introduce different research methodologies in this research, and detailed measures for each construct presented in the research model.

CHAPTER 7

RESEARCH METHODOLOGY

This chapter will first introduce the field site for this research, then discuss in detail different research methodologies used to help answer research questions and test hypotheses.

7.1 Field Study Site

The "Asynchronous Learning Networks (ALN) Research Community" was selected for this field trial. ALN is a form of e-learning that combines self-study with substantial, rapid, asynchronous interactivity with others (http://www.aln.org). The community is hosted on a website called the WebCenter for ALN Effectiveness Research (http://www.alnresearch.org, also see Figure 7.1). This community is fairly small, with about 170 registered members by the end of summer, 2002. The community size gets 420 by Dec. 2003. The potential membership of the ALN research community is estimated to be from one to two thousand people, consisting of current researchers publishing in the field, faculty members who teach online courses, and graduate students doing research in the field. The overall objective of this community is to increase the quality, quantity, dissemination, and application of results of research on the effectiveness of ALN. The goal is to:

- Synthesize existing knowledge on effectiveness of learning networks
- Provide a space for members to share: research methods, theoretical frameworks, and instrumentation for assessing online learning
- Discuss emerging issues and events
- Create new knowledge

Although the initial release of the WebCenter for the community was in August 2001, currently there is not much active participation on the community site; most users use the site as a digital library to get research articles or resources. There are very few discussions in the discussion forum and few members' contributions. The community site did not fulfill its task to connect researchers together. This situation of the site provides a good place for trying out the framework for stimulating members' engagement in the community and encouraging online interpersonal relationship development.

Figure 7.1 Homepage of the ALN Webcenter

7.2 Research Methodology

In order to understand members' behavior in an online research community and test our research model with the selected online research community, both qualitative and quantitative methods are used.

7.2.1 Protocol Analysis

In order to achieve a reasonable objective usability, protocol analysis is used to get feedback on the usability of the system; approximately 12 subjects were recorded. The system then was refined based on the feedback gathered, and tested again with the "thinking out loud" method.

The subjects were selected from graduate students and other adult users of the Internet. Both novice and expert subjects were selected based on their experience with computers and the Website. Detailed results of protocol analysis are reported in Appendix B.

Planned Procedures:

1. Generate appropriate tasks to test the usability of different system features on the website.
2. Develop questionnaires or interview guides
3. Pilot questionnaires as part of coursework, not for publication
4. Obtain approval from human subjects committee (IRB)
5. Select subjects
6. Run the pilot protocol analysis study (2 subjects)
7. Modify task and questionnaire
8. Real protocol analysis study (6 subjects)
9. Analyzing tape and screen events
10. Make changes based on results
11. Repeat protocol analysis with four new subjects

The summarized results are shown in Table 7.1:

Table 7.1 Results from Protocol Analysis

CATEGORY	POSITIVE FEEDBACK	NEGATIVE FEEDBACK
Information	The site has lots of good contents	Help is missing.
Presentation	The presentation of information, especially articles is very good.	
	Some pages and layout are very attractive to the user.	Sometimes, the color appears too dull and sometimes there are multiple colors on the page that does not make it professional
	Maintaining the history of a user makes it very helpful.	
Navigation		Article titles do not have hyperlink, which created confusion.
	Menu at the top of the user's home page is very good	Menu scrolls up when the user scrolls down the home page.
	All information and menu items are available on the home page and make it easy to go to any link.	Accessing the home page is not easy from links. Some links do not have a link back to the home page.
		Some tab entries are confusing (ex DB Entry), as they do not make implicit sense to the user. Same is true for some details on the registration page (How did you hear about the site)
Login		After Registering into the site, the user is not directly logged into the site.
Discussion Forum	The organization is good, the icons look good	
	Fast reply feature is highly appreciated.	Menu scrolls up when the user scrolls down the home page.
Project/People	Presentation of people and project team is good	Users tend to look into people tab when looking for project team details
	Good representation using Internet Explorer	Not sure how it is presented on other browsers.
Search	Search feature is available on the user's home page after he/ she has logged in. Advanced search is available for users who want to filter their searches.	Search does not appear on the regular home page. Advanced search feature does not enable multiple filters. The user has to toggle between multiple criteria.

Summary of recommendations is developed after consulting and deliberating with all the evaluators (see Table 7.2).

Table 7.2 Summary of Recommendations from Protocol Analysis

CATEGORY	RECOMMENDATION
Navigation	Pointing the mouse on a button or a tab should display some information about the button. Drop boxes or selection menus can replace the text areas in the registration form.
	Links should be provided on the article title, which opens up the article. A thread on the side of the screen will also make the site more navigable.
	Menu can appear on a different frame to make it always visible even if the user scrolls to the bottom of the menu.
Appearance	Playing around and trying different color combinations will yield an optimal color scheme, which is professional, and reduces stress, fatigue and dullness in the users since most users of this system are expected to use the system for long durations.
	Frames/No-frames option will increase the spectrum of users
Organization	People can have a section on members and another section on project personnel.
Registration	Creating a registration should bring the user directly bring to the home page

7.2.2 Survey

A short survey was used to measure constructs related to individual aspects and also self-reported usage. This short survey contains 24 five scale questions and two open-ended questions. A short survey is chosen due to the situation of a field study. As the questionnaire response is totally dependent on volunteers and the subjects are not located in the same area, it is not appropriate to use a long survey questionnaire. The survey was distributed at the annual ALN conference, 2003 and also online from the ALNResearch website. The survey was first presented to three Ph.D. students in IS for the check on

external validity. The Ph.D. students were asked about their understanding of the questions. Their answers were compared to the understanding of the author and appropriate adjustments were made accordingly.

The questionnaire was then presented to the NJIT human subjects review board to get approval. The approved consent form and questionnaire can be found in Appendix A and B. The following shows constructs that have been measured by the questionnaire.

7.2.2.1 General Information. The following general information has been collected from the questionnaire (see Table 7.3) in order to gather information to test the "Other Factors". Question about subject's role in ALN is not an exclusive question, subjects can choose multiple columns that fit for their individual situation.

Table 7.3 General Information from Questionnaire

Information	Questionnaire item
Identity	Name: _____ Title: _____ Gender: M___ F___ Email: _____
Internet connection	I usually connected to the Internet using: (a) Modem dialup (56k) (b) Cable Modem/DSL (c) T1
Internet usage	I use the Internet about _____ hours a day (a) Less than 1 (b) Less than 2 (c) 2-4 (d) 4-6 (e) more than 6
Role in ALN	Your roles in ALN (choose all that apply) (a) Researcher (b) Educator (c) Student (d) Administrator (e) Other _____

7.2.2.2 Expectation Relevancy. The following two questions are adopted and modified from TAM2 measurement of job relevance (Davis and Venkatesh, 2000). They are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.4).

Table 7.4 Expectation Relevancy Scale

Item No.	Question / Statement
9	I can get materials that are important to me from the WebCenter.
24	The ALNResearch WebCenter has something useful for my work

7.2.2.3 Output Quality. The following two questions are adopted and modified from TAM2 measurement of output quality (Davis and Venkatesh, 2000) to fit in the current research context. They are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.5).

Table 7.5 Output Quality Scale

Item No.	Question / Statement
1	I am satisfied with the quality of the ALNResearch WebCenter content.
20	The quality of the information I get from the WebCenter is high.

7.2.2.4 Result Demonstrability. The following three questions are adopted and modified from TAM2 measurement of result demonstrability (Davis and Venkatesh, 2000) to fit in current research context. They are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.6). One of the items is a negative question.

Table 7.6 Result Demonstrability Scale

Item No.	Question / Statement
2	Overall, I feel the benefits of using the WebCenter system are apparent to me.
12	It is **difficult** for me to explain why using the WebCenter may or may not be beneficial.
19	I have no difficulty telling others about the results of using the WebCenter system.

7.2.2.5 Attitude. The following three questions are adopted and modified from measurement of attitude toward online environment (Utz, 2000). They are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.7).

Table 7.7 Attitude towards the Online Community Scale

Item No.	Question / Statement
13	I **DO NOT** trust anyone online.
16	I **DO NOT** plan to post anything on the site or communicate with anyone through the WebCenter.
17	For me, the WebCenter is JUST a place where I can get information.

7.2.2.6 Intention to Continue Use. The following two questions are adopted and modified from TAM2 measurement of intention to continue use (Davis and Venkatesh, 2000). They are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.8).

Table 7.8 Intention to Continued Use Scale

Item No.	Question / Statement
7	If I can access the WebCenter, I will continue to use it.
21	I **DO NOT** plan to use WebCenter any more.

7.2.2.7 Sense of Community. The following three questions are selected and modified from measurement of sense of class community (Rovai, 2002). They are constructed as

Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.9).

Table 7.9 Sense of Community Scale

Item No.	Question / Statement
4	I trust other members in the ALNResearch WebCenter
18	I always feel isolated in the ALNResearch WebCenter
22	I feel connected to others in the ALNResearch WebCenter

7.2.2.8 Perceived Ease of Use. The following three questions are adopted and modified from TAM2 measurement of perceived ease of use (Davis and Venkatesh, 2000). They are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.10).

Table 7.10 Perceived Ease of Use Scale

Item No.	Question / Statement
6	The interface of the ALNResearch WebCenter is clear and understandable.
15	I found it easy to complete what I want to do in the ALNResearch WebCenter
23	It requires **significant** effort to use the ALNResearch WebCenter features

7.2.2.9 Perceived Usefulness. The following four questions are adopted and modified from TAM2 measurement of perceived ease of use (Davis and Venkatesh, 2000) and also questions from online communities research (Hiltz, 1984). They are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side (See Table 7.11).

Table 7.11 Perceived Usefulness Scale

Item No.	Question / Statement
5	The WebCenter has provided me with guidance, reference, or other information useful to my work.
8	The WebCenter has broadened my understanding of my research area.
11	I find the ALNResearch WebCenter is **NOT** useful at all for my work.
14	Using the ALNResearch WebCenter improves my work.

7.2.2.10 Preliminary Networking. The following three questions are made to check the preliminary networking inside the community. Preliminary networking was treated as a measure of social engagement in this study. Here, preliminary networking means the pre condition for people to interact with each other. For example, people need to know another person first before they can start talking to them. The first two questions are constructed as Likert-type scales anchored at five points with "Strongly agree" on the left hand side and "Strongly disagree" on the right hand side, and the last question is a "yes/no" question (see Table 7.12)

Table 7.12 Preliminary Networking Scale

Item No.	Question / Statement
3	I might contact someone in the people directory who shares research interests with me.
10	The WebCenter helps me know who else is doing research in my research field.
25	I have contacted with people who I met from the WebCenter by email, phone or other media.

7.2.2.11 Self-reported Usage. The following three questions are adopted and modified from study of online community (Hiltz 1984) to evaluate self-reported usage. (See Table 7.13).

Table 7.13 Self-reported Usage

	Frequency of Use			Value				
	(1) Never Used	(2) Occasiona -lly	(3) Frequently	(1) Cannot Say	(2) Useless	(3) Slightly Useful	(4) Fairly Useful	(5) Extremely Valuable
Empirical Studies								
Other Studies								
Research Instrument								
Research Resources								
Tutorials								
People Directory								
Discussion Forum								
Contribution Forms								

7.2.2.12 Users' Expectations. The following two open-ended questions are used to get members real expectation (see Table 7.14).

Table 7.14 Open-Ended Questions for Expectation

Item No.	Question / Statement
1	What are your expectations when you decided to join the ALNResearch WebCenter?
2	What other kind of features would you like to see added to the ALNResearch WebCenter?

7.2.3 Web Usage Mining

Web usage mining is used in this research to measure user's behavior in the community system. As the dependent variable for this study, members' engagement was measured through a web usage mining technique. Members' engagement was judged by members' behavior within a certain period of time. Members' behavior was captured in the server log file. Mining through the log file was planned to categorize members' activities, analyze their social engagement and information engagement.

The study of Web server logs in order to uncover visitor usage patterns has been widely used in e-commerce and Web usability studies, although it has evolved over time through quite a variety of names, such as Web traffic monitoring, clickstream analysis, Web e-metrics, and Web usage mining. Each of these terms refers to "the application of data mining techniques to discover usage patterns from Web data" (Srivastava et al., 2000).

The raw data captured in Web server logs cannot be directly used for web usage mining. They need to be processed to construct several data abstractions, notably users, server sessions, episodes, click-streams, and page views. These data abstractions then can be used to generate different usage patterns.

For this study, customized Web server logs are used as data source instead of standard Web server logs. The reason for using customized Web logs will be explained in the next chapter. Hits, page view, visitor, and server session are constructed from the data source to perform web usage mining in order to both discover user behavior pattern and to measure engagement. In this research, social engagement activities have been defined as any page access to attempt knowing about other people, or sharing information with other people; information seeking behavior have been defined as page access that aims for seeking information. From the server side, users activities inside the website are captured after they log into the site (see sample log file in Figure 7-1).

```
10:44 f0301641161068561365469 1068561365422 1068565470391 visitor01 resources_research_instruments
10:48 f0301641161068561365469 1068561365422 1068565691359 visitor01 resources_learning_theories
10:49 f0301641161068561365469 1068561365422 1068565756172 visitor01 resources_methods
10:50 f0301641161068561365469 1068561365422 1068565804281 visitor01 resources_journals
10:52 f0301641161068561365469 1068561365422 1068565964187 visitor01 resources_evaluation
10:54 f030829931068566067016 visitor02 NewSession******************************
10:54 f030829931068566067016 1068566066984 1068566089328 visitor02 paper_db_entry 31
11:16 f0301084851068567392016 visitor03 NewSession******************************
11:16 f0301084851068567392016 1068567391984 1068567411469 visitor03 home
11:17 f0301084851068567392016 1068567391984 1068567426406 visitor03  paper_abstract 158
11:18 f0301084851068567392016 1068567391984 1068567518719 visitor03 paper_abstract 142
11:19 f0301084851068567392016 1068567391984 1068567579703 visitor03 paper_abstract 148
11:41 f0301503611068568545328 visitor04 NewSession******************************
11:41 f0301503611068568545328 1068568545266 1068568918891 visitor04 resources_research_instruments
11:43 f0301503611068568545328 1068568545266 1068569016156 visitor04 resources_learning_theories
11:44 f0301503611068568545328 1068568545266 1068569047406 visitor04 project_people
11:44 f0301503611068568545328 1068568545266 1068569085328 visitor04 project_people_proposal
11:46 f0301934121068569210047 visitor05 NewSession******************************
11:46 f0301934121068569210047 1068569209984 1068569217609 visitor05 home
```

Figure 7.2 Sample Log File

In order to measure engagement, all activities have been divided into two categories: social engagement and information seeking, based on page view. For most page views, one point has been given as a measure of engagement. For those pages that need more effort to read or input, two points have been given for engagement score. Then, all the points for social engagement and information seeking for a user are summed up for a certain period of time to count as their engagement. Table 7.15 shows some example of page views. Detailed results will be discussed in the next chapter.

Table 7.15 Examples of Page Views

Page View #	Start time (HH:MM)	Page View Description	Category
2	9:45	home	
3	9:48	paper_abstract	1
4	9:57	paper_full_text	1
5	10:14	paper_full_text	1
6	10:14	paper_abstract	1
7	10:15	paper_abstract	1
8	10:16	paper_full_text	1
9	10:21	paper_abstract	1
10	10:22	paper_abstract	1
11	10:22	paper_abstract	1
12	10:25	paper_full_text	1
13	10:26	paper_abstract	1
14	10:27	paper_full_text	1
15	10:28	paper_abstract	1
16	10:29	paper_full_text	1
17	10:43	people_directory	2
18	10:43	project_people	2
19	10:43	forum_list	2
20	10:44	resources_research_instruments	1
21	10:48	resources_learning_theories	1
22	10:49	resources_methods	1
23	10:50	resources_journals	1

7.3 Data Collection and Measures

Data used to test the research model are captured from web server logs, while the survey questionnaire was answered by members of the community.

A summary of measures for each variable is shown in table 7.16. The corresponding consent form can be found in Appendix A. The survey questionnaire can be found in Appendix B.

Table 7.16 Summary of Measures

Variable	Measures
Social Engagement	New registered users
	People directory page view
	Forum page view
	Paper comments page view
	Project page view
	Contribution page view
information seeking	Empirical study page view
	Other study page view
	Tutorial page view
	Resources page view
Communication activities	Original postings in discussion forum
	Replies
	Contributions of articles
	Comments on articles
Intention to continue use	Survey questionnaire
Expectation relevancy	Survey questionnaire
Output Quality	Survey questionnaire
Result Demonstrability	Survey questionnaire
Sense of community	Survey questionnaire
Attitude	Survey questionnaire
Other Factors	Survey questionnaire
Perceived usefulness	Survey questionnaire
Perceived ease of use	Survey questionnaire
General behavior in the community	Web analytics

7.4 Summary

This chapter introduces research methodologies that were used in this research to answer research questions and test research hypotheses that are presented in Chapter 6. The selected research methods include protocol analysis on recommender system features, survey, and web usage mining with web server log data. Each of the chosen methods has constraints, they must complement one another to provide a complete view, to make sure that the results cover not only the quantitative evidence, but also qualitative evidence.

The following four chapters will discuss data analysis and results obtained from different methods. Chapter 8 will present the results from web usage mining on user behavior inside the community place. Chapter 9 will show the summary of questionnaire responses and check the validity and reliability of the measures. Chapter 10 will test the hypotheses using statistical analysis and answer research questions from both log files and questionnaire responses. Chapter 11 will introduce the Partial Least Square (PLS) method and use it to test the integrated research model.

CHAPTER 8

MEMBERS' BEHAVIOR IN THE ONLINE RESEARCH COMMUNITY

As described in the last chapter, several methods are used for this study. This chapter contains a description of the web usage mining method used in this study and the results. The findings are split into three sections: an overview of community system usage, distribution of information seeking and social engagement activities, and individual's usage pattern.

8.1 Web Usage Mining Methods

Web usage mining is the application of data mining techniques to discover usage patterns from Web data. Here, this method is used to help understand users' behavior inside an online research community.

In Web mining, data can be collected from server side, client side, and proxy servers. All the collected data can be classified into several categories (Srivastava et al. 2000)

- Content: The real data in the Web page. E.g. texts or graphs that form a web page.
- Structure: Data that represent the organization of the content. E.g. hyper-links that connect one page to another.
- Usage: Data that describes the usage pattern of Web pages. E.g. IP address, page references, session number, etc.
- User profile: Data that provides demographic information about users of the Web site.

In this research, we will only use data collected from server side. In order to do Web mining, the raw data from Web server logs has to be first reconstructed as data abstractions in order to do Web usage mining. The following terms are items that are usually captured as data abstractions for Web usage mining (World Wide Web Consortium, 1999):

- **hit** –a server request for a physical file. This is often a misleading indicator of Web activity, since the hit could be the request for an image, frame, JavaScript module, or Web page
- **page view** – the combination of every file or element that contributes to a browser display at one time. It is sometimes referred to as a "logical page," and it is essentially what you see when you look at a Web site
- **clickstream** – sequential series of page view requests
- **visitor** – single individual accessing files from one or more Web servers through a browser
- **cookie** – data sent by a Web server to a Web client, to be stored locally by the client and sent back to the server on subsequent requests
- **server session** – clickstream of page views for a single visitor for a particular Web site, also known as a visit

There are three processes in Web usage mining: preprocessing, pattern discovery, and pattern analysis. The following sections will describe preprocessing and its result, pattern discovery and analysis with the results.

8.2 Preprocessing

Preprocessing is a process to convert the usage, content, and structure information in Web server logs and the web site itself into data abstractions for pattern discovery.

Two kinds of web server logs were used in the web usage mining. One is the standard log from the Apache server which will be called "raw log" in the rest of the document. The other one is the log created from the WebCenter application, which records userid, page visited, time visited, and session number after a user logs in the WebCenter. This log will be called the "customized log" in the rest of the document.

Usage preprocessing is argued to be the most difficult task in Web usage mining because of the incompleteness of the available data in raw log. Usually, only the IP address, agent, and server side click-stream are captured in Web server logs. It is very hard to identify each visitor (See Table 8-1 for example). For example, several users may access the Web through the same proxy server, which ends up with a single IP/multiple user, or one visitor can access the Web from different machines, which ends up with single user/multiple IP. Also, there are lots of noises in raw log, for example, the first two lines in Table 8.1 are accesses from googlebot. This information should not be counted for the user usage. Because of the above reasons, in this study, the raw log will be only used for providing overall picture of the WebCenter access pattern. Customized log will be used for detailed analysis of usage data.

Table 8.1 Sample from Raw Logs

64.68.82.58 - - [27/Mar/2004:05:17:24 -0500] "GET /JSP/People/../user_login.jsp HTTP/1.0" 200 4420 "-" "Googlebot/2.1 (+http://www.googlebot.com/bot.html)"

64.68.82.201 - - [27/Mar/2004:05:18:02 -0500] "GET /HTML/question.html HTTP/1.0" 404 311 "-" "Googlebot/2.1 (+http://www.googlebot.com/bot.html)"

203.76.253.39 - - [27/Mar/2004:05:18:28 -0500] "GET /HTML/AssessmentTutorial/ HTTP/1.1" 200 616 "http://www.google.com.au/search?q=assessing+learning&ie=UTF-8&oe=UTF-8&hl=en&meta=" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"

203.76.253.39 - - [27/Mar/2004:05:18:28 -0500] "GET /HTML/AssessmentTutorial/contents/top.html HTTP/1.1" 200 775 "http://www.alnresearch.org/HTML/AssessmentTutorial/" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"

203.76.253.39 - - [27/Mar/2004:05:18:29 -0500] "GET /HTML/AssessmentTutorial/contents/left.html HTTP/1.1" 200 2894 "http://www.alnresearch.org/HTML/AssessmentTutorial/" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"

In order to overcome this problem, customized log files are used instead of server default log files. As users have to log into the ALNResearch website in order to see the content, whenever a user logs in, his/her behavior is recorded into the customized log file. By using customized log files, the problem of identifying visitors and sessions can be easily solved. Table 8.2 shows part of the raw log file. Table 8.3 is a sample session of page views from one visitor identified from the customized log file. Table 8.4 shows a sample of a session summary.

Table 8.2 Sample of Customized Logs

10:44 f0301641161068561365469 1068561365422 1068565470391 visitor01 resources_research_instruments

10:48 f0301641161068561365469 1068561365422 1068565691359 visitor01 resources_learning_theories

10:49 f0301641161068561365469 1068561365422 1068565756172 visitor01 resources_methods

10:50 f0301641161068561365469 1068561365422 1068565804281 visitor01 resources_journals

10:52 f0301641161068561365469 1068561365422 1068565964187 visitor01 resources_evaluation

10:54 f030829931068566067016 visitor02 NewSession**********

10:54 f030829931068566067016 1068566066984 1068566089328 visitor02 paper_db_entry 31

11:16 f030108485 1068567392016 visitor03 NewSession**********

11:16 f030108485 1068567392016 1068567391984 1068567411469 visitor03 home

11:17 f030108485 1068567392016 1068567391984 1068567426406 visitor03 paper_abstract 158

11:18 f030108485 1068567392016 1068567391984 1068567518719 visitor03 paper_abstract 142

11:19 f030108485 1068567392016 1068567391984 1068567579703 visitor03 paper_abstract 148

Table 8.3 Sample of One Visitor Session of Page Views

Page View #	Start time (HH:MM)	Page View Description
1	9:45	home
2	9:48	paper_abstract
3	9:57	paper_full_text
4	10:14	paper_full_text
5	10:14	paper_abstract
6	10:15	paper_abstract
7	10:16	paper_full_text
8	10:21	paper_abstract
9	10:22	paper_abstract
10	10:22	paper_abstract
11	10:25	paper_full_text
12	10:26	paper_abstract
13	10:27	paper_full_text
14	10:28	paper_abstract
15	10:29	paper_full_text
16	10:43	people_directory
17	10:43	project_people
18	10:43	forum_list
19	10:44	resources_research_instruments
20	10:48	resources_learning_theories
21	10:49	resources_methods
22	10:50	resources_journals

Table 8.4 Sample Session Summary

Session #:	Visitor:	Start time:	Total time:	# of pages viewed:	# articles accessed:
1	Visitor1	2:21	0:05	36	0
2	Visitor2	2:26	0:00	2	0
3	Visitor3	2:43	0:20	48	6
4	Visitor4	2:47	0:09	15	8
5	Visitor5	3:09	0:11	10	0
6	Visitor6	3:37	0:05	23	4
7	Visitor7	3:48	0:18	10	0
8	Visitor8	4:26	0:06	11	1

Content preprocessing involves converting the text, image, and other files into forms that can be used in a Web Usage Mining process. Usually, this process performs

content mining such as classification or clustering page views. Page views can be classified not only according to their topic, but also based on their intended use. Page views can be intended to convey information, gather information from users, achieve navigation, or some combination. Based on the context of this research, two classifying methods have been used. One method is to classify the page views into two categories: social engagement and information seeking activities. The other method is to classify page views based on their intended use: convey information, gather information, and navigation . According to the content of the WebCenter web site, all the pages can be divided into eight categories: empirical research (1), other research (2), resources (3), tutorials (4), contributions (5), discussion forums (6), people directory (7), and project related (8). Table 8.5 shows the classification based on the first method with the structure category id, and Table 8.6 shows the classification based on the second method.

Table 8.5 Page View Categorized by Engagement

Category	Page view description	id[1]	Category	Page view	id
Information Seeking	paper_abstract	1	Social engagement	contribute_instruments	5
	paper_comments_display	1		contribute_books	5
	paper_db_entry	1		contribute_links	5
	paper_full_text	1		contribute_others	5
	paper_new_articles	1		contribute_papers	5
	paper_search_database	1		contribute_studies	5
	published_research	1		contribute_thesis	5
	books	2		forum_create_thread	6
	books_abstract	2		forum_post_message	6
	other	2		forum_reply	6
	other_abstract	2		quote	6
	other_comments_display	2		post_comments	6
	other_full_text	2		forum_search	6
	other_summary	2		forum_search_results	6
	qualitative	2		forum_list	6
	thesis	2		forum_messages	6
	theories	2		forum_replies	6
	thesis_abstract	2		paper_comments_display	1
	technologies	2		other_comments_display	2
	technology	2		people_directory	7
	resources_evaluation	3		people_show_profile	7
	resources_journals	3		project_people	8
	resources_learning_theories	3		project_people_proposal	8
	resources_methods	3			
	resources_other	3			
	resources_research_instruments	3			
	tutorials	4			
	tutorials_assessment_tutorial	4			
	forum_search	6			
	forum_search_results	6			
	forum_list	6			
	forum_messages	6			
	forum_replies	6			
	paper_comments_display	1			
	other_comments_display	2			

[1] Id refers to the structure id that was described above

Table 8.6 Page Views Categorized by Intended Use

Category	page view	Category	page view
convey information	paper_abstract	get information from user	contribute_instruments
	paper_comments_display		contribute_books
	paper_db_entry		contribute_links
	paper_full_text		contribute_others
	paper_new_articles		contribute_papers
	books_abstract		contribute_studies
	other_abstract		contribute_thesis
	other_comments_display		forum_create_thread
	other_full_text		forum_post_message
	other_summary		forum_reply
	thesis_abstract		quote
	resources_evaluation		post_comments
	resources_journals		maintain
	resources_learning_theories	navigation	home
	resources_methods		published_research
	resources_other		forum_search
	resources_research_instruments		forum_search_results
	tutorials		paper_search_database
	tutorials_assessment_tutorial		books
	forum_messages		other
	forum_replies		qualitative
	people_directory		thesis
	people_show_profile		theories
	project_people		technology
	project_people_proposal		technologies
	Paper_aln_study_contents		forum_list

8.3 Pattern Discovery and Analysis

Pattern discovery is the process of using various mining techniques to discover all the patterns from the preprocessed data. Pattern analysis is the step to filter out uninteresting rules or patterns from what has been discovered in the pattern discovery process. Techniques that can be used in the pattern discovery phase include:

- Statistical Analysis: by analyzing the session file, different kinds of descriptive statistical analysis (frequency, mean, median, etc) on page views, viewing time, and visitors.

- Association Rules: by relating pages that are most often referenced together in a single server session, different association rules may be discovered. This has frequently been used for e-commerce sites to reveal associated product pages. It may also serve as a heuristic for prefetching documents to reduce user-perceived latency when loading a page from a remote site.
- Clustering: by grouping together a set of items having similar characteristics, clusters of users and pages may be discovered. Clustering of users tends to group users who perform similar browsing patterns; clustering pages tends to group pages having related content.
- Classification: by mapping a data item into one of several predefined classes, a profile of users belonging to a particular class may be developed. Classification is usually done by inductive learning algorithms such as decision tree classifiers, naïve Baysian classifiers, k-nearest neighbor classifiers, etc. For example: a result from classification may be: 40% of users who put an order in /Product/Movie are in the 20-30 age group.
- Sequential Patterns: tries to discover inter-session patterns such that the presence of a set of items is followed by another item in a time-ordered set of sessions. This method is also often used in the e-commerce domain to predict future visit pattern.

For the purpose of this research, statistical analysis will be used mainly to describe users' behavior patterns in an online research community.

8.3.1 Overall Usage Summary

The following table shows the overall usage summary from log data collected over a seven month period (10/27/2003-5/25/2004) on the ALNResearch WebCenter site. During the observation period, because of the winter holiday and the server upgrading work, there are eleven days that had zero or near zero visits. These days have been removed from the analysis of the customized log data.

Table 8.7 Overall Summary of WebCenter Usage from Customized Log

	Total	avg/visitor	avg /day	max /day	min /day
No. visitors	507		5.6	21	1
No. page views	21,669	42.7	112.8	1,127	16
No. sessions	1,397	2.75	7.27	29	1
No. articles accessed	5,879	11.6	30.6	280	1

Table 8.7 shows that there are altogether 507 members who visited the WebCenter, and research articles were accessed 5,879 times over the observation time. The average of page views per day is about 113 with an average of 5.6 visitors per day. This shows that when people visited the community space, they spent substantial effort on it.

One thing that needs to be mentioned here is that the customized log file only records dynamic pages with the extension of .jsp. All html files that are directed from jsp files have not been recorded. For example, from tutorial.jsp file, a user can access a series of .html files on different topics,, but these accesses were not recorded in the customized log files. From the raw logs, there are more than 2,000 every day after the new system was implemented. Figure 8.2 shows a summary diagram for hits during the period of 10/28/2003-4/18/2004 and Table 8.8 shows the summary statistics of site usage from the

original web server log. (Due to the formatting of the graph, data from 4/19/2004-5/25/2004 are not included).

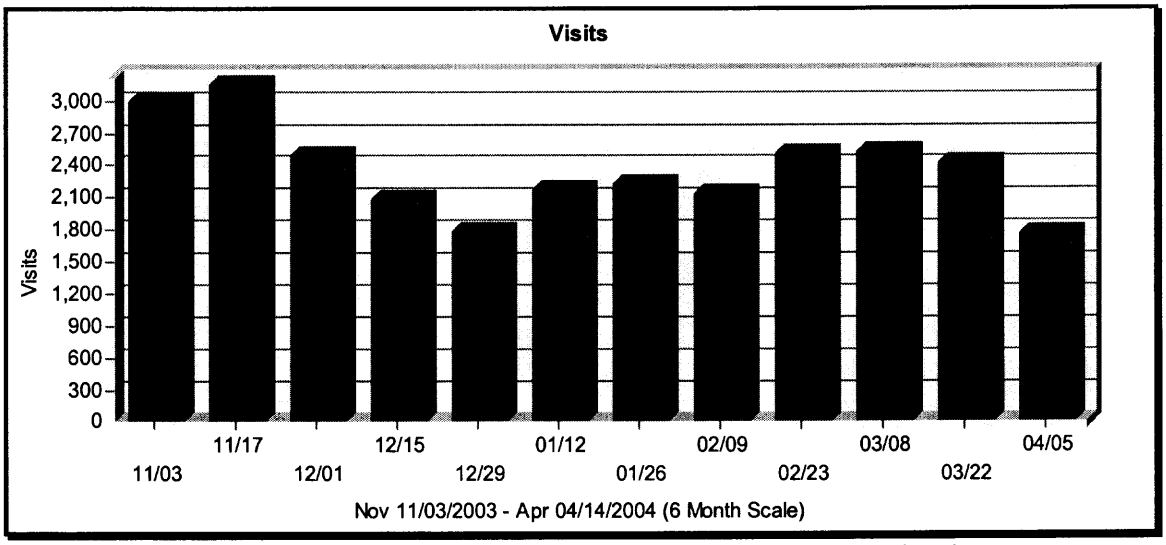

Figure 8.1 General Statistics of Site Visits (From raw logs)

Table 8.8 General Statistics of Site Visits

General Statistics - Report Range: 11/03/2003 00:00:00 - 04/14/2004 23:59:59		
Hits	Entire Site (Successful)	379,984
	Average per Day	2,316
	Home Page	3,764
Page Views	Page Views	80,714
	Average per Day	492
	Average per Unique Visitor	8
	Document Views	20,723
Visits	Visits	28,315
	Average per Day	172
	Average Visit Length	00:12:16
	Median Visit Length	00:00:03
	International Visits	0.00%
	Visits of Unknown Origin	100.00%
	Visits from United States	0.00%
	Visits Referred by Search Engines	4,093
	Visits from Spiders	4,725
Visitors	Unique Visitors	9,736
	Visitors Who Visited Once	7,446
	Visitors Who Visited More Than Once	2,290

From table 8.8, it is obvious that number of hits is much larger than number of page views. The reason is that hit here is defined as a single action on the Web server as it appears in the log file. A visitor downloading a single file is logged as a single hit, while a visitor requesting a Web page including two images registers as three hits on the

server; one hit is the request for the .html page, and two additional hits are requests for the downloaded image files. While the volume of hits is an indicator of Web server traffic, it is not an accurate reflection of how many pages are being looked at.

Page views are hits to files designated as pages. Supporting graphics and other non-page files are not counted. Within page views, document views are defined as hits to pages that are not dynamic pages or forms. From Table 8.7, document pages took almost one-third of the total pages views and this information is not recorded in the customized log file. "Visits" is the number of times a visitor came to the site. . If a visitor is idle longer than the idle-time limit, WebTrends assumes the visit was voluntarily terminated. If the visitor continues to browse your site after they reach the idle-time limit, a new visit is counted. The default idle-time limit is thirty minutes. Visitor in Table 8.7 means a user who logs into the system, all visits from the same user are counted as just one visitor. In Table 8.6, the corresponding measure is session. However, sessions idle-time limit is two hours. Unique visitors in Table 8.7 mean individuals who visited the web site during the report period. If someone visits more than once, they are counted only the first time they visit. Because a lot of visitors (such as crawlers) only visited the homepage and did not log into the system, the number in Table 8.7 is very different from the number in Table 8.6.

Figure 8.2 and Table 8.9 also show the overall usage of the website based on the raw log, yet from the activity perspective. In table 8.9 and figure 8.2, "Bytes" means number of bytes of data transferred from the server to visitors during the specific time interval.

Table 8.9 gives more clear information on summary of activities every two weeks. One interesting finding from this table is that the bytes transferred from server to visitors are not consistent with page views. For example, during 02/23/2004-03/07/2004, the bytes transferred are 526.85M which is much more than bytes transferred during 03/08/2004-03/21/2004 although page views during 02/23/2004-03/07/2004 are less than 03/08/2004-03/21/2004.

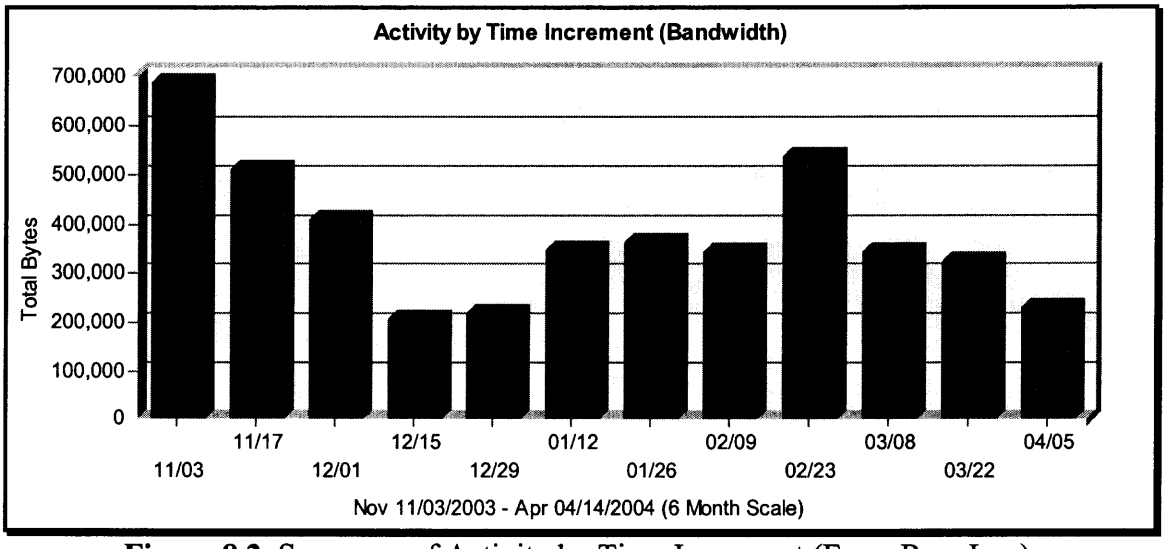

Figure 8.2 Summary of Activity by Time Increment (From Raw Log)

Table 8.9 Summary of Activity by Time Increment

Time Interval	Hits	Page Views	Bytes	Visits
Mon 11/03/2003 – Sun 11/16/2003	81,055	14,565	664,08 M	2,997
Mon 11/17/2003 – Sun 11/30/2003	46,929	10,348	493.15 M	3,149
Mon 12/01/2003 – Sun 12/14/2003	26.416	6,029	397.82 M	2,497
Mon 12/15/2003 – Sun 12/28/2003	16,029	3,963	203.74 M	2,081
Mon 12/29/2003 – Sun 01/11/2004	18,922	4,879	213.04 M	1,792
Mon 01/12/2004 – Sun 01/25/2004	27,830	5,481	337.99 M	2,179
Mon 01/26/2004 – Sun 02/08/2004	30,160	6,370	354.24 M	2,237
Mon 02/09/2004 – Sun 02/22/2004	27,182	6,102	335.06 M	2,138
Mon 02/23/2004 – Sun 03/07/2004	27,926	5,936	526.85 M	2,515
Mon 03/08/2004 – Sun 03/21/2004	32,104	6,420	333.85 M	2,535
Mon 03/22/2004 – Sun 04/04/2004	29,832	7,185	313.88 M	2,422
Mon 04/05/2004 – Sun 04/18/2004	15,599	3,436	224.90 M	1,773
Total	**379,984**	**80,714**	**4,398.65 M**	**28,315**

Figure 8.3 shows the most popular pages on the website based on the raw log. Table 8.10 gives more detailed information on visits to these most popular pages.

