Arkose: A Prototype Mechanism and Tool for Collaborative Information Generation and Distillation

By

Kevin Kyung Nam

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Information) in the University of Michigan 2010

Doctoral Committee:

Professor Mark S. Ackerman, Chair
Professor Thomas A. Finholt
Assistant Professor Mark W. Newman
Assistant Professor Kai Zheng

© Kevin Kyung Nam
All Rights Reserved
2010

Dedication

To Jei Yern

Acknowledgements

I am deeply appreciative of many people who have supported and encouraged me throughout the writing of this dissertation. Although I can only mention a few in this limited space, this dissertation could not have happened without their support.

My deepest gratitude goes to my advisor Mark S. Ackerman whose guidance and training have been invaluable in successfully completing my dissertation. Without his encouragement and patience, I could not have possibly finished this. He is truly a model researcher, advisor, and teacher.

I thank my committee members and co-authors, Tom Finholt, Mark Newman, Kai Zheng, and Lada Adamic, for helping me shape the research problems and plans and giving me insights and guidance. Their passion and enthusiasm have been enormously inspirational.

I would like to mention my intellectual brothers and sisters, especially Mark Ackerman's SocialWorlds research group. Jun Zhang, Xiaomu Zhou, Jina Huh, Ben Congleton, Jiang Yang, Ayşe Büyüktür, and Tao Dong have been wonderful colleagues and friends for several years. I have also greatly benefited from my associations with many doctoral students at the School of Information. Emilee Rader and Libby Hemphill gave me helpful advice as senior students. John Lin tremendously supported me with his thoughtful prayers. Many others have helped me mentally, physically, and spiritually.

I also wish to thank the support staff at the School of Information, especially Sue Schuon, who has always been tremendously friendly and helpful. I am very fortunate to have been in the SI community.

Finally and most importantly, I wish to thank my family. My parents, parents-in-law, and sister Jinny deserve much appreciation for supporting and encouraging me throughout this journey. My deepest love and thanks go to my wife Jei Yern. She stood by me and believed in me even when I had doubts about myself. I could not have done it without her.

Table of Contents

List of Figures

List of Tables

List of Appendices

Abstract

Arkose: A Prototype Mechanism and Tool for Collaborative Information Generation and Distillation

By

Kevin Kyung Nam

Chair: Mark S. Ackerman

The goals of this thesis have been to gain a better understanding of collaborative knowledge sharing and distilling and to build a prototype collaborative system that supports flexible knowledge generation and distillation. To reach these goals, I have conducted two user studies and built two systems.

The first system, Arkose 1.0, is a prototype collaborative distillation system for a discussion space, which provides a set of augmentative tools to facilitate the filtering, structuring, and organizing of discussion information. Arkose 1.0 supports editors to distill a discussion space incrementally and collaboratively, and allows a gradual increase in the order and reusability of the information space.

The study of an online question-answering community, Naver Knowledge-iN, investigates users' knowledge sharing behaviors in a large online question-answering community. Through the analyses of a large quantity of question/answer pairs and 26

user interviews, the study analyzes the characteristics of knowledge generation and user participation behavior and gains insights into their motivations, roles, usage and expertise. It reveals that the limiting nature of the reply interfaces of Knowledge-iN leads to the accumulation of simple and easy questions and answers. This tendency is encouraged by the point system that rewards users who answer many questions quickly.

Arkose2[1] is designed and implemented based on the lessons and insights gained in building Arkose 1.0 and examining Naver Knowledge-iN. Arkose2 provides a host of additional interaction mechanisms and supportive tools over Arkose 1.0 that assists users to flexibly generate knowledge and distill and organize it better.

Finally, the evaluation of Arkose2 reveals a number of insights, issues, and lessons about users' distillation activities of discussion spaces and features of Arkose2.

These together provide valuable lessons and insights for the architecture and features of the next generation collective intelligent system.

[1] In this thesis, Arkose 1.0 refers to the first generation of Arkose and Arkose2 refers to the second generation.

Chapter 1

Introduction

Community Information Generation

Communities know a great deal. It is clear that groups of people, especially large or Internet-scale groups, have a greater understanding of a problem and its issues than any individual or even a select committee. With the advent of the World Wide Web and Internet technologies, many communities use asynchronous virtual interactions as a way to generate and share knowledge, and it has become common to hold a large group discussion online. For example, this could be used for a brainstorming on product ideas or discussion on new technology deployed within a corporate setting. Governments, institutions, and universities could discuss various future plans for organizational changes in order to reach a "shared mind".

A successful knowledge building community depends on continuing interactions among the users and accumulation of quality information that help its members solve a problem. The collective intelligence being created is undoubtedly critical for the success of a community. Popular examples include mailing lists, discussion forums, wikis, and more recently large online communities for general purpose question-answering. These tools allow large-scale collaboration in which members voluntarily share their knowledge in a variety of topics, provide access to expertise, and help create a support community with their ongoing activities. The content and structure of a knowledge building community are often organic and evolving to fit the needs of its members.

While each of the tools has its strength in knowledge production and sharing, they do not handle certain types of information very well. It is relatively difficult to keep track of past information in a mailing list because fragments of information are sent to a user. An online discussion forum can allow users to conduct an in-depth discussion, but the conclusion and relevant information are often buried within other non-essential information. In contrast, the questions and answers in a large-scale question-answering community are generally simple compared to mailing lists or discussion forums, but they are contextualized to fit the users' needs. A wiki, on the other hand, is good at organizing and structuring information under a topic. However, a wiki is not generally very responsive to individual's ongoing information needs.

Each knowledge building community specializes in the types of information and the kinds of operations users can carry out. The difference among the knowledge building communities may be due to the type of knowledge production they envision. Some factors that shape the information accumulated in a knowledge building community may be the expertise of the members and culture of the community. While easy beginner's questions are generally accepted in a general purpose question-answering community, they are often ignored or even frowned upon in a more specialized and technical forum. A community of certain expertise may expect new users to "do their homework" before posing a question in the community.

Yet another important factor may be the site's interaction mechanisms and user interfaces that inherently limit the kinds of information it can generate and handle. For example, an online discussion forum allows users to carry out asynchronous threaded conversations, but the lack of facilities that support the creation of a summary of ideas makes it difficult to save a meaningful conclusion or "convergence of the conversations" Hewitt (2001) calls it. He states that the current reply-based interaction model for discussion threads "imposes a bias in favor of divergent, branching discourse." He argues that the absence of consolidating facilities allows threads to grow "unchecked", making it more difficult to draw coherent ideas. The shape of the conversation itself may be affected by the affordances that the interfaces provide (Wang et al. 2008; Hewitt 2003). These factors affect the generation of information in an online community.

Community Information Distillation

Another important concern is the reuse of the generated information. Communities may generate useful information or reach a shared understanding in problem solving activities, whether it is an answer to a user's specific question, or general information applicable to a large audience of the community. In any case, users might want to revisit or reuse that information and understanding. A standard problem with online information repositories is that once use has ceased either by deadline or by neglect, a site is often a bramble of ideas and topics, too large and unwieldy for its information to be successfully reused. A process of filtering, structuring and organizing of the information, or the process of *distillation*, is needed (Ackerman and McDonald 1996; Ackerman et al. 2003). Distillation is the process of creating a more concise and organized form of information. It is more than just text summarization; rather, it includes sorting through large corpus of text for interesting topics, pruning away redundant and off-topic discussions, identifying interesting authors, different points of view, and ultimately making the information more reusable for later purposes. Similar approaches have been studied such as the concept of "condensation" (Brewer 2000) of a mailing list to extract useful information out of conversations.

Ideally, the distillation process would be performed by a variety of natural language processing methods to minimize the use of precious human resources. Techniques such as automatic summarization (Radev and Hovy 1999), discourse analysis (Passonneau and Litman 1997), and sentence structure parsing (Klein and Manning 2001) have been gradually improving, but they do not yet provide the human level cognitive abilities necessary for the aforementioned distillation tasks. As revealed in the evaluation of Arkose2, even human readers found it more difficult to distill information in which they did not have enough background knowledge or expertise. Thus, I view distillation as a system augmented process, guided and directed by human editors, rather than an entirely automated one.

Arkose: A System for Collaborative Information Generation and Distillation

My thesis concerns user interactions in knowledge building and sharing communities, and the design and development of new interaction mechanisms and new collaborative tools for information generation and distillation. The thesis consists of three of my studies. The first is the original implementation of Arkose (Nam and Ackerman 2007) (or Arkose 1.0 in this dissertation) which provides tools for gradual distillation of information from a completely unstructured state (e.g., a collection of raw discussion threads). The second is a study of a general purpose question-answering community Naver Knowledge-iN (Nam et al. 2009) in which I examine the characteristics of knowledge generation and user participation behavior through quantitative analysis of question-answer data from the site and qualitative analysis of user interviews. Finally, Arkose2 is a prototype collective intelligence system for both flexible knowledge generation and distillation that builds on the lessons and ideas from the first two studies. Arkose2 enables collaborative distillation and utilization of reusable knowledge in more flexible ways than Arkose 1.0. (See Table 5-1 for comparison.) Arkose2 builds on the understanding of the issues in knowledge sharing in an online community and experience in implementing the first generation of Arkose. It tries to overcome some of the weaknesses of previous online discussion forums, question-answering communities, and wikis to enable better knowledge generation and distillation, which may ultimately help reuse collaboratively generated information. While Arkose 1.0 allows a *post hoc* distillation by a small number of editors after the discussion is completed, Arkose2 concerns both flexible generation of information as well as ongoing distillation by the community as a whole.

In summary, the purpose of this dissertation is to design and develop a set of augmentative tools that support flexible generation and distillation of information in an organization or online community and examine various aspects of collaborative information sharing and distillation of discussion spaces. Two implementations, Arkose 1.0 and Arkose2, are built on four design principles; *incremental summarization, incremental diagenesis, ongoing distillation,* and *information linkage.* They demonstrate novel user interfaces and interaction mechanisms that support users' knowledge generation and distillation. Two user studies, Naver Knowledge-iN and the evaluation of Arkose2 examine user behaviors in knowledge sharing and distillation. The findings from these systems and evaluation studies have significant implications for future collaborative knowledge systems.

Dissertation Outline

The dissertation is organized into 7 chapters including this introductory chapter. Chapter 2 is a literature review where I define terminology and discuss various knowledge building communities. First, I will go over different types of popular community information repositories including their characteristics and challenges. I will then discuss what others have done to overcome the challenges. Literature that served as the bases of the design choices of Arkose is discussed next. Chapter 3 describes the first generation of my prototype collaborative information distillation system, Arkose 1.0. The chapter discusses some of the problems in information distillation and reuse that Arkose 1.0 addresses, its design principles, and implementation details.

Chapter 4 is a study of an online question-answering community Naver Knowledge-iN in which I analyze a large number of question-answer data and user interviews to characterize its knowledge generation and user interaction behaviors. The interviews illustrate the limitations of the user interfaces provided by the site, and how they shape and limit users' interactions with one another. The lessons from this study helped shape the design of the prototype Arkose2 system.

Chapter 5 discusses the second generation of Arkose. Arkose2 is both a collective information system as well as collaborative information distillation system. In this chapter, I describe its design principles, conceptual architecture, usage scenario, and implementation details.

Chapter 6 is the evaluation and analysis of Arkose2. Here I discuss a number of insights, issues, and lessons about users' distillation activities and Arkose2.

In Chapter 7, I conclude with the design implications and future work.

Chapter 2

Literature Review

Information Generation and Distillation in Online Knowledge Building Communities

The advancements of the computer and Internet technologies have enabled many new opportunities for managing knowledge in an organization. Ackerman and Halverson (2004) discuss three technical possibilities that the new technologies have brought for knowledge management. First, the notion that knowledge pieces can be stored and reused has led to the creation of knowledge *repositories* where knowledge created in an organization would be saved for future usage. Second, *computer-mediated places* have allowed a large number of people to create and share information through many mechanisms. These include virtual help providing systems (e.g. Answer Garden (Ackerman 1994) and Answer Garden 2 (Ackerman and McDonald 1996)), mailing lists, online discussion forums, wikis, and more recently general purpose question-answering communities. Finally, they discuss a vision of *ad-hoc groupings* of people in order to form flexible groups of people to solve contingency problems.

Indeed, there has been an explosion of new online communities enabling new ways for sharing knowledge and providing help to its members. Many types of these communities exist. Some are virtual and residing entirely online, while some are integrated with an existing physical (e.g., IBM My developerWorks Forums[1]) or another virtual community (e.g., eBay Discussion Boards[2]) as a supplementing information repository. These knowledge communities and communities of practice have been well studied (Preece 2000; Wenger 1998).

The underlying technology for each knowledge community also defines what information and knowledge are accumulated, how the interactions among the users are shaped, and what challenges it faces. Popular examples of community information repositories include mailing lists and Usenet, online discussion forums, wikis, and more recent general purpose question-answering communities. These have been popular places for users to generate and share knowledge. For example, every day, people put significant effort into online discussions (e.g., with Slashdot as in Lampe and Resnick 2004). The WorldJam experiment put IBM employees from around the world for a limited time into an online community for brainstorming (Millen and Fontaine 2003). Similarly, earlier work includes iDIAG/Forum, a prototype for creating community brainstorming (Ackerman et al. 2003). More recently, ultra large community question-answering sites such as Yahoo! Answers and Naver Knowledge-iN have gained popularity as a venue for sharing and getting knowledge. Wikis are another popular form of online community, albeit a specialized form. Wikis, such as Wikipedia, have been used to collaboratively create structured and organized information under various topics. Only some results from online community discussions are suited to wikis, however. (compendium entries were a primary form found in Hansen et al. 2007).

---

[1] https://www.ibm.com/developerworks/forums/index.html
[2] http://forums.ebay.com/category/Ebay-Discussion-Boards/2001

The amount of information accumulated in these online communities is often enormous, which makes it difficult to find relevant information. However, very little work has been done to distill and reuse the information within the communities. These knowledge communities and other collaborative help providing tools are my focus in this literature review.

Definition

In this dissertation, I use the term *Online Knowledge Building Community* to describe any community whose main purpose is to generate and share information and knowledge among its members through computer mediated spaces in order to satisfy each other's information needs. This includes both virtual communities that reside entirely *online* and physical communities that have online counterparts where members of an existing organization meet and share knowledge.

Many types of virtual communities exist (See Porter 2004 for a typology of virtual communities), but not all of them focus on knowledge sharing and helping members with their informational needs. People have used many terms to describe these communities, some of which are broadly defined. Some have used the term *knowledge building communities* (e.g., Hoadley and Kilner 2005), *online knowledge communities* (e.g., Vries and Kommers 2004), *online support communities* (e.g., Hansen 2007), *community question answering* (Liu et al. 2008a), and *communities of practice* (Wenger 1998). These communities are also often defined by the computer mediated spaces and technologies they employ, which include, but not limited to, mailing lists, newsgroups, discussion forums, wikis, and more recently general purpose question-answering sites. There are also collaborative help providing systems that build organizational memory (e.g., Answer Garden (Ackerman and Malone 1990) and Answer Garden 2 (Ackerman and McDonald 1996)).

For the success of a knowledge building community, one may need to look at all aspects and dimensions of an online community such as the membership, incentive system, policies, expertise, culture and technology of the community. I focus on the technical dimension in this dissertation by designing and developing new user supportive capabilities. These include novel user interfaces, new interaction mechanisms, and supportive capabilities that help give users more flexibility in generating, distilling, and organizing information.

Thus when the emphasis lies on the computer media and technologies of online knowledge building communities, I also use the term *Community Information Repositories* to refer to them. None of the two terms perfectly describes the knowledge communities. Often, an existing community uses a collaborative technology to complement its operations; for example, the use of an internal wiki for the members to share ideas in an organization. In this case the aspect of the community regarding its membership, visibility, and rich interactions among the members clearly matter much, and the technology aspect is only part of the entire concern. On the other hand, an entirely online community whose main purpose is sharing information and knowledge may not have as much emphasis on the interactions among the users. However, calling these communities a community information repository might sound as if they are just a place where a community's information accumulates with no regard to the interactions among the users. Clearly this is not the case with many online knowledge building communities (Ridings 2004). While the degree of a sense of community users feel may vary (e.g., in our study of an online question-answering community, Naver Knowledge-iN, users stated they felt little sense of community within Knowledge-iN, but stronger sense of community in smaller and specialized cafés), it is not just about information exchange or accumulation. Virtual discussions entirely devoted to sharing news such as Slashdot (http://www.slashdot.org) have been shown to exhibit an emerging social network-like structure from the user activities (Gómez et al. 2008).

Thus, I use both online knowledge building communities and community information repositories interchangeably in this dissertation, using the former when the community aspect is more emphasized and the latter when the computer media or technologies are discussed.

Outline

The literature review consists of the following. First, I briefly describe four representative community information repositories, which are mailing lists and Usenet, online discussion forums, wikis, and large online question-answering communities. Then I discuss several characteristics of the communities as well as the challenges they face. The discussion about past approaches to the challenges follows next. I conclude the literature review with a discussion of the relevant literature that helped shape the design choices for Arkose.

Types of Community Information Repositories

While there may be other forms of online knowledge communities in existence (for one, the specialized and closely gated *cafés* as a knowledge sharing venue, discussed in our study of Naver Knowledge-iN (Chapter 4)) the following four online community repositories are by far the most representative and widely used. There may also be different shapes and forms of these communities, but I believe my discussion sufficiently covers important points that are common to them.

Mailing Lists and Usenet

A mailing list is a collection of email addresses used to send information to multiple users (Wikipedia 2009c). Similarly, a newsgroup is a repository for messages from members of a Usenet system that distributes asynchronous, threaded conversations. People started using mailing lists and newsgroups soon after the beginning of the Internet in the 1970s, especially in the scientific community (Zakon 2010). They are also called a discussion list, because it is often used to discuss a solution to a problem. For example, remotely located organizations have used emails and message archives to find and share information among the employees (Finholt 1993). The functionality of newsgroups is similar to web based discussion forums, but the underlying technology is different. While a web based discussion forum can be read through a web browser, newsreader software (often integrated with an email client, e.g., Mozilla Thunderbird) is used to access newsgroup data (Wikipedia 2009e). Normally, a newsgroup is devoted to the discussion of one topic. For example, *alt.os.linux* is one of many news groups that discuss the Linux operating systems.

Mailing lists have been popular as a means for users to exchange information and help one another to solve problems that arise in a domain or around a product. For example, beginning users of a product or new comers in a domain may ask questions to learn their way around them, while more experienced users may provide an answer or search for an existing solution. Hundreds of thousands of mailing lists have been in use, covering many areas such as health, technology, education, law, and so forth.

While mailing lists often become a good source of information for users, they do have shortcomings. Because a message is often sent to every subscriber in the list, a user may be overwhelmed by messages that are not necessarily interesting to him. A popular mailing list can have hundreds of messages daily, making it difficult for a user to keep up with the ongoing events. Valuable knowledge and information, in an archived mailing list are often buried among other non-essential information such as off-topic conversations, casual conversations, and flame wars. It can make it very hard to find relevant information relevant to a user's current situation.

Discussion Forums

An online forum is also called an Internet forum, web based discussion forum, or message board, similar to bulletin board systems in the past (Wikipedia 2009a). Currently, the most common way to access information in a discussion forum is through a Web browser.

Many online discussion communities have been created in order to support their members by allowing them to share information on a particular topic and help one another. Online forums are similar to a mailing list in terms of the purpose of the site, but they are also different because a user can visit the forum and browse selective information rather than receive all the members' activities as in a mailing list.

Online support forums cover a wide range of topics. Some popular topics include technical forums that deal with programming languages, operating systems and software products, and health support forums that discuss diseases, symptoms, medications, and help patients with a similar disease or problem. Often an online support forum has a number of experts in the community that provide much of the answers and help to other members. Members can get information provided by others (Preece et al. 2004) or also share their personal experiences and help forming an opinion and a decision. For example, discussion forums have been studied as a useful tool for collaborative learning (Stahl 1999) and used as a valuable supplemental tool for users to discuss products and services in commercial sites.

General Purpose Question-Answering Communities

Recently, large online question answering communities have become popular places to share and get information (Nam et al. 2009; Adamic et al. 2008). The main purpose of these sites is to provide help to their members to ask and answer questions among one another under a variety of topics.

These ultra large question-answering communities differ from online discussion forum in several ways. While the idea of a question-answering community is simple, the power comes from the vast amount of information accumulated within the system. The amount of the questions and answers a community gets each day are much larger than a discussion forum. For example, Naver Knowledge-iN estimated in 2007 that it was getting 110,000 answers to 44,000 questions every day. The variety of the topics covered by the community is much larger as well. Whereas most online forums or mailing lists individually focus on one or few specific domain areas (such as "programming" or "diabetes"), general purpose online question-answering communities allow users to ask and answer questions over a much larger range of topics. For example, Yahoo! Answers and Naver Knowledge-iN have more than two or three thousand categories and sub-categories for question topics, and the number of questions and answers is tens of millions and is still growing. Another difference from an online discussion forum is the nature of the "conversation" afforded by the user interfaces. In an online discussion forum, user interactions are indeed a "discussion" in which users talk freely among themselves continuing a conversation in a threaded manner. In online question-answering communities, the user interactions are question asking-answering activities without much back and forth between the asker and answerers.

Wikis

Wikis have gained popularity as a new information creation and repository system. First wikis were designed and developed by in the mid 1990s (Leuf and Cunningham 2001). A noticeable difference of wikis from other online community repositories is that the content is readily editable by any user thus any one can potentially be an editor or author. Whereas user interactions occur through question asking and replying and discussing activities in mailing lists or discussion forum, users in a wiki collaboratively add topical contents to the space and edit them as needed. Thus, there is little interaction among the users in wikis, though a discussion forum can be supplemented to support a backchannel conversation for discussing topics in conflict. Wikis work well in organizing information under a topic.

Wikis are distinct from other community information repositories in the authorship and version control (Emigh and Herring 2005). Any user may potentially contribute to the site. To add content to a wiki page, a user simply goes into the editing mode provided by the wiki system, add new content or modify existing one and the changes are saved by the system. The changes made are usually tracked by the system in a log of users' activity details such as the IP address of the user, user ID, changed content, time of change, and so forth. This is to control the potentially disruptive behavior of vandalizing wiki pages. Although potentially disruptive "holy wars" may occur in wiki pages (especially on topics such as politics, religion, products, and so forth), the editing interfaces of a wiki support a "higher level of consensus building" (Viégas et al. 2004) than a discussion forum, because contents may be deleted by those who disagree. One other difference from a discussion forum is that many pages are interconnected with hypertext links.

Wikis are a "perplexing phenomenon" (Forte and Bruckman 2006) in that the content is created and edited collaboratively, and it "self corrects" itself through the community's monitoring and restoring. For example, one of the most successful wikis is Wikipedia (http://www.wikipedia.com) that is an open encyclopedia collaboratively created by tens of thousands of users. Even though any user can modify the content without even logging into the system, the information in Wikipedia has been found just as reliable compared to traditional encyclopedias (Giles 2005) and often been cited in the mainstream press (Lih 2004). Viégas et al. (2004) found in their analysis of vandalism in Wikipedia that most disruptions occurred in May 2003 were fixed within a few minutes by the community members.

A wiki has been found to be an ideal repository for summarizing and collecting information by topic (Hansen 2007). Wikis' value as an information management tool (e.g., Nakanishi et al. 2009; Hester 2010) and educational tool (Forte and Bruckman 2006) has also been examined and found to be successful.

Characteristics of Community Information Repositories

In the previous sections, I briefly described the background of some of the most popular community information repositories. In the following sections, I examine characteristics of the repositories - some common and some unique to each type.

Information Interfaces and Information Types

Most of the community information repositories have simple user interfaces. Mailing lists and newsgroups are often viewed with a newsreader client (e.g., Microsoft Outlook Express), and discussion forums and wikis are primarily used through a Web browser. While it is conceivable that other modes of communication may be implemented (e.g., voice based communication), the majority of information activities occur through textual user interfaces.

The interactions in most community repositories are simple. Users in a discussion forum may post a question (e.g., MedHelp[3]), or initiate a discussion topic (e.g., Slashdot[4]), and reply to previous posts in order to provide answers or conduct a discussion. The majority of conversations are often *threaded*. Similarly in a general purpose question-answering community, the two major activities are question posting and answering providing. A difference is that the posting activities are more limited because the asker cannot reply to his own question and a user may post only one answer. This is partly due to the incentive system that rewards users based on the best answers chosen. Modifications to the simplistic user interactions have been added to many of the sites in

[3] http://www.medhelp.org/forums/list
[4] http://slashdot.org/

order to allow better communication among the users. (e.g., Naver Knowledge-iN where users can comment outside of the answers, and similarly in StackOverflow[5].)

The major activities on a wiki are users' addition, deletion, and modification of contents. As discussed in the earlier section, any user can freely change existing contents thus users in a wiki collaborative build and shape information on a topic. So there is little explicit interaction among the users, but a log is kept to track all the changes and users in some wiki systems can conduct a backchannel conversation to discuss topics.

Thus the affordances of the user interfaces in each community information repository partly shape the types of interactions users have and the kinds of information that accumulates in the community. This has implications for what information needs can be met in each community, which is discussed in the next section.

Meeting Users' Information Needs

The simple nature of the user interfaces in many community information repositories lowers the barrier for new users to start conversing and become active in a community quickly. The generation of information from users' activities is fairly fast, which makes these information repositories ideal places for asking and searching for information. However, the user interfaces are also constraining in some ways.

For example, current large question-answering communities meet their user' information needs by allowing other users to provide answers to their questions quickly, and thus the systems are very responsive to individual user's information needs. Since an answerer must interpret the asker's specific situation and needs, answers are often highly contextualized to be helpful for the user. However, question-answering sites such as

[5] http://stackoverflow.com/

Yahoo! Answers and Naver Knowledge-iN make it difficult for the asker and answerers to interact, since a user can leave only one post. This makes an answer in general a one-shot attempt and is sometimes short and incomplete.

In a discussion forum, the asker and answerers can interact with one another and keep reshaping the question and answer better, the serial nature of the forum interfaces does not explicitly support this feedback loop. Any clarification of the question is often intermixed with the actual answers. There are other problems as well. A preliminary informal study of several technical forums I found that most questions answered in a general question-answering community were relatively simple and easy, and the quality of information was generally low. Others have also found quality of many answers in a question-answering community to be low (e.g., Jeon et al. 2006; Liu et al. 2008b) This may be again due to the affordances of the interfaces that do not allow complex interactions among the users. Furthermore, because there is no facility to organize and save distilled information of the answers, information in a question-answering community is scattered all over in the system, and it is very hard to get conclusive information.

Wikis, on the other hand, strive to provide generalized information on a topic that is organized and structured by peer editors. As discussed, the quality of information can be high on a popular wiki site such as Wikipedia (Giles 2005) due to very responsive peer corrections (Viégas et al. 2004; Kittur et al. 2007) and policies that encourage citations of the source material. Wikis often provide objective and general information, which may not meet an individual user's specific and contextualized information needs. As well, the contents may not be responsive to users' emerging needs. Most "consumers" of the information in a wiki are not the creators or editors, but the readers of the articles. Thus, information needs of users may not be adequately reflected in the topics covered in a wiki or in the way an article is written.

Although various types of community information repositories have been invaluable resources for information generation, search, and reuse, each of the repositories has strengths and weaknesses in some ways. In this dissertation I focus on designing new user interfaces and interactions mechanisms that address some of these issues, as well as providing supportive capabilities that these community information repositories lack.

User Participation and Contribution

User participation behavior and contribution patterns have gained great attention and research interest, because a knowledge building community's success often lays on its users' continual participation and contribution. A community needs a constant influx of new users for it to sustain its usefulness. While a few knowledge building communities are monetary based (e.g., Google Answers[6]) the work is done by users' voluntary participation in the majority of knowledge building communities. While Arkose2 is policy and incentive system agnostic, generation and distillation of information is done by users' voluntary work. Thus understanding participant and contribution patterns in a community will be important concerns when it is deployed in an organization or community.

The patterns of participation for online communities in general have been studied in a range of studies. For example, Wenger (1998) discussed the importance of different

[6] http://answers.google.com/

roles in online communities and how they affect community formation and continuation. Preece et al. (2004) studied lurker behavior in different online forums. In a given community only a small proportion of the users actually post something, and the majority of the users come to the community but just read and search for answers. As Preece et al. found, these lurkers were not selfish free-riders; valid reasons existed that kept them from participating, such as unfamiliarity with the community and difficulty using the interfaces. Lurkers are potentially contributing users. By lowering some of the barriers in the design of a community, it might get more participation.

Substantially fewer studies of general purpose question-answering communities exist. In general, Welser et al. (2007) found that there are certain users who handle a disproportionate share of answering. Zhang et al. (2007a) examined the Java Forum community and found clusters of users who ask, users who answer, and users who do both. Adamic et al. (2008) examined Yahoo! Answers, however, found far more separation of answerers and askers. These answerers are critical to the success of Q&A sites, since they provide the answers that draw askers. Yet finding enough answerers, especially for ultra-large sites and especially if answerers are a distinct set of users, is likely to be an interesting challenge.

What motivates users to participate in the knowledge building communities? Prior work has uncovered many of the motivations and incentives for answering questions in online communities in general. Yu et al. (2007) discuss motivations such as active learning, self-enhancement, reciprocity, reputation, enjoyment of helping others, self-protection, moral obligation and the advancement of the virtual community. Within an organizational context, Constant et al. (1996) found altruism to be a strong motivation for answering questions, while strong social ties were not. Butler et al. (2007) in an analysis of listservers, found that users participated to obtain otherwise inaccessible information and visibility in social relationships.

Most question-answering communities implement a variety of incentive systems to encourage users to ask and answer questions. Now defunct Google Answers (http://answers.google.com) was an online knowledge market that let users post monetary rewards from $2 to $200 for answering a question (Wikipedia 2009b). Yahoo! Answers and Naver Knowledge-iN have a point system where answerers earn user posted points when answering questions and providing a best answer. A user's rank in the community is determined by the number of points the user has, which is clearly visible especially for a few top users in the site.

And different incentives and openness of participation also appear to affect the quality of the answers and user behavior. Harper et al. (2008) found by comparing different Q&A systems that rewarding answerers with money and increasing the rewards induced higher quality answers. Yang et al. (2008a) found that offering more money for a solution to a task correlated strongly with attracting more views for the task, and correlated weakly with increased task participation.

Challenges

In the previous section, I discussed some of the characteristics of community information repositories. These characteristics influence that shape of the user interactions, type and quality of knowledge accumulated, and users' information searching behavior. The characteristics of the communities also present challenges in some cases. In this section, I examine some of the challenges these knowledge building communities face.

One of the biggest problems of popular knowledge building communities is that they are too popular. The huge amount of the influx of information into community repositories can be problematic. Large newsgroups receive many new messages every day, but many of them are of little value (Fiore et al. 2002) and the ratio of useful information to noise is too low (Resnick et al. 1994). While newsgroups usually focus on one specific topic and often refactor a topic into more topics when it becomes too large (Resnick et al. 1994) the amount of knowledge constantly accumulating in the system may become daunting. The low quality of answers accumulating in general purpose online question-answering communities is evident. For example, Jeon et al. (2006) find in their quality analysis of 1,700 question answer pairs, about 1/3 of the answers have quality problems and 1/10 of the answers are bad. Similarly, Liu et al. (2008b) find the not all user selected best answers are indeed of high quality. The low quality of information coupled with the large amount of information may be quite problematic. Having to wade through messages and filter out unnecessary information might become too problematic for users to continue to use a system.

One the other hand, the nature of unstructured conversations in a mailing list or online forums makes it difficult to extract useful information. Many email lists have web interfaces, which make it easier to search for information and browse the history of discussion. While discussions in an online forum are more structured and threaded, the linear presentation of a threaded conversation still makes it hard to location useful information. Moreover, the current user interfaces of online community repositories make it difficult to present other inherently important data about discussion threads. Not only is the content of the discussion data difficult to extract (Brewer 2000), but meta-information such as user interaction data and a user's topical interest (Fiore et al. 2002), and temporal and logical sequences of the flow of messages (Smith and Fiore 2001) are not readily available in the user interfaces.

One other challenge is the location of expertise. Experts in a community not only provide valuable information and help other members, but they often do maintenance work such as editing, moderating and correcting of information. Thus being able to locate experts in a community may help with resource allocation (e.g., experts might answer difficult questions while novices easier ones) and quality control of the information.

Approaches for improvements

In the previous sections, I discussed the characteristics of community information repositories and their challenges. Here I discuss past approaches for improvements in several areas.

Reuse and Distillation

Current online discussion forums and general purpose question-answering communities do not provide an easy way to distill information and reuse it for future questions. Although it is possible to user the given user interfaces to create summaries (for example, write a summary post at part of the discussion discourse), it is by no means supported explicitly and it is quite difficult to summarize and reuse information.

Characteristics	Q&A / discussion forums	Wikis
Type of information generated	Relatively short factual answers, opinions, solutions to a specific problem	Summarized article on a topic
Reusability	Low	High (organized, structured, citations)
Visibility of users' information needs	High (Individual user's needs are visible)	Mid (Popular topics quickly generated)
Responsive to users' information needs	High (Individual user's needs are met)	Low (Information needs of the general public)
Meets individual's specific information needs	High	Low
Information collaboratively generated	Low	High
Finding answer to information needs	Easy (answerers "bring" information to a question)	Somewhat easy
Generalizability	Low (much is contextualized, specific, and interpreted information to meet asker's specific situation)	High (strives to be objective)
Quality of information	Low (many answers to easy questions)	High (organized, structured, citations, revised)
Organization of information	Low (scattered all over)	High (related information under the same topic, links to related topics)
Role differentiation	Yes (asker - answerer)	Not really (editors)
Effort needed in knowledge creation	Low to Mid	High

Table 2-1 Comparison among popular community information repositories

Resources in a help-providing community, whether they are users' time, attention, or cognitive capabilities, are limited. Constant influx of new users, and questions and topics in an online community repository is critical for the continuing existence of the knowledge building community. New users may feel ignored and quickly abandon a community if their questions and topics are not handled in a timely manner by the community. However, as it was discussed in earlier sections a common problem in many online community repositories is the sheer amount of information. Reusing exiting solutions and information that the community has already found might help provide a solution to duplicate questions, thus ameliorating the problem of resource allocation. For example, Liu et al. (2008b) estimated that 78% of the best answers in several online question-answering communities are reusable in answering future questions that are similar. They developed a multi-document summarization (summarization of multiple documents together into one summary) technique to summarize different types of answers in order to reuse them for later questions. Jeon et al. (2005) developed methods to retrieve semantically similar questions that may already contain answers.

As the amount of information generated through online media increases exponentially, it has become enormously hard to locate information and reuse it. Many have tried to reduce the information space by extracting and saving only the pieces with high value, reconstructing information as needed. One example is knowledge "liquidation and crystallization" by Hori et al. (Hori et al. 2004; Amitani and Hori 2003) They describe methodology and systems that deconstruct information sources into small pieces with editor provided keywords, and later reconstruct the information sources into new knowledge by combining different pieces. Based on these concepts, Amitani and Edmonds (2007) develop a method for sequencing information pieces with related context information. Other approaches include condensing the information space into efficiently searchable extracts (Brewer 2000), methods for discussion retrieval from mailing list archives (Kolla and Vechtomova 2007), more control over the recipients of the messages (Masui and Takabayashi 2003), and integrating a mailing list with a wiki (Eto et al. 2005). Others have tried to use the knowledge accumulated in a mailing list to build a question-answer system (Watanabe et al. 2004).

However, systems that allow human users to collaboratively distill a discussion space are rare. One system similar to Arkose is the WBT system (Helic et al. 2003) that allows students to summarize a discussion space in an educational setting. Two of the features of the WBT system are adapted in Arkose2. WBT allows a moderator (a teacher) to create a conceptual schema to which students can make "contributions" to the concepts in the schema. The creation of a skeletal structure and separation of the roles are useful in supporting distillation and share the work load and they are supported in Arkose2 as well. WBT has several limitations. First, only the moderator can create a schema, and this may not effectively reflect the changing topics in the discussion space on time. It may be important for anyone to create and update schemas or templates as the discussion progresses, because distillation of a large discussion space may require many users working collaboratively.

Second, when a student in WBT wants to make a contribution to a concept, he has to manually scroll down and search for the entry for the concept in the schema tree. This may hinder students' distillation process, especially when the schema is large and they have to scroll through many entries. Grouping related posts and logically linking together schema would eliminate manual searching.

WBT also lacks other features that might help distillation, such as visual aids, information retrieval capabilities, and organizational tools.

In summary, the focus of my thesis largely lies on computer-mediated places as well as knowledge repositories, and how to improve knowledge generation and distillation over what was available in previous community information repositories. Different mechanisms and technical possibilities support different modes of information generation, sharing, and reuse. As there is a gap between the social requirements of collaborative knowledge management and available technical capabilities (Ackerman and

Halverson 2000), none of these community information repositories is perfect, but each has strengths and weaknesses.

Spatial Hypertext Systems

Hypertext systems allow related documents to be interconnected and navigated through the links. This helps logically divide content into structured sub-sections as well as associate with external information related to the content. However, the document centered view of hypertext systems led to a navigation problem that could confuse users, especially when the users were not familiar with the material (Bernstein et al. 1991). Systems such as NoteCards (Halasz et al. 1987) and StorySpace (Joyce 1991) tried to ameliorate the problem by displaying navigational maps of the hypertext documents. These systems allowed users to browse the documents through a network map rather than a document view. Other systems such as gIBIS (Conklin and Begeman 1988) and Aquanet (Marshall et al. 1991) further allowed users to formalize the information and relationships with typed nodes and links.

The map-centered view of documents allowed the relationships among documents to be represented in multiple ways. Not only could the links in the map show an explicit relationship between two connected nodes, other properties such as the relative location, shape, size and color of the information objects could be used to assign a meaning (Shipman 1999). Users preferred implicit expressions of relationships through these properties over explicit specifications through formalization (Marshall and Rogers 1992), and spatial hypertext systems such as VIKI (Marshall et al. 1994) was developed to better support these interactions.

The map view of documents and use of implicit properties of data objects in hypertext systems have influenced the design of data presentation and manipulation in Arkose2. The current thread view of an online discussion forum makes it difficult for users to understand the overall shape of the discussion space and relationships among different posts and subtopics. It is important in distillation of a discussion space that users easily understand these properties of the space, because they may not be casually reading through posts from top to bottom, but may be actively moving around in the discussion space searching for and manipulating information. These tasks require support for quick access and flexible navigation that previous hypertext systems tried to provide. For example, it may be useful to visualize discussion posts with graph nodes and links, forming a tree structure representing a discussion discourse. This could help users easily figure out where an important topic might be located and how different topics might be connected, and understand the progress of summarization. It might also allow users to quickly move to a location of interest. Visual hypertext systems inspired several features of Arkose2, including the map view. As well, Arkose 2 uses properties of data objects to assign meanings to them as in many hypertext systems.

While the map view of documents in hypertext systems is an improvement over the document view, these previous hypertext systems lack some features helpful in supporting distillation of a discussion space. In a map view, each post is represented as a small graphical node and reading the details of individual posts is difficult especially if the content is large. While earlier hypertext systems had ways to see the entire content such as viewing a data object in a special viewer or maximizing the object, these may not be sufficient in a distillation task where users may need to quickly read through several related posts at a time. A thread view would help by showing multiple posts at a time and by allowing users to attach posts in the thread view to summaries for later retrieval. Features such as the thread view would help support both understanding the overall shape of the discussion space and easily accessing detailed content. These hypertext systems also lacked content analysis and information retrieval techniques, making it potentially overwhelming for a user to process and keep track of a large amount of information at a time. The map view in hypertext systems helps users understand the information space, but it can be further supported by automatic capabilities that locate related information.

Discussion Support

Collaborative work and discussions are inherently social processes. In conducting a conversation of discussion with others, three things might help. The first is being aware of what is going with others around the user, and the second is understanding how the discourse is being shaped.

First, in interacting with other users, Ackerman and Starr (1995) discuss social indicators as important for collaborative systems because other people's activities help frame users' own "goals, motivations, and actions." They argue there are several social indicators that could be useful and the capability for them should be reflected in design of collaborative applications. Similarly, Erickson et al. (1999) argue "Socially Translucent Systems" will help users create "coherent discussions, observe and imitate other's actions, engage in peer pressure, create, notice and conform to social conventions. " A few prototype systems have been developed for these purposes, including Babble (Erickson et al. 1999) and Chat Circles (Donath et al. 1999) that use visual techniques to show social activities of the users in the conversation And other systems that keep track of user contributions in a discussion using machine learning techniques (McLaren et al. 2007; Groot et al. 2007).

Awareness of others' activities could also be achieved through the history of the objects in a shared space as well. Hill et al. (1992) argue that digital information accumulates "computational wear" by the history of interactions among the author and readers. Similarly, annotations left by users on an information object (Goldberg et al. 1992) help users understand what others have previously done or understood. Similarly in Arkose2, users can leave various meta-information in the discussion space and summary space. This includes questions, comments, and notes about the discussion and summary, which might guide users in further discussing on a topic or distilling a discussion space.

Second, allowing users to understand how a discussion is being shaped may also be useful. For example, Wegerif et al. (2009) used a graph-matching algorithm to recognize different types of discourse in a discussion and Ito et al. (2007) developed a system that presented relevant keywords and other statistics during a discussion to help the participants with different backgrounds. Yet others used visual representations of the concepts in a discussion to support collaborative knowledge building. Suthers et al. (2008) examined the following hypothesis among others:

> *H1: Collaborative knowledge construction is more effectively supported by environments that make conceptual objects and relations explicit.*

They found strong evidence supporting the hypothesis in their experiment where a graph representation of users' knowledge construction was found to be more effective than a textual representation, and the graphical representation led to more knowledge construction and more convergence of ideas. This has significance for Arkose2 in two ways. Their findings may indicate that the graph representation of the Visual Navigator may support users' discussion discourse better by showing them the shape of the discussion and relationship among the posts. This is partly supported by participants' statements in the evaluation of Arkose2, where they found the Visual Navigator more useful in understanding the space when there were many discussion posts, and used the Visual Navigator to find potential subtopics and important posts. Similarly, by allowing users to explicitly specifying structure and relationship among the summary nodes, summary creation may be more effective.

Previous approaches do not provide a support for handling complex questions or more flexible discussion discourse as Arkose2 does. Arkose2 allows users to have a backchannel conversation, break a topic into subtopics, and distill discussion space as the conversation is in progress. These should help users conduct a more flexible and guided discussion.

Navigating and Searching

As discussed, the amount of information in a community repository and locating useful information present a challenge. To ameliorate this problem, several newsgroup readers have been augmented with text analysis capability to support navigation and search in the newsgroup repositories. These include automatic retrieval of related posts through user feedback (Ozaku et al. 2000), and identification of topic changes in threads (Uchimoto et al. 1997).

Some have used visualization techniques to show various conversation activities and user interactions to help understand the shape of a discourse. These include Netscan (Smith 1997), Loom (Donath et al. 1999), and Discourse Diagrams (Sack 2000b). Others have combined textual analysis and social relations among the users. (e.g., Sack 2000a; Smith and Fiore 2001 and Sack 2000b.)

Presenting discussion related data in multiple ways in a single view has been a popular approach among information interfaces designers. There is evidence that representing a discussion in a tree structure is useful for reading news (Smith and Fiore 2001) and the ability to tell the location and size of branches is helpful. Arkose2's Visual Navigator also allows users to take advantage of the tree representation of a discussion discourse. As found in the evaluation of Arkose2, this feature was useful for some participants to identify important subtopics and groups of posts discussing the same topic.

Arkose2 helps navigating and searching by having users connect relevant information. This includes the linking generalized information and contextualized information, discussion topics with wikis, and topics and subtopics in order to allow users to better navigate the information space.

Filtering out Noise

In the previous section, various mechanisms tried to support navigation and search in an information space. Some have implemented various filtering mechanisms to weed out irrelevant information. There are largely two types of filtering; content-based filtering and collaborative filtering. Content-based filtering uses natural language processing techniques or information retrieval techniques to select certain information and discard the rest. Examples of the techniques include calculating similarity values among the terms or documents such as Latent Semantic Indexing (Deerwester et al. 1990) and term weighing in document vectors (e.g., Salton et al. 1983; Salton and Buckley 1988; Foltz and Dumais 1992) as well as using machine learning algorithms such as the Bayesian analysis (Belkin and Croft 1992). Other techniques such as automatic clustering of documents (e.g., Kim and Lee 2000; Avrachenkov et al. 2008) and a software agent to filter personalized information (Albayrak et al. 2005) have been used.

On the other hand, collaborative filtering is achieved through users' activities of commenting, voting, using, or otherwise rating information objects or matching information with authors with similar interest and information needs (the matching system is often called a recommender system). This takes advantage of users' knowledge, expertise and interpretation of information that are otherwise not possible with automatic techniques. For example, the importance or usefulness of information content may be annotated by the users who used it (Goldberg et al. 1992). Maltz' collaborative filtering system for Usenet (Maltz and Ehrlich 1995) allowed users to send other users "pointers" to the documents they found interesting. More recently, social tagging services (e.g., Delicious (http://delicious.com) utilize user provided tags to categorize documents. Some have used the social tags in recommending documents (e.g., Guan et al. 2010)).

However, there has been little work in filtering out noise and extract only important information in a discussion space. Ideas are often buried deep within other non-essential information, and a conclusion to a discussion is sometime left to the reader to determine. Arkose2 supports distillation of a discussion space in order to reuse information from a discussion space. In order to assist users' distillation tasks, Arkose2 uses both a content-based and collaborative methods. The underlying automatic clustering of posts (see Figure 5-16) clusters related posts in a discussion space. Arkose2 also allows users to "vote" for the usefulness or interestingness of posts and summaries when they distill and save them for their own use. This may help other users to socially navigate (Dourish and Chalmers 1994) the information space by filtering out unpopular information and following what others have found useful and informative.

Locating Experts

Finding expertise is both difficult and important aspect in a knowledge building community. It is important to be able to locate authors that contribute informative and useful information in the knowledge community. Experts not only provide useful information, they do other maintenance for the continuation of the community (Hansen 2007).

Earlier expertise finders include Expertise Recommender (McDonald and Ackerman 2000) as well as studies such as Ackerman et al. (2002) and Ackerman and Halverson (2004). Zhang et al. (2007b) also developed an expertise finding tool. Similarly, others have used content based analysis in order to find experts. For example, Streeter and Lochbaum (1988) used Latent Semantic Indexing (Deerwester et al. 1990) techniques to analyze project report to find expertise.

Recently, expertise location in social networks has gained much interest, and is considered "the killer app for enterprise social computing" (Lupfer 2010) by some. Examples of expertise finding in social networks include (Yu and Singh 2003) and (Zhang et al. 2007a) that devised expertise finding algorithms that utilized properties of networks. Similarly, Yang et al. (2008b) developed an expertise finder that used the link patterns of Wikipedia.

While Arkose2 does not have expertise finding capabilities, a future implementation may benefit from being able to locate experts for a given topic. Since expertise is a relative concept (Ackerman et al. 2002), finding experts and their relative expertise level presents an opportunity for more efficient handling of users' questions, as examined in Zhang et al. (2007b). This will also allow Arkose2 to more effectively ask users to perform tasks such as distilling a discussion space on a certain topic, organizing different types of information, or answering a question.

Improving Participation

While some online communities get enough activity, not all communities are as successful. Lack of participation reduces the usefulness of the community, and users may eventually abandon it.

There are several ways to motivate users to participate or put in effort. An explicit incentive system such as a point system has been popular in general purpose question-answering communities. In these communities, users are rewarded for performing specific tasks that require users' effort or participation such as answering a question and being selected as the best answer provider, or logging into the system. (Adamic et al. 2008; Nam et al. 2009)

One way a collaborative system may use to encourage user's voluntary effort is to align user work and benefits. The alignment between effort and benefit becomes an important factor in the adaption of the system and garnering users' voluntary work. Grudin argues that many collaborative systems have failed because there is disparity between users who do the work and users who get benefit (the famous “who does the work and who gets the benefit?” question) (Grudin 1987; Grudin 1988). In some collaborative systems, users’ voluntary work towards their own benefits and usage becomes useful information others in the community can use. For example, in the GroupLens system, a user is encouraged to rate information truthfully, because otherwise the system’s usefulness is affected for that user (Resnick et al. 1994).

Arkose 1.0 and Arkose2 are incentive system agnostic, and may be combined with an incentive system that works for a particular community in a field usage. However, Arkose2 uses a similar approach with GroupLens or a social tagging system in that users’ summarization and tagging of posts created for their own use collectively become guidance data for other users. The usage data also allows Arkose2 to select users to perform particular tasks. Users who saved certain information may be interested in knowing more about the topic and may be more likely to do maintenance work on it such as distilling and organizing. As well, Arkose2 allows users to leave explicit requests for task such as a request for further discussion or provision of evidence for an argument. These together should encourage users to put in work and benefit the community as a whole.

Design of Arkose

Arkose 1.0 and Arkose2 build on previous streams of HCI and CSCW research. The first is the design rationale, an area of considerable interest in the late 1980s and early 1990s. The hope was to reuse design and decision understanding. The design rationale research stream (e.g., Moran and Carroll 1996; Buckingham Shum 1996) examined both languages and representations (e.g., Lee 1990; MacLean et al. 1990) as well as interfaces (e.g., Conklin 1992) for supporting design history and explication.

These previous approaches provided well-defined structures for a discussion discourse, but the imposed formal structure limits how participants can discuss topics. Most online forums do not place restrictions on how conversations should progress. A thread on an online discussion forum is initiated as needed, and it often forks into other related (or sometimes unrelated) topics. The free nature of online discussions allows a naturalness of interaction, and is likely to be one of the reasons online forums have flourished. Thus, imposing formal rules and structure in discussion raises barriers for participants. Farnham et al. (2000) supports this point in their findings that an explicit structure imposed in conversation is interruptive and restricting.

The issue of imposing formal structure is exacerbated when the participants are required to make an upfront decision about what the information structure should be. Incremental formalization (Shipman and McCall 1994; Shipman and Marshall 1999) as an approach allows information to be gradually formalized over time. (The authors define formalization as "the process of identifying machine-processable aspects of information", but in general it includes allowing intermediate levels of structure and use of formal representations.) This reduces users' pressure to initially commit to a specific format and organization. Building on this idea, Arkose 1.0's approach in distillation of a large informal information space allows a gradual increment in the organization and reusability of an online community's information. Arkose2 further improves over this by allowing distillation to occur as a discussion is in progress. I will discuss how editors of a discussion space can do their work incrementally and collaboratively in the usage scenario of Arkoses. This reduces the overhead required in creating a permanent design. By allowing the overall distillation structure to be built gradually and flexibly, users may put in as much effort they can afford at any time.

Conclusion

Many virtual knowledge production systems have provided ways to share knowledge and reuse them. In the previous sections, I have discussed what these systems are and some of their strengths and weaknesses. Arkose focuses on providing supportive tools and mechanisms to flexibly generate useful information through users' question-answering and discussion activities as well as distilling and saving activities that some of these other systems lacked. It also builds on the ideas explored and lessons learned in previous online help-providing systems such as Answer Garden (Ackerman and Malone 1990 Ackerman 1994), the Zephyr Help Instance (Ackerman and Palen 1996), and Answer Garden 2 (Ackerman and McDonald 1996) as well as our study of Naver Knowledge-iN (Nam et al. 2009).

While previous systems mostly handled either generalized (e.g. wikis) or situational (e.g. Yahoo! Answers) information, my work seeks synergetic values out of handling both types of information when necessary in an attempt to meet the information needs of a community as well as individuals. Arkose2 enables synergetic efforts between knowledge generation and ongoing distillation, building on the *incremental diagesis* principle from Arkose 1.0.

Chapter 3

Arkose 1.0[7]

Introduction

In this chapter, I discuss the system augmented approach to distillation with the first generation of Arkose, which provides a set of augmentative tools to facilitate the filtering, structuring, and organizing. Online discussions such as a large-scale community brainstorming often end up with an unorganized bramble of ideas and topics that are difficult to reuse. A process of *distillation* is needed to boil down a large information space into information that is concise and organized. I take a system-augmented approach to the problem by creating a set of tools with which human editors can collaboratively distill a large amount of informal information.

Two design principles, which I define as incremental diagenesis and incremental summarization, help editors flexibly distill the informal information. Arkose 1.0 was built as a demonstration of these principles, providing the necessary tools for distillation. These tools included a number of visualization and information retrieval mechanisms, as well as an authoring tool and a navigator for the information space. They would support a gradual increase in the order and reusability of the information space and allow various levels of intermediate states of a distillation. Arkose 1.0's visualizations would allow

[7] Arkose 1.0 was published in 2007. Nam, Kevin K. and Mark S. Ackerman (2007). Arkose: reusing informal information from online discussions. *In Proceedings of the 2007 international ACM conference on Supporting group work*, Sanibel Island, Florida, USA, 137-146

editors to quickly understand the discussion space, as well as function as a substrate for gradually transforming a bramble of nodes into more concise and organized summaries, a process I call incremental diagenesis. The provided authoring tool would permit easy creation and modification of the summaries, and allow incremental summarization, a process in which summaries would be incrementally constructed and distilled. Arkose 1.0 was augmented with information retrieval mechanisms and visual aids to help editors quickly identify important topics and relations among posts, authors of those posts, and summaries. While the selected scenario for the system was an online discussion, Arkose 1.0 would potentially be useful at distilling other types of informal information that have similar problems.

The chapter proceeds as follows. First, I discuss the problems of a typical discussion forum and the need for distillation in more detail. This is followed by a distillation scenario using our Arkose 1.0 system to present some of the main features provided. I then discuss the design principles upon which Arkose 1.0 was built and end with the technical details of the system.

The Need for Distillation

As access to the Internet becomes more ubiquitous and the infrastructure for publishing and discussing people's ideas proliferates, it has become common to hold a large group discussion online. Most online forums do not place restrictions on how conversations should progress. A thread on an online discussion forum is initiated as needed, and it often forks into other related (or sometimes unrelated) topics. The free nature of online discussions allows a naturalness of interaction, and is likely to be one of the reasons online forums have flourished. While the free nature of most online

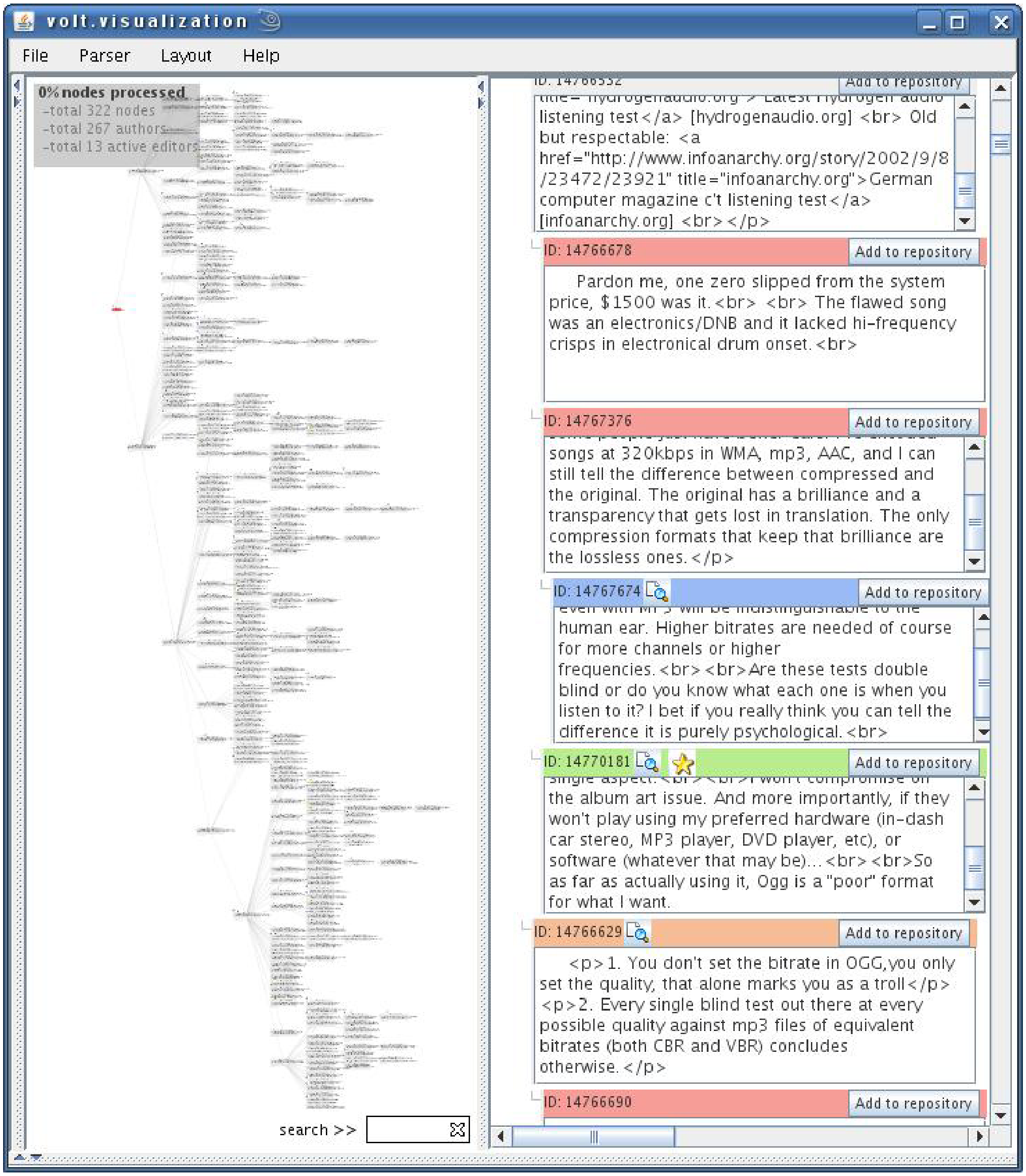

Figure 3-1 An overview of the navigator. The left column graphically presents the discussion space in a tree structure, which also functions as a substrate for distillation. The tree is fully zoomable and draggable with an online search capability for easy navigation. The right column displays each post in a more web forum like interface for maximum readability.

discussion forums may have contributed to their popularity, at the same time, it has been one of the reasons a typical discussion space is left unorganized and unstructured.