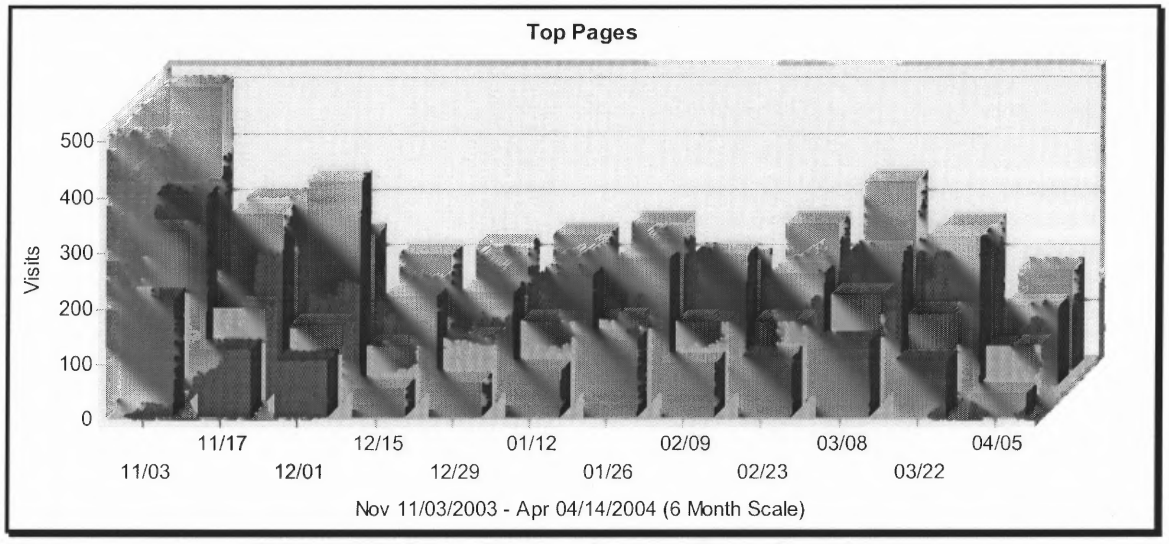

Figure 8.3 Most Popular Ppages (From Raw Log)

Table 8.10 Top Page Statistics

	Pages	Views	% of Total Views	Visits ▼	Avg. Time Viewed
1	**@ ALN WebCenter loading... @** http://alnresearch.org/	3,767	4.66%	3,261	00:01:56
2	http://alnresearch.org/JSP/Empirical_Research/published_research.jsp	8,937	11.07%	2,880	00:01:39
3	http://alnresearch.org/JSP/Empirical_Research/papers_header.jsp	4,226	5.23%	1,911	00:01:06
4	http://alnresearch.org/JSP/redirect.jsp	6,116	7.57%	1,294	00:00:08
5	http://alnresearch.org/JSP/Other_Research/other_papers_header.jsp	2,907	3.60%	1,261	00:01:30
6	http://alnresearch.org/JSP/Empirical_Research/papers_frame.jsp	2,089	2.58%	1,092	00:01:51
7	http://alnresearch.org/JSP/Other_Research/other_papers_frame.jsp	2,064	2.55%	962	00:02:26
8	**Gender Differences in Asynchronous Learning in Higher Education: Learning Styles** http://alnresearch.org/Data_Files/articles/full_text/blum.htm	963	1.19%	908	00:00:27
9	**Structured Interviews** http://alnresearch.org/HTML/AssessmentTutorial/Strategies/StructuredInterviews.html	994	1.23%	893	00:01:15
10	**Pretest/Posttest Comparison** http://alnresearch.org/HTML/AssessmentTutorial/Strategies/PrePostComparison.html	930	1.15%	863	00:02:39
11	**Resources** http://alnresearch.org/Html/resources.htm	1,069	1.32%	849	00:10:28
12	http://alnresearch.org/JSP/People/people_directory_show_profile.jsp	2,257	2.79%	836	00:02:43
13	http://alnresearch.org/JSP/Other_Research/thesis.jsp	1,067	1.32%	755	00:01:37
14	http://alnresearch.org/JSP/Empirical_Research/paper_abstract_frame.jsp	2,147	2.66%	692	00:00:12

Table 8.10 Top Page Statistics (Continued)

Pages	Views	% of Total Views	Visits	Avg. Time Viewed	Pages
15	http://alnresearch.org/index.jsp	1,245	1.54%	652	00:01:24
16	http://alnresearch.org/JSP/Empirical_Research/paper_full_text_frame.jsp	2,041	2.52%	627	00:00:20
17	**Introduction** http://alnresearch.org/HTML/AssessmentTutorial/contents/introduction.html	692	0.85%	620	00:01:24
18	http://alnresearch.org/JSP/user_login.jsp	947	1.17%	605	00:00:36
19	http://alnresearch.org/JSP/Contributions/contributions_frame.jsp	721	0.89%	576	00:05:35
20	http://alnresearch.org/JSP/People/people_directory.jsp	1,853	2.29%	538	00:00:38
Subtotal For the Page Views Above		**47,032**	**58.26%**	**N/A**	**N/A**
Total For the Log File		**80,714**	**100%**	**N/A**	**N/A**

Table 8.10 indicates that besides visits to the homepage, pages under "Empirical Research" are the most popular, which indicates the importance of this feature. Pages under "Tutorials" are also among the most popular pages, and this information was not captured in the customized log since they are not dynamic pages.

Figure 8.4[2] shows the changes of total page views per day over the seven months time from the customized log files.

[2]Here in order to show the scale in smaller units, data from 11/6/2003 (total visits=1127) have been removed as an outlier.

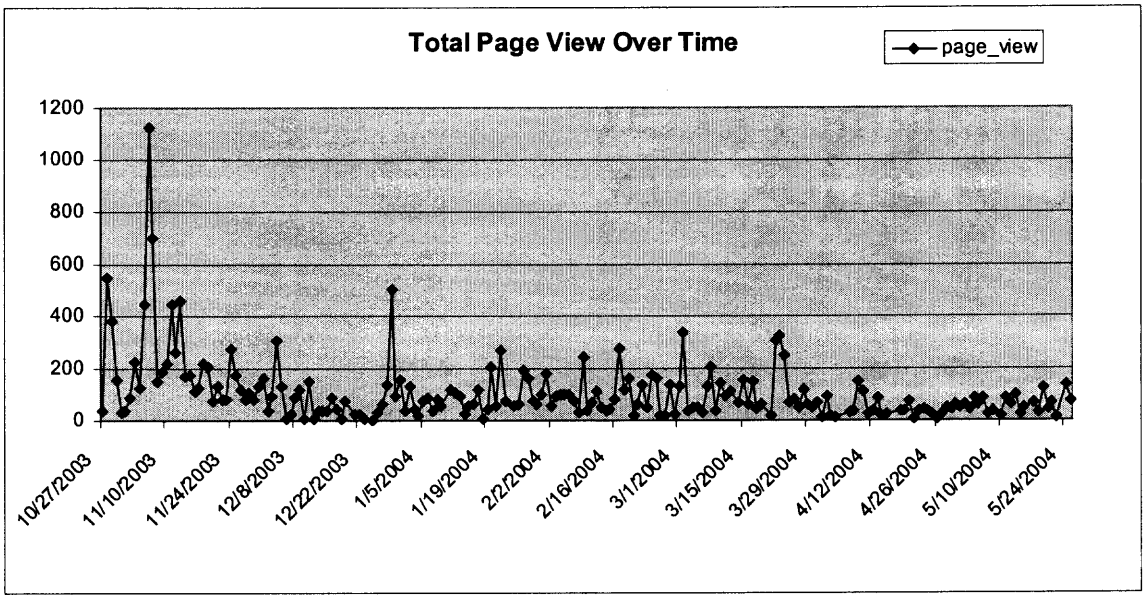

Figure 8.4 Total Page Views Over Time (From customized log)

Figure 8.4 shows that except for several very high hit days, the daily visits to the site are in 0-300 range, and are quite stable over time. Figure 8-5 shows total sessions per day over the observation time from the customized log files.

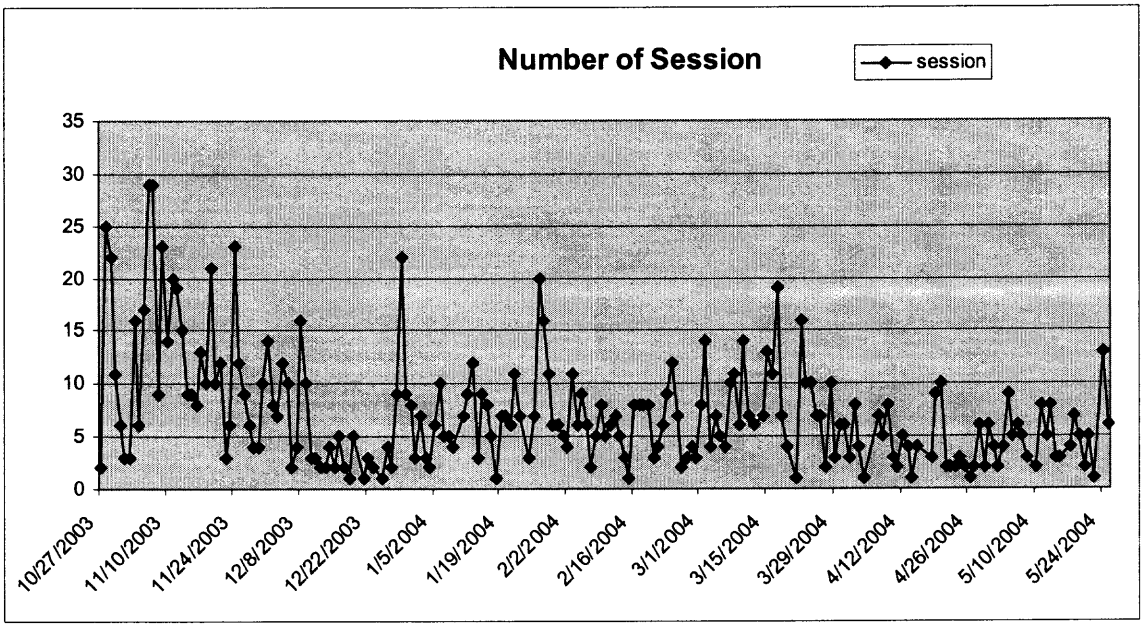

Figure 8.5 Total Sessions Per Day Over Time (From Customized Log)

Figure 8.6 shows the number of articles pages visited per day over the observed time period from the customized log files. It shows that there are accesses to articles everyday, and it falls in a stable pattern for most of the days.

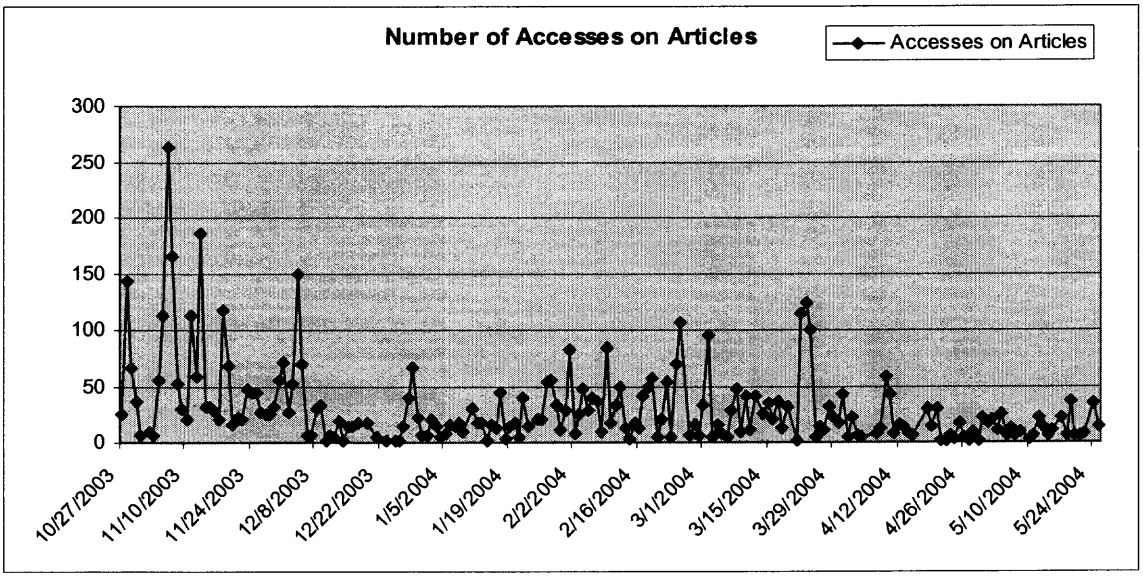

Figure 8.6 Number of Articles Visited Over Time (From Customized Log)

Figure 8.7 shows number of visitors per day over the observed period of time. Again, a stable pattern of visits can be observed here.

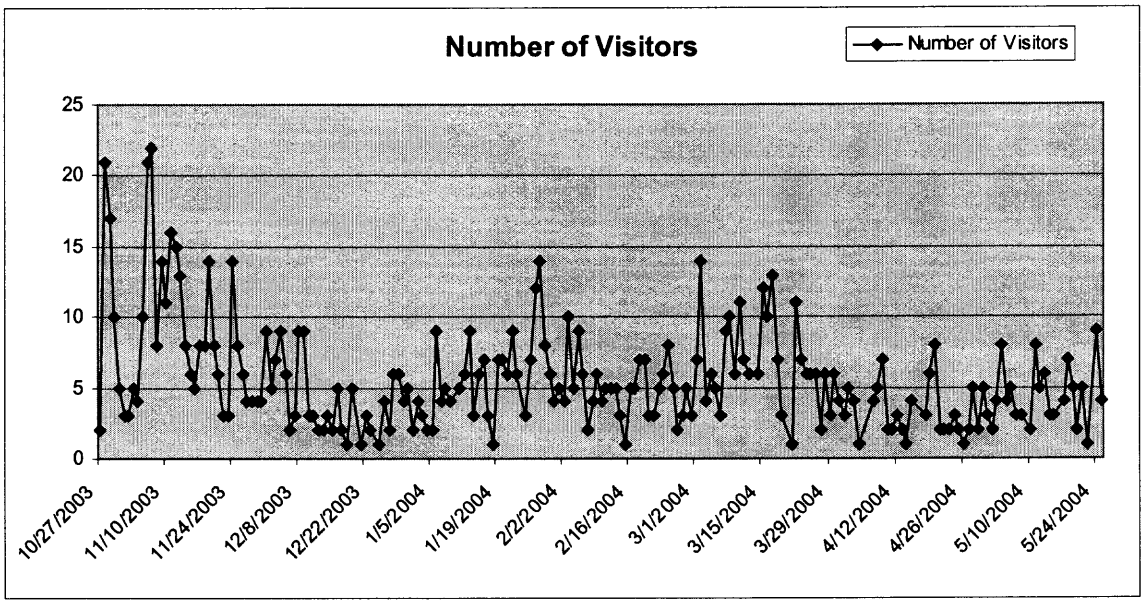

Figure 8.7 Number of Visitors Over Time (From Customized Log)

8.3.2 Visits on Different Content Pages Over Time

In this section, user behavior patterns based on the web site content will be discussed. As described before, all page views can be classified into eight categories according to their content. Figures 8.8 to 8.14 show visits on different categories over time from the customized log files.

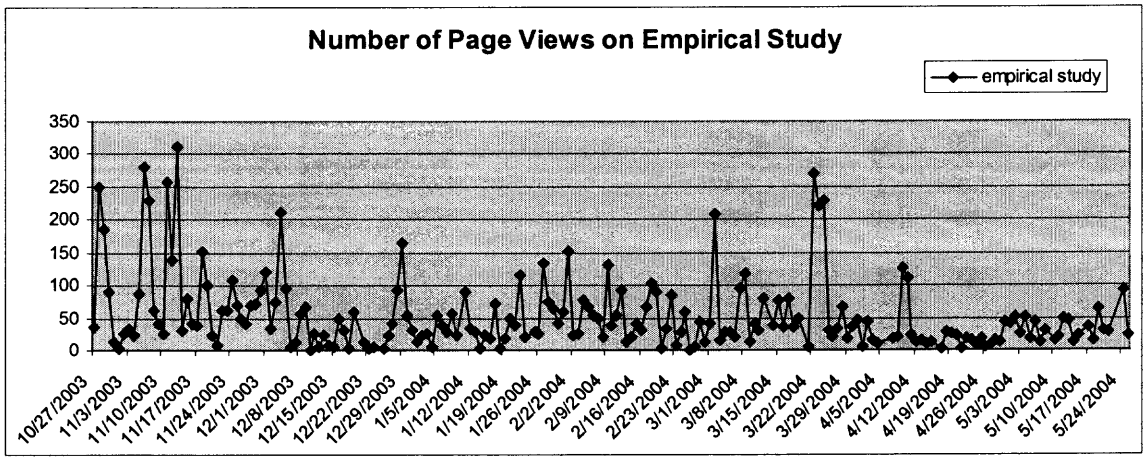

Figure 8.8 Visits of Empirical Studies Over Time (From Customized Log)

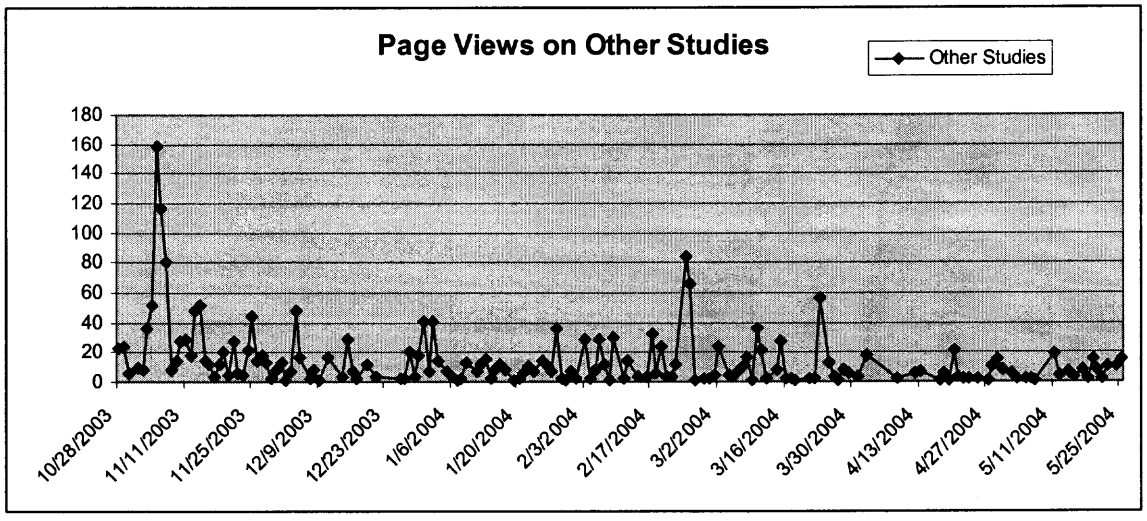

Figure 8.9 Visits of Other Studies Over Time (From Customized Log)

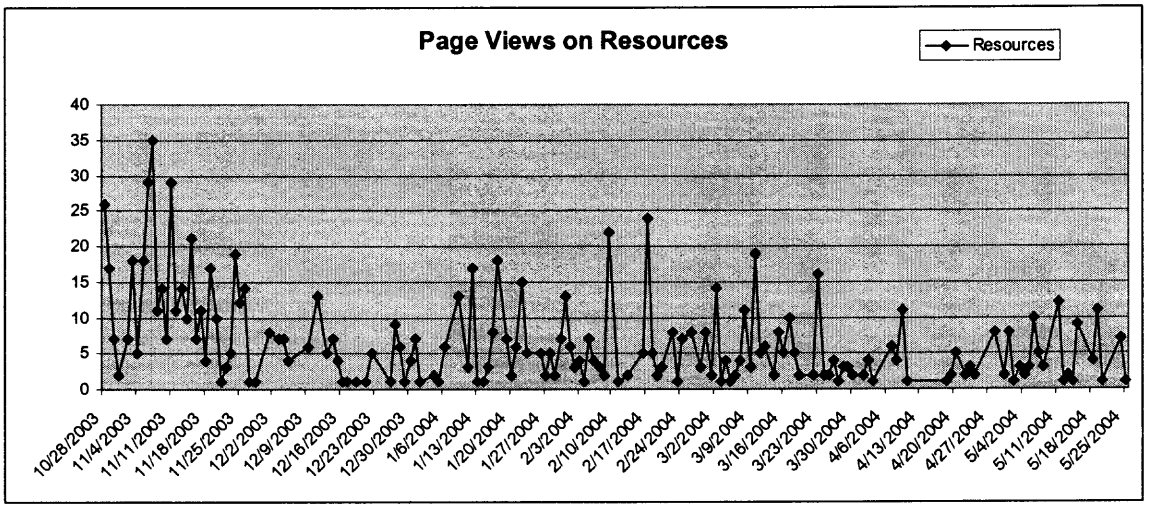

Figure 8.10 Visits on Resources Over Time (From Customized Log)

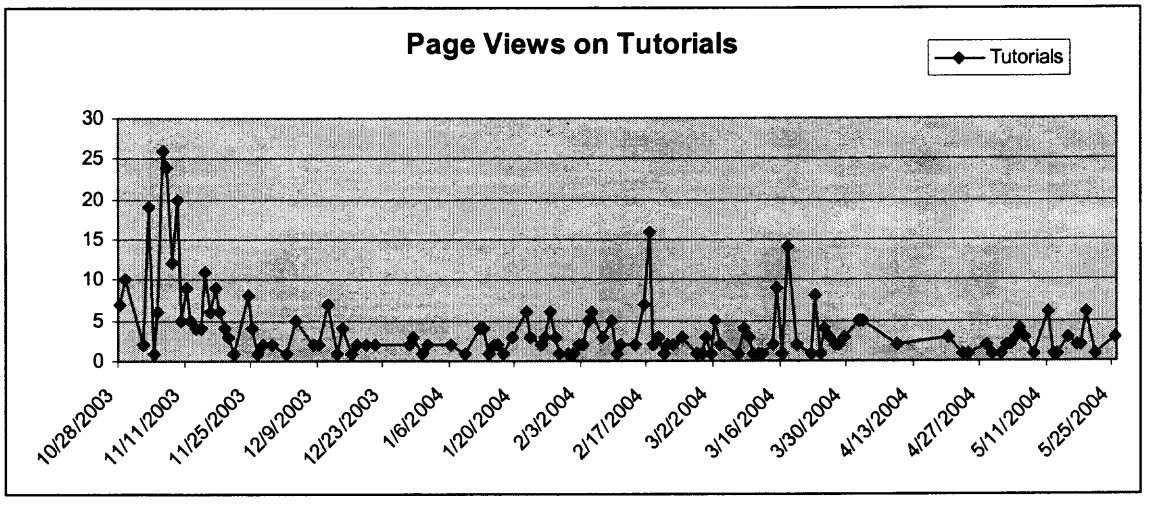

Figure 8.11 Visits on Tutorials Over Time (From Customized Log)

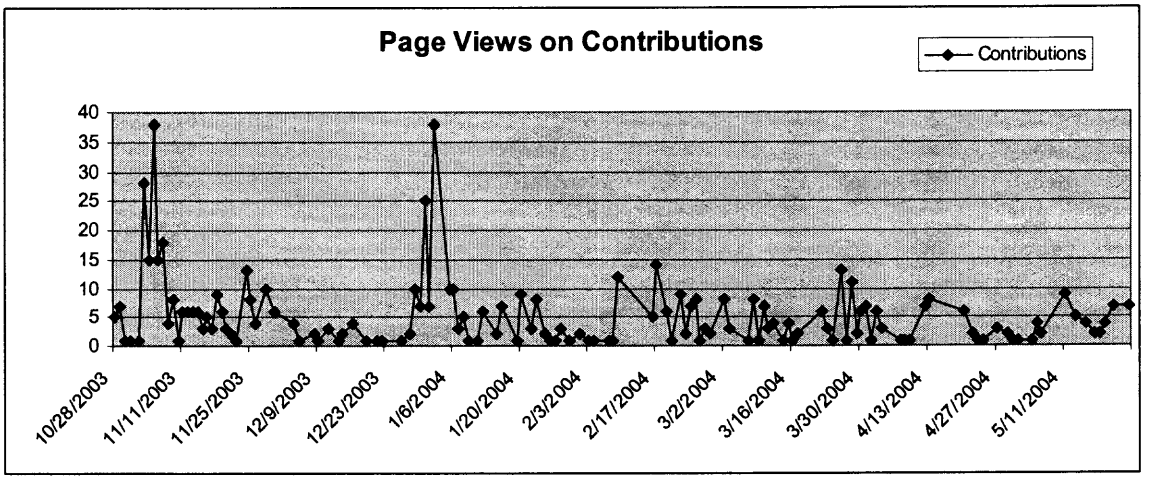

Figure 8.12 Visits on Contributions Over Time (From Customized Log)

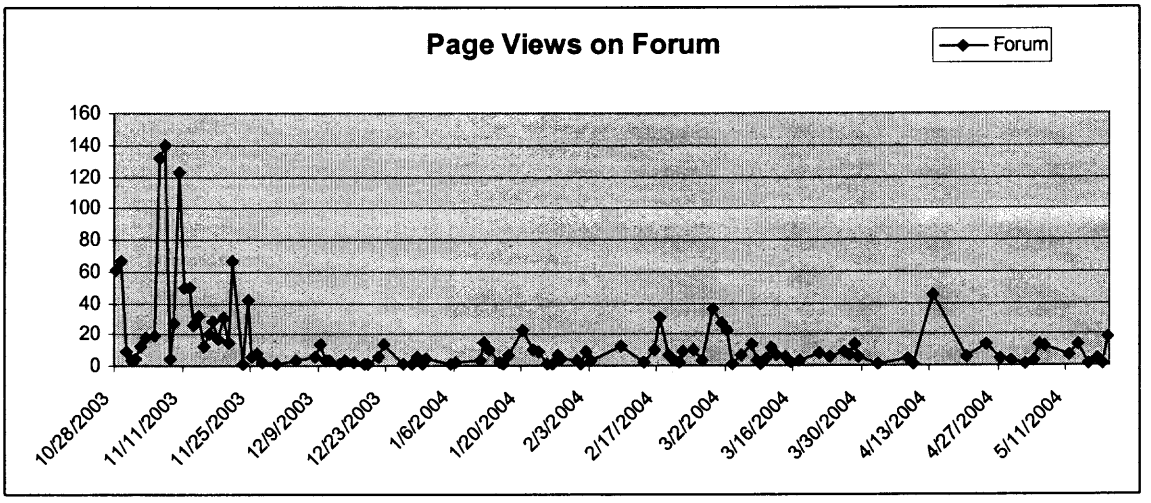

Figure 8.13 Visits on Forums Over Time (From Customized Log)

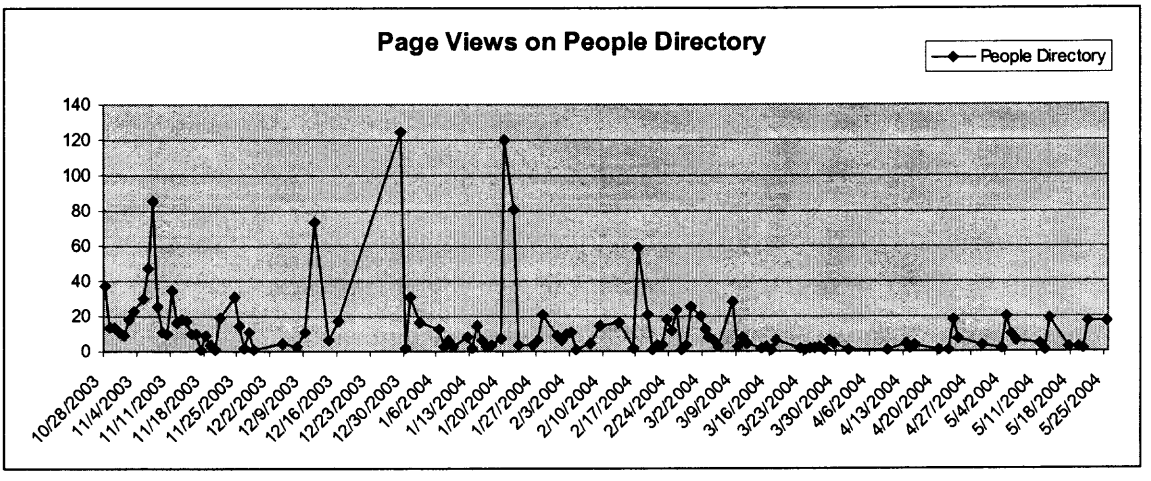

Figure 8.14 Visits on People Directory Over Time (From Customized Log)

From the above diagrams, we can have a general idea of what people are doing when they visit this community. In order to show the distribution of different types of page views visited by users, a pie chart of page view distribution is shown in Figure 8.15.

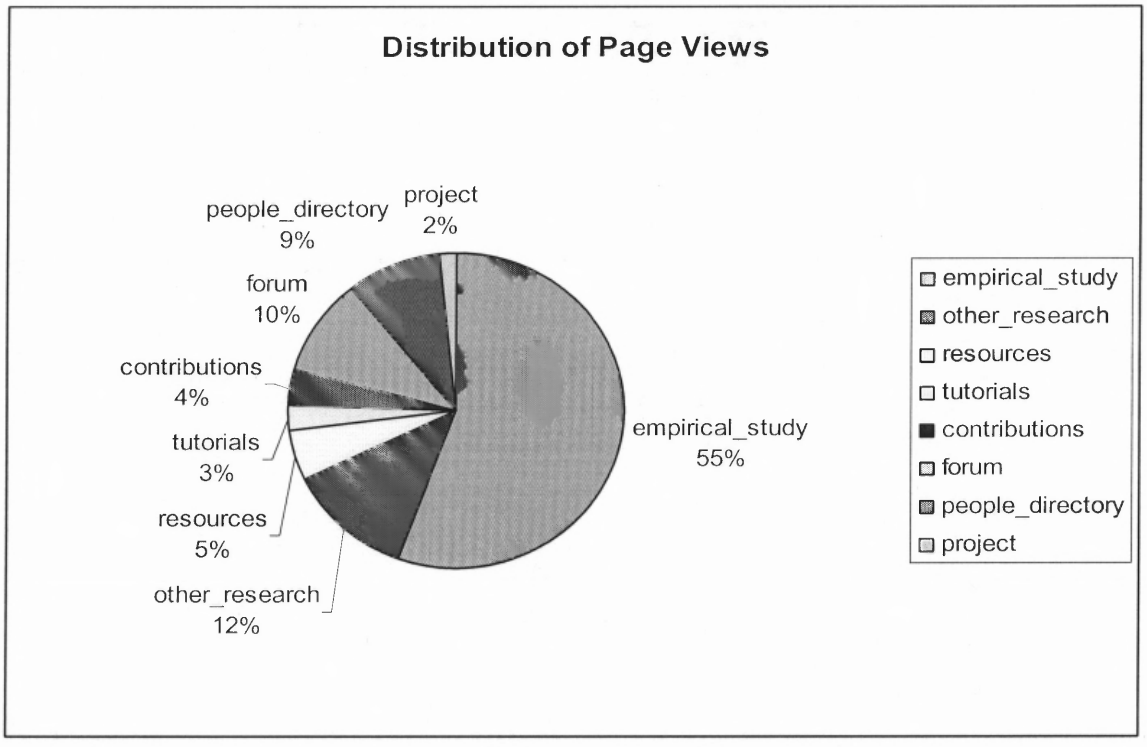

Figure 8.15 Pie Chart of Page View Distribution on Contents (From Customized Log)

From the pie chart, it is very clear that over 50% of the activities fell into the "empirical studies" category, which shows that this is the most important content category. Visits on "other research", "forum", and "people directory" are all at 9%-12% level.

8.3.3 Usage Based on Intended Use of Page Views

As mentioned before, we can also classify page views according to their intended use (see Table 8.4). Table 8.11 shows the total visits counted for each type of page view. Table 8.12 is the total number of visits for the three categories of intended use. Figure 8.16 is the pie chart obtained from the summary data.

Table 8.11 Total Visits for Each Type of Page Views

Category	Page View	Visits	Category	Page View	Visits
convey information	paper_abstract	1838	get information from user	contribute_instruments	685
	paper_comments_display	272		contribute_books	
	paper_db_entry	1096		contribute_links	
	paper_full_text	1830		contribute_others	
	books_abstract	104		contribute_papers	
	other_abstract	136		contribute_studies	
	other_comments_display	103		contribute_thesis	
	other_full_text	302		forum_create_thread	35
	other_summary	95		forum_post_message	82
	thesis_abstract	43		forum_reply	42
	resources_evaluation	69		post_comments	59
	resources_journals	122		maintain	218
	resources_learning_theories	112		home	2119
	resources_methods	96	navigation	published_research	4775
	resources_other	38		forum_search	20
	resources_research_instruments	616		forum_search_results	
	tutorials	395		paper_search_database	865
	tutorials_assessment_tutorial	100		books	235
	forum_messages	576		other	158
	forum_replies	512		qualitative	140
	people_directory	930		thesis	675
	people_show_profile	875		theories	176
	project_people	273		technology	193
	project_people_proposal	37		forum_list	575
	paper_aln_study_content	64			

Table 8.12 Number of Visits Categorized by Intended Use

Category	Number of visits
Convey information	10634
Get information from user	1121
Navigation	9931

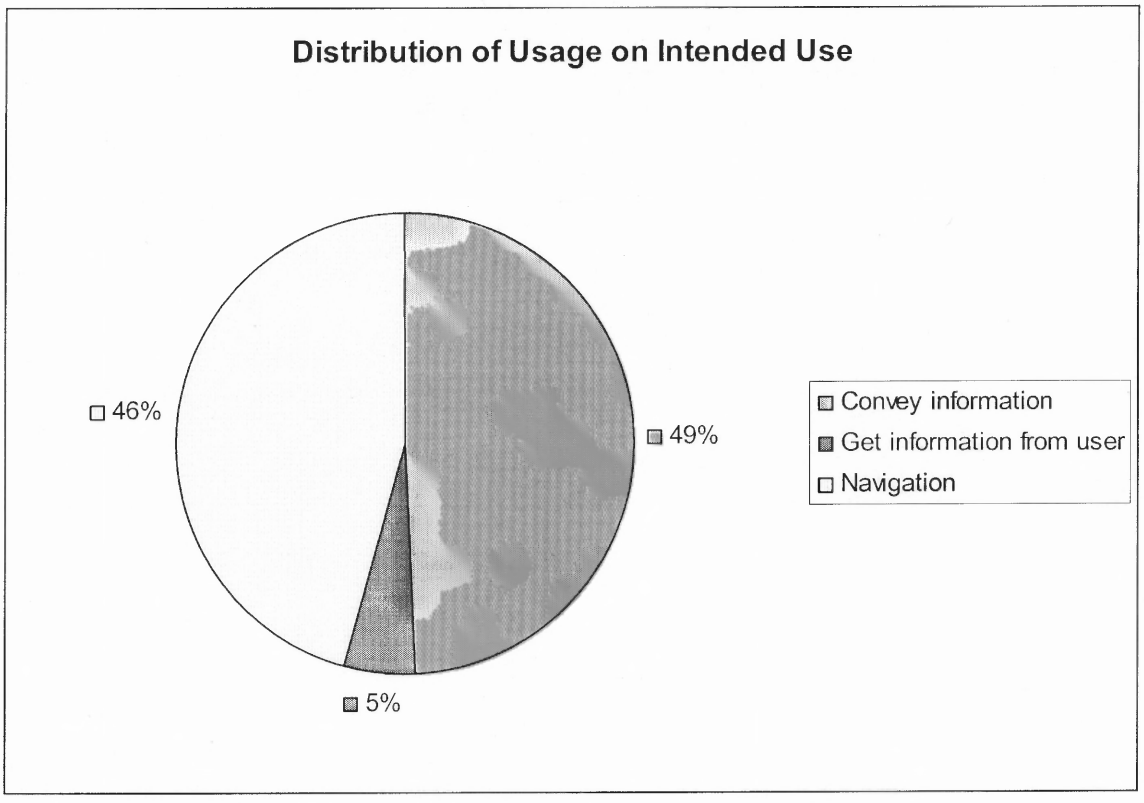

Figure 8.16 Distribution of Usage on Intended Use

According to the pie chart above, half of the visits go to pages that convey information, almost half of the visits go to navigation pages, and only 5% of the visits are from pages that are designed to get information from the user. From the chart, we can see that the navigation pages have been visited very frequently, this also shows how important the navigation strategy is to an information intensive site.

8.3.4 Members' Engagement Over Time

In this section, members' behavior patterns are described from engagement aspects. As mentioned before, we classify all the page views into social engagement and information seeking categories (see Table 8-4 for the classification of the categories). Total number of visits to social engagement pages and information pages can be obtained from Table 8-11. Table 8-13 shows the total page views for social engagement and information seeking.

Figure 8.17 shows the pie chart of the distribution of information seeking and social engagement.

Table 8.13 Number of Visits Based on Engagement

Category	Number of visits
Social Engagement	5077
Information Seeking	16355

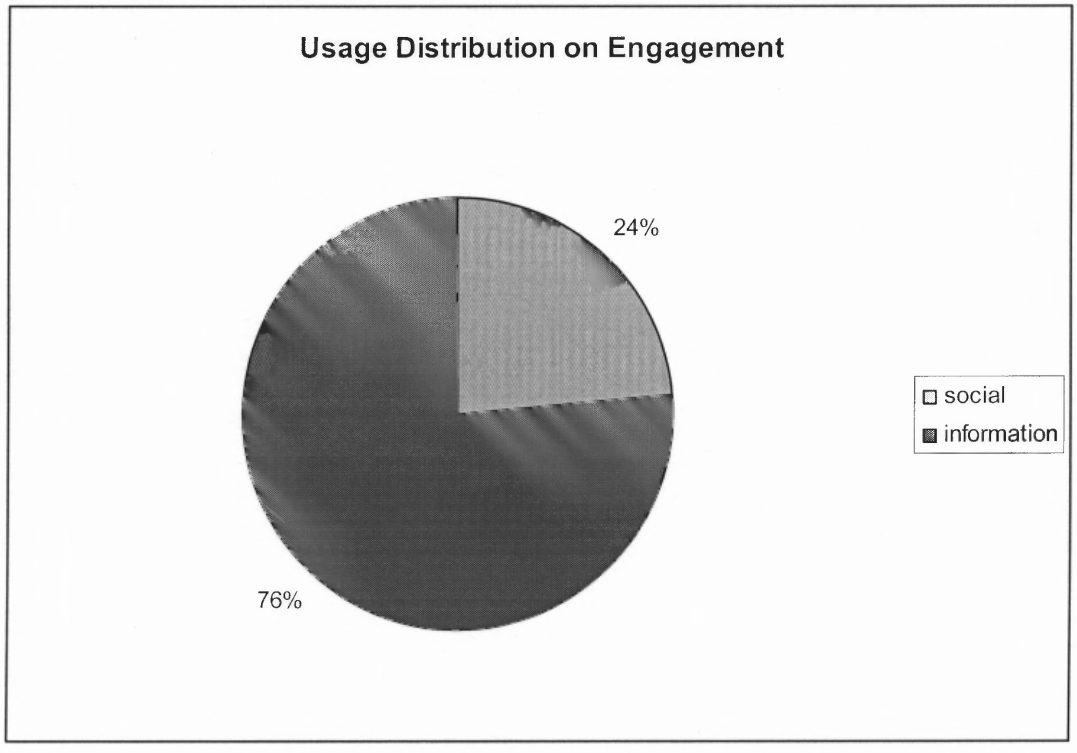

Figure 8.17 Engagement Distribution

From the total number of visits, it is obvious that most users' engagement is focusing on information seeking activities. In order to get a better understanding of engagement, Figure 8.18 shows the comparison of social engagement and information seeking activities on a daily basis.