Consequently, there is a greater need for an explicit distillation process during and/or after the discussion. To provide distillation facilities, several problems must be addressed and support for their solution or amelioration is required. They include:

1) As previously mentioned, discussion spaces in general are hard to comprehend. This is especially true after the discussion space has grown considerably with dozens of topics and hundreds or thousands of posts. Interesting topics and ideas are often scattered around and buried deep in discussions, and are not easy to locate unless the user reads a considerable amount of posts.

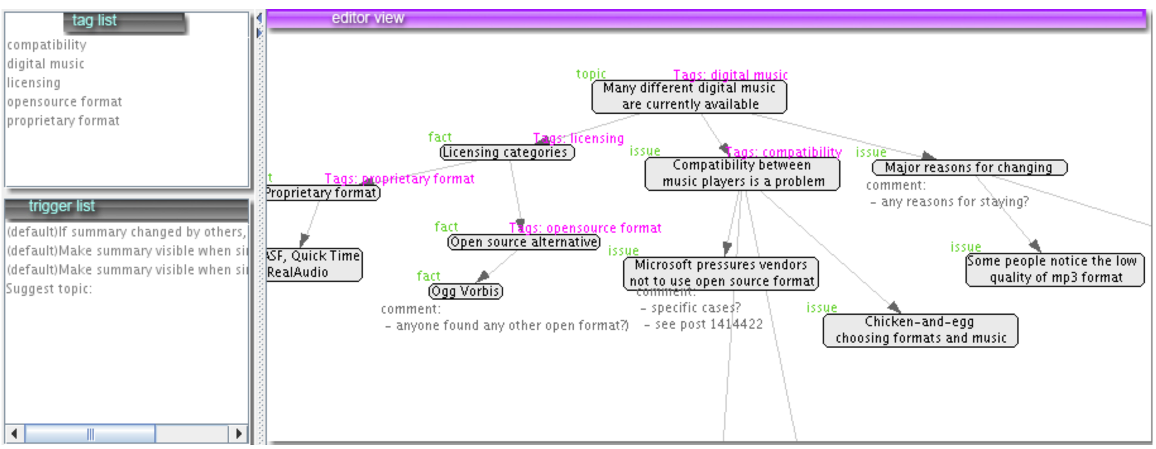

Figure 3-2 An overview of the authoring tool. An editor can create a summary and modify an existing one. Each node is associated with an editor customizable type. Notice that there are "meta-discussions" taking place (noted as "comment"). The trigger list (bottom left column) displays user specified trigger conditions, which are used to notify the editor when an associated event occurs. The graph is fully zoomable and draggable, with creation and deletion capabilities of nodes and links. An editor can also import multiple existing summaries and merge them.

2) Some topics in the discussion space are duplicates. This may or may not be deliberate; users may not realize a topic has been discussed in another part of the discussion space and start a new thread. Users may intentionally start a duplicate thread to push their ideas. In either case, duplicate discussion threads waste users' time and

effort spent in the discussion space, and they only compound the incomprehensiveness of the space discussed in 1).

3) Some topics will be socially problematic, controversial, or off-topic. These posts make it harder for users to participate effectively. Even if users adhere to a strict code of conduct, and all the discussion topics and posts are meaningful and valid, the discussion space will still need some organizing for reuse. For example, a policy discussion at the end needs to report people's points of view for policy makers. The current shape and form of discussion forums are not suitable for such a report.

My approach to the problems is through distillation. By distillation I mean the process of creating a more concise and organized form of information. It is more than just text summarization; rather, it includes sorting through large corpus of text for interesting topics, pruning away redundant and off-topic discussions, identifying interesting authors, different points of view, and ultimately making the information more reusable for later purposes. I view distillation as a system augmented process, guided and directed by experienced human editors, rather than an entirely automated one.

I next turn to a usage scenario that will be used to illustrate what I believe are the necessary design principles for distillation support.

Distillation Scenario

Setting

The basic scenario for Arkose 1.0's use is that a university has recently held an online discussion forum on various topics on the future of the university. The forum was open to the various interest groups in the university to reflect different views and ideas. After two weeks of lively discussions on numerous topics, the university temporarily closes the forum and commissions four employees to distill the discussion space. Jack is assigned as one of the four editors. While residing in the same building, the editors are physically separated into their respective offices. In addition, not all of the editors can work on the report at the same time due to schedule conflicts. Any number of editors from one to four would simultaneously use Arkose to perform distillation tasks.

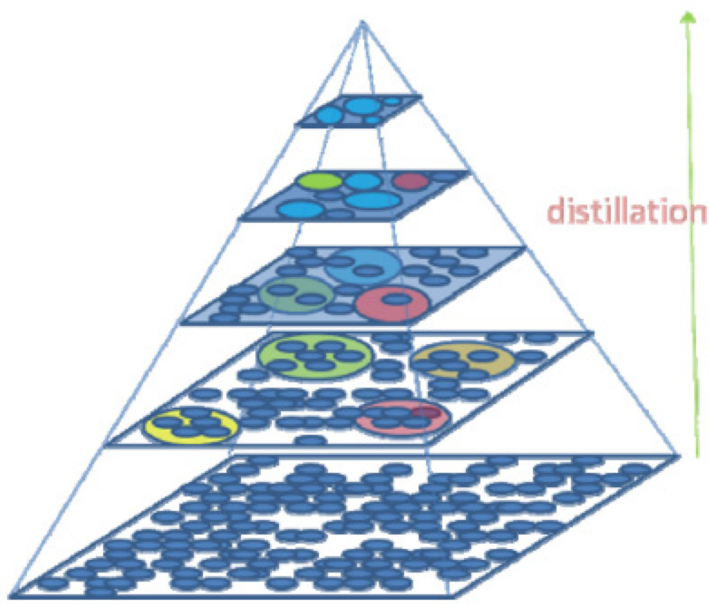

Figure 3-3 The conceptual view of Incremental Diagenesis. As distillation progresses the unorganized bramble of discussion space (depicted as the bottom plane) is gradually replaced by summaries and other meta-information. The small circles represent discussion posts while the aggregate represent some stage of distillation process.

Visually supported

Jack runs Arkose 1.0 and finds the online discussion space presented in a graphical tree form in the navigator (Figure 3-1). He sees by the instant message that other editors, Ken and Lisa, are currently online, while Matt is offline. Initially, the discussion space is completely zoomed out as an overview. Using a mouse, Jack zooms in and drags the tree around to scan through the posts. He quickly looks at a few top nodes of the tree, as it is often the case that the top nodes contain main topics. Jack can also get a sense of the topics in the space without reading the entire content of the posts; Arkose 1.0 has automatically found keywords from the discussion space and attached them to the relevant posts. Jack is distilling a specific part of the discussion space, and that is visually indicated with colored aggregates of the posts with tags (Figure 3-4). After reading a few posts in a thread, Jack finds an interesting discussion taking place. He decides to distill this part of the discussion space, and selects the posts and tags them as “Being worked on”. This creates a yellow aggregate around the posts and reduces the size of the posts to inform others of the status (Figure 3-4, part 1).

Initial summary creation

The creation of an aggregate automatically copies the posts into the authoring tool (Figure 3-2) with which Jack creates a summary structure. He creates a topic node and a few subsequent “issue” nodes and “fact” nodes, and links them together. He also assigns tags to some of the nodes in the summary structure that reflect their contents. But before being able to finish the entire summary, Jack realizes he has a meeting in five minutes and stops the process. He exports his summary and tags it as “Initial work started” and “Do not modify”. This changes the color of the aggregate to a light green (Figure 3-4, part 2) and informs the editors of the summary’s status.

Incremental summarization

After coming back from the meeting, Jack resumes the process. He imports the previous summary into Arkose 1.0's authoring tool, and continues working on the summary structure. At this point, he feels some other editor with more expertise in the domain area should review the summary he has created. He leaves a question on the node, and exports the summary as "Attention needed" (Figure 3-4, part 3). This changes the aggregate color to red and enlarges the posts to give more visibility. The requested information is appended into a message area in Arkose where everyone can read it. Another editor, Lisa, sees the request and decides to help Jack by adding some reference points to the question. Jack is then notified of this, and comes back to this summary to complete it. After Jack feels the distillation is finished, he exports it as "Closed" (Figure 3-4, part 4). This is indicated as a blue aggregate with the original posts reduced in size to indicate there is little need to reopen them (however, an editor can reopen any part of the discussion space if necessary). This replacement of the original posts with a summary node not only condenses the discussion space, but it also makes the information more organized and understandable. Jack then moves on to another part of the discussion space to continue distillation.

Author network

Jack has found an author with a very interesting point of view. First, he moderates the post up with a star symbol (Figure 3-4) to indicate the node has insightful information. He is interested in finding out more about what the author has discussed, and opens up a visual aid, "Author Network" (Figure 3-6). The Author Network visualizes the conversational activities among the authors. Jack can look at the keywords and contents of the posts between the author and others to quickly grasp the topics.

Merging summaries

Meanwhile, Ken has been working on his summary without realizing there already was a summary with similar content. Obviously, there were two discussion threads on the same topic, and the editors did not realize this because the threads were physically apart. The information retrieval mechanisms within Arkose 1.0 not only try to find duplicate posts within the discussion space, but it also compares a newly created summary with existing ones (details below). As Ken is creating the summary, Arkose 1.0 notifies him of the possible existence of a similar summary. Ken reopens the suggested summary and its original discussion thread, and finds that it would be better if the two were combined. Ken then imports both of the summaries and merges them together. Some of the parts are duplicates and are deleted. Some other nodes are linked together to form a bigger and more complete summary. Ken exports the new summary out to the navigator. Since Jack's summary has been changed by another editor, Jack is notified of this.

Keyword farm

Matt, who was initially offline, now joins the other editors. Rather than starting a new summary, he decides to work on existing ones and imports one of the summaries created by another editor but not completed yet. After some more distillation work he exports the summary with a "Closed" tag to indicate the summary is complete. The status panel of Arkose shows that 10 % of the discussion space has now been distilled. At this time, Arkose compares the keywords within the created summaries with the keywords of the discussion space it found in the beginning of the user's session. Some of the keywords that Arkose thinks are important have been used in the summaries and tags; while others have not. Arkose notifies the editors about these unused keywords in "Keyword Farm" (Figure 3-5). Keyword Farm visually presents keywords with two types of information: one is the machine calculated probable importance score of each keyword and the other is the actual usage data of each keyword in editor-created summaries and tags. Matt can easily tell the type of distillation process in which each keyword has been used. He reads the posts associated with suggested keywords in Keyword Farm, and decides a summary indeed needs to be created around some of the keywords. He then looks at the discussion threads on the keywords and proceeds with the distillation process.

Ending distillation

After one week, the distillation process has reached an end. The original discussion space in the navigator has been transformed into mostly completed summaries (represented as blue aggregates in Figure 3-4, part 4), some partially processed summaries (represented as green aggregates in Figure 3-4, part 2), and other meta-information such as editor-created tags, comments, and question-and-answers. The summaries are stored in an XML format. The collection of the editor-created summaries is presented with a style sheet format such as CSS or XSL, to form a report of the discussion space.

Reuse

The summary report is forwarded to the university's policy makers. It contains the important topics and their arguments in an orderly fashion. It also shows a list of topics that have been discussed sufficiently, as well as topics that did not generate enough or sufficient discussion. These have been identified by the editors through the distillation process. The policy makers decide to reopen the discussion forum for another few days to mainly discuss the insufficiently discussed topics. This time, the discussion space not only contains the original threads, but also attached to them are the editor-created summaries and meta-information. Returning discussion participants do not have to read the entire posts again to understand the topic; rather, they could read the attached summaries and meta-information. Authors are invited to join the open topics according to their associated keywords computed in Author Network. After a few days of discussion, the forum is again closed and distillation of the newly added information once again takes place.

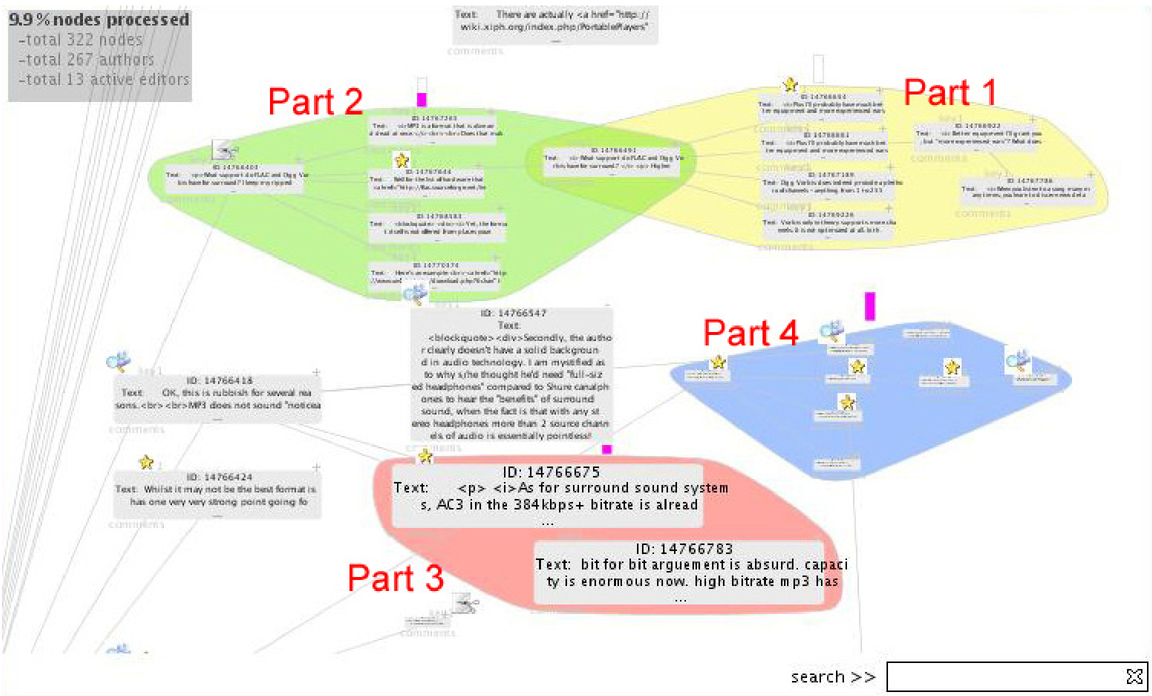

Figure 3-4 Incremental diagenesis in progress. This is a zoomed in view of the navigator (Figure 3-1). Each aggregate represents a different stage of distillation. Posts and summaries have varying visibility (represented in their sizes and colors) according to their status. Note that the red “Attention needed” aggregate has bigger sized posts while the blue “Close” (or completed) aggregate has reduced sized posts. Tags and comments can be directly left on the space. Gradually, the original discussion space is transformed into distilled summary outcomes. (The text “Part 1” through “Part 4” are for reference purposes only.)

Design Principles

In the previous section, I looked at a distillation scenario in detail. Many of the required features were enabled by two design principles upon which Arkose 1.0 was built. These principles set Arkose 1.0 apart from previous systems.

Incremental summarization

One feature lacking in earlier design rationale systems was allowing users to not specify up front what the information structure should look like. In other words, later changes to the existing structure or schema were either not supported or difficult to do in earlier systems. My approach to the problem follows from Shipman and McCall's incremental formalization (Shipman and McCall 1994). As mentioned, incremental formalization allows information piece to be gradually formalized. Thus, it adds much flexibility in creating and modifying information structure and takes a burden of initial commitment off the users.

While my main goal in distillation is not the formalization of the information (into, say, semantic web statements), I adopt the idea to support a flexible distillation process, which I call incremental summarization. Its advantages are:

Low overhead cost. The idea of incremental summarization is simple; yet, it has many ramifications in the way editors perform their tasks. An editor has the capability to modify, extend, and merge or divide existing summaries. The editor does not have to worry about the final structure of the summary, thus much less coordination with other editors is needed especially about the format of structure of the corpus. This reduces the overhead cost of initial structuring of a summary.

More thorough summaries. Topics related to each other may be discussed separately in different places in the discussion space. Since they are separated, an editor may not realize there are more discussion threads on the topic being summarized and later discover them after having finished a rudimentary summary. As part of incremental summarization, an editor can extend an existing summary with newly found topics and evidence, making the summary more complete.

Better expertise distribution. Another way incremental summarization may help is by better distributing editors' expertise. Each editor may have a different level of expertise in any given subject area, and that may affect the quality of the summaries the editor creates. An editor can tag a summary as incomplete and attach a comment that asks for help in a specific area. This information would then be listed where it is visible to everyone. The editor may, of course, directly ask a particular editor through an instant message or email if the required expertise is known. Thus, the summary can be incrementally summarized, allowing a more effective expertise distribution.

Incremental diagenesis

The second design principle comes from the fact the distillation process starts from a large unorganized information space and ends with a smaller and tighter summary space; a process I call incremental diagenesis. Diagenesis is the conversion of sediment into rock, connoting a loosely scattered large amount of information being transformed into a more concise and organized state. This is quite different from earlier systems in that most of their processes start from an empty space. In the previous distillation scenario, the discussion space has been used as a substrate for distillation. This transformation of the space is a gradual process where at any given time the space consists of heterogeneous information entities: the original posts, meta-information such as editor assigned tags and comments, post scores and keywords, and various stages of editor-created summaries of the discussion threads.

Figure 3-3 shows the conceptual view of the transformation of the information space through distillation. The small circles represent individual posts or discussion threads that have not been distilled. The bigger encompassing circles indicate that some distillation process has been applied. Different colors imply various stages of distillation. The crosscuts of the pyramid represent the stages of the information space as distillation progresses from bottom to top. At the beginning of distillation (depicted as the bottom

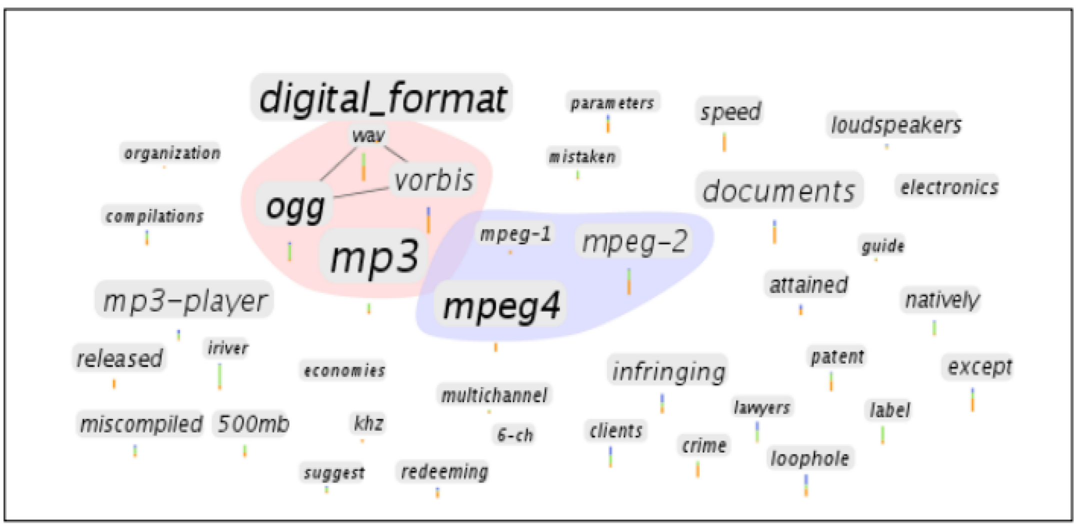

Figure 3-5 A partial view of the Keyword Farm, a visualization of keywords in the discussion space. This visual aid helps editors quickly understand the topic space and allows them to rearrange and group the keywords. It can suggest editors of potentially important topics that have not been covered yet.

plane), the space is essentially raw discussion data imported from an online discussion forum. At this stage, the space is less organized and contains some duplicates and information with little value. It may consist of dozens of discussion threads with hundreds to thousands of posts attached to them and meta-information such as moderation scores or comments. The existing meta-information is presented as notes attached to appropriate places in the space.

As distillation progresses, more and more raw discussion data are transformed into distilled summaries. (Note that different topics can be at different levels. One topic may be completely distilled while another is still intermediate.) This reduces the size of the information space (as depicted at the top of the pyramid), with the information more organized and structured. In the previous distillation scenario, the discussion posts have been replaced with a smaller summary structure. The initially unorganized space gains order and reusability through the distillation process. The process is fully visible in the navigator (a zoomed in partial view of the progress is shown in Figure 3-4).

Both incremental diagenesis and incremental summarization would help editors flexibly distill informal information. Arkose 1.0 was built as a demonstration of these principles, providing tools and mechanisms that would allow a gradual increase in the organization of the informal information.

In the previous sections of the paper, the details of the navigator, authoring tool, and visual aids that constitute Arkose have been purposefully left out to concentrate the presentation on the distillation process and its requirements. I now discuss the technical details of Arkose in the next section.

Technical Details

Arkose 1.0 was built in Java, along with the Swing user interface toolkit and the Prefuse toolkit (Heer 2007). Prefuse provides a rich set of visualization and interaction features with animation and search capability.

Arkose 1.0 itself was developed to be discussion forum agnostic. In other words, any text based online discussion forum could be used by Arkose 1.0, provided a parser could convert the forum into the XML-based TreeML format (Fekete and Plaisant 2003). One of the forums supported is the iDIAG/CyberForum system described in Ackerman et al. 2003.

Arkose 1.0 consisted of four major components: the navigator, the authoring tool, and visual aids. Each is covered in turn below.

The navigator

The navigator was relatively straightforward, but necessary. The purpose of the

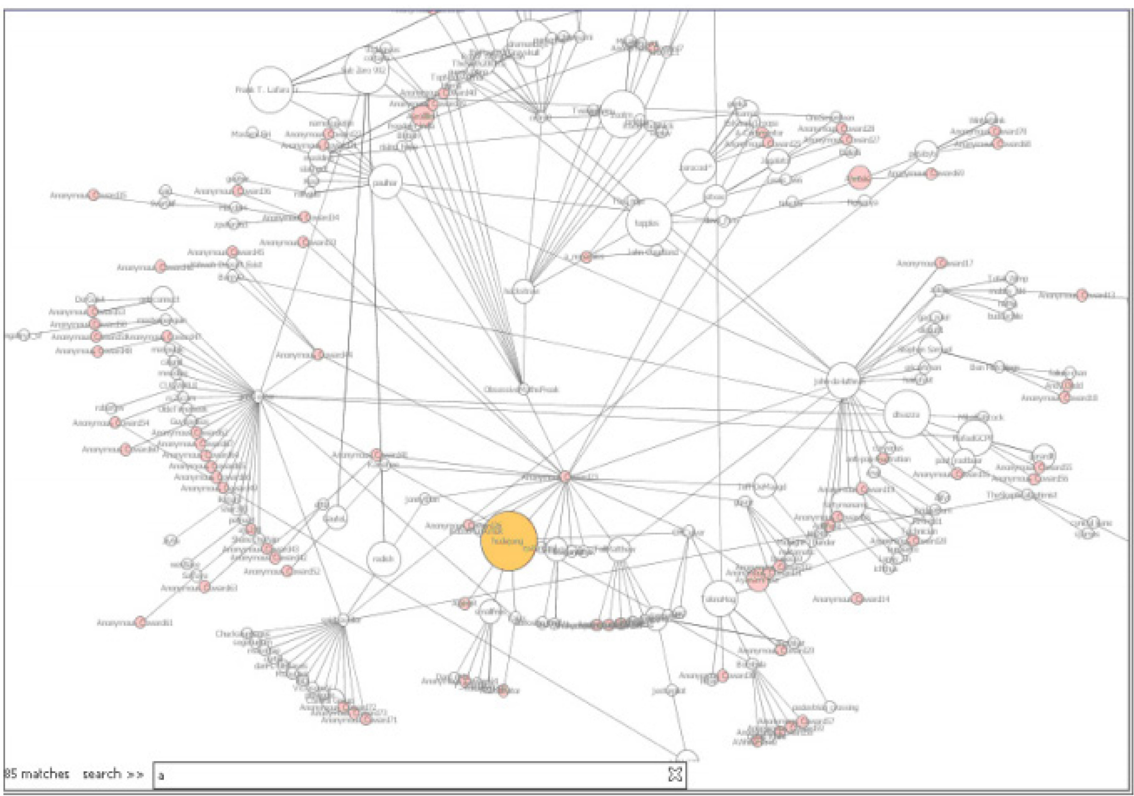

Figure 3-6 An overview of the Author Network. The size of the circle indicates the number of posts a particular author of the discussion space has left. The edges indicate who replied to whose posts, with keywords with high tf-idf values that are common to both authors (when clicked). Keywords found in the posts are associated with respective authors.

navigator (Figure 3-1) was two-fold. One was to visually present the original discussion space in order to solve the partial-view problem inherent in existing web based forums. This would help editors better understand the discussion space by providing an overview and allowing them to focus on specific parts as necessary. The discussion space also served as a substrate for the distillation process. Distilled summaries, tags, various forms of meta-information, and comments were added to the space as the distillation progresses, gradually transforming the space into a more organized and reusable state. The navigator made all the activities of editors visible for a more effective collaboration. The discussion space was fully zoomable and draggable and provided online search capability of the content. The navigator supported multiple visibility levels of posts according to their current distillation status.

The authoring tool

As with the discussion space in the navigator, individual summaries were also incrementally shaped. An editor would create and modify summaries of the discussion space in the authoring tool (Figure 3-2). As mentioned, a summary might be updated as many times as needed by different editors until it was deemed to be complete. Adding and deleting to and from a summary was as simple as connecting and disconnecting edges in the summary graph. This would also allow easy merging of multiple summaries.

Each node in a summary graph would be a typed entity that indicated its role. For example, in the example in Figure 3-2, the types were "topic", "issue" and "fact". These types were customizable, allowing editors to add new ones and modify existing ones. When an editor needed to create a new type, one could be created without any restrictions. However, when an existing type was modified (for example, changing "fact" to "evidence" to better reflect the role), the editor was asked to specify an explanation or justification for the change, which could be read by other editors later.

Each summary was assigned a distillation status when it was exported back into the discussion space in the navigator. The original posts were aggregated with a summary. At the moment, there were four types of status indicating the distillation progress: "Being worked on", "In Progress", "Attention needed", and "Closed." A summary with the "Being worked on" status indicated that it had been imported into one of the editors' authoring tool. This would tell other editors that they should not modify it. "In Progress" meant that initial work had been done, and the current work was exported back to the discussion space. An editor might freely import the summary and work further on it, in which case the status would become "Being worked on" to prevent other from working on it. A summary could also have a more detailed progress indication. Since a summary might be revisited many times before its completion, some indication of its relative progress might be helpful. An editor could specify this by dragging a slider to set how full a bar icon was. At each revisit, the editor would raise up the slider to fill the bar. When the summary was completed, the bar would be fully filled. The idea behind having this kind of secondary indication was to help editors quickly understand the status of summaries in progress. This way, the number of colors for different summary statuses could be kept minimal (currently four) while providing richer information. "Attention needed" was usually left with an editor comment from the editor. This might be a question or concern that would require another editor's help. "Closed" indicated the summary was complete, but Arkose 1.0 allowed editors to work on it further if they wished.

The bottom left column of Figure 3-2 shows the "trigger condition" entries. A trigger condition was a rule used by an automatic process in Arkose. There were three default trigger conditions implemented; 1) notify the editor when another editor modifies the summary, 2) notify the editor when a summary with similar tags is found, and 3) notify the editor when a summary with similar node content is found. The content similarity was calculated by the cosine similarity of term vectors of the nodes in the summary.

As the editor worked on a summary, the automatic process would compare its content and tags with existing summaries to find whether similar summaries already existed so as to minimize duplicate work. Only three pre-defined triggers were available. In future, ways to allow editors to write simple menu-driven scripts to specify various trigger conditions might be helpful. Allowing editors to post and share these customized scripts would facilitate stronger coordination.

The use of visual aids

As well, Arkose consisted of several visual aids that were critical to supporting incremental summarization and incremental diagenesis. These included a visualization of keywords in the discussion space that I called the "Keyword Farm", and a visualization of conversational activities of the posting authors that I called the ″Author Network".

The following subsections discuss these visual aids in detail.

Keyword farm

It is often helpful to quickly see the keywords used in an information space to grasp what topics are being discussed. Recent social bookmarking sites such as del.icio.us implement new ways of utilizing meta-information by displaying user created tags in such a way that visibility varies according to the importance of the tags. I created a similarly purposed visualization, the "Keyword Farm" (Figure 3-5). The Keyword Farm visualized selected keywords from the discussion space so editors could more easily see what needed to be done.

First, I integrated WordNet (Fellbaum 1998) using the MIT Java WordNet Interface to help identify synonyms to allow editors to easily build a quasi-ontology. An editor might replace a number of semantically similar keywords with an overarching new word, or specify a relationship between two keywords.

Second, the visualization utilized both machine calculated data and actual usage data of the keywords to give editors some useful functionality. One was a standard information retrieval technique to calculate a word's probable importance value, or the term frequency times inverse document frequency (*tf*idf*) to give editors statistically computed keywords. (See Appendix B: Proportional tagging for more detail on *tf*idf*.) The top *p* keywords from the original discussion space were picked according to the following:

> First, the words in each post were tokenized, filtered with a stop word list, porter stemmed (Porter 1980), and formed a term vector for the post. After all the unique words were gathered from the entire nodes, each word's term frequency (how many times the word appears in the entire space) times inverse document frequency (inverse of the number of documents in which the word appears) was calculated.

The *tf*idf* scores were visually represented as the size of the words in the Keyword Farm. Thus, the bigger the word was the more likely it was significant. In addition to showing the importance values of individual keywords, a matrix of word frequency was used to indicate groups of words frequently appeared together in the discussion space. Upon selecting a keyword in the Keyword Farm, lines visually connected the word and its associated words to form a graph with the frequency information presented as the varying thickness of the lines. In addition to the machine calculated keywords, editors could add a new keyword or delete an existing one.

The Keyword Farm also suggested important topics that had not yet been covered by the editors, providing for incremental diagenesis. This was done through the second type of information represented in the Keyword Farm, which was each word's actual usage in summaries and tags. The graph bar under each word indicated the word's actual usage. As editors organized the discussion space and created summaries of discussion threads, some of the keywords (raw or as described in the quasi-ontology) were included in them. By visually presenting the usage information, editors could have a better understanding of the work progress, and identify topics that were sufficiently covered and topics that needed further organization. In the Keyword Farm, all the keywords started from the ground initially and as they were used by editors the graph bars under the keywords grew taller so as to push the keywords upward. Colored sections in the bar indicated specific usage cases. For example, the top blue portion represented how frequently the word had been used in a topic in a summary, the second green portion represented the frequency with which the word had been used in editor-specified tags in various places in summaries.

After distillation has progressed to some degree (the current threshold is when every 10% of the posts are distilled), a number of keywords would have been used and thus their graph bars in the Keyword Farm have grown accordingly. However, the Keyword Farm might notice that some of the keywords that it thought were important (i.e., words that have high *tf*idf* scores) had been used very little in summaries and tags. This might indicate that editors simply overlooked these topics, the portion of the discussion space had not been distilled yet, or the *tf*idf* values for the keywords did not correctly represent their actual importance. When notified by the Keyword Farm, an editor might check to see whether the suggestion was a valid one. The editor could then initiate a new distillation process over the discussion space where the keywords were relevant, or the editor might simply turn off the suggestion if it was not correct. Thus, not only more value was gradually added to the Keyword Farm, but it also helped editors incrementally organize the discussion space.

Author network

Another supportive technique was visually presenting the conversation activities of the authors in the discussion space. One of the goals of distillation, as discussed earlier, was to identify authors with interesting ideas. If is often helpful to know how active a particular author is and what the author's messages are about. The Author Network visualized the information in a social network, where an editor could search for authors and their discussion contents and keywords. Each circle in the network represented an individual author. The size of the circle indicated how many posts the author had written, thus showing magnitude of the author's activity, and the links or edges between authors in the visualization represented conversation activities. Keywords from their conversations were visually attached on an edge, so an editor could quickly scan through what the authors talked about. As the distillation progressed and editors identified authors with interesting ideas, tags could be added to emphasize those authors.

Conclusion

In this chapter I have discussed Arkose 1.0 with its support for distillation and reuse of informal information, especially from an online community's brainstorming and discussions. I presented two design principles *incremental summarization* and *incremental diagenesis* that allow a more flexible distillation process.

While Arkose 1.0 was a good distillation tool, it was limited to only an existing and completed discussion space. Although the summaries created from the discussion space could potentially be readmitted into the discussion space and further discussion would be possible based on them, it was not supported naturally. A better approach would be to allow distillation to occur gradually while users were creating and sharing knowledge through a discussion, which is why I later created Arkose2. In order to create a new knowledge generation tool, I first needed to understand how users actually created and shared knowledge and helped one another in a knowledge building community. At the time, new types of knowledge building communities were gaining much popularity but little work had been done on them. These were ultra large question-answering communities such as Yahoo! Answers that dealt with topics of an extraordinary variety. I studied one of the earliest and largest of such communities, Naver Knowledge-iN, in order to learn about users' participation behavior and motivations as well as limitations in the user interfaces that limited knowledge generation and sharing.

In the next chapter, I discuss my study of Naver Knowledge-iN.

Chapter 4

Naver Knowledge-iN[8]

Introduction

Large general-purposed community question-answering sites are becoming popular as a new venue for generating knowledge and helping users in their information needs. The Web has clearly led to new forms of knowledge production on a scale never seen before. One of the most interesting forms is the question and answer (Q&A) community. Q&A communities serve an important role in informal knowledge production, one focused on users helping one another.

This chapter seeks to understand the underlying user participation behavior and strategies in the ultra-large Naver Knowledge-iN (KiN) service. KiN, started in South Korea in 2002 Wikipedia 2004, has an average of 110,000 answers to 44,000 questions asked every day and 4.5 million daily visitors (Kim 2007), and has been a commercial success for Naver. My goal is to understand users' behavior in knowledge generation and sharing through question-answering activities. Further, I investigate what factors exist, especially in the user interfaces, that limit certain knowledge creation and sharing activities. These understandings will help me design more flexible discussion support and distillation capabilities in Arkose2. This chapter describes the features of

---

[8] This study was published at CHI 2009. (Nam, K. K., M. S. Ackerman and L. A. Adamic (2009). Questions in, knowledge in?: a study of naver's question answering community. *In Proceedings of the 27th international conference on Human factors in computing systems*, Boston, MA, USA, 779-788

Knowledge-iN (KiN) and our data collection and analysis methods: a mix of quantitative analysis and qualitative interviews. Following an overall description of the system, I focus on the top answerers on KiN and their characteristics. I then present an analysis of the users' motivations and the observed allocation of expertise. I conclude by discussing design implications for future online Q&A systems.

Interaction in Knowledge-iN

Naver (http://www.naver.com) is the most popular portal site and search engine in South Korea, set up by NHN Corporation in 1999. At the time, most of the documents on the Web were still written in English, which made the development of a search engine in Korean very difficult. NHN started its KiN service in 2002 to overcome the lack of Web documents in Korean (Wikipedia 2004). By having users ask and answer many questions. The main activities on KiN are searching, asking, and answering. The KiN main page aims to capture users' attention while at the same time satisfying information needs and acknowledging the most active users. The page lists some of the most "popular" questions (questions that people voted as useful), questions that need an answer, and questions for which an answer needs to be selected by the askers. The site also presents its visitors with the list of top categories where people can post a question and answer, and provides a list of real time search words that reflect the current interest of the users. Some of the top answerers in the site are also posted on the main page.

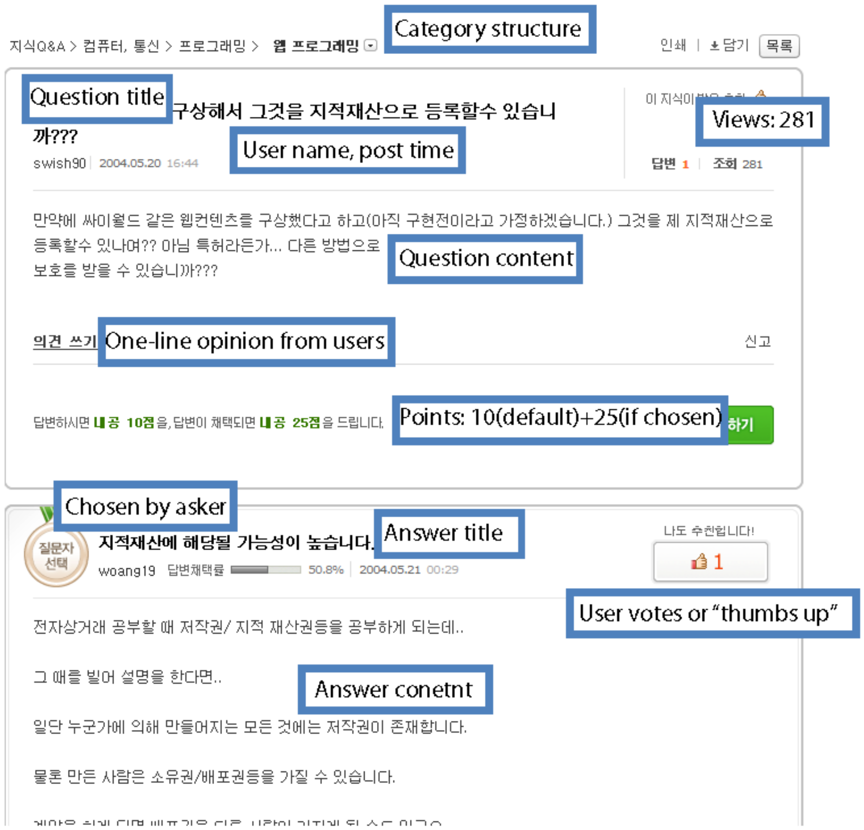

Figure 4-1 An example of a question-answer exchange.

The format of KiN constrains how a question-answer exchange proceeds. An entry is created by a user's posting of a question, and subsequent answers to the question follow in the same entry (see Figure 4-1) Since a user can only post one answer to a specific question, it makes Usenet or forum type discussions difficult. While users may leave a one line "comment" to the question or any of the answers, and an answerer can rebut to an objection, the structure is flat and not threaded as often seen in a forum like Slashdot. A user may also directly email or send a message to another user to ask a question.

KiN has accumulated over 60 million user-generated posts in 14 top categories and over 3,500 sub-categories. As of our data collection, the three largest top categories were ’Computer / Telecommunication’ (4.58 million entries), ’Entertainment and Arts’ (2.58 million entries) and ’Education’ (2.53 million entries), reflecting the interest of Korean Internet users. Other top categories include ‘Games’, ‘Business / Economy’, ‘Shopping’, ‘Society/Culture’, ‘Health/Medicine’, ‘Family/Life’, ‘Travel/Leisure’, ‘Sports’, ‘Local Q&A’, ‘Advice Q&A’, and ‘Juniver KiN’ (for people under the age of 14).

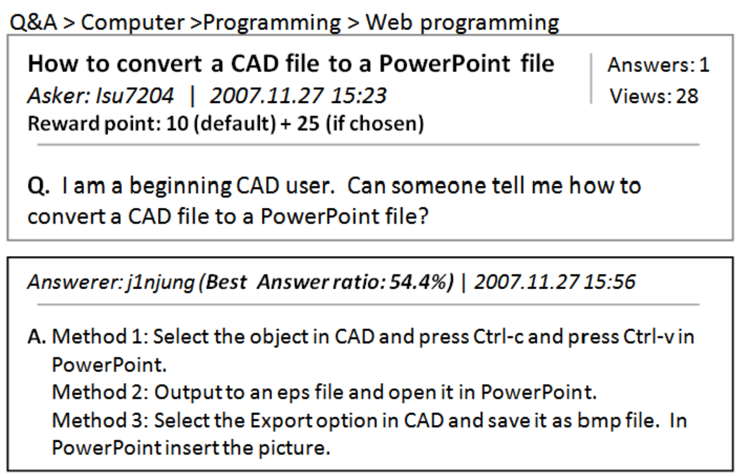

Figure 4-2 A typical question in Software.

There is a huge range of question types, from factual questions (e.g., “Who was the first president of Korea?”), to procedural questions (e.g., “How do you build a computer?”), opinion oriented questions (e.g., “Who is the better singer, A or B?”), task oriented question (e.g., “Can you write a program in C to do X?”), and advice seeking questions (e.g., “I broke up with my girlfriend. What should I do?”). The prevalence of question type differs by category. The Medical and Finance categories are full of advice seeking questions. Categories such as Java or C/C++ have many task-oriented questions, and an example is shown in Figure 4-2. The question received three similarly helpful answers and one “junk” post about a current event.

In the Movie and Singer categories, many of the questions are opinion-garnering ones that often turn into a flame war. The following question in the Singer category garnered an explosive 277 answers, the one chosen as best being humorous rather than informative:

```
Q. Do you think the new singer Shiny will be a hit?

A. What? How come someone younger than me can be a singer?
They are all fake. Who’s Blackbeat? Park Hyun-Bin is the
best.

Other answers: 276
```

Many of the other 276 answers are similarly opinionated and flippant, and some of them have more than 100 one line “comments” that a user can leave regarding the answer. While questions that generate this many answers are relatively rare, they are more likely to occur in discussion categories such as Entertainment. In contrast, in our data of 2.6 million questions spanning 15 categories, the average number of answers per question was 1.7 and about 59% of the questions received only one answer.

Askers in KiN can ask just about any question they can come up with. Naver assists them by providing an automatic category suggestion feature that can guess relevant categories based on the keywords in the question. KiN rewards answer with a point system that allows users to advance from a status of “lowlife” to “hero” or “god”. As of the data collection, an answerer gets 10 points for providing an answer, and 25 additional points for being selected as the best answer. The asker can post up to 100 extra points in hopes of receiving faster and better answers. Users receive points not only for answering questions, but also for various activities in KiN. For example, logging in everyday is 3 points and voting for an answer is 1 point, and so on. Users with a large amount of points are ranked and listed on the site. The top 1,000 users from the entire site are listed on the KiN People page, and each sub-category lists its own top 10 users on its page.

Data Collection

The data collection was comprised of two complementary parts: automated crawling of publicly accessible question-answer pairs on the Knowledge-IN website (www.kin.naver. com) and phone interviews. The crawl of millions of questions and their answers allowed us to gather comprehensive statistics on user behavior, while the phone interviews with the most active users yielded insights into the driving motivation behind the observed behavior.

The Web interface of Knowledge-iN allows only viewing of up to a certain number of questions (1,500 as of the data collection) per topic category, which can be as little as a few days' worth of data. To overcome this limitation, I devised a search method using frequently used terms in questions in each category that I believe retrieves over 99% of Knowledge-iN's questions with answers. I collected approximately 2.6 million questions and 4.6 million answers from 15 categories from a five year period between 9/1/02 and 12/31/07. This excludes questions without answers, which are removed from the system after a period of time.

I supplemented the question and answer data with telephone based interviews of 26 KiN users. These users were a mix of heavy and moderate answerers, along with two askers. I found these users by asking people who had recently posted a question or answer (so as to form critical incident interviews), and by asking people who were in the top 1,000 answerers. I had a very low response rate (less than 2%), and as a result this group cannot be considered representative. Nonetheless, their responses were extremely useful in understanding the site, and if taken appropriately, may lead to insights about those who answer on KiN.

The interviews were conducted by me and were semi-structured, meaning that the interviewer asks specific questions but was then free to follow up as appropriate. The interviews lasted between 20 and 45 minutes. They were recorded, and then translated and transcribed. I used standard qualitative coding techniques, based in Miles and Huberman (Miles and Huberman. 1994), to code and analyze the transcripts. These techniques allow for the appropriate categories of analysis to emerge from the data; as such, our results should be considered generative rather than confirmatory. I then compared with the crawled data which revealed mutually supporting evidence for the role of the point system, users' strategy and desire to answer unanswered questions, and the behavior of correcting existing answers. As the reader will see below, I believe that the combination of the two was analytically fruitful.

Patterns in Participation

Those who ask don't answer

In KiN, the users are largely divided into askers and answerers, with only 5.4% both asking and answering in the same category. The Naming category (for suggesting baby names) had the smallest overlap (1.9%) and the Singer category had the largest (10.4%). As a result of the separation of the asker and answerer role, there is very little within-category reciprocity in KiN. This lack of reciprocity is reflected in the absence of reported community by interviewees. As we will see, answerers tend not to know other participants, and rarely exchange personal email.

Category	Num Asker	Num Answerer	Answer by top 1 %	Avg. answer by top 1 %
C/C++	37145	15960	33.5%	223.53
Car repair	26618	11735	41.4%	234.90
Fashion	117934	142482	22.1%	48.67
Finance	148346	72251	52.8%	344.13
Java	10437	3986	31.4%	194.28
Laws	152221	107832	44.4%	152.45
Linux	6723	3390	36.3%	159.09
Medical	417243	381210	35.1%	123.17
Movies	75082	93794	27.4%	61.49
MS Windows	105442	59378	35.6%	138.43
Naming	35048	46808	29.7%	55.89
Singers	134002	203726	23.4%	63.02
Stocks	24352	11982	43.6%	226.97
Translations	313637	155686	37.7%	207.35
Web design	49816	33052	34.9%	110.32

Table 4-1 Number of askers and answerers from the 13 categories. Most categories have more askers than answerers, except for the Fashion, Movie, Naming, and Singer categories.

I also note an overwhelming tendency of users to specialize. Even though Naver contains over 3,500 different subcategories, many users focus on one or a mere handful. Among the top 1,000 repliers, over 52% the users had more than half of their answers in a single category. Among those, 30% had more than three quarters of their answers in a single category. It is possible to maintain this focus either by consistently browsing questions in just a single category, or by specifying a particular category as an "interest" and being shown questions specifically in that category.

Top answerers

Because there is such a strong separation of roles between askers and repliers, I decided to focus on the answerers only. It is among this group of users that I find the heavy tail in activity: a number of users have answered tens of thousands of questions, but it is rare for an asker to ask more than a couple of hundred questions. It is therefore unlikely that a highly active asker is individually significantly contributing to the activity of the site, while a "top answerer" may very well be helping hundreds to thousands of others. Given that the point incentives in Naver are to answer as many questions as possible (with points awarded and no cost, save for effort expended, in providing a poor answer), I wanted to develop a measure that would accurately reflect the quality, and not just the quantity of a person's answer. I use a γ or guru score, described in Appendix A, to measure the answerers' performance against the random chance of their answers being selected as best.

I compared the most active 1% of the users (by the number of the answers provided) with the rest of the users. Their contributions account for 22% to 52% of all answers ever given (Table 4-1). The top answerers not only produce many answers, but the quality of the answers, measured by the guru score is better than the average (Table 4-2).

There are important differences in categories that may affect top answerers' behavior. First, there is little competition for the best answer within a question. It varies among the categories: C/C++ and Java are the least competitive with the average of 1.42

and 1.30 answers per question respectively, and Movies and Singers are the most competitive with 2.23 and 2.57 answers per question. This may be due to the type of expertise sought in each category. In categories where many questions require hierarchical expertise (i.e., to know an advanced topic such as thread programming, one would be sufficiently experienced in basic programming) may have fewer qualified answerers than categories that require ”flat” knowledge (i.e., one does not need to know about all singers to answer about a particular singer.). Additionally, one “correct” answer to a programming question may be sufficient, but several opinions may be welcome about a singer.

Category	Guru score top 1%	Guru test
C/C++	0.0884	-0.172
Car repair	0.0230	-0.132
Fashion	0.0465	-0.0451
Finance	-0.0454	-0.0454
Java	0.0484	-0.117
Laws	0.151	-0.176
Linux	0.172	-0.0950
Medical	0.0573	-0.0859
Movies	0.0738	-0.147
MS Windows	0.0428	-0.157
Naming	0.0278	-0.116
Singers	0.0765	-0.0732
Stocks	0.0902	-0.154
Translations	0.102	-0.199
Web design	0.103	-0.164

Table 4-2 Guru scores for top answerers compared to the rest. All differences significant at $p < 0.001$

Motivation for Participation

What is most essential to the success of sites such as KiN is the participation of its users, but this is also one of the biggest puzzles. What motivates users to essentially provide knowledge services for others, when there is very little in the way of explicit rewards?

Most of my understanding of why people participate in KiN comes from the interviews. In the interviews, users gave a number of reasons. By far, the most often stated reasons were the wish to help others, to learn and review material, or to participate as a hobby. Several people also pointed to participating for business reasons. These are not mutually exclusive motives: Most interviewees gave multiple reasons.

These motivations echo prior literature (e.g., Constant et al. 1996; Butler et al. 2007), as might be expected. However, two of the motivations, participating as a hobby and learning, have implications for the way participation by answerers was intermittent. And also the learning had some interesting cross-cultural implication, as will be discussed in a later section. Furthermore, motivations were often co-mingled with interviewees' discussions of KiN's game-like points and their quest for those points. I will discuss each in turn below, and relate it our analysis of the question-answer dataset as appropriate.

Altruism and helping others

When asked why they participated, many users said to help others, by providing knowledge that others did not have. Altruism, then, was the most oft provided answer. The type of help was dependent on the answerer's expertise. One doctor stated:

> Since I was a doctor, I was browsing the medical directories [in KiN]. I found a lot of wrong answers and information, and was afraid they would cause problems. So I thought I'd contribute in fixing it hoping that it'd be good for the society. [Sangmin]

(All responses have been anonymized.)

Another expert stated:

> Many people in Korea have incorrect information about social security, and I was a bit frustrated because I work in the area. So I started out to explain it to people. Sometimes people sent me a thank-you email message for answering, and it motivated me more. [Youngsoo]

Another answerer, who participated in the Translation category, was not as expert. She instead stated:

> I try to answer so that regular people can share knowledge, rather than technical knowledge. ...Someone needs it, and I have the ability to do it, and it'll be a service to society. [Mirae]

Altruism was a very common response for our interviewees. It may seem, at first glance, as though these claims of altruism are merely "public" statements, that is, socially-sanctioned responses that people give as a matter of routine to those they do not know well. There is evidence in our data that suggests that I saw some "backstage" responses (i.e., not merely "public" responses); I will discuss this below further. However, this is deeply resonating with Korean culture, as one of our interviewees insightfully noted:

> I think there's something about Korean people that they take time to write about something when there's no real benefit to them. If you take the GRE for example, one doesn't really need to share information, but on Hackers [a famous Korean site for sharing standardized exam information], people share a lot of information. Some people make their own report and study material [available] without anything in return. [Garam]

Business motives

Yet another interesting motivation also discussed in interviews was promoting a user's business through answers. This goes on in many online worlds - not just KiN. For example, it is a recognized problem in review and recommendation sites (Dellarocas 2003).

In a more implicit manner, this seemed quite acceptable to our interviewees. One person reported wanting to be considered an expert for the status in his online e-commerce community. Another user reported getting offers to publish books or other tour guides based on his activity in the Tourism category of KiN. At its most extreme, two interviewees reported getting solicited by other online communities attempting to garner online participation.

A more explicit manner of promoting a business was reported in the Medical and Finance categories. (It may exist elsewhere as well; this is a limitation of our interviewing.) Said one interviewee:

> I've been working as an insurance agent for 9 years. I started answering in Knowledge-iN as part of my business activity. In the evening, I answered questions to solicit potential clients.... So when I'd leave an answer, I'd say I would meet with you face-to-face to talk about more details and give you advice. [Taein]

Two interviewees stated that they had originally started on Naver to gain clients, but they found it to be less valuable than they had hoped. Instead, they stayed as a hobby and for altruistic reasons.

Learning

Many interviewees reported wanting to gain further understanding or to maintain their current understanding of a topic. This included reviewing what they knew before or extending their knowledge by explaining it to others. One interviewee said:

> My first intention [in answering] was to organize and review my knowledge and practice it by explaining it to others. [Taein]

Others reported learning through practice:

> Answering questions helps me study. I can learn from answering [in Translation]. I get to review what I used to know such as vocabularies and idioms. [Minhyuk]

Still others reported that explaining a topic to others maintained and perhaps extended their understanding. In a few cases, interviewees reported active learning. Two answerers, both in the C/C++ category, reported taking programming questions as practice problems to learn more about the language.

Review may strike some as less important than active learning. On the other hand, review of material is an essential part of Korean educational processes (much more so than American or most European systems), and so this motivation is heavily resonate with Korean culture.

Hobby and personal competence

Many interviewees reported that they viewed answering questions as a hobby - something to do when they had spare time, as this user did:

> But in the evening after work, my kids are asleep, and I don't just want to stare at the TV. I go to Naver, and it's fun to answer. It's interesting to know what questions are coming up. I keep repeating [going there]. [Kisoo]

This implies less of a sense of obligation on the part of the participants and more of a casual interest. One interviewee, however, did note that his involvement was heavier:

> Yes [I answer everyday]. I am addicted (laughs). [Nami]

Several other interviewees, as one might expect with any community, spent from morning to night online during some period. For these people, their involvement might be obsessive; others reported continued involvement not from compulsiveness, but from a continued sense of mastery and competency. When interviewees talked about their trajectory of participation over time, many people talked about a strong desire to earn points and advance in level, to promote their businesses and professional lives, and to help others early in their KiN "careers." After the initial period, it appeared that, while the altruism stayed or grew stronger, many participated more as a hobby (which reflected their casualness of participation). As we will see later in the paper, this resulted in often intermittent participation over time.

Points

The point system in Naver is another source of motivation for users. It also interacts sharply with the other forms of motivation. As mentioned, when an asker posts a question, he or she may give up to 100 additional points hoping to get a faster and better answer. Since some of the top answerers enjoy high visibility and sometimes a celebrity-like reputation on the site, points can be a driving reason for participation.

Interviewees often dismissed the points, as this user did:

> I don't really care about the points. Points don't affect me in looking at questions or leaving an answer. [Mirae]

Most interviewees echoed this. But she goes on to say:

```
Although points do not affect whether I answer a specific
question or not, it's fun to see them accumulate.
```

This pattern was also typical in our data. For example, another interviewee stated:

```
I don't care about the points. [But] It's fun to see points
accumulate and my character level up [increase to the next
level]. [Jeyeon]
```

This interacted with helping others. Interviewees felt that others would trust them more or consider them greater experts if they were higher levels, such as this interviewee:

```
I felt I needed to be a high rank so people would trust my
answers more. That way more people would look at my
answers. [Sangmin]
```

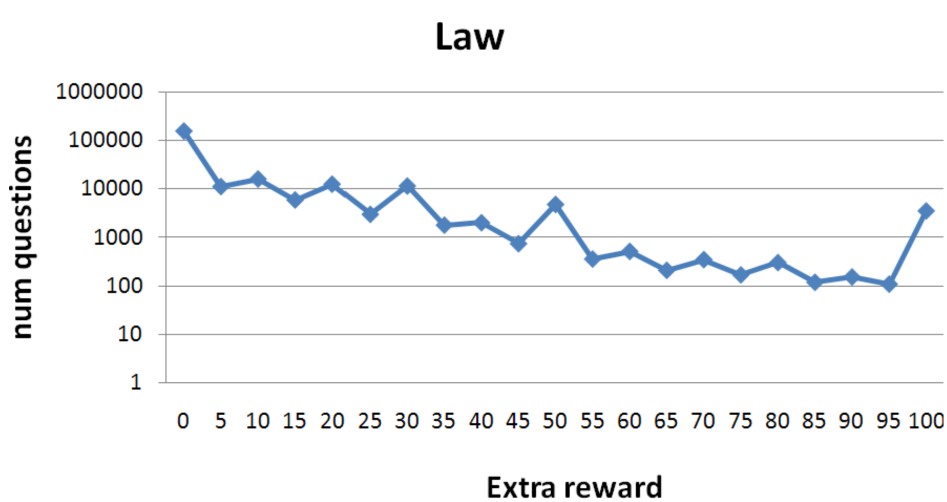

Figure 4-3 Typical number of questions per expected rewards. Other categories exhibit a similar distribution.

Overall, interviewees claimed they did not care about the points, but points did seem to have an effect, often weak, on their activities.

> I'd be lying if I said [points] had no effect. Points are really nothing if you think about it, though. But you care about it. [Sunhan]

That users are weakly motivated by points is also evident in the number of answers a given point reward attracts. First, I describe the distribution of points offered across all questions in our dataset. Over 60% of all questions across the categories have only the minimal reward points. The number of questions with a given number of points decreases gradually as the number of points increases with the exception of bumps at 50 and 100 points (Figure 4-3). These two bumps, along with small peaks at increments of 10, may be indicative of askers thinking in terms of coarse granularity of points (i.e., small, medium, and large rewards).

By observing the number of replies at each point level, I find that users are slightly more motivated to answer questions with higher point awards. In most categories, the average number of answers to a question gradually increases as the expected reward goes up (Figure 4-3). The additional motivation to answer questions with higher awarded points was mentioned by several users, though primarily as a secondary motivation following answering unanswered questions. Many also mentioned weighing the amount of time a question would take to answer against the number of points offered.

While users may be motivated by higher numbers of points, the points offered are also used an indication of how motivated the asker is to obtain an answer. One interviewee noted:

> Usually questions with points do not seem frivolous. I feel like answering questions with points, not because of the points, but because those questions are more detailed and seek realistic help. [Taein]

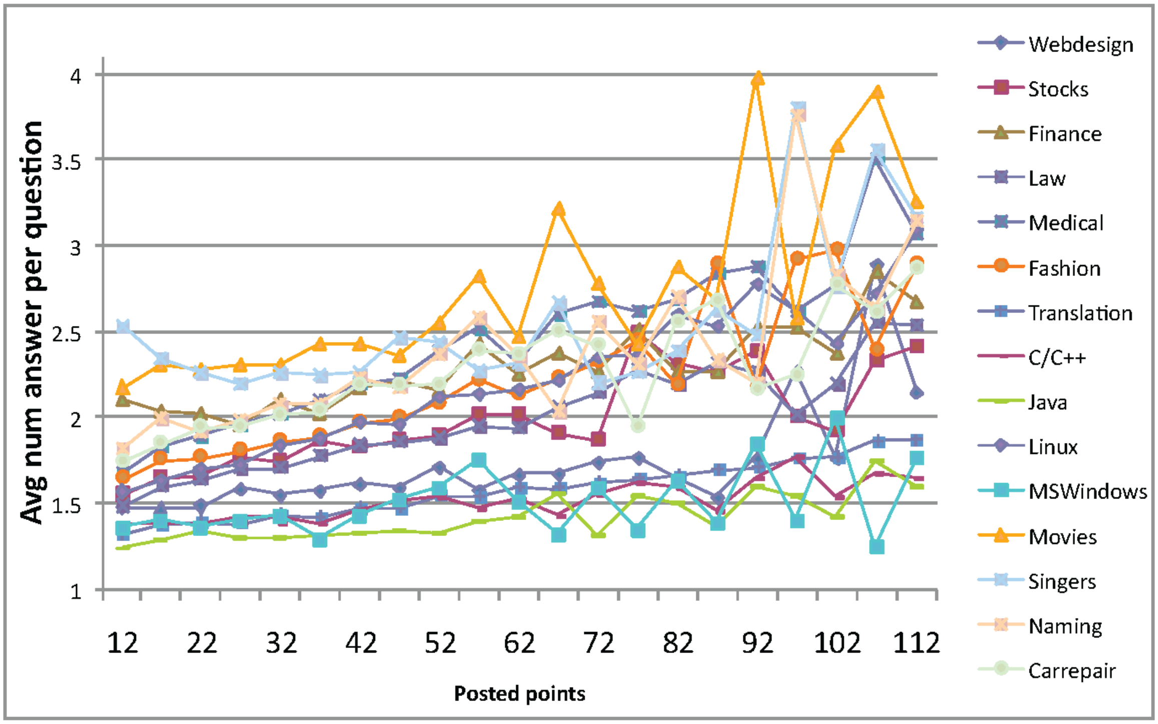

Figure 4-4 Average number of answers for a question offering a given number of points as reward.

Therefore the attraction of answerers to questions with higher awards may be in part due to the perceived need of the asker, in addition to the interestingness of the question. The points offered are a surrogate for the perceived need. A related observation was made in an analysis of Yahoo! Answers that showed the good answers tended to correspond to good questions (Agichtein et al. 2008).

Allocation of Expertise

As mentioned, KiN is a highly popular site whose community has collaboratively generated over 60 million questions and answers, with approximately two thirds of the questions receiving an answer. But the success of such a system hinges not just on the volume of questions answered. In this section, I examine three critical characteristics for any Q&A site - what knowledge level users can expect, whether problematic or erroneous responses are corrected, and how well the necessary expertise is covered across time and categories. I cover each in turn.

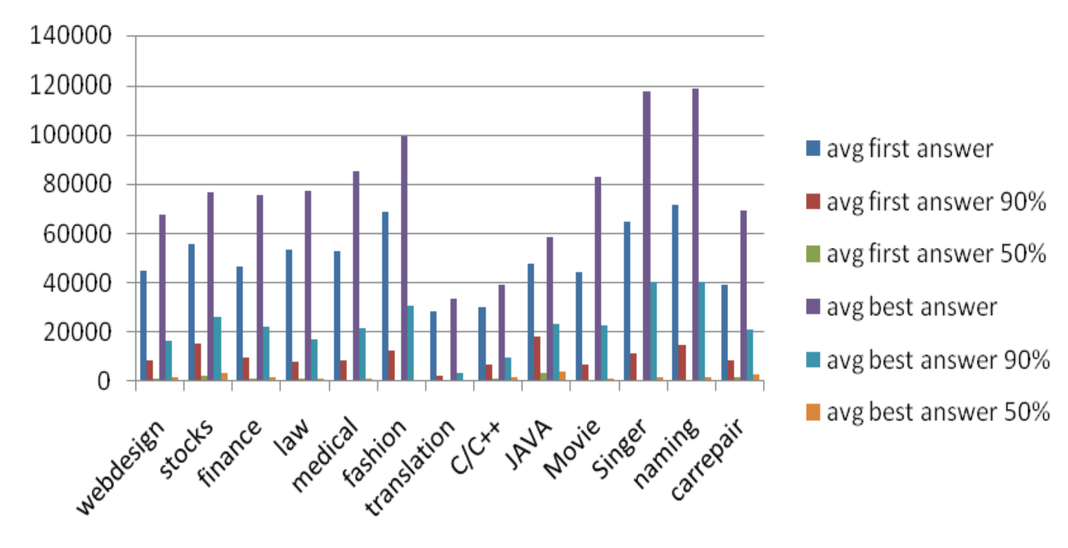

Figure 4-5 Delay before a first answer and a best answer. Across the categories, 50% of the questions received the first answer within 15 minutes and a best answer within 30 minutes.

Knowledge Level and Quality

There appears to be a level of knowledge that users can expect in answers on KiN. All of the interviewees stated that the information in KiN was useful for getting information on commonsense knowledge, current events, basic domain knowledge, advice and recommendations from people, and diverse opinions. Our examination of the site confirms this view. As one interviewee stated:

> It's very useful for getting everyday information. The most useful information is recipe or directions. [Sunhan]

Some interviewees were able to point to certain types of information that could not be found in KiN:

```
It's hard to get professional knowledge. [Jinoh]

Knowledge-iN is not for very domain specific, technical
information. [Hyeil]
```

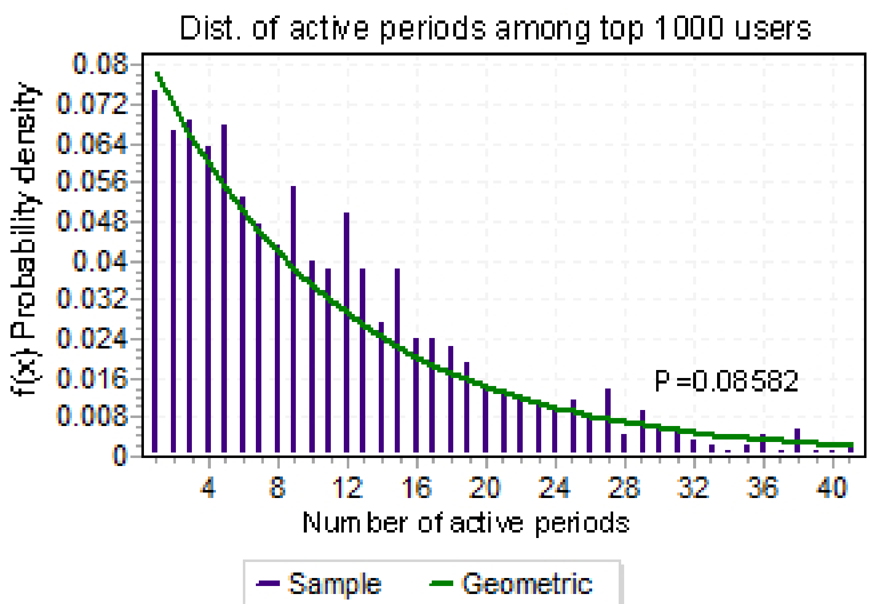

Figure 4-6 Distribution of active periods among top 1000 users.

Many of the heavy answerers in our interview pool stated that online cafés were a better source of detailed information. Cafés is a type of popular online community in South Korea where people with similar interests can join and perform various activities such as posting new information, uploading and downloading materials, and discussing related topics. Sometimes the members have offline meetings.

This inclination towards a level of answer appears to be heavily affected by two factors in how answerers pick candidate questions. First, people tend to want to answer quickly. Some of this results from the point system. Said two users:

```
I look at high points first. But I answer all questions.
[Eunjin]
```

and

```
To higher point ones I provided more detailed answers
(laughs). [Sanggyu]
```

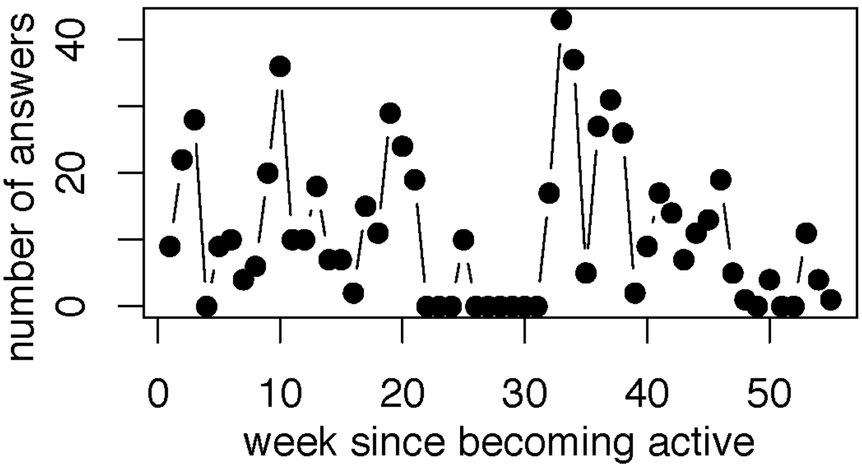

Figure 4-7 Weekly activity levels of a user.

However, the answerers are also pressed for time. Some answer when out of work or when their job responsibilities are light, but others try to cover many questions quickly:

```
If I'm answering when I am taking a break at work, or I am
answering at home in the evening after work, I don't have
much time. A longer answer takes more time. [Sunhan]
```

Time pressure, of course, interacts with the desire to obtain points.

Second, my interviewees tended to answer questions for which either they already knew the answer or they had to look up only minor additional information:

```
I only try to answer with what I know. If I know somewhat,
but not entirely, then if I answer it may be incorrect.
It might hurt the asker. [Minjae]
```

Others were willing to answer questions slightly beyond their expertise:

```
If I am not sure about the answer, I skip [it]. Sometimes
I study further and provide an answer. [Manki]
```

While the expert professionals mentioned answering in depth, they also added two additional constraints. Answering in depth required substantial time, which they seldom had. As well, askers did not provide the relevant detailed information to make recommendations or suggest diagnoses. These kinds of interactions were, in their opinion, more prevalent in the more intimate cafés mentioned above.

Variable	Estimate (se) * 100
Num answers	-0.001 (0.002)
Num active weeks	0.170*** (0.043)
Num active periods	-0.863*** (0.205)
Time span (weeks)	0.047 (0.025)

Table 4-3 Regression model of users' guru score based on temporal activity. * indicates $p < 0.001$ $R^2 = 0.04$**

This constant churn of mid-level questions and answers is reflected in our interviewees' view that KiN lacks a sense of community, again in contrast with the close-knit atmosphere of the cafes. All of the interviewees mentioned that they did not interact with anyone in KiN other than through asking and answering, and that there was neither a sense of community in existence nor the possibility of forming a community in the future. No one mentioned detailed interactions with other users.

Correcting inaccurate information

If there are many erroneous answers, as interviewees believe, then it is important to a well functioning site that those answers be corrected. Indeed, interviewees reported often that they supplemented another person's answer when it was incorrect.

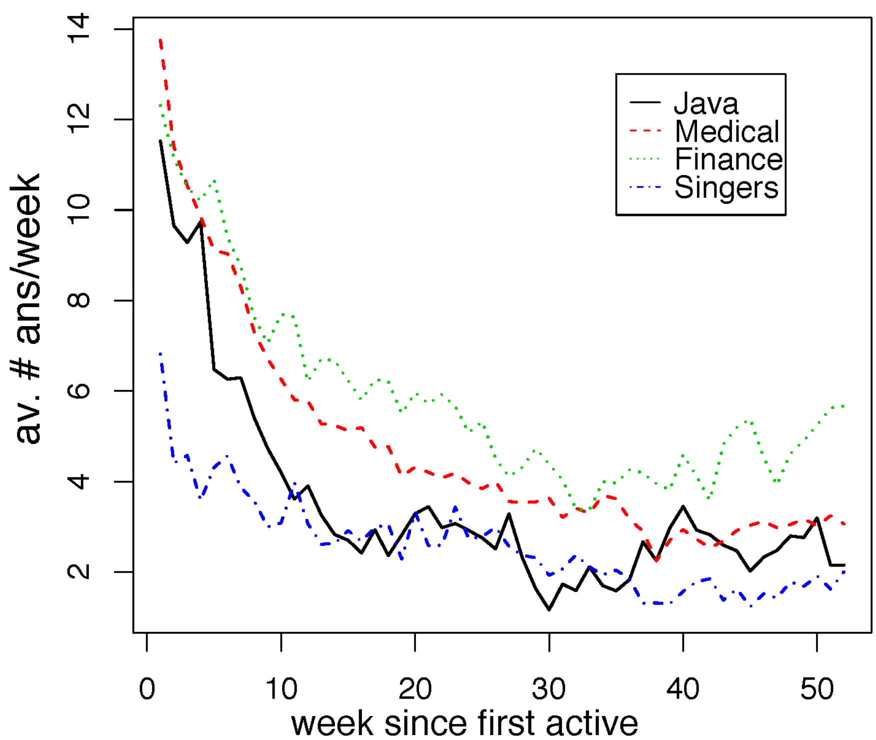

Figure 4-8 Weekly contributions averaged over users who posted > 100 answers and became active more than a year prior to the crawl.

I see evidence of this in the reply patterns. In the Java and C++ categories questions that receive between two and five replies (I chose to omit long threads that may represent discussions), the next-to- last question is chosen as best answer 51% of the time (in C++, 50% in Java), and the last question 69% (66%) of the time ($\chi^2 = 974$, $p < 10^{-16}$). Note that more than one answer can be selected as best. Even in the Singer category, where arguably people are frequently expressing opinions, the last answer is selected as best 50% of the time, compared to 40% for the next to last answer. This indicates that later replies, especially the last reply, are likely a correction or improvement upon earlier answers.

Further evidence that users evaluate the quality of the previous answer before posting their own is found in the fact that answers posted by users who have a history of good answers are less likely to prompt additional answers. Specifically, if the first answer is given by a user with a guru score, the number of subsequent replies is reduced ($\rho = -0.11$, $p < 10^{-16}$). This suggests that other users, upon seeing a good answer, are less likely to submit another. The above two observations are consistent with a strategy where users elect to answer questions where their expertise is needed, but tend to avoid wasting effort on already adequately answered questions.

Intermittent Participation

In addition to providing answers at an expected level of knowledge and quality, the necessary expertise has to be at hand on the site. This can occur only if there are sufficient answerers to handle the workload. However, I find that individual users are highly intermittent in their participation.

One user commented:

```
...since the last quarter of 2007, I became the head of the
team and the range of my work changed, so it was difficult
to find time. [Mike]
```

Another user, whose activity patterns over the span of a year are shown in Figure 4-7 mentioned new familial obligations, and loss of access to a computer at home (having to pay for Internet access at Internet cafés instead) as reasons for weeks of inactivity. When I consider the average number of answers posted per week for all users who had posted at least 100 answers in the Java and other categories, I see a steady decay in activity from the very first week (see Figure 4-7). This is due to many users starting out with a burst of activity, but ceasing all or most activity within a matter of weeks.

One might expect that users who remain on the site and answer thousands of questions do so by participating regularly. The surprising result was that even among the most active users, intermittency was the norm. I defined an inactive period to be one where the user posts no answers for one or more weeks. Figure 4-6 shows the distribution in the number of active periods, separated by inactive periods. Although users typically have a few active periods, quite a number of users have dozens of active and inactive periods. Note that an active period may encompass a single week up to years. A user with a single active period may have only used the system briefly, or has been using the system so regularly, that he never became inactive.

That many users are leaving after a short period, or attending to questions only intermittently, appears to have relatively little impact on the rate of questions being answered and the quality of those answers. The number of answers per question remains steady, even as many users join and leave. For example, in the medical category, the number of questions per week fluctuates by 25%, due to seasonality in the data, but the weekly ratio of answers to questions varies only 8%. Nevertheless, the users do not always meet the demand for answers with a proportional supply. An increase in the number of questions is correlated with fewer answers per question ($\rho = -0.29$, $p < 0.001$).

Looking at 1,000 users who answered the most in the 15 categories I crawled, those who were active a greater number of weeks also answered more questions ($\rho = 0.44$, $p < 0.001$), but the time span from initial to final post was less correlated with volume ($\rho = 0.12$, $p < 0.001$). This again indicates that many users' activity need not be contiguous.

Indeed, as I learned from the interviews, it often occurs as a hobby, in front of the television, when one has few job responsibilities, or otherwise has free time. From a regression, shown in Table 4-3, I can see that being active for a greater number of weeks has a weak positive correlation with quality, while being more intermittent has a weak negative impact. This implies that more committed and consistent users are more likely to provide good answers.

Design Implications and Conclusion

Naver's Knowledge-iN, an Internet-scale Q&A site, with its 60 million questions and answers, is arguably one of the most successful sites of its kind in the world. The success of KiN and other Q&A forums masks to what extent providing answers by ordinary users about all kinds of questions on such a scale is a difficult task. Expertise must be organized and allocated, which is especially difficult when it is discretionary and unpaid. In this thesis, I analyzed approximately 2.6 million questions and 4.6 million answers, and interviewed 26 participants. I found that answerers' motivations were a combination of altruism, learning, participating in a hobby that brought a sense of competence, business motives, and questing for points. I also saw, through a quantitative analysis, how these motivations led to critical features of KiN's question answering - a range of knowledge that was displayed and sought, the ability to correct answers, and coverage.

The level of questions in KiN appears to be relatively low. Some factors affecting this were the user interaction mechanism implanted in KiN that only allowed one posting per user. This prevented answers from being shaped collaboratively and more thoroughly, leading to the accumulation of easy questions and answers. The incentive system that encouraged users to answer many easy questions quickly rather than difficult questions also contributed to the low quality of the information in the community. Arkose2 will tackle this problem by providing more flexible user interfaces that allow continual shaping of the question and answers, break a complex problem into smaller sub-problems, and further organize the discussion space.

The observed distinction between answerers and askers may imply separating user roles in distillation or organization of information might be fruitful in Arkose2. In Arkose2, users with more background or experience who know the topic well can create summarization scaffolding, in which other users may fill with summaries. This would allow distillation to be done more effectively, sharing workloads among the users.

In the next chapter, I discuss my implementation of Arkose2 based on the lessons from Arkose 1.0 and the Naver Knowledge-iN study.

Chapter 5

Second Generation Arkose

Introduction

A successful knowledge building community depends on continuing interactions among its users and accumulation of quality information that help its members solve a problem. The collective intelligence being created is undoubtedly critical for the success and continuing existence of a community.

Based on the design lessons from Arkose 1.0 (Chapter 3) and understanding of user interactions in a knowledge building community from the Naver Knowledge-iN study (Chapter 4) and other literature, this chapter concerns the design and development of new interaction mechanisms and collaborative tool for flexible information generation and distillation. I name the new tool Arkose2. Arkose2 builds on the lessons and ideas from the two of my previous studies. It enables collaborative distillation and utilization of reusable knowledge in more flexible ways than Arkose 1.0 (See Table 5-1 for comparisons).

Arkose2 tries to overcome the weaknesses of question-answering communities, discussion forums and wikis to enable better knowledge generation and distillation. While the original Arkose 1.0 allows gradual distillation of information from a completely unstructured state (e.g., a finalized discussion thread), Arkose2 encourages the generation of better information and ongoing distillation of it.

<table>
<tr><th colspan="2">Arkose 1.0</th><th colspan="3">Arkose2</th></tr>
<tr><th>Functionality</th><th>Capability</th><th>Additional functionality</th><th>Additional capability</th><th>Contribution</th></tr>
<tr><td>Interface for grouping posts</td><td rowspan="4">Incremental distillation of discussion space</td><td>Interface for creating sub topics and sub questions</td><td rowspan="3">Handle complex questions that require multiple answering cycles</td><td rowspan="3">Better handle questions that require collaborative multi user answering that other question-answering forums or wikis did not do well</td></tr>
<tr><td>Interface for replacing posts with summary</td><td>Move posts around, show posts in the original thread</td></tr>
<tr><td>Progress indicator for distillation process</td><td>Separate interaction interface for question formulation</td></tr>
<tr><td>Question posting, answering among editors</td><td>Interface widgets such as polling and survey</td><td rowspan="3">Help organize and search information space</td><td>Allow customized information collection widgets that fit the needs</td></tr>
<tr><td>Keyword farm that shows important keywords and their usage in distillation</td><td rowspan="2">Guide distillation process</td><td>Proportional tagging to simulate TF*IDF values for tags</td><td rowspan="2">New mechanisms for tagging and organizing information</td></tr>
<tr><td>Automatically find posts with similar topic in the discussion space</td><td>Organic categorization for changing content</td></tr>
<tr><td rowspan="3">Author network that shows conversation interactions among authors</td><td rowspan="3">Help identify interesting authors and topics</td><td>Automatically create topic entries in a wiki as discussion progresses</td><td rowspan="3">Store valuable information from discussion spaces in a wiki to reuse the information. Create and link contextual information to a wiki for better information navigation and use</td><td rowspan="3">Linking different types of information systems is not new (e.g. combining a mailing list with a wiki) Arkose2 allows creation and linking of contextualized information to a wiki, so a wiki is both a information repository and starting points for further discussion</td></tr>
<tr><td>Allow searching and linking related information</td></tr>
<tr><td>Interfaces to add contextualized information to a wiki</td></tr>
</table>

Table 5-1 Comparison between Arkose 1.0 and Arkose2.

In many knowledge systems, links to other related information are sometimes used to help a user acquire more information. These related links are usually found by keywords or content analysis. In Arkose2, users may connect generalized information with contextualized information. For example, a domain expert could write a summarized article on car brakes in the wiki. This article would be generalized and objective. Other users may associate specific instances of applied information (for example, "how to change car brakes") in order to add practical or tacit knowledge. Further information may be added as users express specific information needs and start a new discussion forum out of the wiki article. The article then may be updated with answers and discussion for the users' questions.

The combination of the general, objective information and applied, contextualized information may make the information both highly reusable by the general population of the users alike and at the same time responsive to the emerging needs of individuals.

In addition, Arkose2 allows users to create a customized view of the information. Arkose 1.0 did not. One shortcoming of a wiki or discussion forum is that every user sees the same content in the same format. Some users may want to save certain topics in an order not used in the original wiki or discussion forum. They may also want to change part of the information by adding new information or deleting unnecessary information. This may even create an information summarization market where users may request other users to summarize part of information or gather related information together and reward those who provided good content.

Design Principles

Arkose2 includes the design principles of Arkose 1.0, the *incremental summarization* and *incremental diagenesis*. In this section, I discuss two additional design principles of Arkose2.

Ongoing Distillation

As briefly discussed in the previous section, Arkose2 is both a flexible information generation system as well as ongoing distillation system. It focuses on providing users with flexible user interfaces that help accumulate more quality knowledge with less effort. It tries to make the distillation tasks easier by allowing *ongoing distillation* through supportive mechanisms such as summarization scaffolding, visual aids, and information retrieval capabilities. Arkose2 enables synergetic efforts between knowledge generation and ongoing distillation, which builds on the *incremental diagesis* principle from Arkose 1.0. Ongoing distillation has two advantages over it. First, whereas Arkose 1.0 distills a discussion space after the discussion is done (*after-the-fact* distillation), distillation can be performed for a smaller discussion space at a time while it is still progressing in Arkose2. This may make it easier for users to handle. As found in the evaluation of Arkose2, participants felt distillation was much more difficult when the discussion space was larger and more complex. As well, the summarization scaffolding that breaks the discussion spaces into smaller topic groups was found to be very helpful because a participant only needed to focus on a small portion at a time. Thus, ongoing distillation while the discussion is in progress may be easier and more efficient than distilling the entire space at once. Second, the discussion discourse itself may benefit from the distilled information (Ackerman et al. 2003). When the discussion space grows very large with a number of ideas and information accumulated, it becomes hard to grasp the overall state of the discussion. A user might unwittingly start a new discussion on a topic that other users are already discussing elsewhere in the discussion space. Newcomers to the discussion space may need to read much of the posts before they can start participating in the discussion. A distilled outcome that consists of summaries of the posts, discussions pointed out as incomplete, and claims found to lack examples and evidence may guide future discussions among the users.

While this ongoing distillation is achievable by Arkose 1.0 to some degree, it is not supported explicitly and naturally, and it would be somewhat difficult to integrate the distillation outcome with a disconnected external site as the discussion progresses. Arkose2 supports distillation of a completed discussion as Arkose 1.0 did, but it gains additional benefits by allowing a discussion to occur within the system.

Information Linkage

Arkose2 ties different types of related information together for better understandability and navigability. This includes connecting topics and subtopics, semantically related topics, generalized information, and contextualized information. Metcalfe's law (Wikipedia 2009d) states that the value of a network increases proportionally to the square of the number of linked objects. Linking different types of information may also have advantages. First, it helps discover information. A user looking at one topic may follow linked subtopics and example cases that would be difficult to find otherwise. Second, linked and structured information may be more useful and acceptable to the community (Judge 2007) and may be easier to understand the overall structure as found in the evaluation of Arkose2. Third, it provides an opportunity for organic organization of the topics by allowing related topics to grow together. In the current question-answering or discussion sites, a question or topic a user enters is mostly treated independent of the existing topics in the system. Arkose2 allows a user to initiate his question or topic from an existing topic, which allows easy linkage among the topics.

Over time, this may lead to an organic growth of related topics because of how a network grows with preferential attachment, allowing social navigation as well as semantic navigation (Dourish and Chalmers 1994) (See more detail on this in the Topic Nebula section.)

Since the linkage of information is nothing new, as the current hypertext systems or Wikipedia implement it, I do not claim this as a new and novel idea. Rather, I take this design as a necessary component for Arkose2 from which it can gain an additional benefit.Arkose2 is designed based on these two principles supported through information visualization and information retrieval techniques. In the next section, I discuss the conceptual architecture of Arkose2. Its implementation details based on the conceptual architecture follow it.

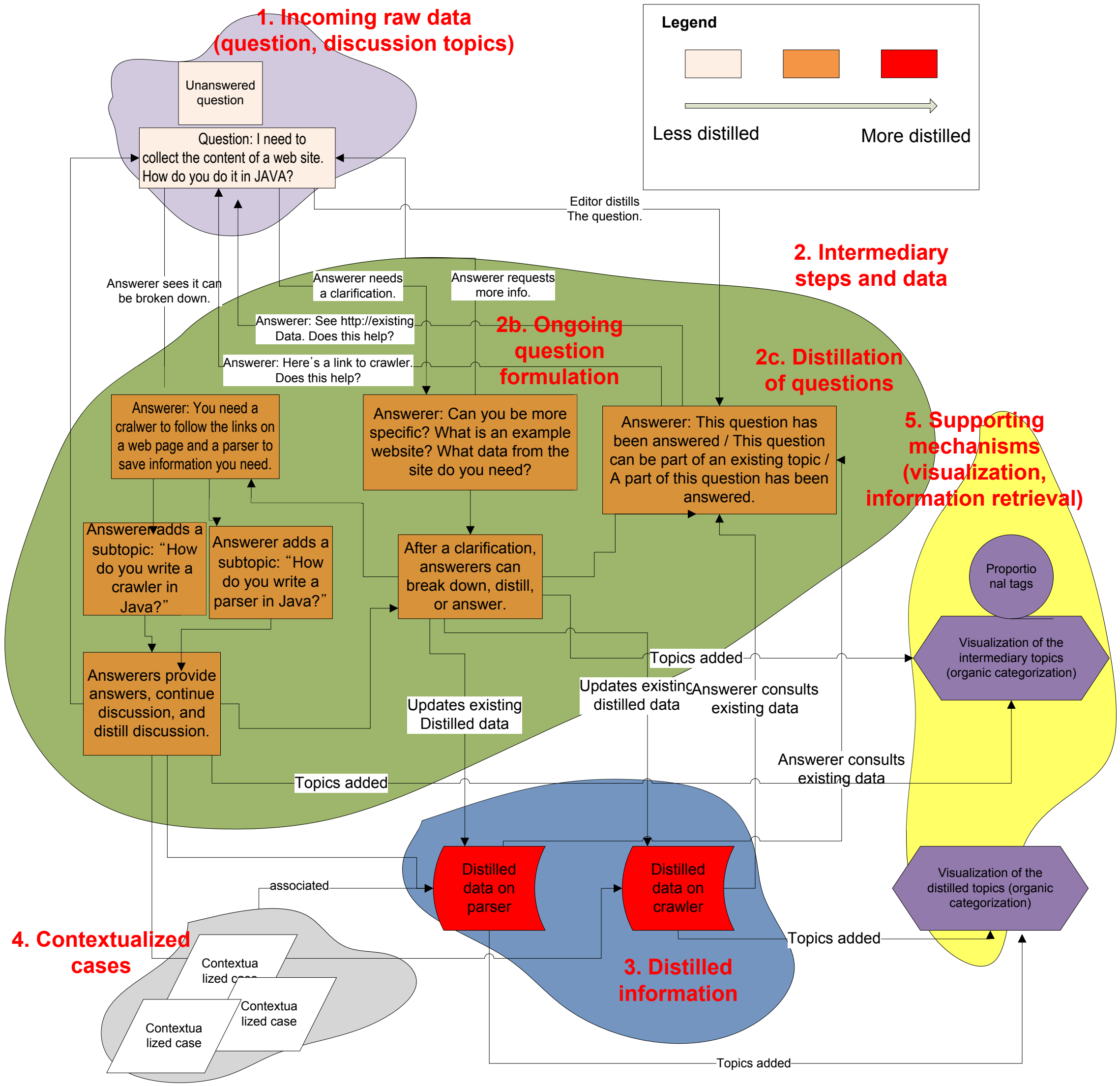

Figure 5-1 The conceptual architecture of Arkose2. 1) Incoming data, 2) Discussion space where question asking and answering, and distillation occur, 3) Distilled information in the Summary Editor and/or wiki, 4) Related information connected to the distilled space, 5) Various visual aids and supportive tools.

Conceptual Architecture

Arkose2 has five different components. Firstly, it handles the incoming data. Question-answering and discussion data may be imported from an existing community as with Arkose 1.0. Or users may start a discussion directly in Arkose2 for more flexible knowledge generation. Secondly, the visual discussion space allows users to handle complex questions, formulate questions, and distill the discussion as it is happening. Thirdly, distilled information may reside in the Summary Editor Space, a wiki or personal view. The discussion spaces and wikis are linked together in such a way that distilled information from the discussion spaces are stored in a wiki and users may continue the discussion or start a new one from the wiki. Fourthly, distilled information is connected with various meta-information such as contextualized cases of generalized information, helping to make the information multi-faceted and hopefully more useful. And lastly, a variety of visual aids and information retrieval capabilities support users' discussion and distillation activities.   In this section, I will go over each subpart of the system and discuss its features and relevant scenarios.

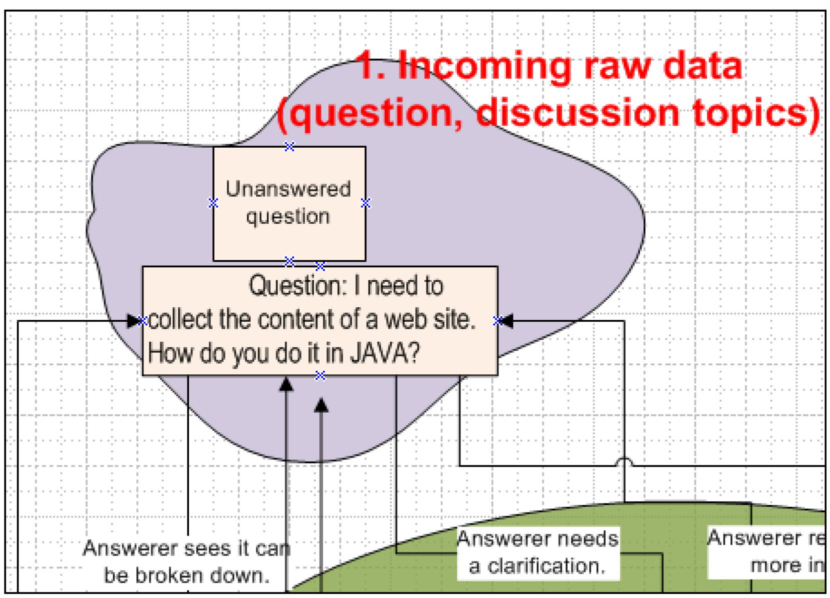

Figure 5-2 Incoming question or discussion data into the system.

Incoming Raw Data (Questions and Discussion Topics)

As with Arkose 1.0, Arkose2 can import existing discussion materials from an external system given that an intermediary parser program is written that translates the discussion space into the XML based TreeML format (Fekete and Plaisant 2003). But the real benefit of Akose2 may be fully experienced when a discussion or question-answering activities occur within Arkose2. This enables flexible information generation through the provided user interfaces with the support of a variety of aids. Ongoing distillation as the discussion occurs may also guide further discussion as discussed in the previous section.

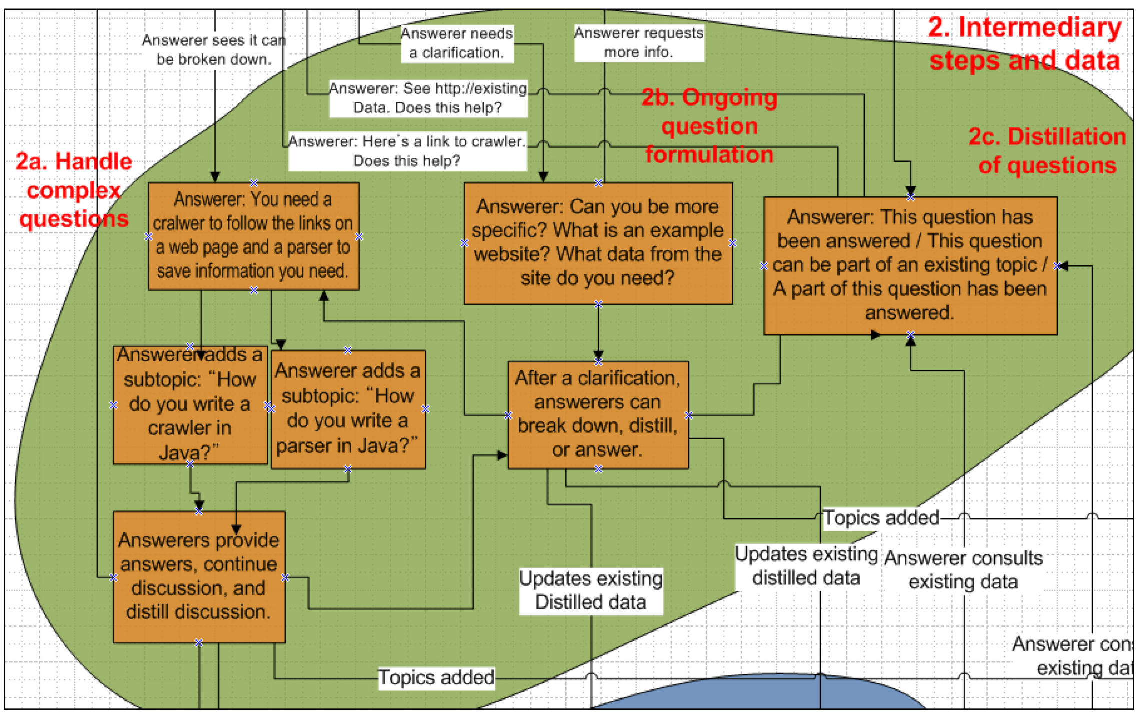

Figure 5-3 Conceptual view of the activities in the discussion space in Arkose2. This is where a complex topic is broken into smaller subtopics (2a), an unclear question is clarified (2b) and the distillation of the questions and answers is performed.

Section 1 in Figure 5-1(as well as the partial view in Figure 5-2) conceptualizes the data entry point to Arkose2. Incoming data include questions or requests that a user enters. As discussed in the introduction, Arkose2 can handle different types of questions, some of which not handled well by a traditional question-answering community or a wiki. Examples of questions that can be handled by Arkose2 include, but not limited to:

a. *A simple factual question: "What is Java Eclipse?"*

b. *A how-to or procedural question: "How do you debug a program in Eclipse?"*

c. *An opinion-seeking question: "What is the best IDE for working with Java?"*

d. *A complex topic with potential subtopics: "I need to collect the content of some web sites. How do you do it in Java?"*

The type of questions and discussions that accumulate in a community may differ according to the domain area, expertise level, the size, culture and norm of the community. Arkose2, as a prototype system, stays agnostic about these differences among the communities and the policies that govern information generation behaviors. Rather, it focuses on what it allows the users to do once the data is entered in the system. One notable advantage of Arkose2 over other knowledge building communities is the explicit support for handling complex questions such as *d*. This is discussed more in detail in the *Handling Complex Questions* section.

Intermediary Steps and Data

This is where data goes through continual transformation through partial answering, question and answer distillation, feedback loops, and potential hand-offs. Figure 5-3 conceptualizes some of the activities for the discussion data Arkose2 supports within the system. Many paths of transformation are possible once a question enters the system. A question may be sufficiently answered by one or two users if it is a simple question. When an asker does not provide enough information for the question to be sufficiently answered, an answerer may request the asker for more information (2b). An answerer may break a question into manageable sub-questions or sub-topics within the discussion space or they may feed back into the system to start a new topic (2a). The subtopic, if it is a substantially large topic, may start its own discussion space or may be moved to a new space as it grows. This deconstruction of a topic into smaller subtopics may make it easier for users to understand the discussion and distill it.

Handling Complex Questions

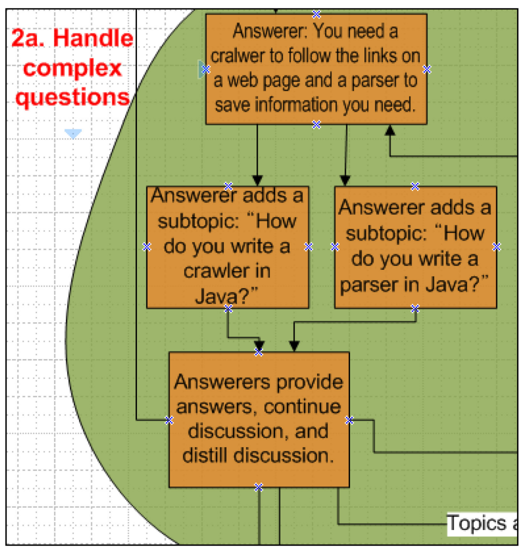

Figure 5-4 Handling complex questions. A question may be broken down into subtopics and discussion may continue for each subtopic. If a subtopic itself requires much discussion, it may be further broken down or reside in a separate discussion space.

Consider the question, "I need to collect the contents of some web sites. How do you do it in Java?" There are two things to consider for this type of a question. One is that it does not provide enough information about the asker's situation. Under what context and requirements is the user asking the question? (This will be discussed in the next section, *Ongoing Question Formulation*.) Another concern is that the question may require more information to be sufficiently answered. For example, an answerer may suggest that the question is actually a two-part question about a Java crawler and a Java parser. Another user may also suggest that knowledge about both HTML DOM structure and regular expressions is needed in order to build a successful Java parser. In a question-answering community, this type of question is not handled well; a question that requires too much effort to answer is often ignored or given an insufficient, short answer. This again, may be due to the culture of a particular community that frowns upon users who ask questions out of the norm or who did not "do their homework" before asking. But the limiting nature of the interactions provided by a question-answering community, as well as the linear and partial presentation of a discussion in a forum, make it difficult for the users even when the questions are considered acceptable. While not every question is suitable for this process, certain types of questions may be broken down into smaller questions or topics. Arkose2 provides user interface mechanisms (Figure 5-4) that allow users to handle complex questions by breaking them into manageable size pieces and feed each sub-question or sub-task back into the system. More details about the implementation can be found in the Handling Complex Questions section.

Ongoing Question Formulation

It is not trivial for a novice user to formulate his question sufficiently in order to search for the right information. An asker may not know what information is necessary for an answerer to be able to provide an answer. A question such as "I can't compile my Java code. What's wrong?" does not tell enough information about the user's specific situation. An editor may request additional information, such as the source code, error message, or version of Java environment. The process may occur several times until the user gets a satisfactory answer. This interaction between an asker and an answerer can usually be handled in an existing discussion forum, but not in a way that is cleanly separated. Since a discussion forum presents its data in a linear and threaded format, these interactions are intertwined with actual answers. This makes harder to extract the needed information without further organizing. Furthermore, the user interfaces in a question-answering community such as Yahoo! Answers limit the number of the posts a user can leave under a question, which makes the interaction among the users very difficult.

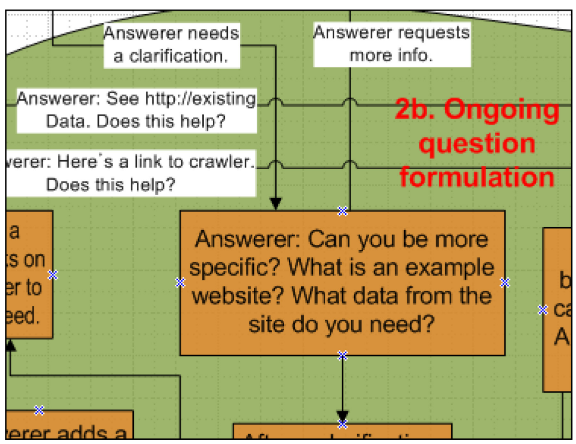

Figure 5-5 Ongoing question formulation. A question without sufficient information may be further developed through the users' interactions.

Another type of question formulation is re-formulation of the question. It may be possible that the asker is asking the wrong question. For example, an editor may say "What you really want to ask is this…", and resubmit the question in its re-formulated form. A novice in a domain may not know how to formulate the right question or the right vocabulary, and this may be an area where other more experienced users can help, instead of just answering the question.

The process of question formulation can be seen in an existing discussion forum; in Arkose2 however, it is augmented with flexible user interfaces that separate it out from

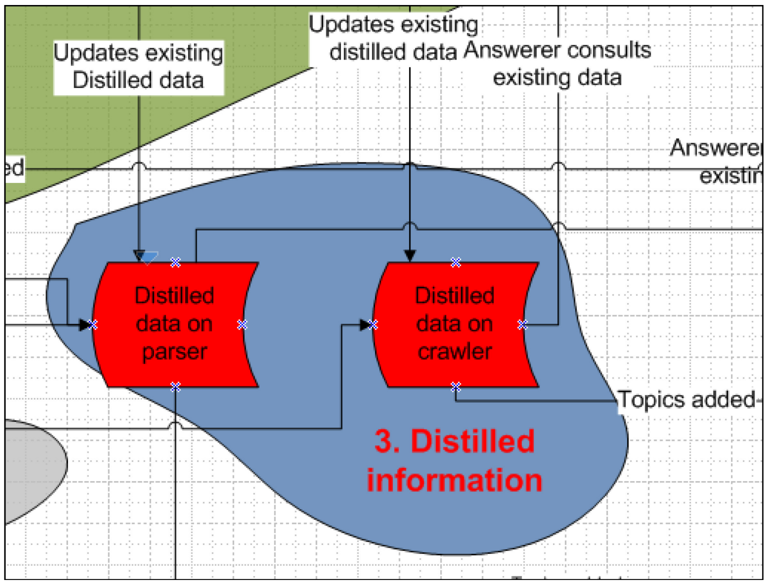

Figure 5-6 Distilled information may reside in the summary editor, wikis, and personal wikis. Contextualized information, original discussion, and other meta-information may be linked.

the actual answers. Arkose2 gives users flexibility to add answers in any stage of a question formulation. It would be too rigid if users had to wait until the question was finalized. It would also be unclear when the asker provided enough information for the question to be finalized. Rather, Arkose2 allows users' answers and clarifications to happen at the same time but separately, gradually making the separation clearer if needed. While the question is being formulated, other processes such as answering, distilling, and breaking the question down into smaller topics may occur as well.

Distillation of Questions and Answers

Through the process of ongoing distillation, multiple pieces of distilled information on a topic may exist in the system. To ameliorate this condition, Arkose2 allows a user to open up a wiki page on the topic and save the distilled information.

The distilled information object can also have a topic request feature. During the distillation of the discussion space, topics may be identified that need further discussion or more data. A user may specifically ask other users to fill in the "holes" in the topic by feeding the request back into the system (an *active* request) or by leaving a note or comment in the distilled summaries (a *passive* request). For example, the user may ask others to provide a summary, outside resources, code examples, and so forth. This will help make the distilled information object responsive to the emerging needs of the community, which is difficult to do in a traditional discussion forum.

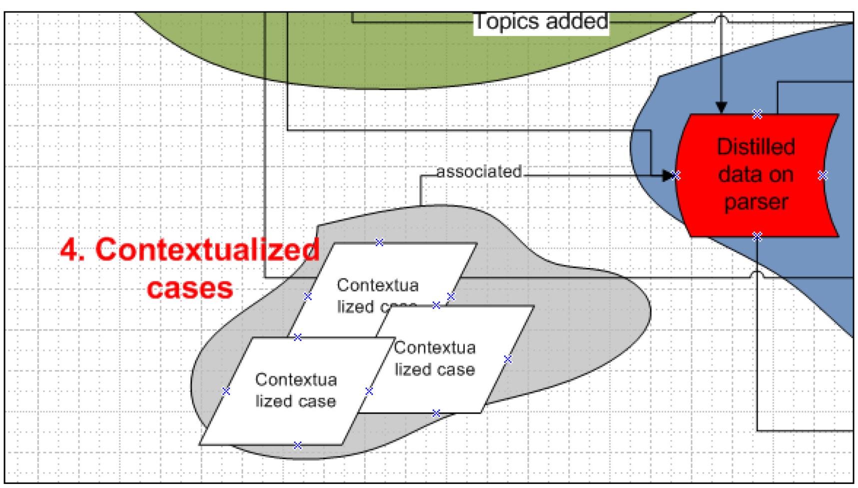

Figure 5-7 Specific contextualized cases are linked to general information. For example, several discussions on changing brakes on specific car models may be linked to a wiki on car brakes.

Contextualized Cases

Along with distilled information, there are many "instances" of specific cases in the system, which are related to a topic. For example, a number of users may ask questions on installing the Java SDK for an operating system. While the general instructions should cover most cases, there are some special cases that require further instructions. For example, different distributions of the Linux operating system may require a special "tweak" in order for it to work. Conversations generated by users may be used to meet personalized needs (Hansen 2007). Contextualized information may be much more valuable than non-contextualized information, if the context is a correct one (Ackerman and Malone 1990 Ackerman and McDonald 1996). Associating specific installation cases along with a general instructions page may help users better by providing contextualized information on the topic.

Even within a specific case, there may be many different scenarios. For example, compiling an open source program on a Linux operating system may require a specific explanation for each of the OS versions. If there are many requests for a specific case, it can also be turned into a more generalized information object for those users. Specific cases may be a result of the diverse questions that users ask. Or a user may notice an emerging topic and anticipate their needs.

Thus, there are two ways contextualized information and generalized information evolve together. One is from the generalized information to the attaching of contextualized information (a *top-down* approach), and the other is from many related specific cases emerges a general topic (a *bottom-up* approach). Arkose2 supports both. There are several places this can be done. One is when a user searches and initiates a question or topic in the Topic Nebula (Figure 5-25) or in a discussion space or a wiki page. An example is provided in the usage scenario in the Usage Scenario chapter.

Supportive Mechanisms

Users' activities discussed so far are assisted by a number of supportive features in Arkose2. These include information visualization and information retrieval techniques. The visual aids help ease the cognitive burden of a user when he navigates in the information space or distilling information. A variety of data are presented in a meaningful way that helps the user keep track of the progress and related information. The underlying information retrieval capabilities such as search with keywords, content analysis, automatic clustering and other flexible tools help search and location information. I will first present a usage scenario followed by details of the implementation in the following sections.

Usage Scenario

In this section, I will briefly go over a usage scenario of Arkose2 in order to better illustrate how different parts work together to support user activities.

The Problem

John is a Linux newbie who wants to install OpenSuse Linux on his Windows XP laptop. Although he has installed Linux before, it was Kubuntu, a different distribution of Linux. As well, he installed it on an empty computer, not one that already had an operating system. These factors may complicate the situation and may lead to a loss of data if he does not plan the installation carefully.

Finding Relevant Topics

He runs Arkose2 that has been used by a Linux community for some time. He searches for the words *install Linux* in the Topic Nebula UI (Figure 5-25) that returns visual representations of questions, discussion topics, and wiki entries that have the search terms. He can easily tell which ones are questions, personal wiki pages, or wiki pages by the shape of the topic node. He can also estimate the potential usefulness of each topic by looking at the size of it. The bigger the topic node, the more posts, entries, and information are in it. Some topics are linked together as specific cases of a general topic, or contextualized cases, or as subtopics to another topic. This helps John understand the overall structure of the topics space and how different topics are related.

Posting a Question

He browses a few seemingly related topics and gets some useful information on installing Linux, but does not find any information about how he might install Linux on a computer with another operating system. He selects a topic "How to install Linux" and selects the option "Post a question/topic related to this topic." In current question-answering systems or discussion forums, each question asked is independent of other questions. In Arkose2, however, a user may initiate a question from an existing topic. This automatically relates John's question to the existing topic and a link is created. As John does not know how his question might evolve as users answer his question, he does not specify the relationship type yet. He asks the question in Arkose2.

Discussion Space Interactions

The next day John returns to Arkose2 and sees his question, and indeed there are many replies to the question in the discussion space (Figure 5-8). Some users left clarifying questions (Figure 5-5) in order to learn John's specific type of operating system and laptop. Other replies also provide links to some web pages that may have some of the drivers. One user left a post that he might need to partition his hard drive and save the Master Boot Record in order to safely install Linux. The user has created two subtopics in the discussion space, one for each (Figure 5-4). Some users left information under the subtopics, which John reads and finds useful. He notices these two subtopics automatically show up and are connected to his question as "subtopics" in the Topic Nebula.

Distilling and Saving Information

He feels he has enough information to finally install Linux on his computer. He knows he will have to do it again in future, so decides to save some of the information he found useful in the discussion space. He creates a few summary nodes in the Summary Editor Space (Figure 5-8) and saves that to his wiki page (Figure 5-23). He wants to organize his information space, and tags some of the posts using the Proportional tagging UI (Figure 5-21). As he creates these tags, the posts are categorized under the tags. When some categories contain more posts than the threshold John has set, Arkose2 notifies John to re-categorize them into new sub-categories that Arkose2 has suggested to him. This helps him better organize his information space (Figure 5-19; Figure 5-20).

Further Searching and Distilling

John is about to install Linux on his computer. Then he remembers to back up his hard drive before writing anything over it. His hard drive is quite big and he is not sure what the proper way to back it up is. He revisits Arkose2 and searches for backup options. Indeed, it is a popular topic and there are many discussions, questions, and wiki entries in the Topic Nebula. By looking at the size of the topics, he finds that a particular topic has garnered many replies. He decides to read further and goes to the topic's discussion space (Figure 5-8).

Using the Supportive Features

The discussion space, as he found in the Topic Nebula, has a large quantity of posts. While John is interested in reading most of the posts, he wants some guidance in picking out important posts. He notices that some other user has already grouped several posts together under one topic "Backup software". That user has requested that someone to make a summary of the group since it contains useful information (Figure 5-15). The scaffolding makes it easier for John to quickly tell what some of the subtopics are. While John might come back to the topic and summarize eventually, he wants to continue looking for information on backup options for large hard drives.

While part of the discussion space has scaffolding, most of the posts are not organized yet. Instead of reading one by one from top to bottom, John turns on the automatic clustering feature of Arkose2 (Figure 5-16). Arkose2 has highlighted posts with similar content in the same color grouping. While not all posts are clustered correctly, it gives John a reasonable idea which posts are relevant and how to navigate from post to post. The tree representation of the discussion posts also gives him an idea where a topic branches off.

He finds some helpful information about backup options. He summarized a few findings and saves them into his personal wiki page.

Organizing

He goes back to the discussion space of his question. Users have posted more replies in the meantime. One subtopic created by a user, the partitioning of hard drive, has gained so many posts that it was crowding the space. John decides that the subtopic needs its own discussion space. He selects the option for giving a subtopic its own discussion space, and then a link node replaces the posts under the subtopic (Figure 5-12).