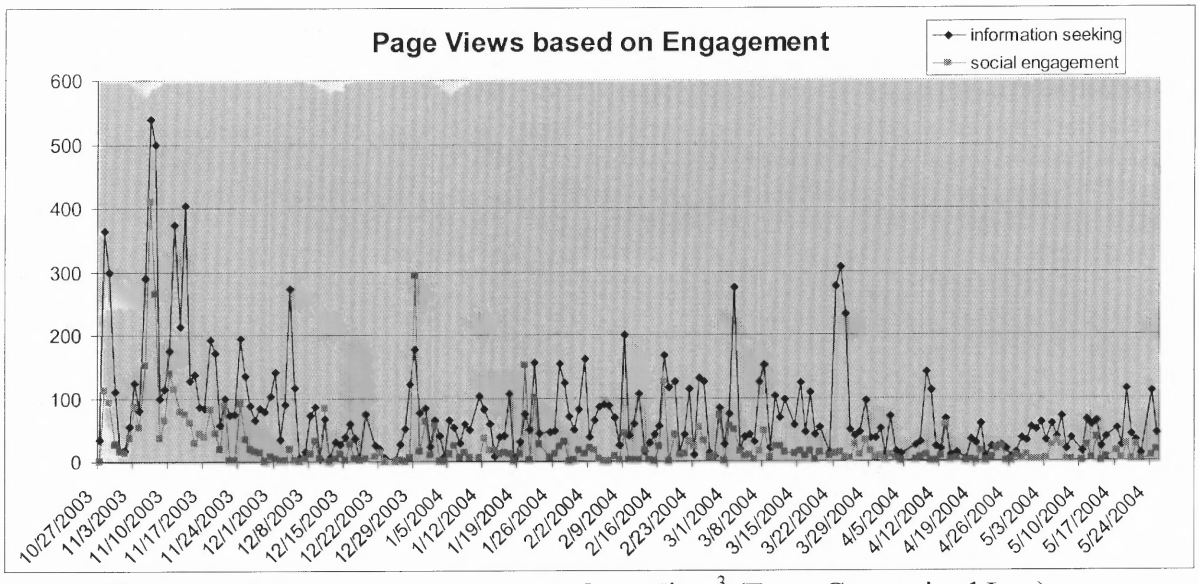

Figure 8.18 Members' Engagement Over Time[3] (From Customized Log)

In order to get a better measurement of engagement, each page view has been assigned a score. The score is decided on the basis of time and energy that users need to spend on a page in order to achieve their goals. Most page views are assigned 1 point, pages that shows db_entry and full text in the information seeking category are assigned 2 points, pages that relate to posting information to the community space are assigned 2 points. After we used the score instead of number of hits, we get the engagement over time as in Figure 8.19.

[3] In order get a better scale, in this diagram, we also remove the data of 11/6/2003 from the source data.

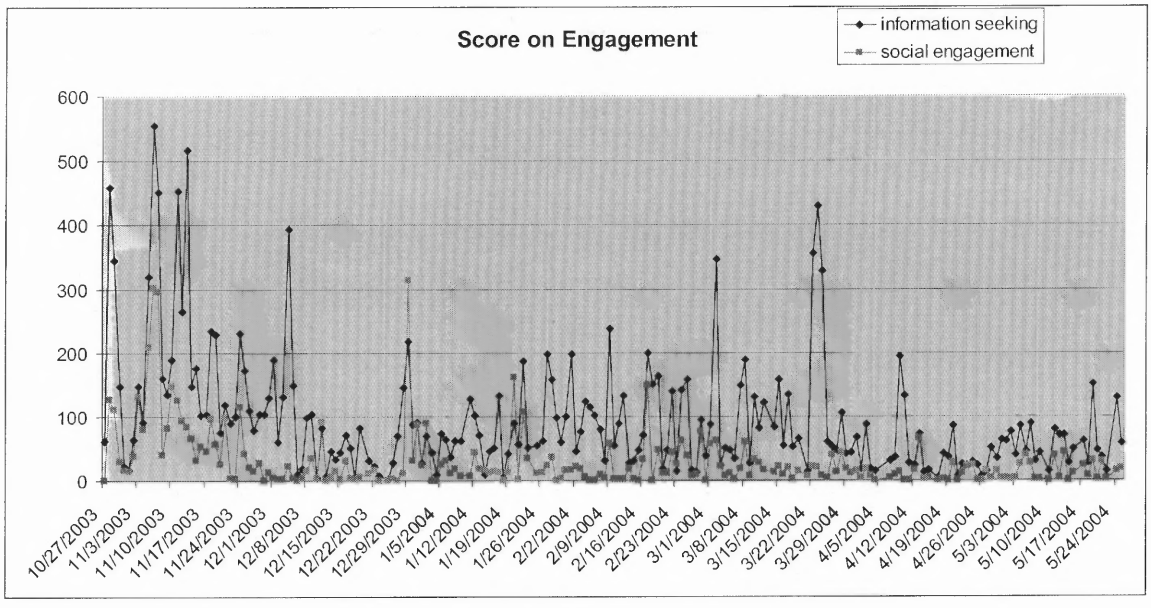

Figure 8.19 Engagement Over Time Based on Score (From Customized Log)

Comparing Figure 8.18 and 8.19, there is not much difference, which indicates that both measurements of engagement are equivalent. We can notice from the chart that except for three dates, engagement in information seeking is always greater than social engagement, which again confirms that the information seeking is dominant in such an online research community. However, this doesn't mean that features designed for social engagement are not useful at all. On most days, social engagement is still observed, although the efforts spent on it are less than efforts spent on information seeking activities.

8.4 Information on Visitors of the Community Site

This section is focused on the characteristics of the visitors of the ALNResearch site. Table 8.14 and Figure 8.20 shows the visitors information by number of visits.

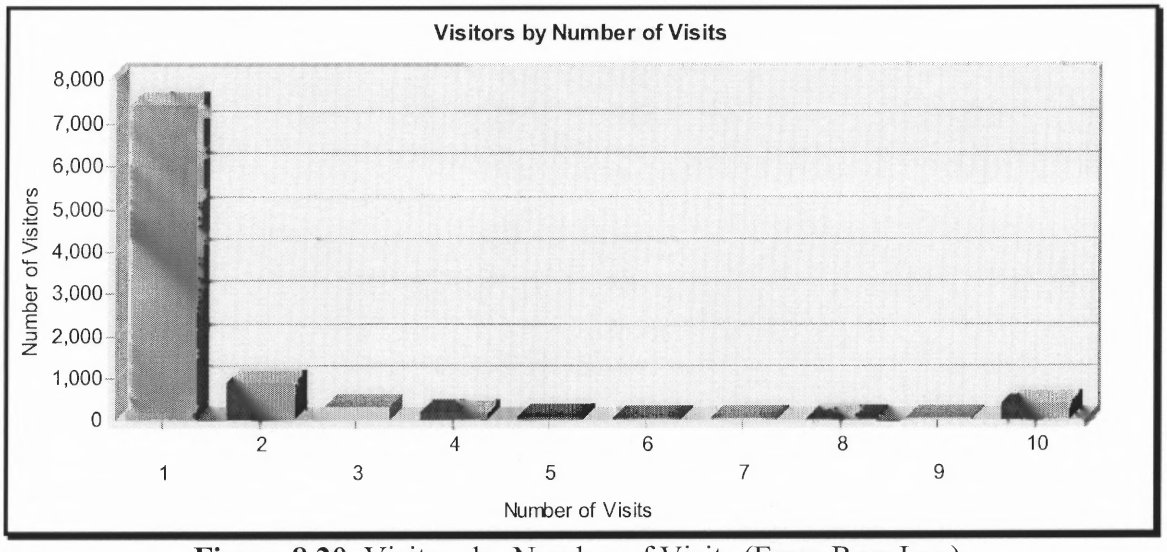

Figure 8.20 Visitors by Number of Visits (From Raw Log)

Table 8.14 Visitors by Number of Visits

Number of Visits	Number of Visitors	% of Total Unique Visitors
1 visit	7446	76.47%
2 visits	911	9.35%
3 visits	351	3.60%
4 visits	187	1.92%
5 visits	129	1.32%
6 visits	86	0.88%
7 visits	80	0.82%
8 visits	85	0.87%
9 visits	54	0.55%
10 or more visits	407	4.18%

Figure 8.20 and table 8.14 shows that the majority of visitors only pay one visit to the site, which means they are either crawlers or people who happen to access the homepage. Visitors who visit multiple times are more likely to be members. It can also be found from Table 8.14 that visitors either only pay 1-2 visits to the site, or they visit more than 10 times during the observation period (407 visitors). Those users who do more than 10 visits are members of the WebCenter.

Figure 8.21 and Table 8.15 shows the geographic region information about visitors.

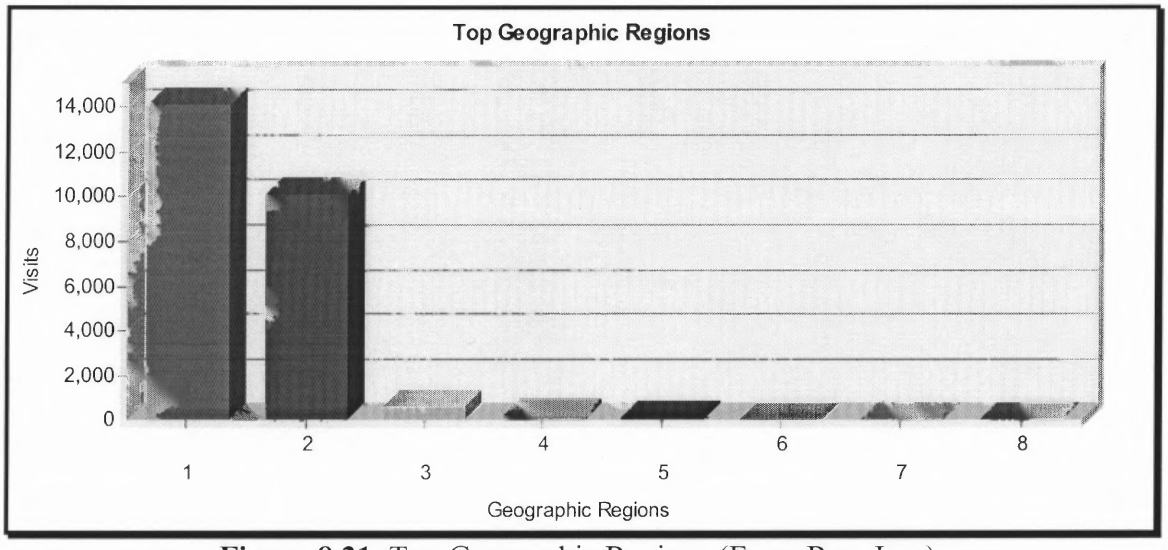

Figure 8.21 Top Geographic Regions (From Raw Log)

Table 8.15 Top Geographic Regions

	Geographic Regions	Visits
1	North America	14,124
2	Region Unspecified	10,132
3	Western Europe	655
4	Asia	252
5	Australia	140
6	Middle East	73
7	Northern Europe	57
8	Eastern Europe	47
9	Pacific Islands	36
10	South America	27
11	Region Not Found	21
12	Southern Africa	15
13	Eastern Africa	13
14	Central America	8
15	Caribbean Islands	7
Total for the Geographic Regions above		**25,607**

The above figures and tables show that the ALNResearch website attracted visitors from all over the world. Although the majority visits are from North America, there are still lots of visits from Europe and Asia which shows that it actually connects researchers all over the world who are interested in the topic.

Figure 8.22 and Table 8.16 show visitors by time increment. It shows information for total users and unique users during the specific time period.

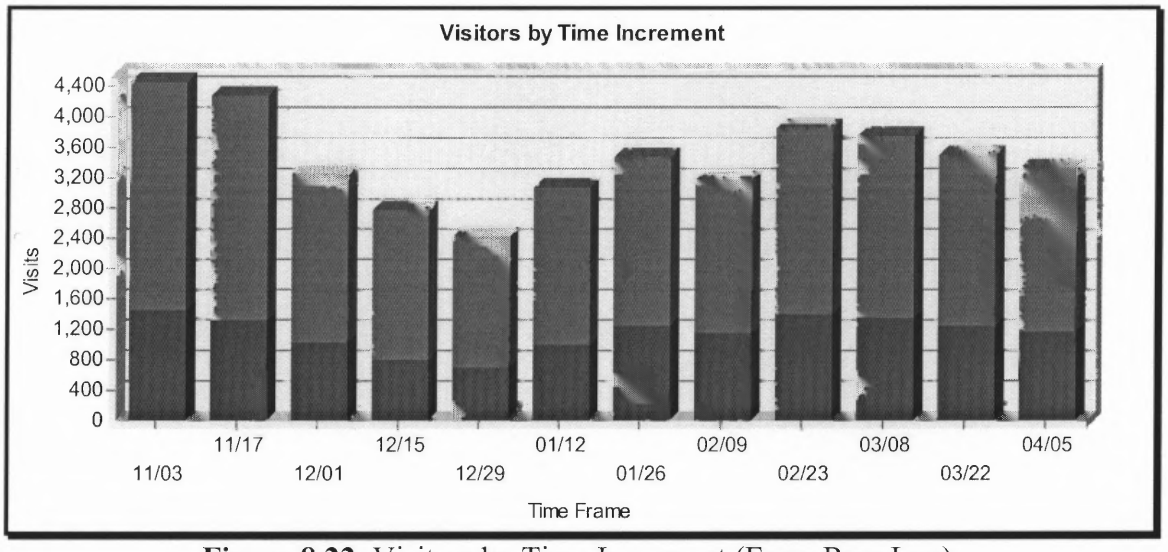

Figure 8.22 Visitors by Time Increment (From Raw Log)

Table 8.16 Statistics on Visitors by Time Increment

Time Frame	Visitors	Unique Visitors	First-Time Visitors	Average Visit Length	Visitor-Minutes
Mon 11/03/2003 - Sun 11/16/2003	3,021	1,453	0	00:04:04	12324:04
Mon 11/17/2003 - Sun 11/30/2003	2,950	1,337	0	00:03:02	8975:17
Mon 12/01/2003 - Sun 12/14/2003	2,215	1,040	0	00:03:30	7764:04
Mon 12/15/2003 - Sun 12/28/2003	1,956	819	0	00:03:43	7293:30
Mon 12/29/2003 - Sun 01/11/2004	1,719	711	0	00:03:12	5518:31
Mon 01/12/2004 - Sun 01/25/2004	2,086	1,003	0	00:02:57	6155:25
Mon 01/26/2004 - Sun 02/08/2004	2,216	1,261	0	00:02:16	5057:05
Mon 02/09/2004 - Sun 02/22/2004	2,014	1,167	0	00:02:43	5503:47
Mon 02/23/2004 - Sun 03/07/2004	2,443	1,412	0	00:02:04	5072:50
Mon 03/08/2004 - Sun 03/21/2004	2,399	1,368	0	00:02:27	5896:13
Mon 03/22/2004 - Sun 04/04/2004	2,254	1,265	0	00:02:59	6743:27
Mon 04/05/2004 - Sun 04/18/2004	2,138	1,194	0	00:02:48	6009:37
Averages	**2,284**	**1,169**	**0**	**00:02:58**	**6859:29**
Totals	**27,411**	**14,030**	**0**	**00:35:45**	**82313:50**

Figure 8.22 also shows that number of visitors also becomes quite stable after 1/20/2004. Table 8.16 shows both total number of visitors and number of unique visitors during the specific period.

Before the implementation of the new system (before 10/28/2003), the ALNResearch community had 184 registered members. As of 5/25/2004, there are 765 registered members in the community, which means during the six months after the implementation of the new system, the community size increased dramatically. However, most of the members do not visit the community place very often.

8.5 Summary

In this chapter, the research question; "what do people do in online research communities?" has been answered through analyzing members' behavior patterns in an online research community over a four month period of time. The following findings summarize the analysis:

- The ALNResearch WebCenter has been regularly visited by members, although only a small to moderate number (mean of 6 or 176, depending on what one considers a "visit") per day. This might be due to the small size of the community and the fact that new entries to the data base are not contributed very often. It may also be because this is almost entirely a "pull" web site, with little "push" though email notification to notify members when new items of interest may have been added.

- The main purpose of the ALNResearch WebCenter is to disseminate a new field of knowledge to its members and to enable them to use this knowledge in improving their research. From the summary data, there are all together 5,143

accesses to articles over about six months, which is a reasonable number considering the number of visitors. From the distribution of visits based on the content of the page, about 50% of the visits are on "empirical studies" and 12% of the visits are on "other studies," which are the most important contents of this community. From the analysis of members' engagement, 76% of the engagement is focused on information seeking activities. All these show that the ALNResearch WebCenter achieves its purpose of distributing knowledge to members.

- Although social engagement seems to be a much smaller part compared to information sharing, visits to "discussion forums" and the "people directory" still take 20% of the total visits, which cannot be ignored. It also indicates that people want to know who else is also in the community, and what other people said in the community space.

CHAPTER 9

INDEX VALIDATION

In this chapter, descriptive data and correlations between different constructs will be analyzed from the survey responses. First the demographic information will be described, followed by the factor analysis for the survey data, and finally the correlations between constructs and dependent variable "real usage" will be presented.

9.1 Demographic Information

In the survey questions, the demographic related information includes gender, role in ALN field, connection speed, and hours spent on the Internet per day. Table 9.1 shows basic descriptive data for these variables:

Table 9.1 Descriptive Data for Demographic Information

Characteristics	Type	Frequency	Percent
Gender	Female	43	52.44%
	Male	39	42.76%
Roles in ALN	Researcher	51	62.19%
	Educator	46	56.10%
	Student	33	40.24%
	Administrator	29	35.36%
	Other	7	8.54%
Internet Connection	Modem (56K)	5	6.1%
	DSL/Cable	39	47.56%
	LAN/T1	38	46.34%
Hours spent on the Internet per day	Less than 2	5	6.1%
	2-4	18	21.95%
	4-6	24	29.27%
	more than 6	35	42.68%

In the questionnaire, when asked about "roles in ALN", users can choose as many roles as apply. Thus the cumulative percent for this question is more than 100%. The above table shows that gender of the subjects who answered the survey is roughly balanced, the majority of the subjects use high speed Internet connections, and most of

them spend more than 4 hours online per day. The majority are engaged both research and education on ALN.

9.2 Descriptive Data on Self Report Usage and Value

This section will show the descriptive results from questionnaire questions that related self-reported use frequency and perceived usefulness of different system features (see Table 9.2 and 9.3 for results).

Table 9.2 Descriptive Results of Self-reported Use Frequency on Different System Features

System Feature	Answer Category	Number	Percentage
Empirical Studies	Never	3	4.11
	Occasionally	42	57.53
	Frequently	28	38.36
Other Studies	Never	13	18.06
	Occasionally	47	65.28
	Frequently	12	16.67
Research Instrument	Never	15	20.83
	Occasionally	46	63.89
	Frequently	11	15.28
Research Resources	Never	8	10.96
	Occasionally	46	63.01
	Frequently	19	26.03
Tutorial	Never	41	56.16
	Occasionally	26	35.62
	Frequently	6	8.22
People Directory	Never	30	41.10
	Occasionally	39	53.42
	Frequently	4	5.48
Discussion Forum	Never	43	60.56
	Occasionally	25	35.21
	Frequently	3	4.23
Contributions	Never	46	63.99
	Occasionally	24	33.33
	Frequently	2	2.78

Table 9.2 indicates that except for "Contribution", other features are used occasionally by most users . More than half of the subjects use the "people directory" frequently, which gives a strong indication of social needs of members in the ALNResearch community.

Table 9.3 Descriptive Results of Self-reported Use Frequency on Different System Features

System Feature	Answer Category	Number	Percentage
Empirical Studies	Can't Say	4	5.41
	Useless	0	0
	Slightly Useful	7	9.46
	Fairly Useful	34	45.95
	Extremely Useful	29	39.19
Other Studies	Can't Say	13	18.06
	Useless	0	0
	Slightly Useful	10	13.89
	Fairly Useful	37	51.39
	Extremely Useful	12	16.67
Research Instruments	Can't Say	14	19.18
	Useless	1	1.37
	Slightly Useful	11	15.07
	Fairly Useful	28	38.36
	Extremely Useful	19	26.03
Research Resources	Can't Say	6	8.33
	Useless	1	1.39
	Slightly Useful	8	11.11
	Fairly Useful	39	54.17
	Extremely Useful	18	25.00
Tutorials	Can't Say	31	43.06
	Useless	1	1.39
	Slightly Useful	15	20.83
	Fairly Useful	20	27.78
	Extremely Useful	5	6.94
People Directory	Can't Say	27	36.99
	Useless	0	0
	Slightly Useful	21	28.77
	Fairly Useful	22	30.14
	Extremely Useful	3	4.11
Discussion Forum	Can't Say	33	46.48
	Useless	3	4.23
	Slightly Useful	14	19.72
	Fairly Useful	18	25.35
	Extremely Useful	3	4.23
Contributions	Can't Say	39	54.17
	Useless	1	1.39
	Slightly Useful	12	16.67
	Fairly Useful	18	25.00
	Extremely Useful	2	2.78

Table 9.3 indicates that all systems features have been perceived at some level of usefulness. Among all the features, more people rated "Empirical Studies" as extremely useful than other features. There are a lot of ratings of "Can't Say" that indicate that the

subjects haven't had enough experience with the corresponding features to give a clear answer.

9.3 Questionnaire Scale Validation

Factor analysis was used to assess construct validity of the survey questionnaire items. For rotation method in factor analysis, PROMAX has been chosen to carry out the evaluation. PROMAX is one of the oblique methods, which can be used when there are correlations between constructs. It is also the most commonly used rotation method for this type of construct. For this study, as all the questionnaire items are related to users' intention to continue use, PROMAX is the appropriate rotation method for factor analysis. Table 9.4 shows the initial loadings after a rotation of PROMAX with all items.

Table 9.4 Initial Factor Loadings With All Items

	F1	F2	F3	F4	F5	F6	F7
Q5	0.53						
Q8	0.80						
Q14	0.54						
Q11		0.57					
Q9	0.96						
Q24	0.42	0.44					
Q1	0.43						0.42
Q20			0.46				
Q2	0.40						
Q12			0.46				
Q19			0.68				
Q7		0.51					
Q21		0.67					
Q6						0.68	
Q15						0.71	
Q23						0.31	
Q4					0.58		
Q18					0.82		
Q22							
Q13			-0.41			0.37	
Q16				0.80			
Q17				0.42			
Q3							0.68
Q10				0.41			

Q3, Q13, and Q22 are removed to reduce the number of factors and to increase interpretability of the factors. Table 9.5 shows the factor loadings for the finalized items which will be included for further analysis. Table 9.6 shows the variance explained by these factors.

Table 9.5 Factor Loadings for Finalized Items

	F1 (perceived usefulness)	F2 (intention to continue use)	F3 (perceived ease of use)	F4 (result demonstrability)	F5 (sense of community)	F6 (preliminary networking)
Q5	0.53					
Q8	0.82					
Q14	0.47					
Q9	0.86					
Q1	0.68					
Q2	0.49					
Q10	0.52					
Q11		0.51				
Q7		0.62				
Q21		0.70				
Q6			0.73			
Q15			0.74			
Q23			0.67			
Q12				0.73		
Q19				0.45		
Q20				0.34		
Q24				0.54		
Q4					0.58	
Q18					0.69	
Q16						0.70
Q17						0.53

Table 9.6 Variance Explained by the Factors

Factor	Variances explained
F1 (Perceived Usefulness)	5.90
F2 (Intention to continue use)	3.65
F3 (Perceived Ease of Use)	4.59
F4 (Result Demonstrability)	4.03
F5 (Sense of Community)	3.65
F6 (Attitude)	1.99
Final Communality Estimates	12.04

The level of internal consistency reliability of the constructs is tested using Cronbach's alpha. (See Table 9.7 for results). Normally, constructs with Cronbach's

alpha greater than 0.7 are considered reliable. Although Cronbach's alpha of "sense of community" and "Attitude" did not reach 0.7, they are close enough to 0.7 to be accepted as a reliable constructs for further analysis.

Table 9.7 Scale Reliability

Constructs	Cronbach Alpha	Reliability
Perceived Usefulness	0.88	Reliable
Intention to continue use	0.73	Reliable
Perceived Ease of Use	0.79	Reliable
Sense of Community	0.63	Reliable
Result Demonstrability	0.72	Reliable
Attitude	0.67	Reliable

After validity and reliability checks, there are five constructs left from questionnaire items for further analysis. Table 9-8 shows the questionnaire items that remain for further analysis.

Table 9.8 Questionnaire Items to Be Analyzed

Measure	Item Number	Question / Statement
Perceived Usefulness	1	I am satisfied with the quality of the ALNResearch WebCenter content.
	2	Overall, I feel the benefits of using the WebCenter system are apparent to me.
	5	The WebCenter has provided me with guidance, reference, or other information useful to my work.
	8	The WebCenter has broadened my understanding of my research area.
	9	I can get materials that are important to me from the WebCenter.
	10	The WebCenter helps me know who else is doing research in my research field.
	14	Using the ALNResearch WebCenter improves my work.
Intention to continue use	7	If I can access the WebCenter, I will continue to use it.
	11	I find the ALNResearch WebCenter is NOT useful at all for my work.
	21	I DO NOT plan to use WebCenter any more.
Perceived Ease of Use	6	The interface of the ALNResearch WebCenter is clear and understandable.
	15	I found it easy to complete what I want to do in the ALNResearch WebCenter.
	23	It requires significant effort to use the ALNResearch WebCenter features.
Sense of Community	4	I trust other members in the ALNResearch WebCenter.
	18	I always feel isolated in the ALNResearch WebCenter.

Table 9.8 Questionnaire Items to Be Analyzed (Continued)

Measure	Item Number	Question / Statement
Result Demonstrability	12	It is difficult for me to explain why using the WebCenter may or may not be beneficial.
	19	I have no difficulty telling others about the results of using the WebCenter system.
	20	The Quality of the information I get from the WebCenter is high
	24	The ALNResearch WebCenter has something useful for my work.
Attitude Towards Online Environment	16	I DO NOT plan to post anything on the site or communicate with anyone through the WebCenter.
	17	For me, the WebCenter is JUST a place where I can get information.

9.4 Test of Normality

Normal distribution is the assumption for parametric measures like Pearson's R, T-Test, or ANOVA test. This section will use the goodness of fit test to measure the normality of the sample data. There are several methods for testing normality, for example, SAS software provides the Shapiro-Wilk test, the Kolmogrov-Smirnoc test, Cramer Von Mises test, and Anderson-Darling test. According to the SAS manual, if the sample size is over 2000, the Kolmogrov test should be used. If the sample size is less than 2000, then the Shapiro test is better. However, some researchers argue that the Shapiro test was originally designed for a sample size up to 50. To examine the normality for a sample size between 51 and 1999, other methods like the Anderson-Darling test are recommended (Yu, 1997). As the sample size for this research is 74, theAnderson-Darling test is chosen (see Table 9.9 for results).

Table 9.9 Results of the Anderson-Darling Test

Measure	Anderson_Darling A-Sq	Pr>A-Sq	H0: Sample distribution fits into normal distribution (at significance level p=0.05)
Perceived Usefulness	0.58	.122	Not rejected
Intention to continue use	5.86	<.005	Rejected
Perceived Ease of Use	1.17	<.005	Rejected
Sense of Community	2.09	<.005	Rejected
Result demonstrability	1.17	<.005	Rejected
Attitude	1.76	<.005	Rejected
Number of pages visited	2.92	<.005	Rejected
Information Seeking Score	1.79	<.005	Rejected
Social Engagement Score	11.91	<.005	Rejected

From the above table, the only variable that follows normal distribution is "Perceived Usefulness". All other variables are not validated for the normality assumption, which means either these variables need to be transformed in order to fit normality, or non-parametric measures have to be used for further data analysis.

9.5 Summary

This chapter shows descriptive data on demographic information obtained from the survey. It also checks the validity, reliability, and normality of all variables for further analysis. After factor analysis using PROMAX rotation, some of the original designed constructs are combined, and some of them are split into different constructs. Among questionnaire constructs, "Output Quality", "Preliminary Networking", and "Expectancy relevance" have to be removed because they haven't appeared as a separate construct from the factor analysis. Based on normality test, only "Perceived Usefulness" follows a normal distribution. All other variables failed the normality distribution, which indicates that further analysis needs to use non-parametric measures or the raw data needs to be transformed to fit the normality assumption for parametric measures. Table 9.10 summarizes hypotheses that can be tested after the validity and reliability test.

Table 9.10 Validated Measures

Hypotheses	Variables	Measurement
H3a. Result demonstrability will have a positive effect on perceived usefulness of a community place.	Result demonstrability Perceived Usefulness	Questionnaire Items
H3b. Result demonstrability will have a positive effect on perceived ease of use of a community place.	Result demonstrability Perceived Ease of Use	Questionnaire Items
H3c. Result demonstrability will have a positive effect on intention to continue use a community place.	Result demonstrability Intention to continue use	Questionnaire Items
H3d: Result demonstrability will have a positive effect on number of hits	Result demonstrability Number of Hits	Questionnaire Items
H4a: People who think the community place is just for information will have less social engagement activities then people who do not agree with it.	Attitude Social engagement	Questionnaire items
H4b. People's attitude towards online community will have a positive effect on perceived usefulness of the community place.	Attitude Perceived Usefulness	Questionnaire items
H4c. People's attitude towards online community will have a positive effect on intention to continue using of the community system	Attitude Perceived Ease of Use	Questionnaire items
H4d: People who think the community place is just for information will have a lower level of perceived sense of community.	Attitude Sense of Community	Questionnaire items
H4e: People who think the community place is just for getting information will use the awareness tools less than others.	Attitude Self report usage on system features	Questionnaire items
H4f: People who think the community place is just for getting information will have a lower perceived usefulness of awareness tools.	Attitude Perceived usefulness on system features	Questionnaire items
H5a: Perceived sense of community will have a positive effect on perceived usefulness of the community place.	Sense of Community Perceived Usefulness	Questionnaire Items
H5b: Perceived sense of community will have a positive effect on perceived ease of use of the community system.	Sense of Community Perceived Ease of Use	Questionnaire Items
H5c: Perceived sense of community will have a positive effect on intention to continue using the community place.	Sense of Community Intention to continue use	Questionnaire Items
H5d: Perceived sense of community will have a positive effect on social engagement activities.	Sense of Community Social Engagement	Questionnaire items Web Server Logs
H6a: Perceived usefulness will have a positive effect on intention to continue using the community place.	Perceived Usefulness Intention to continue use	Questionnaire Items
H6b: Perceived ease of use will have a positive effect on intention to continue using the community place.	Perceived Ease of Use Intention to continue use	Questionnaire Items

Table 9.11 Validated Measures (Continued)

Hypotheses	Variables	Measurement
H6c: Perceived ease of use will have a positive effect on perceived usefulness	Perceived Ease of Use Perceived Usefulness	Questionnaire Items
H6d: Perceived usefulness will have a positive effect on number of pages visited	Perceived Usefulness Number of pages visited	Questionnaire Items Web Server Log
H6e: Perceived ease of use will have a positive effect on number of pages visited	Perceived Ease of Use Number of pages visited	Questionnaire Items Web Server Log
H6f: Intention to continue use will have a positive effect on number of pages visited	Intention to continue use Number of pages visited	Questionnaire Items Web Server Log

The next chapter will focus on testing hypotheses and answering research questions. Chapter eleven will use the Partial Least Square (PLS) method to test the integrated research model that is presented in Chapter six.

CHAPTER 10

BIVARIATE CORRELATION ANALYSIS AND ANSWERING RESEARCH QUESTIONS

In this chapter, all research questions will be examined, and hypotheses with validated measures will be tested from further data analysis. Some of the questions are measured by non-parametric methods, and others are measured by parametric methods, depending on the nature of the sample data used. A non-parametric statistical test is a test whose model does not specify conditions about the parameters of the population from which the sample was drawn. A parametric statistical test requires certain assumptions about the parameters of the population from which the research sample was drawn. The meaningfulness of the result of a parametric statistical test depends on the validity of these assumptions. One of these assumptions is that the sample needs to follow a normal distribution.

10.1 Test of Hypotheses

As tested in the previous chapter, except for "Perceived Usefulness," sample data for all other variables failed to follow the normal distribution. Thus, Spearman's rank correlation test is used instead of Pearson's correlation for correlation tests, as the latter requires both variables to follow normal distribution. As discussed by Altman (1991), Spearman's correlation coefficient gives as much information as Pearson's correlation coefficient and is of wider validity. Non-parametric ANOVA is used to test the difference in population.

10.1.1 Result Demonstrability

Result demonstrability is measured by four questionnaire items and a higher number denotes a higher level. Table 10.1 shows the correlation between "Result Demonstrability" and other variables.

Table 10.1 Correlation Between Result Demonstrability and Other Variables

Variable	Result demonstrability	
Perceived Usefulness	Correlation	0.58
	significance level	<.0001
Perceives Ease of Use	Correlation	0.57
	significance level	<.0001
Intention to continue use	Correlation	0.48
	significant level	<.0001
Number of pages visited	Correlation	0.08
	significance level	0.48 (not significant)

Table 10.1 shows that "Result Demonstrability" has a significant correlation with "Perceived Usefulness", "Perceived Ease of Use", and "Intention to continue use". However, it doesn't have a significant correlation with real usage data. Furthermore, non-parametric analysis of variance (ANOVA) is conducted for further analysis to look at whether there is a relationship between level of perceived result demonstrability and perceived usefulness. Because there are more than three classes for "Result demonstrability," the Kruskal-Wallis test is used to compare groups.

Table 10.2 Non-Parametric ANOVA for Perceived Usefulness on Result Demonstrability

Result demonstrability	N	Mean
13	6	23.50
14	2	23.50
15	9	26.88
16	13	29.38
17	11	30.09
18	12	29.33
19	7	31.43
20	9	31.67
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=29.36, df=7, Pr>Chi-Sq <.0001		

Results from table 10.2 indicate:

From bivariate analysis, Result demonstrability is found to have a positive influence on perceived usefulness of a community place.

Table 10.3 Non-Parametric ANOVA for Perceived Ease of Use on Result Demonstrability

Result demonstrability	N	Mean of Perceived Ease of Use
13	6	10.00
14	2	9.00
15	9	11.55
16	13	11.69
17	11	11.45
18	12	13.00
19	7	12.71
20	9	14.00
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=26.39, df=7, Pr>Chi-Sq = .0004		

Table 10.3 and correlation test results from Table 10.1 indicate:

Result demonstrability is found to have a positive influence on perceived ease of use of the community system.

Table 10.4 Non-Parametric ANOVA for Intention to Continue Use on Result Demonstrability

Result demonstrability	N	Mean of Intention to continue use
13	6	13.17
14	2	12.00
15	9	12.89
16	13	13.92
17	11	13.73
18	12	14.42
19	7	14.71
20	9	14.77
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=21.40, df=7, Pr>Chi-Sq = .0032		

Table 10.4 and correlation test results from Table 10.1 indicate:

Result demonstrability is found to have a positive influence on intention to continue use a community place.

Table 10.5 Non-Parametric ANOVA for Number of Pages Visited on Result Demonstrability

Result demonstrability	N	Mean of Square root of number of Pages accessed
13	6	5.51
14	2	4.41
15	9	7.30
16	13	6.02
17	11	8.42
18	12	7.51
19	7	8.83
20	9	5.81
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=4.47, df=7, Pr>Chi-Sq = .7238		

Result from Table 10.5 and correlation test results in Table 10-1 indicate:

Result demonstrability do not have any influence on number of pages accessed in the community place.

The bivariate test on "Result Demonstrability" shows that perceived result demonstrability of the community place has significant positive influence on members' perceived usefulness, perceived ease of use, and also intention to continue using the system space. However, it doesn't directly affect members' real usage in the community place.

10.1.2 Attitude towards the Online Environment

Attitude towards the online environment are measured by two questionnaire items. One of them is "For me, the WebCenter is JUST a place where I can get information." The relationship with the response to this question and the social engagement activity is tested using non-parametric ANOVA (Table 10.6 shows the result). Again, the Kruskal-Wallis test is used for the purpose. As the number of answers for "1" and '5" is quite small, people whose response was "1" are grouped together with people whose response were "2". Higher scores show disagreement with the statement.

Table 10.6 Non-parametric ANOVA for Attitude and Social Engagement

Attitude	N	Mean of Social Engagement
2	26	14.15
3	22	7.27
4	25	43.04
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=6.48, df=2, Pr>Chi-Sq = .039		

The result shows that the difference between different levels of attitudes toward the community place did make a difference in degree of social engagement at the .05 significance level:

H4a. People who think the community place is just for information will have less social engagement activities than people who do not agree with this statement.

Supported

The correlations between attitude and other variables are shown in table 10.7.

Table 10.7 Correlation Between Attitude and Other Variables

Variable	Attitude toward online space	
Perceived Usefulness	correlation	0.32
	significance level	.0059
Intention to continue use	correlation	0.28
	significance level	.001
Social Engagement	correlation	0.02
	significance level	0.87
Perceived Sense of Community	correlation	0.34
	significance level	.002

Table 10.7 shows that there is a positive correlation between attitude and perceived usefulness of the community place; attitude and intention to continue using the community place. There is no direct relation between attitude and social engagement activities. In order to check whether different levels of attitude will cause differences in perceived usefulness, perceived ease of use, and intention to continue use, the Kruskal-Wallis test is used for analysis of variance.