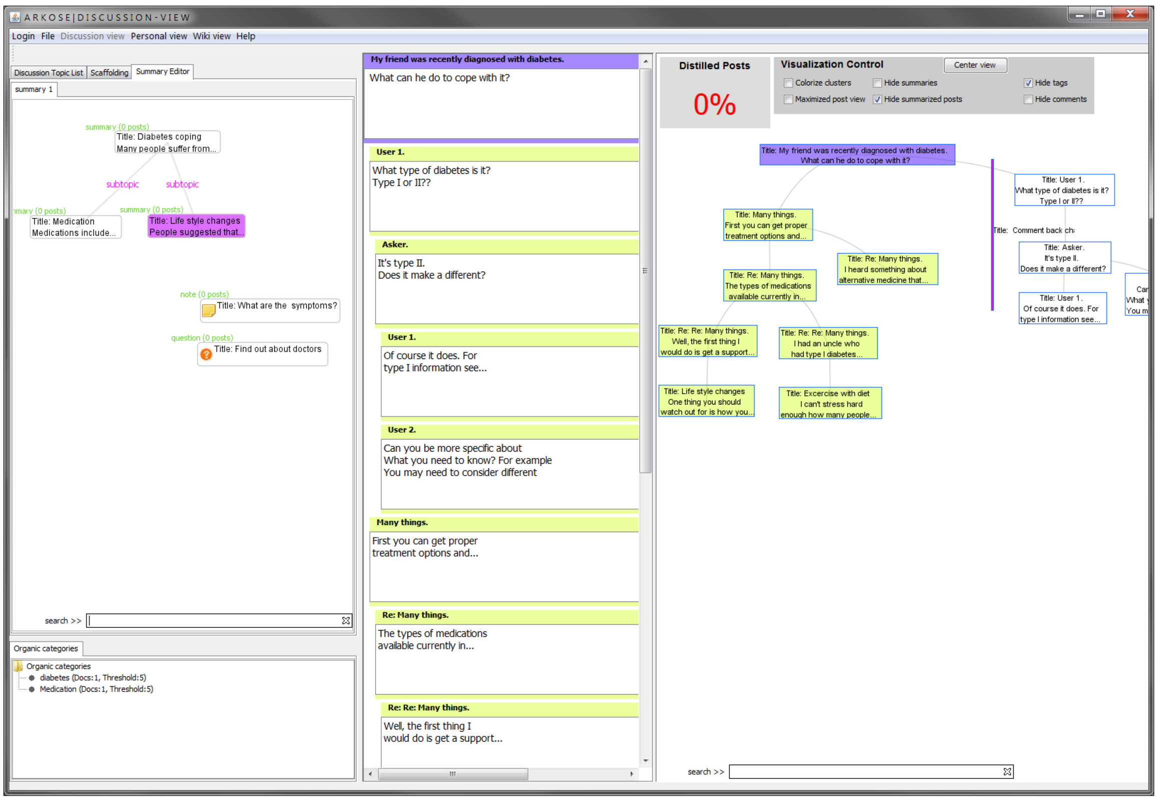

Figure 5-8 The main view of Arksoe2. Left column: Topic list, Summary editor. Center column: the Thread view. Right column: the Visual navigator. Wiki pages and personal wiki pages, as well as other supportive capabilities are accessible through menus and popups.

After a few days, John has another related question and opens up the Topic Nebula. He sees his question and the subtopics created by another user reside in the Nebula. It is easy for him to specify the relationship now. He links his question as a subtopic to the existing topic from which he initiated the question. He also sees a few related questions on installing Linux on specific laptops, and links them as contextualized cases of the general topic (Figure 5-25).

Discussion

The above scenario describes an experience a user might have, and it points out the capabilities and usefulness of various components of Arkose2. The supportive capabilities help make information generation, distillation, and organization more effective. In the following sections, details of each component are discussed.

Implementation Details

In this section, I discuss the implementation details of Arkose2 based on the conceptual architecture explained in the previous sections. First, I explain the technical details of Arkose2 including the programming language and libraries used. Second, I describe the components and features of Arkose2 that fall into three capabilities (See Table 5-1). The first capability allows handling of complex questions that require multiple answering cycles. The features for the capability include user interfaces that allow breaking of a topic into subtopics and a “back channel” conversation separate from the actual answering activities. The second capability is the better organization of information by allowing proportional tagging and organic categorization of changing content. And the third capability is the linkage of information by allowing different types of information to be connected such as linking topics with subtopics, topics with supporting evidence and examples, contextualized information with generalized information, and with wiki pages and discussion spaces that reside separately.

Technical details

Arkose2 is programmed in Java and uses a variety of components. For its visualization, Arkose2 uses the Prefuse toolkit (Heer 2007) that provides basic graphing capabilities and the Processing toolkit (Fry and Reas 2001) that provides an easy way to create customized visualization. Prefuse and Processing focus on information visualization and presentation, and they lack some capabilities needed in order to support discussion generation and distillation. Therefore, several key functionalities were added in Arkose2 on top of the basic functionalities in these toolkits. These include a data editing capability, support for more data types and user interaction capabilities, connections to a backend database, and integration of content analysis capabilities.

Some of the underlying information retrieval capabilities in Arkose2 are provided by the LingPipe library (alias-i 2008) . LingPipe is a toolkit for processing textual information using computational linguistics techniques. While Arkose2 uses LingPipe as an underlying content analysis tool, it also has the flexibility to work with other Information Retrieval toolkits. It may be possible to improve Arkose2's performance by replacing the LingPipe in future.

A backend MySQL database stores all the data in Arkose2. The interactive nature of Arkose2 makes it very difficult to store and load partial data as needed from a database. For example, when a discussion or summary node is created or links are created among them, only the changed parts should be sent to the database. There are t the data types, and the mappings among them, of which Arkose2 needs to keep track.

Arkose2 uses the Java Persistence library (O'Conner 2007) to handle the database interactions. In the real world settings where Arkose2 would be used by many more users and with more complexity, it would be prudent to use a library more suited to the heightened complexity, for example, the Hibernate library (http://www.hibernate.org/).

In the following sections I discuss the three capabilities of Arkose2 and their corresponding features.

Visual Navigator

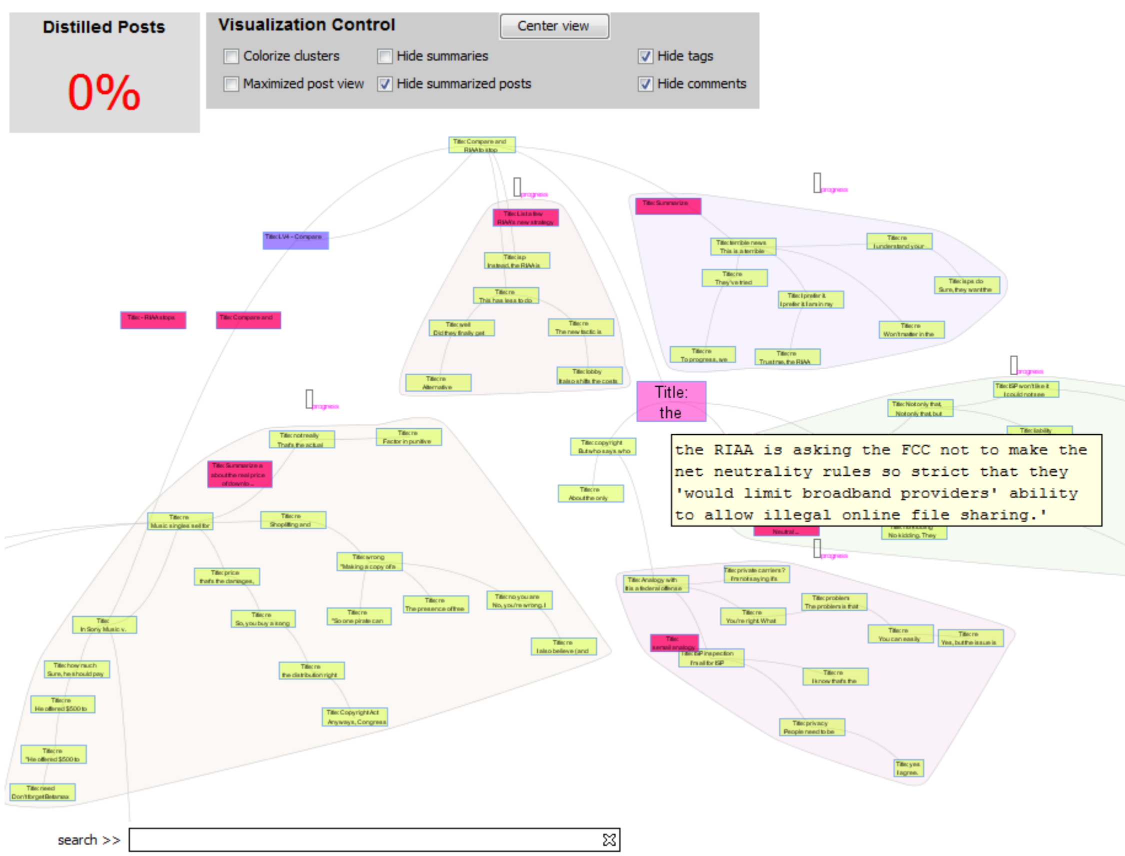

Figure 5-9 A partial view of Visual navigator. It supports discussion discourse, summarization scaffolding, editing, and a variety of progress awareness. Raw discussion data, user interactions through comments, summaries, keywords, and other meta-information co-exist in the space.

The Visual Navigator (Figure 5-9) is the substrate in which discussion posts, summaries, and various meta-information co-exist and are gradually transformed into a distilled state by the users. Users can read, manipulate, and organize discussion data in the Visual Navigator. It is integrated with the Thread View (middle column in Figure 5-8), which presents the threaded conversation in a similar way as an online discussion forum does. The Thread View provides very simple user interfaces with which users could be already familiar, and it allows a quick view of several posts at once. The Visual Navigator, on the other hand, supports richer interactions and operations in it but is also more complex. The design rationale is that users might prefer the simpler Thread View when performing a simple task or reading a small discussion space (as found in Smith and Fiore 2001 and also in the evaluation of Arkose2), while they might find the

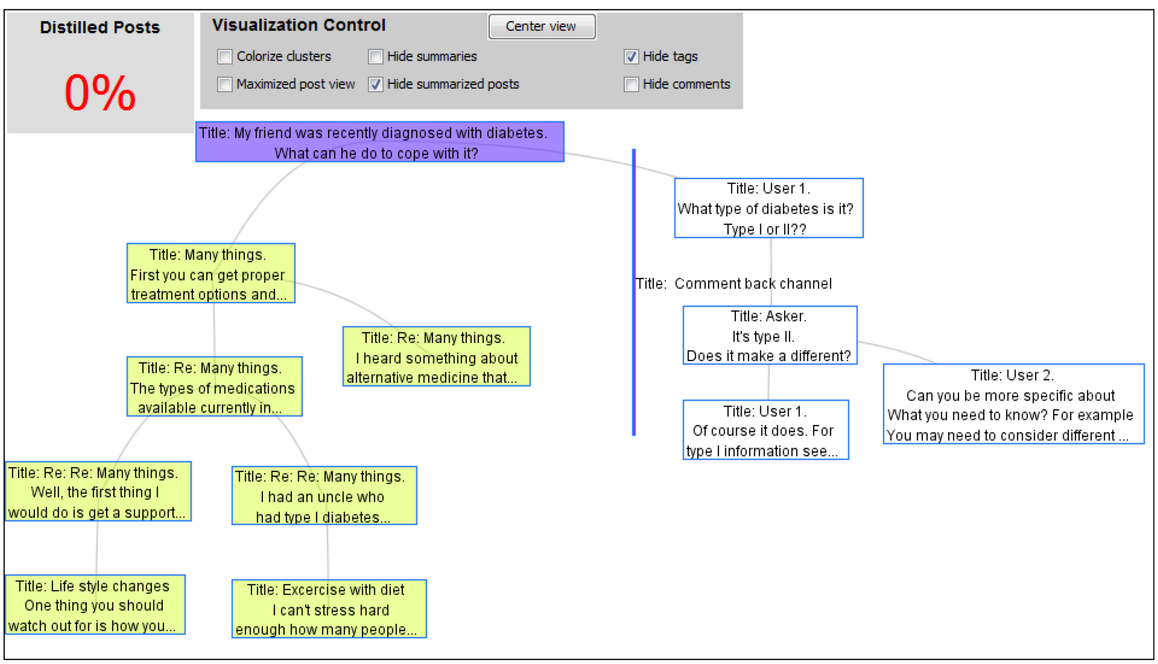

Figure 5-10 Conversation "back channel" for question clarification and formulation. On the left, users may continue discussing answers. On the right, clarification may be done separately from the actual discussion.

flexibility of the Visual Navigator more useful when the discussion space become bigger and more complex. The combination gains an additional benefit by visually connecting together the discussion data in both views as users should find it helpful to see the same data in multiple ways (Smith and Fiore 2001).

The Visual Navigator is also one of the main components of Arkose 1.0. The Visual Navigator in Arkose2 has several improvements and changes over the previous version. First, it now allows a discussion to occur in the Navigator (e.g., Figure 5-10, Figure 5-11) rather than simply presenting a discussion space as Arkose 1.0 did. This gives Arkose2 an opportunity for a flexible generation of discussion data by integrating the discourse activities with a variety of supportive capabilities, and synergetic efforts with ongoing distillation. The visual navigator hosts a number features and tools that help make users' distillation and discussion activities easier. In the following sections, I discuss these supportive features in turn.

Question Clarification and Formulation

Arkose2 cleanly separates the actual answering posts from other comments and clarification as opposed to current discussion forums that present them intermixed and in a linear fashion. Although the interactions among the asker and other users may not always be clearly separable from the answering activities, having a back channel where users can discuss other matters allows interactions to shape better. The back channel conversation may occur at any point in the discussion, not just for clarifying the initial question.

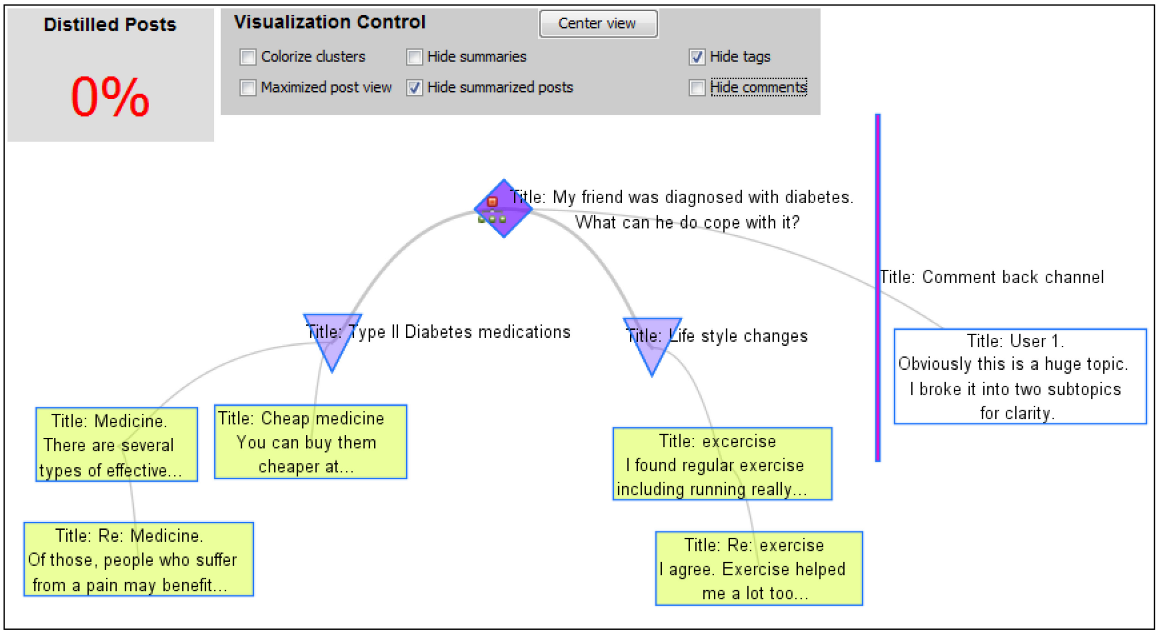

Figure 5-11 Handling a complex question. The interfaces allow breaking a question or topic into subtopics. This allows a more organized discussion. The diamond shape represents the original question that is divided, and each of the triangles below represents a subtopic. A user may break any subtopic into more than one sub-subtopics.

Suppose an asker entered the following question in Arkose2.

```
My friend was recently diagnosed with diabetes.  What can
he do to cope with it?
```

Since the asker does not really provide the details of his situation, users need further clarification to answer it sufficiently.

Handle Complex Questions

Arkose2 sets itself apart from other knowledge systems by explicitly supporting complex questions. A complex question may be broken into several subtopics for a more organized discussion. The following interactions show how it can be done in Arkose2.

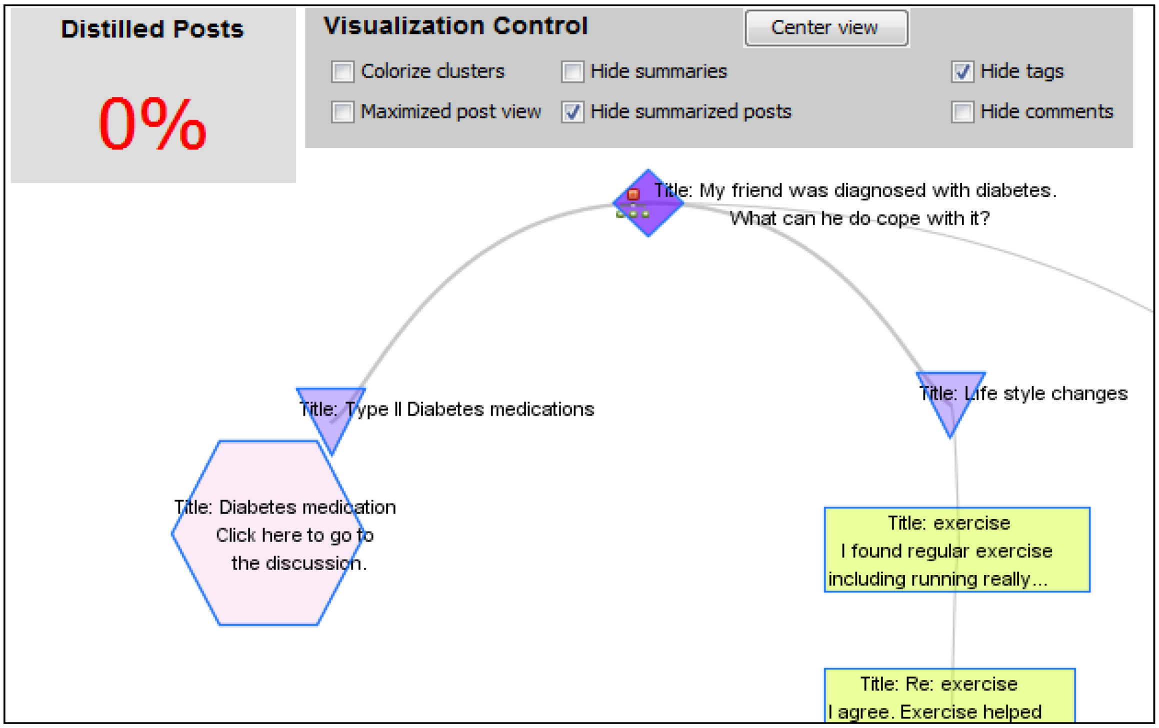

Figure 5-12 As a subtopic becomes a huge discussion on its own, it is replaced with a link to its own discussion space. Clicking the link node (represented as a hexagon) will move the user to the discussion space.

Consider the earlier question the asker posted in Arkose2.

```
My friend was recently diagnosed with diabetes.  What can
he do to cope with it?
```

Users may realize this is not a question for which a simple answer is sufficient. There are many subtopics to consider were the question to be answered fully. In real world settings, a question that requires too much information might be ignored or answered with insufficient information. But occasionally there may be a question that the community feels might benefit the community and needs an in-depth discussion. It is likely that the question would have several subtopics, and a user may break the original question into them in order to better discuss it.

As will be discussed in the Topic Nebula section, Arkose2 ties different types of information together to help a user navigate and understand the topic space better. When creating subtopics from a topic the relationship is maintained and saved for future use. For example, a user searching for information about diabetes medication not only gets that discussion space, but also its parent topic (i.e., "how to cope with diabetes") and any other sibling or children topics.

Arkose2 also allows the subtopic to have its own discussion space. If the medication subtopic in the diabetes discussion garners a huge amount of discussion with many subtopics, it might be better to turn it into a new discussion space and linking it from the original discussion space. This way, a potentially very large discussion may be supported without worrying about crowding the space. A user can see the topic topology in order to determine where in the topic space he is in, or navigate to a related discussion space using the link.

A user may also see where in the related topic space he is. This may also help him understand the topic space.

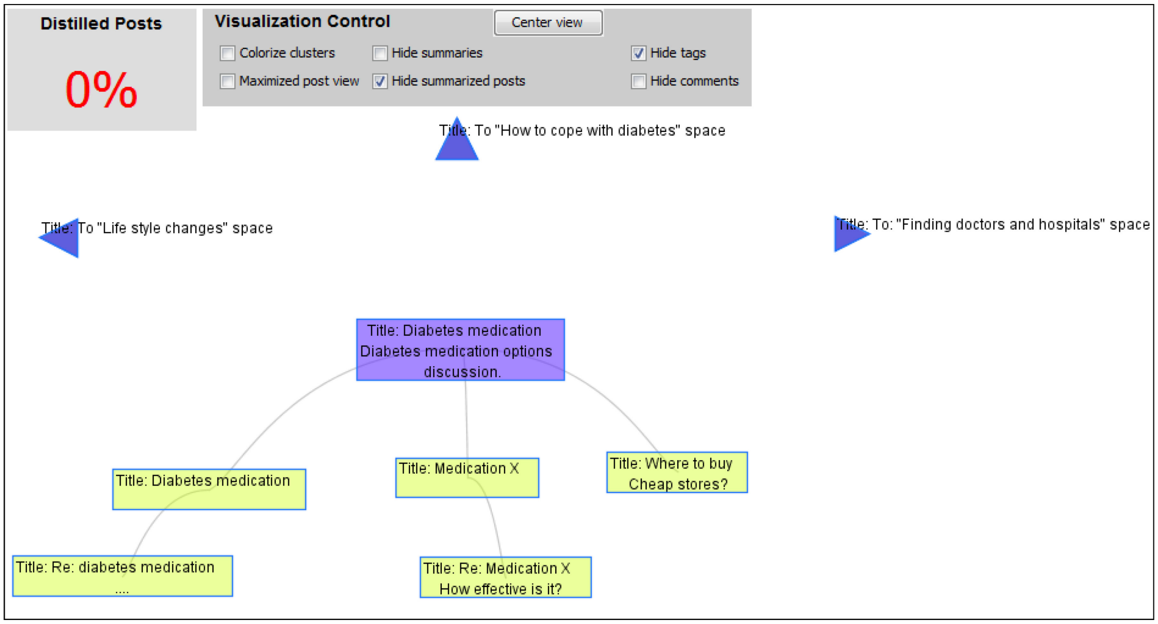

Figure 5-13 A subtopic with its own discussion space is shown with links to the parent discussion space as well as sibling discussion spaces. The peripheral awareness of related topics may help users keep the discussion relevant.

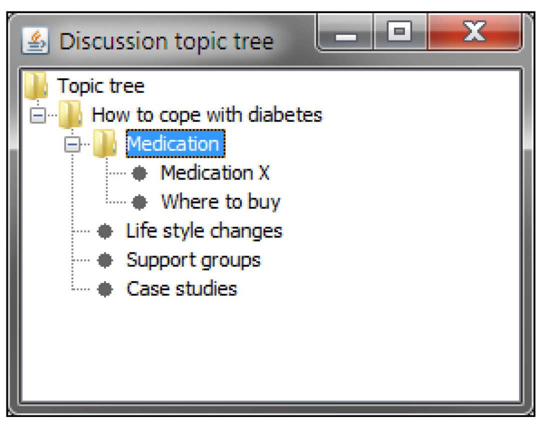

Figure 5-14 Discussion topic tree. It helps a user to understand and navigate the topic space better.

Scaffolding

When a discussion space is large and complex, sharing the workload may make it easier for the users to distill the space. A user who is experienced or has expertise in the domain area may know how to organize the space better, so as to provide guidance to others. Or a user who has read a certain portion of the discussion space that is under the same topic may feel the posts he has read may be summarized together. Both of these reasons may lead a user to create scaffolding.

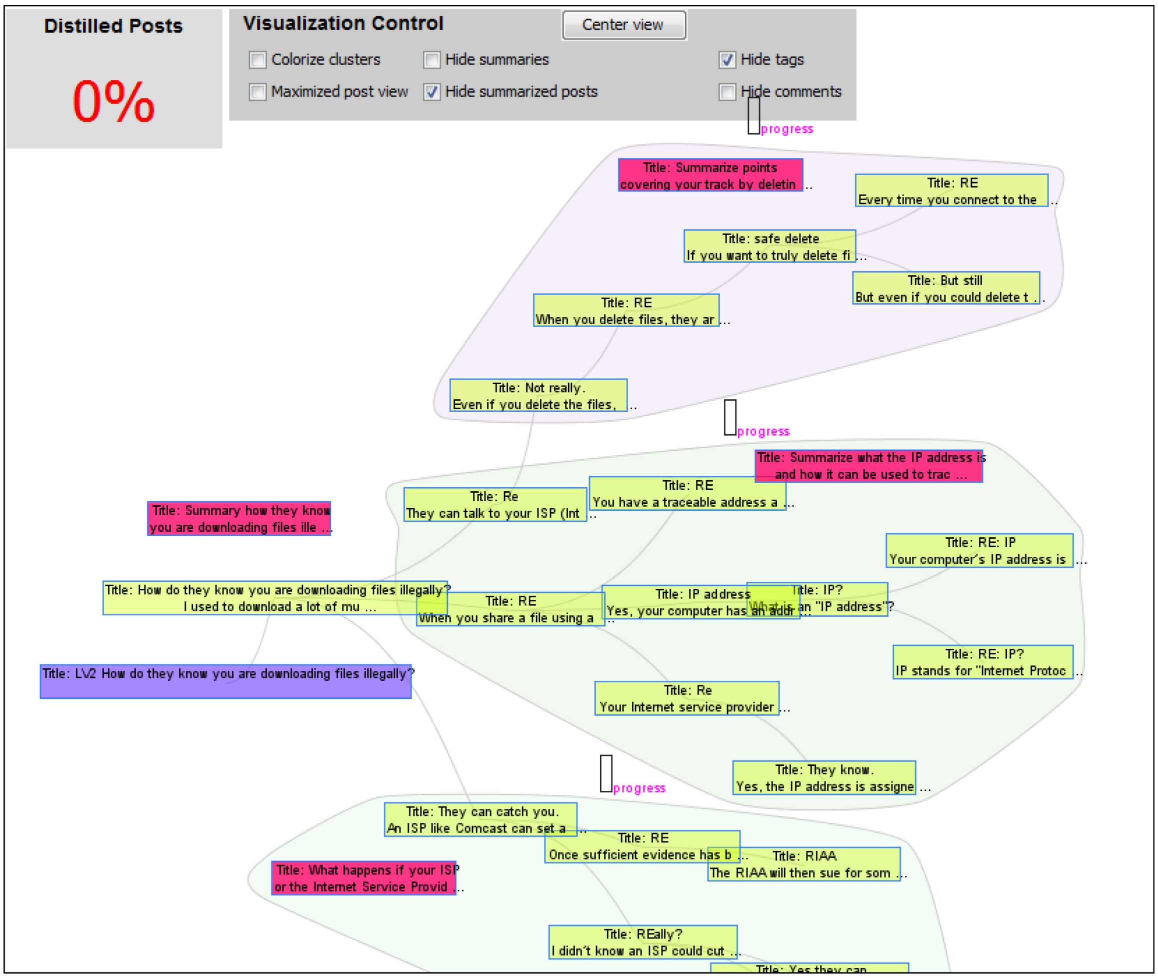

Figure 5-15 Scaffold groupings and requests. A user may group related posts together to summarize them under one topic. Scaffolding is an incremental process. Arkose2 allows an easy update of existing scaffolding.

There are a couple advantages of having a scaffolding structure, as revealed in the evaluation of Arkose2. First, scaffolding may break the discussion space into smaller subparts users can distill individually. This reduces the complexity of the distillation task, so no one user needs to handle the space all by himself. While there are some posts that may fall into more than one subtopic, users may be able to focus on each scaffold grouping and then make connections among them as needed as revealed in the evaluation. The second advantage is that users can quickly scan through the scaffolding nodes and tell what subtopics exist in the discussion space without reading many posts.

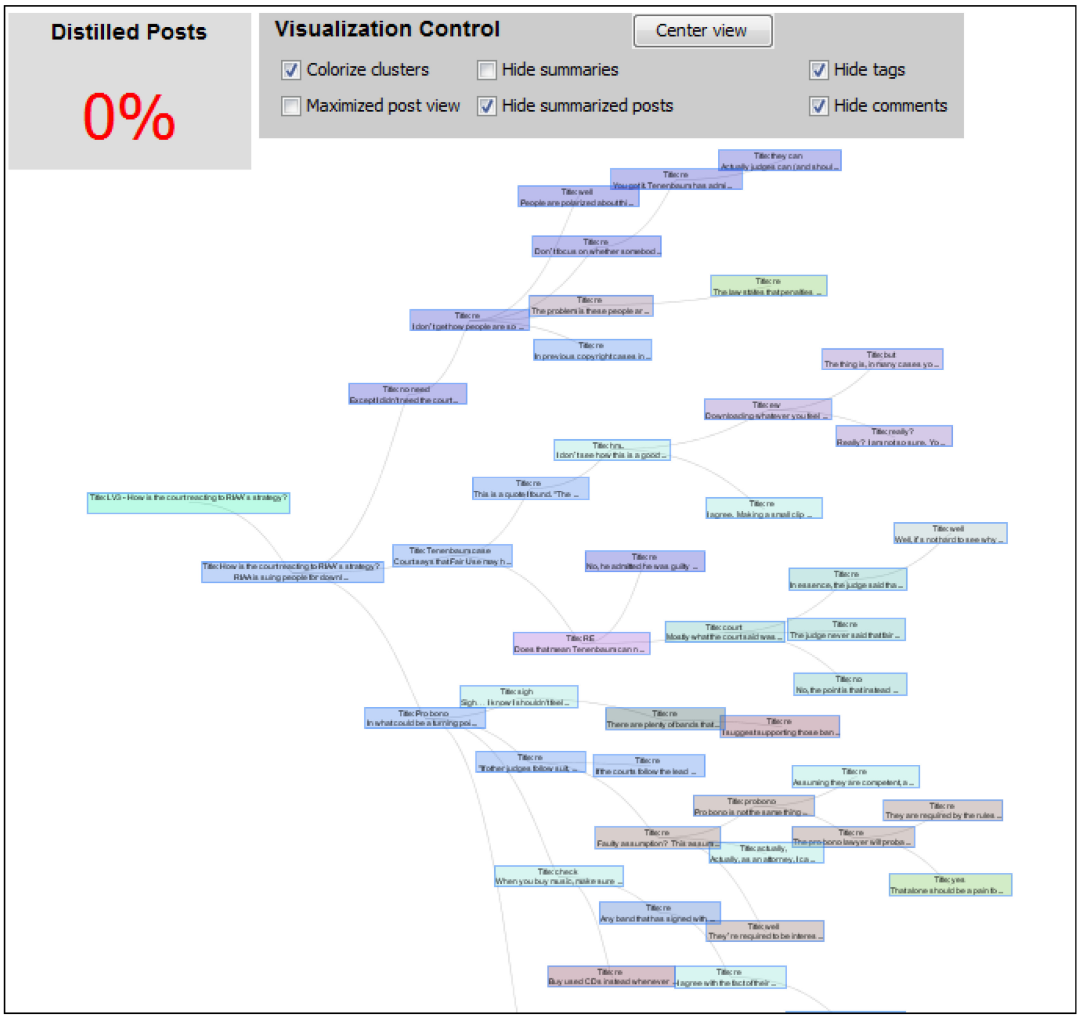

Figure 5-16 Automatic clustering of the posts in the visual navigator. This may help a user to understand and navigate in the discussion space.

Arkose2 supports an easy creation of scaffolding groups and nodes that visually groups related posts together and leaves a request for summarization. However, some users may disagree with the way the scaffolding is grouped, or may believe the subtopics are not important enough. As revealed in the evaluation, scaffolding is an incremental task that needs to update as new posts come in or users come up with a better way to organize the discussion. Arkose2 makes it easy to update the scaffolding.

As well, users might want to organize the summary space differently and wish to create alternative scaffolding structures. I am currently investigating ways for Arkose2 to support building multiple scaffoldings for a discussion space.

Automatic Clustering of the Posts

When the discussion space is large and complex, understanding and navigating the space become much harder. The tree representation of the discussion data provided by the visual navigator may help a user in looking for existing topics and understanding the overall shape of the discussion. For example, a post to which several children posts reply to may indicate it introduces an important or interesting topic. The overall shape of the tree may help divide the space roughly into related groups of discussion.

Arkose2 further supports the user by automatically clustering related posts together. This is done by calculating the cosine similarity of the contents of the posts, and binning related posts in groups. Posts in each group are highlighted in the same color when a user turns on the automatic clustering feature in the visual navigator, giving the user an idea which posts are related. Since the clustering process is done by a statistical algorithm, the accuracy of the result can vary. However, some users may find this feature helpful in understanding and navigating the discussion space, as revealed in the evaluation of Arkose2. Currently, only one type of clustering based on the content similarity values is used in Arkose2. I am investigating ways to implement more clustering methods, such as tag-based or author-based.

Figure 5-16 shows the clustering feature in action. As mentioned statistically similar contents are highlighted in the same color. Note that much of the grouping follows the natural discussion branches indicating a reasonable accuracy of the algorithm. Clustering gives a user an opportunity to discover related posts in other sub-threads.

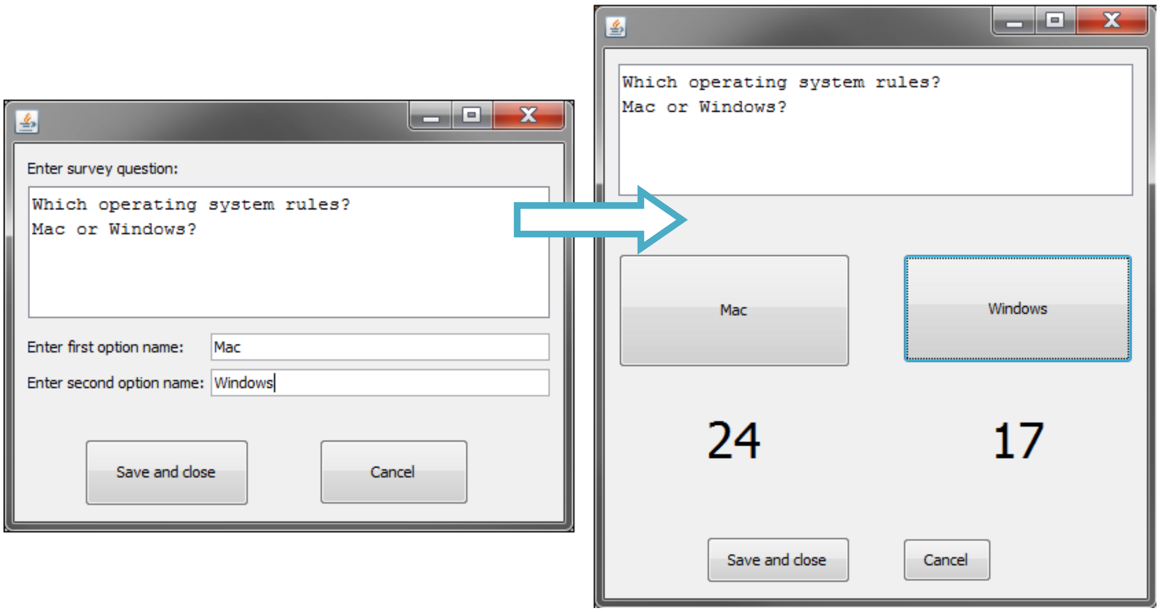

Figure 5-17 Survey widget creation in action. The left dialog is the survey creator, which automatically creates the right interfaces that allow users to vote for a choice by clicking one of the buttons. The result will be added to the discussion space in a post form.

Discussion Widgets

In this section I discuss a few other features within the Visual Navigator that can help users' tasks.

One is a discussion widget that can potentially replace the discourse of a certain discussion type. These may include a survey, rating, polling, and so forth. In a current discussion forum format, gathering and counting opinions from the discussion participants are cumbersome. Some discussion sites allow a sidebar polling feature where users may vote for a topic. However, this is not readily embeddable to the relevant section of the discussion itself. Arkose2 features several voting and polling widgets that can replace many posts that collect preferences. For example, a user may want to add a survey widget to gather users' preferences over two options. In future, more types of widgets may be created and used in the discussion space. Organic Topic Categorization

An information system may need to categorize information objects for organizing and structuring them. Developing a good categorization scheme *a priori* is hard because one cannot completely anticipate what information and how much of it will be created in the system. Even if one has developed an adequate categorization scheme, a static scheme may not meet the future needs of the users and changing content of the information system.

Arkose2 tries to ameliorate this problem by having an "organic" topic categorization. It is "organic" because the categorization scheme evolves based on the changes in the information space. For example, let us assume there are only the following three questions for Java in the system at the moment:

i. What is a good book on JAVA concurrency?

ii. I need to write a Swing GUI for a simple calculator. Where can I see some code?

iii. How do you program socket interactions in Java?

Since there are only three Java related questions in the system, it may be sufficient to have just one category, 'Java', to categorize all three the questions. However, as more questions enter the system the categorization scheme may need to *refactor* them in order to better organize information. If there are a number of concurrency related documents, it would make sense to create a "Java – Concurrency" sub-category.

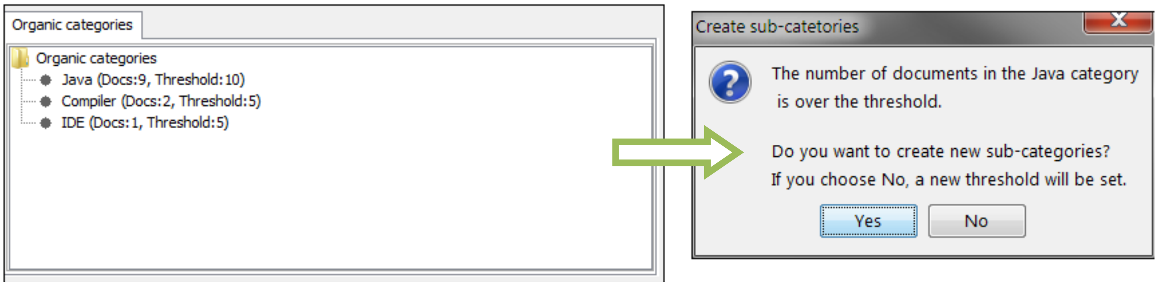

Figure 5-19 Organic categorization notifier in action. Once a category hits the threshold (in this example, *Java*), it asks the user to re-categorize.

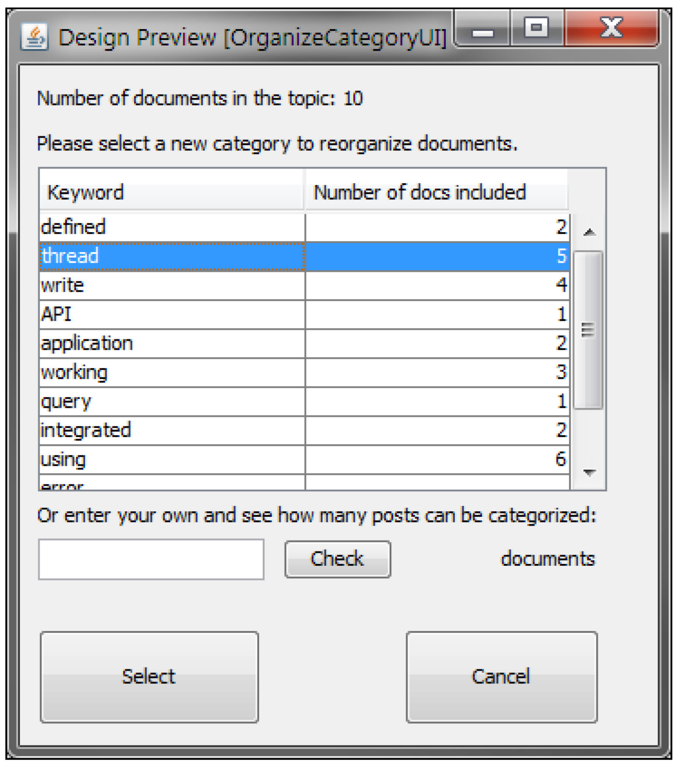

Figure 5-18 New category selection. Some words are suggested for a new category. A user may also try his own keyword. (Data shown are simulated for an example.)

While the decision to create a new category or a sub-category is ultimately up to a human user, the system can keep track of changes in the information profile of the system and bring it to the users' attention when it concludes there should be a new category. Without an automatic supportive feature, it would be very difficult to manually re-categorize the user's documents. First, a user may not know when he should re-categorize, and it can be difficult to determine this as well. Although he could go through all the categories and check whether some categories have too many documents in them, this is very cumbersome. Second, the user may be able to locate a category he wants to re-organize, but he may not know what new categories make sense for the documents he has. Third, even if the user knows what category to re-organize and what new categories to use, he cannot readily tell what documents will be affected and what the final

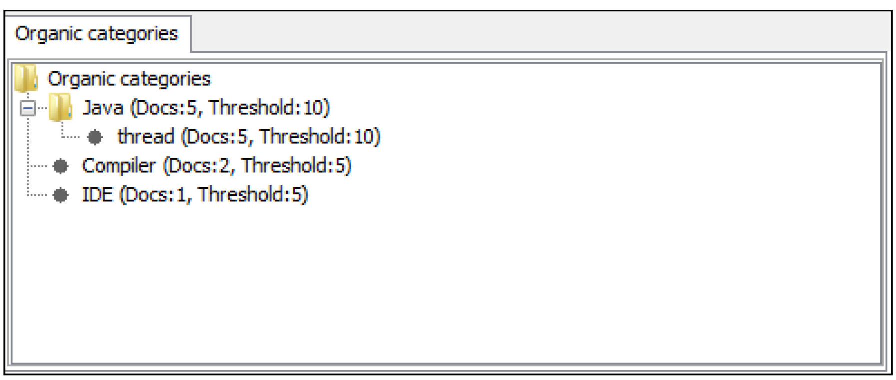

Figure 5-20 After re-categorization. A new subtopic *thread* includes the documents.

organization would be.

Arkose2's organic categorization system should ease the burden of having to re-organize manually. It ameliorates all three problems of manual re-organization. First, the system keeps track of how many documents are in each category, and when the number exceeds the user set threshold, it notifies and suggests the user to re-organize. The current implementation lets a user specify the thresholds, but a more complex algorithm could be specified to adjust the thresholds according to the size of the information space and number of categories. Second, a user may not know what new categories to use, but the system can help by suggesting keywords. The keywords are found the following way. First, the system performs a content analysis of the documents in the entire category and calculates the *tf*idf* values for each word in the documents. Then it presents the top 10% (this may be adjusted if it retrieves too many or too few keywords.) of the keywords in order. Using the *tf*idf* values makes sense because a term's weight is proportional to the number of occurrences in the document and also inversely to the number of documents in which it appears. Thus a term that shows up only few documents has a bigger weight, which has more distinguishing power. Each keyword shows how many documents can be separated by the keyword. For example, if there are 10 documents in *Java* and the word *thread* shows up in 5 of them, then the new category *thread* will have organized 5 documents, leaving 5 documents in *Java*. Note that the documents in *thread* are still categorized as *Java*. The 5 documents would have *Java – thread* category whereas the other 5 documents would just be *Java.*

If none of the keywords does a reasonably good job of re-categorizing, a user may enter his own keyword and check how much it separates the documents.

Proportional Tagging System

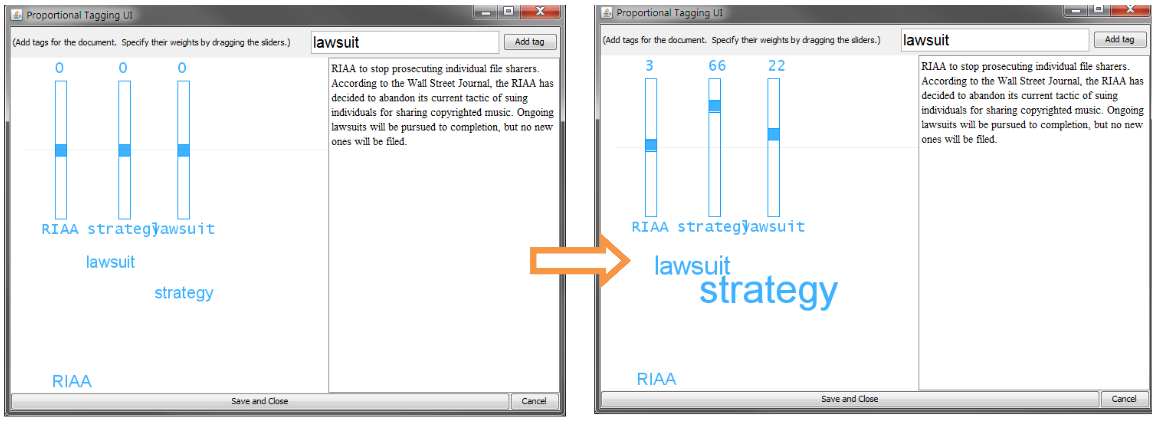

Figure 5-21 The proportional tagging system. When the tags are first created they weigh equally. A user may drag the bar up or down to give more and less weight. As the bar is dragged the size of the tag and value continuously change to reflect the motion. Note the relative size changes among the tags. The bottom section acts as a tag cloud.

A tagging system is often used to organize information and retrieve it later. Existing tagging systems allow a user to tag an information object with several keywords (Guan et al. 2010). However, the keywords are weighed equally (or rather, there is no easy way to weigh each keyword differently) which means the tags may not represent the information object in a meaningful way. For example, consider a document on Java concurrency. A user may tag the information with ‘java’ and ‘concurrency’, but may want to emphasize the ‘concurrency’ part. Perhaps the user feels the information should be under-represented when he later searches for general Java documents. Proportional tagging is similar to assigning *tf*idf* (Term Frequency times Inverse Document Frequency) values to the tags. The *tf*idf* value of a word increases as the number of the times the word appears increases and as it appears in fewer documents. It tells what words are important in searching for the document. (See Appendix B: Proportional tagging for more detail.)

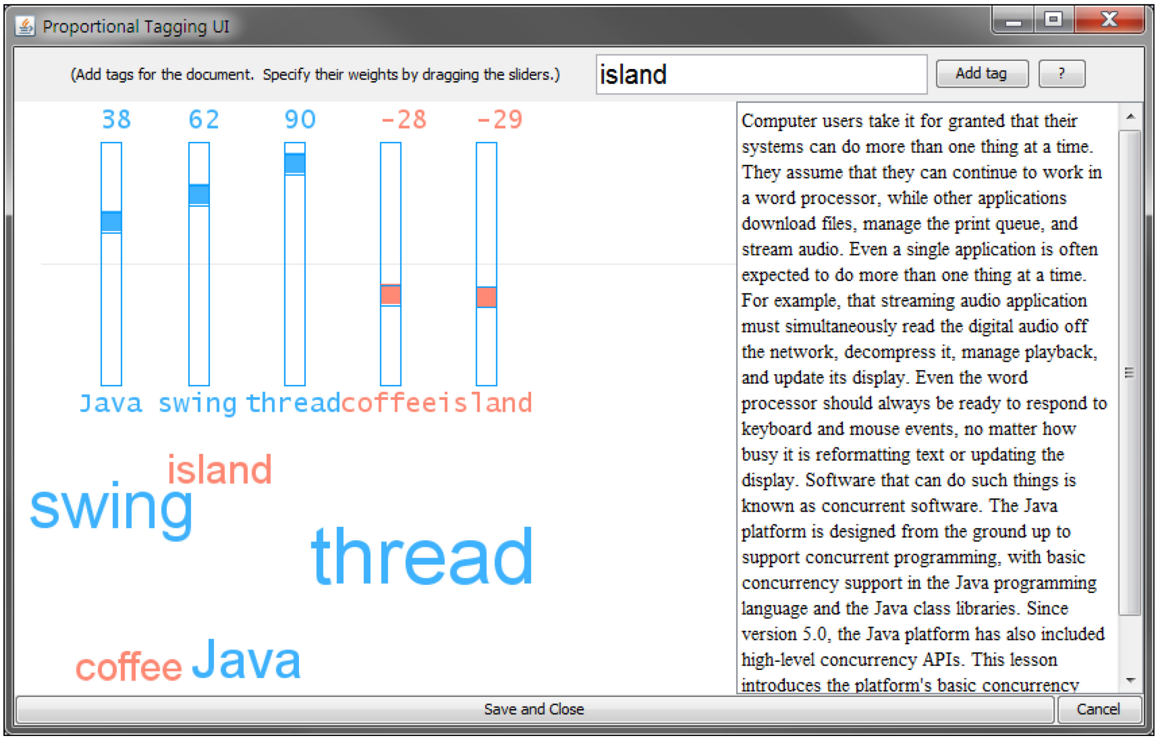

Figure 5-22 Negatively weighing tags in the proportional tagging system. Tags may be weighed negatively to be used not to match certain words.

Arkose2 provides a user with a simple but very interactive GUI that allows a user to visually weigh each tag differently if he wants. The default will be an equal weight for all the keywords.

A user may choose to represent the tags with differently from the default setting. To increase or decrease the weight, the user can drag the bar associated with each tag (Figure 5-21). Also, a user may weigh a tag *negatively* (Figure 5-22). In the proportional tagging GUI, the blue words are represented as tags with a positive weight (between 0 and 100) and the red ones with a negative weight (between -1 and -100). The size of the word represents the absolute value of the weight. As the user drags the bar below 0, the color of the word changes to red and the size continues to grow according to the value. These weights are used when calculating a similarity value among the tagged documents.

Although the current version uses the cosine similarity values when matching documents, a more effective algorithm may makes a better use of the extra information a user supplies.

Wiki and Personal View

Arkose2 also provides a rudimentary wiki functionality to save and organize information. The information may come from a discussion space, summaries, and other wiki pages. One advantage of Arkose2's wiki over other Internet based wikis is that each object in the wiki page is draggable, thus more readily rearrangeable. This is similar to the user experience in a page of Microsoft OneNote™. Although a user can rearrange information in a traditional wiki page by cut-and-paste, Arkose2's wiki allows a more natural rearrangement of information within the page.

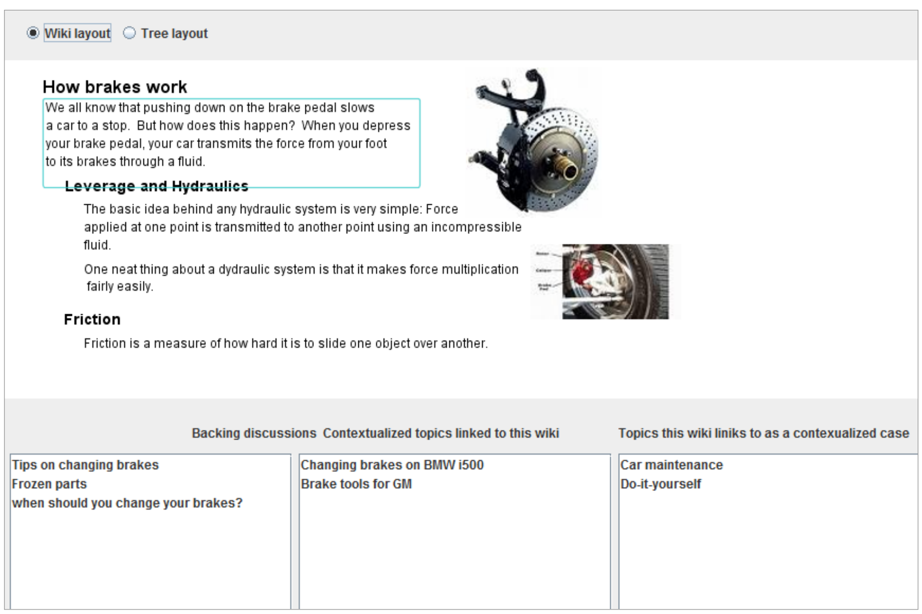

Figure 5-23 A wiki page. Information may be saved from summaries or users may directly edit on the wiki. Note that the headers, paragraphs, and images movable objects (similar to a page in Microsoft OneNote™). User connected contextualized cases and other related discussion spaces and wikis are shown (data simulated).

Wiki pages are connected with other information spaces in Arkose2. For example, Figure 5-23 shows a list of discussion from which the information came, contextualized cases linked to the wiki page, and other topics to which the wiki page may be linked. (Data shown are simulated in the example.) This allows users to easily navigate to other related information spaces by providing them and to be aware of how information spaces are structured.

Depending on the user preferences, a wiki page may show data in the paragraph view (Figure 5-23) for readability or the tree view (Figure 5-24) for easily understanding the structure.

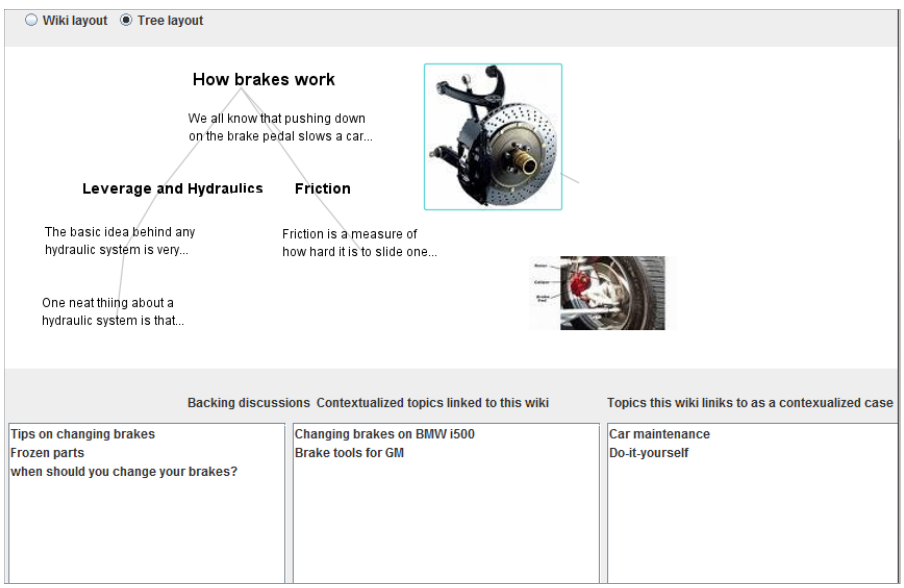

Figure 5-24 A tree layout view of the wiki. The headers and paragraphs form a tree.

Similarly, a user has his own personal wiki space. While the information on a general wiki page has the same structure for every user who sees it, a user may rearrange information according to his preferences in the personal wiki pages. A problem with wikis or discussion forums is that everyone is forced to view the same content in the same format. A user may want to selectively choose discussions on certain topics or re-order information to suit the needs. This was evident in the evaluation of Arkose2 where some participants expressed strong preferences for their own organization of information over what was given to them. A personal wiki view that allows a user to flexibly organize customized contents might be helpful in keeping the information relevant to the user.

Allowing users to have a customized user view may in turn benefit the main Arkose system as well. A user may organize information on a topic that is copied from other parts of Arkose2, for example, he can bring in content about how to deal with depression. In the discussion spaces and wikis of Arkose2, there may be discussions and questions about medications, different treatment options, and people's testimonies in various distillation stages. Some may be in a question-answering format, some may be semi-distilled, and others may be wiki pages. Since copying the content into the personal view is indicative of the user's interest on the topic, it is likely the user wants to see further changes on the topic as users distill the information more. It also provides an opportunity for Arkose2 to request the user perform some work on the topic. For example, if the information the user saved is tagged with task requests created by other users such as a request for further distillation or organization, the user will be asked to work on them. The user will, of course, may choose not to work on any of the tasks. However, users interested in a topic will be more likely to work on it than a random user. In the evaluation of Arkose2, some participants revealed that they were more likely to distill information that they found interesting. Participants in general also expressed increased difficulty in dealing with information with which they were not familiar (see more detail in Chapter 6).

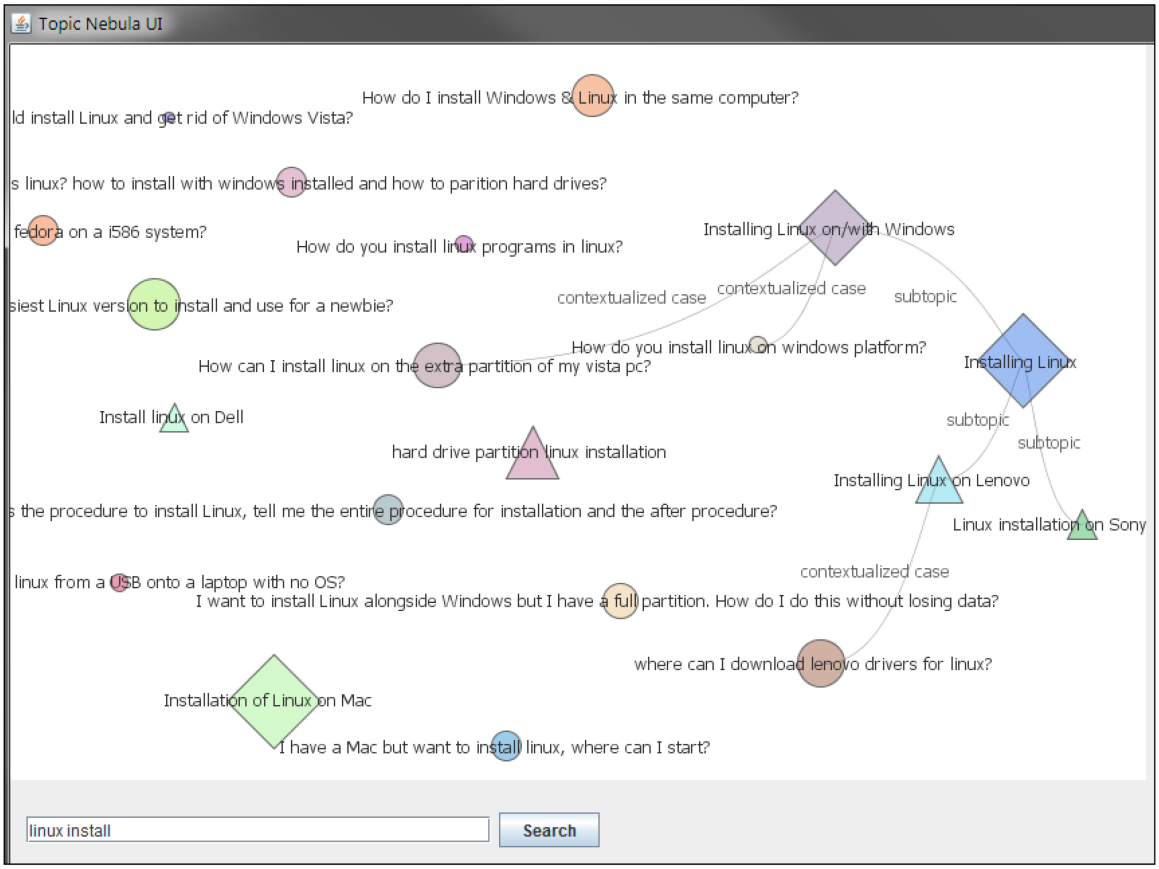

Figure 5-25 The Topic Nebula UI for searching and navigating topic space. Circles represent questions or discussions, triangles represent personal wiki topics, and diamonds represent wiki topics. Some of the questions are linked as contextualized cases while some others as subtopics. Note the size differences among the topics. It represents the number of posts, replies, and/or paragraphs in each topic space, indicating how much information a user may potentially find. (Data simulated)

Further, some of the user selected rearrangement and customized information in the personal view may be used by others as well. One may envision a collection of individual user's organization of information may even form a type of market place in which users are rewarded for creating useful distillation of information. Users may even explicitly post a distillation task for certain part of the discussion space, similar to a task oriented market place such as TaskCn (Yang et al. 2008a). While the current version of Arkose2 does not support this feature, it would be relatively easy and straightforward to add the functionality.

Topic Nebula

Lastly, Arkose2 allows users to easily search and navigate the *topic space*. A topic space in Arkose2 consists of all the questions and topics answered and discussed, from which a user can go to the discussion space of a question or topic. A visualization named the Topic Nebula visually presents topics and searched topics. The Topic Nebula is used in three ways. First, a user may navigate the topic space in order to see what topics are in the system. Second, the user may use the Topic Nebula to distill the topics. For example, if he sees many related topics he may organize them by linking them together under an overarching title. And third, the user may use the Topic Nebula to connect contextualized cases with generalized topics.

For example, there may be questions and discussions on installing different versions of Linux operating systems as well as a wiki topic on installing Linux in general. The topics are *floating* in the Topic Nebula without any connection to one another. A user may choose to organize them by connecting several of specific cases of Linux installations as contextual cases of the wiki article.

In another potential scenario, there might be several questions about installing different versions of Linux, but no general guide. A user can summarize the cases and write a wiki page including the cases, which is the *bottom-up* approach discussed in the Usage Scenario section. Similarly, if a user sees a general wiki page but not any specific information on installing his version of Linux, he may initiate a question from the wiki page (*top-down*). Through these activities topics in Arkose2 are gradually organized by the users' activities, which in turn may help the future users of Arkose2.

This may lead to the interesting possibility of organizing information by harvesting users' normal asking activities. It is well known that artificial networks such as a social network, the Internet, and a citation network exhibit a power-law degree distribution because of preferential attachment (Barabási and Albert 1999; Dorogovtsev and Mendes 2003). New nodes come into the network and connect to existing nodes not in a random way, but they prefer nodes with many connections. This behavior leads to "the rich gets richer" where a large group of nodes grows faster. Similarly, the growth of the pages in Wikipedia (http://www.wikipedia.org) is found to show this behavior. When a user browses topics in the Topic Nebula, he may prefer topics with more objects such as subtopics and contextualized cases attached to them because there may be a bigger chance to finding useful and diverse information in that space. Not only that, it is likely he may initiate his own question from the topic contributing to the growth of the space. A user may choose to post his question without connecting it to any exist topic, letting it stand alone. But a large group of connected topics may gain more attention than an individual topic because of its visibility in the Topic Nebula. It may attract users with relevant interest and expertise more because they are already in that space looking for relevant information.

Over time, this may lead to an organic grouping of topics where topics form networks and grow into a topic space of their own. Thus users may not need to manually distill related topics as much but take advantage of the natural formation of topics in the Topic Nebula.

Summary of Features and Contributions

In this chapter, I discussed the design and implementation of collaborative information generation and distillation system, Arkose2. The features of Arkose2 fall into three capabilities. The first is handling complex questions that require multiple answering cycles. For this, user interfaces for breaking a complex topic into sub-topics are implemented. Users can also conduct a backchannel conversation separately from the actual discussion. These help a discussion to occur more flexibly.

The second is organizing and searching in the discussion and information space. The proportional tagging system allows users to weigh tags differently in order to provide new ways to match relevant information. The organic categorization helps organize documents in a constantly changing information space, and the discussion widgets such as polling and survey widget help organize the discussion space better.

Finally, the third capability is saving and connecting various information types, such as generalized information with contextualized information and topics with subtopics.

My contributions in this chapter are the design and implementation of new interaction mechanisms and supportive features. Especially, I believe the proportion tagging system and organic categorization are new and innovative and they should help users' organize information in an unintrusive way. The tools may become even more effective as new algorithms are developed to utilize the extra information better.

Another contribution of my work in Arkose2 is that I designed and showed a way for a discussion to progress in a more organized way, yet without imposing a formal structure ahead of the time. As discussed in Arkose 1.0 (Chapter 3), imposing an *a priori* structure as to how the conversation should be shaped was limiting and stifling. I instead allow a small amount of structure to be built in the discussion *as* the needs come up, which allows users to freely conduct a discussion while helping them to do it more coherently. The features of organization in a discussion include dividing a topic into subtopics, separating a backchannel conversation from the actual discussion, and organization of several discussion types through discussion widgets.

And finally, my other contribution is putting all the components together in a coherent manner with other supportive tools such as the scaffolding, various visual aids, and underlying information retrieval techniques that work together well in support of user activities.

In the next chapter, I evaluate Arkose2's distillation support capabilities in action.

Chapter 6

Evaluation of Arkose2

In this section, I discuss the user evaluation of distillation with Arkose2. The focus of the evaluation is not on determining a quantitatively-based conclusion about the usefulness or effectiveness of the system. Rather, it is a generative approach that focuses on the issues, difficulties, and user perceptions in distillation using Arkose2 through a qualitative analysis. The evaluation will also test whether Arkose2 can support users to create summaries, examine what features are helpful and what are not useful, and describe users' approaches in summarization of a discussion space.

Outline

This evaluation chapter is outlined as the following. First, I present the research questions asked, and briefly describe the design claims and findings from the evaluation. These will be more fully explored in later sections. Next, I describe the evaluation settings; the participants, data preparation, evaluation conditions, and data collection. I then go into detailed descriptions of participants' tasks, followed by the summarization details of what participants did in the evaluation. Next I walk through each of the design claims and discuss whether the claims were supported in the evaluation as well as any issues discovered. Finally, I present users' experience with Arkose2 and limitation and tradeoffs, and conclude the chapter with a summary of findings and implications.

Research Questions

Several important questions are examined in the evaluation of Arkose2. First, it examines whether users can sufficiently summarize a discussion space using Arkose2. Distillation of information takes much effort. A user would need to read and understand the posts in the discussion space, locate important topics and supporting evidence, organize different topics in a coherent way and create structured summaries by linking information. The evaluation examines whether Arkose2 supports participants to create summaries.

The second question is how the complexity and difficulty of a discussion space affect distillation tasks. As the discussion space becomes larger and more complex, what challenges do the participants have and how do they affect their summarization tasks while using Arkose2?

The third question is how the support capabilities of Arkose2 help users' summarization tasks, especially when the complexity of the discussion space increases. Arkose2 has various visual aids, interactive features, automatic algorithms that cluster similar posts together and search capabilities, and understanding how these features help or do not help users' tasks will provide insights to the success of future supportive tools.

As well, any issues in the design and implementation of Arkose2 will be discussed.

Claims and Findings

Based on the research questions, several design claims about distillation using Arkose2 can be made. They are:

1. Participants can create suitable summaries for the discussion spaces using Arkose2. They can create the summaries incrementally.
2. Summarization scaffolding in Arkose2
   a. Scaffolding can assist participants in creating summaries in Arkose2.
   b. Support for scaffolding in Arkose2 can allow specific kinds of collaborative distillation.
   c. Scaffolding will be more helpful when the discussion space is larger and more complex.
3. Design features of Arkose2
   a. Visual Navigator will assist participants in creating summaries.
      i. The tree representation of the discussion space will be helpful in several ways.
      ii. Hiding summarized posts will be helpful in keeping track.
      iii. The underlying automatic clustering of the posts will be helpful.
   b. Linkage among different views will assist participants in navigating the discussion space.

These will be discussed in terms of several findings from the evaluation that will be discussed below. They are:

1. There were sufficient differences among the different levels in terms of the complexity and amount of information.
2. Several factors influenced what details and information summaries had.
3. Participants' summarization approaches varied.

4. There were several factors that made understanding and summarizing a discussion space difficult.
5. There were issues with Arkose2 that made summarization tasks more difficult.

Evaluation Settings

In this section, I describe various components of the evaluation settings such as the participants, discussion data to be summarized, tasks settings, and data collection.

Solicitation and Participants Detail

I sent out a recruiting email to the School of Information mailing list (si.all.open@umich.edu) which consisted of the faculty, students, and staff member of the school. The initial email had information about the tasks involved (reading discussion threads and creating structured summaries) and the age requirement of at least 21 years old. Seventeen people responded with an interest for participation, upon which I sent them an individual email with more details about the evaluation, compensation, and available times. Two responders had scheduling conflicts and 15 participated in the evaluation in the end. They consisted of 11 School of Information Masters students, two Informatics undergraduates, one engineering graduate (but who took SI courses), and one former Political Science undergraduate (see Table 6-1 for user demographics).

A fee of $60 was offered for a full participation that would last about 4 hours, and a rate of $10 per hour was offered if a participant had to stop in the middle for any reason and did not finish the session. To encourage the participants to create quality summaries, I offered an additional reward of $20 for the participant who would create the best summaries based on my judgment. When it was closer to the scheduled evaluation date, I

Name (replaced)	Age	Occupation or School/ Major	Years in school (college+)	Experience w/ visual/interactive application	Experience with summarization
Phil	23	SI[1] / IPOL[1]	5	None	2 years, academic articles
Emily	25	SI / CI[1]	5	None	Book reports for school, progress report for work, otherwise minimal
Brian	24	Political Science (graduated)	6	Charts/graphic organizers, "Inspiration" for essay outlines	Books and articles throughout college
Ellen	22	Informatics / Life Sciences undergrad.	3	Not sure	Textbook materials, 3 chapters per week
Rob	37	SI / PI[1]	8	None	limited experience summarizing journal articles in undergrad
Daniel	27	Engineering graduate	Didn't say (assuming 5+)	None	Only articles on printed paper
Matt	25	SI / HCI[1]	College + one semester of SI	Used visual based operating systems for 18 years	Books and news articles for various school activities
Lisa	25	Informatics undergrad.	5	None	Books, articles since grade school
Josh	26	SI / IAR[1]	9	Yahoo answers, presentation tools	For exams, competitions
Morgan	28	SI / HCI	6	Photoshop, Visio	school assignments
Erin	34	SI / ARM[1]	11	None	Books and articles, 10 years
Gabriel	22	SI / IPOL	4	Some, not sure.	Average, mostly with books/articles
Justin	26	SI / SC[1]	6	None	Fairly small, handful summaries of books throughout 18+ years of education
Chris	49	Former journalist, SI / HCI	6	Photoshop, desktop publishing programs for 20 years, Movie editor program.	Edited text and written headlines, 20 years. Summarizing web pages, 6 years
Travis	23	SI / LIS[1]	5	None	Summarization as a history major in college

Table 6-1 User demographics based on participants' responses.

sent them a reminder email with directions to the evaluation location along with the consent form.

I took their names, major, and specialization[9] at the evaluation session. Additional demographic information about their age, years in school (college+), experience with a visual or interactive application (such as a mind mapping program), and experience with summarization were later collected through email communication. Most participants did not have any experience with a visual application similar to Arkose2, and their summarization experience was mostly school or work related. One participant, Chris, was a former journalist who worked at a newspaper. He had significantly more experience with summarization in a variety of textual materials and with desktop publishing programs. See Table 6-1 for details about their demographic information.

Data Preparation

In this evaluation participants were to summarize a set of discussion spaces collected from existing online discussion forums and question-answering sites. The set included four discussion spaces, each consisting of a topic question posted by a discussion initiator, subsequent posts replying to the topic question and to one another, answers to the questions, and other information provided by the community members. (See Appendix D for the discussion data in each level and condition.) The discussion information was collected from one or more forums or question-answering sites and was modeled following Webb's Depth of Knowledge model (Webb 2005).

[9] School of Information's Master program has sub-specializations such as Human Computer Interaction, Information Policy, Social Computing etc. in which students specialize.

Question type	Description	Example
Level 1	Recall and reproduction – requires only a rote response, uses a well-known formula that involves one step procedure	- Recall or recognize a fact, term, or property. - Represent in words or diagrams a scientific concept or relationship - Perform a routine procedure
Level 2	Skills and concepts – consists of mental processing beyond simple recalling, involves more than one step	- Specify and explain the relationship between facts, terms, properties - Formulate a routine problem given data and conditions - Organize, represent and interpret data
Level 3	Strategic thinking – requires reasoning, planning, using evidence, and a higher level of thinking. Drawing conclusions from observation.	- Identify research questions and design investigations for a scientific problem - Solve non-routine problems - Develop a scientific model - Form conclusions from experimental data
Level 4	Extended thinking – requires high cognitive demands and relating several connections among ideas. Requires complex reasoning, experimental design and planning.	- Investigate a problem for an extended period of time from problem definition, experimental planning, to analysis of data and creating conclusions.

Table 6-2 Webb's Depth of Knowledge

Webb's depth of knowledge defines four basic types of questions that guide educators in creation of learning objectives and assessment of students' learning experiences. It has been used widely in the education community and education departments (e.g., WCER 2007; ETS 2008 Hess 2004; Hess 2005) to design curricula. Table 6-2 shows the types of questions and examples from the depth of knowledge.

In order to model each discussion space to the Depth of Knowledge, I used two heuristics in measuring the complexity of a discussion space. One is the nature of a question itself that might indicate the complexity of the information. For example, the question "How should the government reform the medical insurance policy?" would in general call for discussions far more complex than the question "How do you pick the right insurance provider?"

Another heuristic for determining the complexity of a discussion space is the number of sub-topics and posts. A discussion on a topic can be arbitrarily complex – for example, answers to a seemingly simple question such as "Are you using SPF records?" on Slashdot.com (http://ask.slashdot.org/story/09/12/17/2359241/Are-You-Using-SPF-Records) garnered more than 360 posts as of 12/23/2009. Obviously, it is not a simple yes-or-no question. As more sub-discussions occur, more information and ideas are brought in leading the complexity of the discussion space to increase regardless of the question type.

I used a combination of the two heuristics when preparing the discussion data for each level. I believe this approach of approximating the complexity of a discussion space was successful as indicated in the evaluation of Arkose2 (see the Summarization details section for more detail). Each discussion space was selected and modified to fit the level. According to the model, level 1 questions would be simple question-answer pairs (e.g., "Who is the first president of the United States?") that might not require further distillation, and were thus omitted from the evaluation settings. The discussion space was created by combining several related topics together. A level 2 discussion space had about 3 sub-threads with 20 posts. A level 3 discussion space had about 8 sub-threads with 50 posts. A level 4 discussion space had about 14 sub-threads with 80 posts. (The numbers of the sub-threads are an approximation as it is not always easy to tell where one thread ends and another starts. I used the number of major branches in the discussion tree to count them because often a coherent discussion happens within a large branch.) The length of a post varies across the levels as well. In level 2 discussion most posts were short, while they were longer in level 3 and even longer in level 4. The longer length of the posts in the later levels means there is potentially more information in each post, contributing to the complexity. (See Appendix D: Discussion threads in the evaluation for the discussion posts used in the evaluation.)

The following is a list of candidate questions found in Yahoo! Answers and Slashdot, which were used to create each discussion space for the evaluation. I chose topics in illegal music file sharing and the RIAA (Recording Industry Association of America)'s reaction, as most college or graduate students should know them at least to some degree. As found in the evaluation, all of the participants were familiar with the general topic, though the familiarity with a specific sub-topic varied from participant to participant. Representative sub-threads from the following discussion threads were combined in a way that fit the number of questions in each level specified above. The next section shows the questions and discussion topics from which the evaluation data was prepared.

Candidate Questions and Discussion Threads

Level 2: Topic – How would RIAA or ISP know you are downloading illegally?

A combination of

"How do you know you are caught with the RIAA?"

(http://answers.yahoo.com/question/index;_ylt=Ak1ZTyWJUua0qlvlIonQwHUjzKIX;_ylv=3?qid=20080118233749AAdriD4)

"Exactly how would my ISP know what files I download?"

(http://answers.yahoo.com/question/index;_ylt=AkFeuyoXvqfK.5md.ny6WtIjzKIX;_ylv=3?qid=20080708212204AAu4vYH)

"Is there a download limit which makes me appear suspicious to my ISP or the Government?"

(http://answers.yahoo.com/question/index;_ylt=Asb49dQFEGvOTy9EuaS7._ojzKIX;_ylv=3?qid=20061125091409AAF1ZiB)

Level 3: Topic – How is the court reacting to RIAA's strategy?

A combination of

"Court says fair use may hold in some RIAA cases"

(http://yro.slashdot.org/story/09/12/09/1835205/Court-Says-Fair-Use-May-Hold-In-Some-RIAA-Cases)

"Court appoints pro bono counsel for RIAA defendant"

(http://yro.slashdot.org/story/09/07/17/189208/Court-Appoints-Pro-Bono-Counsel-For-RIAA-Defendant)

"Fair use defense dismissed in Sony v. Tenenbaum"

(http://news.slashdot.org/story/09/07/27/2356219/Fair-Use-Defense-Dismissed-In-SONY-V-Tenenbaum)

Level 4: Topic - Compare and contrast the positions of those for and against RIAA's strategies.

A combination of

"Constitutionality of RIAA damages reinstated"

(http://yro.slashdot.org/story/10/01/04/195208/Constitutionality-of-RIAA-Damages-Challenged)

"Antitrust case against RIAA reinstated" (http://yro.slashdot.org/story/10/01/15/1329200/Antitrust-Case-Against-RIAA-Reinstated)

"$1.9 million award in Thomas case raises constitutional questions" (http://yro.slashdot.org/story/09/06/19/1550232/19-Million-Award-In-Thomas-Case-Raises-Constitutional-Questions)

"DoJ sides with RIAA on damages" (http://entertainment.slashdot.org/story/07/12/04/213231/DoJ-Sides-With-RIAA-On-Damages)

"RIAA wants limits on Net Neutrality so ISPs can police file sharing" (http://yro.slashdot.org/story/10/01/16/160257/RIAA-Wants-Limits-On-Net-Neutrality-So-ISPs-Can-Police-File-Sharing)

	No scaffolding	Maximum completion time	Scaffolding	Maximum completion time
Level 1	N/A	N/A	N/A	N/A
Level 2	Yes	25 min.	Yes	15 min.
Level 3	Yes	50 min.	N/A	N/A
Level 4	N/A	N/A	Yes	90 min.

Table 6-3 Evaluation settings for the levels. One level 2 with scaffolding, one level 2 without scaffolding, one level 3 without scaffolding, and one level 4 with scaffolding were prepared. In total, the summarization tasks add up to 3 hours.

The other condition besides the level of the discussion space was the presence or absence of summarization scaffolding. Two types of distillation tasks were prepared. One was distillation of a discussion space with predefined summarization scaffolding and the other without. I envision that experienced Arkose2 users or users with more domain expertise can create a skeletal structure (scaffolding) for distillation tasks that other users can fill. The scaffolding may be incrementally constructed as well. An experienced user may write an initial structure upon which other users may build on or add contents to it.

Level 4 without scaffolding was omitted from the evaluation because it was likely to be too difficult and time consuming for one participant to handle it at a time. Given the time constraint, I chose a mix of settings (see Table 6-3) that I believed would give insights to the research questions and would still be somewhat doable for the participants in one session. The four tasks themselves were calculated to add up to about three hours in total. With interviews after each task and instructions and demonstration, each session took about 4 hours.

Task Settings

Upon arriving, the participant read and signed the consent form. I explained the overall purpose of the evaluation and described the task details to the participant. I gave the participant a few minutes to try out features of Arkose2 with some demo discussion data. All the actual tasks used different discussion data from the demo data. I then explained that the tasks would be a collaborative effort in the real world settings although he would be summarizing the discussion space by himself in the session. I instructed the participant that he should pretend he was working with others and feel free to leave comments and questions for others in the summarization tasks. And I asked him to create structured summaries rather than one paragraph summary for the entire discussion space and "think aloud" during the task. However, the first participant stated that he could not do both simultaneously and he was allowed to just concentrate on the summarization tasks. The rest of the participants were given a choice of "thinking aloud" or doing it silently, but except for two participants who spoke throughout the tasks, all the other participants did the tasks without talking about what they were doing. (The details of the tasks are discussed in the Task details section.)

After each of the fours tasks was finished, I interviewed the participant for 5-20 minutes and asked about what he did, why he did it, and walked through from the beginning of the task. He was also asked to make comments, express his thoughts, and ask questions during the tasks and interviews. The interview was semi-structured in that I started with some initial questions but allowed new questions brought up as the interview was shaped. In total a participant was interviewed for about one hour. (See Appendix C for interview questions.)

The participant went through each task in the following order:

For the first three participants,

1) Level 2 with scaffolding (20 min.)

2) Level 2 with **no** scaffolding (20 min.)

3) Level 3 with **no** scaffolding (45 min.), and

4) Level 4 with scaffolding (70 min.)

For the rest of the participants:

1) Level 2 with **no** scaffolding (20 min.)

2) Level 2 with scaffolding (20 min.)

3) Level 3 with **no** scaffolding (45 min.), and

4) Level 4 with scaffolding (70 min.)

The first three evaluation sessions were conducted in two days, after which the result and evaluation protocol were reviewed. The order of test settings was changed so that a participant would start with level 2 with no scaffolding. The original testing order with level 2 with scaffolding as the first task was meant to give the participant some idea on what summarization was and to test whether a participant would be able to build at least some summary with the help of scaffolding. The first three evaluations indicated participants could easily create summaries using Arkose2, and the subsequent evaluations were conducted with level 2 with no scaffolding as the first condition in order to give a participant a chance to use his own strategy without being influenced by the scaffolding. The switch of the order did not produce any noticeable difference.

	Level 2 w/o scaffolding (Max. 20 min.)	**Level 2 with scaffolding (Max. 20 min)**	**Level 3 w/o scaffolding (Max. 45 min)**	**Level 4 with scaffolding (Max. 70 min)**
*Phil	NF (*2/3*)	NF (*3/3*)	NF (*7/8*)	NF (*8/14*)
*Emily	18	NF (*2/3*)	25	42
*Brian	14	14	43	61
Ellen	NF (*3/3*)	20	43	66
Rob	NF (*3/3*)	NF (*3/3*)	NF (*8/8*)	NF (*12/14*)
Daniel	NF (*2/3*)	NF (*3/3*)	NF (*8/8*)	52, ST (*10/14*)
Matt	12	18	33	59
Lisa	11	15	22	45
Josh	17	NF (*3/3*)	48	59, ST (*11/14*)
Morgan	18	NF (*3/3*)	39	**60, ST (*11/14*)
Erin	10	11	24	47
Gabriel	15	NF (*3/3*)	NF (*4/8*)	NF (*7/14*)
Justin	NF (*3/3*)	20	NF (*7/8*)	NF (*6/14*)
Chris	13	15	43	50, ST (*9/14*)
Travis	NF (*3/3*)	19	39	48

Table 6-4 Task completion time. NF: Not finished on time, ST: Stopped before finishing, Numbers: time took. (m/n) states m sub-topics were covered out of n sub-topics. (*: for these participants, Level 2 with scaffolding was the first task then level 2 without scaffolding. **: mistakenly stopped 10 minutes early)

Data Collection

What was Collected

During each task a participant's actions were observed and notes were taken to record the starting and finishing times and any notable actions by the participant such as how much of the discussion space he read before creating the first summary and what features of Arkose2 he tried and used. Questions that were formed during the task were also written down so I could ask the participant in the interview afterwards.

A participant's summarization outcome, along with other meta-information, was saved in real time into Arkose2's backend MySQL database, preserving all of the data the participant produced. The database contents were dumped into sql files and saved as well.

The participant's voice and summarization details were recorded using a screen capture program, which allowed me to go back and see the details of the summarization steps later. About 3 to 4 GB worth of video was captured for each participant. Audio was extracted and edited to remove any silent part. The entire voice recordings were transcribed and color coded to indicate which part of the transcripts had particular information.

Participants' demographic information was also collected. The name, major, and specialization were recorded at the evaluation (age confirmation was done only to check a participant was over 21.). Later, age, years in school, experience with any visual or interactive application, and experience with summarization were collected through email communication.

Task Completion

A few participants did not finish some of the tasks on time. People's reading speed and comprehension vary and I expected some would perform better than others. The most telling part in the tasks was the first 20 or 30 minutes or so during which a participant decided on the summarization approach and other factors were fully shaped. The rest of the time was mostly spent in repeating the same strategy, which did not give much new information.

Table 6-4 shows each participant and his task completion times. If a task was finished before the allotted time, the number of minutes is recorded. (Any number smaller than the maximum time means the task was completed.) Tasks that were not completed are marked **NF** (i.e., Not Finished). In this case, I measured a rough estimate of summarization coverage of the discussion space by revisiting the discussion space and counting how many sub-threads were included in the summaries. In the table, NF (*m/n*) means *m* sub-topics out of *n* total sub-topics were covered in the summaries. This is inclusive in that NF (3/3) means the participant was working on the last sub-thread but did not complete it. Thus, it indicates he was very close to finishing even though he did not.

Four hours of evaluation (three participants took a few minutes of break once, but others did every task and interview in one sitting) took a toll on some participants. Three participants became tired toward the end of their level 4 task, (they either stated they were tired or it was obvious it became harder for them to focus.) In those three cases, I rather wanted to get interview data out of them than more summaries, so the task was stopped for an interview. In all three cases the participants had done at least 50 minutes of summarization (out of 70 minutes allowed), which I felt produced enough telling data. These cases are noted as **ST** (i.e., Stopped) in the table and an approximate summarization coverage was also estimated. In one case, I mistakenly stopped one participant's level 4 task at 60 minutes rather than at 70 minutes. (Noted as **ST in the table.) In all cases, enough summaries were produced to examine all aspects of the analysis.

Any issues found in using Arkose2 were collected as well. In the first three evaluation sessions, a bug in the Visual Navigator discovered which did not highlight related scaffolding nodes in the Summary Editor Space. It was still possible, but made it more difficult to locate the corresponding nodes. It was fixed for the remaining 12 evaluations.

Four participants found a usability issue when using the Visual Navigator to navigate the discussion space and read the content of the posts. The drag-and-drop operation within the Visual Navigator was not for attaching source posts to a summary but for moving the posts around in the discussion space. However, the participants mistakenly tried to drag the post from the Visual Navigator to the Summary Editor Space, which was not supported. However, the participants could work around this by remembering first to locate the post in the thread view and then drag and drop from there, and this did not affect the participants' summarization tasks much.

One other issue was the view for the editor box in summary creation was partially blocked when a summary node was created near a border. Participants could work around this by creating a summary node closer to the center or by first creating the node and further editing it in the Advanced Summary Editor. This was not fixed for the evaluation because it would require a considerable change in the underlying system.

Data Analysis

I originally planned on comparing the amount of the summaries for each level and among the participants to see whether there would be any important findings that emerged. I later moved away from this because a direct quantitative comparison would not be very meaningful or interesting. First, not every participant finished every task on time (see Table 6-4). This made it difficult to compare among the tasks and among the

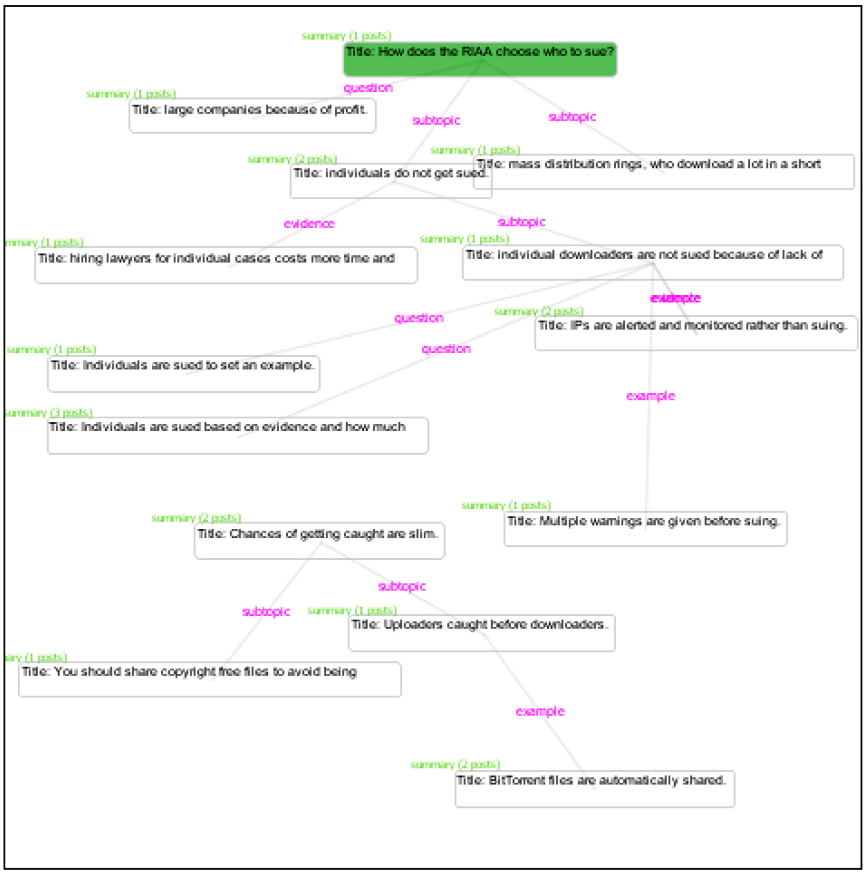

Figure 6-1 Participant Emily's summaries for level 2 without scaffolding.

participants. Second, participants' goals and approaches were different and produced varying amount of summaries (compare Brian's summaries (Figure 6-7) with Emily's (Figure 6-1)). While it *is* interesting to see some participants were so different in their summarization, the analysis would be more meaningful in conjunction with a discussion of their approaches and goals rather than by simply comparing the numbers. Thus I will compare and contrast participants' actions and their summarization outcomes throughout the discussion of the design claims and other findings in the chapter. In the next section, I discuss participants' tasks details, followed by summarization details that revealed their goals, approaches, and the type of information saved.

Task Details

A participant summarized four different discussion spaces. In this section, I describe the details of different tasks and what a participant did. I use two level 2s to

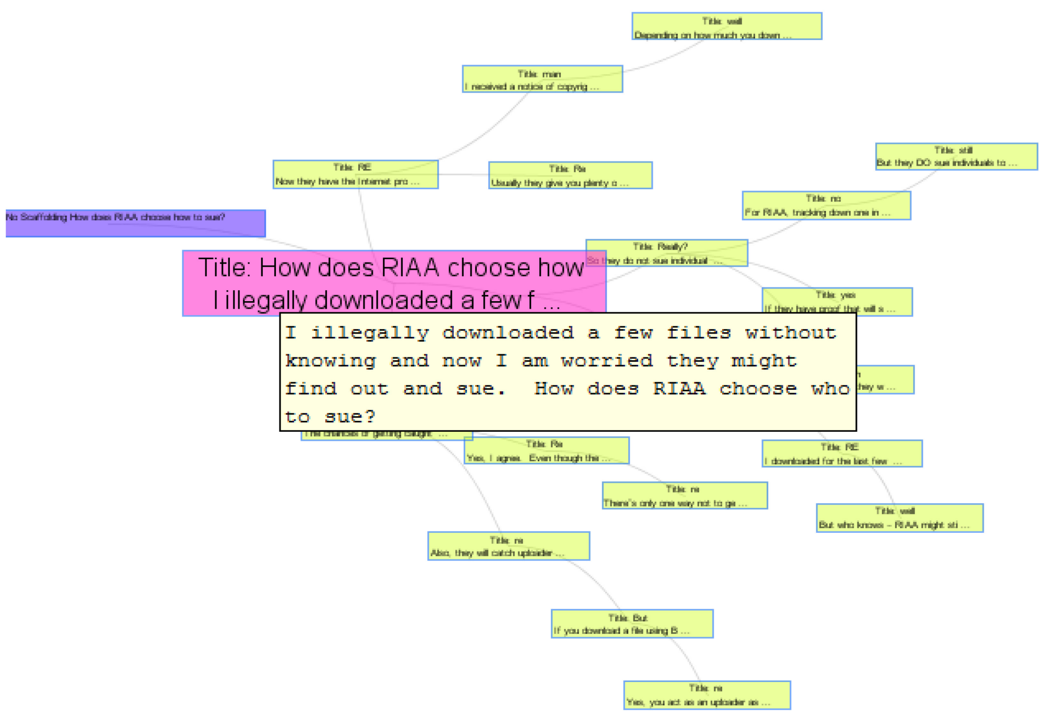

Figure 6-2 Level 2 discussion space without scaffolding in the Visual Navigator. Hovering the cursor over a post is one of several ways to view the content.

discuss the details in this section since they are small and each will illustrate task differences in the scaffolding/no scaffolding conditions. After explaining general details, I use the summaries of one of the participants to walk through what he did in order to provide an example case of what participants actually did.

Without scaffolding

Upon the beginning of the task, the participant was presented with a level 2 discussion space without scaffolding (Figure 6-2) in the Visual Navigator in the right column of Arkose2. The Thread View is in the middle column and the Summary Editor Space is in the left column (See Figure 5-8 for an overview). Level 2's topic is "How does RIAA choose to sue?" The following is a summary of the instructions I gave the participant.

- Read the discussion threads in the space and create structured summaries in the Summary Editor Space. (And I explained that structured summaries were a tree or graph structure in which summary nodes and other nodes were linked together, rather than a paragraph or page of summary.)
- The summaries should cover all the important topics covered in the discussion space as well as other information the participant thought was important, useful, and/or interesting. Use any summarization approaches that work for you.
- Imagine this is being done collaboratively with other users and try to leave any meta-information in the summary space such as questions, comments, and notes as necessary.

- The nodes (including summary, note, question, and comment nodes) created in the summary space should be linked together whenever there is a valid relationship among them.
- Add any "source" posts (the posts for which a summary is created) to the summary by drag-and-dropping the posts to their summary node which will create links between the summary and original posts.

  We will talk about what you did after the task is done. Please remember the approaches you used and any issues or questions as they come up.

Any clarification needed by the participant was also made at this time.

When the participant read the discussion posts, he had two choices. He could navigate and read the content of the posts by hovering the mouse cursor over them (Figure 6-2) or read them in the Thread View (Figure 5-8). Upon deciding to create a summary node for the posts he read, the participant would double click the Summary Editor Space (or right click and select the option from a popup menu) to open up an editor box (Figure 6-3) to type in the title and content of the summary. Upon finishing, a graphical box (Figure 6-4) representing the summary node would show up in the Summary Editor Space. The participant would then drag a post from the Thread View and drop it on the summary node to include the post to the summary. (This can be repeated to include many posts.) This would create a logical link between the two so when either was clicked the other would highlight, indicating the connection. Note that the left top corner of the summary node in Figure 6-3 says "summary (0 posts)." When the participant drag-and-drops posts to it, the number is updated (Figure 6-4) and the posts in the Visual Navigator become temporarily invisible to show the progress (Figure 6-5).

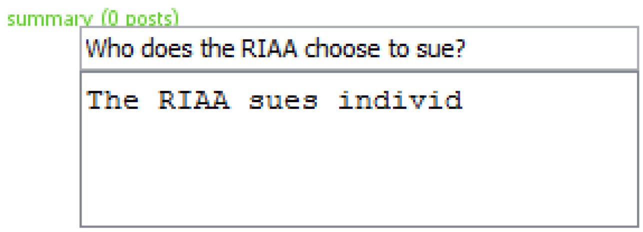

Figure 6-3 Initial summary node creation. The top box is for the title and the bottom box is for the content. Currently there is no post attached as indicated by "summary (0 posts)."

Figure 6-4 Four source posts are added to the summary. It shows 4 source posts have been attached. Clicking this summary node highlights the four posts. Details of the posts can be seen in the Advanced Summary Editor.

The participant could update the summary node as needed by double clicking it and opening up the Advanced Summary Editor (Figure 6-6). It has a bigger space for adding and organizing content, and the bottom half shows the source posts attached to the summary.

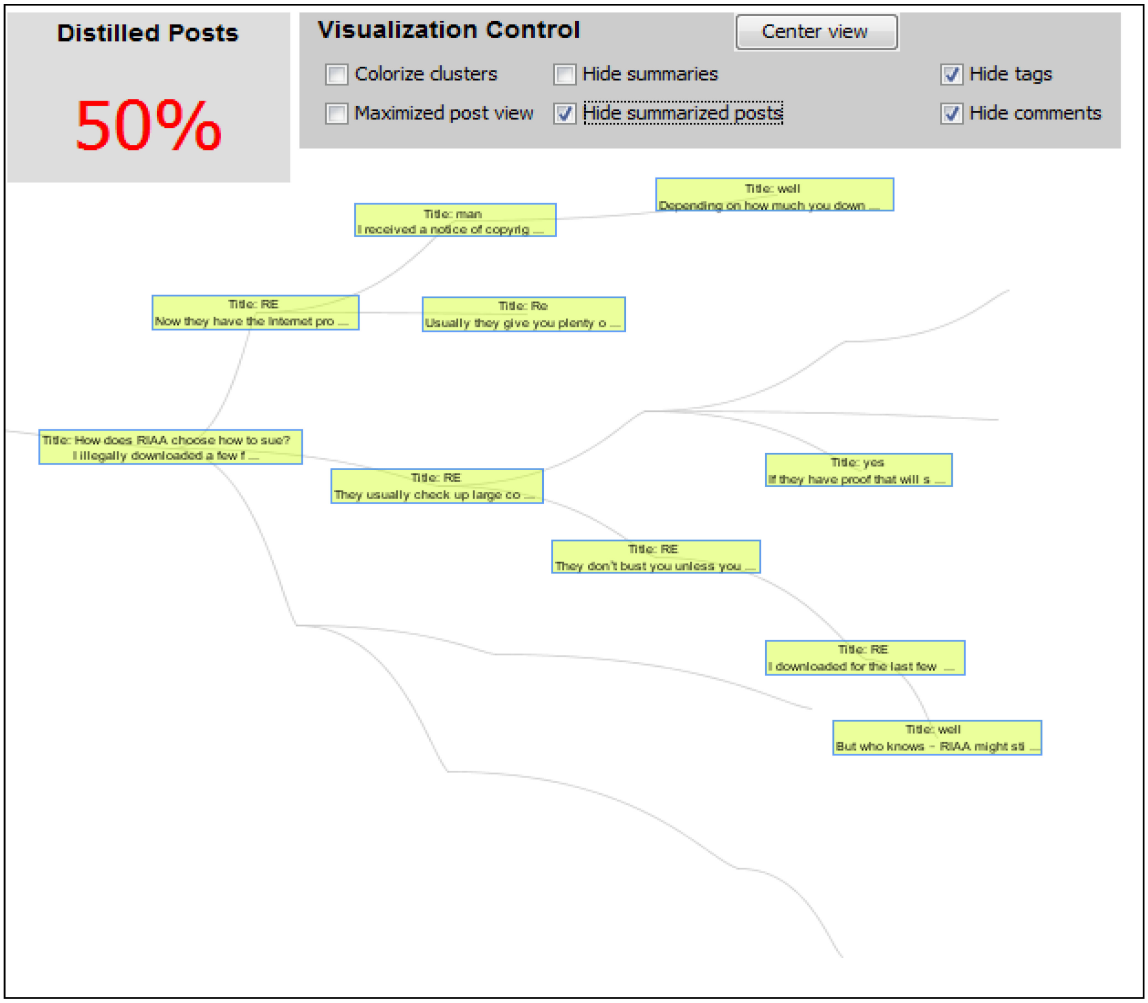

Figure 6-5 Summarization progress in the Visual Navigator. When a post is attached to a summary (thus summarized), it becomes temporarily invisible in the Visual Navigator. This option may be turned off to show every node.

Similarly, the participant could create a comment, question, or note node when he needed one. He could also create links among the nodes by selecting in the popup menu and specify the relationship.

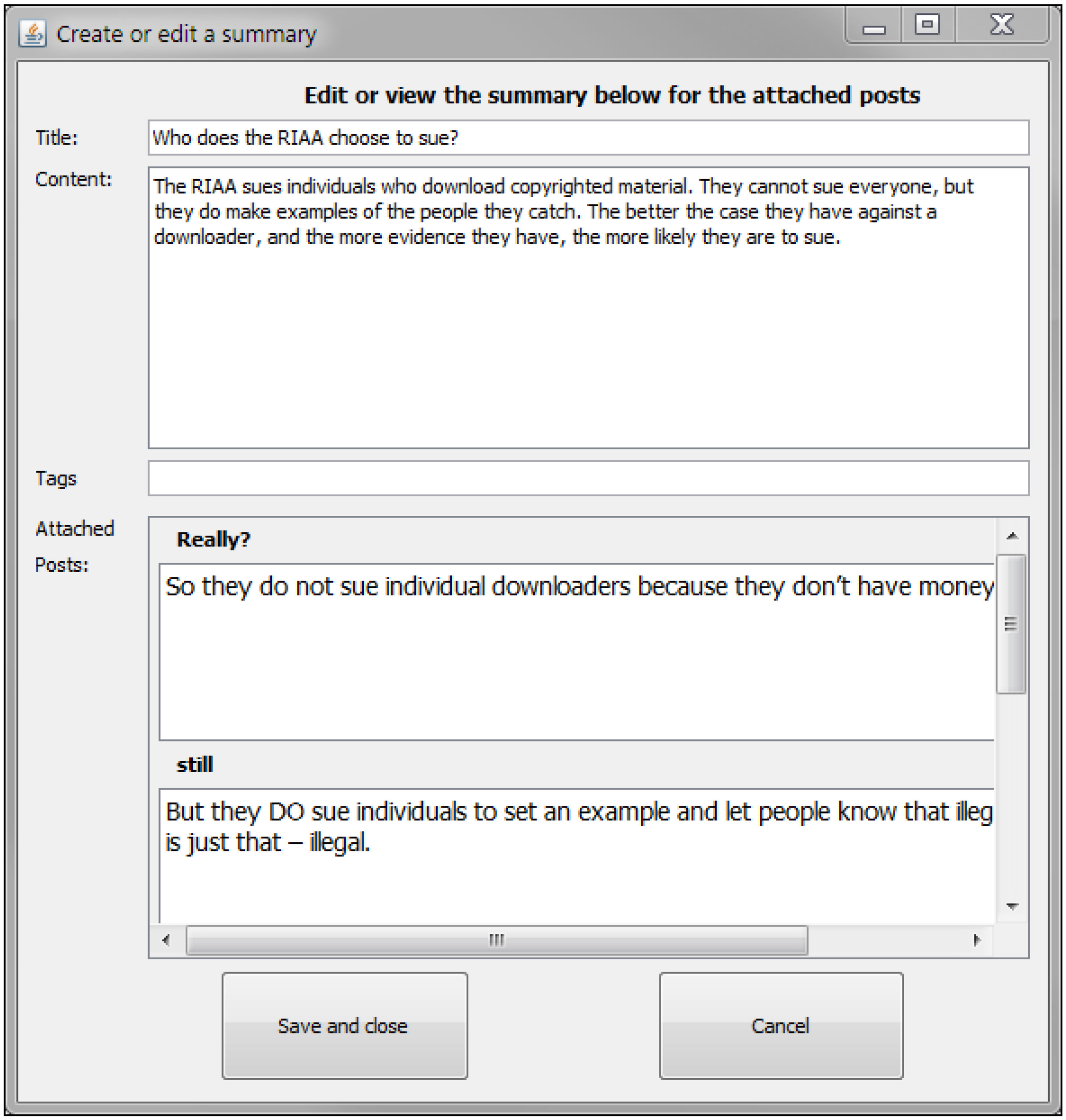

Figure 6-6 Advanced summary editor. A participant can add or organize more content. The bottom screen shows attached source posts.

The above description shows what a participant would do to create summary node and connect one another to create a structured summary. Figure 6-7 shows the actual level 2 summary of one participant, Brian. Here are the steps Brian took to create the summaries (recreated from the video recording).

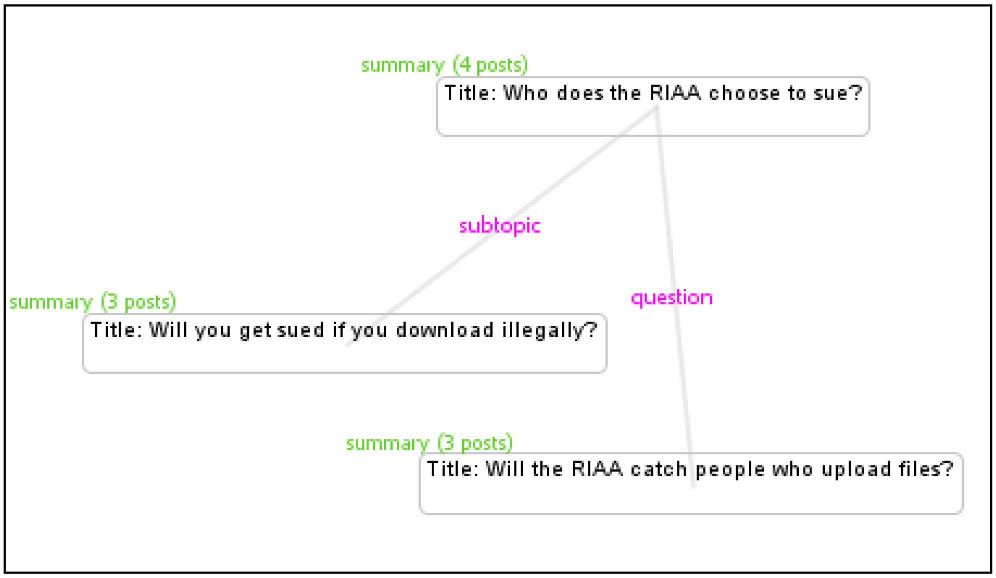

Figure 6-7 Participant Brian's summary nodes. Hovering mouse over each post shows the content. Relationship among the posts is specified. Each summary node also includes "source" posts.

- Upon starting the task, he spent about two minutes to scan through all the posts in the Thread View, from top to bottom. (He later stated he wanted to get a good overall idea what was going on.)
- Then he created the first summary node (the top node in Figure 6-7) with the title "Who does the RIAA choose to sue?" typed in, but with *no* content. He went up and down in the Thread View to find source posts and drag-and-dropped them into the summary node one by one to attach them to the summary node.
- He then created the second summary node (the middle node in Figure 6-7) with the title "Will you get sued if you download illegally?", again with no content. He added source posts to the summary node in the similar manner.
- After browsing some posts again the Thread View, he created the third node with the title "Will the RIAA catch people who upload files?" again with no content,

but only source posts attached to it. At this point, three separate summary nodes were created.

- Then he double clicked the top summary node to open it up in the Advanced Summary Editor (Figure 6-6). He read through the attached source posts, and started writing the summary content.
- He repeated this process for the other two summary nodes as well. When the summary contents were all done, he started creating a link among the nodes. For the top node and the middle one, he specified "subtopic" and for the top node and the bottom one, he chose "question" to indicate the relationships. He added a question "try to find out more info about this" in the summary.

The following shows the content of his summaries (see Figure 6-7 for the graphical representation in Arkose2).

Top node:

```
Title: Who does the RIAA choose to sue?
Content: The RIAA sues individuals who download
copyrighted material. They cannot sue everyone, but
they do make examples of the people they catch. The
better the case they have against a downloader, and
the more evidence they have, the more likely they are
to sue. (4 source posts attached)
```

Middle node:

```
Title: Will you get sued if you download
illegally?
Content: There are millions of downloaders, so
the chances of getting caught are small.
However, the only way to know for sure that one
won't get caught is to not download at all.
(3 source posts attached)
```

Bottom node:

Title: Will the RIAA catch people who upload files?
Content: Programs like BitTorrent automatically use downloaded files to upload files to others.

One post stated that the RIAA will catch uploaders before catching downloaders. (try to find out more info about this)
(3 source posts attached)

I believe Brian's summaries covered the discussion space sufficiently. He pointed

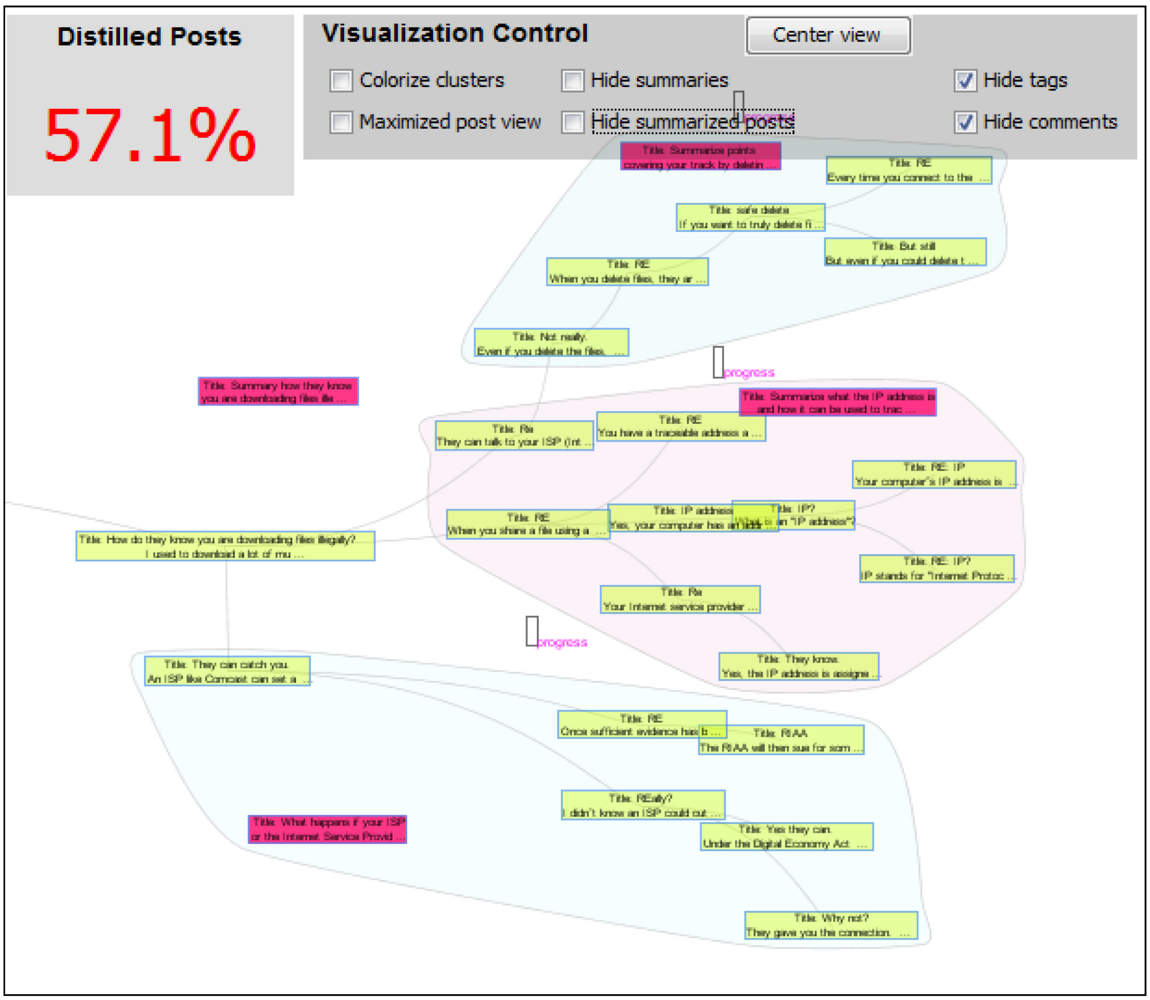

Figure 6-8 Level 2 discussion space with scaffolding. Each group represents a scaffolding group. The red nodes are scaffold instructions for the groups. As the posts in the group get summarized, the progress bar attached to each group is updated.

out all the major topics and included many source posts. Next, I will describe the task details for a level 2 with scaffolding.

With Scaffolding

Almost the same instructions were given to a participant with a discussion space with scaffolding. The only extra instructions were regarding the scaffolding:

- Someone else already created a summarization structure in which you should fill in with summaries and other meta-information.
- For each scaffolding node, related discussion posts are grouped together in the Visual Navigator.

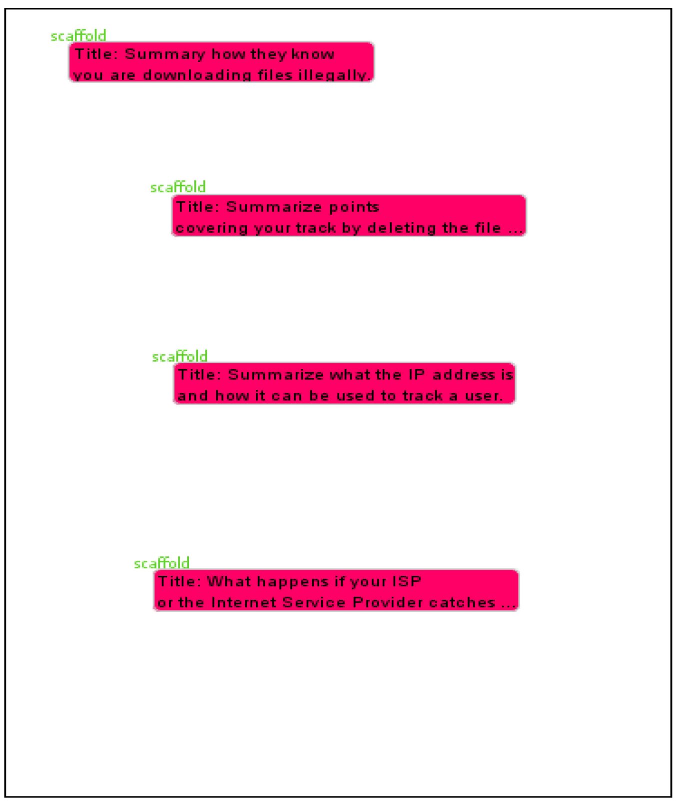

Figure 6-9 Scaffolding nodes in the Summary Editor. These are connected with the scaffolding groups in the Visual Navigator.

- If you don’t find the scaffolding suitable, you may choose to change it or create your own.