Table 10.8 Non-Parametric ANOVA for Perceived Usefulness on Attitude

Attitude	N	Mean of Perceived Usefulness
4	5	29.8
5	17	27.23
6	9	28.77
7	17	26.88
8	15	30.2
9	6	30.83
10	4	33.50
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=16.63, df=6, Pr>Chi-Sq = .0108		

The above ANOVA results and the correlation test show that:

People's attitude towards online community is found to have a positive influence on perceived usefulness of the community place.

Table 10.9 Non-Parametric ANOVA for Intention to Continue Use on Attitude

Attitude	N	Mean of Intention to continue use
4	5	13.60
5	17	13.47
6	9	13.44
7	17	13.65
8	15	14.20
9	6	14.50
10	4	15.00
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=8.92, df=6, Pr>Chi-Sq =.1776		

The above ANOVA results and the correlation test in Table 10.7 show that:

People's attitude towards online community is found to have a positive effect on intention to continue using the community place.

As proposed in Chapter six, attitude towards the online community also influences members' sense of community, the Kruskal-Wallis test is used to test this hypothesis with question 17 ("For me, the WebCenter is just a place where I can get information") and sense of community (see Table 10.10 for the result).

Table 10.10 Non-Parametric ANOVA for Sense of Community on Attitude

Attitude	N	Mean of sense of community
1	4	7.50
2	23	7.26
3	22	7.64
4	18	8.72
5	6	8.66
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=13.52, df=4, Pr>Chi-Sq =.009		

The above ANOVA result and shows that:

H4d: People who think the community place is just for information have a lower level of perceived sense of community.

Supported

In order to look at whether people's attitude towards the online community will influence their perceived usage and value of the awareness tools, Kruskal-Wallis Test is used to compare perceived usage (1 is "never" and 3 is "frequently") and value (1 is "can't say," 2 is "useless," 5 is "extremely useful") on "people directory" feature among different levels of response to question 17.

Table 10.11 Non-Parametric ANOVA for Attitude and Perceived Usage on Awareness Tools

Attitude	N	Mean of usage of "people directory"
1	4	2.0
2	22	1.41
3	19	1.47
4	17	1.88
5	5	2.0
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=10.92, df=4, Pr>Chi-Sq =.027		

The Spearman's correlation between q17 and self reported usage of "people directory" is 0.26 (p=0.03). This ANOVA test indicate that:

H4e: People who think the community place is just for getting information use the awareness tools less than others.

Supported

Table 10.12 Non-Parametric ANOVA for Attitude and Perceived Usefulness of the Awareness Tools

Attitude	N	Mean of perceived usefulness of "people directory"
1	3	4.00
2	14	3.21
3	9	3.55
4	13	3.61
5	4	4.25
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=12.09, df=4, Pr>Chi-Sq =.0167		

The Spearman's correlation between q17 and self_reported value of "people directory" is 0.28 (p=0.06). Table 10-12 and correlation test indicate that:

H4f: People who think the community place is just for getting information have a lower perceived usefulness of awareness tools.

Supported

From the correlation test and analysis of variance, people's attitude towards online community has a positive effect on perceived usefulness of the community place. Although attitude has a positive correlation with intention to continue use at the .01 level, different levels of attitude did not make a difference in intention to continue use. This result indicates that in such a knowledge-based community, how people feel about the online environment does not have a significant role in their behavior intention. People who treated the online community as a place just for getting information have less perceived sense of community; they use the awareness tools less frequently and perceive less. This result confirms the research model in which attitude directly relates to perceived sense of community and perceived usefulness of the awareness tools, but not to intention to continue use.

10.1.3 Sense of Community

Sense of community is measured by two questionnaire items. Table 10-13 shows the correlation between sense of community and other variables.

Table 10.13 Correlation Between Sense of Community and Other Variables

Variable	Sense of Community	
Perceived Usefulness	correlation	0.39
	significance level	.0006
Perceived Ease of Use	correlation	0.38
	significance level	.0009
Intention to continue use	correlation	0.33
	significance level	.0045
Social Engagement	correlation	0.16
	significance level	0.18

Besides correlation analysis, non-parametric ANOVA with Kruskal-Wallis Test is conducted for further analysis (see Table 10.14 to Table 10.17 for results).

Table 10.14 Non-parametric ANOVA for Perceived Usefulness Classified by Sense of Community

Sense of Community	N	Mean of Perceived Usefulness
6	10	27.30
7	18	26.56
8	12	28.83
9	19	30.00
10	10	31.70
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=14.58, df=4, Pr>Chi-Sq =.0056		

Together with the result from correlation analysis, the above results show that:

Perceived sense of community is found to have a positive influence on perceived usefulness of a community place.

Table 10.15 Non-parametric ANOVA for Perceived Ease of Use Classified by Sense of Community

Sense of Community	N	Mean of Perceived Ease of Use
6	10	10.70
7	18	11.28
8	12	12.00
9	19	12.26
10	10	13.60
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=13.12, df=4, Pr>Chi-Sq =.0107		

Together with the result from correlation analysis, the above results show that:

Perceived sense of community will have a positive effect on perceived ease of use of a community place.

Although from a theoretical point of view, perceived sense of community should not have any influence on perceived ease of use of the system, the current bivariate analysis shows an opposite result. This result will be checked against the PLS analysis described in the next chapter to see whether this relationship holds when considering the entire model.

Table 10.16 Non-parametric ANOVA for Intention to Continue Use Classified by Sense of Community

Sense of Community	N	Mean of Intention to continue use
6	10	12.90
7	18	13.06
8	12	14.42
9	19	14.11
10	10	14.60
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=16.50, df=4, Pr>Chi-Sq =.0024		

Together with the result from correlation analysis, the above results show that:

Perceived sense of community is found to have a positive effect on intention to continue use of a community place.
Supported

Table 10.17 Non-parametric ANOVA for Social Engagement Classified by Sense of Community

Sense of Community	N	Mean of Social Engagement
6	10	8.70
7	18	14.94
8	12	36.25
9	19	27.21
10	10	18.60
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=3.28, df=4, Pr>Chi-Sq =.51 (not significant)		

Result from Table 10.17 shows that:

Perceived sense of community does not have a positive effect on social engagement activities of a member.

After checking the original question designed for perceived sense of community, it is not hard to understand why it does not affect real usage activities. The sense of community construct includes two questionnaire items

- I trust other members in the ALNResearch WebCenter
- I always feel isolated in the ALNResearch WebCenter

Both questions are asking about people's feeling for the community place, thus it's more reasonable that these feelings will affect users' perceived usefulness rather than real usage.

In this section, the relationships between sense of community and other variables are studied. The results show that sense of community has a significant positive effect on perceived usefulness of the community place, perceived ease of use of the community place and intention to continue using the community system. However, perceived sense of community doesn't have a direct relation with social engagement activity.

10.1.4 Basic TAM Model

In this section, variables in the basic TAM model will be tested using both correlation analysis and non-parametric one-way ANOVA.

10.1.4.1 Perceived Usefulness. Perceived usefulness of the community place is measured by seven questionnaire items. Table 10.18 shows the correlation between perceived usefulness and other variables.

Table 10.18 Correlation Between Perceived Usefulness and Other Variables

Variable	Perceived Usefulness	
Perceived Ease of Use	correlation	0.59
	significance level	<.0001
Intention to continue use	correlation	0.56
	significance level	<.0001
Number of pages visited	correlation	0.23
	significance level	.05
Information seeking	correlation	0.19
	significance level	.11 (not significant)
Social engagement	correlation	0.11
	Significance level	.36 (not significant)

Table 10.18 shows that perceived usefulness has a strong correlation with perceived ease of use, intention to continue use, and number of pages visited. However, it doesn't have a strong correlation with information seeking and social engagement

activities separately. To further understand the relationship between perceived usefulness and perceived ease of use, intention to continue use, and number of pages visited, Kruskal-Wallis test is conducted. Perceived usefulness is measured by the sum of responses to the seven questionnaire items. It results in too many levels of the measure. In order to reduce the classes of perceived usefulness, the initial classes have been grouped by dividing the score by 7 and taking the closest integer. After this process, three classes are established. Table 10.19 to Table 10.21 shows results of Kruskal-Wallis test with the three levels of perceived usefulness

Table 10.19 Non-parametric ANOVA for Perceived Ease of Use on Perceived Usefulness

Perceived usefulness	N	Mean of Perceived Ease of Use
3	9	10.11
4	46	11.61
5	18	13.67
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=21.48, df=2, Pr>Chi-Sq <.0001		

The above results show that:

Perceived usefulness of the community place has a positive influence on perceived ease of use of the community place.

Table 10.20 Non-parametric ANOVA for Intention of Continued Use on Perceived Usefulness

Perceived usefulness	N	Mean of Intention to continue use
3	9	12.88
4	46	13.63
5	18	14.83
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=15.58, df=2, Pr>Chi-Sq =.0004		

The above results show that:

Perceived usefulness of the community place has a positive effect on intention to continue using the community place.

This result is consistent with the result from TAM model. Perceived usefulness of the community place is very important in determine users intentions on continued use.

Table 10.21 Non-parametric ANOVA for Number of Pages Visited on Perceived Usefulness

Perceived usefulness	N	Mean of number of pages visited
3	9	59.00
4	46	62.08
5	18	77.44
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=2.92, df=2, Pr>Chi-Sq =.2316 (not significant)		

The above results show that:

Perceived usefulness of the community place does not have a positive effect on real usage of the community place.

From Table 10.21, although different levels of perceived usefulness did make a difference on mean number of pages visited, this difference is not significant. One possible reason is that we use 1-5 levels to measure perceived usefulness, which number of pages visited is measured by a ratio scale.

10.1.4.2 Perceived Ease of Use. Perceived ease of use is measured by three questionnaire items and a higher value represents a higher level. Table 10.22 shows the correlation between perceived ease of use and other variables in the basic TAM model.

Table 10.22 Correlation Between Perceived Ease of Use and Other Variables

Variable	Perceived ease of use	
Perceived Usefulness	correlation	0.56
	significance level	<.0001
Intention to continue use	correlation	0.41
	significance level	.0003
Number of pages visited	correlation	0.22
	significance level	.0658
Information seeking	correlation	0.19
	significance level	.11 (not significant)
Social engagement	correlation	0.07
	Significance level	.56 (not significant)

The above table shows that perceived ease of use has a strong correlation with perceived ease of use, intention to continue use, and a weak correlation with number of pages visited. However, it doesn't have a strong correlation with information seeking and social engagement activities separately. To further understand the relationship between

perceived ease of use and perceived usefulness, intention to continue use, and number of pages visited, the Kruskal-Wallis test is conducted. Perceived ease of use is measured by the sum of responses to the three questionnaire items. It results in too many levels of the measure. In order to reduce the levels of perceived ease of use, the initial classes have been grouped by dividing the score by 3 and rounding it to the closest integer. After this process, three classes are established. Table 10.23 to Table 10.25 show results of the Kruskal-Wallis tests with the three levels of perceived ease of use.

Table 10.23 Non-parametric ANOVA for Perceived Usefulness Classified by Perceived Ease of Use

Perceived ease of use	N	Mean of Perceived Usefulness
3	9	25.89
4	36	27.72
5	28	31.04
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=17.53, df=2, Pr>Chi-Sq =.0002		

The above results show that:

Perceived ease of use of the community place has a positive influence on perceived usefulness of the community place.

Table 10.24 Non-parametric ANOVA for Intention to Continue Use by Perceived Ease of Use

Perceived ease of use	N	Mean of Intention to continue use
3	9	13.11
4	36	13.50
5	28	14.50
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=12.07, df=2, Pr>Chi-Sq =.0024		

The above results show that:

Perceived ease of use of the community place has a positive influence on intention to continue using the community place.

The above results are consistent with the TAM model, in which perceived ease of use has a positive influence on both perceived usefulness and behavioral intention. However, from previous studies, perceived ease of use should not take an important role

after the system is adopted. This bivariate analysis result will be checked against the PLS analysis in the next chapter which takes all constructs into account.

Table 10.25 Non-parametric ANOVA for Number of Pages Visited by Perceived Ease of Use

Perceived usefulness	N	Mean of Number of Pages Visited
3	9	38.11
4	36	68.32
5	28	70.14
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=3.69, df=2, Pr>Chi-Sq =.1578 (not significant)		

Results from Table 10.25 and correlation test show that:

Perceived ease of use of the community place does not have a significant positive influence on real usage of the community place.

The correlation analysis and ANOVA results show that perceived ease of use has a positive effect on both perceived usefulness of perceived ease of use. It has a positive correlation with real usage of the community place, but different levels of perceived ease of use do not make significant difference on members' real usage of the system.

10.1.4.3 Intention to Continue Use. Intention to continue use is measured by three questionnaire items; higher value shows higher intention to continue using the community place. Table 10.26 shows the correlation between intention to continue use and other variables.

Table 10.26 Correlation Between Intention to Continue Use and Other Variables

Variable	Intention to continue use	
Perceived Usefulness	correlation	0.59
	significance level	<.0001
Perceived Ease of Use	correlation	0.41
	significance level	.0003
Number of pages visited	correlation	0.17
	significance level	.14 (not significant)
Information seeking	correlation	0.10
	significance level	.40 (not significant)
Social engagement	correlation	0.12
	Significance level	.29 (not significant)

Table 10.26 shows that "intention to continue use" has a significant correlation only with "perceived usefulness" and "perceived ease of use"; it doesn't have a significant correlation with real usage of the community place as expected in the research model. The Kruskal-Wallis test is conducted to further analyze the relationship between intention to continue use and perceived usefulness, and relationship between intention to continue use and perceived ease of use (see Table 10.27 and Table 10.28 for the results).

Table 10.27 Non-parametric ANOVA for Perceived Usefulness by Intention to Continue use

Intention to continue use	N	Mean of Perceived Usefulness
11	4	26.00
12	10	25.50
13	11	27.09
14	14	28.72
15	34	30.61
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=23.55, df=2, Pr>Chi-Sq <.0001		

Results from Table 10.27 and correlation test show that:

Higher intention to continue using the community place is related to a higher perceived usefulness of the community place.

Table 10.28 Non-parametric ANOVA for Perceived Ease of Use by Intention to Use

Intention to continue use	N	Mean of Perceived Ease of Use
11	4	10.50
12	10	10.80
13	11	11.27
14	14	11.57
15	34	12.79
Kruskal-Wallis Test (Chi-Square Approximation) Chi-Sq=12.23, df=2, Pr>Chi-Sq =.0157		

Results from Table 10-28 and correlation test show that:

Higher intention to continue using the community place is related to a higher perceived ease of use of the community system.

The bivariat analysis of the basic TAM model constructs shows that TAM model relationships hold true in this research context. This means that members' perceived usefulness of the community place and their perceived ease of use of the community

system significantly influence their intention to continue using the community place. Perceived ease of use of the community system has a positive influences on perceived usefulness of the community place. Some new findings are:

- Perceived usefulness also positively influence the perceived ease of use of the community system.
- Intention to continue using the community place also positively affects perceived usefulness of the community place.
- Intention to continue using the community place also positively affects perceived ease of use of the community system.

10.2 Answering Research Questions

In this section, research questions that were proposed in chapter six will be answered.

10.2.1 Effect of Awareness Tools

The following research question is asked about the effect of awareness tools on members' engagement.

RQ1: Will the support of awareness help to encourage social engagement in the community?

It has been broken into two questions. They will be answered separately using both Web Server log data and the survey questionnaire results.

RQ1a*:* Will members use the awareness tools to get to know other people?

This question can be answered by looking at how much of the usage has been on checking other members' information, which relates to "people_directory" page views in the log files. Figure 10.1 shows the number of page views on "people_directory" over the observation period. Figure 10.2 shows the distribution of different categories of the page

views. From the distribution figure, 11% of the page views are on the "people directory", which is a relatively important portion of all page views. This result shows that members in the community are trying to get to know other people in the community. Although there is not much interaction in the community place, the behavior of checking other members' information shows a tendency of people's willingness to get to know other members. In the long run, this will lead to more social engagement in the community place.

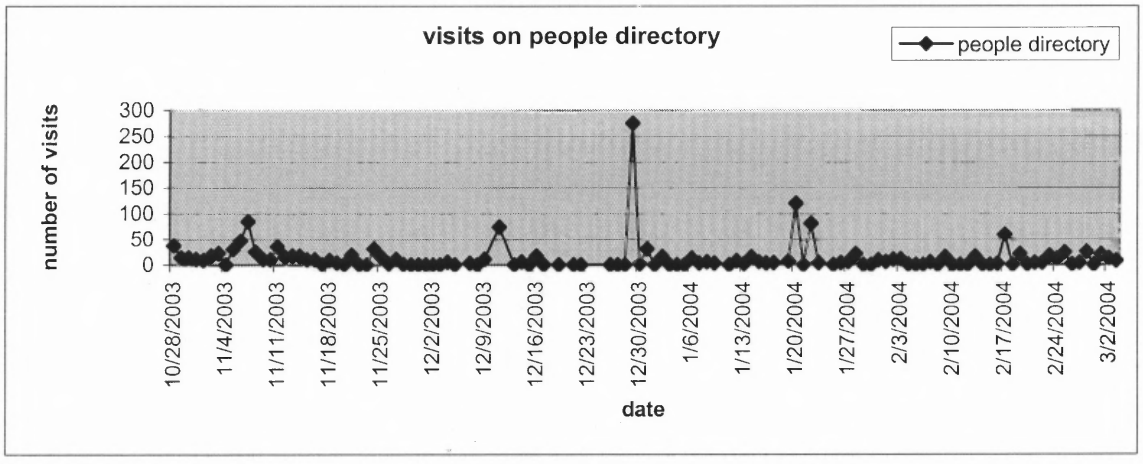

Figure 10.1 Number of Page Views to "People Directory" Over the Observation Period

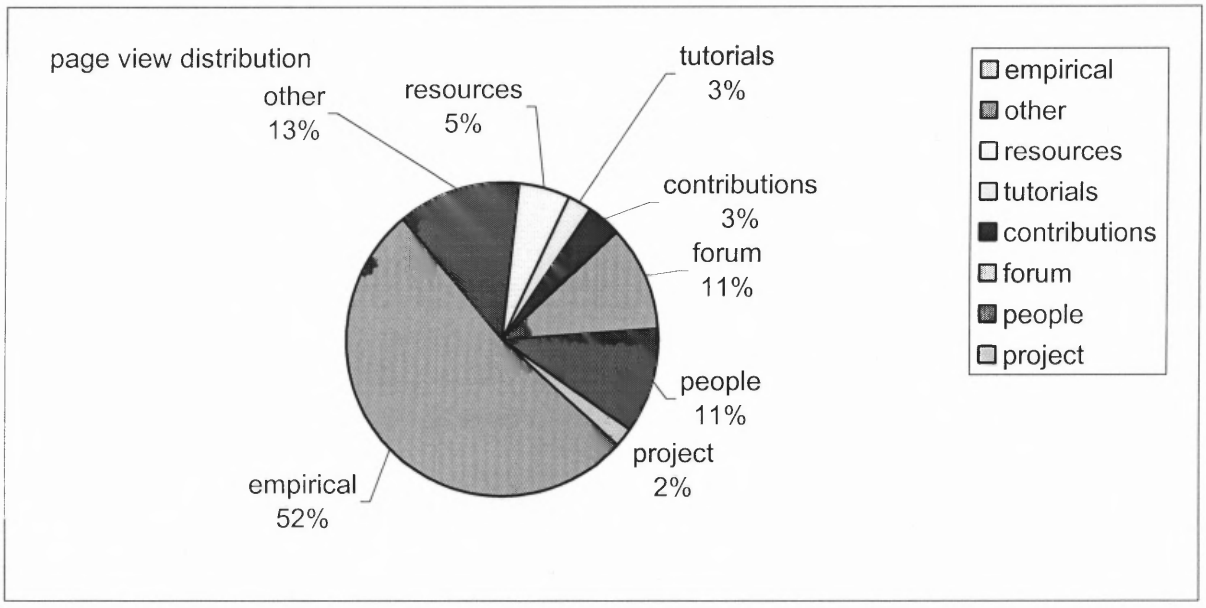

Figure 10.2 Distribution of Page Views on Different Categories

RQ1b: Will members start networking with other members with the information they obtain from the awareness tools?

This will be examined through the answers to the question "I have contacted people who I met from the WebCenter by email, phone or other media" in the questionnaire. Among 78 valid responses, 11 of them answered "yes" and 67 answered "no", which shows that 14% of the subjects started networking with other members they got to know from the community place. Although this is not a very high percentage, when taking into account the short observation time after the implementation of the tools, the result is promising. It gives a sign of intention to socialize and broaden individual social networks.

10.2.2 Effect of Recommender System

One research question has been asked about the effect of the recommender system.

RQ2: Will the implementation of a recommender system make community members engage more in information seeking activities?

It also has been divided into two sub questions. They are answered separately below.

RQ2a: Are there more activities after implementing the recommender system than before its implementation?

The recommender system was implemented on 12/20/2003. In order to compare the difference in activities before and after the implementation, the average article pages accessed per day is compared with data for before and after the implementation of the recommender system. As there is not much activity during Christmas time, the data from 12/22-12/26 were removed from the analysis. Table 10.29 shows the result of average

article pages accessed per day before and after implementing the recommender system. Figure 10.3 shows number of page views on articles accessed per day over the observation time. Table 10.30 shows the number of average article pages accessed daily per week during the observed period, and Figure 10.4 gives the corresponding diagram. The weekly information gives a better image of changes of article accesses over time.

Table 10.29 Articles Accessed Per Day

	Before implementation of the recommender system	After implementation of the recommender system
Articles accessed per day	49.1	27.7

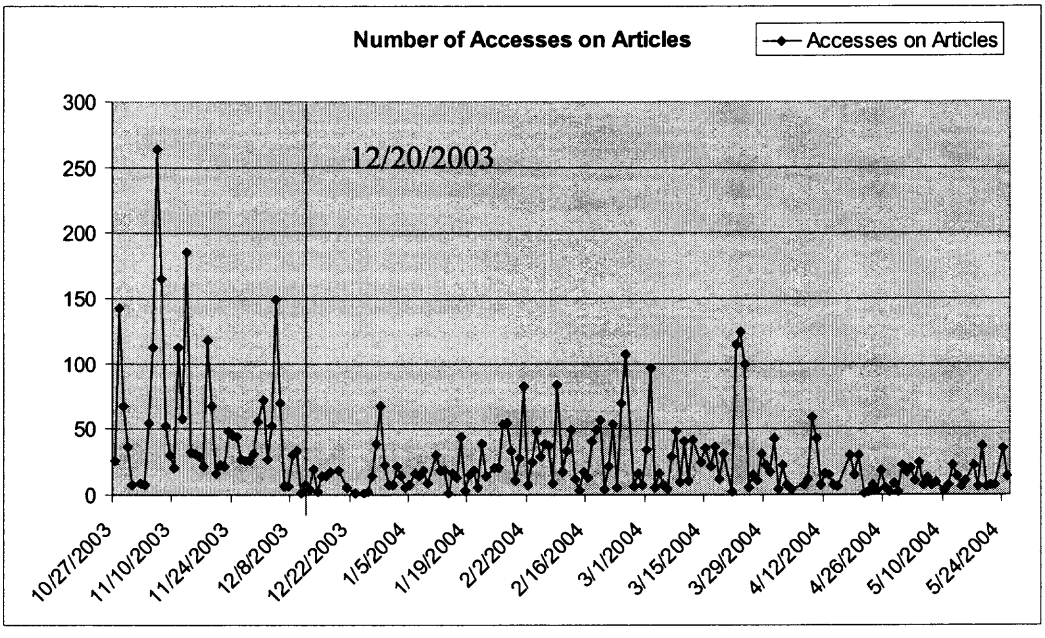

Figure 10.3 Articles Accessed Per Day Over the Observed Period

Table 10.30 Weekly Average Accesses to Article Pages Per Day

Date	Number of accesses to articles	Number of article accesses from the recommendations
11/2-11/8	665	56
11/9-11/15	468	114
11/16-11/22	296	73
11/23-11/29	247	56
11/30-12/6	433	140
12/7-12/13	100	33
12/14-12/20	65	4
12/21-12/27	23	6

Table 10.30 Weekly Average Accesses to Article Pages Per Day (Continued)

Date	Number of accesses to articles	Number of article accesses from the recommendations
12/28-1/3	178	20
1/4-1/10	70	21
1/11-1/17	140	32
1/18-1/24	94	23
1/25-1/31	221	40
2/1-2/7	269	44
2/8-2/14	207	38
2/15-2/21	202	61
2/22-2/28	258	85
2/29-3/6	170	56
3/7-3/13	181	57
3/14-3/20	160	41
3/21-3/27	372	30
3/28-4/3	147	69
4/4-4/10	126	14
4/11-4/17	53	44
4/18-4/24	88	33
4/25-5/1	76	25
5/2-5/8	93	15
5/9-5/15	63	19
5/16-5/22	86	16

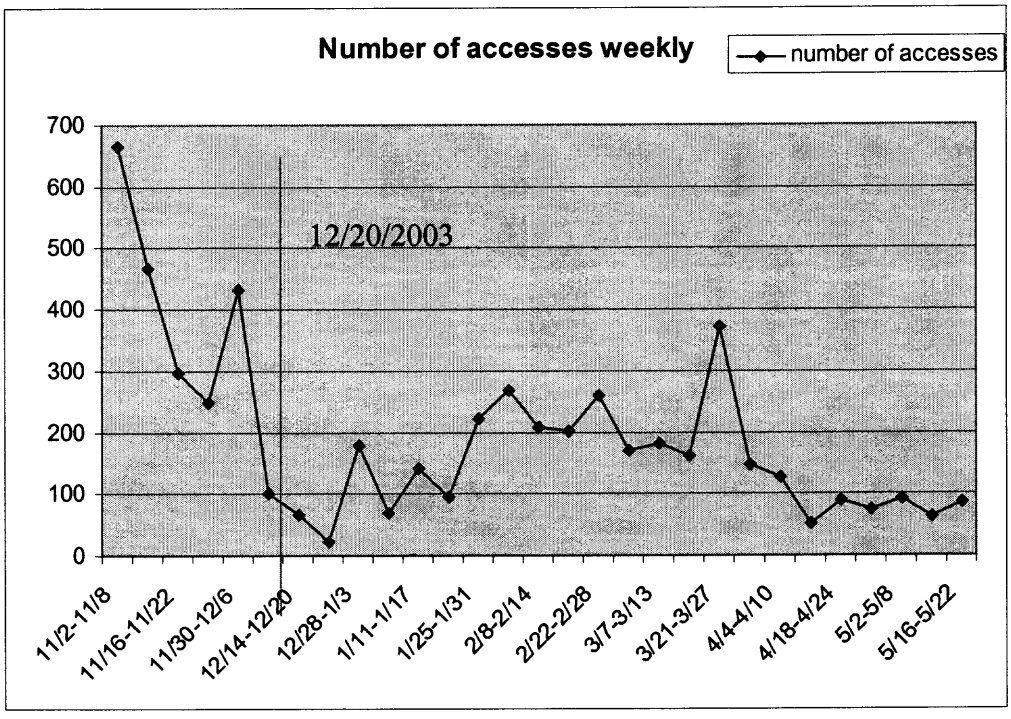

Figure 10.4 Weekly Average Accesses to Articles Per Day

From table 10.29, it seems that average accesses to article pages per day is more before the implementation of recommender system than after the implementation. By checking the weekly average data, the highest average accesses to articles pages is in Nov, during which the advertisement of the ALNResearch WebCenter had been sent to two different mail-lists and also the ALN annual conference was held. Thus, it is not surprising to see the average number of accesses is high. December and January are the months that have least accesses to article pages, which is the time of finals and winter vacation for academics. From the third week of January, which is also the start of the new semester, the number of accesses has increased and keeps very stable. However, there is not enough evidence to show the effect of the recommender system on number of accesses to articles. One reason might be because no push technology (like email notification) has been used to make the recommendations reach the users. Members can only check the recommendations when they log into the system. However, the above table and figures shows that the number of accesses is correlated closely to advertisement and academic calendar.

RQ2b: Will the recommender system help members pick up more useful information from the community knowledge base?

Content analysis on the web server log was conducted in order to answer this question. The number of times that users choose to read an article suggested by the recommender system is counted. Among all 5,879 page views of accessing articles, 1267 of them were accessed from the recommender system feature, which is 22% of all page views of accessing articles. Although the total number of page views decreases after December 2003, the percentage of accesses from recommender system increases. In order

to get recommendations from the system, the user must read a sufficient number of articles (at least 5) in the knowledge base. Thus, 22% is a promising number considering the short period of implementation of the recommender system, and the small number of articles in the knowledge base.

10.2.3 Other Factors

Besides the constructs that are presented in the research model, whether other factors such as gender and role in ALN have an effect on real usage or perceived usefulness of the community place is also an interesting question. With the questionnaire items, because subjects can choose as many roles as fit their situation, the test on comparison of different roles is not valid. However, from the title of the member and their behavior, some evidence can be found on the relationship between roles and real usage.

10.2.3.1 Gender. First, effect of gender on perceived usefulness is studied. As perceived usefulness is following a normal distribution, a T-test can be used to see whether there is any difference in perceived usefulness between females and males. Table 10.31 shows the result from the t-test.

Table 10.31 Result of T-test of Perceived Usefulness on Gender

	N	Mean	Std
Male	38	29.23	3.51
Female	35	28.26	3.91
T=1.13, Pr>|t| = .26 (not significant)			

Results from Table 10.31 show that there is no difference in perceived usefulness between females and males.

As perceived ease of use, intention to continue use and number of pages visited do not follow a normal distribution, the Wilcoxon test is used as a non-parametric test to

test the difference between males and females (see table 10.32 to table 10.36 for the results).

Table 10.32 Result of T-test of Perceived Ease of Use on Gender

Perceived ease of use	N	Mean
Male	38	11.82
Female	35	12.06
Wilcoxon test Z=0.36, Pr>|Z| = .73 (not significant)		

Table10.33 Result of T-test of Intention to Continue Use on Gender

Intention to continue use	N	Mean
Male	38	13.81
Female	35	13.86
Wilcoxon test Z=-0.17, Pr>|Z| = .86 (not significant)		

Table 10.34 Result of T-test of Number of Pages Visited on Gender

Number of pages visited	N	Mean
Male	38	72.42
Female	35	57.97
Wilcoxon test Z=0.04, Pr>|Z| = .96 (not significant)		

Table 10.35 Result of T-test of Information Seeking Activities on Gender

Information seeking	N	Mean
Male	38	52.71
Female	35	52.14
Wilcoxon test Z=0.65, Pr>|Z| = .51 (not significant)		

Table 10.36 Result of T-test of Social Engagement Activities on Gender

Social engagement	N	Mean
Male	38	25.81
Female	35	17.80
Wilcoxon test Z=0.08, Pr>|Z| = .93 (not significant)		

From the results in above tables, there is no difference in perceived usefulness, perceived ease of use, intention to continue use, and all real usage categories between males and females.

10.2.3.2 Title. In order to check whether there is any difference between members' title and their usage in the community place, members with different titles and their behavior are shown in table 10.37.

Table 10.37 Relationship Between Members in Different Title and Their Behavior

	N	Mean of number of pages visited	Mean of information seeking	Mean of social engagement
N/A	12	149.41	100.25	60.92
Ph. D. student	4	105.25	85.00	56.75
Ms	2	99.00	86.50	15.50
Administrator	20	36.25	31.10	8.40
Dr.	6	46.00	18.67	27.00
Researcher	4	22.00	12.50	7.25
Programmer	1	41.00	40.00	0
Consultant	3	108.00	127.33	22.33
Founder	1	3.00	2.00	0
Student	1	52.00	43.00	16.00
Professor/Instructor	12	51.05	53.89	8.21
Mr.	2	54.50	42.50	17.50
Evaluator	1	30.00	32.00	0
Teacher	1	8.00	6.00	0
Specialist	1	16.00	11.00	9.00
Designer	1	25.00	25.00	25.00

Results from Table 10.37 indicate that:

- The modal category for members who answered the questionnaire is administrator. However, their engagement in the community place is quite low. This might be due to their time constraints.
- Besides people who did not provide information on their titles, Ph.D. students are the group that spends more time and effort in the community place both for information seeking activities and social engagement activities.
- Consultants do a lot of information seeking activities, and they do not spend much time in social engagement activities.
- Teachers and evaluators come to the community place just for seeking information.

- Professors and instructors spend more of their time in the community for information seeking than social engagement activities.

10.2.3.3 Time Spent on the Internet Per Day. From basic information that collected from the user, the connection speed to the Internet doesn't vary much. People use either Cable/DSL or T1 connection. (check Table 9-1 for detail). The time a person spent on the Internet per day varies from "less than 2 hours" to "more than 6 hours". Whether the differences in time people spent on the Internet will make any difference in people's perception of the community place and their real usage is tested by Kruskal-Wallis test (see Table 10.38).

Table 10.38 Kruskal-Wallis Test Between Time Spent on the Internet and Other Variables

Time spent on the Internet	N	Mean of perceived usefulness	Mean of perceived ease of use	Mean of intention to continue use	Mean of number of acesses
Less than 2 hours	5	26.40	11.60	13.60	42.60
2-4 hours	16	28.68	11.87	14.06	54.37
4-6 hours	23	27.91	11.43	13.43	45.87
More than 6 hours	28	29.96	12.39	14.07	91.50
		Chi=6.33 Pr>Chi = .0966	Chi=2.90 Pr>Chi = .41	Chi=3.34 Pr>chi=.34	Chi=2.01 Pr>Chi=.57

10.2.4 How Well does the Community Place Fit Individuals Expectations

In chapter six, research question four asks about how well the community place fits members' expectations. Two open-ended questions in the questionnaire are related to this research question (see Table 10.39 for detail).

Table 10.39 Two Open-ended Questions in the Questionnaire

Item No.	Question / Statement
1	What are your expectations when you decided to join the ALNResearch WebCenter?
2	What other kind of features would you like to see added to the ALNResearch WebCenter?

For question 1, the summarization of the answers is shown in Table 10.40.

Table 10.40 Members' Initial Expectations and Corresponding System Features /Community Contents

Initial expectations	No.	Corresponding community contents/system features
Get references in ALN research	1	Empirical study, other study
Find empirical studies	4	Empirical study, other studies related to ALN, and search function
Share information. Share research. Collaborate with others. Learn from the people. And if it is possible teach with my experiences.	2	Contribution form, people directory
Learn more ALN research and get to know who is doing what in this field	3	Empirical and other studies related to ALN. People directory
Find quality research articles and research instruments related to ALN	19	Empirical and other studies, research instruments
Find Updated research on ALN	6	Still needs input more updated studies
To conduct join research, publish or invite for publishing results of the empirical studies in my university's online journal, in due course. To meet more researchers and educators in DE field. To find out what can be done cooperative project.	1	The community place provide people directory for getting to know other members, it also provides members' contact information for collaboration. The discussion forum also provides a place for interaction.
Rapid access to ALN literature	1	Gather most recent empirical studies related to ALN
Find others with similar research agenda, collaboration	2	The community place provide people directory, with member' interests and contact information
Find materials to current course work or thesis	4	Empirical studies, other studies and resources
Find some specific research	3	Empirical studies, other studies and resources.
Get to know top researchers and hot topic in ALN	1	Video clip with top researchers about hot topic in ALN
Obtain background information on ALN research and pedagogy for online teaching and to find literature resources for use when preparing my own papers and grant applications	1	Empirical studies, other studies, resources, and discussion forums
Curiosity - ALN is not my line of research. I was just looking at different types of elearning research related web sites.	1	
To encourage more faculty. To use the research to ground my own interests in ALN To distribute research findings to faulty beginning in ALN	1	

The above table shows that people join the community with different expectations. They can be categorized as either expecting to get information related to ALN research or to get to know other people in the field and collaboration. Most of the people join the

community for the purpose of accessing more information, but there are still a number of people who want to know other people in the field and if possible start collaboration with other members. These answers confirm the proposition stated at the beginning of this dissertation that online community is not only a good form of knowledge gathering space, but also a place for people with similar interests to meet and collaborate.

Open-ended question 2 asked about what other features members wish to add into the community system. People who answered this question provide the following expectations:

- Better keyword search functions.
- Add more full text research articles
- Different sorting methods on list of articles
- Add IM or chatroom functions
- Email notifications for new contributions
- A collection of conference talks/posters that are submitted & screened, so that we have access to more current work being conducted by ALN researchers.

The above suggestions show that although the site was originally designed for asynchronous communication and pull technologies only, there is some desire for synchronous communication functions to help connecting people more cohesively and push technology for regular updates on the contents of the community place.

In the questionnaire, the value of the WebCenter to members was not directly asked. There are still some members who expressed their appreciation for the site in the answer to the open-ended questions. Following are some quotes from users responses:

"I was pleasantly surprised to have free access to research articles."

"Very informative WebSite."

"Really like the public domain resources page."

"It looks like a wonderful resource. Thanks."

"I discovered the WebCenter online and it appeared to be an excellent repository of empirical research. It did provide some excellent materials."

"What am I expecting are already there in this site."

"I would find some specific research. My expectations have been exceeded."

"Your presentation at ALN was very good and I enrolled when I got back."

"I was hoping to be able to access the areas and find useful information....which I did. Thank you very much!"

"I was looking for empirical research studies and I found the studies on this site of high quality."

"So far, what I find EXCEEDS my expectations. I'm surprised that I'm finding instruments there."

"I think you have done a great job with it."

The above quotes show that members of the WebCenter really like it and appreciate the availability of such a community space.

10.3 Summary

This chapter tests hypotheses and answers questions presented in chapter six. Some of the hypotheses cannot be tested due to the validity of the constructs. Research question three, which asks about the effect of community activities also cannot be answered because not enough communication occurred during the observation period to answer this question. Research question six, which asks about the usage change over time, is answered in chapter eight and is not repeated here.

Table 10.41 summarized all the results of testing of hypotheses.