The scaffolding instruction nodes appear in two views. The first as groupings in the Visual Navigator (Figure 6-8) and the second as scaffolding nodes in the Summary Editor Space (Figure 6-9). These nodes are interconnected; in other words, clicking a node in one view centers the other view on its counterpart. Each scaffolding node has some instruction about what needs to be done. For example, the very top scaffolding node is “Summarize how they know you are downloading files illegally.” Another scaffolding node says “Summarize what the IP address is and how it can be used to track a user.”

The following is what Brian did in this level.

- After scanning through the discussion space in the Thread View, he read the first few scaffolding instruction nodes in the Summary Editor Space.
- He clicked the top scaffolding node in the Summary Editor Space, which highlighted its counterpart node in the Visual Navigator. He then moved his mouse over to the Visual Navigator to view the content of the posts in the grouping.
- He created a summary node in the Summary Editor Space and typed in the title “P2P programs” and in the content “P2P programs are traceable”.
- He drag-and-dropped two posts into the summary to attach them.

- He then opened the summary in the Advanced Editor and added more content based on the attached posts. He added “A computer's IP address is visible when sharing files.”
- He then moved onto the next scaffolding node, and repeated similar steps and created a summary node under each scaffolding node.

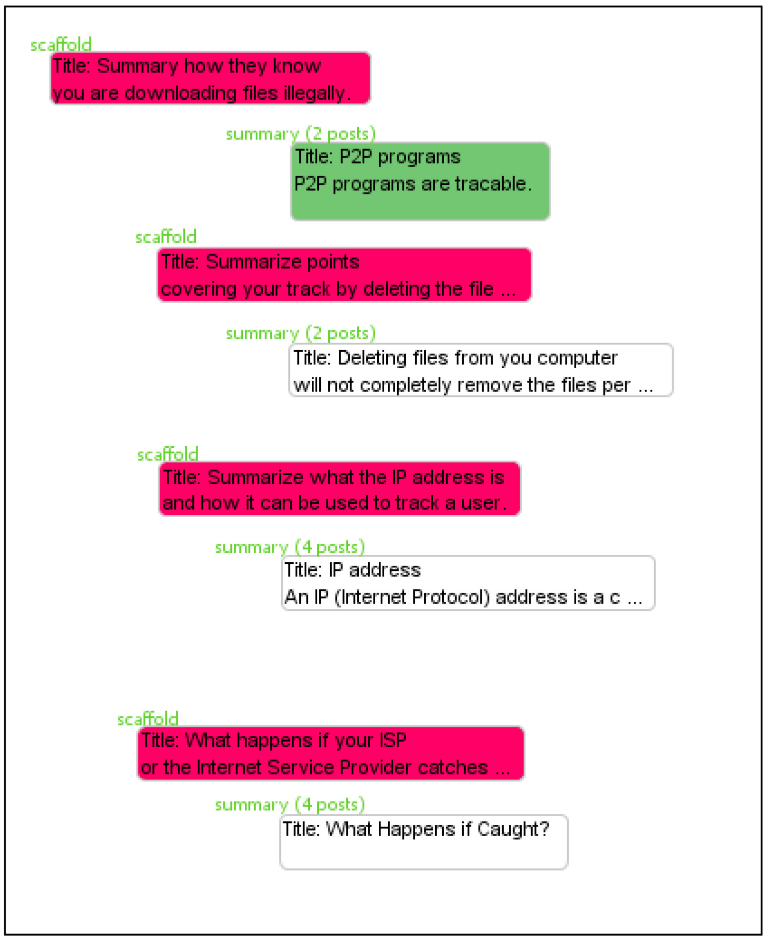

Figure 6-10 Summaries created for the scaffolding structure.

The following is the summaries he created.

(Scaffold) Summarize how they know you are downloading files illegally.

```
Title: P2P programs
Content: P2P programs are traceable. A computer's IP
address is visible when sharing files.
(2 source posts attached)
```

(Scaffold) Summarize points about covering your track by deleting the files.

```
Title: Deleting files from your computer
Content: will not completely remove the files
permanently. Even though they are not visible, they
are still saved on the computer.
(2 source posts attached)
```

(Scaffold) Summarize what the IP address is and how it can be used to track a user.

```
Title: IP address
Content: An IP (Internet Protocol) address is a
computer's address that is given to it by an Internet
Service Provider (ISP). Each computer has a different
IP address.

IP addresses are logged and time-stamped. This is
also true for downloaded files, and because IP
addresses are different for each computer, and given
out by ISPs, they are traceable.
(4 source posts attached)
```

(Scaffold) What happens if your ISP or the Internet Service Provider catches you?

```
Title: What Happens if Caught?
Content: If an ISP catches you downloading files
illegally, you can have your internet connection
removed, a search warrant issued for files on your
computer, and eventually you will be sued by the RIAA.
These lawsuits can be for large amounts of money, but
usually there are court settlements. These fines are
typically in the $5,000 to $10,000 range.
(4 source posts attached)
```

While not every participant's approaches were the same, the above descriptions of two level 2 tasks illustrate what a participant was asked to do in each task, and what actions he could perform and what features he could use to create summaries. Level 3 without scaffolding and level 4 with scaffolding were similarly done in terms of what a participant was asked to do. However, the size and complexity increase in the later levels meant changes in the summarization approaches and reliance on the supportive features of Arkose2 for many of the participants, as I will discuss in later sections.

Summarization details

In this section, I describe the details of discussion data participants summarized, their summarization approaches, and details of the summaries and meta-information created.

Discussion levels

The discussion thread data was modeled following Webb's Depth of Knowledge that defined four basic types of questions based on the knowledge and skills required to answer them. For Webb, level 3 is more complex and difficult than level 2 and level 4 more than level 3 too. This was confirmed by the participants in the evaluation.

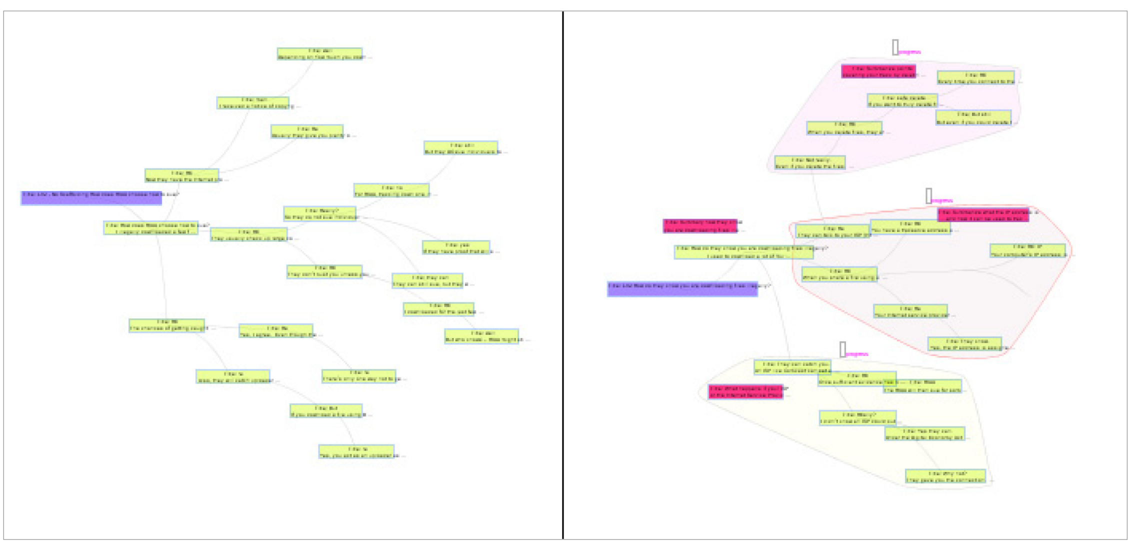

Figure 6-11 Level 2 discussion space without scaffolding (left) and level 2 with scaffolding (right).

Participants stated the two level 2s were easy to summarize. (See Figure 6-11 for the two level 2s.)

> It was fairly straightforward and the discussion wasn't very fragmented. [Matt]
>
> [It was] not really difficult. [Gabriel]

There were two level 2s prepared; one with scaffolding and one without one. While participants felt both of them easy to do, many stated level 2 with scaffolding was slightly easier because the scaffolding broke the discussion space into smaller subsections and told the participants what to do. (More discussion on this in the Scaffolding section.)

The level of difficulty participants felt in level 2 quickly increased in level 3 which had more than twice as many posts and more subtopics. The content was more interrelated and more in-depth as well. Three participants stated the following:

> [Level 3] is definitely harder. It's more of a multi-threaded conversation. [Chris]
>
> This one is more difficult. […] Since there were so many different groupings and pairings that I was seeing, it was a little bit harder than the first time that I did it, a little less straightforward. [Gabriel]
>
> […] there's just more of it and it was more interconnected than the past information's been. […] The difficulty in general was higher because the thoughts from each post were more in-depth. [Brian]

Level 3 was harder for several reasons. It had more posts, had no scaffolding that many participants found helpful in level 2, and the discussion was more in-depth and more scattered around. Similarly, level 4 was much more difficult to summarize than the level 2s. Most participants also felt level 4 was more difficult than level 3, while two participants mentioned they were similar probably because of the scaffolding. Level 4 had many more subtopics and posts than level 3. (See Figure 6-11 and Figure 6-12 for an overview of the four settings.)

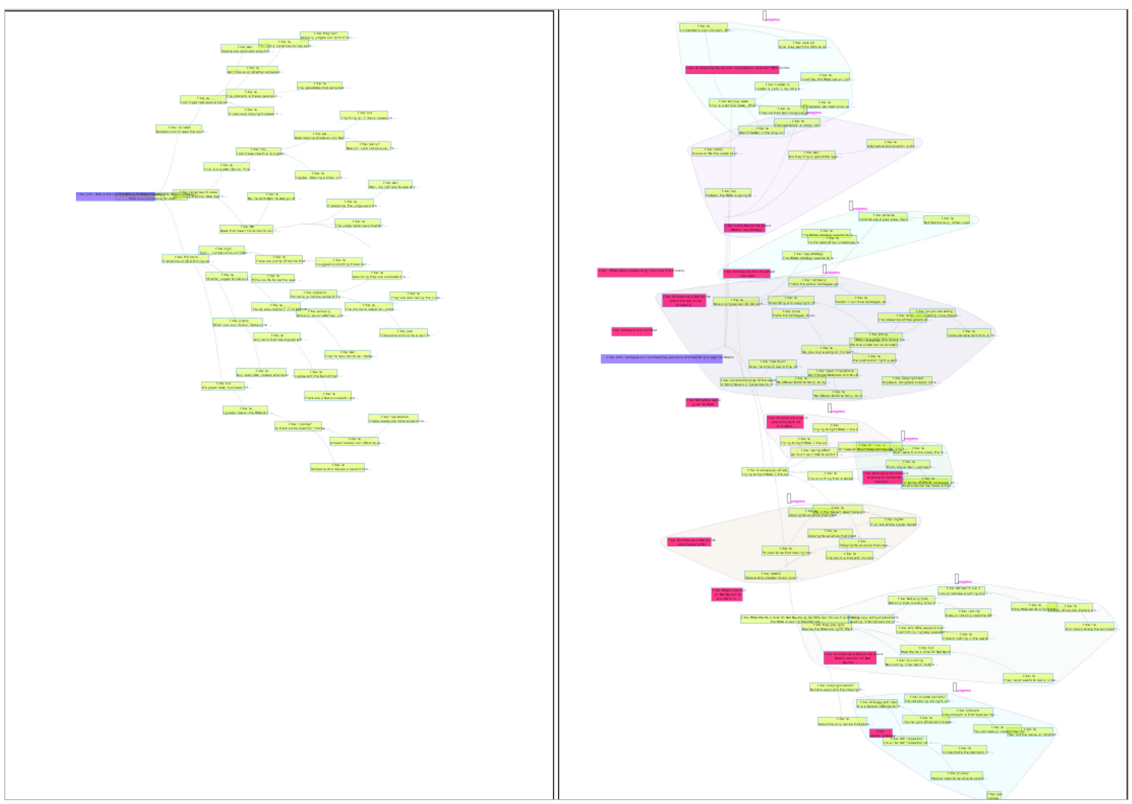

Figure 6-12 Left: level 3 without scaffolding, Right: level 4 with scaffolding

> [Level 4] was probably the hardest. A lot more topics [were] being discussed. Some of them just overlap. [Phil]

> I think just the level of complexity was much higher and there were several points I was confused about what argument was being made by who. [Justin]

I think even with the scaffolding, this was more difficult. [Erin]

I felt like it was as in-depth as [level] 3. […] I think it would have been more difficult [without scaffolding]. Absolutely. [Matt]

For all participants, the difference among the levels was clear and large enough. The changes in the size and complex of the discussion spaces influenced some participants' approaches in the later levels. I next discuss the participants' summarization approaches and their changes according to different settings.

Summarization Approaches

The number of summary nodes created and details included in them varied among the participants. While there may be no "golden rule" about what is a good summary for a discussion thread, one may argue that a good summary succinctly points out most or all the substantial subtopics discussed in the discussion space, with enough detail for the subtopics to make sense to a reader. For such a summary to be useful in a discussion forum setting, it should reduce the amount of information a newcomer has to read in order to understand the broad topics discussed. A good summary should also allow the user to access the original posts for which the summary was made and it would point out topics that would need further discussion, evidence, and/or supporting examples.

All of the participants could write summaries that more or less reduced the original discussion spaces. Thirteen out of fifteen participants produced summaries in which each summary node would summarize several discussion posts, while the other two participants mostly created summary nodes that were close to one-to-one to the discussion posts, summarizing one or two posts at a time. The amount of detail included in each summary node varied among the participants as well. For example, in the Task

Details section, Brian created three summary nodes (Figure 6-7) for level 2 discussion while Emily created more than a dozen (Figure 6-1) for the same discussion. However, each of Brian's summary nodes contained much more text than Emily's summary nodes did.

Through observation and participants' statements about their approaches I categorized their summarization approaches into the following:

1) Read or skim through some or the entire discussion space. As a new subtopic emerges while reading several posts under a similar topic, create one or more summary nodes and attach source posts to them. Move on to another part of the discussion space and continue look for a new subtopic.

2) Initially create almost empty summary nodes with just the title and source posts as place holders. As going through the discussion space again, fill in the structure with content and source posts. This is similar to creating summary scaffolding.

3) Create a summary node for one or two directly connected posts, going through the discussion space from top to bottom. This is similar to approach 1, but while approach 1 summarizes several posts into one summary node, approach 3 is close to "one summary for one post". The way a participant finds related posts differs as well. In approach 1, a participant navigates more diverse locations in the discussion space to find related posts, but in approach 3, a participant only looks at posts in local proximity. This approach is very linear and narrowly localized, without really considering the information being summarized with respect to other parts of the discussion space.

Often, a participant would use more than one approach in a task. Thus it is difficult to say a participant's approach was entirely one or another. Rather, one's strategy seemed to be a mix of one or two, though ten out of fifteen participants primarily used approach 1, three participants approach 2, and the other two participants primarily used approach 3.

I will now describe three participants' summarization as an example in describing each summarization approach.

First, approach 1: participant Erin used the following steps to summarize the level 3 discussion without scaffolding.

- She spent the first 5 minutes going through the posts in the Thread View from top to bottom. After she got to the bottom of the Thread View and read the last post, she went back to the top again.
- Then she created a summary node for the first post and drag-and-dropped the first post into it.

```
Title: Court responses to RIAA Lawsuits
Content: This topic is all related to the legal
aspects of RIAA lawsuits, especially how the courts
have reacted to them.
```

Figure 6-13 The first three summary node and note nodes. Several posts are summarized under each node.

- She read the next few posts and created a note node and drag-and-dropped three consecutive discussion posts into the note.

```
Title: Copyright Law
Content: Discussion of whether or not it is legal or
illegal to share files even if you are simply
creating copies of your own files.
```

- After that, she opened the note node in the Advanced Summary Editor (see Figure 6-6 for an example) to view the attached posts and the summary content she created. She then closed it and created a link between the summary node and the note node she created and specified the relationship as "subtopic".
- Instead of summarizing the posts linearly from top to bottom, she then moved to the Visual Navigator and navigated there. During the interview, she stated the tree structure helped her figure out which posts were related and where the major optics were located. The rest of the task was done by navigating the posts in the Visual Navigator not sequentially but spatially. After reading a few posts she created a note node and drag-and-dropped related posts. (To her, a summary node was a big "topic node" while note nodes were used as subtopic summary nodes.) She also updated the existing summary node and note nodes as new information was found. She repeated this process until the task was done. (Figure 6-14)

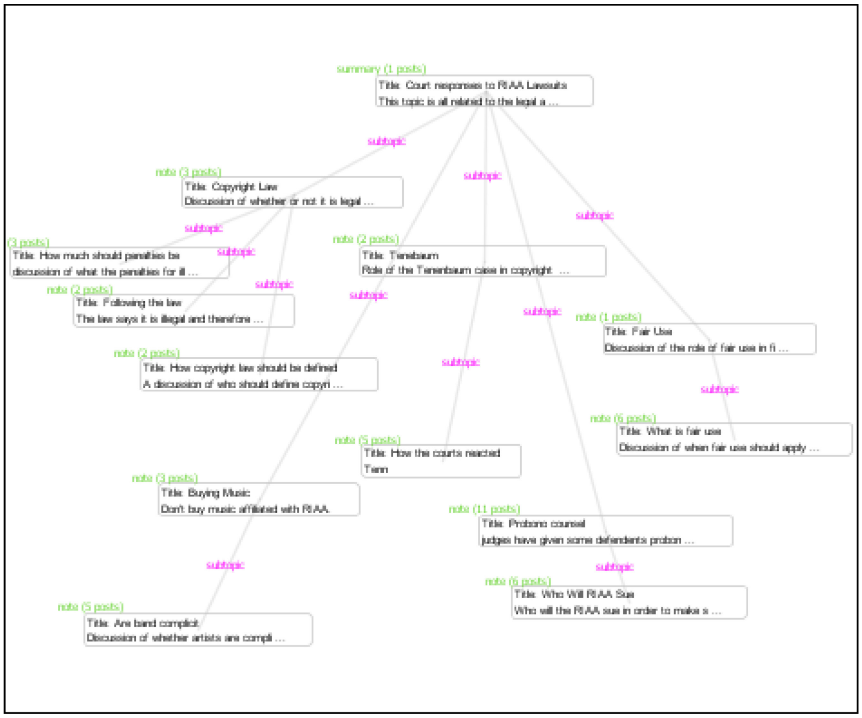

Figure 6-14 Erin's summary structure for the level 3 discussion. Each summary node (or note node that she used) summarizes several posts at a time (2 to 11 posts).

Approach 3 produces many summary nodes due to almost one-to-one summarization of the discussion posts. Thus it does not reduce or distill the original information as much as approaches 1 and 2, but retains more of the original information, both the content and structure. Approaches 1 and 2 reshape and filter out much information and compact the discussion space more. Approaches 1 and 2 also require more reading-at-once and manipulation of the posts before a summary node is created, while approach 3 only looks at locally co-located posts and the cognitive load required does not seem as high. Consequently, the two participants who used approach 3 found scaffolding less useful in their approach because the instructions of most scaffolding nodes asked for a task similar to approach 2 (e.g., "Summarize 3-4 major points about X from these 20 posts). They very much contrast with other participants who found scaffolding very helpful and useful. While the two participants did not follow the scaffolding, they still thought scaffolding was helpful in a way because it showed how many subtopics existed and what the approximate divisions of the space were. (More on this in the Scaffolding section.)

Each summarization type may be useful in its own way. A very compact summary may be good for pointing out big subtopics relatively fast, while a detailed summary may be useful to get more information without reading the original discussion thread. One can envision both summary types coexisting in Arkose2. Some users may point out major information in each post, while other users may combine, filter, and organize those local level summaries into a more global summary.

The above described the approach 1. Now, the approach 2: participant Brian's approach was to create a skeletal structure with source posts first and then write summary contents later. This is essentially scaffolding building.

- As soon as he read the first few posts, he created a summary node with just the title, and drag-and-dropped the posts into the summary.

  ```
  Title: Tenenbaum Trial
  ```

- He repeated this process until he reached the last post. (See Figure 6-15.)
- Then he went through each empty summary node he created. He opened the summary node in the Advanced Editor where all the source posts for the summary appear, and he read the posts again and added summaries to the content.
- He also added more posts to the existing summaries, iteratively updating them. He repeated the process until done.

- He then created links among the summary nodes with relationship such as "subtopic", "evidence", and "example".

Figure 6-15 Brian's initial summary structure for the level 3 discussion. Each summary is almost empty with just the title and source posts. He then iteratively updated the summary content based on the attached source posts.

And, approach 3: participant Emily's approach was to create a very short summary node for almost every post starting from the top post to the bottom post. For example, the first few of her summary nodes and original posts are below: (P) indicates discussion post and (S) indicates a summary.

```
(P)How is the court reacting to RIAA's strategy?  How is
the court reacting to the RIAA's strategy for suing
individuals for downloading files illegally.

      (S)How is the court reacting to the RIAA's strategy
      for suing individuals for downloading files illegally.

(P)Tenenbaum case. Court says that Fair Use may hold in
some RIAA cases.  In Sony BMG Music v. Tenenbaum, the
defendant admitted liability at this trial and the Court
```

```
has sided with the RIAA and granted them a monetary award.
However, the Court recognized there are some cases that
might constitute fair use, such as creation of MP3 files
from audio CDs exclusively for saving the music in another
place.

      (S) Fair use may hold in some cases (such as creation
      of Mp3 files for saving in multiple locations).

(P) This is a quote I found. "The Court, deeply concerned
by the rash of file-sharing lawsuits, the imbalance of
resources between RIAA and individuals being sued, and the
upheaval of norms of behavior brought on by the Internet,
did everything in its power to permit Tenenbaum to make his
best case for fair use."
(P) I don't see how this is a good thing, making 'file
sharing' and 'fair use' synonymous.
(P) I agree.  Making a small clip of a copyrighted work so
the work could be discussed is one thing, but crying 'fair
use' at the last minute of your lost defense only makes you
look desperate.

      (S) File sharing vs. Fair use
```

Note that the last summary consists of just keywords describing the three posts for which the summary is created. For most posts, she summarized the content of each post

Figure 6-16 Brian's final summaries for level 3 discussion. After the initial structure, he iteratively updated the content and source posts.

into a summary node, rather than combining several posts into one cohesive summary. It was hard for her to process several posts together. She stated,

> It would just take me more time for forethought or to plan to create the structure. […] I can’t mentally hold that much information in my head. [Emily]

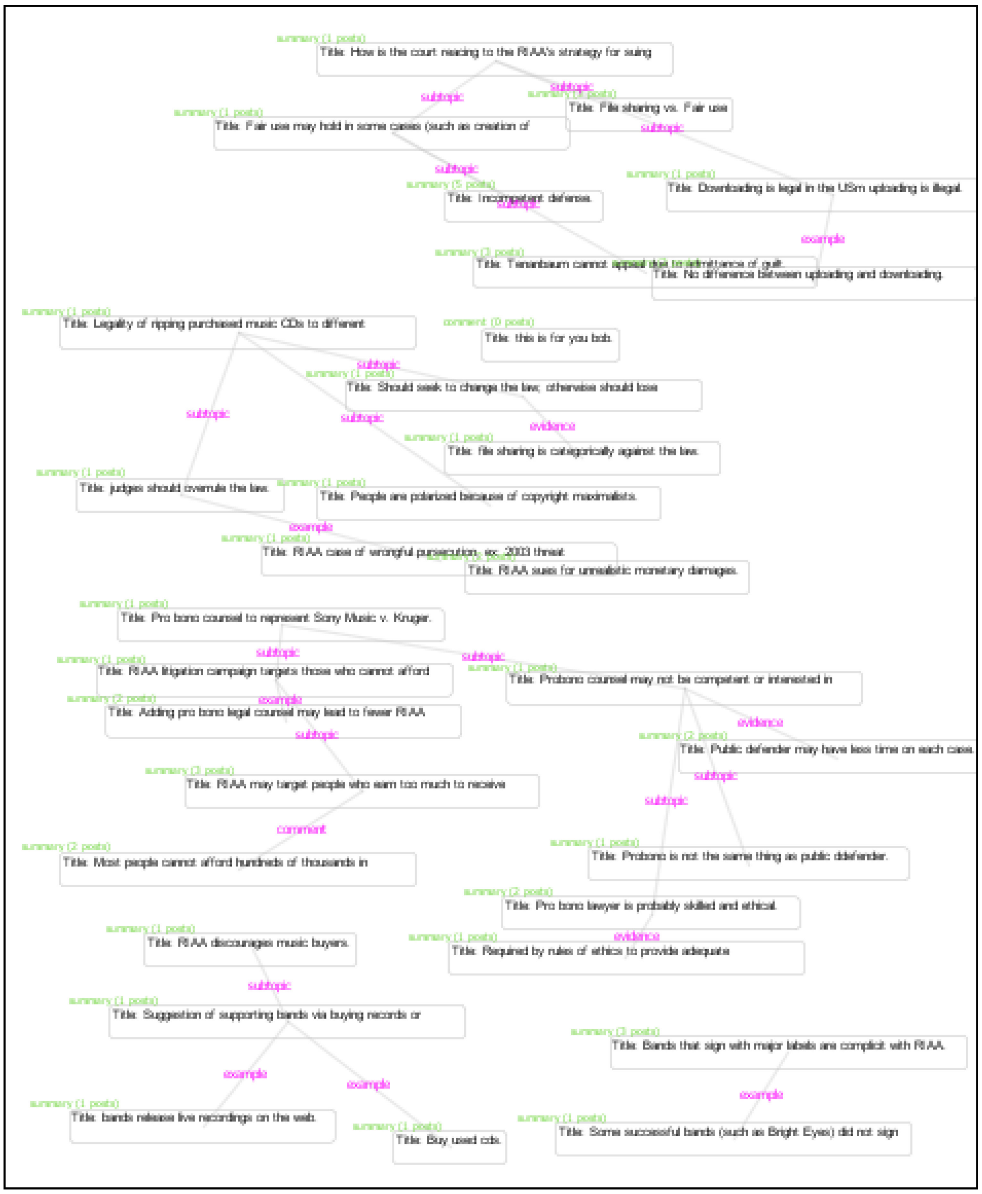

Figure 6-17 Emily's summary structure for level 3 (Summarization approach 3). One summary node covers only one or two discussion posts. Compare to Brian’s approach 2 (Figure 6-16) and Erin’s approach 1 (Figure 6-14).

Another participant, Josh, used a similar approach of creating localized summaries for one or two discussion posts. Brian, Matt, and Daniel primarily used approach 2 of skeletal structure and filling in it iteratively. The rest primarily used approach 1.

	Approach 1	Approach 2	Approach 3
Number of participants who primarily used the approach	10	3	2

Table 6-5 Number of participants who primarily used the approach. Note that it is sometimes a mix of approaches as well.

Approach 1 and 2 produce summaries that compress the discussion space much because each summary node summarizes several posts. Approach 3 produces one summary per one post most of the time. Both summarization types may be useful in their own way. A very compact summary may be good for pointing out big subtopics relatively fast, while a detailed summary may be useful to get more information without reading the original discussion thread. One can envision both summary types coexist in Arkose2. Some users may point out major information out of each post, while other users may combine, filter, and organize those local level summaries into a more global summary.

Summary Details

Most participants summarized to create an objective view of the discussion space. Twelve participants (of which two also stated they summarized for themselves) stated that they were summarizing for other users so that they might benefit from the summaries.

> […] by reading my summary people will be able to understand [quickly] what was going on in the entire conversation, so they'll see what kind of topics discussed and what are the main key points of the discussion. […] And if the user wants to look more detail then they are allowed to [read] related posts. [Daniel]

> My goal is to make things readable, easy for people to quickly scan through and select what they want, and not select what they don't want. [Chris]

> I was trying to summarize so that anybody that wanted to go to this forum would be able to look at my summarization and get idea about the conclusions that were reached or pointed towards. [Matt]

This led them to think about the expertise levels and background of their potential readers when creating the summaries.

> (When asked why the participant included particular information) Because I think that might be relevant as more and more people, particularly maybe older people are getting involved in online forums and copyrighted file sharing discussions. Like… somebody in my parents' age perhaps might feel more that way, than somebody of my generation. And so, they might think it's relevant. [Matt]

As well, the participants who summarized for other users tried to be inclusive and objective when including information. When asked whether the summary structure and content would be different if they were summarizing just for themselves, two stated they would. They stated the summaries would be less objective and thorough in coverage, and rather more selective and fitted to personal needs.

> For myself? Yeah, definitely [different]. I would probably not care about what other people's preferences… […] So I might probably spend more effort on things that I don't know, something that I really want to know. [Phil]

Two participants stated they were summarizing more for their own use. While there was no noticeable difference in the summarization process, the participants stated they would be more detailed and carefully done if they were summarizing for others,

> (When asked what would change when summarizing for others) I would make the [summary] title more descriptive for others. Because for some of the titles, I did not put in complete sentences, but I just put like a word or two and then a dash and a word or two. [Ellen]

> I think maybe if I summarized for other people, I would probably a lot more explicit and careful about the language I use. [Erin]

These findings indicate that the target audience or users of the summaries would influence the details of the summaries. Participants stated that performing summarization of a discussion thread would be useful to others to understand the topics in the discussion space, recall information later, filter out unnecessary information, and generally reduce the amount of information.

> I like that you can create an outline… It's kind of hard to get any information out of this type of [discussion] format. [Tom]

> I like the idea of doing [summarization] because a lot of it becomes repetitive and hard to filter through because of the structure. So I like the idea of getting to the gist of what people are saying. [Morgan]

> Once you get up to fifty posts and some of [the posts] are somewhat lengthy, I think [summaries] would be more useful that there's a lot more information here and it would take a lot of time to read through it all, so having summaries would be useful. [Justin]

But at the same time a summary of a discussion space does not replace it, but acts more of a structured pointers to more content to look up.

> I don't think [a summary] is a good replacement for the posts, but I think it's a good, kind of an index in the sense that I think it's a quick way to look for certain topics that you are more interested, but I couldn't study for a test off of these summaries. I think I thought of this more as an index where I could click on it and it would go to the thing. [Emily]

> I think it's like having a table of contents for a book. It's the same kind of idea and it would help people to navigate to different parts. [Erin]

> People post fresh questions as opposed to going and searching for it. But if they had a structure where it was a little easier to navigate, they might not. [Morgan]

A discussion thread may have different types of information. Within the conversations in the discussion space, participants identified and distinguished more objective “factual” information and subjective “opinions (arguments)” when asked what information was saved in the summaries. Factual information tend to be information related to keywords such as copyright, Net neutrality, file sharing, pro bono, and so forth, while opinions included emotional comments, suggestions, advice, and so forth. Participants stated they generally saved more factual information than pure opinions. Even when they saved opinions, they had to have some important information or supporting evidence in them.

> Basically, the first priority is for the facts, and the second for the opinions. If an opinion is informative like if it’s telling something related, then I’ll save it. [Josh]

Often for participants, there was not a clear line between what was factual information and what was opinion.

> It’s a little hard to tell where the line is between an opinion and fact. If this person said it, do I trust him or what? It’s a little harder to filter them out this time (level 4). They are all mixed up together. [Gabriel]

When saving opinions, participants stated they did or would save opinions by listing each side of the arguments (either in the same summary node or by separating each out) and put the “gist” of the arguments along with any supporting evidence.

> “Well, many people believe and think this, and other people think that.” […] And in some cases, I just split it up and created two different summaries [for each opinion]. [Justin]

> (When asked how the participant handled conflicting view points) I actually kept them together. One argument with a pro and a con.[…]It's better to keep together. You want one argument. [Chris]

> […] these are all opinions, basically. And I just left a note saying this is all questionable. We don't really have any evidence. [Justin]

The nature of the topic of a discussion seemed to influence what types of information is gathered within the discussion space, and in turn they affect what information is saved in the summaries. Five participants stated that in level 2s whose topics were "How do they know if you are downloading files illegally" and "How does the RIAA choose who to sue?", factual information was saved more or more important than opinions. This observation changed somewhat in level 3 and 4, where the topics were "How is the court reacting to RIAA's strategy?" and "Compare and contrast the positions of those for and against RIAA's strategies." in which the participants stated opinions seemed either more saved or more important than in level 2s. The nature of the topics of level 3 and 4 may have brought in more in-depth discussion among the posters with more arguments whereas more factual information existed in level 2s.

> (In level 2) I think it's more important to summarize with the facts because [the discussion] is very factually based, so I would say that. [Rob]

> (In level 2) This is more of fact based information, I think. Not really opinion based. (In level 3) This time there are more opinions. [Daniel]

> (In level 3) It's certainly more argument based than the previous [level 2]. It just seemed like the posters first of all were more opinionated about [the topic], and it's generally less objective nature especially take the first subtopic, I mean, "It shouldn't be illegal", or the anger [towards the RIAA], that's purely an opinion. [Travis]

Three participants also mentioned summarizing facts was easier than summarizing opinions.

> I think facts are easier [to summarize] than opinions. [Morgan]
>
> I think some of the opinions sometimes are really difficult to summarize, because you'd have an opinion in there that just didn't fit in with anything. [Erin]
>
> (When summarizing opinions) if you separate them out, you lose the context of the argument. And if you put them together, you also potentially collapse it too closely so that you don't get the two different sides of it. [Justin]

Participants stated the importance of information and worthiness of saving was estimated in several ways. One was when they knew the topic well, and already knew what was important and interesting information to save. Another was the keywords used in the discussion. One participant stated that even though she did not know what pro bono lawyers or Net Neutrality meant, they seemed important and summarized the discussion surrounding them.

In summary, there were a variety of factors that influenced the details and information in the summaries created by the participants. Participants had in mind how the summaries would be used and who would use them, and these affected the shape the summaries. The types and difficulty of information and participants' familiarity with the topics also influenced what information was saved and how it was saved in the summaries.

Incrementally Constructing Summaries

Summarization was an incremental process. Participants sometimes updated or changed summaries that they had created as more information was gathered.

I actually changed some of [the summaries later]. I thought that there was redundancy between the two. [Ellen]

Two participants used note nodes or comment nodes in the Summary Editor Space for collecting related source posts into them (as a temporary storage) so they could create a summary based on the source posts.

I just started making [summary nodes] and set them aside. And then, when I finished pulling out posts and putting them into my summaries, I went back and organized the thoughts. [Brian]

I just feel like summary would probably happen after I re-look at [my comment nodes] and do it one more time […]. I think a summary would be like going back through what I created and understanding if my comments are valuable, if they are reliable, credible, and then making the summaries. [Lisa]

This has some implications for the design of an application such as Arkose2 that the application needs to provide a workspace-like feature where a user may collect information, manipulate it, and come back to unfinished work, allowing the work to be done incrementally. After it's created the user can further make changes to the node by double clicking it (or by selecting an entry in the right-click popup menu), which will bring up Figure 6-18. This feature was helpful for the participants who wanted to collect related posts together and then start creating a summary.

Not only was individual summarization incrementally done, participants understood that summarizing a discussion space as a whole could be done collaboratively with other users. A discussion space might have posts on several topics, of which some topic may not have gotten enough attention. It would be useful to guide the direction of a discussion if the participants or moderators of a discussion thread could point out topics that would need more discussion, claims that would evidence, and information that would need examples, and so forth. Throughout the summarization tasks, four participants left meta-information in the summary space for other users (and also for themselves for later) with their judgment on the completeness of a discussion, asking questions when some information did not make sense, and recommending further actions such as "need more examples."

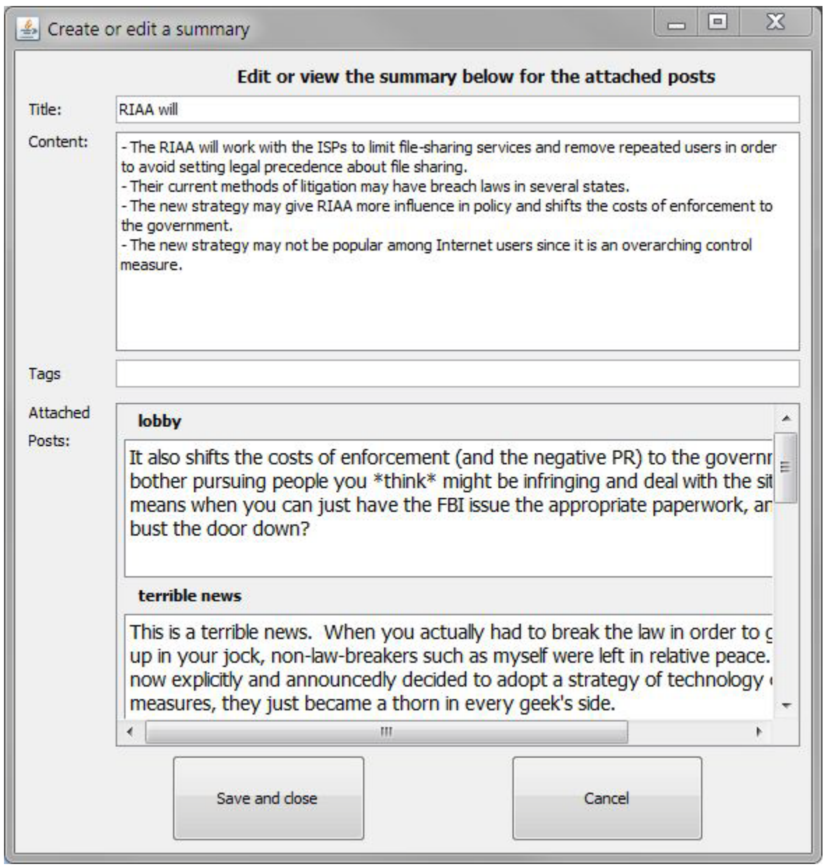

Figure 6-18 The Advanced Summary Editor GUI. User can add more detail as well as read the attached "source" posts.

> […] In the end I added a subtopic on "there is no discussion about what people qualify for a pro bono." [Phil]
>
> I added [..] that experiential evidence is helpful. And then, I asked a question,[…] "What other experiences have you all had?" as if I am talking to the other people and, "Do you know of anyone yourself who has been sued [by

the RIAA]?" […] trying to evoke more of that type of [discussion]. Justin]

> (When asked why the participant was posing a new question for the discussion) I think that might help bring the discussion more full circle to better answer the question that was originally posed. [Matt]

Matt also mentioned this moderator's role would be better if done organically, but it would be difficult if he were forced to come up with a task.

> It would be nice if [giving the people in the discussion space a direction or instruction] happened organically. In [level 2], I posted a question that was my own question. If it happened organically, […] That would be great, because it would be more natural. Obviously it would be harder to force myself to formulate a question to stimulate their discussion. I think it would feel more like a chore. [Matt]

Another participant expressed a similar sentiment.

> Rather than trying to get people to do what you want [by telling them what to do], the best thing would be if I want to comment more on the subject, write something on the subject, […] and participate in the forum than just, "Hey you guys, would you mind talking about?" [Gabriel]

Three participants commented that whatever information they felt was difficult or could not be confirmed might be handed off to other editors since in a real world situation the summarization would be done collaboratively.

> If I am not familiar with [the information], I had to write notes [rather than creating summary nodes]. And those will be things that I, or a research partner would have to go back and find the resources. [Brian]
>
> I assume someone is going to double check [the correctness of information], like it is a collaborative effort. You try to write down what is… you think is best and then you allow someone else to look at it and then double check it and I assume if you have a lot of people [looking at it] then [it will] converge to some right information. I think that's how wiki works anyway. [Daniel]

> I thought that it would be better to include [certain discussion] and let somebody else decide if it was pointless to include it or not, rather than not include it. [Matt]

One participant stated that it would be even easier to work on top of a summary that was started by someone else than to start from scratch.

> I feel that it would be easier to add onto an existing summary, because my only task will be what the previous person has missed and I can just give additional points. So that would be easier. [Phil]

These statements indicate that summaries can be incrementally and collaboratively built using Arkose2.

Where does difficulty come from?

As discussed in earlier sections participants felt summarization more difficult when the discussion space was larger and more complex. In general, participants had to explore the discussion space more, read more posts, organize more ideas into a summary, and kept more information in head at the same time. This led to relying more on the supportive capabilities of Arkose2 such as various visual aids, search features, and automatic clustering, as discussed in earlier sections.

> I feel there are more sub-topics and it's hard for me. It would just take me more time to forethought or plan to create the structure. […] I can't mentally hold that much information in my head. [Emily]

There were also other factors that made it more difficult for a participant to summarize information effectively. The most frequent reason was the participant's lack of background knowledge in the topics being discussed or specific terminology used.

> I would say that [level 3] was maybe twice as hard for various reasons. One, I don't know the topic area and

there is a lot more legal stuff in this. […] I think sometimes if the meaning is ambiguous, it's also very difficult [to summarize]. [Emily]

> I would say if the person [summarizing the discussion thread] doesn't know anything about the topic, it might be difficult to summarize it. Based on my knowledge I can extract what's important and not. [Daniel]

When summarizing information with which participants were not familiar, they would often use more of the original sentences found in the posts, or just grouped several posts under a very simple summary, retaining more of the original information.

> For the things I did not know, I included more details and then actually quoted some of the sentences [in the summary]. But for the things I already knew, I just included points [about them]. I actually wrote in my own sentences [the things I knew well], but more copy and paste [the things I didn't know well]. [Ellen]

It is normal in a discussion thread that many people bring in different aspects of a topic. Often people's arguments go off in many directions even when they are discussing under a common topic. And sometimes related discussions are going on in different places in the discussion space. The scatteredness of information contributed in the difficulty some participants felt when summarizing.

> It's hard to summarize because, they are not writing these discussions like if they were writing books. So it's hard to extract information. [Daniel]

> [Level 3] was somewhat [more difficult]. [Posts] are under the same topic and they are related to each other, even though they are very far apart from each other in the tree. [Gabriel]

One participant found summarizing for arguments under a topic was generally more difficult when they did not reach a clear conclusion.

> It's hard to be completely objective. Because there haven't been any hard and fast rule set [about the topic]. […] I think that it's still kind of up in the air about

> exactly how copyright infringement will ultimately be defined. [Matt]

The majority of people's summarization experiences have been on summarizing an article or book, rather than a conversation or discussion (see Table 6-1 for participants' prior summarization experiences). Five participants stated that an article or paper had a clear structure built-in in the content such as the abstract, introduction, chapters, and so forth whereas a discussion thread did not.

> Because a conversation can go off many trails and you can't really control a conversation in the same way a book or article, something someone has thought a lot about. It's been read and edited. They are already divided into sections for you, so I would just use their own [structure]. [Gabriel]

> I feel like it's more fragmented and more schizophrenic. It's more prone to be all over the place than a book or an article. You usually have some sort of a structure. [Justin]

All of them stated summarizing a discussion thread was more difficult than summarizing an article or a book, because of the lack of a structure often found in an article or book.

> Oh, yeah. [Summarizing a discussion thread] is far more difficult. Far more difficult, because you have to impose your structure yourself. [Erin]

> Basically, with [a discussion thread] you've got lots of tiny abstracts. Whereas with a book or article, if the authors did a good job, one paragraph is on one subtopic and that makes it easier. And obviously a more difficult thing about a thread is people can just post whatever they wanted in whatever order as we saw in some of the threads. So that makes it more difficult. [Travis]

In summary, there were several factors that made summarization of a discussion space more difficult. The lack of background knowledge in the topic made it harder to know what information to save. Issues specific to distilling a discussion space existed as well. Often, discussions that included many opinions were difficult to summarize, especially when there was no clear conclusion. The lack of a content structure in a discussion space also meant imposing an explicit structure when summarizing a discussion.

Design claims discussion

Several claims about the helpfulness of Arkose2's features were examined in the evaluation. First, participants could create summaries incrementally using Arkose2. Second, scaffolding supported their summarization tasks. And third, specific features of Arkose2 helped the activities in summarization. I discuss these in turn.

Creation of Summaries using Arkose2

I believe all participants could create reasonable summaries for the discussion spaces using Arkose2. Participants' summarization goals and idea of an ideal summary varied. For example, Chris, the former journalist, stated his goal was to write summaries that were "readable, easy for people to quickly scan through and select what they want." He wanted to write simple summaries that would grab readers' attention and convey the meaning at the same time. For example, he wrote the following summary for three discussion posts talking about the RIAA's switch in strategy not to go after individual copyright violators (also see Figure 6-19 for the actual representation in Arkose2). ((S) denotes a summary node, (P) a post.)

```
(S) Title: What's the reason?
     Why did the RIAA switch tactics?
    (attached posts: 3)
```

Although his summary does not contain much information, clicking it highlights the following attached source posts in the Thread View, so a user who wants to find out more about it can read them quickly.

```
(P) This has less to do with the RIAA deciding to switch
tactics in enforcing copyrights and it has more to do with
the RIAA not wanting a legal precedent set about file
sharing.

(P) Did they finally get some legal advice?  I mean, their
current methods have apparently at least been in breach of
investigative laws in several states and they may still end
up in mess because of it, but ending the thing will at
least lessen the exposure…

(P) Alternative explanation is that they have actually
understood that extortion is bad.. nah.. not likely.
```

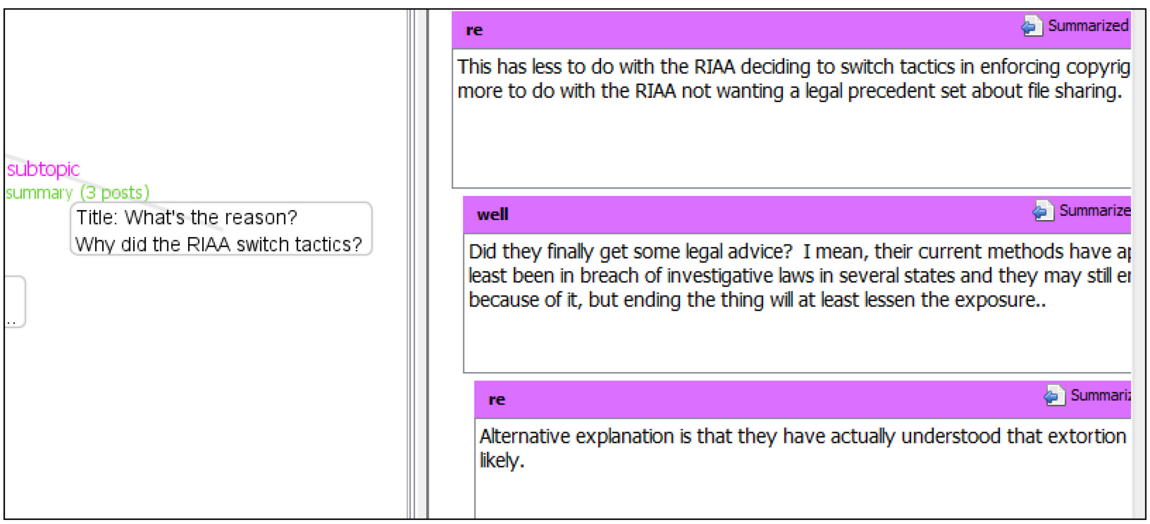

Figure 6-19 One of Chris' summary nodes (on the left). Clicking the summary node highlights attached source posts as shown. His goal was to write a simple "headline" that can guide users to quickly find information they want.

For the same three posts, another participant Emily created three summary nodes, one for each post:

```
(S) RIAA does not want to set a legal precedent about file
sharing.
(attached post: 1)
```

```
(S) Methods are a breach of investigative laws.
(attached posts: 1)

(S) Not likely that RIAA understands that extortion is bad.
(attached posts: 1)
```

And another participant, Brian, created one summary that included other related posts that would answer the question. His summary included 5 posts that were in different places in the discussion space:

```
(S) Why change tactics?
It seems that the RIAA has stopped prosecuting individuals
for two reasons. One, that they would prefer to spend their
resources on behind-the-scenes lobbying instead of
unpopular trials. Secondly, they may not want to risk
losing a case. If they lose, and a precedent is set that it
is unconstitutional to limit the distribution of
information, it could become impossible to sell music at
all.
(attached posts: 5)
```

The three participants' summaries varied very much. Chris created a "headline" for several posts to grab attention but without much information, Emily created three summary nodes with each containing some information from the posts attached, and Brian created one summary that organized information from several posts in a coherent, structured way. It is very hard to objectively determine which summary is the best, since the purpose of the summaries participants had in mind was all very different. Each of them, however, could be useful in its own way and co-exist in the summary space in a complementary way. Readers of the discussion space could use Chris' headlines to quickly tell what subtopics existed and decide which part of the discussion space they might want to read further. They could use Emily's summaries to quickly skim through long posts and read just the gist of each post, and use Brian's structured and coherent summary to get an overall conclusion to several discussion posts.

	Level 2 w/o scaffolding	Level 2 with scaffolding	Level 3 w/o scaffolding	Level 4 with scaffolding	Average
Phil	0.50	1.00	0.82	0.54	**0.72**
Emily	1.00	0.50	1.00	1.00	**0.88**
Brian	1.00	1.00	1.00	1.00	**1.00**
Ellen	0.83	1.00	1.00	1.00	**0.96**
Rob	0.83	0.83	0.94	0.82	**0.86**
Daniel	0.50	0.83	0.94	0.68	**0.74**
Matt	1.00	1.00	1.00	1.00	**1.00**
Lisa	1.00	1.00	1.00	1.00	**1.00**
Josh	1.00	0.83	1.00	0.75	**0.90**
Morgan	1.00	0.83	1.00	0.75	**0.90**
Erin	1.00	1.00	1.00	1.00	**1.00**
Gabriel	1.00	0.83	0.44	0.46	**0.68**
Justin	0.83	1.00	0.81	0.39	**0.76**
Chris	1.00	1.00	1.00	0.61	**0.90**
Travis	0.83	1.00	1.00	1.00	**0.96**
Average	**0.89**	**0.91**	**0.93**	**0.80**	**0.88**

Table 6-6 Participants' summarization coverage of the topics.

	Level 2 w/o scaffolding	Level 2 with scaffolding	Level 3 w/o scaffolding	Level 4 with scaffolding	Average
Phil	0.30	0.19	0.17	0.18	**0.21**
Emily	0.70	0.10	0.62	0.59	**0.74**
Brian	0.25	0.19	0.16	0.16	**0.19**
Ellen	0.66	0.48	0.38	0.30	**0.46**
Rob	0.36	0.75	0.30	0.31	**0.43**
Daniel	0.30	0.17	0.15	0.18	**0.20**
Matt	0.15	0.24	0.32	0.20	**0.23**
Lisa	0.30	0.19	0.18	0.13	**0.20**
Josh	0.65	0.92	0.92	0.67	**0.79**
Morgan	0.45	0.34	0.30	0.21	**0.33**
Erin	0.4	0.38	0.26	0.31	**0.34**
Gabriel	0.35	0.29	0.82	0.21	**0.42**
Justin	0.66	0.19	0.42	0.53	**0.45**
Chris	0.15	0.24	0.24	0.30	**0.23**
Travis	0.66	0.29	0.38	0.22	**0.39**
Average	**0.42**	**0.39**	**0.37**	**0.30**	**0.37**

Table 6-7 The ratio (number of summary nodes / number of discussion posts)

Although it is very difficult to devise an objective measure for the quality of a summary, I use a simple score tuple to show an approximation of how good the participants' summaries were. Here I only use the number of sub-topics covered and reduction of the original space as the components, but one may create an equation that considers human evaluations on coherence, cohesion, and reiteration (Kintsch et al. 2000) or uses a more complex natural language process (e.g. Lin and Hovy 2003).

	Level 2 w/o scaffolding	**Level 2 with scaffolding**	**Level 3 w/o scaffolding**	**Level 4 with scaffolding**	**Average**
Phil	0.14	0.10	0.06	0.04	**0.08**
Emily	0.27	0.49	0.14	0.10	**0.25**
Brian	0.09	0.07	0.01	0.01	**0.05**
Ellen	0.30	0.18	0.08	0.06	**0.16**
Rob	0.13	0.28	0.07	0.08	**0.14**
Daniel	0.11	0.06	0.02	0.02	**0.05**
Matt	0.10	0.12	0.06	0.03	**0.08**
Lisa	0.09	0.06	0.01	0.01	**0.04**
Josh	0.22	0.29	0.22	0.13	**0.22**
Morgan	0.22	0.15	0.04	0.04	**0.11**
Erin	0.21	0.20	0.07	0.07	**0.14**
Gabriel	0.13	0.14	0.18	0.04	**0.13**
Justin	0.30	0.10	0.11	0.13	**0.16**
Chris	0.09	0.15	0.07	0.08	**0.10**
Travis	0.19	0.13	0.06	0.04	**0.11**
Average	**0.17**	**0.17**	**0.08**	**0.06**	**0.12**

Table 6-8 Ratio (number of summary words / number of discussion post words)

	Level 2 w/o scaffolding	**Level 2 with scaffolding**	**Level 3 w/o scaffolding**	**Level 4 with scaffolding**	**Average**
Phil	0.22	0.15	0.11	0.11	**0.15**
Emily	0.48	0.77	0.38	0.35	**0.50**
Brian	0.17	0.13	0.09	0.09	**0.12**
Ellen	0.48	0.33	0.23	0.18	**0.31**
Rob	0.25	0.51	0.18	0.19	**0.28**
Daniel	0.20	0.12	0.08	0.10	**0.12**
Matt	0.12	0.18	0.19	0.12	**0.15**
Lisa	0.19	0.13	0.10	0.07	**0.12**
Josh	0.44	0.61	0.57	0.40	**0.50**
Morgan	0.33	0.25	0.17	0.12	**0.22**
Erin	0.31	0.29	0.17	0.19	**0.24**
Gabriel	0.24	0.22	0.50	0.13	**0.27**
Justin	0.48	0.15	0.27	0.33	**0.30**
Chris	0.12	0.20	0.16	0.19	**0.16**
Travis	0.43	0.21	0.22	0.13	**0.25**
Average	**0.30**	**0.28**	**0.23**	**0.18**	**0.25**

Table 6-9 Reduction score (average of Table 6-7 and Table 6-8)

First, I use the same method of counting the number of subtopics covered as I do in estimating the task completion (see Table 6-4). If a participant did not finish the task on time, the subtopic he was working on but did not complete is counted as a half. So if a participant was working on the last subtopic in level 2 that had 3 subtopics, the score would be 2.5/3 = 0.83. See Table 6-6 for the participants' summary coverage scores.

Second, I calculate the reduction rate (the ratio between the summary space and discussion space) using two measures. One is the number of summary nodes over the number of the discussion posts and the other is the number of the words used in the summaries over the number of the words in the discussion space. The reasoning for these two measures is that some participants created a smaller number of summary nodes with more detail in each while other users created many more summary nodes but each contained less detail. To reflect these two factors, I simply calculate the average of the two for the reduction rate. One may weigh each differently to emphasize one factor over the other. In summary, the tuple (coverage, reduction) score for the quality of summaries is:

$$
\begin{aligned}
&Score = \left(\frac{num\_topics\_summarized}{num\_total\_topics}, \beta_1\left(\frac{\#summary\_nodes}{\#discussion\_posts}\right) + \beta_2\left(\frac{\#words\_in\_summary\_nodes}{\#words\_in\_discussion\_posts}\right)\right)\\
&, where\\
&\beta_1 + \beta_2 = 1.0\\
&and\\
&0.0 \le \beta_1, \beta_2 \le 1.0
\end{aligned}
$$

Equation 6-1 The equation for calculating the summary score tuple that considers coverage and reduction of the discussion space.

By varying the weights one can calculate scores that emphasize different factors in summarization. For the analysis in this thesis, I simply use equal weights. The following table shows the summarization scores calculated following Equation 6-1 for each participant in each task (the reduction ratio is adjusted to reflect only the summarized sub-topics).