Table 10.41 Results for Testing of Hypotheses

Hypotheses	Variables	Result
H3a. Result demonstrability will have a positive effect on perceived usefulness of a community place.	Result demonstrability Perceived Usefulness	**Supported**
H3b. Result demonstrability will have a positive effect on perceived ease of use of a community system.	Result demonstrability Perceived Ease of Use	**Supported**
H3c. Result demonstrability will have a positive effect on intention to continue use a community place.	Result demonstrability Perceived Ease of Use	**Supported**
H3d: Result demonstrability will have a positive effect on number of hits	Result demonstrability Number of Hits	Not supported
H4a: People who think the community place is just for information will have less social engagement activities than people who do not agree with it.	Attitude Social engagement	**Supported**
H4b. People's attitude towards online community will have a positive effect on perceived usefulness of the community place.	Attitude Perceived Usefulness	**Supported**
H4c. People's attitude towards online community will have a positive effect on intention to continue using the community place.	Attitude Intention to continue use	Not Supported
H4d: People who think the community place is just for information will have a lower level of perceived sense of community.	Attitude Sense of Community	**Supported**
H4e: People who think the community place is just for getting information will use the awareness tools less than others.	Attitude Self report usage on system features	**Supported**

Table 10.41 Results for Testing of Hypotheses (Continued)

Hypotheses	Variables	Result
H4f: People who think the community place is just for getting information will have a lower perceived usefulness of awareness tools.	Attitude Perceived usefulness on system features	**Supported**
H5a: Perceived sense of community will have a positive effect on perceived usefulness of the community place.	Sense of Community Perceived Usefulness	**Supported**
H5b: Perceived sense of community will have a positive effect on perceived ease of use of the community place.	Sense of Community Perceived Usefulness	**Supported**
H5c: Perceived sense of community will have a positive effect on intention to continue using the community place.	Sense of Community Intention to continue use	**Supported**
H5d: Perceived sense of community will have a positive effect on social engagement.	Sense of Community Social Engagement	Not supported
H6a: Perceived usefulness will have a positive effect on intention to continue using the community place.	Perceived Usefulness Intention to continue use	**Supported**
H6b: Perceived ease of use will have a positive effect on intention to continue using the community place.	Perceived Ease of Use Intention to continue use	**Supported**
H6c: Perceived ease of use will have a positive effect on perceived usefulness.	Perceived Ease of Use Perceived Usefulness	**Supported**
H6d: Perceived usefulness will have a positive effect on number of pages visited.	Perceived Usefulness Number of pages visited	**Supported**
H6e: Perceived ease of use will have a positive effect on number of pages visited.	Perceived Ease of Use Number of pages visited	Not supported
H6f: Intention to continue use will have a positive effect on number of pages visited.	Intention to continue use Number of pages visited	Not supported

Besides the above results, the analysis also shows that:

Higher intention to continue using the community place is related to a higher perceived usefulness of the community place.

Higher intention to continue using the community place is related to higher perceived ease of use of the community system.

Higher perceived usefulness of the community place is related to perceived ease of use of the community system.

These results are not included in our research model. They show an important feedback loop in the TAM model for adapting to an online community place.

The analysis for answering research questions shows that there is no difference in perceived usefulness or real usage for male and female members. Ph.D. students are the

user group who spent the most time and energy in the community place. People join the community expecting to obtain more information on their research topic and also know more people in the field and find possible collaborations in the community. Currently, the community system doesn't provide enough capability to connect people together. Members suggest adding some synchronous system features such as Instant Messenger (IM) or chat-room features to the community place. Members also hope to get email notification of updates on community contents in order to be able to know the updates of new knowledge in the community place.

Overall, the ALNResearch community place fulfilled its function as a place to hold related field knowledge and provide ways for people to get to know each other. However, the capability for connecting people is still weak and the knowledge base still needs to be expanded with more updated field knowledge. From the observation, very few contributions have been made by people outside of this research project, which suggests that such a research community needs regular maintenance and updating for its knowledge base. The initial expectation of having members themselves maintain the content of the community place is not happening and may not happen until the knowledge base is more complete and a specific strategy is carried out to motivate members to contribute. For example, a points-based system can be implemented so that if members do not contribute, they might not be able to use the knowledge base after a certain period of time.

In the next chapter, the Partial Least Square (PLS) method will be introduced and the results of using it to test the overall research model will be presented.

CHAPTER 11

TESTING THE RESEARCH MODEL

This chapter starts with the introduction of the Partial Least Square (PLS) path model method. This followed by description of the preparation of the data for PLS modeling and the testing of the proposed research model.

11.1 Partial Least Square Path Analysis

The PLS path analysis has been gaining a lot of interest among IS researchers in recent years (Compeau and Higgins 1995; Aubert, Rivard and Paltry 1994; Chin and Gopal 1995; Mathieson, Peacock, and Chin 2001; Yoo and Alavi 2001). When working in an area that has relative scarcity of theoretical knowledge, PLS path analysis with latent constructs is a useful and flexible tool for statistical modeling. It is especially useful when used in exploratory research rather than a confirmatory research situation.

PLS was developed in the 1960's by Herman Wold as an econometric technique for the estimation of path models involving latent constructs indirectly measured by multiple indicators (Word 1975, 1982. 1985), but it was used mostly among chemical engineers and chemometricians. Being a component-based structural equations modeling technique, PLS is similar to regression, but simultaneously models the structural path and measurement paths. PLS can be a powerful method of analysis because of the minimal demands on measurement scales, sample size, and residual distributions (Chin et al. 1996).

The following description of PLS models is taken from "State-of-art on PLS Path Modeling through the available software" (Chatelin, Vinzi and Tenenhaus 2002). A PLS

model is described by (1) a measurement model (or outer model) relating the manifest variables to their latent variable and (2) a structural model (or inner model) relating some endogenous latent variables to other latent variables. In the outer model, the relationship between the manifest variables and the latent variables can be detected in two ways. One is reflective way in which each latent variable ξ_j is indirectly observable by a set of manifest variables χ_{jh}. Each manifest variable is related to its latent variable by simple regression:

$$\chi_{jh} = \pi_{jh0} + \pi_{jh1}\xi_j + \epsilon_{jh}$$

where ξ_j has mean m_j and standard deviation 1. The other way is the formative way in which the latent variable ξ_j is generated by its own manifest variables. The latent variable ξ_j is a linear function of its manifest variables:

$$\xi_j = \sum_h \varpi_{jh}\chi_{jh} + \delta_j$$

For the inner model, every latent variable is related to every other latent variable by a simple regression:

$$\xi_j = \beta_{j0} + \sum_i \beta_{ji}\xi_i + \nu_j$$

A latent variable, which never appears as a dependent variable, is called an exogenous variable. Otherwise, it is called an endogenous variable. The causality model must be a causal chain, which means that there is no loop in the causality model.

Compared to other well known covariance fitting approaches, such as Maximum Likelihood, the component-based PLS avoids two serious problems: inadmissible solutions and factor indeterminacy (Fornetll and Bookstein 1982). The basic distinction

between these methods is whether it is used for theory testing or for predictions (Anderson and Gerbing 1988). When there exists strong theory and the goal is to further test and develop the model, covariance based full-information estimation methods are more appropriate (e.g. Maximum Likelihood or Generalized Least Squares). However, these methods are not good for prediction due to the indeterminacy of factor score estimations. For application and prediction a PLS approach is more suitable. It is assumed that all the measured variance is useful variance to be explained. Since the approach estimates the latent variables as exact linear combinations of the observed measures, it avoids the indeterminacy problem and provides an exact definition of component scores. Using an iterative estimation technique (Word 1982), the PLS approach provides a general model which encompasses, among other techniques, canonical correlation, redundancy analysis, multiple regression, multivariate analysis of variance, and principle components.

As a consequence of using an iterative algorithm consisting of a series of ordinary least squares analyses, it does not presume any distribution form for measured variables. Further more, it does not require large sample size (Barclay et al. 1995; Fornell and Bookstein 1982). Because of the minimal requirement on data and capability of testing complex models (Fornell, Lorange, and Roos 1990; Fornell and Bookstein 1982), PLS has been used widely in various research areas recently.

The validation of the model can be measured by cross-validated R-square between each endogenous latent variable and its own manifest variables. The significance levels of the regression coefficients can be computed (Chatelin et al. 2002).

The most popular PLS application in the IS field has been PLS-GRAPH, developed by Dr. Wynne W. Chin. PLS-Graph is based on LVPLS, which was developed in the 1980s by Dr. Lohmöller. It was a DOS-based program. Mr. Yuan Li of the University of Southern California developed a GUI interface for LVPLS, and named it PLS-GUI. PLS-Graph also uses LVPLS as the base. It extended LVPLS by adding assessment of the significance level of PLS parameters using Bootstrap and Jackknife resampling procedure (Cotterman and Senn 1992). The bootstrapping procedure is selected in this research to assess the significance level of coefficient. The bootstrap samples are built by resampling with replacement from the original sample. The procedure produces samples consisting of the same number of cases as in the original sample. Both standard error and T statistics of path coefficient are given as the result of the bootstrapping procedure.

11.2 PLS Results for the Research Model on Survey Items

PLS-Graph 3.0 beta version developed by Dr. Chin has been used for PLS analysis in this research. All 81 questionnaire responses has been used for the analysis, responses with missing values has been dealt by the PLS algorithms. Before explaining the result from PLS analysis, the original proposed relationships between constructs and corresponding hypothesis is presented in Figure 11.1. The meaning of abbreviations in the figure are :

RD: Result demonstrability

OQ: Output Quality

ER: Expectation Relevancy

AT: Attitudes

SC: Sense of Community

RU: Real Usage

PEU: Perceived Ease of Use

PU: Perceived Usefulness

IU: Intention to continue use

PN: Preliminary Networking

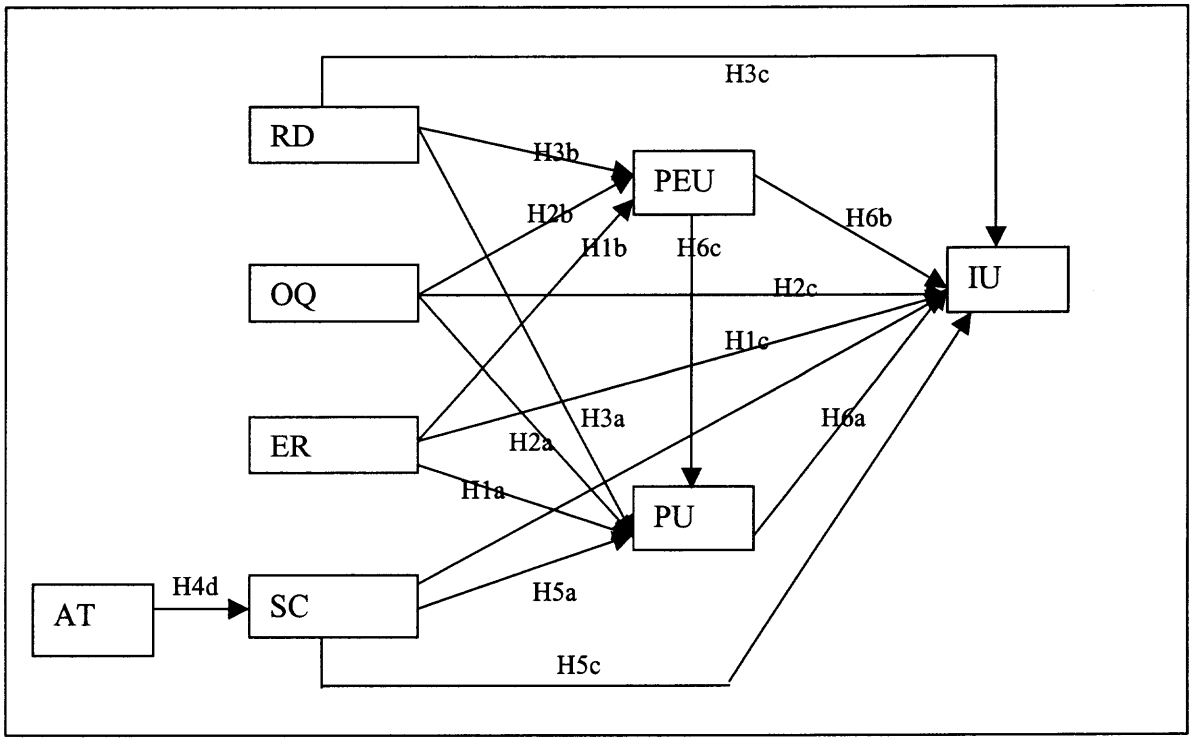

Figure 11.1 Proposed Research Model and Hypotheses

11.2.1 Test of Measurement Model

The measurement model is tested by examining (1) internal consistency, (2) individual item reliability, and (3) convergent and discriminant validity.

In this research, reflective measures are used for outer model. In this case, rather than using Cronback's alpha, which represents a lower bound estimate of internal consistency due to its assumption of equal weighting of items, a better estimate can be obtained using the composite reliability (Chin 1998, Fornell and Larcker 1981).it is not

influences by the number of items in the scale as opposed to alpha. It is based on the ration of construct variance to the sum of construct and error variance.

Table 11.1 Composite Reliability for Latent Variables

Constructs	Composite Reliability
Output Quality	**0.895**
Expectation Relevancy	**0.870**
Result demonstrability	**0.824**
Attitude	**0.862**
Sense of Community	**0.802**
Perceived Ease of use	**0.876**
Perceived Usefulness	**0.867**
Intention to continue use	**0.933**
Preliminary Networking	**0.773**

The result in the above table was obtained from PLS-GRAPH 3.0 Beta version, obviously all composite reliabilities are greater than 0.7. This result indicates that the construct variance is accounts for at least 70% of the measurement variance.

Individual item reliability is assessed by examining the loadings of the manifest variables on their corresponding latent variables. A rule of thumb (Rivard and Huff 1988) is to accept items higher than 0.7, which implies that there is more shared variance between the latent variable and its manifest variable than error variance. However, it is also very important to retain as many items as possible from the original scale. For this

research, items with loadings higher than 0.65 are accepted. From the original questionnaire items, Q3 and Q22 are removed in order to achieve the individual item reliability. Table 11.2 shows the individual item reliability of all manifest variables (AT1 and SC3 has been removed due to the low loading).

Table 11.2 PLS Loadings

Variables	*OQ*	*RD*	*AT*	*SC*	*PEU*	*PU*	*IU*	*PN*
OQ1	0.89							
OQ2	0.91							
RD1		0.82						
RD2		0.80						
RD3		0.72						
AT2			0.93					
AT3			0.80					
SC1				0.96				
SC2				0.65				
PEU1					0.92			
PEU2					0.91			
PEU3					0.67			
PU1						0.79		
PU2						0.81		
PU3						0.69		
PU4						0.85		
IU1							0.94	
IU2							0.93	
PN1								0.71
PN2								0.87

Convergent validity can be obtained from average communality measure from the PLS-GUI results (see Table 11.3 for details).

Table 11.3 Convergent Validity

Constructs	Average Communality
Output Quality	0.81
Result demonstrability	0.61
Attitudes	0.75
Sense of Community	0.67
Perceived Ease of Use	0.71
Perceived Usefulness	0.62
Intention to continue use	0.87
Preliminary Networking	0.63

In order to evaluate discriminant validity, Fornell and Larcker (1981) suggested a comparison between the Average Variation Extracted (AVE) and the variance shared

between the constructs (the squared correlations between the constructs). AVE is obtained by the sum of loadings squared, divided by the number of items in the constructs. The AVE should be higher than 0.5 (Rivard et al. 1993). The number on the diagonal should be higher than the elements off diagonal. Table 11-5 shows the comparison of the square root of AVE and correlation between constructs.

Table 11.4 Discriminant Validity

Variables	OQ	ER	RD	AT	SC	PEU	PU	IU	PN
OQ	**0.9**								
ER	0.567	**0.877**							
RD	0.68	0.667	**0.781**						
AT	0.256	0.323	0.269	**0.87**					
SC	0.452	0.285	0.367	0.257	**0.822**				
PEU	0.684	0.549	0.632	0.327	0.424	**0.84**			
PU	0.704	0.782	0.742	0.287	0.384	0.681	**0.79**		
IU	0.663	0.663	0.652	0.321	0.457	0.525	0.728	**0.935**	
PN	0.539	0.539	0.499	0.452	0.438	0.537	0.614	0.555	**0.795**

In table 11.4, values on the diagonal are the square root of AVE values, and values off-diagonal are correlations. The largest correlation off-diagonal is 0.74. The lowest AVE value is 0.79. Hence, the smallest on-diagonal value is larger than largest off-diagonal value, which shows that both convergent and discriminant validity appeared to be satisfactory.

11.2.2 Assessment of the Structural Model

After checking validity and reliability, the adjusted data can be used to assess the structural model. Figure 11.2 shows the path correlations of the hypothesized relationships (n=81). The estimated path effects are given with their significance level. A bootstrapping procedure was used to assess the level of significance of the coefficient computed by PLS. T-values were given as a result of bootstrapping from PLS-Graph, it is then changed into significance level.

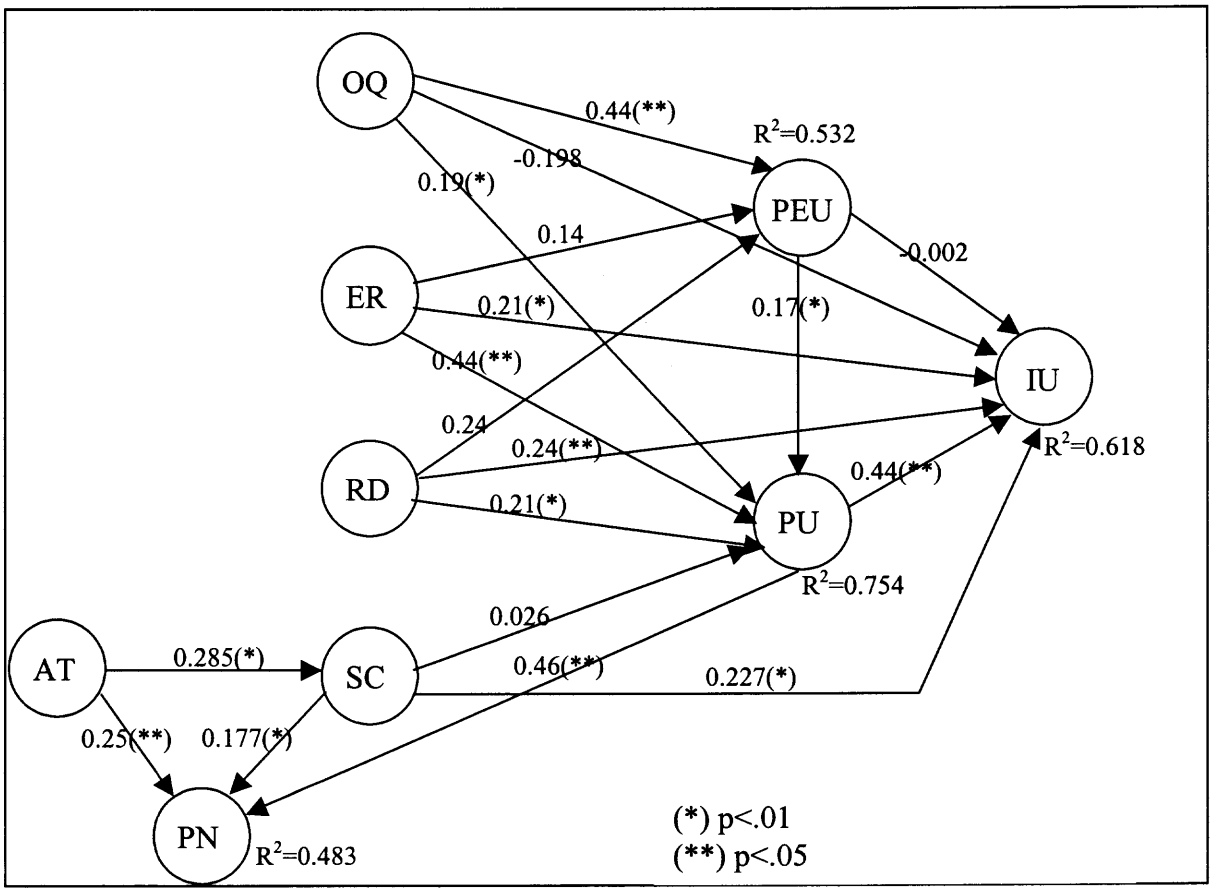

Figure 11.2 Path Diagram of the Original Research Model

As can be seen in Figure 11.2, the percentage of variance explained (R^2) of "Perceived Usefulness" was 75%, the percentage of variance explained of "Perceived Ease of Use" was 53%, the percentage of variance explained of "Intention to continue use" was 62%, and the variance explained of "Preliminary Networking" was 48%. These values were very satisfactory considering the complexity of the research model. Surprisingly, "Perceived Ease of Use" doesn't directly influence "Intension to Use" as proposed in the TAM model. Also "Sense of community" doesn't directly influence "Perceived Usefulness". Relationship between "Output Quality" and "Intension to Use" is negative instead of positive, but it is not significant. "Result demonstrability" is not significantly influence "Perceived Ease of Use". Figure 11.3 shows the path analysis result after removing all insignificant links.

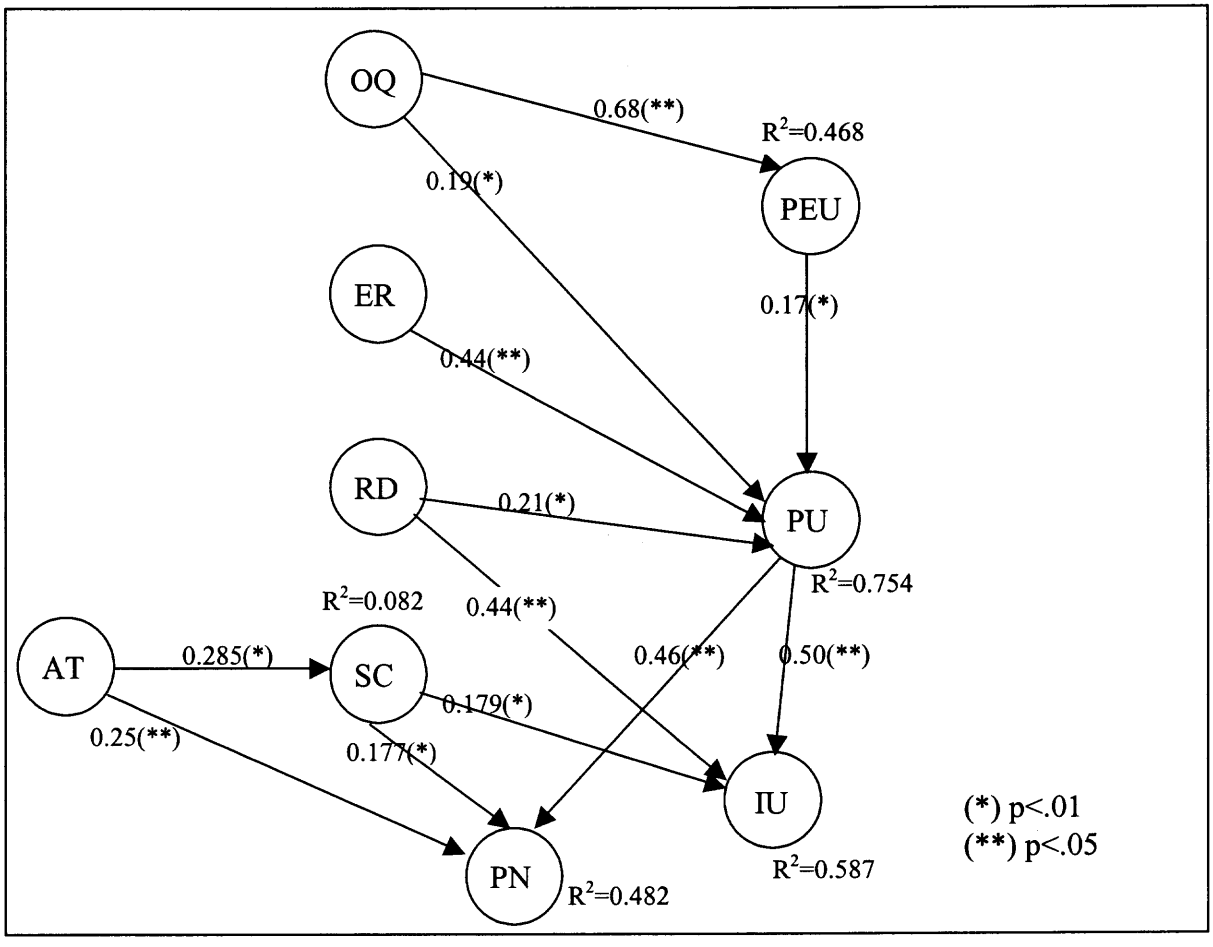

Figure 11.3 Simplified Model with All Significant Relations

After removing the insignificant links and rerun PLS-Graph, R^2 did not change much from figure 11.2, which also confirms that the removal links were not very important.

"Expectation Relevancy" has a significant effect on "Perceived Usefulness" with a path coefficient of 0.44 ($p<0.01$). This result is consistent with the TAM2 model. However, "Expectation Relevancy" does not directly influence "Intention to continue use" as expected in hypotheses. The results relate to "Expectation Relevancy" indicate that:

H1a. Expectation relevancy has a positive influence on perceived usefulness of a community place.

Supported

H1b. Expectation relevancy doesl not have a positive effect on perceived ease of use of a community system.

Supported

H1c. Expectation relevancy has a significant positive influence on intention to continue using a community place.

Not Supported

The relationship between expectation relevancy and intention on continued use indicates that although expectation relevancy serves as a crucial factor in users' decision-making process to eliminate unrelated alternatives, it does not significantly influence continued usage after users adopt the community place. This result also confirms that factors that influence pre-adoption and post-adoption are different (Agarwal and Prasad, 1997; Karahanna et al., 1999).

"Output Quality" has a very significant effect on "Perceived Ease of Use" with a path coefficient of 0.68($p<0.01$), and it also has a significant effect on "Perceived Usefulness" with a path coefficient of 0.19 ($p<0.05$). This result is different from the TAM2, in which "Output Quality" only influences "Perceived Usefulness". In this research context, "Output Quality" contributes more to "Perceived Ease of Use" than "Perceived Usefulness". It indicates that when people judge whether an online community system is easy to use or not, the quality of the community content influences people's perception of the usability of the system. The above results indicate:

H2a. Output quality has a positive effect on perceived usefulness of a community place.

Supported

H2b. Output quality does not have a positive effect on perceived ease of use of a community system.

Not Supported

H2c. Output quality has a positive effect on intention to continue using the community place.

Not supported

The result on relationship between output quality and perceived ease of use of a community system is very interesting. As output quality measures the content and perceived ease of use measures the interface of the community place, they should not have any significant relationship. Current results show that in this type of community, which draws attention from users mainly due to its content, the content may also influence peoples image of the system interface. The relationship between output quality and intention to continue use again shows that factors that influence users' choice of adoption may not influence their intention to continue use. However, whether output quality will have a significant influence on an online community's initial adoption needs future study.

"Result demonstrability" has a positive effect on "Perceived Usefulness" with a path coefficient of 0.21($p<0.05$). It also has a positive effect on "Intention to continue use" with a path coefficient of 0.44 ($p<0.01$). "Result demonstrability" had not effect on "Perceived Ease of Use". Interestingly, it has a stronger relationship with "Intention to continue use" than "Perceived Usefulness". While in the TAM2 model, "Result demonstrability" only is an indicator of "Perceived Usefulness". It doesn't have a direct effect on "Intention to continue use" The above results indicate:

H3a. Result demonstrability has a positive effect on perceived usefulness of a community place.

Supported

H3b. Result demonstrability does not have a positive effect on perceived ease of use of a community system.

Supported

H3c. Result demonstrability has a positive effect on intention to continue using the community place.

Supported

The above result is consistent with Agarwal and Prasad's (1997) study. Result demonstrability has a significant positive effect on both perceived usefulness of the community place and intention to continue using the community place. This result indicates that how well a community place shows its function in helping members' achieve their goals is very important in sustainability of the community usage.

"Sense of Community" has positive relationships with both "Intention to continue use" and "Preliminary Networking" at 0.05 significance level. However, it doesn't have a direct effect on "Perceived Usefulness". These results indicate that:

H5c. Perceived sense of community has a positive effect on perceived usefulness of the community place.

Not Supported

H5d. Perceived sense of community has a positive effect on perceived intention to continue using the community place.

Supported

H5e. Perceived sense of community has a positive effect on perceived Preliminary networking.

Supported

The lack of relationship between perceived sense of community and perceived usefulness of the community place again confirms that in a knowledge-based community, information needs take much more account than social needs. Even if the sense of community is not very high, people will still think the community place is quite useful. This would be very different from communities whose purpose is letting members help each other.

"Attitude towards the online community" has a positive effect on "Sense of Community" with a path coefficient of 0.285 ($p<0.05$). It also has a positive effect on "Preliminary Networking" with a path coefficient of 0.25 ($p<0.01$). The above results confirm the following hypotheses:

H4d: People who think the community place is just for information will have a lower level of perceived sense of community.

Supported

H4c: People's attitude towards online community will have a positive effect on preliminary networking.

Supported

The above results are consistent with the results obtained from bivariate analysis, see Table 10.8 to Table 10.10 for comparison.

The results on basic the TAM model constructs are not consistent with the original model. In the original model, both "Perceived Usefulness" and "Perceived Ease of Use" are directly related to "Intention to continue use". However, in this research context, only "Perceived Usefulness" has a direct effect on "Intention to continue use" with a path coefficient of 0.50 ($p<0.01$). Although bivariate analysis shows that perceived ease of use has a significant correlation with intention to continue use, from the PLS analysis, "Perceived Ease of Use" doesn't have a direct influence on "Intention to continue use". This result is consistent with some other studies on technology acceptance in which perceived ease of use became insignificant after the system was initially adopted or when users were experienced users (Agarwal and Prasad 1997, 1998; Tompson et al. 1991, 1994; Karahanna et al., 1999; Venkatesh et al., 2003). "Perceived Usefulness" also positively related to "Preliminary Networking" with a path coefficient of 0.48 ($p<0.01$). These results indicate:

H6a: *Perceived usefulness has a positive effect on intention to continue using the system.*

Supported

H6b: *Perceived ease of use does not have a positive effect on intention to continue using the system.*

Supported

H6c: *Perceived ease of use will have a positive effect on perceived usefulness.*

Supported

H6e: Perceived Usefulness will have a positive effect on Preliminary Networking

Supported

11.3 Testing Research Model including Real Usage Data

After testing the research model based on all subjective measures (survey response), real usage data is added to the model to test relationships between subjective constructs and people's real behavior. As real usage contains only one manifest variable, the measurement model will be the same as previously. After adding real usage (RU) data and rerun PLS-Graph, the only significant link that was found is between "Perceived Usefulness" and "Real Usage". All other construct did not have any direct effect on the "Real Usage" data. PLS result of the structural model is shown in Figure 11.4.

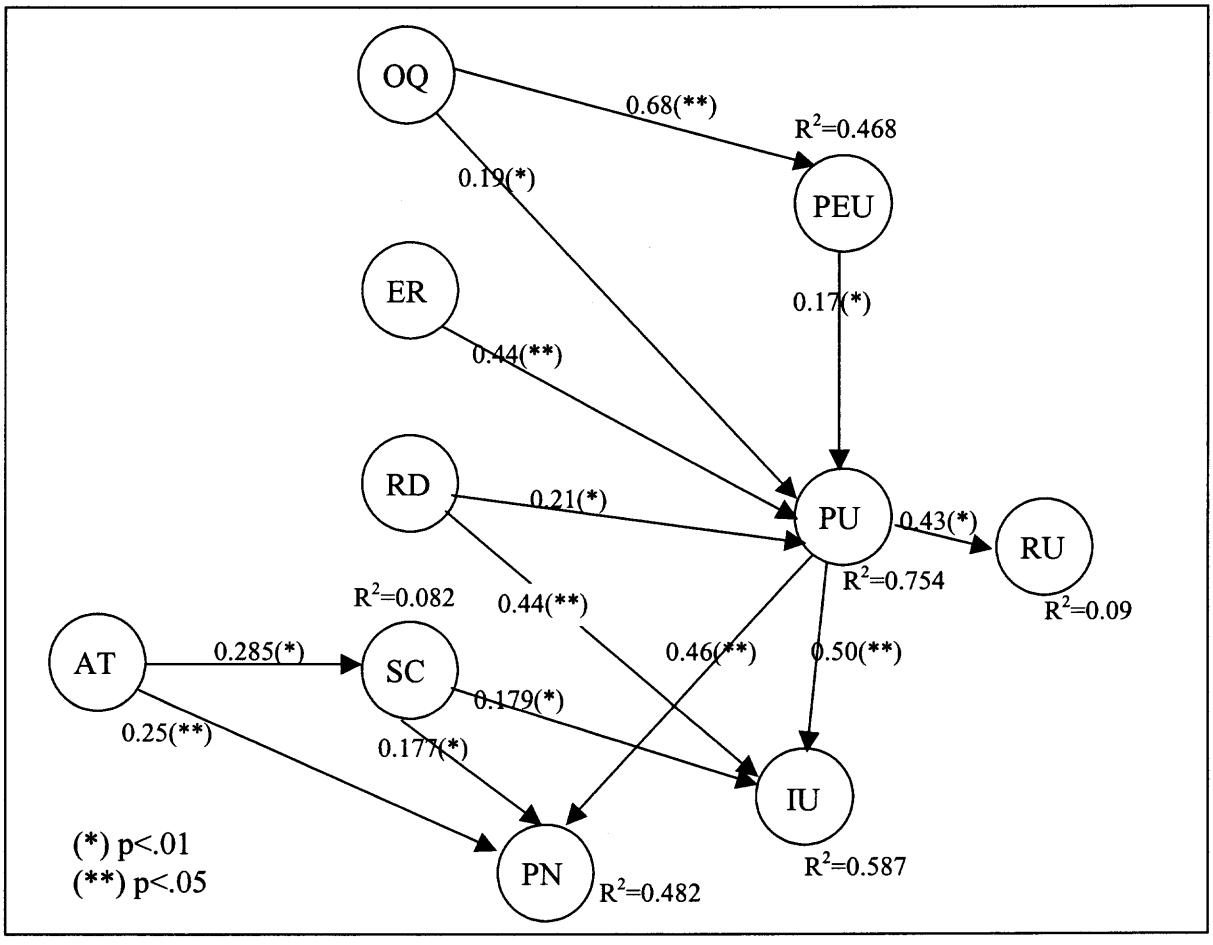

Figure 11.4 Path Analysis with Real Usage Data

Figure 11.4 shows the result after removing insignificant coefficient. The variance of real usage is not satisfactorily explained from the model (R^2=0.09). This finding is consistent with the result of association test in Chapter 10, where there seems to be no relationship between real usage data and subjective measures. From the PLS result including real usage data, the following hypotheses can be tested:

H1d. Expectation relevancy will have a positive effect on real usage of a community place.

Not Supported

H2d. Output quality will have a positive effect on real usage of a community place.

Not Supported

H3d. Result demonstrability will have a positive effect on real usage of a community place.

Not Supported

H5b. Perceived sense of community will have a positive effect on member's social engagement activities.

Not Supported

H6d: Intention to continue using the system will have a positive effect on real system usages.

Not Supported

H6f: Perceived Usefulness will have positive effect on real usage of the community system.

Supported

11.4 Summary

In this chapter, the complete research model was tested by Partial Least Square Path modeling with PLS-Graph 3.0 beta. The model was found to explain relationships among subjective measures satisfactorily, but not on successively explaining real usage data (R^2=0.09). The results suggest that "Output Quality" represents an important source of "Perceived Ease of Use" of the community system (R^2=0.47). "Output Quality", "Expectation Relevancy", and "Result demonstrability" are important indicators for "Perceived Usefulness" of the community space (R^2=0.75). "Perceived Usefulness", "Result demonstrability" and "Sense of Community" represent sources of "Intention to continue use" (R^2=0.58). Both similarity and differences were found between the TAM model and the proposed model for factors that influence users' selection of an online community space.

First, different from the TAM model where "Output Quality" is only an indicator for "Perceived Usefulness", "Output Quality" was found to contribute more to "Perceived Ease of Use" than "Perceived Usefulness" in this research context. These results indicate

that in an online research community, people regard the quality of the community content as very important. It even influences the perception of usability of a community system.

Secondly, different from the TAM mode where both "Perceived Ease of Use" and "Perceived Usefulness" directly influence the "Intention to continue use", only "Perceived Usefulness" directly influences the "Intention to continue use", while "Perceived Ease of Use" only has a positive effect on "Perceived Usefulness". This result further confirms that when considering participate in an online community, the content of the community system is more important then the usability of the community system.

The following table summarized all the results of testing of hypotheses.

Table 11.5 Results for Testing of Hypotheses

Hypotheses	Variables	Result
H1a: Expectation relevancy has a positive effect on perceived usefulness of a community place	Expectation Relevancy Perceived Usefulness	**Supported**
H1b: Expectation relevancy doesn't have a positive effect on perceived ease of use of a community system	Expectation Relevancy Perceived ease of Use	**Supported**
H1c: Expectation relevancy has a positive effect on intention to continue using the community place	Expectation Relevancy Intention to continue use	Not Supported
H1d: Expectation relevancy has a positive effect on real usage of the community space	Expectation Relevancy Number of pages visited	Not Supported
H2a: Output Quality has a positive effect on perceived usefulness of a community place	Output Quality Perceived Usefulness	**Supported**
H2b: Output Quality does not have a positive effect on perceived ease of use of a community system	Output Quality Perceived Ease of Use	**Not** Supported
H2c: Output Quality has a positive effect on intention to continue using the community place	Output Quality Intention to continue use	Not Supported
H2d: Output Quality has a positive effect on real usage of the community space	Output Quality Number of pages visited	Not Supported
H3a. Result demonstrability has a positive effect on perceived usefulness of a community place.	Result demonstrability Perceived Usefulness	**Supported**
H3b. Result demonstrability does not have a positive effect on perceived ease of use of a community system.	Result demonstrability Perceived Ease of Use	**Supported**

Table 11.5 Results for Testing of Hypotheses (Continued)

Hypotheses	Variables	Result
H3c. Result demonstrability has a positive effect on intention to continue use a community place.	Result demonstrability Perceived Ease of Use	**Supported**
H3d: Result demonstrability has a positive effect on number of hits	Result demonstrability Number of pages visited	Not supported
H4b. People's attitude towards online community has a positive effect on perceived usefulness of the community place.	Attitude Perceived Usefulness	Not Supported
H4c. People's attitude towards online community has a positive effect on intention to continue using the community place.	Attitude Intention to continue use	Not Supported
H4d: People who think the community place is just for information has a lower level of perceived sense of community.	Attitude Sense of Community	**Supported**
H4g: People's attitude towards the online community has a positive effect on perceived preliminary networking	Attitude Preliminary Networking	**Supported**
H5c: Perceived sense of community has a positive effect on perceived usefulness of the community place.	Sense of Community Perceived Usefulness	Not Supported
H5d: Perceived sense of community has a positive effect on intention to continue using the community place.	Sense of Community Intention to continue use	**Supported**
H5e: Perceived sense of community has a positive effect on perceived preliminary networking.	Sense of Community Preliminary Networking	**Supported**
H6a: Perceived usefulness has a positive effect on intention to continue using the community place.	Perceived Usefulness Intention to continue use	**Supported**
H6b: Perceived ease of use does not have a positive effect on intention to continue using the community place.	Perceived Ease of Use Intention to continue use	**Supported**
H6c: Perceived ease of use has a positive effect on perceived usefulness.	Perceived Ease of Use Perceived Usefulness	**Supported**
H6d: Intention to continue use has a positive effect on number of pages visited.	Intention to continue use Number of pages visited	Not Supported
H6e: Perceived usefulness has a positive effect on number of pages visited.	Perceived Usefulness Number of pages visited	**Supported**
H6f: Perceived Usefulness has a positive effect on perceived preliminary networking.	Intention to continue use Number of pages visited	**Supported**

It seems that there is no strong correlation between real usage data and people's perceptions, which suggests that other methods such as interviews may be used to better understand relationships between people's real behavior and their perceptions. The other reason might be time limitation. As mentioned in previous chapter, many people who

responded to the survey questionnaire are administrators. They might be too busy to spend a lot of time in the community space although they might want to use it more. This factor should be considered in future research on this topic.