	Level 2 w/o scaffolding	**Level 2 with scaffolding**	**Level 3 w/o scaffolding**	**Level 4 with scaffolding**	**Average**
Phil	(0.50, 0.22)	(1.00, 0.15)	(0.82, 0.11)	(0.54, 0.11)	**(0.72, 0.15)**
Emily	(1.00, 0.48)	(0.50, 0.77)	(1.00, 0.38)	(1.00, 0.35)	**(0.88, 0.50)**
Brian	(1.00, 0.17)	(1.00, 0.13)	(1.00, 0.09)	(1.00, 0.09)	**(1.00, 0.12)**
Ellen	(0.83, 0.48)	(1.00, 0.33)	(1.00, 0.23)	(1.00, 0.18)	**(0.96, 0.31)**
Rob	(0.83, 0.25)	(0.83, 0.51)	(0.94, 0.18)	(0.82, 0.19)	**(0.86, 0.28)**
Daniel	(0.50, 0.20)	(0.83, 0.17)	(0.94, 0.08)	(0.68, 0.10)	**(0.74, 0.12)**
Matt	(1.00, 0.12)	(1.00, 0.18)	(1.00, 0.19)	(1.00, 0.12)	**(1.00, 0.15)**
Lisa	(1.00, 0.19)	(1.00, 0.13)	(1.00, 0.10)	(1.00, 0.07)	**(1.00, 0.12)**
Josh	(1.00, 0.44)	(0.83, 0.61)	(1.00, 0.57)	(0.75, 0.40)	**(0.90, 0.50)**
Morgan	(1.00, 0.33)	(0.83, 0.25)	(1.00, 0.17)	(0.75, 0.12)	**(0.90, 0.22)**
Erin	(1.00, 0.31)	(1.00, 0.29)	(1.00, 0.17)	(1.00, 0.19)	**(1.00, 0.24)**
Gabriel	(1.00, 0.24)	(0.83, 0.22)	(0.44, 0.50)	(0.46, 0.13)	**(0.68, 0.27)**
Justin	(0.83, 0.48)	(1.00, 0.15)	(0.81, 0.27)	(0.39, 0.33)	**(0.76, 0.30)**
Chris	(1.00, 0.12)	(1.00, 0.20)	(1.00, 0.16)	(0.61, 0.19)	**(0.90, 0.16)**
Travis	(0.83, 0.43)	(1.00, 0.21)	(1.00, 0.22)	(1.00, 0.13)	**(0.96, 0.25)**
Average	**(0.89, 0.30)**	**(0.91, 0.28)**	**(0.93, 0.23)**	**(0.8, 0.18)**	**(0.88, 0.25)**

Table 6-10 Summarization score tuples. (Coverage, Reduction) value for each task.

While the summaries varied, I believe all of the participants in the evaluation produced reasonable amount of summaries that covered many of the subtopics and details of the discussion spaces. (More examples of summaries created by the participants can be found in the Summarization Details section.) On average, participants covered close to 90% of the subtopics and reduced the discussion spaces to one quarter of the original size.

I performed paired t-tests to see whether the complexity and size and the presence of scaffolding made any difference. Since there are missing conditions from the task settings (e.g., level 3 with scaffolding and level 4 without scaffolding were omitted.), I can only compare the following three task conditions:

1) Scaffolding condition: between level 2 w/o scaffolding and level 2 w/ scaffolding

2) Complexity change in no scaffolding tasks: between level 2 w/o scaffolding and level 3 w/o scaffolding

3) Complexity change in scaffolding tasks: between level 2 w/ scaffolding and level 4 w/ scaffolding

The following table shows the result from paired t-tests for coverage, with the mean and STD.

	P(T<=t) two-tail	**Mean coverage**		**Std deviation**	
		1st variable	2nd variable	1st variable	2nd variable
1) Lv2 w/o – Lv2 w/	0.72	0.89	0.91	0.17	0.14
2) Lv2 w/o – Lv3 w/o	0.46	0.89	0.93	0.17	0.15
3) Lv2 w/ – Lv4 w/	0.13	0.91	0.8	0.14	0.22

Table 6-11 Paired t-tests for coverage. There is no significant difference in the test conditions.

And, paired t-tests for space reduction, with the mean and STD:

	P(T<=t) two-tail	**Mean reduction**		**Std deviation**	
		1st variable	2nd variable	1st variable	2nd variable
1) Lv2 w/o – Lv2 w/	0.70	0.30	0.28	0.13	0.20
2) Lv2 w/o – Lv3 w/o	0.07	0.30	0.23	0.13	0.15
3) Lv2 w/ – Lv4 w/	0.01	0.28	0.18	0.20	0.10

Table 6-12 Paired t-tests for reduction. Larger discussion spaces are reduced more.

While there was no statistically significant change in coverage between the task conditions, it is encouraging that the higher levels were reduced progressively more. The reasons were not investigated in the evaluation, but two factors might have contributed to the increased reduction in level 3 and especially level 4. First, participants stated that later levels had more opinions than in level 2s and they often summarized opinions with "meta-information" that described the views, for example, "some people support view A while others view B", rather than listing arguments from every post. In other words, opinions could have been more "compressible".

Second, the large amount of information in level 3 and 4 might have affected participants' decision as to how much detail to include in the summaries. Although the amount of the summaries that participants created generally increased in higher levels because there was more information to cover, it was less than proportional to the increased amount of the discussion information. Participants might have been compelled to create more concise summaries in order to cover larger spaces within the given time or to reduce the amount of the overall work they had to do, as stated by one participant during the interview. However, further examination is needed in order to investigate the reasons for the reduction changes.

Summarization Scaffolding in Arkose2

Arkose2 allows a user of the system to group a few posts together and leave a summarization guide. For example, a user might read discussion posts about copyright and feel that there is some important information to be saved. He can visually group the posts and leave a scaffold node to ask others to summarize the part. This can be done collaboratively. The user does not need to read everything, but just the part in which he is interested. Instead of leaving an instruction for others, he may also group the posts and summarize for his own use. By allowing users to create scaffolding on top of the discussion, it effectively breaks the discussion space down in to smaller subsections.

Two design claims about Arkose2's scaffolding support are made in the evaluation. The first is that scaffolding assisted participants in creating summaries. The second claim is support for scaffolding would allow specific kinds of collaborative distillation, namely the separation of the roles.   I will discuss each in turn.

Claim 1: Scaffolding assisted participants in creating summaries.

There were two tasks that had summarization scaffolding; level 2 and 4. (See Figure 6-11 and Figure 6-12 for the visual representation of discussion spaces with scaffolding.) All of the participants stated that scaffolding was helpful one way or another in the summarization tasks. From their statements, I found two ways scaffolding helped.

1. Scaffolding helped a participant understand important subtopics in the discussion space by pointing them out in the scaffold nodes, e.g., "Summarize what the main arguments are in the discussion of copyrights".
2. Scaffolding broke down the discussion space into smaller subsections and allowed a participant to quickly look at related posts.

For one participant, Ellen, scaffolding made difference in her summarization process. The following is the steps she took in level 2 *without* scaffolding (see Figure 6-20 for the completed state).

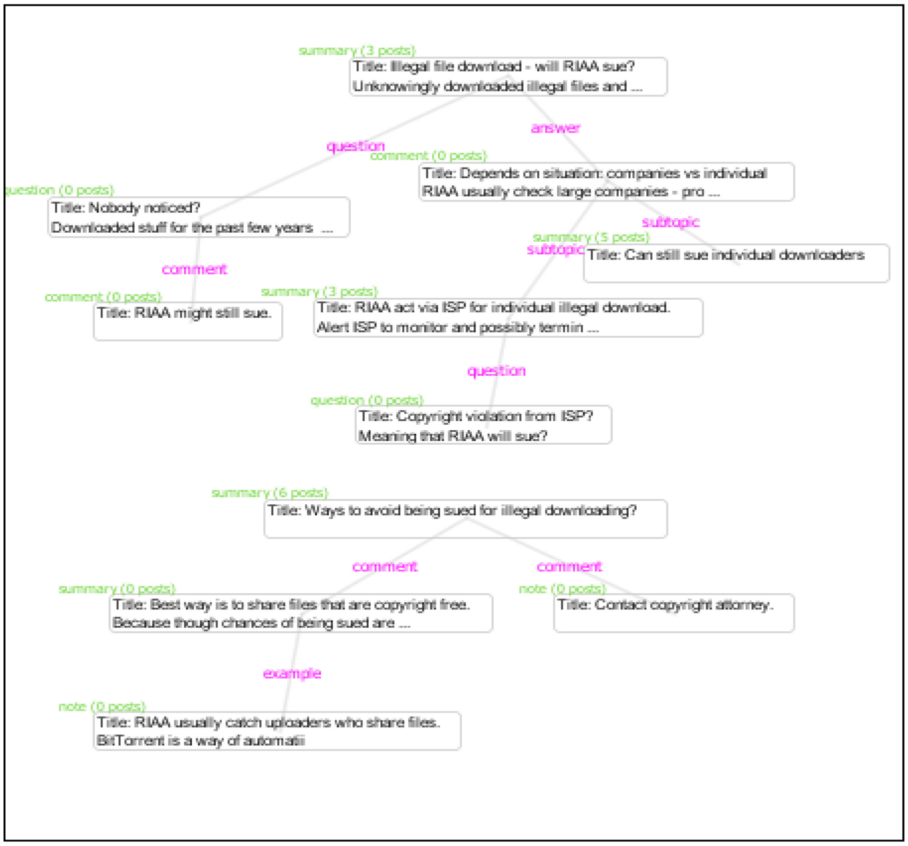

Figure 6-20 Ellen's summary structure for level 2 without scaffolding

1. She primarily used approach 1 of reading the entire discussion space and creating summaries as new topics are found. She spent the first two minutes scanning through the discussion posts in order to understand the overall topics.
2. She then created a summary node, added summary content, and attached source posts by drag-and-dropping.
3. As more nodes are created, she connected them with relationship types.
4. She repeated this until she covered the entire discussion space.

In level 2 with scaffolding, this was changed. The following is a detailed account of what she did (from the notes and video recording).

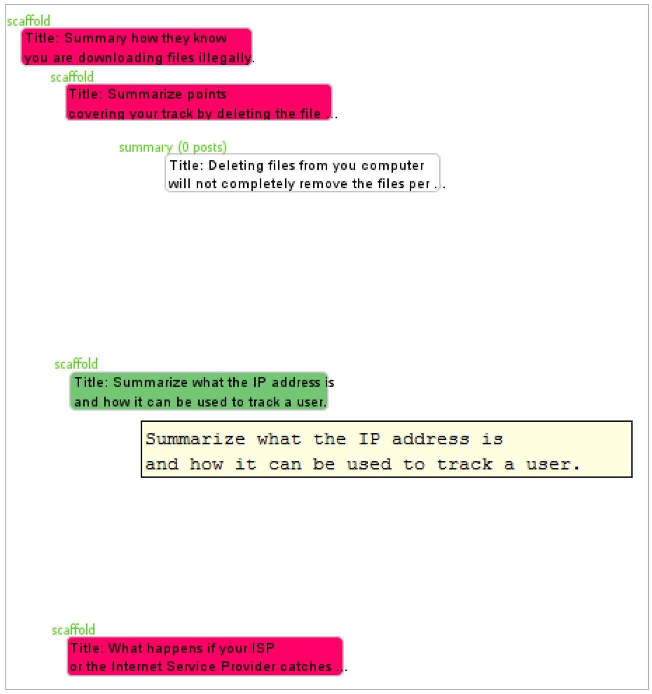

Figure 6-21 Level 2 Scaffolding in Summary Editor. The content is viewable by hovering the mouse over the scaffolding node, as shown here.

Instead of reading through the entire discussion posts, she first read the four scaffolding nodes (Figure 6-21) one by one. Then, as she was reading the posts and finding posts that fit under each scaffolding node, she created a summary node underneath the scaffolding node. As she was finding more posts, she put them into summaries under each scaffolding node. This was similar to “binning” the posts under a relevant scaffolding node.

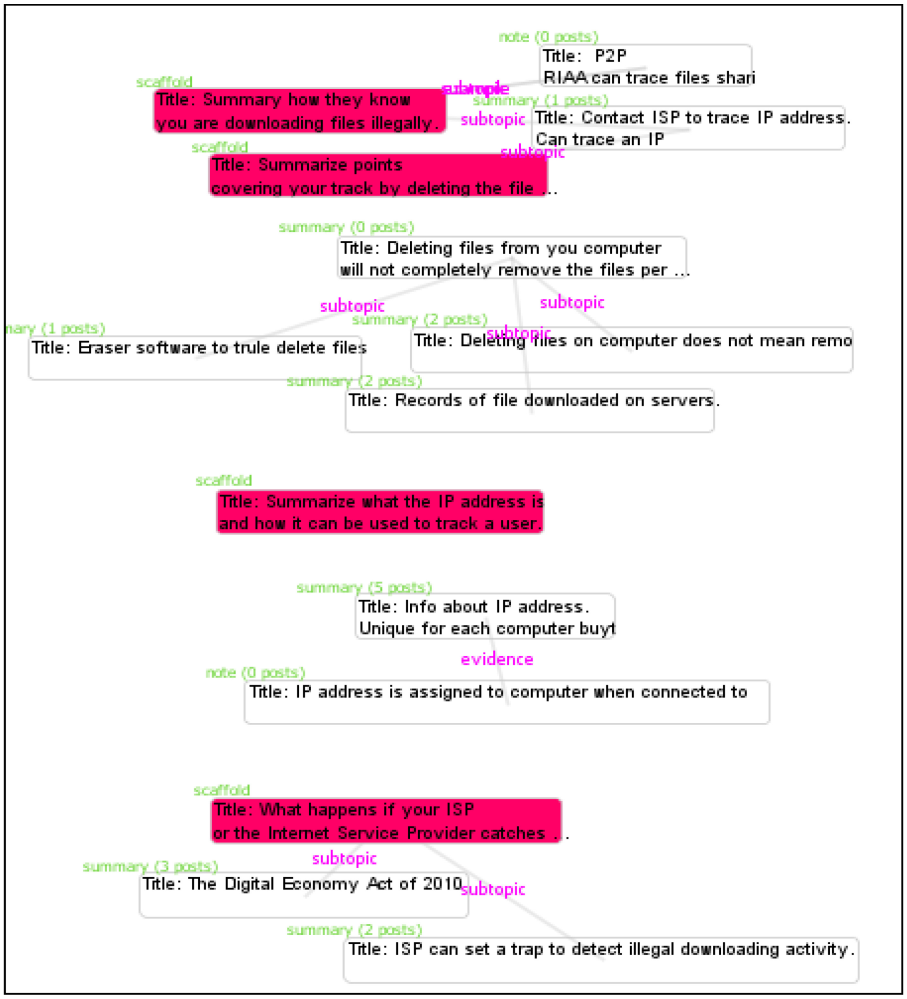

Figure 6-22 Ellen's summary structure for level 2 with scaffolding.

In the interview afterwards Ellen stated that the scaffolding helped her figure out the subtopics.

```
The scaffolding solved the topics for me, [and it] broke
the thread to different topics so I just need to find
relevant posts.  [Ellen]
```

And similarly, Lisa and Gabriel:

```
This time, reading the thread first was a little silly
since they already told you some major scaffolds, what the
biggest issues are. [Lisa]

Instead of just reading through and finding the topics on
my own, there were a lot of things already grouped together.
[Gabriel]
```

Erin's statement supports her summarization process observed in the evaluation. Instead of reading the entire discussion space to find out what subtopics existed as she did in her level 2 discussion without scaffolding, she could just read the scaffolding nodes to find out important topics.

Ellen's process of filling in each scaffolding node was also used by other participants, Erin and Matt.

> The scaffolding sort of helped. I was filling in details, so the basic structure was already there and I just had to fill in. [Erin]
>
> […]I kind of remembered the gist of each scaffold topic. And so when I saw a post I could put it in the scaffolding summary. [Matt]

Matt's comment may indicate that a prior knowledge of all the topics in the discussion might make it easier for a user to decide what to do with individual posts. One post on a new topic might not justify the creation of a new summary node until more related posts are found and a substantial discussion is formed. A user, without knowing whether the new post is part of an important topic, might have to defer the decision and revisit the post again when other related posts are found. Thus "binning" posts to existing categories might be easier than having to create a new topic. (Of course, even in this case if the user found a post that did not fit any category (scaffolding node), he would have the same problem.) This has implications for collaborative distillation in that users who want to organize topics might build scaffolding, while others may fill in the structure.

All participants stated or agreed that scaffolding also helped break down the discussion space into smaller, almost independent subsections. The scaffolding groupings are defined in the Visual Navigator (see Figure 6-12) and posts are grouped under scaffolding. While the scaffold groupings were not completely independent from each other as some posts discussed similar topics, participants stated they could individually summarize each group in turn.

> Yeah. That's what I liked about the scaffolding was that it made it instead of one big task, it divided it into ten different tasks. And then, so you could compartmentalize each topic. [Matt]
>
> […] you just start with one of the scaffolds and kind of find all of the posts that are important to that. [Lisa]
>
> I went right to the groupings and then focused on each grouping rather than looking at the big... everything and trying to find the groupings for myself. [Morgan]
>
> Scaffolding enables me to, together with [the visual navigator], go through related posts a lot faster. [Ellen]

The statements were supported by what they did in the distillation tasks. Nine participants mostly followed the groupings of the scaffolding when reading and summarizing. For example, Chris summarized each scaffolding group one by one independently from others. Here are the steps he took in level 4 discussion space with scaffolding (see Figure 6-23 for the final state) :

- He first read the first few scaffolding nodes in the Summary Editor Space.
- Then, he clicked the scaffolding node, which highlighted the corresponding scaffolding node in the Visual Navigator. A group of posts in the Visual Navigator was already included with the scaffolding node indicating they all fell under a similar topic.
- He read each post in the group one by one, and then created summary nodes in the Summary Editor Space under the scaffolding node.
- After he summarized all the posts in the group, he then moved on to another scaffolding group to repeat similar steps. While he explored other areas of the

discussion space from time to time, he mostly stayed within each grouping when summarizing.

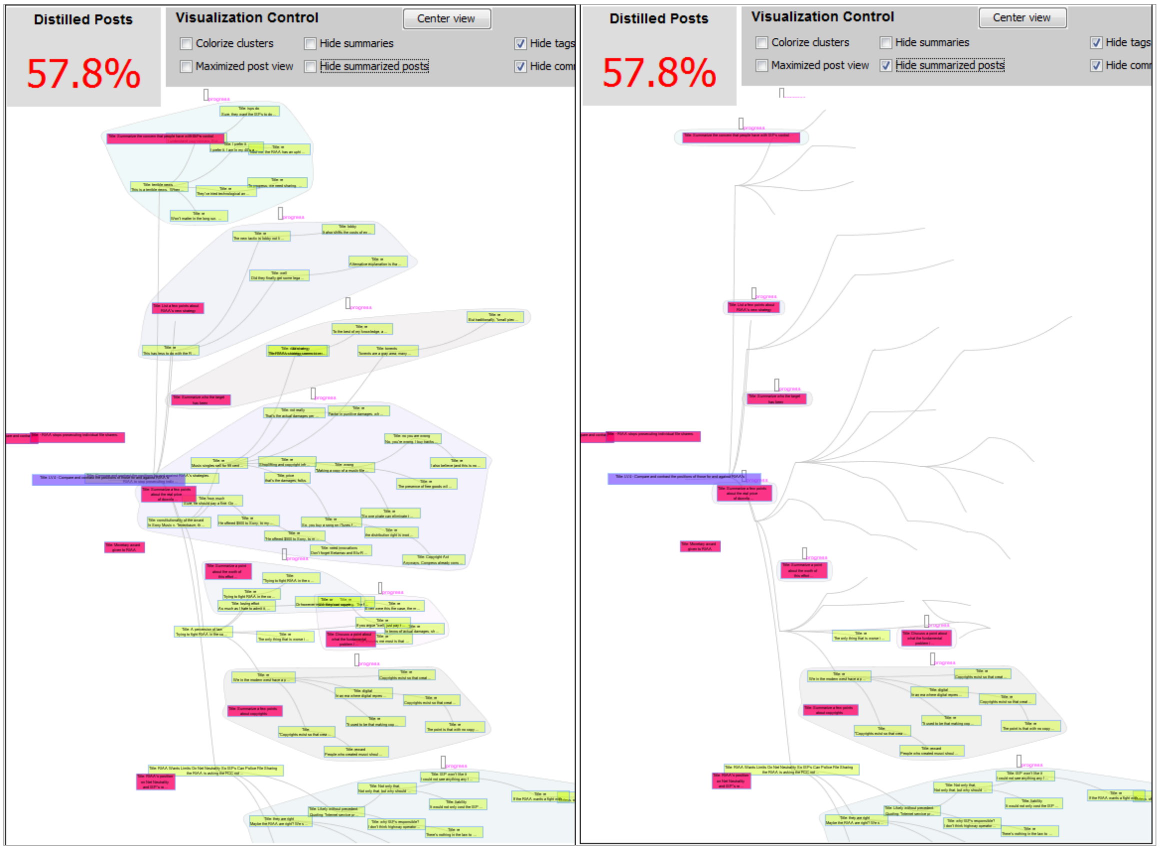

Figure 6-23 Chris' level 4 discussion space before (by turning off the hiding feature) and after summarization. Summarized posts are temporarily invisible. Chris mostly summarized each scaffolding group one by one.

While all participants agreed grouping of the posts in the scaffolding was a useful feature, it had limitations. Two participants thought scaffolding was "rigid" and "constraining."

> I felt like the structure was more rigid with the scaffolding. […] In a sense there was already so much scaffolding, I was more hesitant to add my own topics because it seemed like there was like a clearer idea and a clearer purpose from the beginning with the scaffolding. [Matt]

> [Scaffolding] was both helpful and constraining. Because some of them weren't exactly the way I would've scaffolded it. [Chris]

And not all participants agreed in the way the posts were grouped together. Six participants found some posts in the grouping could have been included in another scaffolding group and did not really navigate within each scaffolding group when reading related posts, but instead mostly navigated linearly from top to bottom in the Thread View or jumped from place to place in the Visual Navigator. Participants who thought some posts belonged to another scaffolding group stated that they summarized those posts under the scaffolding node they thought would fit better. Further, two participants stated they wanted to organize the discussion space in their own way.

> It’s also very hard when they already have it broken down into scaffolds, because you feel like you should be doing that. Although it’s a good place to start, maybe what I should’ve done is put all the scaffolds up here and just gone with how I would’ve organized it. [Lisa]
>
> People organize it differently. […] I would’ve made the title about money and copyright and how those are all corresponded to each of the scaffold titles. [Lisa]
>
> I think that’s true that I might have [updated the scaffolding]. I might have made the scaffolds differently had I made the scaffolds. But I also think it’s more difficult to organize everything when I haven’t read everything. [Justin]

One participant, Travis, did just that. While he consulted the original scaffolding structure, in the end he created his own scaffolding structure and summaries that he thought better reflected the discussion space (see Figure 6-24). Again, this has implication for collaborative distillation that some users may be able to build an overall structure that others may fill in the details, while some may have trouble with the provided scaffolding created by others. Arkose2 allows users to freely change the scaffolding instructions and groupings of the posts. But it is likely there might be some conflict over what would be the correct scaffolding, similar to the "edit war" in Wikipedia.

In summary, claim 1 that scaffolding in Arkose2 helped users was supported by the participants' statements and their summarization process. Scaffolding helped participants understand the topics in the discussion space, and it broke the space into smaller sub-spaces. However, some users did have problems conforming to the provided scaffolding as discussed; either the scaffolding should serve as guidelines, or not all users may be able to accommodate this though process.

Figure 6-24 Travis created his own scaffolding structure.

Claim 2: Support for scaffolding in a system such as Arkose2 would allow specific kinds of collaborative distillation, namely the separation of the roles.

Arkose2's support for scaffolding may allow collaboration among users to take on different roles. Users with more experience with the topic or read more discussion space may create a scaffolding structure, while other users may fill in the structure. In the evaluation, three participants explicitly used approach 2 of creating a skeletal structure of the topics (see Figure 6-15 and Figure 6-24 for examples).

> I assume there's different topics within the discussion thread so I tried to make a bunch of different categories […] sort of like what we did in scaffolding, you know. (When asked whether the participant was creating his own scaffolds) Yeah, pretty much. [Daniel]
>
> (When asked whether the participant was building scaffolding) Yes, exactly. I feel like, maybe not for all of [the scaffolds], but I think I was a little bit more specific. I think these scaffoldings tend to be like my others ones [that I created]. [Travis]
>
> (When asked whether the participant was building scaffolding) I think so. I think that's how I tend to outline. That's how I like to do research. [Brian]

Other users of Arkose2 may find filling in the structure easier, as indicated by one participant.

> The scaffolding sort of helped. I was filling in details, so the basic structure was already there and I just had to fill in. [Erin]

Ten participants who used approach 1 could point out important topics in their summaries as well.

Scaffolding may also be created incrementally and updated as needed. Arkose2 allows collaborative updating of the scaffolding structure and content. Participants could identify scaffolding nodes that they felt were not suitable. Two participants found that a scaffold grouping actually had more subtopics than what the scaffold was instructing to summarize. Updating a scaffolding node simply meant opening it up in the Advanced Summary Editor and changing the text in it. Three participants actually changed the instructions of some of the scaffolding nodes in the Summary Editor Space. One of them stated,

> People organize it differently. […] I would've made the title about money and copyright and how those are all corresponded to each of the scaffold titles. [Lisa]

These together imply that scaffolding can be a viable solution for collaborative distillation by allowing separation of roles among the users and by shaping it incrementally.

However, two problems may exist in supporting the separation of the roles. The first is finding enough users for each task. Scaffolding building is likely to require more experience and background knowledge to do it well, because the user would need to know what was a coherent and important topic among the posts, and how to divide sub-topics effectively. It might also require more effort than just filling in an existing structure, as stated by some participants. Finding the right users and motivating them to do the work might mean the need for expertise location (e.g., Zhang et al. 2007b) and matching tasks with users within Arkose2. I may need an effective incentive system for the community using Arkose2 as well. Arkose2 supports a task request feature that allows users to leave a specific request in the summary nodes.

The second problem is the "ownership" of the scaffolding. At the moment, Arkose2 does not give users any ownership of the scaffolding structure they may have created. Anyone can modify anyone else's scaffolding, similar to the editing process in a wiki. One participant, Chris, did not want others to be able to modify the scaffolding once created, since he wanted more control over the way it was built. It is likely there might be conflicts over the structure of scaffolding as discussed earlier. Users who fill in the structure may find it inadequate and may want to change it. One solution may be allowing multiple scaffolding layers, even with private scaffolding that can be locked by the creator.

Claim 3: Scaffolding is likely to be more useful when the size and complexity of the discussion space are larger.

All participants agreed that scaffolding was more helpful when there were more posts and subtopics in the discussion space. Three participants stated they relied more on the groupings already in place as the complexity of the discussion space increased in level 4.

> I couldn't have done it without [scaffolding]. I was like, when you first showed [scaffolding] to me, I was like, why do I have to this, right? And now doing this one, the more posts you have, the more helpful [scaffolding] is. It's like the visual representation of what… and I am also a very visual learner. So for me to understand certain things, I have to have an image to have everything connect. [Erin]
>
> I think [scaffolding] would be more useful in level 4. The more information that there is, it seems the better it is to have that. [Brian]
>
> [Without scaffolding] I think it would have been more difficult. Absolutely. [Matt]

All participants agreed scaffolding was more helpful in level 4 than in level 2, and the task completion time may reflect that. Table 6-4 shows the task completion times for the participants in their tasks. For 8 participants, it took one to six minutes *longer* to complete the level 2 discussion with scaffolding. While the reasons were not investigated during the evaluation, there may be a couple explanations as to why this may be. First, when the discussion space is small, just introducing a few more nodes to read might have a visible effect on the completion time. Level 2 discussion space has 20 posts and level 2 with scaffolding has 5 extra scaffolding nodes. This is an increase of 25% more nodes to read and obviously there is an overhead cost of introducing scaffolding into a discussion space. Also, a user would need to formulate a strategy for the scaffolding instructions, which may take time. Second explanation might be that participants needed to time to get familiarized with the scaffolding since it was the first time they saw it.

	Level 2 w/o scaffolding (Max. 20 min.)	Level 2 with scaffolding (Max. 20 min)
Phil	30	24
Emily	18	30
Brian	14	14
Ellen	24	20
Rob	24	24
Daniel	30	24
Matt	12	18
Lisa	11	15
Josh	17	24
Morgan	18	24
Erin	10	11
Gabriel	15	24
Justin	24	20
Chris	13	15
Travis	24	19
Avg.	**18.9**	**20.4**

Table 6-13 Adjusted task completion times for level 2 tasks. Highlighted ones are the participants who took longer with scaffolding.

To examine the changes in the task completion times between the level 2s, I adjust the completion times for unfinished tasks by estimating how much time it would have taken to finish the unfinished task based on the time it took for the completed portion. For example, if a participant took 20 minutes to finish 2 out of 3 sub-topics, I estimate that it took 10 minutes per sub-topic, and thus 30 minutes for the entire task.

This is a reasonable estimation since each sub-topic has a similar size. The following table shows the adjusted times.

On average, level 2 with scaffolding took 1.5 minutes longer than level 2 without scaffolding. For the 8 participants whose completion times increased with scaffolding, the average increase was 5.9 minutes (with $p = 0.002$). And for the 5 participants whose completion times decreased with scaffolding, the average decrease was 5 minutes (with $p = 0.0004$).

There is some evidence that scaffolding in level 4 cut down the completion time, since the participants who did not have to spend initial several minutes to scan through the entire discussion space saved time by immediately jumping into scaffolding groups. However, there was no task condition for level 4 without scaffolding, thus a direct comparison is not possible. Further evaluation for the scaffolding's effects on the completion time may be needed to reveal the extent to which scaffolding affects the completion time.

Design Features of Arkose2

Arkose2 has various visual aids and supportive features. The Summary Editor Space, Thread Viewer, and Visual Navigator are all interconnected and clicking or changing a node in one place locates and highlights related objects in other views. Participants found various aids helpful in reading and organizing information, and navigating in the discussion space. In this section, I discuss what features were found useful and how they were used to support participants' summarization tasks. Any issues found are discussed as well.

Claim: Features of Arkose2 assisted participants in creating summaries incrementally.

Visual Navigator

The Visual Navigator had several features that participants found helpful in creating summaries. First, it helped participants understand the discussion space more easily as stated by four participants who found the overall tree representation of the discussion thread very useful in finding potential "topic posts" and related posts. The tree representation makes is obvious which posts have many children posts, and where approximate groups of discussion exist.

This was further supported by allowing users to navigate the discussion space in any of the depth-first, breadth-first, and random order. When there are many posts (level 3 had about 50 posts and level 4, about 80.), reading them in a linear fashion (and in depth-first order) in the Thread View may quickly become a limiting factor. It becomes very hard to go back and forth among the posts that are far apart, and coming back to the previous post requires scrolling through multiple posts in order to locate the one. The Visual Navigator solves this problem by visually presenting the discussion thread in a tree structure with more spatial information for each post. A user can zoom-in and out, and drag the entire space or a few nodes. Because posts are presented spatially, it is easier to recall the location of a particular post. The following participants used the structural shapes as heuristics in locating important information.

> Just looking at [the tree structure], I would say this right here is a big intersection, this right here, and this right here (all pointing at posts that have many children). And then there's more than three links, so I think if there's more than three links coming off of something, that's probably either a question or [subtopic]. [Lisa]

> You can actually see which one is a child of which one. Because in the straight linear view you don't get that. [Chris]
>
> It's helpful to see the whole map and see the whole context. [Travis]
>
> I was able to use this visual structure to see how some of the posts up there were related to some of the posts down here, and to group them together… [Erin]

Two participants stated they used the Visual Navigator to keep track of the progress and their current location in the discussion space.

> I used [the visual navigator] more because it was hard to keep track of what I was doing. There was a lot more, and I needed some method to track where I was reading and what I was doing. [Daniel]
>
> I can visualize where I am in the… I guess in the topic cloud. [Matt]

The Visual Navigator can hide all the summarized posts in order to give a sense of progress. The "hiding feature" that temporarily makes summarized posts invisible (see Figure 6-23 for an example) was useful in telling what posts are remaining and need summarization, as stated by two participants.

> I found the visualization more helpful to tell me which ones I haven't summarized. [Emily]
>
> The hiding helps a lot. Yeah. Because I know how many I've got left. [Chris]

These visual aids follow the model-view-controller architecture, in that the same data or status is shown in different ways in different places of Arkose2.

However, the Visual Navigator may not be necessary when the discussion space is small. In the level 2 without scaffolding, only two participants used it extensively, while the others did not use or used very little. Level 2 had only about 20 posts, and the participants were able to just use the Thread View to read the posts in the discussion space. However, participants relied more on the Visual Navigator and other supportive features in a later task with more complexity, as I will discuss in a later section.

The Summary Editor Space

The Summary Editor Space helped the participants to incrementally create structured summaries. How do structured summaries compare to a less structured one such as one page summary of the discussion space done in, for example, a text editor? While a page of summary may still have some structure such as paragraphs and titles, explicitly specifying relationship among different summary nodes and separating information into various types of nodes (a summary, meta-information for summaries such as a note, comment, and question node) was found to be very useful.

> (When asked whether structuring summaries is important) Yes, definitely. People see visually, and when you look at a graph you definitely know what's important. [Lisa]

> [Separating information out] was really helpful. It was really helpful to have it outside of the summary. So that I know if there is a note, I'll [know the summary's not complete]. [Tom]

Separating out summaries and different types of meta-information also allows a user to attach source discussion posts to their respective summaries. With a page of summary, it would be more difficult to indicate from where a particular summary came. One participant stated this would be useful to look back at the original information.

> [Attaching posts to a summary is useful] that they were related to whatever the main topic that I identified. [More useful] when you're looking back later. [Erin]

A couple of participants stated that specifying relationship between two summary nodes was somewhat difficult when their relationship was not clear, but many participants found adding links and specifying relationship easy to do and helpful to have.

Another benefit of creating a structured summary was that the spatial information of the location of the nodes helped participants to remember where particular information was. Participants often rearranged the node structure in a way that made sense to them, by clustering related summary nodes more closely, and putting nodes in different levels to give them a hierarchy. While a page of summary with titles and paragraphs has some spatial information, but it is mostly linear (i.e., chapter 2 is below chapter 1, and so forth). The tree or graph structure of the summaries participants created was more useful in organizing and remembering information.

The Advanced Summary Editor (Figure 6-18) in the Summary Editor Space has a number of features that support these activities for creating structured summaries. As well, it allows the structured summaries to be built incrementally. Figure 6-18 shows how a summary node is displayed in the Advanced Summary Editor. Participants could update the summary as new information was found, added and removed the source posts, and created new summaries that reflected the discussion better. The Summary Editor Space also allowed participants to leave meta-information such as questions, comments, and notes for other users, which would allow a collaborative distillation of the discussion space and construction of summaries.

However, the size of the Summary Editor Space was problematic, especially for the participants who created many summary nodes and links. Although the space can be zoomed and panned to give a user more space for work, it quickly became crowded as participants created more summaries. One participant kept the number of the summary nodes small, but added more bullet points within each summary detail.

It looked a little messy if you have too many different [summary nodes] spread out. I actually broke it down to [bullet] point forms, so it's not necessary to have too many different summary nodes. [Phil]

In summary, the Summary Editor Space and Advanced Summary Editor helped participants create structured summaries incrementally, but the size of the available space was problematic for some participants.

The interconnected-ness of different views

The interconnected-ness of different views in Arkose2 also was used much in the summarization tasks. The Summary Editor Space allows a user to create structured summaries (see Figure 5-8, Figure 6-7, Figure 6-9 for examples), connect source discussion posts to them, and add various meta-information in the summary space such as a comment, question, and note. It maintains connections to the Thread View and Visual Navigator. For example, clicking a summary node highlights all the discussion posts in the Thread View that users included in the summary node. Clicking a scaffolding node in the Summary Editor Space finds the scaffold and its grouping in the visual navigator.

In the Thread View, each discussion post keeps track of what summaries are made for each post. When a post is added to a summary, its status icon indicates it has been summarized. Clicking a post in the Thread View highlight all the summary nodes to which the post belongs and locates and highlights its visual representation in the Visual Navigator space. One participant stated its helpfulness.

I would just float the mouse over to read the posts [in the visual navigator], so I didn't have to scroll up and down [in the thread view]. And then I make my [summary] node, and in order to add a comment I just used the click-feature which was really nice, just instantly highlights it and you could drag it over really quickly. [Travis]

Automatic clustering

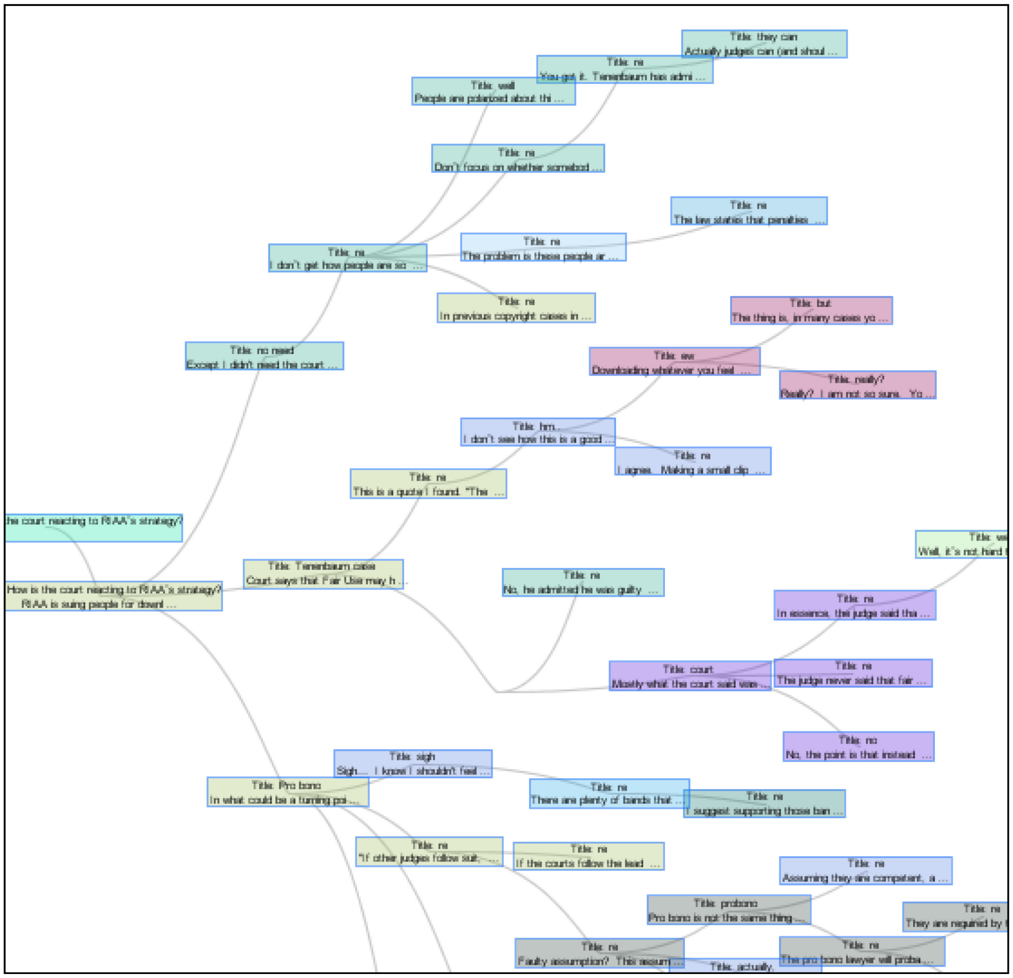

Figure 6-25 Automatic clustering in action. Posts in the same color are statistically found to be similar in content.

On the other hand, only four participants tried the automatic clustering feature, even though it was explained in the initial task instructions. By default it was turned off, and a user had to select a check box in the control box above the discussion space to activate it. Chris and Brian, who used the Visual Navigator extensively throughout all four tasks, used the clustering feature and found the automatic clustering helpful.

> First thing I did was colorize the clusters just to get the graphic representation of what computer thought was related, which was fairly accurate. Probably 70-75%? Probably a little less than that because there were some of them even I had hard time categorizing. But still, for a

> machine, not bad. And then, I picked related colors and after I read through a few and got a just what I thought was the main point of what the comment was, I created a [summary] node. [Chris]
>
> And actually, I really liked the... I like the color clusters. It makes it little easier for me to see like which goes together. [Brian]

However, two participants who tried the clustering found it confusing and thought it did not work as expected. Although it is likely posts with similar content would be under the same discussion thread, similar posts may be found in other parts of the discussion space. But it may be difficult to follow posts in different discussion threads since it breaks the natural discourse of users' conversation. This was too confusing for one participant, Ellen.

> I tried to use this [clustering feature] to identify the clusters that were similar, but I couldn't really follow the conversation. So I just went back and then read from top to bottom [in the thread view]. [Ellen]

And a usability issue kept one participant from using the feature. In the Visual Navigator, a node is highlighted when a user moves the mouse cursor over it or clicks it. This made it easy to locate the current post, but at the same time it made it difficult to track which color cluster to which the post belonged.

> In order to figure out what the initial [cluster] color was, I'd have to click off [the post]. That process seemed a little bit too laborious for me. [Morgan]

Using a different method to present the clusters might work better. For example, varying the shape of the clustered nodes or temporarily locking the highlighting feature might work. Similarly, using colors for clusters might not work well with users with color blindness. (The same goes for any feature in Arkose2 that uses a color to assign a meaning.) One participant stated that he was color blind, but was able to distinguish all the color clusters in the Visual Navigator or any other parts in Arkose2 because it used a high contrast color scheme. But this may present a potential problem in the field because many of Arkose2's features are augmented with the use of a color. Further investigation on a proper color scheme or other alternatives may be needed before Arkose2 can be used widely.

As with any automatic algorithm, clustering in Arkose2 is not perfect. Even human editors might not be able to group the posts *perfectly*, as mentioned by Chris in the above statement. This can lead to confusion among users especially if their expectation of the clustering is high. The problem is compounded when the clustering is done on conversations rather than documents. Although some posts may be very similar in content, they may belong to separate threads and the disconnection in the conversation may lead to further confusion.

In summary, the automatic clustering was not helpful or not needed for most participants, while two participants found it useful.

More Useful with Complexity

The Visual Navigator was used more by the participants in later levels where the discussion space was much bigger and more complex. Through participants' statements and observation, I found thirteen out of fifteen participants did not use the Visual Navigator or used it only a little in level 2s (though they generally used it more in level 2 with scaffolding because the scaffolding groups were defined in the Visual Navigator) but eleven of fifteen used it considerably more in level 4. All of the participants stated that the Visual Navigator was more useful when the discussion space had more posts and had a higher complexity, (Even the two participants who did not find it useful enough to use it mentioned the feature would be more useful when the discussion space was bigger.) Level 4's Visual Navigator with scaffolding was helpful. The following shows the changes some participants felt.

> [Justin]
>
> [Level 2]: I didn't use [the visual navigator] at all except since you told me that [summarized posts] disappear and I saw that at the end.
>
> [Level 3]: Well, [the discussion space] is much larger. It has a lot more information. So I found that [the visual navigator] really helpful over here because […] it's much easier to view it this way, the branching [of the discussion tree], rather than having to view the indentation [of the thread view]. So I found that it's easy just for me to read through everything by clicking on the branches and following.
>
> [Ellen]
>
> [Level 2]: I didn't really use [the visual navigator], because I don't know what I should do with it. (When asked whether reading the posts in the thread view was good enough) Yes.
>
> [Level 4]: For this one, I actually depended heavily on [the visual navigator] to get to the points that are related and to create structured. Because it's long and I do not really want to go through each of [the posts].
>
> [Emily]
>
> [Level 2]: It's helpful to know how [the discussion posts are] connected but even then… Actually, I found the visualization more helpful to tell me which ones I haven't summarized. So, it was less helpful to help me with my summarization but more as a reminder for me.
>
> [Level 4]: [Just having the thread view] would be way too hard, yeah. So [the visual navigator] is actually super helpful when you get into really, really large data sets. But when [the number of the posts] is like ten, I'd rather just look at [the thread view].

These findings have implications for the design of future Arkose. As found in the participants' statements, users might start looking for supportive features and relying more on them as the complexity grows. The same indicates that when the discussion space was small, users might find having all the supportive features on the screen unnecessary or even bothersome in performing their tasks. A better approach would be presenting only the essential features needed to discuss and distill at the early stage of a topic, and gradually show more features as the discussion space grows. (An idea similar to the "training wheel" concept (Carroll and Carrithers 1984).) This feature would especially be helpful for beginners of Arkose2 since users may have harder time understanding all the features and information of Arkose2 at once.

The training wheel concept was initially considered, but not implemented for the evaluation. The reason was that I first needed to find out what essential features users would use in their summarization tasks, and could not arbitrarily decide which features to show and which ones to hide. Second, the training wheel feature might be better suited in a field test where participants would use Arkose2 for an extended period of time and have a chance to get familiarized with it. The evaluation was done in a compressed time period with specific timed tasks, and switching features before participants had enough experience with Arkose2 might be too confusing.

In summary, I found enough evidence that supported the claim that the supportive features were more helpful when the complexity of the problem space increased. The number of the participants who used the Visual Navigator increased from two in level 2 to 11 in level 4. All participants stated that the features were or would be more useful in the later levels where the size and complexity of the discussion space dramatically increased.

Experience with Arkose2

In this section, I summarize and discuss what features of Arkose2 worked and what did not work throughout the evaluation tasks and in the interviews. These include issues the participants found as well as my own observation and reflection.

Usability Issues in using Arkose2

Most of the issues found by the participants were usability issues. In the first three sessions, there was a bug that did not highlight connected scaffold nodes in the Summary Editor Space and Visual Navigator. While participants could create summaries without that feature, it was cumbersome to locate related scaffolding nodes manually. It was fixed for the subsequent 12 sessions, and did not seem to have affected the outcome.

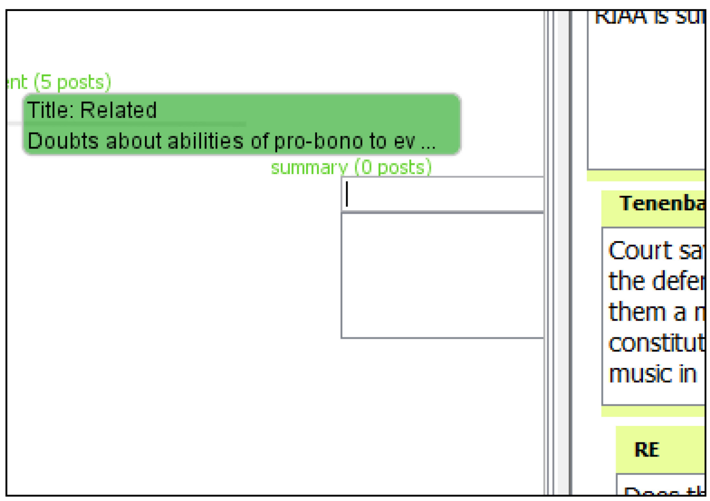

Figure 6-26 The problem of creating a summary node too close to the border. The editor is partially blocked by other components.

Another issue occurred while a user clicked on the Summary Editor Space to create a summary node, which opened up simple boxes for specifying the title and content of the summary. When a user was creating a summary node too close to the right border of the Summary Editor Space, part of the editor boxes were hidden and typed words are thus invisible (Figure 6-26). Participants were able to work around this by either creating a summary node closer to the center of the Summary Editor Space by mentally imagining the hidden part and typing, or by typing just enough for the available space and reopening the summary node in the Advanced Summary Editor for further editing. (See Figure 6-6) Since this was only a minor nuisance and fixing this problem would be considerable work that would change many parts of the program, it was left as is during the evaluations.

A user can drag a post from the center Thread View to the adjacent Summary Editor Space and drop it onto a summary node to attach it as a source post. It became a minor problem. Some participants tended to drag the visual representation of a post into the Summary Editor Space, not the post in the Thread View. Participants could, however, work around this by remembering to first locate the corresponding post in the Thread View and drag it from there. I believe this is a matter of user preferences rather than an incapacitating issue. One participant commented that he liked grouping the posts according to his preferences in the Visual Navigator. Other users moved around the posts in the Visual Navigator during the distillation tasks. One reasonable approach to overcome the issue may be to provide a user with an option to choose a drag-and-drop option.

One participant thought the zooming in and out using the mouse wheel was “counterintuitive” for him, since he expected going forward would zoom in and not out. Currently, the underlying Prefuse toolkit zooms in as the mouse wheel is moving away from the objects and towards the user (as if pulling things closer to the user), but a popular application such as the Google Maps (http://maps.google.com) uses the opposite direction (as if moving closer to the object). Depending on users' prior experience with other visual applications, it might be initially confusing, but all participants were able to work with the way the zooming function was implemented in Arkose2.

As well, Arkose2 froze twice during the evaluation. On both occasions, simply restarting it restored the state when it froze and participants could continue summarizing because the data was saved in real time. Later it was identified that the participants were trying to add more than one link (edge) between two summary nodes. The underlying program saves the nodes and edges in the database, and especially the edge is saved with the pair of the source node and target node as a primary key. When the participant created another edge on top of an existing one, it threw an exception because the program was trying to save a primary key that already existed. (A primary key must be kept unique in the database (MySQL in this case).) It was later fixed to show a warning message if a participant tried to add more than one edge, or add an edge to itself (i.e., the source node and target node are the same).

Features that did not work as expected or not used

There were a few issues and limitations that I learned as well. The obvious lesson is that the usability of a feature is critical for it to be actually used. One participant stated that she stopped using the automatic clustering because the mouse-hover highlighting hid the color of the cluster, which made it difficult to see to what group the current post belonged. For Arkose2 to be successful in a field test, several iterations in further development and design refinements may be needed.

Users needed better ways to track the summarization status. While the Visual Navigator has status information such as what proportion of the discussion space is summarized, this was not as useful in keeping track of the progress of summarization, since a user may have found particular posts that did not have any information and thus were not worthy to save. In this case the progress status would still show the discussion space as being summarized less than 100%, which may not be true. Many participants did not use the status information because it was not helping them in their summarization process. Participants often stated they were done (and their summaries did cover the entire space) even though the status did not reach 100%. A better representation of summarization progress may be useful in helping users keep track of the summarization progress of the discussion space. One approach may be to count the number of the sub-threads being summarized rather than individual posts. This is the way I estimated participants' summary coverage when they did not finish a tasks. (See Table 6-4 for examples.)

The automatic clustering feature had problems too. Only four participants tried it and the accuracy of the result and a usability issue kept two participants from using it. Two ways might help with the problem. One is to allow users to change the clustering so they can correct the result. Another is to cluster *groups* of posts rather than individual posts, since this may preserve the continuation of the discourse better. The group clustering may also be more useful when there are many more posts in the space, and finding related groups makes more sense than looking for one related post.

As well, users needed more help with crowded spaces, and the Summary Editor Space may need improvements. As users create more summary nodes and links among them, the space becomes crowded. For example, one participant, Emily, created a huge number of summary nodes and links, and it became almost impossible to rearrange the structure visually nicely. (See Figure 6-17 for her level 3 summaries.) Too many posts and links made it harder to organize them after they were created. Emily stated that her approaches changed in the later tasks in that she incrementally organized the structure as she was creating the nodes and links, and did not wait until everything was created as she did in earlier tasks. However, even in this way it was becoming problematic as the participant was adding more objects in the Summary Editor Space.

Also, a temporary work-space feature for users' intermediary work would be useful. A couple of participants used note nodes or comment nodes to collect posts and content that they would work on later. One participant stated that she was hesitant to attach source posts to a summary without planning much, because otherwise she might need to remove them from the summary node. Committing a post to a summary required some thought for her, and she wanted a temporary space where adding posts was not permanent. Note nodes and comment nodes had a sense of permanency and adding posts to them meant more commitment. It would be more flexible to provide a space that users feel free to erase.

A considerable limitation of this evaluation is that it only examined the distillation support capabilities of Arkose2. Given the limited time and resources, the evaluation was designed to focus on the issues, problems, and users' summarization approaches and experience in distilling a discussion space with Arkose2. As such, features of Arkose2 that were mostly for conducting a discussion were not examined. These included handling complex questions, question reformulation, dividing a topic into sub-topics, and creating wikis, though these features were available to the participants if they wanted to use them. These features will be evaluated in another field study examining the discussion support side of Arkose2.

As well, participants mostly used the features that were immediately needed for the distillation tasks. (i.e., the features discussed in earlier sections.) When asked why some features were not used, participants stated they did not need them for the tasks they were doing. For example, the automatic clustering feature was only tried by four participants as discussed earlier. And only one participant tagged a few of his summary nodes as he was creating summaries. (Note that the proportional tagging was initially part of the discussion features and was only available for tagging discussion posts. The participant who tagged a few of his summary nodes used non-weighted tags. No one else tagged a summary or a post. The proportional tagging system was available for summary nodes for the remaining 12 evaluations.)

One reason participants did not use some of the features might be that the evaluation consisted of timed tasks, and it did not give the participants much room to explore different features of Arkose2. However, some participants could gradually find new features or ways to do a task as they were using Arkose2. These included finding out multiple ways to create and edit a summary node, delete a summary node or links, remove attached posts, and using keyboard shortcuts for copying and pasting text. This gradual discovering of the features is encouraging, and I believe users might find the organization support features of Arkose2 such as tagging, categorization support, and wikis more useful in the long term use of Arkose2 where the amount of the information would be larger and the need for organizing the information space would be greater. A field test of Arkose2 would be more appropriate to examine the features not used in the evaluation.

Finally, Arkose2 was both more helpful and more limiting as the discussion space became larger and the complexity increased. As discussed earlier, participants found the supportive features in the Visual Navigator and the interconnected views more helpful in the later sections where the discussion space was more complex, and the features assisted the participants' summary creation. However, for two participants who created a summary node for each post, distilling a large space also meant creating more summary nodes and adding more details, and it became problematic. One is the Summary Editor Space became very crowded with many summary nodes and other meta-information, and organizing the graph structure became harder and recalling the location of a particular summary would be harder. The current size of the Summary Editor Space was problematic especially for the participants who used summarization approach 3 that created many one-to-one summary nodes. Although the Summary Editor Space is infinitely zoommable and draggable, it becomes harder to read the text as users zoom out to show more nodes at once. Allowing the Summary Editor Space to have its own popup window and change its size might help users better. An automatic layout feature might ameliorate a user's having to organize the structure by himself.

Another problem is that as the discussion space became larger, tracking which posts users already read would be more difficult. Participants sometimes looked through some posts again when looking for new posts they did not read. This may be redundant and time consuming unless users intentionally wanted to read them again. Allowing users to mark a post that they already read would be helpful in assisting them to navigate the information space.

What was liked

A few participants also commented on what they liked about Arkose2. One participant commented the system's usefulness and increasing demand for such a system.

> I'd say for this task, it'll be more important as more and more academic conversations become more numerous, you are going to need to do these summaries, and especially the [Visual Navigator], I think, is extraordinary. [Travis]

One participant stated he liked the visual and supportive features of Arkose2, such as various status reminders and automatic clustering.

> I thought that the progress bars and the percentage [indicator] were interesting and it made me think about if I was putting enough in. So I found it useful in that way. And actually, I really liked the color clusters. It made it a little easier for me to see which went together. [Brian]

Even with a minimum amount of instructions at the beginning of a session, participants were able to use the features well. (However, more detailed instructions were given as needed when a participant was using a certain feature. Rather than presenting the details of the features all at once, I presented just enough details to get the participant started and provided further details along the way.) They often discovered new capabilities along the way such as finding more than one way to do the same work, using Ctrl+C for copying text, learning how to use various GUIs and understanding what the status indicators meant and so forth. However, two participants seemed confused by what the Visual Navigator represented, but after further explanation they seemed to understand correctly. One participant expressed that Arkose2's features were easy to learn.

> I feel like it was pretty easy to learn how to use the system in the time that I'd been here. [Matt]

While there were several issues in using Arkose2 as discussed in earlier sections, no participant complained of any major issue that prevented them from distilling a discussion space.

What features worked

In most cases, I find the Visual Navigator was successful in supporting users' summarization tasks. It helped with understanding of the topics, navigating in the discussion space, and sub-divided the space into smaller parts. I believe the inclusion of scaffolding feature worked very well. Scaffolding was found to support users' summarization tasks and allowed collaborative distillation of the discussion space. In future implementation of Arkose, the Visual Navigator will also be an important component as well. The Thread View was not very useful by itself, but it might still be useful in reading a few posts quickly. The combination of the two worked very well as found in the evaluation. Participants used the Thread View for level 2 while many used the Visual Navigator for level 4, adapting to the changing situation.

Limitations and tradeoffs

Because the evaluation participants were recruited from the School of Information, they may not represent users of different organizations that may use Arkose2. Depending on the technical background and prior experience with online forums and content management systems, users may have different experiences using Arkose2. The focus of the evaluation is not on finding a quantitative conclusion as to the usefulness or effectiveness of the system. Rather, it focuses on the issues, difficulties, and user perceptions in distillation using Arkose2 through qualitative analysis.