The next chapter will summarize all the findings from this research, discuss the results, and give future research directions.

CHAPTER 12

SUMMARY AND DISCUSSION

This chapter will summarize the major findings of this study, discuss the results in terms of theoretical and practical implications, discuss the limitations of this study, and possible future research directions.

12.1 Summary of Conclusions

The main objectives of this study are to:

- Understand users behavior in a knowledge-based research community
- Design and evaluate technologies that serve members' social and information needs in a knowledge-based community
- Propose and test a research model on factors that influence members' engagement in a knowledge-based research community

12.1.1 Members' Behavior in the Community

The following conclusions are found from the web usage mining with the web server log.

- There were over 5,800 articles accessed during a four month period by members in the community. The average page views per day are 113 by members during the four month period from 10/28/2003 to 5/25/2004. Considering an average of only six members visit the community place per day, the above number shows that members who visit the community place have spent considerable time and effort in the community space.
- From 10/28/2003 to 4/14/2004, there are 379,984 successful hits to the community place, the average per day is 2,316. These hits include web crawlers,

and users who visited the homepage or accessed the html pages in the website without login as a member.

- Among all page views on dynamic pages from members, 55% of them are on "Empirical Research" pages. Visits to "Other Research", "People Directory", and "Discussion Forums" took 10% each. This information shows that the "Empirical Study" feature is the most useful feature of all. It also shows that although information seeking is the major activity in the community place, members still show their needs for social engagement as they try to get to know other members in the community. The ratio between information seeking activities and social engagement is 7:3.
- Visitors who come to the community place are from all over the world. Members have been increased greatly since the implementation of the new system. The number of registered members grew from 184 (before 10/28/2003) to 765 (5/25/2004). However, as there is no push technology (e.g. email notification) on new contents of the community place, most members do not visit the community site very frequently.

12.1.2 Effectiveness of the Designed Technology

In order to better serve members information and social needs, a recommender system and awareness tools have been designed and implemented. The awareness tools were implemented on 12/27/2003 and the recommender system was implemented on 12/20/2003.

12.1.2.1 Awareness Tools. Although there are not many postings from the members to the community place, members check the "people directory" frequently (10% of all page

views) to look at other people's research interests. Messages posted in the discussion forums or comments on research articles have been visited by members occasionally.

In the survey questionnaire, one question asks whether the subject started to contact people they got to know in the ALNResearch WebCenter. Among 78 valid responses, 11 of them answered "yes" and 67 answered "no", which shows that 14% of the subjects started networking with other members they got to know from the community place.

Although there is not much interaction in the community place among community members, the preliminary networking activity such as looking at other's profiles, checking others' messages, or contacting other members through email shows the need for such awareness tools in a research community. When asking about members' expectations in the questionnaire, some subjects expressed their expectations to get to know other people in the same field and start collaboration if possible. Some members also desire adding synchronous communication tools such as chat-room on the ALNResearch website to communicate with others.

12.1.2.2 Recommender System. After the new system was implemented, a message was sent to two research email lists in November 2003 to announce the existence of the WebCenter. Thus, visits during November were the highest among all. The recommender system was implemented on 12/20/2003. Because December and January are final exam period and winter vacation, there were not many visits during these two months. Visits to the ALNResearch community have been stable since the end of January. Because of the above reason, use comparison to articles before and after implementation of the recommender system is not reliable.

Content analysis on the web server log shows that among all 5,879 page views of accessing articles, 1267 of them were directed from the recommender system feature, which is 22% of all page views of accessing articles. One reason for not using the recommender system is because users must read a sufficient number of articles (at least 5) in the knowledge base before they can be included. The other reason is that there are only 170 empirical articles and about 50 other research articles or books in the knowledge base, which is not big enough to show the effectiveness of the recommender system. 22% is a promising number in consideration of the short period of implementation of the recommender system, and the small number of articles in the knowledge base.

12.1.3 Testing Hypotheses

Relationships among all individual related constructs are tested with Partial Least Square (PLS) Path Modeling method. This method is selected because it doesn't require large sample size, and it does not require normal distribution of the sample data. The following diagram shows the result from PLS analysis:

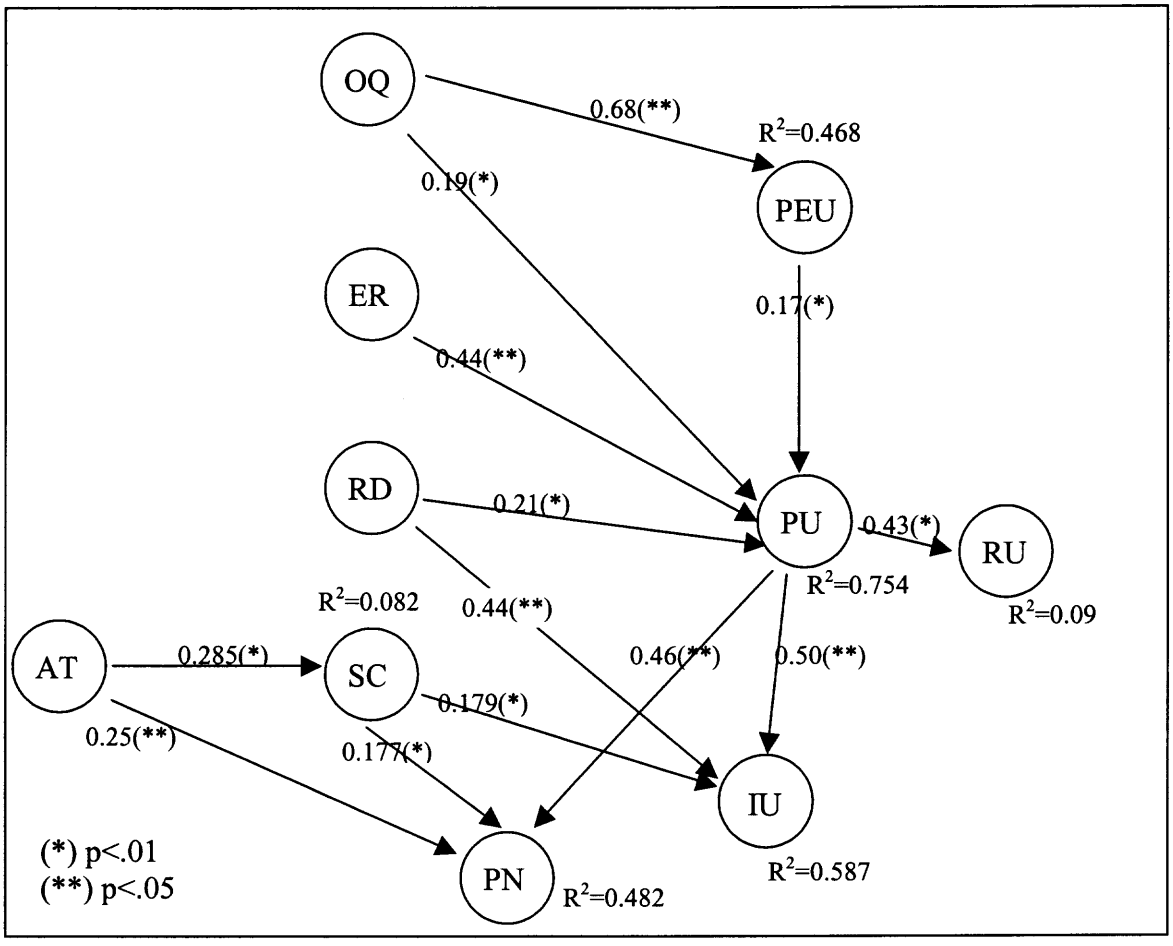

Figure 12.1 Structure Model from PLS Analysis

The following conclusions are drawn from the structure model test results:

Expectation relevancy as an important component in people's decision making is found to have a significant positive influence on perceived usefulness of the community place. This result is consistent with the extended TAM model (Venkatesh and Davis, 2000) in which expectation relevancy functions as an indicator for perceived usefulness of the technology.

In this study, output quality is found to have a significant positive influence on both perceived usefulness of the community place and perceived ease of use of the community system. The relationship with perceived usefulness is consistent with extended TAM (Venkatesh and Davis, 2000). However, the relationship between output

quality and perceived ease of use is very surprising as output quality measures the content of the community place, while perceived ease of use measures the interface of the community system. This finding shows that people's perception about the content of the community place may influence their impression about the community system. In other words, if they find the content is of high quality to them, then the interface may seem to be very friendly to them too.

People's attitude toward the online environment is found to have a significant positive influence on their perceived sense of community level and also their perceived preliminary networking activities. This is not surprising as peoples' initial attitude toward the online community may influence their willingness to post or to communicate with others in the community.

Perceived sense of community is found to have a significant positive influence on both perceived preliminary networking and intention to continue using the community place. The relationship between perceived sense of community and perceived preliminary networking is reasonable, as if people who have a higher level of sense of community would logically be more likely to intend to start preliminary networking with other members in the community. The relationship between perceived sense of community and intention to continue using the system confirms that the need for being a community, not just information seeking, is important in people's intention on continued use.

Among the three constructs in the basic TAM model, perceived ease of use is found to have a significant positive influence on perceived usefulness of the community place, and perceived usefulness is found to have a significant influence on intention to continue using the community place. This result is consistent with most studies on

technology acceptance or innovation diffusion. The relationship between perceived ease of use and intention to continue using the community place is not shown as significant. Although this result is different from the TAM and TAM2 model, a lot of empirical studies did find similar results between perceived ease of use and intention on post-adoption (Agarwal and Prasad 1997, 1998; Tompson et al. 1991, 1994; Karahanna et al., 1999; Venkatesh et al., 2003). The function of perceived ease of use decreased dramatically after technology was initially adopted.

In this study, the dependent variable – real usage - is not explained satisfactorily. Only 10% of the variance is explained. One possible reason may be the different measuring scale. All other constructs are measured using an interval scale (likert-scale type); only real usage is measured by a ratio scale. If the number of pages visited be converted into an interval scale, it might cause a different result. Another reason is that there are other constraints that may influence users' real usage, e.g. time constraints, accessibility, etc.

Some additional hypotheses were tested by paired correlation and one-way ANOVA. They are listed below:

- H4a: People who think the community place is just for information will have less social engagement activities than people who do not agree with this view.
- H4e: People who think the community place is just for getting information will use the awareness tools less than others.
- H4f: People who think the community place is just for getting information will have a lower perceived usefulness of awareness tools.

The above results show that there are strong correlations among individual factors themselves (from questionnaire data), but only "perceived usefulness" has a significant effect on real usage of the community place. This is different from the original TAM model as the variable which correlated with real usage is the "intention to continue use" in the original TAM model. Also, different from the original TAM model, perceived ease of use does not have a direct effect on intention to continue using the community place.

For both PLS modeling and individual association tests, we did not find any significant predictor for real usage data. Only 9% of the variance of real usage data can be explained by the proposed model. "Perceived Usefulness" is explained very well in the model with 75% of variance been explained. "Intention to continue use", "Perceived Ease of Use", and "Preliminary Networking" also has a satisfactory level of variance explained by the model.

12.2 Contributions and Limitations

12.2.1 Contributions

This study has made the following contributions to the online community area.

1. This study is one of the few studies that use real usage data to understand users' behavior in a knowledge-based community that does not have much interaction among the members.
2. This study is the first study which adopts the TAM model to explain factors that influence users' decision of staying in an online community.
3. This study shows the usefulness of such knowledge-based online communities for the research community. A lot of members in the community expressed their

appreciation of the site. They thought this is such a great resource for their research. It shows that researchers need a knowledge-base which contains updated and synthesized knowledge. They also need a place to get to know other researchers in the same research area who are not in the same geographic region.

4. This study shows that in such a research community, information seeking is the dominant activity. Thus, features that may help information seeking such as search engines are very important in this type of community system.

5. This study shows that there are no gender differences in behavior patterns in such online communities.

6. This study shows that members' attitude towards a research online community will directly influence their social engagement in the community place. Organizations who are interested in forming such communities should specifically address the credibility of the community place, and carry out strategies to increase trust between members and the community.

7. This study empirically tests a research model to explain members' intention on continued use of a community place.

8. This research shows the importance of providing awareness tools in an online community to serve members' social needs.

9. This research shows a framework for building a knowledge-based community system, and suggests possible features for such a community system.

10. From PLS analysis, "Output Quality" is a strong predictor for "Perceived Ease of Use" which indicates that the quality of community content will have an impact

on users' perception of the community system. "Perceived Ease of Use" does not directly affect "Intention to continue use", while "Result demonstrability", "Perceived Usefulness", and "Sense of Community" are directly influencing "Intention to continue use". These results tell us that the content of a knowledge based research community and its atmosphere of being a community is very important to attract its members instead of the system interface. Thus, the designer and administrator of such communities should spend more effort in maintaining high quality and updated content, and creating real community atmosphere in virtual space.

11. This research shows that in such a knowledge-based community, it is hard to get members to contribute to the knowledge base. In order to keep updated field knowledge, such a community system needs daily maintenance on contents by field experts. In order to increase the interaction inside the community, other strategies such as a reward system should be implemented.

For the ALN community, this research helps to improve its cohesiveness, enhance knowledge dissemination within the community, provides guidance to those beginning to conduct research in this field, increase opportunities for collaboration among researchers, and improve the quality of work in this community.

For general IS research, this study applied and tested the technology acceptance model in a new context. It also validated measures of the constructs in this new research context, with modifications to fit the application.

12.2.2 Limitations

Despite the above contributions, this research also bears the following limitations. First, because this is a field study in a small community, there was no pilot study to test the survey questionnaire. Although a rough external validity check had been done before distribution of the questionnaire, there are still confusions in the statements of questionnaire items that caused the final constructs not to form as designed.

Secondly, the response rate of the questionnaire is very low since the subjects are totally volunteers to complete the questionnaire. Among 350 members who have been asked, only 80 of them completed the questionnaire, which caused the power of statistical tests to be very low.

Third, as this is a field study, the subjects and their behavior cannot be controlled. During the academic final period and winter break, there were not many activities in the community place. During the initial release of the new system, advertisement messages were sent to users of two research email-lists, which caused lots of visits during the initial months. Thus, the comparison of usage among different months is not reliable.

Fourth, the customized log file did not capture all activities in the community. The customized log file only records visits to dynamic pages, not static pages like html files.

Fifth, there is not enough data to show influence of the recommender system. After removing data from winter break, there is only one-month usage data to measure the recommender system. Also, the questionnaire did not contain any questions concerning users' perception of the recommender system. The questionnaire can be

modified to include those question items and we will keep collecting data for future papers.

Finally, this is a study of only one research community and one specific software environment. Several replications particularly with larger knowledge bases and community size, would be necessary in order to generalize the findings.

12.3 Future Research

One of the biggest concerns about this study is that we do not have enough data to really check the effectiveness of a recommender system in such a knowledge-based system. Also, from both open-ended questions and the real usage of the community place, there appears to be a strong need to send notifications when new content has been added to the community place. For the next step, we need to look at other similar knowledge based communities, such as the research community in the astronomy field. We need to check both features and strategies that work and don't work for this type of community combined with lessons from the ALNResearch community. These lessons should then be put into upgrading the existing software. For example, members should be able to request notification of new contents or new recommendations through email; and the search engine should support search on more fields. In order to test the effectiveness of system features, other than log file data, qualitative data such as interviews should be used to gain a better understanding.

This research has developed a system model for knowledge-based communities and a research framework to present factors that may influence members' engagement in a research community. However, the system model and research framework have been tested only using one field site. In order to offer more generality, similar research

communities with similar system structures should be tested. By the end of this research, we already received several requests for the same system model to be used for other research communities. The implementation of those communities will help in testing the generality of the proposed system features and research framework.

By using PLS, we found some very interesting relationships among constructs. For example, "Perceived Ease of Use" is not related to "Intention to continue use" as suggested in the TAM model; "Output Quality" has a stronger impact on "Perceived Ease of Use" than "perceived Usefulness". Whether these relationships hold in general or just for this particular research community needs to be verified on other similar communities.

In this research, we did not find a strong relationship between people's subjective perceptions and their real usage. Only 9% of the variance of real usage data has been explained by the model. One reason maybe because different measure scales were applied for intention and real usage. Real usage data could be converted into a interval or ordinal scale and test the model again. Besides intentions there might be other factors that influence real usage, e.g. time constraint, facilitating conditions (Venkatesh et al., 2003). In order to further understand the gap between people's perceptions and their real use, qualitative research methods like interviews would be useful. Semi-structured interviews can be conducted with selected members who answered the survey to find out why there was a big gap between their perception and their real usage. The results from the interview should be able to help us further extend our research model. The newly discovered factors then could be used to develop a measurement scale, and be tested in another study to form a model for real usage.

In this research, we failed to motivate members to contribute to the community place. Other strategies, such as a points-system can be implemented in such research communities to motivate members to contribute to the knowledge base. In our future research, motivation strategies can be applied to the community to measure changes in members' interaction, to further explore interpersonal relationship development in a research online community, and to suggest ways to help researchers collaborate in a virtual place.

APPENDIX A

CONSENT FORM

NEW JERSEY INSTITUTE OF TECHNOLOGY
323 MARTIN LUTHER KING BLVD.
NEWARK, NJ 07102

CONSENT FORM TO PARTICIPATE IN A RESEARCH STUDY

TITLE OF STUDY: Online Community for Asynchronous Learning Networks (ALN)

RESEARCH STUDY: Survey on Participating in the WebCenter for Asynchronous Learning Networks Effectiveness Research

I, _____, have been asked to participate in responding to a survey under the direction of Yi Zhang (Doctoral candidate, Information Systems, New Jersey Institute of Technology)

PURPOSE:

The objective of this research project is to investigate possible software functionalities to increase the use of the website, the cohesion of the ALN community, and the communications among community members.

DURATION:

My participation in this study will last for **5-15 minutes**.

PROCEDURES:

I have been told that, during the course of this study, the following will occur:

Filing out the short questionnaire

PARTICIPANTS:

I will be one of about 100 participants to fill out the questionnaire.

EXCLUSIONS

I will inform the researcher if any of the following apply to me:

No Exclusions

RISK/DISCOMFORTS:

I have been told that the study described above may involve the following risks and/or discomforts:

The investigator believes that there are not any physical or other risks associated with this research. However, there could be risks and discomforts that are not yet known.

CONFIDENTIALITY:

Every effort will be made to maintain the confidentiality of my study records. Officials at NJIT will be allowed to inspect sections of my research records related to this study. If the findings from the study are published, I will not be identified by name. My identity will remain confidential unless disclosure is required by law.

PAYMENTS FOR PARTICIPATION:

I have been told that I will receive no compensation for my participation in this study.

CONSENT AND RELEASE

I fully recognize that there could be little risks that I might be exposed to by volunteering in this study which are inherent in participating in any study. I understand that I am not covered by NJIT's insurance policy for any injury or loss I might sustain in the course of participating in the study.

RIGHT TO REFUSE OR WITHDRAW:

I understand that my participation is voluntary and I may refuse to participate, or may discontinue my participation at any time with no adverse consequence. I also understand that the investigator has the right to withdraw me from the study at any time.

INDIVIDUAL TO CONTACT:

If I have any questions about my treatment or research procedures I may discuss them with the principal investigator. If I have any addition questions about my rights as a research subject, I may contact:
Richard Greene, M.D., Ph.D., Chair, IRB (973) 596-3281

SIGNATURE OF PARTICIPANT

I have read this entire form, or it has been read to me, and I understand it completely. All of my questions regarding this form or this study have been answered to my complete satisfaction. I agree to participate in this research study.

Subject: Name:_____ Signature:_____

Date:_____

SIGNATURE OF READER/TRANSLATOR IF THE PARTICIPANT DOES NOT READ ENGLISH WELL

The person who has signed above,
_____, does not read English well,
I read English well and am fluent in (name of the language)
_____, a language the subject understands well.
I have translated for the subject the entire content of this form. To the best of my knowledge, the participant understands the content of this form and has had an opportunity to ask questions regarding the consent form and the study, and these questions have been answered to the complete satisfaction of the participant (his/her parent/legal guardian).

Reader/
Translator: Name:_____ Signature:_____
Date:_____

SIGNATURE OF INVESTIGATOR OR RESPONSIBLE INDIVIDUAL

To the best of my knowledge, the participant,
_____, has
understood the entire content of the above consent form, and comprehends the study. The participants and those of his/her parent/legal guardian have been accurately answered to his/her/their complete satisfaction.

Investigator's Name:_____ Signature:_____

Date:_____

APPENDIX B

QUESTIONNAIRE

Survey Questionnaire on the WebCenter for ALN Effectiveness Research
(www.ALNResearch.org)

General Information
Name: _____

Title: _____

Gender: M___ F___

I usually connected to the Internet using:
(a) Modem dialup (56k) (b) Cable Modem/DSL (c) T1

I use the Internet about _____ hours a day.
(a)< 1 hour a day (b) Less than 2 hours (c) 2-4 hours (d) 4-6 hours (e) more than 6 hours

Questionnaire
[1] I am satisfied with the quality of the ALNResearch WebCenter output
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[2] Overall, I feel the results of using the WebCenter system are apparent to me.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[3] I might contact someone in the people directory who shares research interests with me.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[4] I trust other members in the ALNResearch WebCenter
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[5] The WebCenter has provided me with guidance, references, or other information useful to my work
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[6] The interface of the ALNResearch WebCenter is clear and understandable.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[7] If I can access to the WebCenter, I will continue to use it
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[8] The WebCenter has broadened my understanding of my research area
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[9] I can get things that are important to my interests from the WebCenter.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[10] The WebCenter helps me know who else is doing research in my research field.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[11] I find the ALNResearch WebCenter is NOT useful at all for my work.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[12] It is difficult for me to explain why using the WebCenter may or may not be beneficial.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[13] I **DO NOT** trust anyone online
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[14] Using the ALNResearch WebCenter improves my work.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[15] I found it easy to complete what I want to do in the ALNResearch website.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[16] I **DO NOT** plan to post anything on the site or communicate with anyone through the WebCenter.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[17] For me, the WebCenter is **JUST** a place where I can get information.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[18] I **ALWAYS** feel isolated in the ALNResearch WebCenter.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[19] I have no difficulty telling others about the results of using the WebCenter system.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[20] The quality of the information I get from the WebCenter is high
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[21] I **DO NOT** plan to use the WebCenter any more
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[22] I feel connected to others in the ALNResearch WebCenter.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[23] It **REQUIRES** significant mental effort to use the ALNResearch WebCenter
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[24] The ALNResearch WebCenter has something useful for my work.
Strongly agree | 1 | 2 | 3 | 4 | 5 | Strongly disagree

[25] I have contacted with people who I met from the WebCenter by email, phone or other media.
Yes ____ No ____

[26] How valuable and useful do you currently find each of the following features or capabilities?

	Frequency of Use			**Value**				
	(1) Never Used	(2) Occasiona -lly	(3) Frequently	(1) Cannot Say	(2) Useless	(3) Slightly Useful	(4) Fairly Useful	(5) Extremely Valuable
Empirical Studies								
Other Studies								
Research Instrument								
Research Resources								
Tutorials								
People Directory								
Discussion Forum								
Contribution Forms								

Open-ended Questions:

1. What are your expectations if you decide to join the ALNResearch WebCenter?

2. What other kind of features would you like to see added to the ALNResearch WebCenter?

APPENDIX C

RESULTS FROM PROTOCOL ANALYSIS[4]

Goal of the study

The usability study is being undertaken for the ALN Research web based system available at www.anlresearch.org. This study aims at analyzing the usability of the site by undertaking a protocol analysis study. The study will focus on the navigation features, the presentation of the website, a few features of the system like search and discussion forum and will eventually present a set of recommendation, which can be used as a basis for improving the usability of the site.

Mechanics of the study

The study uses the protocol analysis method to make its analysis and recommendation. A set of subjects will be presented with a set of tasks to be accomplished and will be instructed to use the "thinking aloud" protocol to express their thoughts, feelings and actions while performing the tasks. The session with the subjects is recorded and later transcribed for performing the analysis. This step will provide the complete view of the user's first hand reaction to using the system. The user is also presented with a set of post-test questionnaire which provides the post-mortem view of the subject. The responses of the subjects along with their session details are considered as the inputs for compiling the results.

Introduction

ALN Research website was developed as a research implementation of a recommender system. This Web Center for Learning Networks Effectiveness Research is intended to become the cyberspace home for a knowledge-building community of researchers in this area and a knowledge base for anyone who wants to learn about it.

The goal of the ALN research program is to increase the quality, quantity, and dissemination of results of research on the effectiveness of Asynchronous Learning Networks (ALN). It synthesizes existing knowledge and creates new knowledge about the methods and findings of research on the determinants of effectiveness of ALN, and to make the results available worldwide via a project web site. The major activities of the system include a series of online knowledge bases that are regularly updated and available through the project web site to researchers, faculty, the press, and the public. The website provides a comprehensive listing of research that is being done in the area of ALN Learning Effectiveness. The system will also build/strengthen an ALN evaluation research community that will create and share improved research methods, theoretical frameworks, and instrumentation for assessing the outcomes of online learning.

The system also makes recommendations to the user for new documents. The system is web based and is available at http://www.alnresearch.org. Users using the system have to create a login and password to use the system. The system is available free of cost. The

[4] The protocol analysis is done as a project in a graduate course: "Evaluation of Information Systems", all investigators are students in the course.

web-based system also has a discussion forum. This discussion forum can be used to exchange thoughts and ideas about the research topics. Users can use this system to interact with other users. Users can also upload and maintain their personal details including their photographs.

Project Overview

In the current project, the usability features of the ALN Research system are being tested using protocol analysis. The target user community of this system is the research community. The site is being tested for its usability in terms of speed, ease of use, ease of searching information, general navigation, appearance, locating work of other people and their information.

The test is subdivided into a set of tasks, which includes the generic tasks that a user may perform when accessing this website. The set of tasks include creating a user id and password for the system, locating papers corresponding to a topic, looking up the discussion threads, locating people at the website, locating the system maintainers and related activities.

Our understanding is that most people will visit the site with similar activities.

Project Schedule

The milestones for the various activities of the project are as follows:

Table C.1 Project Schedule

WEEK	DATES	TASKS
Week 1	Feb 5-11	Form group, select software and become familiar with it. Working as a group, devise your task and instructions. Find subjects to interview.
Week 2	Feb 12-18	Carry out first Protocol Analysis, submit summary report to the group (including recommendations for changes), and decide on any modifications to your task, instructions, or whatever. This must be done by the end of the week. If you have not done the first one by the second week, your group will not be able to finish on time.
Week 3	Feb. 19-25	Complete final protocol analysis. Analyze your individual results and share with group.
Week 4	Feb. 26-Mar. 4	Compare results and prepare overall group comments/recommendations. Submit final group document (one report per group) in class at 6pm on March 4, 2004.

Protocol Instruments

In addition to a tape recorder or a recording device, to record the thinking aloud activity of the subjects, the following instruments was implemented for the protocol analysis:

Consent Form

The consent form is signed by all the participants/subjects in the test. It is an acknowledgement form describing the goal of the test, the investigators involved, and assuring confidentiality to the participants. The personal information of the participants is

kept confidential and is replaced by numbers or subscripted letters to identify them. Only the investigators of the study have access to the consent forms and the true identity of the participants.

General Instructions

A set of guidelines are prepared and explained to the participants prior to the testing of the system. The guidelines include the instructions that each subject was given prior to the start of the protocol analysis exercise. The instruction script used is provided in Appendix D as a part of the introduction script.

Tasks

A set of tasks was devised to test the usability of the system. These tasks are the actual instructions that the subject has to perform when he/she is visiting this web-based system for our protocol analysis. These tasks were created taking into considerations that most users of this site would perform similar activities. The tasks were set up to adequately represent the features that we wanted to test. In addition, it was a way of ensuring that all subjects performed the same tasks in the same order. The initial set of tasks provided to the pilot subject and the final set of tasks which were revised after the pilot interview are provided as a part of Appendix D.

Interview Questions

Post interview questions were developed and attached to the back of the Instructions page. For consistency, all subjects were to answer the exact same questions. These set of questions were devised to get a post session feedback from the subjects. The questions were formed in such a way that it would help the subject to elaborate on his/her experiences while performing the tasks. The questions were administered immediately after the tasks were completed to ensure that the information in the short term memory of the user is not lost. The interviewer could also question any additional question if he/she wanted to clarify any activity of the subject. These sample set of questions and the answers of the subjects are appended in Appendix E.

SUMMARY ANALYSIS REPORT(S)

Test results observed by : investigator1

Test Results for Subject "A1"

Test Date: 2/21/04

Background of Subject:

Subject A! is a female in her late 50's. She works for UPS and has a low to moderate web experience. She uses her computer mostly for e-mail, playing games, and some web surfing.

Summary of Experience:

The subject was originally frustrated trying to register on the web site. Part of the problem was that our Task #1 said to create a user id and password. It did not use the word register. It took several tries before she realized she had to register first.

Positive Aspects of the System:

- Liked the History feature of the site. It remembers your name and the articles you have reviewed and shows you more on that subject.
- Liked the tabs at the top of the page. Made it easy to get around.

Difficulties encountered with the System:

- Difficulty Registering for the site.
- Difficult in finding Search
- Process of getting back to the home page of the site. Subject thought it was cumberson.

User Comments and Suggestions:

- The word **Register** should be in a large font and written on the Top of the box.
- If you click on the underlined title of an article it should open without going to the abstract icon.
- The web site needs more color, it looks a little dull.

Evaluator comment and Suggestions:

The subject had some difficulty navigating the site. The frustration with registering had a somewhat negative impact on her overall experience. Also, the subject felt she would not be likely to have need of this kind of a web site.

Test results observed BY: investigator2

Test Results for Subject "S1"

Test Date: 2/22/04

Background of Subject:

Subject S1 is a male in the age group of 25-35. He works as a software architect in a telecommunication company. His experience with the web in general and research sites in particular can be considered as high.

Summary of Experience:

The subject did not face problems in normal navigation of the system. But the subject had little frustration with the fact that there was no clear indication that he had to log into the site to do any search. He was happy with the results given back by the system and gave very detailed description about each of his actions. He also gave suggestions at screens regarding display of information and presentation of information.

Positive Aspects of the System:

- The subject was impressed with the layout of most screens. He felt that the site was attractive.
- The subject like the discussion forum. The organization was appreciated.
- The subject appreciated the project page detailing the research assistants and faculty and liked the information and links that this page provided.
- The subject liked the details in the retrieved articles and their presentation.
- The subject also liked the option available to choose the abstract or full text of a given article.

Difficulties encountered with the System:

- The subject was not happy with the extra number of windows he had to close to get back to the home page.
- He felt that the home page of the site should somehow convey that a login and password are required to access articles on the site.
- The subject prefers seeing the same home page when he clicks on the home page link rather than different home page presentations for logged in users and users who have not logged in.
- There was inconsistency of whether a URL opens in the sane window or a new window.
- The subject felt that the site was dull.

- The subject was confused with the DB entry on a few pages, since it does not clarify what DB entry means.

- The subject felt that all articles should be consistent in factors such as whether a printer-friendly version is available or not.

- The subject was frustrated when a few of the hyperlink of the article titles would not work and a few did not provide a hyperlink. He feels that the titles of the articles should have a hyperlink to display the article.

User Comments and Suggestions:

- Different and pleasant color combinations should be tried out to increase the appearance of the site.

- Navigation should not be so confusing. Hyperlinks on articles should always work.

- Consistency should be maintained in all the articles like having a printer-friendly version as well as every page should have a link back to the home page.

- The user should be able to sort the list of the searched articles.

- The users should be able to point their mouse over the buttons (for ex, DB Entry) and be able to get some information about the button.

- The people tab on the home page should have a section for the faculty and research assistants as it is not very intuitive for users to click on projects to find the faculty.

- Lectures provided on the site can be made downloadable. As low-speed internet users may not always prefer to read a lecture online.

- Check boxes on the login pages are preferred as compared to text areas for information like preference, how the user came to know about the site and similar features.

- Feedback to the site can be available on the home page along with the email to web master tab. Feedback need not be in the discussion forum.

- Search feature should be available on the main home page.

- The site should have a frames/ no-frames option for people who intend to use the system from a browser which does not support frames.

Evaluator comment and Suggestions:

The subject had no difficulty in accomplishing the tasks. But he expected the site to mention at some point that login was a necessity. The subject may visit the site in future and use its features.

The subject showed some fatigue by the time task 3 and 4 commenced. That is evident form the difference in the length he spoke. The subject spent 7 minutes and 9 minutes for tasks 1 and 2 and spent less than 2 minutes each on task 3 and 4. The subject was advised to take a break after that before he started task 5. It is probably because the number of tasks is too many. The number of tasks can be reduced for additional studies.

The subject provided insightful information regarding the web pages. He needed no prompting once the sessions had started.

The subject felt that the download of the real player file took a lot of time since he was using a dial up connection. Also, he mentioned that there was a fair chance that some users did not have real player. Therefore, the observer feels that it may be a better idea to have a player embedded in the site for lectures or have a link to download the presentation slides, if not the complete lecture.

The subject felt very frustrated because some of the links on the article titles were not opening up the articles. The fact that the font and presentation of the titles looked very close to that of a regular hyperlinked text, it resulted in more confusion. The better option would be to present the titles as hyperlinks to the actual articles. In the event that such a hyperlink is not possible, it is necessary to indicate the same on the page either explicitly (write in a parenthesis after the title) or implicitly (change the font and presentation color).

The subject liked the color of the site in general but found it too dull. At times he preferred the blue color of the site and at other times he commented on it saying it was too dull. It is possible that though the site looks presentable, the users get bored after prolonged usage. Trying out different color combinations and presentation may help in finding an optimal presentation style of the web pages.

Test results observed by : investigator3

Summary of the first protocol analysis done on alnresearch

In pursuit to carry out first protocol analysis, I choose a subject who had very good knowledge of the computers. We started with the basic task of the registering with the website.

I requested a friend of mine who is a PhD student to be my subject. I explained him the purpose & requirement of the study. He gave his consent by signing on the consent form. My subject's domain of computer application knowledge is decent and considered well above average.

We met at our scheduled time at my residence (I had to call him at my place, because his internet connection was not working properly), I read out the introduction and clear description of the series of tasks that I wanted my subject to accomplish. The subject was little unclear, so he requested to read the paper himself. After browsing thro' it, he nodded to go on with the task.

Tasks:

Though we had outlined six tasks to be accomplished during the experiment, I asked my subject to take his time and do it one by one at his pace. I will call my subject just a "Subject" in my discussion ahead.

First Protocol Analysis:

Subject opened the website by putting in the URL in the address toolbar on my notebook. It gave him a window of the homepage with the different links. This got him a bit confused. There was option of existing users and a small link for register. He asked me for the directions. I revived his memory that I cannot actively help, but advised to concentrate on what is written on the instructions page he had. After some time of concentrated reading him again looked at the homepage, and found the link for the "register". This really got him really energized, like he had won a major feat. Anyways, a new page on the screen opened with many fields requiring the name and the password and other details from the subject. He handled this with ease, but complained about the number of fields he had to fill out. Also he was confused with the "ALN specific information", specifically with "Your Roles in ALN" where he had to fill out the "other" field by typing "visitor". Eventually, he was done with filling out his information. He clicked on "Submit request" to get himself registered. And that's when he got really confused he got a screen prompt saying to check his email. He wasn't sure if he had to check his email first before logging into the website, and complained that the message should say that for further information check your email.

The next task was to search for articles/papers about how successful virtual teams are. This task was the most confusing task for the subject. The confusion started with the question what's the difference between article and paper do they refer to the same thing. The subject then entered his Login user name and password in order to use the system. Then the subject located the search field and entered "success of virtual teams" and 12 articles came up. He tried to rate the articles after reading the article's abstract and was complaining about the rating part saying he can't rate the article from just reading the abstract. The subject couldn't find the comments for the articles so he decided to move to next task.

The third task was to "find the discussion forum in the web site and Look at what other people feel about this web site". The first thing the subject did was search for a link that will allow him to do the task; finally he clicked on discussions which gave him a new page with the list of different forums on it. After going through the list he clicked on "Discussion on the site" link, which gave him a new page with a list of topics on it. Again he complained about the low number of people who had commented, and decided to click on "Feedback on the interface for the research articles database" and went through the only message that was there and its reply. He agreed with the message saying that he also couldn't see any comment about the article and decided to go back and click on another topic; "Feedback on the Experimental Design and other Tutorials". After going through some of the messages the subject moved to the next task.

Task4 was to find the names of the faculty and research assistants working on this project. Eventually after deliberating for about 3-4 minutes, he was able to click on the "Project" tab and then "People" and was finally inside. After going through the list of names he moved to the next task.

Task5 needs to go to lecture # 2 in tutorials and view the acceptance slide. Go back to home. This task was supposed to be an easy one, but the subject had some delays locating Lecture #2, because the lecture numbering of the lecture was not specific enough.

All this took about 50 minutes; the subject was a bit bored and tired. After the session, I asked our question from the subject, which I recorded. I felt the need of revising our questionnaire based on my experience. And I will be putting forward my suggestions.

Experience & Observations to improve our experiment procedure:-

It was in a way good that I choose an "above average" computer user in my first PA. I could start identifying certain problem areas in the alnresearch. Alnresearch is not a kind of website which a common computer user (Common computer user means the one, whose primary usage is Email & Net Surfing). So I think my choice of the first subject was appropriate.

Recommendation for the changes:

I think our introduction to the task was all right, but I recommend some changes in our task list. Like I had time issue with my first subject, and we were not able to complete all six tasks. I propose to decrease the number of tasks and to inform our subjects that it might take more than 50 mins to complete the experiment.
Moreover I propose to modify task 2; mentioning both "articles" and "papers" in the task created some confusion to my subject, so I suggest using one term. Moreover I think the rating part of the task was unnecessary; it was time consuming and doesn't add to the usability of the website.
Also another issue was in task5; the issue was that it was hard to find what lecture #2 refers to, I think mentioning the lecture by its name is better is a better approach.

CHANGES MADE WITH REASONS AND DISCUSS

The following changes were made to the task list after discussing with the group.

- The task list was very lengthy. One subject could not complete the task list and another subject showed signs of fatigue. Therefore, the number of tasks was brought down to 5.

- Since the task 1 did not have the register word explicitly, it was confusing. Hence, it was decided to mention the word "register" in the final version of task 1.

- Task 2 looked like very detailed. It was confusing to ask subject to rate articles. One subject actually read a few articles and ranked them from 1 through 5. Therefore, the second part of task 2 was removed.

- Real player was pointed out to be a problem particularly for users using a dial up connection. This could also be a problem if the user did not have a real player installed on their system. Also, the task list had to be cut short. Therefore, it was decided to eliminate task 5 which required the user to go to a particular slide on a particular lecture note.

An annotated list of the final tasks is being presented below.

Final Task List

Task1
CHANGED FROM: You are a new visitor to the alnresearch.org web site. Create your user id and password.

CHANGED TO: You are a new visitor to the alnresearch.org web site. Register, create your user id, and password.

Task 2
CHANGED FROM: you want to find a few articles relating to how successful virtual teams are and papers related to the same. Use the simple search feature to find appropriate articles.
Find a few appropriate articles and rate them. Find an article that has comments posted.

CHANGED TO: You want to find a few articles relating to how successful virtual classrooms are and papers related to the same. Use the simple search feature to find appropriate articles. Find a few appropriate articles and rate them. We removed the last part of this task because most articles do not have comments posted for them.

Task3.
CHANGED FROM: Find the discussion forum in the web site. Look at what other people feel about this web site.

CHANGED TO: Find the discussion forum in the web site. Look at what other people feel about this website.

Task4.
CHANGED FROM: Find the names of the faculty and research assistants working on this project (ALN research)?

CHANGED TO: Find the names of the faculty and research assistants working on this project (ALN research).

Task5:

CHANGED FROM: Go to lecture # 2 in Tutorials and view the acceptance slide. Go back to home.

CHANGED TO: We removed this task because Real One Player must be installed on your computer to accomplish this task.

Task6:
CHANGED FROM: Find *XXXX* and choose the abstract of "A Comparison of Student Outcome & Satisfaction between Traditional & Web Based Offerings" Now view the full text and return to home.

CHANGED TO: Find XXXX and choose the abstract of "A Comparison of Student Outcome & Satisfaction between Traditional & Web Based Offerings" Now view the full text and return to home.

FINAL ANALYSIS AND RECOMMENDATIONS FOR SOFTWARE CHANGES

Analysis and recommendations by investigator1

Test Results for Subject "M2"

Test Date: 2/24/04

Background of Subject:

Subject M2 is a female in her mid 20's. She works in a law office and uses office suite. She also uses proprietary software to complete real estate closing forms. She uses the web for e-mail and surfing. She has graduated from college and is familiar with research.

Summary of Experience:

The subject enjoyed the site and was able to navigate quite easily. She was able to complete all of the tasks without undo difficulty.

Positive Aspects of the System:

- Felt the site had a lot of good information
- Liked the pop-ups that gave additional information
- Subject liked the way her name was displayed on the site.
- Subject liked the pictures of individuals that were provided

Difficulties Encountered with the System:

- Difficulty finding the search feature

User Comment and Suggestions:

- Subject liked the site. She felt it had a lot of good information. She said, "I knew what I was looking for by finding a person, and or subject matter. The abstracts were useful."
- Several of the links associated with the people involved in the project were not activated. Subject felt they should be.
- The search feature should be made more visible. The word search should be in large letter sat the top of the box.

Evaluator comment and Suggestions:

The subject was able to navigate the web site easily. She seemed to enjoy clicking trough all the tabs at the top of the page to learn about the site.

Analysis and recommendations by investigator2

The analysis and recommendation are provided after the test on subject 2.

Test Results for Subject "S2"

Test Date: 2/28/04

Background of Subject:

Subject S2 is a male in the age group of 25-35. He works as a software developer in a company which develops tools for the financial markets. He uses the computer on a daily basis at work and is very familiar with generic and research websites.

Summary of Experience:

The subject had no problems and performed the tasks without any hesitation. He also in fact was inquisitive enough to actually try out a few extra navigations to let me know his comments.

Positive Aspects of the System:

- The subject liked the registration link just over the login area, which made it more presentable.
- The color and presentation of the registration form especially was appreciated.
- Having the search feature on the home page of a user was acknowledged by the subject.
- The subject liked the presentation style of the articles which included the author information, date of publication and other relevant details.
- The subject liked the rating feature of the site.
- The discussion forum was appreciated for its contents. The subject especially liked the icons and the fast reply feature.

Difficulties Encountered with the System:

- The subject did not find a few pages very professional as it included colors like pink, magenta and orange.
- At the first point, when the subject created his login id and password, he expected to be directly logged into the site after his registration was complete. Since that was not the case, he was not happy.
- The Web Learning Center icon in the center of the user's homepage was wasting a lot of real space. The subject found that objectionable.
- The absence of a hyperlink on the searched articles and the non-functioning of certain hyperlinks was another problem that the user felt.

- On the discussion forum, the menu scrolled up if the subject tried to scroll down the forum postings.

- The user expected the list of faculty and research assistants to appear under people rather than projects.

- The user expected multiple filters (based on author, keyword, date) to be simultaneously available in the advanced search feature.

User Comment and Suggestions:

- The subject wanted the ALN Research home page to be displayable on one screen rather than having scroll bars.

- The user should be automatically logged in after creating his registration information when he/ she first visit the web site.

- The rating scale of an article should be moved below the article itself as it becomes more intuitive for the user to first read the article and then rate it.

- The menu on the discussion forum can be separated out on a different frame so that it does not scroll up when the page is scrolled down.

- The discussion forum can be made more user-friendly by adding a search button to enable searching information when the number of discussions in the forum increases.

- The Project tab should list project details along with the list of faculty and research assistants. On the same note, the people tab should have two sub-tabs, one for members and other for project team.

- In the filter of the advanced search feature, the user should be able to filter based on multiple criteria.

- The topics on the discussion forum have to either left aligned or right aligned. Central aligned topics do not give the feel of a discussion title.

Evaluator comment and Suggestions:

The subject had no difficulty in understanding the tasks or directions for performing the tasks. The subject was very vocal and did not have to be prompted during the protocol analysis.

The subject provided suggestions at different point in the protocol analysis.

The subject was not very comfortable when some of the links on the article titles did not open up though they resembled hyperlinks. It is suggested that the titles are presented as

hyperlinks to the actual articles. In the event that such a hyperlink is not possible, it is necessary to indicate the same on the page either explicitly (write in a parenthesis after the title) or implicitly (change the font and presentation color).

The subject liked the color of most of the pages on the website based system. At some points, like the discussion forum, he was not happy to see pink, magenta and orange colors, as he felt that the usage of these colors did not make the site look professional. Trying out different color combinations and presentation may help in finding an optimal presentation style of the web pages.

Analysis

After conducting two protocol analyses and analyzing them individually, the following table presents the analysis of the two subjects in a tabular form.

Table C.2 Analysis Results from Investigator 2

CATEGORY	POSITIVE FEEDBACK	NEGATIVE FEEDBACK
Information	• Article presentation and details was well presented. • Ability to choose abstract or full text of an article was a good feature.	• Rating scale appears above article.
Presentation	• Screen layout is well presented. • Project page with faculty and research assistants was well presented. • Rating system is very helpful.	• Some comments on colors being dull or un-professional. • DB entry was not self-explanatory and no help available for the same. • The system returned to the main page after the registration process was complete.
Navigation	• Registration site is well presented.	• Number of windows to be closed to return to the home page • No link to home page from various pages. • Hyperlinks are not present or do not work.

Table C.2 Analysis Results from Investigator 2 (Continued)

CATEGORY	POSITIVE FEEDBACK	NEGATIVE FEEDBACK
Login	• Registration was good	• Home page should have information that login is required
Discussion Forum	• The discussion forum in general is well appreciated. • Icons and fast reply feature was appreciated.	The topics in the discussion forum were centrally aligned.
Project/People	• The information regarding the project teams was sufficiently detailed.	• Users invariably clicked on people to look for project team.
Search	Has a simple and advanced search feature.	• Search feature not available on the home page. • Advanced search does not have multiple filters.

The users were in general very comfortable using the system. After the initial instructions, none of the subjects needed any prompting during the test. Since both the subjects are involved in some sort of research even in their work, they were very enthusiastic when given this set of tasks.

The only place both the subjects were stuck was in the task to look for the project faculty and research assistants. The subjects were trying to locate this information in the people section rather than the project section. Therefore, one of the suggestions was to include the team members as a part of the people tab.

The subjects were particularly pleased to look up the articles. The presentation of the publication along with the details like author, date of publication, source etc were highly appreciated. The subjects also provided valuable information regarding the general web layout based on their experience at work as well as other web sites.

Recommendation:

Based on the analysis in the previous section, the following recommendations are being proposed. The recommendations are presented in a tabular form based on the feature that the recommendation addresses (See Table C.3).

Table C.3 Recommendations from Investigator 2

CATEGORY	RECOMMENDATION
Navigation	• Pointing the mouse on a button or a tab should display some information about the button. • Drop boxes or selection menus can replace the text areas in the registration form. • Links should be provided on the article title, which opens up the article. • Home page scroll bar can be removed. • Menu can appear on a different frame to make it always visible even if the user scrolls to the bottom of the menu.
Appearance	• Playing around and trying different color combinations to reduce stress, fatigue and dullness and make it more professional. • Frames/No-frames option will increase the spectrum of users
Organization	• People can have a section on members and another section on project personnel. • Lectures or presentation slides can be made downloadable • A feedback link can be provided on the home page. • An online player to avoid delays can replace real player. • Rating of the articles can move below the article being rated.
Registration	• Creating a registration can directly bring to the home page without sending him/her back to the main page.
Search	• Search button can be put on the main home page, which directs the user to the registration form/login if he/she already does not have a login. • Advanced search can be upgraded to incorporate multiple filters rather than having to toggle between filters. • Sort facility may be made available on the result page of the search request to sort the information based on the number of hits, year of publication of other columns.

Analysis and recommendations by investigator3

Second Protocol Analysis Summary

The subject for my second protocol analysis is an PhD student (chemistry), who has a moderate expertise with computer usage. He has to use the common tasks like, word editors, spreadsheet and some chemistry softwares for his research.

I explained him the purpose & requirement of the study. He wanted to know the details about the study, which I promised that I would let her know, once we go ahead with the experiment. I explained that it is required for our research purposes that I provide him with the details about the study, only after we have his consent to participate. I explained this restriction is to avoid any sort of bias in the experiment. I reassured him about the option of opting out of the experiment at any stage he wishes, if he is not satisfied with it. This reassured him. And he gave his consent by signing the consent form.

We again met at our scheduled time. I gave him the introduction and clear description of the series of tasks that I wanted him to accomplish. After browsing through the document and few questions, he nodded to go ahead with the task. I requested him that I will be taking some printshots of the screen for my documentation purposes while during the experiment and if he gets distracted, he let me know that.

Observations & Recommendations

As I carried out the experiment on two different subjects, who were computer users at different levels of usage. Both were computer literate. But my second subject was confined to somehow limited usage for his research. He is accustomed to same set of computers routines everyday. Whereas my first subject, was a well trained user in his previous profession and now uses the computer on daily basis. He is extensive user of the Information Systems in his school, work place and uses the computer for many of his personal matters. This almost little diversity in my subjects posed some interesting commonality. But as a final observation, I opine "alnresearch.org" is a system, which is suited for special purpose of research for special users.

Both of the subjects were intrigued by the functionality provided by the system, after we talked about the system & experiment after our sessions. I am basing my observations & recommendations more on my first session, because I feel my first subject, owning to her computer experience, gave more insights than the second one.

Observation & Recommendation I

No presentation of help

Though, of my own experience with alnresearh.gov, I know that alnresearh.gov supposedly targets special users the system doesn't provide any help what so ever. I think, tool like help or a manual should be available, and added to the links on top of the page.

Observation & Recommendation II

No Descriptions

The links have little description for them for a new user. I feel, there should be description of functionality for each link. Both the subjects had problems choosing the right link on top of the page. So descriptions would increase ease of usage.

Observation & Recommendation III

Provision of better display

I think the system needs to have a better presentation. Links will serve better if they were put on the left of the page always visible and in a threaded from, for example you have "People" as a link and as a sub-link somewhere else.

FINAL GROUP CONCLUSIONS AND RECOMMENDATIONS

The above individual results were analyzed and discussed to arrive at the final conclusions of the project. The final group conclusions have been tabulated as positive and negative feedback from the users. These analyses are classified based on the particular functionality it affects. The focus of the analysis has been the presentation of the web site, the contents in it, the ease of using functionalities like search features and discussion forum. Also, inherent properties like the user's feel of the website have been included.

The best feature probably was that the website has lot of information assimilated in it. It involved a lot of patience on the part of the subjects to complete this study. Frequent users of websites tended to look for easier options like search features. Some subject went beyond the given task to show how the same task can be completed in different manners.

The most frequent comment was regarding the presentation and the navigation. The presentation included the web page layout, information presentation, color combination and use of scroll bars and menus. The navigation included links to other information, ease of locating information, able to move back and forth between web pages and additional features of the web site.

Surprisingly none of the subjects wanted to know what ALN stood for. A few subjects did not try to go beyond the specific task allocated whereas, a few others enthusiastically spoke about each page and links on the page.
The recommendations are presented are presented after the analysis section.

Table C.4 Summary of Analysis Results

CATEGORY	POSITIVE FEEDBACK	NEGATIVE FEEDBACK
Information	The site has lots of good contents	Help is missing.
Presentation	The presentation of information, especially articles is very good.	
	Some pages and layout are very attractive to the user.	Sometimes, the color appears too dull and sometimes there are multiple colors on the page that does not make it professional
	Maintaining the history of a user makes it very helpful.	

Table C.4 Summary of Analysis Results (Continued)

CATEGORY	POSITIVE FEEDBACK	NEGATIVE FEEDBACK
Navigation		Article titles do not have hyperlink, which created confusion.
	Menu at the top of the user's home page is very good	Menu scrolls up when the user scrolls down the home page.
	All information and menu items are available on the home page and make it easy to go to any link.	Accessing the home page is not easy from links. Some links do not have a link back to the home page.
		Some tab entries are confusing (ex DB Entry), as they do not make implicit sense to the user. Same is true for some details on the registration page (How did you hear about the site)
Login		After Registering into the site, the user is not directly logged into the site.
Discussion Forum	The organization is good, the icons look good	
	Fast reply feature is highly appreciated.	Menu scrolls up when the user scrolls down the home page.
Project/People	Presentation of people and project team is good	Users tend to look into people tab when looking for project team details
	Good representation using Internet Explorer	Not sure how it is presented on other browsers.
Search	Search feature is available on the user's home page after he/ she has logged in. Advanced search is available for users who want to filter their searches.	Search does not appear on the regular home page. Advanced search feature does not enable multiple filters. The user has to toggle between multiple criteria.

Recommendations

The summary of the recommendation after consulting and deliberating with all the evaluators was developed. The recommendations of the system are classified and presented under sub categories. This helps easy comprehension by the maintainers of the website to prioritize and plan their next steps of actions.

Table C.5 Summary of Recommendations

CATEGORY	RECOMMENDATION
Navigation	Pointing the mouse on a button or a tab should display some information about the button. Drop boxes or selection menus can replace the text areas in the registration form.
	Links should be provided on the article title, which opens up the article. A thread on the side of the screen will also make the site more navigable.
	Menu can appear on a different frame to make it always visible even if the user scrolls to the bottom of the menu.
Appearance	**Playing around and trying different color combinations will yield an optimal color scheme, which is professional, and reduces stress, fatigue and dullness in the users since most users of this system are expected to use the system for long durations.**
	Frames/No-frames option will increase the spectrum of users
Organization	People can have a section on members and another section on project personnel.
Registration	Creating a registration should bring the user directly bring to the home page

The results of this system are being presented to the maintainer of this website for records and enhancements of the web site.

Conclusion

This test was undertaken as a requirement for the protocol analysis for as a part of a course requirement. But during the experiment and analysis and discussions the usefulness of this method became very apparent. There are potential benefits of using this approach especially in determining the usability features of systems.

Being in the academic fields, some aspects of a system become very obvious to us. The developers of a system are at a disadvantage because of their expertise to discover potential problem areas. Therefore, it becomes necessary to seek the help of subjects in a test like the protocol analysis to determine the actual usability of a system. In our study, many subjects put forth interesting points, which may have escaped our attention in the

regular development process. The beauty of protocol analysis is that it combines the basic attribute of humans to be able to talk aloud into the actions they are doing and eventually contribute towards the technical enhancement of a system.

The evaluators of the ALN Research system enjoyed the experience of undertaking a protocol analysis experiment in conducting the experiment, evaluating and generating feedback.

In the current setup of testing the ALN Research website, the objective was to study the usability of the system. Additional expert studies need to be conducted to test advanced features like the recommendation of the system and the rating significances.

APPENDIX D

PROTOCOL ANALYSIS INTRODUCTION SCRIPT

Task directions and documentation

Thank you for volunteering to assist in evaluating ALNResearch website. Your feedback will be used to make specific recommendations to the Webmaster for improvements to the site. We will be evaluating the ease of use for the prospective user of this Web site.

I would like to encourage you to talk constantly throughout the usability test.

I am going to give you a series of things to look for on the web site. Your task is to tell me what you are thinking as you look and navigate through each screen of the system. Use your mouse to navigate this site.

- For each screen **please tell me** all of the actions that you think you would be able to accomplish based on what you see.
- This includes clicking on your selections or pressing keys on the keyboard.
- Read the selections out loud and talk about what you think would happen if the action is taken.
- You should speak up about anything that you are certain of or is confusing. I will not be able to help you unless you are completely stuck. The goal of the study is to see how easy the site is to use.
- If you do not like something about the site, please speak up and describe what you do not like and why.
- If you see something you especially like, please speak up and describe what you like and why.
- While you are talking, I will be recording your response. This information will be used to analyze the web site. The goal is to improve the web site so it will be "user friendly."
- Although, I will not be able to answer any of your questions during this exercise, I may point to an item on the screen that I would like you to talk about further.
- Please tell me before you go to the next screen so I can record where you are or prompt you for more information if needed.
- If you forget to "talk out loud" I will prompt you with some open ended questions such as Tell me what you are thinking about as you work. Why do you suppose it did that? What do you think happened there?

- We are evaluating the web site, not your abilities. Be sure to tell us everything you see on the screen and what you are thinking.

- At the end of the usability test I will ask you some questions about your impressions and feelings about the web site.

I will now demonstrate the "Talking Out Loud" technique to you while I am logging onto the Web Site. www.alnresearch.org

Task direction and documentation

Initial tasklist

- Task 1 - you are a new visitor to the alnresearch.org web site. Create your user id and password.
- Task 2 - you want to find a few articles relating to how successful virtual teams are and papers related to the same. Use the simple search feature to find appropriate articles. Find a few appropriate articles and rate them. Find an article that has comments posted.
- Task3. Find the discussion forum in the web site. Look at what other people feel about this web site.
- Task4. Find the names of the faculty and research assistants working on this project(ALN research)?
- Task5: Go to lecture # 2 in Tutorials and view the acceptance slide. Go back to home.
- Task6: Find XXXX and choose the abstract of "A Comparison of Student Outcome & Satisfaction between Traditional & Web Based Offerings" Now view the full text and return to home.

Final Task List

- Task 1 - You are a new visitor to the alnresearch.org web site. Register, create your user id, and password.
- Task 2 - You want to find a few articles relating to how successful virtual classrooms are and papers related to the same. Use the simple search feature to find appropriate articles. Find a few appropriate articles and rate them. Task 3 - Find the discussion forum in the web site. Look at what other people feel about this website.
- Task 4 - Find the names of the faculty and research assistants working on this project (ALN research).
- Task 5 - Find XXXX and choose the abstract of "A Comparison of Student Outcome & Satisfaction between Traditional & Web Based Offerings" Now view the full text and return to home.D. Post-task questionnaire and frequency table of results

APPENDIX E

INTERVIEW RESULTS FROM PROTOCOL ANALYSIS

Investigator: investigator2

Subject: 1

Question 1: What did you think about the ALNResearch Web Site?
Subject: First impression is it has lot of information but what I see is that the information is not presented the way it is supposed to be. That can be improved. The user interface which is like the primary key for any person to come to a particular web site and use it fully that needs some attention. Right now, if I look at the web site, in the simplest terms right from the colors used to the way the text is presented and the consistency of the information and appearance.
Navigation is not consistent. Sometimes it opens in the same page, sometimes it opens in a different page. Sometimes the printer friendly version appears, sometimes, it does not appear. Also, instead of simply providing some tabs right, I think it should be much more specific because when u have this web site, we all know it provides information in a particular arena, so it is not very generic. So, when you are providing that kind of information, I ideally I would have expected very specific tabs or check boxes which will tell what is the information which is available and I should be able to search, do something very specific.

Question 2: What did you like best about the site?
Subject: It is very soothing whatever colors were used, though it was only blue that was used across the website, I think some pages looked very good. The best part was the way the articles were presented. I think that is highly appreciated.

Question 3: What did you like least about the site?
Subject: The least is people have to it is not easy to navigate thru the sire, that is one thing and it is very confusing on occasions where I expect a hyperlink, it is no there, I have a hyperlink, I click on it, it does not work. Those kinds of surprises I did not like. And above all, the way they mentioned about that site woks best with Internet explorer or Netscape, they should also have provided some basic information that the user cannot do anything unless and until they click on the register or they log in. I think that kind of basic information is missing.

Question 4: Would you want to come back to this web site again? Why or why not?
Subject: yes definitely yes, I think it has abundant information and it is a very good website to come back and research for articles. And I am sure they are picking up the best articles in the industry.

Question 5: Do you have any suggestions for improving the web site's usability?
Subject: Yes, I have quite a bit of suggestions. One of the first suggestions I would make is as soon as you come into the web site that should be much more presentable.

The centre of the website, which has a big circle has a list of points, ideally I would have expected the list to be very very specific and very clear and provide hyperlinks. The web site should be created in such a way that anywhere in the web site I will be able to click and navigate and go wherever I want. It should not comeback always to the first page or the previous page, that's not what I expect form the web site. Also one more thing I expect from the web site the information that I am looking at, suppose I want to search for articles, I want to search based on people, I think that the search feature should be available right away. Right now what happens is, if I come to the home page, I go to some other page, I come back to the home page, now my home page is changed and that is where I had the search. I think search should be another tab on the web site.

Subject: 2

Question 1: What did you think about the ALNResearch Web Site?
Subject: It provided me the information so I am definitely going to use it but it didn't look professional to me because of the reasons that it is multi color and navigation is little bit haphazard.

Question 2: What did you like best about the site?
Subject: I like best about this site, ah ... the login was good, it was available on the home page. I didn't have to go to separate login page and search was available on the home page as well which is good. I guess that's it.

Question 3: What did you like least about the site?
Subject: Least items are that the navigation menu disappears when I scroll down and the faculty should be listed under people rather than under projects

Question 4: Would you want to come back to this web site again? Why or why not?
Subject: Yes, since it has provided me the content that I need it is definitely useful to me

Question 5: Do you have any suggestions for improving the web site's usability?
Subject: Reduce the number of colors used stick to more professional colors like blue and grays rather that having light pink, magenta and orange all kinds together it looks good.

Investigator: investigator3

Subject: 1

Usability Test for ALN Research

Interview Questions

1. **What did you think about the ALNResearch Web Site?**

If I were to rate the site out total score of 10, I would give it 6. I liked the simple design of the site, but believe that since it is a research site, there should be a brief introduction about the mission of the site listed on the main page. Most of links on the main page show you the research articles, which could little overwhelming for somebody with not much knowledge about this research area. A few lines of summary attached to every article would be helpful to get better picture.

2. **What did you like best about the site?**

The simple design of site is very appealing and so are the colors. The site is quite structured and easy to use.

3. **What did you like least about the site?**

It looks like that site is designed for a particular community. And you have to register yourself to access most of areas on the site. The registration process is not too tough, but is little confusing. Maybe when one clicks on any link, instead of asking for login, if there is small summary on the content of that link, that would be helpful. The page on tutorials is very ill-structured and difficult to use. The natural meaning of tutorial in websites is thaw there would be step wise guidance, its not the case in this site. Also to view a part of tutorial, one need to have real video, instructions for such should be given on the site.

4. **Would you want to come back to this web site again?Why or why not?**

Maybe not. Because the site is built for a particular research community.

5. **Do you have any suggestions for improving the web site's usability?**

Make it more user friendly by making some content available without registering. Search is also not user friendly; there could mini explanation about search on the webpage itself.

Subject: 2

1. **What did you think about the ALNResearch Web Site?**

I think the site is good for its targeted audience, but a bit hard to follow for a new user.

2. **What did you like best about the site?**

I like the idea of the website, sharing, consolidating, and spreading information about effectiveness of Asynchronous Learning Network.

3. **What did you like least about the site?**

I didn't like the registration procedure, and also I didn't like how members were organized on the people's page.

4. **Would you want to come back to this web site again? Why or why not?**

I don't think so, because it doesn't serve my needs or interests. Don't get me wrong I'm interested in a different field, and it's in a very specific subject.

5. **Do you have any suggestions for improving the web site's usability?**

There should be more work done on how the web site is organized, and the flow or transition from one page to the other.

REFERENCES

Ackerman, M., & Muramatsu, J. (1998). Computing, Social Activity, and Entertainment: A Field Study of a Game MUD. *Computer Supported Cooperative Work, 7*(1-2), pp. 87-122.

Adams, D. A., Nelson, R. R., & Todd, P. A. (1992). Perceived Usefulness, Ease of Use, and Usage of Information Technology: A Replication. *MIS Quarterly, 16*(2), pp. 227-250.

Aggarwal, C. C., Wolf, J. L., Wu, K.-L., & Yu, P. S. (1999). Horting Hatches An egg: A New Graph-theoretic Approach to Collaborative Filtering. *Paper presented at the the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, pp. 201-212.

Agarwal, R. J., & Prasad, J. (1997). The Role of Innovation Characteristics and Perceived Voluntariness in the Acceptance of Information Technologies, *Decision Science, 28*, pp. 557-582.

Agarwal, R., & Prasad, J. (1998). A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. *Information Systems Research, 9*(2), pp. 204-215.

Agrawal, R., & Wimmers, E. L. (2000). A Framework for Expressing and Combining Preferences. *Paper presented at the ACM SIGMOD 2000*, Dallas, TX, pp. 297-306.

Ahmad, M., & Ahmad, W. (1999). Collecting User Accessing Patterns for Building User Profiles and Collaborative Filtering. *Paper presented at the International conference on Intelligent User Interface Communications of ACM*, pp. 57-64.

Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman & Hall.

Anderson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin, 103*, pp. 411-423.

Andrews, D. C. (2002, April). Audience-Specific Online Community Design. *Communication of the ACM, 45,* pp. 64-68.

Ansari, A., Essegaier, S., & Kohli, R. (2000). Internet Recommendation systems. *Journal of Marketing Research, 37*, pp. 363-375.

Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Priceton Hall.

Aubert, B. A., Rivard, S., & Party, M. (1994). Development of Measures to Assess Dimension of IS Operation Transactions. *Paper presented at the Fifteenth International Conference on Information Systems*, Vancouver, British Columbia, pp. 13-26.

Bair, J. H., & Mancuso, L. (1985). *The office systems cycle: The process and technology of office automation.* Palo Alto,CA: Hewlett-Packard Company.

Balabanovic, M., & Shoham, Y. (1997). Fab: content-based, collaborative recommendation. *Communication of the ACM, 40*(3), pp. 66-72.

Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Reading, MA: Addison-Wesley.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall.

Barclay, D., Higgins, C., & Thompson, R. (1995). The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Adoption and Use as an Illustration. *Technology Studies, 2*(2), pp. 285-324.

Baym, N. (1995). The Emergence of Community in Computer-Mediated Communication. In S. Jones (Ed.), *Cybersociety: Computer-mediated communication and community* (pp. 138-165). Thousand Oaks, CA: Sage.

Beach, L. R., & Mitchell, T. R. (1987). Image Theory: Principles, Goals, and Plans in Decision-Making. *ACTA Psychologica, 66*(3), pp. 201-220.

Beach, L. R., & Mitchell, T. R. (Eds.). (1990). *Image Theory: A Behavioral Theory of Decision Making in Organizations* (Vol. 12). Greenwich, CT: JAI Press.

Beach, L. R., & Mitchell, T. R. (1998). The Basics of Image Theory. In L. R. Beach (Ed.), *Image Theory: Theoretical and Empirical Foundations* (pp. 3-18). Hillsdale, NJ: Lawrence Erlbaum.

Beals, D. E. (1991). Computer-Mediated Communication Among Beginning Teachers. *T.H.E. Journal*, pp. 74-77.

Belkin, N. J., & Croft, B. W. (1992). Information Filtering and Information Retrieval: Two slides of the Same Coin? *Communications of ACM, 35*(12), pp. 29-38.

Berger, C. R., & Calabrese, R. J. (1975). Some Explorations in Initial Interaction and Beyond: Toward A Developmental Theory of Interpersonal Communication. *Human Communication Research, 2*, pp. 99-112.

Berger, C. R., & Douglas, W. (1981). Studies in Interpersonal Epistemology: III. Anticipated Interaction, Self-monitoring, and Observational Context Selection. *Communication Monographs, 48*, pp. 183-196.

Berghel, H. (1997). Cyberspace 2000: Dealing with Information Overload. *Communication of the ACM, 40,* pp. 19-24.

Bly, S., Steve, R., & Irwin, S. (1993). Media space: Bringing People Together in a Video, Audio, and Computing Environment. *Communications of ACM, 36*(1), pp. 28-47.

Bradner, E., Kellogg, W., & Erickson, T. (1999). The adoption and use of BABBLE: A field study of chat in the workplace. *Paper presented at the ECSCW 1999*, Lyngby, Denmark, pp. 139-157.

Brazelton, J., & Gorry, G. A. (2003). Creating a knowledge-sharing community: if you build it, will they come? *Communication of the ACM, 46,* pp. 23-25.

Breese, J. S., Herckerman, D., & Kadie, C. (1998). Empiricle Analysis of Predictive Algorithms for Collaborative Filtering. *Paper presented at the the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Madison, WI.

Borner, K., & Lin, Y. (2001). Visualizing Chat Log Data Collected in 3-D Virtual Worlds. *Paper presented at the Fifth International Conference on Information Visualization*, London, England, pp. 141-146.

Burgoon, J. K., & Hale, J. I. (1984). The Fundamental Topoi of Relational Communication. *Communication Monographs, 51*, 193-214.

Burgoon, J. K., & Hale, J. L. (1987). Validation and Measurement of the Fundamental Themes of Relational Communication. *Communication Monographs, 54*, 19-41.

Burgoon, J. K., & Saine, T. (1978). *The unspoken dialogue: An introduction to nonverbal communication*. Boston: Houghton Mifflin.

Bultler, B. S. (2001). Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures. *Information Systems Research, 12*(4), pp. 346-362.

Canny, J. (2002). Collaborative Filtering with Privacy via Factor Analysis. *Paper presented at the 25th annual International ACM SIGIR conference on Research and Development in Information Retrieval*, Tampere, Finland, 238-245.

Carey, J. (1980). Paralanguage in Computer Mediated Communication. *Paper presented at the 18th annual meeting of the association for computational linguistics and parasession on topics in interactive discourse*, Philadelphia, PA, 67-69.

Carlson, J. R., & Zmud, R. W. (1999). Channel Expansion Theory and the Experiential Nature of Media Richness Perceptions. *Academy of Management Journal, 42*(2), pp. 153-170.

Chalup, S. R., Hogan, C., Kulosa, G., McDonold, B., & Stansell, B. (1998). Drinking from the fire (walls) hose: Another approach to very large mailing lists. *Paper presented at the 12th Systems Administration Conference (LISA '98)*, Boston, MA, pp. 317-325.

Chandler, D. (1998). *Personal Home Pages and the Construction of Identities on the Web*. Chandler, Daniel. Available: http://www.aber.ac.uk/media/Documents/short/webident.html [2002, Oct 9].

Chatelin, Y. M., Vinzi, V. E., & Tenenhaus, M. (2002). State-of-art on PLS Path Modeling through the available software. *Research Papers in Economics (RePEc)*.

Cherry, S. M. (2002, May). Talk is cheap, text is cheaper. *IEEE Spectrum, 39,* pp. 55.

Chidambaram, L. (1996). Relational Development in Computer-Supported Groups. *MIS Quarterly, 20*(2), pp. 143-165.

Chin, W. W., & Gopal, A. (1993). An Examination of the Relative Importance of Four Belief Constructs On the GSS Adoption Decision: A Comparison of Four Methods. *Paper presented at the 26th Hawaii International Conference of System Sciences*, Hawaii, pp. 548-557.

Chin, W. W., & Todd, P. A. (1995). On the Use, Usefulness, and Ease of Use of Structural Equation Modeling in MIS Research. *Paper presented at the 17th International Conference of Information Systems*, Cleveland, OH.

Churchill, E. F., & Bly, S. (1999). It's All in the Words: Supporting Work Activities with Lightweight Tools. *Paper presented at the GROUP '99, Phoenix*, AZ, pp. 40-49.

Ciskowski, D., & Benedikt, C. (1995). *MUDs: Exploring Virtual Worlds on the Internet.*: Macmillan Computer Publishing.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In J. M. L. L. B. Resnick, and S.D. Teasley (Ed.), *Perspectives on Socially Shared Cognition* (pp. 127-149). Washington DC: American Psychological Association.

Clark, H. H. (1992). *Arenas of Language Use*. Chicago: University of Chicago Press.

Cohen, M., Remiso de Macedo, D., Augustin Jaques, P., & da Costa Mora, M. (2000). Collaborative Learning Tools as Part of An Open Architecture. *Paper presented at the 30th Annual Frontiers in Education Conference*, pp. 6-10.

Cole, J. R., & Cole, S. (1973). *Social Stratification in Science*. Chicago: University of Chicago Press.

Compeau, D. R., & Higgins, C. A. (1995). Computer Self-Efficacy: Development of Measure and Initial Test. *MIS Quarterly, 19*(2), pp. 118-143.

Compeau, D. R., Higgins, C. A., & Huff, S. (1999). Social Cognitive Theory and Individual Reactions to Computing Technology: A Longitudinal Study. *MIS Quarterly, 23*(2), pp. 145-158.

Conhaim, W. W. (2002). Blogging - What is it? *Medford, 19*(3), pp. 3-5.

Connolly, T., Jessup, L. M., & Valacich, J. S. (1990). Effects of Anonymity and Evaluative Tone on Idea Generation in Computer-Mediated Groups. *Management Science, 36*, pp. 97-120.

Constant, D., Sproull, L., & Kiesler, S. (1997). The Kindness of Strangers: On the Usefulness of Electronic Weak Ties for Technical Advice. In S. Kiesler (Ed.), *Culture of the Internet* (pp. 303-322). New Jersey: Lawrence Erlbaum Associates.

Cotterman, W., & Senn, J. (1992). *Challenges and Strategies for Research in Information Systems Development*. Chichester, England: John Wiley & Sons.

Cowan, D., Mayfield, C., Tompa, F., & Gasparini, W. (1998, April). New Role for Community Networks. *Communications of the ACM, 41,* pp. 61-63.

Culnan, M. J., & Bair, J. H. (1983). Human Communication Needs and Organizational Productivity: The Potential Impact of Office Automation. *Journal of the American Society for Information Science, 34*(3), 215-221.

Culnan, M. J., & Markus, M. L. (1987). Information Technologies. In F.M. Jablin, L. L. Putnam, K. H. Roberts & L. W. Porter (Ed.), *An interdisciplinary perspective* (pp. 420-443). Greenwich, CT: JAI Press.

Curtis, P. (1992). Mudding: Social Phenomena in Text-based Virtual Realities. *Paper presented at the 1992 conference on Directions and Implications of Advanced Computing*, Berkeley, CA.

Dai, K., Zhang, S., & Fu, Q. (2001). Real-time 3D Model Sharing Conference System Based on Agent and VRML. *Paper presented at the 2001 IEEE International Conference on Systems, Man, and Cybermetics*, pp. 3355-3360.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), pp. 319-339.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science, 35*(8), pp. 982-1002.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers In The Workspace. *Journal of Applied Social Psychology, 22*, pp. 1111-1132.

Davis, F. D. (1993). User Acceptance of Information Technology: System Characteristics, User Perceptions and Behavioral Impacts. *International Journal of Man-Machine Studies, 38*(3), pp. 475-487.

Davis, F. D., & Venkatesh, V. (1996). A Critical Assessment of Potential Measurement Biases In the Technology Acceptance Model: Three Experiments of Behavior. *International Journal of Human Computer Studies, 45*(1), pp. 19-45.

Delgado, J., & Ishi, N. (1999). Memory-Based Weighted-Majority Prediction for Recommender Systems. *Paper presented at the ACM SIGIR'99 Workshop on Recommender Systems*, Berkeley, CA.

Dennis, A. R., & Wheeler, B. C. (1997). Groupware and the Internet: charting a new world. *Paper presented at the 13th Hawaii International Conference on System Science*, Hawaii, pp. 287-296.

DeSanctis, G., & Poole, M. S. (1994). Capturing the Complexity in Advanced Technology Use: Adaptive Structuration theory. *Organization Science, 5*(2), pp. 121-145.

DeSanctis, G., Poole, M. S., Dickson, M. S., & Jackson, M. (1993). Interpretive Analysis of Team Use of Group Technologies. *Journal of Organizational Computing, 3*(1), pp. 1-30.

Dieberger, A. (1997). Supporting Social Navigation on the World-Wide Web. *International Journal of Human Computer Studies, special issue on innovative applications of the Web, 46*, pp. 805-825.

Donath, J. S. (1996). *Inhabiting the virtual city: The design of social environments for electronic communities.* Unpublished Ph.D., Massachusetts Institute of Technology, Boston.

Donath, J., Karahalios, K., & Veigas, F. (1999, January). Visualizing conversation. *Paper presented at the 32th Annual Hawaii International Conference on Systems Science*, Hawaii

Dourish, P., & Bly, S. (1992). Portholes: Supporting Awareness in a Distributed Work Group. *Paper presented at the the Conference of the Conference on Computer Human Interaction*, Vancouver, pp. 206-207.

Dubrovsky, V. J., Kiesler, S., & Sethna, B. N. (1991). The Equalization Phenomenon: Status Effects in Computer-Mediated and Face-to-Face Decision-Making Groups. *Human Computer Interaction, 6*, pp. 119-146.

Edwards, P., Bayer, D., Green, C. L., & Payne, T. R. (1996). Experience with Learning Agents which Manage Internet-based Information. *Paper presented at the AAAI Stanford Spring Symposium on Machine Learning in Information Access*, Palo Alto, CA.

Egdorf, K., & Rahoi, R. L. (1994). Find a Place Where "we all want to hear it": Email as a Source of Social Support. *Paper presented at the Annual meeting of the Speech Communication Association*, New Orlearns, LA.

Eisner, W. (1990). *Comics and Sequential Art*. Tamarac, FL: Poorhouse Press.

Erickson, T., Smith, D. N., Kellogg, W. A., Laff, M., Richards, J. T., & Bradner, E. (1999). Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of 'Babble'. *Paper presented at the CHI'99 Conference on Human Factors in Computing Systems*, pp. 72-79.

Fishbein, M., & Ajzen, I. (1975). *Beliefe, Attitude, Intention and Behavior: An Introfuction to Theory and Research*. Reading, MA: Addison-Wesley.

Feldman, M. S. (1987). Electronic Mail and Weak Ties in Organizations. *Office Technology and People, 3*, pp. 83-101.

Fornell, C., & Larcker, D. (1981). Structural Equation Models with Unobserved Variables and Measurement Error. *Journal of Marketing Research, 18*, pp. 39-50.

Fornell, C., & Bookstein, F. L. (1982). Two Structural Equation Models: LISERL and PLS Applied to Consumer Exit-Voice Theory. *Journal of Marketing Research, 19*, pp. 440-452.

Fornell, C., Lorange, P., & Roos, J. (1990). The Cooperative Venture Formation Process: A Latent Variable Structural Modeling Approach. *Management Science, 36*, pp. 1246-1255.

Fulk, J., Schmitz, J. A., & Schawarz, D. (1992). The Dynamics of Context-behavior Interactions in Computer-Mediated Communication. In M. Lea (Ed.), *Contexts of Computer-Mediated Communication* (pp. 7-29). London: Harvester Wheatsheaf.

Fulk, J, Schmitz, J. & Ryu, D. (1995) Cognitive Elements in the Social Construction of Technology. *Management Communication Quarterly*, 8, pp. 259-288.

Garvey, W. D. (1979). *Communication, the Essence of Science: Facilitating Information Exchange Among Librarians, Scientists, Engineers, and Students*. Oxford: Pergamon Press.

Girgensohn, A., & Lee, A. (2002). Making Web Sites Be Places for Social Interaction. *Paper presented at the Computer Supported Collaborative Work'02*, New Orleans, Louisiana.

Glance, N., Arregui, D., & Dardenne. (1998). Knowledge Pump: Supporting the Flow and Use of Knowledge in Networked Organizations. In R. P. U. Borghoff (Ed.), *Information Technology for Knowledge Management*. Berlin: Springer Verlag.

Goldberg, D., Nichols, D., Oki, B., & Terry, D. (1992). Using Collaborative Filtering System to Weave an Information Tapestry. *Communications of the ACM, 35,* pp.61-70.

Goldberg, K., Roeder, T., Gupta, D., & Perkins, C. (2000). *Eigentaste: A Constant Time Collaborative Filtering Algorithm* (ERL Technical Report M00/41): UCS.

Good, N., Schafer, J. B., Konstan, J., Borchers, A., Sarwar, B., Herlocker, J., & Riedl, J. (1999). *Combining Collaborative Filtering with Personal Agents for Better Recommendations.* Paper presented at the the 1999 Conference of the American Association of Artifical Intelligence (AAAI-99), 439-446.

Goodhue, D. L. (1995). Understanding the Linkage Between User Evaluations Of the Systems and the Underlying Systems. *Management Science, 41*, pp. 1827-1844.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*, 1360-1380.

Graveline, A., Geisler, C., & Danchak, M. (2000). Teaming Together Apart: Emergent Patterns of Media Use in Collaboration at A Distance. *Paper presented at the IPCC/SIGDOC 2000*, pp. 381-393.

Grier, D. A., & Campbell, M. (2000). A Social History of Bitnet and Listserv, 1985-1991. *IEEE Annals of the History of Computing, 22*(2), pp.32-41.

Grimes, S. M. (2003). All About the Blog. *ACM SIFCAS Computers and Society, 31*(5), pp. 1.

Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004). Mining New Media: Information Diffusion through Blogspace. *Paper presented at the 13th Conference on World Wide Web, New York*, pp. 491-501.

Hackman, J. R., & Oldham, G. R. (1976). Motivation Through The Design of Work: Test of A Theory. *Organizational Behavior and Human Performance, 16*, pp. 250-279.

Hagel, J., & Armstrong, A. (1997). *Net Gain: Expanding Markets through Virtual Communities*. Boston, MA: Harvard Business School Press.

Handel, M., & Herbsleb, J. D. (2002). What is Chat Doing in the Workplace? *Paper presented at the 2002 ACM conference on Computer Supported Cooperative Work*, New Orleans, Louisiana, pp.1-10.

Hartwick, J., & Barki, H. (1994). Explaining The Role of User Participation in Information System Use. *Management Science, 40*, pp. 440-465.

Haveman, H. A. (1993). Organizational Size and Change: Diversification in the Savings and Loan Industry After Deregulation. *Administrative Science Quarterly, 38*(1), pp. 20-47.

Herbsleb, J. D., Atkins, D. L., Boyer, D. G., Handel, M., & Finholt, T. A. (2002). *Introducing instant messaging and chat in the workplace.* Paper presented at the Human factors in computing systems, Minneapolis, MN, pp. 171-178.

Herlocker, J., Konstan, J., Borchers, A., & Riedl, J. (1999). An Algorithmic Framework for Performing Collaborative Filtering. *Paper presented at the 1999 Conference on Research and Development in Information Retrieval.*

Herlocker, J., Konstan, J., and Riedl, J. (2000). Explaining Collaborative Filtering Recommendations. *Paper presented at the ACM 2000 Conference on Computer Supported Cooperative Work*, pp.241-250.

Hesse, B. W., Werner, C. M., & Altman, I. (1988). Temporal Aspects of Computer-Mediated Communication. *Computers in Human Behavior, 4*, pp. 147-165.

Hiemstra, G. (1982). Teleconferencing, Concern for Face, and Organizational Culture. In M. Burgoon (Ed.), *Communication yearbook 6* (pp. 874-904). Beverly Hills, CA: Sage.

Hill, W., Stead, L., Rosenstein, M., & Furnas, G. (1995). Recommending and Evaluating Choices in a Virtual Community of Use. *Paper presented at the 95 Conference on Human Factors in Computer Systems*, Denver, Colorado, pp. 194-201.

Hill, J. L. (1996). Psychological Sense of Community: Suggestions for Future Research. *Journal of Community Psychology, 24*(4), pp. 431-438.

Hiltz, S. R. (1975). *Communications and group decision-making: Experimental evidence on the potential impact of computer conferencing* (Research Report No. 2). Newark, NJ: New Jersey Institute of Technology, Computerized Conferencing and Communications Center.

Hiltz, S. R. (1984). *Online Communities: a case study of the office of the future.* Norwood, New Jersey: Ablex Publishing Co.

Hiltz, S. R., Johnson, K., & Agle, G. (1978). Replicating Bales's problem solving experiments on a computerized conference: A pilot study (Research Report No. 8). Newark, NJ: *New Jersey Institute of Technology, Computerized Conferencing and Communications Center.*

Hiltz, S. R., Johnson, K., & Turoff, M. (1986). Experiments in Group Decision Making Communication Process and Outcome in Face-to-Face Versus Computerized Conferences. *Human Communication Research, 13*, pp. 225-252.

Hiltz, S. R., & Turoff, M. (1978). *The Network Nation: Human Communication via Computer.* Cambridge, MA: MIT Press.

Hiltz, S. R., & Turoff, M. (1981). The Evolution of User Behavior in Computerized Conferencing System. *Communication of the ACM, 24,* pp. 739-751.

Hiltz, S. R., & Turoff, M. (1985). Structuring Computer-Mediated Communication Systems to Avoid Information Overload. *Communication of the ACM, 28,* pp. 680-689.

Hiltz, S. R., Turoff, M., & Johnson, K. (1989). Experiments in Group Decision Making: Disinhibition, Deindividuation, and Group Process in Pen Name and Real Name Computer Conferences. *Decision Support Systems, 5*, pp. 217-232.

Hof, R. D., Browder, S., & Elstrom, P. (1997, May 5). Internet Communities. *Bus. Week.*

Im, I., & Hars, A. (2001). Effectiveness of a Personalized Technology Across Different Domains, User Groups, and Search Modes. *Paper presented at the International COnference on Inforamiton Systems (ICIS) 2001*, New Orleans, LA.

Irvine, S. E., Hein, T. L., & Laughlin, D. (1999). Different Degrees of Distance: the Impart of the Technology-based Instructional Environment on Student Learning. *Paper presented at the 29th Annual Frontiers in Education Conference*, San Juan, Puerto Rico, pp. 13C13/17 - 13C13/11.

Isaaces, E., Walendowski, A., Whittaker, S., Schiano, D. J., & Kamm, C. (2002). The Character, Functions, and Styles of Instant Messaging in the Workplace. *Paper presented at the 2002 ACM Conference on Computer Supported Cooperative Work*, New Orleans, Louisiana, pp. 11-20.

Jeffery, P., & McGrath, A. (2000). Sharing Serendipity in the Workplace. *Paper presented at the the Conference on Collaborative Virtual Environments (CVE)*, San Francisco, pp. 173-179.

Johansen, R., Vallee, J., & Spangler, K. (1988). Teleconferencing: Electronic Group Communication. In R. S. C. L. A. Samover (Ed.), *Small group communication: A reader* (pp. 140-154). Dubuque, IA: William C. Brown.

Jones, Q. (1997). Virtual-communities, Virtual-settlements and Cyber-archeology: A Theoretical Outline. *Journal of Computer-Mediated Communication, 3*(3).

Jones, Q., & Rafaeli, S. (1999). User Population and User Contributions to Virtual Publics: A Systems Model. *Paper presented at the ACM International Conference on Supporting Group Work (Group'99)*, Phoenix, AZ

Jones, Q. (2000). Time to Split, Virtually: Expanding Virtual Publics into Vibrant Virtual Metropolises. *Paper presented at the 33rd Hawaii International Conference on System Science*, Hawaii.

Jones, Q., & Rafaeli, S. (2000). What Do Virtual 'Tells' Tell? Placing Cybersociety research Into a Hierarchy of Social Explanation. *Paper presented at the 33rd Hawaii International Conference on System Science*, Hawaii.

Jones, Q., Ravid, G., & Rafaeli, S. (2002). An Empirical Exploration of Mass Interaction System Dynamics: Individual Information Overload and Usenet Discourse. *Paper presented at the 35rd Annual Hawaii International Conference on System Sciences*, Big Island, Hawaii.

Jung, Y., & Lee, A. (2000). Design of a Social Interaction Environment for Electronic Marketplaces. *Paper presented at the DIS'2000 - Designing Interactive Systems: Processes*, Practices, Methods, Techniques, pp. 129-136.

Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information Technology Adoption Across Time: A Cross-Sectional Comparison of Pre-Adoption and Post-Adoption Beliefs. *MIS Quarterly, 23*(2), pp. 183-213.

Kautz, H., Selman, B., & Shah, M. (1997, March, 1997). ReferralWeb: Combining Social Networks and Collaborative filtering. *Communication of the ACM, 40,* pp. 63-65.

Kerr, E. B., & R., H. S. (1982). *Computer-mediated communication systems: Status and evaluation*. New York: Academic Press.

Kieras, D. E., & Polson, P. G. (1985). An Approach To the Formal Analysis of User Complexity. *International Journal of Man-Machine Studies, 22*, pp. 365-394.

Kiesler, S. (1986). The Hidden Messages in Computer Networks. *Haward Business Review, 64*, pp. 46-54.

Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social Psychological Aspects of Computer-Mediated Communication. *American Psychologist, 39*, pp.1123-1134.

Kiesler, S., Zubrow, D., Moses, A. M., & Geller, V. (1985). Affection in computer-mediated communication. *Human Computer Interaction, 1*, pp.77-104.

Kling, R. (1996). Synergies and Competition Between Life in Cyberspace and Face-to-Face Communities. *Social Science Computer Review, 1*, pp. 50-54.

Kollock, P., & Smith, M. (1999). *Communities in Cyberspace*. London, UK: Routledge.

Kollock, P., & Smith, M. (1994). Managing the Virtual Commons. Cooperation and Conflict in Computer Communities. In S. Herring (Ed.), *Computer-mediated communication* (pp. 109-128). Amsterdam: John Benjamins.

Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., & Riedl, J. (1997, March, 1997). GroupLens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM, 40,* pp.77-87.

Konstan, J. A. (2002, November/December 2002). About Place. *SIG CHI Bulletin,* 3.

Korzenny, F. (1978). A Theory of Electronic Propinquity: Mediated Communications in organizations. *Communication Research, 5*, pp.3-24.

Kraut, R. E., Fish, R. S., Root, R. W., & Chalfonte, B. L. (Eds.). (1990). *Informal Communication in Organizations: Form, Function, and Technology.*: Sage Publications.

Krautz, R., Galegher, J., & Edigo, C. (1990). *Intellectual Teamwork: Social and Technological Bases for Cooperative work*. Hillsdale, NJ: Lawrence Erlbaum.

Kuehn, S. A. (1993). Communication innovation on a BBS: A Content Analysis. *Interpersonal Computing and Technology Journal, 1*(2).

Kumar, R., Novak, J., Raghavan, P., & Tomkins, A. (2003). Dynamic Services and Analysis: On the Bursty Evolution of Blogspace. *Paper presented at the 12th International conference on World Wide Web*, Budapest, Hungary, pp. 568-576.

Lawewnce, J. A., & Mongeau, P. (1996). The Role of Image and Anticipation of Future Interaction in Computer-Mediated Impression Development. *Paper presented at the International Communication Association*, Chicago.

Lee, A., & Girgensohn. (2002). Design, Experiences and User Prefernces for a Web-Based Awareness Tool. *International Journal of Human Computer Studies, 56*(1), pp. 75-107.

Lee, F. S. L., Vogel, D. R., & Limayem, M. (2003). Virtual Community Informatics: A Review and Research Agenda. *Journal of Information Technology Theory and Application, 5*(1), pp. 47-61.

Leonard-Barton, D., & Deschamps. (1988). Managerial Influence In the Implementation of New Technology. *Management Science, 34*, pp. 1252-1265.

Licklider, J. C. R., & Taylor, R. W. (1968). The Computer as a Communication Device. *Science and Technology*.

Loher, B. T., Noe, R. A., Moeller, N. L., & Fitzgerald, M. P. (1985). A Meta-Analysis of The Relation of Job Characteristics to Job Satisfaction. *Journal of Applied Psychology, 70*, pp. 280-289.

Malone, T. W., Grant, K. R., Turbak, F. A., Brobst, S. A., & Cohen, M. D. (1987). Intelligent Information Sharing Systems. *Communication of the ACM, 30,* pp. 390-402.

Maltz, D., & Ehrlich, K. (1995). Pointing the Way: Active Collaborative Filtering. *Paper presented at the 1995 Annual Conference on Human Factors in Computing Systems*, Denver, Colorado.

Markus, M. L. (1987). Toward a Critical Mass Theory of Interactive Media: Universal Access, Interdependence, and Diffusion. *Communication Research, 14*(5), pp. 491-511.

Markus, M. L. (1990). Toward a Critical Tass theory of Interactive Media: Universal Access, Interdependence, and Diffusion. In J. F. C. Steinfield (Ed.), *Organizations and Communication Technology*. Newbury Park, CA: Sage.

Mason, B. (1999). Issues in Virtual Ethnography published in Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities. *Paper presented at the Esprit i3 Workshop on Ethnographic Studies.*, Edinburgh, pp. 61-69.

Mathieson, K., Peacock, E., & Chin, W. W. (2001). Extending the technology acceptance model: the influence of perceived user resources. *ACM SIGMIS Database, 32*(3), pp. 86-112.

Matzat, U. (2002). Academic Communication and Internet Discussion Groups: Transfer of Information or Creation of Social Contacts. *Social Networks (Under Review).*

McGrath, J. E., Arrow, H., Gruenfeld, D. H., Hollingshead, A. B., & O'Conner, M. (1993). Groups, Tasks, and Technology: The Effects of Experience and Change. *Small Group Research, 24*, pp. 406-420.

McMillan, D. W. (1996). Sense of Community. *Journal of Community Psychology, 24*(4), pp. 315-325.

Meadows, A. J. (1974). *Communication in Science*. London: Butterworth.

Middleton, C. A. (1999). *Asynchronous Computer Conferencing in the MBA Classroom.* Paper presented at the 32nd Annual Hawaii International Conference on Systems Science, Hawaii, pp. 1013.

Millar, F. E., & Rogers, L. E. (1976). A Relational Approach to Interpersonal Communication. In G. R. Miller (Ed.), *Explorations in Interpersonal Communication* (pp. 87-104). Beverly Hills, CA: Sage.

Miller, H., & Mather, R. (1998). The Presentation of Self in WWW Home Pages. *Paper presented at the Internet Research and Information for Social Scientists (IRISS'98)*, Bristol, UK.

Moore, G. C., & Benbasat, I. (1991). Development of An Instrument to Measure The Perceptions of Adopting An Information Technology Innovation. *Information Systems Research, 2*, pp. 192-222.

Moore, G. C., & Benbasat, I. (1996). Integrating Diffusion of Innovations and Theory of Reasoned Action Models to Predict Utilization of Information Technology by End-Users. In K. K. a. J. Pries-Hege (Ed.), *Diffusion and Adoption of Information Technology* (pp. 132-146). London: Chapman and Hall.

Moreland, R. L., & Levine, J. M. (1982). Socialization in Small Groups: Temproral Changes in Individual-Group Interactions. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 15, pp. 137-192). New York: Academic Press.

Morningstar, C., & Farmer, F. R. (1991). The lessons of Lucasfilm's Habitat. In M. Benedikt (Ed.), *Cyberspace* (pp. 271-301). Cambridge: MIT Press.

Morris, M., & Ogan, C. (1996). The Internet as Mass Medium. *Journal of Communication (JCMC), 46*(1).

Mostafa, J., Mukhopadhyay, S., Lam, W., & Palkal, M. (1997). A Multilevel Approach to Intelligent Information Filtering: Model, System and Evaluation. *ACM Transactions on Information Systems, 15*(4), pp. 368-399.

Mukjerjee, R., Dutta, P. S., Jonsdottir, G., & Sen, S. (2001). *MOVIES2GO-An Online Voting Based Movie Recommender System.* Paper presented at the Agent'01, Montreal, Quebec, CA.

Nakamura, M. (1999). Design and Implementation of the Smtpfeed to Improve Delivery Speed of Mailing Lists. *IPSJ Symposium* (3), pp. 27-32.

Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, C. (1995). Can Computer Personalities be Human Personalities? *Paper presented at the Human Factors in Computing Systems*, Denver, CO, 228-229.

Nguyen, H., & Haddawy, P. (1998). Applying Decision Theory to Collaborative Filtering. *Paper presented at the AAAI Workshop on Recommender Systems*, Madison.

Nonnecke, B. (2000). *Lurking in Email-Based Discussion Lists.* Unpublished Ph.D. Thesis, South Bank University, London.

Nonnecke, B., & Preece, J. (2000). Lurker Demographics: Counting the Silent. Human Factors in Computing Systems. *Paper presented at the Conference on Human Factors in Computing Systems(CHI'00)*, The Hague, The Netherlands, pp. 73-80.

Nonnecke, B., & Preece, J. (2001). Why Lurkers Lurk. *Paper presented at the Americas Conference on Information Systems*, Boston, pp. 1520-1530.

Oliver, P., Marwell, G., & Teixeira. (1985). A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of the Collective Action. *American Journal of Sociology, 91*(3), 522-556.

Onishi, M., Kagebayashi, T., & Fukunaga, K. (2001). *Production of Vieo Images by Computer Controlled Cameras and its Application on TV Conference System.* Paper presented at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), pp. 131-137.

Ord, J. G. (1989). Who's joking? The Information System at Play. *Interacting with Computers, 1*, pp. 118-128.

Over, R. (1982). Collaborative Research and Publication in Psychology. *American Psychologist, 37*, pp. 996-1001.

Palme, J. (1981). Interactive software for Humans. In B. Shackel (Ed.), *Man-Computer Interaction: Human Factors Aspects of Computer & People*: Sijthoff & Noordhoff.

Palme, J. (1997). Choices in the Implementation of Rating. In R. Alton-Scheidl, Schumutzer, R., Sint, P. P. and Tscherteu, G. (Ed.), *Voting, Rating, Annotation: Web4Groups and other projects: approaches and first experiences* (pp. 147-162). Vienna, Austria: Oldenbourg.

Parks, M. R. (1977). Relational Communication: Theory and Research. *Human Communication Research, 3*, pp. 372-381.

Parks, M. R., & Floyd, K. (1996). Making Friends in Cyberspace. *Journal of Communication, 46*(1), 80-97.

Peek, R. P., & Newby, G. B. (1996). *Scholarly Publishing. The Electronic Frontier.* Cambridge, London: The MIT Press.

Pemberton, D., Rodden, T., & Procter, R. (2000). *GroupMark: A WWW Recommender System Combining Collaborative and Information Filtering.* Paper presented at the 6th ERCIM Workshop, Florence, Italy.

Pennock, D. M., Eric Horvitz, Steve Lawrence and C.Lee Giles. (2000). Collaborative Filtering by Personality Diagnosis: A Hybrid Memory - and Model-Base Approach. *Paper presented at the Sixth Conference on Uncertainty in Artificial Intelligence (UAI-2000)*, Morgan Kaufmann, San Francisco, pp. 473-480.

Plouffe, C. R., Hulland, J. S., & Vandenbosch, M. (2001). Research Report: Richness Versus Parsimony in Modeling Technology Adoption Decisions - Understanding Merchant Adoption of a Smart Card-Based Payment System. *Information Systems Research, 12*(2), pp. 208-222.

Polson, P. G. (1987). A Qualitative Theory of Human-Computer Interaction. In J. M. Carroll (Ed.), *Interfacing Thought*. Cambridge, MA: MIT Press.

Poole, M. S., & DeSanctis, G. (1989). Use of Group Decision Support Systems as an Appropriation Process. *Paper presented at the 22nd Hawaii International Conference on System Sciences*, Hawaii, 149-157.

Poole, M. S., & DeSanctis, G. (1992). Microlevel Structuration in Computer-Supported Group Decision Making. *Human Communication Research, 19*(1), 5-49.

Preece, J. (2000). *Online Communities: Designing Usability, Supporting Sociability*. Chichester, UK: John Wiley & Sons.

Preece, J., & Ghozati, K. (1998). In Search of Empathy Online: A Review of 100 online Communities. *Paper presented at the 1998 Association for Information Systems Americas Conference*, Baltimore, pp. 92-94.

Rafaeli, S., & LaRose, R. J. (1993). Electronic Bulletin Boards and 'public goods' Explanations of Collaborative Mass Media. *Communication Research, 20*(2), pp. 177-197.

Ramakrishnan, N. (1997). *Recommender Systems for Problem Solving Environments*. Unpublished Ph. D., Purdue University, West Lafayette, IN.

Reid, E. M. (1991). *Electropolis: Communication and Community on Internet Relay Chat.*, University of Melbourne.

Reid, E. M. (1994). *Culture Formations in Text-based Virtual Realities*. Unpublished M.A., University of Melbourne.

Reid, E. M. (1995). Virtual Worlds: Culture and Imaginations. In S. G. Jones (Ed.), *Cybersociety: Computer-mediated communication and community* (pp. 164-183). Thousand Oaks, CA: Sage.

Resnick, P., & Varian, H. (1997). Recommender systems. *Communication of the ACM, 40*(3), 56-58.

Rheingold, H. (1993). *The Virtual Community*. MIT Press.

Rice, R. E., & Case, D. (1983). Electronic Message Systems in the University: A Description of Use and Utility. *Journal of Communication, 33*, pp. 131-152.

Rice, R. (Ed.). (1994). *Network Analysis and Computer-Mediated Communication Systems*. Newbury Park, CA: Sage.

Rice, R. E. (1984). Mediated Group Communication. In R. E. R. Associates (Ed.), *The new media Communication, research and technology* (pp. 129-156). Beverly Hills, CA: Sage.

Rice, R. E., & Love, G. (1987). Electronic Emotion: Socioemotional Content in a Computer-Mediated Network. *Communication Research, 14*, 85-108.

Rice, R. E. (1993). Media Appropriateness: Using Social presence Theory to Compare Traditional and New Organizational Media. *Human Communication Research, 19*(4), pp. 451-484.

Rivard, S., & Huff, S. (1988). Factors of Success for End-User Computing. *Communication of ACM, 31*(5), pp. 552-561.

Robertson, S. E. (1977). The Probability Ranking Principle in IR. *Journal of Documents, 33*(4), 294-304.

Rockart, J. F., & DeLong, D. W. (1988). *Executive support systems: The emergence of top management computer use*. Homewood, IL: Dow Jones-Irwin.

Rogers, E. M. (1986). *Communication Technology: The New Media in Society*. New York: Free Press.

Rogers, E. (1995). *Diffusion of Innovations*. New York: Free Press.

Rovai, A. P. (2002). A Preliminary Look at The Structural Differences of Higher Education Classroom Communities in Traditional and ALN Courses. *JALN, 6*(1).

Rucker, J., & Polano, M. J. (1997). Siteseer: Personalized Navigation for the Web. *CACM, 40*(3).

Saito, N., Mizasawa, J., Yamamoto, H., & Yamaguchi, S. (1998). Networked Collaboration and its Support Systems. *Transactions of Information Processing Society of Japan, 39*(10), pp. 2907-2913.

Salton, G., & McGill, M. J. (1983). *Introduction to Modern Information Retrieval.*: McGraw-Hill, Inc.

Sakata, K., & Kurashima, A. (2001). A Mailing List System Enabling Consummatory Communication. *Paper presented at the 2001 Symposium on Applications and the Internet (SAINT 2001)*, San Diego, CA, pp. 515-522.

Sarwar, B. M. (1998). *Using Semi-intelligent Filtering Agent to Improve Prediction Quality in Collaborative Filtering Systems.* Unpublished M.S. Thesis, University of Minnesota, Minneapolis.

Sarwar, B. M., Karypis, G., Konstan, J. A., and Riedl, J. (2000). Application of Dimensionality Reduction in Recommender System -- A Case Study.: Full length paper at ACM WebKDD 2000 Web Mining for E-Commerce Workshop.

Schafer, J. B., Konstan, J., and Riedl, J. (2001). Electronic Commerce Recommender Applications. *Journal of Data Mining and Knowledge Discovery, 5*, pp. 115-152.

Schatz, B. R. (1992). Building an Electronic Community System. *Journal of Management Information Systems, 8*(3), pp. 87-101.

Schoberth, T., Preece, J., & Heinzi, A. (2003). Online Communities: A Longitudinal Analysis of Communication Activities. *Paper presented at the 36th Hawaii International Conference on System Sciences*, Hawaii.

Shardanand, U., & Maes, P. (1995). Social Informaiton Filtering: Algorithms for Automating 'Word of Mouth'. *Paper presented at the CHI'95*, Denver, Colorado.

Sheridan, T. B., & Ferrell, W. R. (1974). *Man-Machine Systems: Informaiton, Control, and Decision Models of Human Performance*. Cambridge, MA: MIT Press.

Short, J., Williams, E., & Christie, B. (1976). *The Social Psychology of Telecommunications*. New York: John Wiley.

Siegel, J., Dubrovsky, V., Kiesler, S., & McGuire, T. W. (1986). Group Processes in Computer-Mediated Communication. *Organizational Behavior and Human decision processes, 37*, pp. 157-187.

Smith, M. (1992). *Voice from the Well: The logic of the virtual commons*. Unpublished Master's Thesis, UCLA.

Smith, M. H., Trajakovic, L., & Rubin, S. (1998). Fuzzybase: An Information-intelligent Retrieval System. *Paper presented at the International Conference on Systems, Man, and Cybernetics*, San Diego, CA, pp. 2797-2802.

Smith, M. A., & Kollock, P. (Eds.). (1999). *Communities in Cyberspace*. New York: Routledge.

Spears, R., & Lee, M. (1992). Social Influence and the Influence of the "social" in Computer-Mediated Communication. In M. Lea (Ed.), *Contexts of computer-mediated communication* (pp. 30-65). London: Harvester Wheatsheaf.

Sproull, L., & Kiesler, S. (1986). Reducing Social Context Cues: Electronic Mail in Organizational Communication. *Management Science*, pp. 1492-1512.

Sproull, L., & Kiesler, S. (1991). *Connections: New Ways of Working in the Neworked Organization*. Cambridge, MA: MIT Press.

Sproull, L., & Faraj, S. (1997). Atheism, Sex and Databases: the Net as a Social Technology. In S. Kiesler (Ed.), *Culture of the Internet* (pp. 35-52). Mahwah, NJ: Erlbaum.

Srivastava, J., Cooley, R., Deshpande, M., & Tan, P. N. (2000). Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explorations, 1*(2), pp. 12-23.

Star, S. L., & Ruhleder, K. (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Informaiton Spaces. *Information System Research, 7*(1), pp. 111-134.

Steinfield, C. W. (1986). Computer-Mediated Communication in an Organizational Setting: Explaining Task-related and Socioemotional uses. In M. L. McLaughlin (Ed.), *Communication yearbook 9*. Beverly Hills, CA: Sage.

Stewart, C. M., Shields, S. F., Monolescu, D., & Taylor, J. C. (1999). Gender and Participation in Synchronous CMC: an IRC Case Study. *Interpersonal Computing and Technology, 7*(102).

Storck, J., & Sproull, L. (1995). Through a Glass Darkly: What do People Learn in Videoconferences? *Human Communication Research, 22*, pp. 197-219.

Swearingen, K., & Sinha, R. (2001). Beyond Algorithms: An HCI Perspective on Recommender Systems. *Paper presented at the ACM SIGIR 2001 Workshop on Recommender Systems*, New Orleans, Lousiana.

Szajna, B. (1994). Software Evaluation and Choice: Predictive Validation Of the Technology Acceptance Instrument. *MIS Quarterly, 18*(3), pp. 319-324.

Tak, W. Y., & Garcia-Molina, H. (1999). The SIFT Information Dissemination System. *ACM Transactions on Database System (TODS), 24*(4), 529-565.

Taylor, S., & Todd, P. A. (1995). Understanding Information Technology Usage: A Test of Competing Models. *Information Systems Research, 6*(2), pp. 144-176.

Tang, J. C., Bonte, A., Raven, M. B., & Isaaces, E. (2000). Instant Messaging: Products Meet Workplace Users. *Paper presented at the 2000 ACM conference on Computer Supported Cooperative Work*, Philadelphia, PA, p 363.

Terveen, L., Hill, W., Amento, B., McDonald, D., & Creter, J. (1997). PHOAKS: a System for Sharing Recommendations. *Communication of the ACM, 40,* pp. 59-62.

Terveen, L., McMackin, J., Amento, B., & Hill, W. (2002). Specifying Preferences Based on User History. *Paper presented at the CHI2002 conference on Human-factors on Computing Science*, Minneapolis, MN.

Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal Computing: Toward a Conceptual Model of Utilization. *MIS Quarterly, 15*(1), pp. 124-143.

Thomas, B. (1997). Real interactivity - Internet discussion lists. *IEEE Internet Computing, 1*(1), pp. 81-83.

Tilley, S. R., & Smith, D. B. (1996). Documenting Virtual Communities. *Paper presented at the 14th annual international conference on systems documentation*, Research Triangle Park, North Carolina, pp. 43-49.

Tornatzky, L. G., & Klein, K. (1982). Innovation Characteristics and Innovation Adoptation-Implementation: A Meta-Analysis of Findings. *IEEE Transaction on Engineering Management, 29*(1), pp. 28-45.

Turkle, S. (1994). Constructions and Reconstructions of Self in Virtual Reality: Playing in the MUDs. *Mind, Culture and Activity, 1*(3), pp. 72-84.

Uhlig, R. P., Farber, D. J., & Bair, J. H. (1979). *The Office of the Future: Communication and Computers*. Amsterdam: North-Holland.

Ungar, L. H., & Foster, D. P. (1998). Clustering Methods for Collaborative Filtering. *Paper presented at the AAAI Workshop on Recommendation System.*

Utz, S. (2000). Social Information Processing in MUDs: The Development of Friendships in Virtual Worlds. *Journal of Online Behavior, 1*(1).

Vallerand, R. J. (1997). Toward a Hierarchical Model of Intrinsic and Extrinsic Motivation. *Advances in Experimental Social Psychology, 29*, pp. 271-360.

Van Gelder, L. (1985). The Strange Case of the Electronic Lover. In C. D. R. Kling (Ed.), *Computerization and Controversy: Value Conflicts and Social Choices* (pp. 364-375). Boston: Academic Press.

Venkatesh, V. (1999). Creating Favorable User Perceptions: Exploring the Role Of Intrinsic Motivation. *MIS Quarterly, 23*(2), pp. 239-260.

Venkatesh, V., & Davis, F. D. (1996). A Model Of the Antecedents of Perceived Ease of Use: Development and Test. *Decision Science, 27*(3), pp. 451-481.

Venkatesh, V., & Morris, M. G. (2000). Why Don't Man Ever Stop To Ask For Directions? Gender, Social Influence, and Their Role in Technology Acceptance and User Behavior. *MIS Quarterly, 24*, pp. 115-139.

Venkatesh, V. (2000). Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion Into the Technology Acceptance Model. *Information Systems Research, 11*(4), pp. 342-365.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension Of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science, 46*, pp. 186-204.

Venkatesh, V., Morris, M., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly, 27*(3), pp. 425-478.

Vessey, I. (1991). Cognitive Fit: A Theory-Based Analysis of The Graphs Versus Tables Literature. *Decision Science, 22*, pp. 219-240.

Vincent, C. (1992). *Collegiality in Cyberspace, Case Studies in Computer Mediated Communication.* Unpublished Master of Social Science, University of Tasmania.

Wang, Y., Zhang, X., & Li, W. (2000). Audio Multimedia Conferencing System Based on the Technology of Speech Recognition. *Paper presented at the 2000 IEEE Asia-Pacific Conference on Circuits and Systems*, Tianjin, China, pp. 771-774.

Wang, C., Griebel, S., & Brandstein, M. (2000, July 30 - Aug 2). Robust automatic video-conferencing with multiple cameras and microphones. *Paper presented at the 2000 IEEE International Conference on Multimedia and Expo (ICME 2000)*, pp. 1585-1588.

Walther, J. (1994). Anticipated Ongoing Interaction versus Channel Effects on Relational Communication in computer-mediated interaction. *Human Communication Research, 40*, pp. 473-501.

Walther, J. B. (1992). Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective. *Communication Research, 19*(1), pp. 52-90.

Walther, J. B., Anderson, J. F., & Park, D. W. (1994). Interpersonal Effects in Computer-Mediated Interaction: A Meta-Analysis of Social and Antisocial Communication. *Communication Research, 21*(4), pp. 460-487.

Walther, J. (1995). Relational Aspects of Computer-Mediated Communication: Experimental Observations Over Time. *Organization Science, 6*(2), pp. 186-203.

Walther, J. B., Slovacek, C., & Tidwell, L. C. (2001). Is a Picture Worth a Thousand Words? Photographic Images in Long Term and Short Term Virtual Teams. *Communication Research, 28*, 105-134.

Wasserman, S., & Galaskiewicz (Eds.). (1994). *Advances in Social Network Analysis.* Thousand Oaks, CA: Sage.

Weedman, J. (1991). Task and Non-task Functions of a Computer Conference Used in Professional Education: A Measure of Flexibility. *International Journal of Man-Machine Studies, 34*, 303-318.

Weick, K. E. (1979). *The Social Psychology of Organizing.* Reading, MA: Addison-Wesley.

Weil, V. (1998). The Information Revolution. A Dose of Reality. *Science Communication, 1*, pp. 136-141.Wellman, B. (2001). Computer Networks As Social Networks. *SCIENCE, 293,* pp. 2031-2034.

Wellman, B. (1997). *Using Social Network Analysis to Study Computer Networks.* Paper presented at the GROUP 1997

Wellman, B., Salaff, J., Dimitrova, D., Garton, L., Gulia, M., & Haythornthwaite, C. (1996). Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community. *Annual Review of Sociology, 22*, pp. 213-238.

Wellman, B. (2001). Computer Networks As Social Networks. *SCIENCE, 293,* pp. 2031-2034.

Whittaker, S., Issacs, E., & O'Day, V. (1997). Widening the Net. *Paper presented at the SIGCHI*, 27-30.

Whittaker, S., Terveen, L., Hill, W., & Cherny, L. (1998). The Dynamics of Mass Interaction. *Paper presented at the ACM conference on Computer-Supported Cooperative Work*, Seattle Washing, pp. 257-264.

Whyte. (1988). *City: Rediscovering the Center.* New York: Doubleday.

Williams, F., & Rice, R. E. (1983). Communication Research and New Media Technologies. In R. N. Bostrom (Ed.), *Communication yearbook 7.* Beverly Hills, CA: Sage.

Withworth, B. (1997). *Generating Group Agreement in Cooperative Computer-mediated Groups: Towards and Integrative Model of Group Interaction.* Unpublished Doctor of Philosophy, University of Waikato.

Word, H. (1975). Path Models with Latent Variables. The NIPALS Approach. In H. M. Blalock (Ed.), *Quantitative Sociology* (pp. 307-357). New York: Seminar Press.

Word, H. (1982). Soft Modeling: The Basic Design and some Extensions. In K. G. Joreskog & H. Word (Eds.), *Systems Under Indirect Observation, Part II.* Amsterdam: Morth Holland Press.

Word, H. (1985). Partial Least Squares. In S. J. Kotz, N. L. (Ed.), *Encyclopedia of Statistical Sciences* (Vol. 6). New York: John Wiley & Sons.

Wynn, E., & Katz, J. E. (1997). Hyperbole over Cyberspace: Self-presentation & Social Boundaries in Internet Home Pages and Discourse. *Information Society, 13*(4), pp. 297-328.

Xu, H. (1998). *Global Access and its Implications: The use of mailing lists by systems librarians.* Paper presented at the ASIS Annual Meeting, pp. 501-515.

Yoo, Y., & Alavi, M. (2001). Media and Group Cohesion: Relative Influence on Social Presence, Task Participation, and Group Consensus. *MIS Quarterly, 25*(3), pp. 371-390.

Yu, C. H. (1997). *Checking Assumptions in Regression.* Retrieved 3/22, 2004, from http://seamonkey.ed.asu.edu/~alex/computer/sas/regression_assumption.html

Zhang, Y., Bot, R. & Hiltz, S. R. (2001). Building a Virtual Community for ALN. *Paper presented at the Seventh Americas Conference on Information Systems*, Boston, MA.

Zhang, Y., & Im, I. (2002). Recommender Systems: A Framework and Research Issues. *Paper presented at the Americas Conference on Information Systems* 2002, Dallas, TX.