While Arkose2 is both a flexible discussion system as well as distillation system, only the distillation part was evaluated in this study. Thus, it has limitations in terms of revealing the overall usefulness or effectiveness of Arkose2 as a system that supports both flexible discussion and distillation. As such, many features relevant to conducting a discussion were not examined. These need to be examined in a field test where Arkose2 would be used for an extended period of time by a number of users.

The distillation tasks took place in a compressed time period compared with the normal timeframe in which these tasks may be performed in an organization. To simplify the study and to focus on certain important features of Arkose2, the study design was individually based. While multiple users will be using the system simultaneously in a real world situation, the basic operations users will perform will be mostly the same as in the study design. Other concerns such as concurrency of operations, data bandwidth, and synchronization of data will need to be addressed in a real world deployment of a collaborative system, which is outside the scope of my thesis.

Another concern may be the learning effect as a participant moves on to the next level. Since the participant would have done a couple distillations before moving to level 3 or 4, the experience would be different for the later levels. However, I believe this should not be a concern as the participant needs to be familiar with the system before handling a more difficult task. In a long term real world deployment of the system, users will gradually become more familiar with the system and more efficient in performing tasks. Thus a slight learning effect from level to level is reflective of the real usage of the system.

Conclusion

In the evaluation of Arkose2, various aspects of summarization of discussion spaces have been examined.

First, all participants were able to use Arkose2 and create structured summaries. In doing so, the participants revealed their approaches in summarization, changes in the approaches based on the complexity of the discussion space, reliance on the supportive capabilities of Arkose2, and how the features supported participants' tasks, and the type of information saved and summarized.

As more and more group discussions are done through online media such as the discussion forum, the need to save and reuse important information from the discussion increases. As the need is becoming more visible, so is the importance of the supportive applications. Any discussion space with a sizeable amount of information would be difficult to summarize without help from other users and supportive features.

Creating summaries has a cost. Locating related posts, extracting important information out of them, and creating a structured summary requires much of users' effort and cognitive work to create a good summary that is beneficial to the community members. Thus, an application such as Arkose2 may not be suitable for summarizing casual conversations. Rather, organizations, companies, and groups of workers in a project may find that in some cases the benefits of saving discussion information exceed the costs of summarization.

All of the participants who were asked to summarize about two hundred posts in the four tasks were able to do a reasonably good job. At the least, an organization may hire a few domain experts to summarize the discussion space, which was the scenario for Arkose 1.0. Participants' work may potentially be reduced by collaboratively summarizing and using various supportive mechanisms as indicated in the evaluation.

In summary, Arkose2 proved to be useful for participants in creating structured summaries for a discussion space, and the evaluation revealed some very important insights as to the success of such collaborative systems in future.

Chapter 7

Conclusion

In this concluding chapter, I summarize the main findings in my three dissertation studies. Based on these findings, I discuss the design implications for a collaborative knowledge generation and distillation system. At last, I briefly discuss the contributions of my dissertation, followed by the future work.

Summary of Findings and Contribution

The goal of my dissertation has been to gain a better understanding of knowledge sharing and distilling in an online community, and to build a prototype collaborative system that supports both flexible knowledge generation as well as distillation. To reach this goal, I have conducted two user studies and built two systems.

The first system, Arkose 1.0, was a prototype collaborative distillation system for a discussion space. I presented the need and requirements for distillation to reuse informal information especially from an online community's brainstorming and discussions. Two design principles, which I defined as incremental diagenesis and incremental summarization, were presented to help the distillation process of the users of the system. Arkose 1.0 is built as a demonstration of these principles, with the necessary supportive tools and capabilities for distillation. These tools included a number of visualization and information retrieval capabilities, as well as an authoring tool and a navigator for the discussion space. Arkose 1.0's design supported a gradual increase in the order and reusability of the information space. However, Arkose 1.0 was limited to the distillation of an already completed or sufficiently progressed discussion, rather than supporting the discussion itself for a better discussion outcome. Incrementally distilling as the discussion is in progress would be more effective and useful. To build a new system that also supported a more flexible discussion, I next studied an online knowledge building community in order to examine users' knowledge sharing behavior and the factors affecting a successful discussion and distillation.

The study of Naver Knowledge-iN helped me understand users' knowledge sharing behaviors in a large online question-answering community. I collected 2.6 million question/answer pairs from fifteen categories between 2002 and 2007 and analyzed the characteristics of knowledge generation and user participation behavior. I also interviewed 26 users to gain insights into their motivations, roles, usage and expertise. I found that altruism, learning, and competency are frequent motivations for top answerers to participate, but that participation is often highly intermittent. A new *guru* measure was devised to calculate user performance in answering questions, and I found that higher levels of participation correlated with better performance.

More importantly, I found that the limitations in the user interfaces, among other factors, influenced user interactions and knowledge creation in the community. The limiting nature of the reply interfaces of Knowledge-iN prevented an answer from being developed further, and led to the accumulation of easy questions and answers. It was difficult to handle complex questions or to separate users' backchannel conversations from the actual answer providing activities, and the tendency for answering easy questions was encouraged by the point system that rewarded users who answered many questions quickly. Users of Knowledge-iN did not have a good way to organize information either. Although they could specify three categories of interest in order to receive questions in the categories, there was no supportive feature that allowed them to save and organize information they found useful. While there was a huge amount of useful information in Knowledge-iN, it was scattered all over in the system without any way to organize and distill them into more useful information. These findings gave me insights with the design of Arkose2. Some of the supportive features were designed to help conduct a more in-depth discussion on a complex topic, and have a separate conversation outside of the answering activities. Users' discussion and distillation activities would need to be supported by many visual and automatic aids. b.

Thus, Arkose2 is designed and implemented based on the lessons and insights gained in both building Arkose 1.0 and the Naver Knowledge-iN study. Arkose2 provides a host of supportive capabilities that help user activities of knowledge generation, distillation and organization. Its tools are designed to reduce the cognitive burden of users when dealing with a large quantity of information, give a more meaningful structure to the information objects in the information space, and personalize a user's informational experience. The supportive tools and various interactive mechanisms allow user activities some of which have been impossible or very difficult in current knowledge building systems.

Much of the previous summarization research has focused on developing algorithms and natural language processing methods that automatically generated summaries for text corpora without human intervention (e.g., Radev and Hovy 1999; Kintsch et al. 2000; Lin and Hovy 2003). While automatic methods can generate reasonable summaries with statistically important words and sentences, they cannot achieve human level summarization that is nuanced and interpreted. In Arkose 1.0 and Arkose2, I take a system augmented approach in which I allow human editors or users to distill a discussion space with supportive tools and mechanisms that assist their tasks.

	Arkose 1.0	Arkose2	WBT
Distill discussion space	Y	Y	Y
Collaborative distillation	Y	Y	Y
Ongoing distillation as discussion in progress	N	Y	Y
Separation of roles	N	Y	Y
Summarization scaffolding or schema	N	Y	Y
Pre-defined scaffolding groups	N	Y	N
Scaffolding or schema required	N	N	Y
Multiple scaffolding	N	N	N
Automatic clustering or classification of posts	N	Y	N
Multiple clustering or classification methods	N	N	N
Support distillation with visual aids	Y	Y	N
Create multiple summaries	N	N	N
Structured summaries	Y	Y	N
Summaries can be tagged	Y	Y	N
Summaries can be categorized	N	Y	N
Summaries can be saved for personal use	N	Y	N
Connect summaries and source data	Y	Y	N
Shows summarization progress	Y	Y	N
Displays discussion space in visual graph	Y	Y	N
Displays discussion space in threaded format	Y	Y	Y

Table 7-1 Feature comparisons for distillation support.

Table 7-1shows the features of distillation support covered by Arkose 1.0 and Arkose2, as well as the WBT system (Helic et al. 2003) that has similar a goal as Arkoses. As discussed in the literature review the WBT system allows students to summarize a discussion space in an educational setting, and it has some similar features such as the creation of skeletal structure ("conceptual schema") and the role of separation, but Arkose2 improves much over what WBT provides.

Arkose2's distillation support is also an improvement over that of Arkose 1.0. In Arkose 1.0, distillation was an afterthought when the discussion was completed. This made it difficult to integrate and reuse distilled information (e.g., summaries of the topics discussed, issues identified, and topics that need further discussion) with the discussion space and guide further discussion. Distillation itself was done by a small number of human editors that did not necessarily participate in the discussion. Arkose 1.0 did not support an explicit role separation in distillation as well. Arkose2 helps users distill a discussion space more effectively through the support for scaffolding and ongoing distillation by the community (see Table 7-1 for the features support by the two Arkoses and WBT system discussed in the literature review).

Finally, the evaluation of Arkose2 revealed a number of insights and lessons about users' distillation activities of discussion spaces. Participants in the evaluation could create reasonable summaries for the discussion spaces using Arkose2, which covered sufficient information about the subtopics and details. Participants used different approaches, some creating very concise summaries that provided an overview of the topics while some including more details from the original discussion spaces. Their approaches also changed as the size and complexity of the information space grew.

Scaffolding in Arkose2 was proved to assist many participants in creating summaries. It would allow an interesting separation of the roles among the users, with some users creating an overall structure that others could fill. However, scaffolding was also found to be rigid and constraining for some participants, and they did not agree with the way the scaffolding was constructed and wanted to change them. Allowing users to create multiple scaffolding might help ease the problem.

As with scaffolding, participants relied more on the supportive features of Arkose2 when the discussion space was larger and more complex. Many participants who did not use certain features in an earlier task started using them more in the later tasks. These features helped participants in several ways. First, the Visual Navigator allowed the participants to understand and navigate the discussion space better, to keep track of the progress, and provided guidelines for related posts. Also, the linkage among different views assisted participants' summary creation by readily presenting related information in multiple views. However, some users had trouble using the Visual Navigator's tree representation or did not find the automatic clustering very helpful. Further design improvements would be needed to make these features more useful.

Several factors also made it more difficult for the participants to summarize the discussion spaces. These included the lack of background information in the topic, the complexity of the discussion space, and usability issues of Arkose2, which are discussed in detail in earlier sections.

Arkose2 allowed an incremental creation of summaries, and participants were able to create summaries incrementally. Many participants gradually constructed and updated summaries as new information was discovered, new subtopics emerged, and a new summary structure was needed.

Finally, issues and problems of using Arkose2 were discovered. Some features did not work as expected or not used at all in the distillation tasks, and usability issues and bugs in the program made the distillation process more difficult for some users. For these and other problems discussed earlier, I presented potential solutions in the discussions. One considerable limitation of the evaluation was that in only examined the distillation support side of Arkose2, and leaving many features untested. The discussion support side of Arkose2 and other organization features will be studied in a field test to examine the usefulness and effectiveness of them.

In summary, I have conducted two user studies (Naver Knowledge-iN and the evaluation of Arkose2) and have built two systems (Arkose 1.0 and Arkose2). In the two user studies I have shown and discussed user behaviors and approaches in knowledge generation, sharing, and distillation. In the implementations of the two systems, I have designed and built new prototype user interfaces, supportive tools, and interaction mechanisms that may help users' discussion and distillation tasks. They, together, provide valuable lessons and insights for the architecture and features of the next generation collective intelligent system.

Design implications

Some findings in analysis of Naver Knowledge-iN and the development and evaluation of Arkose have design implications.

One design implication for a collaborative system may be providing different user interfaces for a small discussion space and a large one, or a switch between the two as the task evolves. Many of the participants in the evaluation of Arkose2 did not use the more elaborate Visual Navigator but just used the simpler Thread Viewer when reading and summarizing level 2 discussion spaces (the smallest of all levels). Thus, while working with those levels, users may only need to look at 2/3 of the entire screen of Arkose2. The Visual Navigator contains a lot of information, such as the tree representation of the discussion threads, various status information, and controls for manipulating the visualization. Presenting these kinds of unnecessary information might not be useful and might even be more burdensome to the users.

One idea for this problem might be to have two the separate interfaces - one for a simple question and answering, and the other for more complex discussion and distillation. The idea is similar to the training wheel user interfaces tested by Carroll and Carrithers (1984). Since new users of Arkose2 are more likely to ask and handle simple questions, showing them just enough user interfaces may be better. As the discussion evolves into a more complex stage, the Visual Navigator may show up on the screen.

Another design implication might be the need for a more explicit support for intermediary work. Arkose is designed to support incremental processes from its conception. The discussion space and summaries can be transformed gradually by users' collaborative work. During the evaluation of Arkose2, it was learned that participants often used an existing feature such as the note node in the Summary Editor Space to save data that they were working on. The note node was designed to act as a comment to a summary (as in, make a note to myself) than to be used as a place for storing lots of information (as in a piece of paper). Arkose2 makes it easy for a user to change the content of a summary node and to collect related posts in the summary, but some users may want a temporary storage that similarly acts. While the summary node and note node can be used for this purpose, having an explicit feature for intermediary work might be better.

And finally, the details of the summaries were partly influenced by the potential users and usage goals of the summaries that the participants had in mind. Summaries created for a personal understanding may be different from ones created as a report of opinions on which a policy decision may be based. The former may be selective, personalized, and incomplete, while the latter needs to be objective and thorough. Thus a feature that indicates the goal of the distillation might be helpful in guiding users' summarization tasks.

Future work

Much work remains for the successful deployment of Arkose2 in a field test. As a couple participants in the evaluation suggested, collaborative summarization might be more efficiently done if performed organically and *in situ,* while the discussion is in place.

> If you could start [summarizing] early before things [become old]… If someone was [summarizing as discussion was occurring], I think I would be more attractive of this style of [summaries]. A lot of times when I'm looking up information for things, people would be like "Oh, this is an old thing". People will post fresh questions as opposed to going and searching for it. But if they had a structure where it was a little easier to navigate, they might not. [Morgan]

For this, deploying Arkose2 in an organization for a field test for a considerable period of time is needed. This will also give me a chance to examine the features for information organization such as the proportional tagging, organic categorization, and linking various contextualized information and generalized information in wikis. I believe, however, Arkose2 needs another development cycle before it can be used in the real world. Arkose2 lacks some of the necessary components needed to be used by many people in a daily operation. Although it saves information in real time, synchronizing data among many users close to real time is very resource intensive, and requires more careful design. This is out of the scope of my thesis, but it needs attention for the future version of Arkose. As well, other dimensions of knowledge building community need to be considered for the use of Arkose2 to be successfully accepted. These include the policy, incentive system and privacy aspects of an online community. What are the right policies regarding information ownership, authority, and moderation for the organization? What incentive system, for example a point based system or monetary reward system, is right for it? And how open should the information activities be? These all need to be decided according to the specifics and needs of the organization in which Arkose2 is deployed.

Another future work may be to examine how distilled information may be used. Can users read the summaries and get a clear overview of the discussion space? Can a summary be created for one purpose and used in another? These are only some of the questions that need to be addresses as distillation systems become more popular and used widely.

Appendices

Appendix A: The Guru measure

Let $b_i = 1$ if the user provided the best answer to question i and 0 otherwise, and n_i be the number of answer to question i. I exclude those questions where $n_i = 1$ because there is no point of comparison about the quality of the answer. A user providing m answers would be expected to give x best answers, where $x = \sum_{i=1}^{m} \frac{1}{n_i}$

This probability takes into account the number of other users answering each question. The guru measure $\gamma = \frac{(\sum_{i=1}^{m} b_i) - x}{x}$, then indicates whether a user's performance is better or worse than chance.

Appendix B: Proportional tagging

*tf*idf weighing*

A document can be represented as a vector of weighted terms:

$$v_d = (w_{1,d}, w_{2,d},\ldots,w_{N,d})^T$$

There may be many ways to give a weight to each term. For example, the weight $w_{t,d}$ for term t can be the number of the term's occurrences in the document, in which case $w_{t,d} = tf_{t,d}$. It can also consider the global distribution of the term, inversely weigh the number of documents in which the term appears. So a term gains more weight if it frequently appear in the document, but infrequently in the entire document space. This weight is called the *tf*idf* value. Thus,

$$w_{t,d} = tf_{t,d} * \log \frac{|D|}{|\{d \in D \mid t \in d\}|}$$

Equation 0-1 Term weight.

, where $|D|$ is the total number of documents in the document space, and

$\{d \in D \mid t \in d\}|$ is the number of documents in which the term t appears.

The cosine similarity between document d_j and query q can be calculated as:

$$sim(d_j, q) = \frac{d_j \cdot q}{\| d_j \| \cdot \| q \|} = \frac{\sum_{i=1}^{N} w_{i,j} * w_{i,q}}{\sqrt{\sum_{i=1}^{N} w_{i,j}^2 * \sqrt{\sum_{i=1}^{N} w_{i,q}^2}}}$$

Equation 0-2 The cosine similarity value between two documents.

, where the similarity value ranges from 0 to 1.0.

Calculating similarity value using proportional tagging

Whereas the *tf*idf* values are statistically calculated, the proportional tagging weighs each term by a user supplied value through the user interfaces (Figure 5-21). The values may be used as weights for the terms, or they may be counted as the term frequency values and counted off with the document frequency value. For simplicity, Arkose2 currently takes the user supplied values as the weights for the terms. Thus measuring the cosine similarity between two documents with proportional tags is straightforward and follows Equation 0-2.

The proportional tagging system also allows a user to weigh a term negatively. (Figure 5-22) This complicates calculating the similarity value between two documents *p* and *q*. When comparing two documents with negatively weighed terms, a special approach is needed when there are t_q and t_p that are the same terms but with opposite signs. For example, if document *p* has the terms (*java* (40), *concurrency* (60), *coffee* (-30))and *q* has (*java* (50), *coffee* (30)), the term *coffee* should discount the similarity value. To handle this, Arkose2 first scans the terms in both documents when calculating the similarity. If the above condition is found, it creates two sets document vectors; one with all the tags with the same sign in both *p* and *q* and another with the opposite sign. Thus, *p*+ and *q*+ have terms with the same sign, while *p*- and *q*- have terms with the opposite signs. Then, the cosine similarity measures between *p*+ and *q*+ and *p*- and *q*-

$$sim(p,q) = sim(p+,q+) - sim(p-,q-)$$

Equation 0-3 The cosine similarity value with documents with terms with opposite signs.

are each calculated. The final cosine similarity value is calculated as the above.

All other cases, where the signs of the terms are equal (either negative or positive), can be simply treated as a normal vector comparison.

Similarly, one may specify how much weight the discount should have. Thus the new equation is:

$$sim(p,q) = w_{+}(sim(p+,q+)) - w_{-}(sim(p-,q-))$$

Equation 0-4 The cosine similarity value for weighted discount.

Note that these similarity calculations may produce a negative value, which is different from the "normal" cosine similarity calculation. This should not matter when comparing documents and ordering them by the similarity value.

While the above equations present a simple but useful way to compare documents with negatively weighed terms, further investigation may be needed to devise an equation that may better utilize the extra information a user can supply with the proportional tagging system.

Appendix C: Semi-structured interview questions

A semi-structured interview with a participant was performed during and after the completion of each distillation task. After each task a participant was asked to compare the experience with previous levels. At the end of the four tasks, questions covering the entire evaluation were also asked. The following is a set of sample questions participants were asked. As the interview evolves, other related questions are also asked.

During a distillation task:

(If the participant seems lost and does not do anything)

`Are you stuck? What are you trying to do right now?`

`Are you unsure how to distill the posts (content-wise)?`

`Are you trying to figure out how to distill it with Arkose (feature-wise)?`

`Are you unsure when to stop?`

`What do you think you should do next?`

`What do you think you can do next?`

(If the participant is excessively trying out different features of Arkose)

`Are you trying to find out what each function does?`

`Are you looking for a particular feature?`

(If the participant is having difficulty using a feature of Arkose)

`Are you having difficulty in using the system?`

`What do you think you can do with feature X?`

After all the distillation tasks:

`Overall, what did you think?`

What problems did you encounter? (Answers were classified such as problems specific to the features of Arkose, distillation itself and other categories, and were probed further if necessary.)

What was your approach in summarization?

Let's talk about one of your distillations, say (pick one).

Did you have a specific goal when you were doing it? Tell me how it unfolded for you.

How did you find the tasks with the scaffolding? What problems did you encounter?

What type of information was easy to distill and what type was harder? (Follow-on: Tell me about them.)

What type of information was saved in the summaries and what type was discarded? (Follow-on: Tell me about them.)

How did you structure your summary? What do the links among the summary nodes mean? How useful do you find creating structured summaries?

How did you find subtopics or important information?

How easy or difficult was distilling discussion spaces? (Follow-on: Tell me about them. What did you find hard?)

How much did you use the visual aids? In what ways did you use them? In what ways were they helpful?

How did the scaffolding affect the summarization experience? In what ways was the scaffolding helpful or not helpful?

How is summarizing a discussion space different from summarizing an article, paper, or book?

How was this task compared to the previous one? What differences did you find?

Appendix D: Discussion threads in the evaluation

The following made up data is used in the evaluation settings. The indentation indicates the depth of the post.

Level 2 without scaffolding

Topic: How does RIAA choose how to sue?

I illegally downloaded a few files without knowing and now I am worried they might find out and sue. How does RIAA choose who to sue?

> They usually check up large companies because that’s where the profit is being made. If you just share a few files once a while, chances are they won’t sue you.
>
>> They don’t bust you unless you’re running a mass distribution ring out of your garage. Just don’t download a lot in a short amount of time.
>>
>>> I downloaded for the last few years on and off. I didn’t get any notice from anybody.
>>>
>>>> But who knows – RIAA might still sue you.
>>>
>>> So they do not sue individual downloaders because they don’t have money?
>>>
>>>> They can still sue, but they will really have to prove that you did it.
>>>>
>>>> If they have proof that will stand up in court, yes, they could.
>>>>
>>>> For RIAA, tracking down one individual, setting up the case and hiring the lawyers costs more time and money than what they will profit.
>>>>
>>>>> But they DO sue individuals to set an example and let people know that illegal downloading is just that – illegal.
>
> Now they have the Internet provider alerted and usually your connection will be monitored and possibly terminated, instead of suing individuals.
>
>> Usually they give you plenty of warnings, and most of the time they go after those who don’t stop. If you just downloaded a while ago, most likely they’ll leave you alone.

I received a notice of copyright violation from my Internet provider. Does this mean the RIAA will sue me?

Depending on how much you downloaded and what evidence they got against you. If you feel threatened, you might consider contacting a copyright attorney.

The chances of getting caught are virtually zero. There are millions upon millions of downloaders worldwide. But if you are worried, don't download illegally.

Yes, I agree. Even though the chances are slim, you might be the unlucky one that gets caught.

There's only one way not to get caught. Only share files that are copyright free.

Also, they will catch uploaders who share files first before they catch downloaders. So if you just downloaded a few files, there's nothing much to worry about.

If you download a file using BitTorrent, you automatically share the file with others.

Yes, you act as an uploader as soon as you download a BitTorrent file, because that's how BitTorrent works. You download from multiple people, and others will in turn download what you have downloaded.

Level 2 with scaffolding

Topic: How do they know you are downloading files illegally?

I used to download a lot of music files using a sharing program. How does RIAA (Recording Industry Association of America) know if you are downloading illegal files?

When you share a file using a P2P program, they can trace where the file came from.

Yes, your computer has an address called the IP address, which is visible when you are sharing files.

What is an "IP address"?

IP stands for "Internet Protocol". Each computer has a unique one, although it can change.

Your computer's IP address is what makes it possible trace any computer, sort of like a real address. You can hide it using a proxy server.

You have a traceable address associated with your computer which can be used to track you down.

Your Internet service provider (ISP) assigns you the IP address when you make a connection. So they know who you are because they have your account with address, phone number, etc.

Yes, the IP address is assigned to your computer when you connect to the Internet, so the server computer knows who's connecting it. Also your ISP, because you are using their services.

They can talk to your ISP (Internet service provider, like Comcast) and trace your IP address. But if you delete everything from your computer, you should be ok.

Even if you delete the files, they are not actually removed, but just marked down to be overwritten.

When you delete files, they are still there – just not visible.

If you want to truly delete files for sure, you should use a special eraser software that overwrites data.

But even if you could delete the files for sure, they still have records of you downloading the files on their server. They log connections, downloading, etc.

Every time you connect to the Internet, your IP address is logged and time-stamped as is the file you downloaded.

An ISP like Comcast can set a trap for you to connect and download illegally. The IP address can tell them who you are. If you are caught, they can disconnect the service.

I didn't know an ISP could cut your connection off for downloading files.

Under the Digital Economy Act of 2010, an ISP can remove you Internet connection for illegal file sharing.

They gave you the connection. I am sure they can disconnect it too.

Once sufficient evidence has been found, they can get a warrant to search your computer. And if they find illegal files, they will sue you.

The RIAA will then sue for some outrageous amount like $50,000 and usually allow you to settle for $5,000 to $10,000.

Level 3 without scaffolding

Topic: How is the court reacting to RIAA's strategy?

RIAA is suing people for downloading files illegally. How is the court reacting to their strategy?

Court says that Fair Use may hold in some RIAA cases. In Sony BMG Music v. Tenenbaum, the defendant admitted liability at this trial and the Court has sided with the RIAA and granted them a monetary award. However, the Court recognized there are some cases that might constitute fair use, such as creation of MP3 files from audio CDs exclusively for saving the music in another place.

Does that mean Tenenbaum can now appeal since saving the music files was his primary purpose of ripping the CDs?

No, he admitted he was guilty at the trial. So he can't appeal.

Mostly what the court said was that the defense wasn't sufficient enough and the court would have been receptive to such fair use tactics but the defense failed to do it.

No, the point is that instead of making a case for specific types of fair use that might be applicable, the defendants' lawyers tried to argue rather broadly that all downloading was fair use, something that the judge couldn't agree to.

In essence, the judge said that they were open to certain arguments being made, but the defendant never tried to make them and instead opted for an incompetent defense.

Well, it's not hard to see why Tenenbaum wound up getting the judgment that he did…

The judge never said that fair use would be a justifiable defense for the case at hand, but that for some future case it could be a justifiable defense.

This is a quote I found. "The Court, deeply concerned by the rash of file-sharing lawsuits, the imbalance of resources between RIAA and individuals being sued, and the upheaval of norms of behavior brought on by the Internet, did everything in its power to permit Tenenbaum to make his best case for fair use."

I don't see how this is a good thing, making 'file sharing' and 'fair use' synonymous.

I agree. Making a small clip of a copyrighted work so the work could be discussed is one thing, but crying 'fair use' at the last minute of your lost defense only makes you look desperate.

Downloading whatever you feel like just to list on your iPod is completely legal in the U.S. UPLOADING the copywritten files is illegal and infringement of copyrights. This is a very important difference that the RIAA doesn't want you to realize.

Really? I am not so sure. You can get into a lot of trouble just by downloading a file.

The thing is, in many cases you are uploading a file as well when you are downloading it, especially if you are doing it with BitTorrent.

Except I didn't need the court to suggest to me it might be legal to rip my music off CDs I purchased to whatever format I choose. If some law made it illegal then there is something wrong with the law.

I don’t get how people are so polarized about this. Look, it’s against the law to infringe copyrighted material. It’s against the law to aid somebody else breaking the law. File sharing therefore is Against the Law.

Don’t focus on whether somebody wins or loses their cases. It is Proper that they lose! It would be Wrong if the law bent so much to allow what is clearly outlawed. Instead, seek to change the law, but don’t expect to get off the hook for doing it until you change the law.

You got it. Tenenbaum has admitted to file sharing of copyrighted songs, he’s lost under current law no matter how you look at it. Judges can’t over rule the law, merely work within it’s boundaries.

Actually judges can (and should) over-rule the law. The judiciary exists as a check on both the executive in the application of the laws as written, and on the legislature in the drafting of laws that are in accord with constitutions and with individual rights. If the judiciary isn't going to over-ride unconstitutional laws, no one else will.

In previous copyright cases in the US, things just aren’t that simple. The RIAA is not to be trusted in court. Their history includes threatening the wrong people such as the 2003 threat against Penn State's Prof. Usher who, with his team of researchers, innocently recorded a song in celebration of their new telescope. How did they get caught in the RIAA's all-too-blind dragnet? Apparently they dared to store an MP3 file containing the strings "usher" and ".mp3" in the filename on a publicly-accessible FTP server and nobody at RIAA thought to listen to the file before launching into litigation threats.

People are polarized about this issue because they sometimes see the needless legal suffering and hypocrisy brought by well-funded copyright maximalists and they don't want those maximalists defining the contours of copyright law alone.

The problem is these people are convicted for INSANE damages. Ok. How much is a song worth? About 99 cents, right? But to the RIAA the can sue for hundreds and thousands for a single song. So what do you think would happen if I stole a CD from Wal-Mart and they found out about it? They would probably charge me a few hundred dollars, and perhaps ban me from the store, etc. They wouldn’t sue me for many thousand dollars.

The law states that penalties should not be outrageous. I think anything more than $20-$30 a song is outrageous. The RIAA did not lose much of anything whenever a song is "pirated".

In what could be a turning point in the RIAA’s litigation campaign, a Michigan judge has decided to appoint pro bono counsel (a private lawyer who will defend you for free) to represent college student Brittany Kruger, who is being sued by the RIAA in Sony Music v. Kruger.

"If other judges follow suit, things will change dramatically." That is because the RIAA’s entire litigation campaign is based upon economic inequality of the parties. Almost none of those sued by the RIAA can afford legal representation, and the RIAA has a huge economic incentive to fight cases to the death.

If the courts follow the lead and appoint pro bono legal counsel in future cases, the RIAA will no longer be able to achieve the easy pickings default judgments and settlements it’s routinely obtained in the past.

Faulty assumption? This assumes the appointed pro-bono counsel is competent and interested in the welfare of his/her client, which may or may not be the case.

They’re required to be interested in the welfare of their client. The issue is more that a public defender (a state provided lawyer) will likely have 40 to 50 cases to worry about at the same time, where a private pro-bono defender may have 2 to 3 cases to worry about. That means more time to spend on each case.

Pro bono is not the same thing as public defender. Public defenders are for criminal cases. This is civil. There are lawyers in big firms who take on cases for the public good. These lawyers have an enhanced sense of social responsibility. Pro bono is short for "pro bono publico" (for the benefit of the public).

The pro bono lawyer will probably be skilled and ethical and not simply out to make a name for himself/herself. The defendant won't get absurd theatricals and stupid gamesmanship, but will get decent fair representation.

That alone should be a pain for the RIAA.

They are required by the rules of ethics to provide zealous representation to their client, even if the client is a pro bono client.

Assuming they are competent, all I can say is that It's about time.

Actually, as an attorney, I can say that it's not always that cynicism-worthy. Many young, bright attorneys choose to work for firms that give them leeway to take on pro bono cases, specifically because they know that they can do some good for the world by taking them. I personally hate the numerous times every week that my phone rings and I think to myself, "I wish I could afford to help this guy for free, because what was done to him is just plain wrong.," and then explain that it's not the type of case I can handle on a contingency fee basis and that it will cost him X dollars per hour. I give every person who calls me the advice not to chase bad money with good, and all too often they take it, much to my chagrin, because I really wanted to help them out.

When you buy music, make sure to check http://riaaradar.com/ [riaaradar.com] to see if the album is from a company that funds the RIAA. If they do, don't buy it and stick it to them a couple dollars of lost earnings at a time.

Buy used CDs instead whenever possible, and if you really want to support the bands themselves, send the difference in cash.

Any band that has signed with a major label since Napster was shut down is complicit.

I agree with the fact of their complicity but not with the degree. Don't punish the artists who are actually creating art just because they are part of a system that gives them no choice for ultimate success other than signing with a major label. Punish the source of the problem, not a fellow victim.

There are a few successful acts that did not sign with the big guys. Bright Eyes comes to mind. In fact, the number 5 album at Amazon right now is Bryan Sutton, which is ranked "Safe" by RIAA Radar. A harder path to be certain, but I'll actually respect those guys.

It’s great news, but doesn't fix the problem.

I guess now all the RIAA will do is shift their efforts to people that earn too much to get Pro Bono, but still don’t earn enough to be able to defend themselves against being hounded with litigation. In fact this is probably most of us.

Is there some specific income limit to receive pro bono counsel? If there isn't, how could someone have enough money to hire a lawyer to defend them (and not qualify for pro bono counsel) and at the same time not have enough money to hire a lawyer to defend them?

Someone who makes a decent middle-class salary may not have the means to add legal counsel to the list of bills. Just because someone makes a decent amount of money does not mean that they have discretionary income to throw around.

Almost nobody can afford to expend hundreds of thousands of dollars in attorneys fees, which is what the RIAA makes sure a contested case will cost.

These cases are more expensive than they need to be because of the RIAA tactics.

Sigh… I know I shouldn't feel this way, but I just don't care any more. The RIAA has worn me out. I hate all music now. I never want to buy any of their crap again.

There are plenty of bands that allow the release of their live stuff on the web and no, we're not talking about crappy Indy artists that you've never heard about before. We're talking real bands that care more about their fans and who actually tour rather than live off the royalties of overprocessed studio.

I suggest supporting those bands by buying their records and/or going to their shows instead of paying for music that sucks and that has no chance of expanding what is freely available out there.

Level 4 with scaffolding

Topic: Compare and contrast the positions of those for and against RIAA's strategies.

RIAA to stop prosecuting individual file sharers. According to the Wall Street Journal, the RIAA has decided to abandon its current tactic of suing individuals for sharing copyrighted music. Ongoing lawsuits will be pursued to completion, but no new ones will be filed.

Instead, the RIAA is going to try working with the ISPs to limit file-sharing services and cut off repeated users. This very surprising development apparently comes as a result of public distaste for the campaign.

This has less to do with the RIAA deciding to switch tactics in enforcing copyrights and it has more to do with the RIAA not wanting a legal precedent set about file sharing.

Did they finally get some legal advice? I mean, their current methods have apparently at least been in breach of investigative laws in several states and they may still end up in mess because of it, but ending the thing will at least lessen the exposure…

Alternative explanation is that they have actually understood that extortion is bad.. nah.. not likely.

The new tactic is lobby not litigate - far worse in the long run since they can keep trying to influence policy and legislation ad infinitem even if they get shot down the first time.

It also shifts the costs of enforcement (and the negative PR) to the government. Why bother pursuing people you *think* might be infringing and deal with the situation via civil means when you can just have the FBI issue the appropriate paperwork, and have them bust the door down?

This is a terrible news. When you actually had to break the law in order to get the RIAA all up in your jock, non-law-breakers such as myself were left in relative peace. Since they've now explicitly and announcedly decided to adopt a strategy of technology control measures, they just became a thorn in every geek's side.

Won't matter in the long run. They can't stop the sharing no matter what they do. But they can keep making life difficult until the public comes to realize sharing is impossible to control and instantly dismisses these ridiculous attempts to do so.

They've tried technological and legal solutions. They've tried appeals to morals and ethics (think of the starving artists), but they've undercut themselves mightily on that one. It's hard enough to lock things up, let alone ideas.

To progress, we need sharing. That's what the patent system was supposed to encourage. Copyright is a little different--it focuses on encouraging production rather than the sharing of ideas. Apparently sharing was expected to be so easily accomplished once copyright expired that they didn't think to provide provisions in copyright law to help sharing along, such as funds for public libraries.

I understand your concern. But the interesting thing about the industry is that it fails utterly to understand all things digital. Completely. It's unreal that they have not a single person who can explain these matters to them.

Sure, they want the ISPs to do their jobs for them. "Stop those pirates! They're on your nets, so you stop 'em!" What they fail to understand is that it is impossible.

I prefer it. I am in my 40's and grew up in the 80's. Piracy was rampant among the geeks then. In the 90's more so. Then the kids that only remember a world of the Internet. Do you really believe that they consider making a digital copy of a file is a crime? That it robs somebody? Remember these kids will be Judges and Lawmakers someday. No matter how much money the RIAA throws at it, it won't help in the long run.

Trust me, the RIAA has an uphill battle. Enough people feel that "stealing music" by downloading MP3's is about as wrong as jaywalking that piracy will not go away.

In Sony Music v. Tenenbaum, the defendant has filed a motion for new trial, attacking the constitutionality of the jury's $675,000 award as being violation of due process. He also argues that the Court's application of fair use doctrine was incorrect.

Music singles sell for 99 cents to $1.50. That's the damages.

Shoplifting and copyright infringement are not comparable. If you shoplift a pair of pants, the store cannot sell them to someone else. The store takes a hard loss of the cost they paid to acquire the pants. Making a copy of a music file, by contrast, does not cause any direct damage to someone selling copies of that file, since they still have the undiminished ability to continue selling copies.

"Making a copy of a music file, by contrast, does not cause any direct damage to someone selling copies of that file, since they still have the undiminished ability to continue selling copies."
Wrong. The presence of free goods will always displace higher cost goods. What that means is if I can download it for free, why would I (or anyone for that matter) pay for it. So one pirate can eliminate the revenue for a work completely.

No, you're wrong. I buy tracks on iTunes and Amazon every day, which I could choose to illegally download or rip from a friend's CD. I even borrow a friend's iPod sometimes, listen to some tracks then go and pay to download the ones I like when I could quite easily copy the (DRM free) tracks straight off the hard drive. I believe in doing my best to comply with the law and pay people for the work they produce.

I also believe (and this is not a contradiction) that copyright terms should be massively shortened, probably to no more than 20 years. The most effective answer to piracy is not more litigation, but to ensuring that music and movies are available cheaply and conveniently through legitimate channels.

"So one pirate can eliminate the revenue for a work completely."
Got any examples where that has happened? I can point out many examples where the copyright holder made the material available for free and still made money by selling the same thing. In fact, I can point to a few examples where they found that giving it away for free increased sales.

> The presence of free goods will always displace higher cost goods.
> Note that I did put emphasis on the word "direct" in the sentence you quoted. You are correct that there are indirect damages, but they are nothing like the direct damages suffered when a physical good is shoplifted, which was the ENTIRE POINT of my comment.

That's the actual damages per download.
Factor in that distributing can lead to many downloads (especially if there is some concept of downstream culpability, and those people all distribute) and one MP3 can easily be downloaded many hundreds or thousands of times, and actual damages can reach those amounts.

> Factor in punitive damages, which can be many times higher than actual damages, and the "correct" damages far exceed 99 cents.

The RIAAs strategy seems to rely on people confusing downloading with uploading; the mere fact that I am downloading a song does not prove any intent by me to redistribute said content. Uploading content, on the other hand, does imply intent to redistribute.

> The RIAAs strategy seems to rely on people confusing downloading with uploading; the mere fact that I am downloading a song does not prove any intent by me to redistribute said content.

> To the best of my knowledge, all the lawsuits have been focused on distributors, not downloaders. So, far from them relying on people confusing the two, they have solely gone after the uploaders.

Torrents are a gray area; many people may not understand that while they are downloading it, the bittorent software is also making small pieces of the file available for others to download. But traditionally, "small pieces" has been considered fair use.

> But traditionally, "small pieces" has been considered fair use.
> Typically, using a small subsection is a necessary, but not sufficent, condition for fair use. Fair use turns on the motivation (commentary, satire, etc.) Since the purpose is to allow the entire file to be recreated, I doubt any reasonable fair use argument could be made.

that's the damages, folks.

> So, you buy a song on iTunes for 99 cents. You think you just bought a license for unlimited distribution, even though the copyright owner expressly refused to sell you one at that price? Why don't you believe in the freedom of contract and the free market?

the distribution right is worth substantially more than the right to make one copy. That's why it's an entirely separate right. Things like the fair use doctrine, the Betamax doctrine, format-shifting and time-shifting, and even the first sale doctrine all apply to that one copy. Mass distribution is another animal altogether.

Anyways, Congress already considered all this, most recently in 1996 when they last amended the Copyright Act and explicitly mentioned file sharing as their justification to raise the statutory damage provisions. The constitutionality argument really has no weight.

Trying to fight RIAA in the courts is a losing effort. RIAA pay politicians handsomely, and generally gets the laws they want. If they temporarily loose in court, they just pay to have the laws changed, and than they win. The draconian penalties as well as the never expiring rights RIAA enjoy is an amazing perversion.

The only thing that is worse is that this can happen in a democracy, and few care.

If you argue "well, just pay the $0.99 on iTunes and stop whining" you misunderstand culture fundamentally. Humans as a species copy. From infants looking at their parents to musicians, architects, engineers and philosophers listening to others, we refine and produce. This is the essence of human culture. That companies can monopolize this flow is damaging to the progress of mankind.

It’s not about copying. The fundamental problem here is not that of copying, but the matter of justice in proportion to the crime. Suppose, for example, we take the RIAA's argument at face value: Because she's shared these 19 songs, the RIAA companies will never make another sale from them. According to the RIAA, she owes them for the lost profits they would have made.

Even were this the case, the maximum cost of these 19 songs is the cost the RIAA paid to the artists to produce them, which isn’t much.

In terms of actual damages, she probably resulted in no lost sales. What the RIAA doesn't understand is that with the exception of the upper-middle and upper classes, most of America has become accustomed to getting their music for free, without paying a dime. If they can't get it for free, they just do without. It is almost never a lost sale.

What disturbs me most is that a jury could be convinced to grant a judgment of a few million dollars against her without any actual proof of infringement. They have no idea how many - if any - downloads actually occurred.

We in the modern west have a problem, and I, for one, do not see an easy solution. It used to be that making copies of creative works was a physical task that was the domain of professionals. As such, enforcing copyrights was relatively easy. As soon as copyrightable creative works were representable in digital form, and computers became capable of copying them trivially, that changed utterly. Copyrights exist so that creators of creative works can be given an incentive to create. Their creations, on the whole, enrich society. That's the basic copyright bargain: You write good books and we, as a society, will insure that you can make a living doing it.

> In an era where digital representations of copyrightable works can be freely, however, the idea of being able to police the copying so that authors of creative works can be fairly compensated becomes impossible. The Copyright system no longer functions properly because conditions in the world have changed irrevocably. I don't have an answer as to how to fix it. Nobody does, because if they did, things would be different.
>
> > Copyrights exist so that creators of creative works can be given an incentive to create. Their creations, on the whole, enrich society. That's the basic copyright bargain: You write good books and we, as a society, will insure that you can make a living doing it. Copyrights exist to serve the public interest. Part of the means by which they work is to give authors an additional incentive to create and publish beyond those which are naturally present. Creation and publication are both important, as unpublished works do so little good for society that they may as well not exist.
> >
> > > The point is that with no copyright, how is any author ever going to earn any money? As soon as a new book was published, anyone could copy it and sell it without the author getting a penny.
>
> "It used to be that making copies of creative works was a physical task that was the domain of professionals. As such, enforcing copyrights was relatively easy. " Copyright came into existence when technology enabled copies to be made cheaply if you wanted many copies.
>
> > "Copyrights exist so that creators of creative works can be given an incentive to create. Their creations, on the whole, enrich society. That's the basic copyright bargain: You write good books and we, as a society, will insure that you can make a living doing it."
> > Actually the theory is more that if there is money to be made from a work the author should get it. It's by no means clear that the possibility of profit plays any great part in motivating people to be creative in the first place.

People who created music should be rewarded, not RIAA.

Copyrights exist so that creators of creative works can be given an incentive to create. Their creations, on the whole, enrich society. That's the basic copyright bargain
That might be the way it works on paper, but how do copyrights on computer software enrich society, I can keep my work excluded from society until 50 (to 100) years after I die, by which time the only hardware that will be capable of running it will be in a museum.

Trying to fight RIAA in the courts is a losing effort. If the defendant wins the statutory damages argument in court on constitutional grounds then it will not have been wasted effort because it means that the copyright holders would have to get the Constitution amended to specifically allow for LARGE statutory damages for copyright infringement. It requires a super-majority vote in both houses of Congress and a super-majority vote of the states to amend the US Constitution.

"Trying to fight RIAA in the courts is a losing effort. RIAA pay politicians handsomely, and generally gets the laws they want."
This is why the drafters of the constitution put the judicial branch into place. No matter how many lawmakers they bribe, the RIAA can never change the constitution, nor can they implement any unconstitutional laws.

As much as I hate to admit it, you're right. It is a losing effort. But the goal is to make sure it's a losing effort for the RIAA as well. The best case scenario for them is that they collect every penny of the $675k. I'm willing to bet they owe their law firm a good chunk of that in billable hours. The longer the defendant drags this out costly it will be form them. At some point, the money gained from the settlement will be far less than the money they owe their lawyers.

Or however much they can squeeze from him as creditors during a bankruptcy.

Sure, he should pay a fine. One in the order of, say, $675, not $675000.

He offered $500 to Sony, to my knowledge, and they turned him down and have now succeeded in the big bucks.

"He offered $500 to Sony, to my knowledge, and they turned him down and have now succeeded in the big bucks."
Sony needed much more to offset the cost of developing and marketing all their proprietary formats... ATRAC, MiniDiscs, and MemorySticks all took a lot of effort you know.

Don't forget Betamax and Blu-Ray. Where would we be today without these amazing innovations? Using VHS and DVD like a bunch of savages, that's where.

The RIAA is asking the FCC not to make the net neutrality rules so strict that they 'would limit broadband providers' ability to allow illegal online file sharing.'

Maybe the RIAA are right? We should block file sharing - let’s start with the ISP the RIAA use, and their drones and employees and the Record Labels. That way they'll never be able to tell what the rest of the internet is sharing.

No kidding. They want limits on Net Neutrality, I want limits on their access to the Federal Government. These are the kind of people that drive the need for election reform.

RIAA Wants Limits on Net Neutrality... ... then it wouldn't be net neutrality, would it? The RIAA is so technologically clueless that I am frequently entertained at the depth and breadth of the stupid things that they say.

They never seem to really understand the concept of anything, but I've gotta say, this may be the first time I've seen them make a proposal that was an oxymoron right out of the box.

Quoting: "Internet service providers should have authority to block subscribers from sharing music and other files without permission of the copyright owner, the RIAA said."

I don't think highway operators in this country have ever been compelled or encouraged to stop grand theft auto, or interstate smuggling of stolen goods... Or that phone companies have been expected to prevent con artists from swindling people out of their money to buy "beach-side" Florida swamp land. Et cetera. This would appear to be unprecedented.

There's nothing in the law to make ISPs filter copyrighted content. What the RIAA wants is to cleverly sneak in the back door and use the FCC's regulatory powers to force ISP level filtering.

Not only that, but why should ISPs care? Because of the increased bandwidth use? They have simpler tools to counter this (like metering the use and having their user pay for it). Independent of what's sent through the cable.

It would not only cost the ISPs money, that's actually a minor concern. The liability is. From "they should have the authority" it's a tiny step to "they have to". And then they'd be liable for any file sharing going on through their networks.

I could not see anything any ISP would fight harder against than that. Because it is nearly impossible to police that without pretty much banning all traffic except. And ponder how many of their users this would piss off and make them quit. All online gamers (because games, especially MMOs, pretty much have to use nonstandard data packets to avoid cheating and use of unauthorized clients) and all VPN users (and those are usually the ones with the fat contracts, think companies and the like) first of all.

If the RIAA wants a fight with the major ISPs, this is the way to do it.

Unless, of course, there's a common interest with a company who owns the ISPs *cough, Time Warner, cough*

And here's where the schizophrenia of large companies set in: Just because Time Warner is content provider and ISP, it doesn't mean that they cooperate. Look at Sony and how their content crippling caused their hardware to stop reading their own content.

But who says who the copyright owner is? The RIAA? The RIAA wants to own all music and make un singed band be locked out of putting their own music on line for free without paying the RIAA for the right.

About the only bands that escape the music industry unsigned are the indies. Most of the rest get their fingers burnt.

It is a federal offense to riffle through someone else's mail. This nonsense by the RIAA and friends is like saying "yeah we agree that FEDEX etc. shouldn't be going through other peoples' mail... except to make sure that people aren't pirating things..." Everyone understands that position to be completely ridiculous so why is it that the concept is so difficult to apply to internet packets etc? Just as your mail is legally protected from being ripped open by others, so should your internet packets. It isn't the job of ISPs to do the RIAA's work nor is it their right to riffle through your online activities at their whim.

You're right. What don't these RIAA jerks get? It's almost like they are operating with their own agenda or something.

The problem is that "opened mail" is difficult to conceal without failing to deliver it. "Inspected Internet packets" on the other hand, look just the same upon delivery.

You can easily inspect mail by using vapor to open the letters.

Yes, but the issue is not whether you can, it’s whether you legally are/should be allowed to.

I'm not saying it's right, but are carriers such as FedEx and UPS bound by those same restrictions. I understand that it's illegal to snoop through USPS mail, but what about private, commercial carriers?

I'm all for ISP inspection of packets, as long as the ISP becomes criminally liable for EVERYTHING that goes over their network.

I know that's the standard line and I would probably agree that, if they were to inspect every packet, then they probably should become liable for the traffic they choose to allow. However, I would not be OK with that setup. You'd quickly see the Internet turn into a for-pay broadcast network.

People need to be able to communicate privately and even anonymously. If it's possible for a society to be arranged to function without privacy and anonymity, then we don't yet have a precedent to look to.

Yes, I agree.

Bibliography

Ackerman, M. S. (1994). Augmenting the Organizational Memory: A Field Study of Answer Garden. *CSCW*, 243-252

Ackerman, M. S. and L. Palen (1996). The Zephyr Help Instance: promoting ongoing activity in a CSCW system. *CHI*: 268-275.

Ackerman, M. S. and T. W. Malone (1990). Answer Garden: a tool for growing organizational memory. *Proceedings of the ACM SIGOIS and IEEE CS TC-OA conference on Office information systems*, Cambridge, Massachusetts, United States, 31-39

Ackerman, M. S. and D. W. McDonald (1996). Answer Garden 2: Merging Organizational Memory with Collective Help. *CSCW*, 97-105

Ackerman, M. S., A. Swenson, S. Cotterill and K. DeMaagd (2003). I-DIAG: From Community Discussion to Knowledge Distillation. *International Conference on Communities and Technologies*,

Ackerman, M.S. and C. Halverson (2004). Sharing Expertise: The Next Step for Knowledge Management. *M. H. Huysman and V. Wulf (Ed.). Social Capital and Information Technology*. Cambridge, MIT Press**:** 273-299.

Ackerman, Mark S. and Christine Halverson (2000). Re-Examining Organizational Memory. *Communications of the ACM* 43(1): 58-63.

Ackerman, Mark S. and Brian Starr (1995). Social activity indicators: interface components for CSCW systems. *Proceedings of the 8th annual ACM symposium on User interface and software technology*, Pittsburgh, Pennsylvania, United States, ACM, 159-168

Ackerman, Mark S., Volker Wulf and Volkmar Pipek (2002). *Sharing Expertise: Beyond Knowledge Management*, MIT Press.

Adamic, L. A., J. Zhang, E. Bakshy and M. S. Ackerman (2008). Knowledge Sharing and Yahoo Answers: Everyone Knows Something. *WWW*, 665-674

Agichtein, E., C. Castillo, D. Donato, A. Gionis and G. Mishne (2008). Finding High-Quality Content in Social Media. *WSDM*, 183-194

Albayrak, Sahin, Stefan Wollny, Nicolas Varone, Andreas Lommatzsch and Dragan Milosevic (2005). Agent technology for personalized information filtering: the PIA-system. *Proceedings of the 2005 ACM symposium on Applied computing*, Santa Fe, New Mexico, ACM, 54-59

alias-i (2008). LingPipe. Retrieved 04/09, 2009, from http://alias-i.com/lingpipe/index.html.

Amitani, S. and K. Hori (2003). Knowledge nebula crystallizer for knowledge lequidization & crystallization- from a theory to a methodology of knowledge management. *Proceedings of Expertise In Design, Design Thinking Research Symposium*,

Amitani, Shigeki and Ernest Edmonds (2007). A method for visualising possible contexts. *KES/WIRN*, 901–908

Avrachenkov, Konstantin, Vladimir Dobrynin, Danil Nemirovsky, Son Kim Pham and Elena Smirnova (2008). Pagerank based clustering of hypertext document collections. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, Singapore, Singapore, ACM, 873-874

Barabási, Albert-László and Réka Albert (1999). Emergence of scaling in random networks. *Science* 286: 509-512.

Belkin, Nicholas J. and W. Bruce Croft (1992). Information filtering and information retrieval: two sides of the same coin? *Commun. ACM* 35(12): 29-38.

Bernstein, Mark, Peter J. Brown, Mark Frisse, Robert Glushko, Polle Zellweger and George Landow (1991). Structure, navigation, and hypertext: the status of the navigation problem. *Proceedings of the third annual ACM conference on Hypertext*, San Antonio, Texas, United States, ACM, 363-366

Brewer, Robert S. (2000). Improving problem-oriented mailing list archives with MCS. *Proceedings of the 22nd international conference on Software engineering* Limerick, Ireland, 95-104

Buckingham Shum, S. (1996). Design Argumenttion as Design Rationale*(Ed.). The Encyclopedia of Computer Science and Technology*. New York, Marcel Dekker, Inc.**:** 95-128.

Butler, B., L. Sproull, S. Kiesler and R. Kraut (2007). Community effort in online groups: Who does the work and why? *S. Weisband. (Ed.). Leadership at a distance***:** 171-194.

Carroll, John M. and Caroline Carrithers (1984). Training wheels in a user interface. *Commun. ACM* 27(8): 800-806.

Conklin, J. (1992). Corporate Memory*(Ed.). Groupware*. San Jose, Morgan-Kaufmann**:** 131-137.

Conklin, J. and M. L. Begeman (1988). gIBIS: a hypertext tool for exploratory policy discussion. *CSCW*, 140-152

Constant, D., L. Sproull and S. Kiesler (1996). The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice. *Organization Science* 7(2): 119.

Deerwester, S., S. T. Dumais and G. W. Furnas (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science* 41(6): 391-407.

Dellarocas, C. (2003). The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms. *Management Science* 49(10): 1407-1424.

Donath, Judith, Karrie Karahalios and Fernanda Viegas (1999). Visualizing conversation. *HICSS-32*, Maui, HI,

Dorogovtsev, S. N. and J. F. F. Mendes (2003). Evolution of Networks: From Biological Nets to the Internet and Www*(Ed.)*. Oxford, Oxford University Press.

Dourish, Paul and Matthew Chalmers (1994). Running Out of Space: Models of Information Navigation. *HCI*, Glasgow, UK,

Emigh, William and Susan C. Herring (2005). Collaborative Authoring on the Web: A Genre Analysis of Online Encyclopedias. *Proceedings of the Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05) - Track 4 - Volume 04*, IEEE Computer Society, 99.91

Erickson, Thomas, David N. Smith, Wendy A. Kellogg, Mark Laff, John T. Richards and Erin Bradner (1999). Socially translucent systems: social proxies, persistent conversation, and the design of \“babble\&rdquo. *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, Pittsburgh, Pennsylvania, United States, ACM, 72-79

Eto, K., S. Takabayashi and T. Masui (2005). qwikWeb: integrating mailing list and WikiWikiWeb for group communication. *In Proceedings of the 2005 international symposium on Wikis*, San Diego, California, 17-23

ETS (2008). Report on the Findings of An External Alignment Study for New York State's Regents Examination in Integrated Algebra. Retrieved 07-02, 2010, from http://www.emsc.nysed.gov/osa/pub/ia-ext.pdf.

Farnham, S., H. R. Chesley, D. E. McGhee, R. Kawal and J. Landau (2000). Structured online interactions: improving the decision-making of small discussion groups. *In*

Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, Philadelphia, Pennsylvania, United States, 299-308

Fekete, J. and C. Plaisant (2003). DTD describing a tree structure for visualization. Retrieved May 26, 2007, from http://www.nomencurator.org/infoVis2003/download/treeml.dtd.

Fellbaum, C. (1998). Wordnet: An Electronic Lexical Database. *C. Fellbaum (Ed.)*. Cambridge, MA, MIT Press.

Finholt, Thomas Albert (1993). Outsiders on the inside: sharing information through a computer archive, Carnegie Mellon University**:** 129.

Fiore, Andrew T., Scott Lee Tiernan and Marc A. Smith (2002). Observed behavior and perceived value of authors in usenet newsgroups: bridging the gap. *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, Minneapolis, Minnesota, USA, ACM, 323-330

Foltz, Peter W. and Susan T. Dumais (1992). Personalized information delivery: an analysis of information filtering methods. *Commun. ACM* 35(12): 51-60.

Forte, Andrea and Amy Bruckman (2006). From Wikipedia to the classroom: exploring online publication and learning. *Proceedings of the 7th international conference on Learning sciences*, Bloomington, Indiana, International Society of the Learning Sciences, 182-188

Fry, Ben and Casey Reas (2001). Processing. Retrieved 04/09, 2009, from http://processing.org.

Giles, Jim. (2005). Internet encyclopedias go head to head. *Nature*(438): 900–901.

Goldberg, David, David Nichols, Brian M. Oki and Douglas Terry (1992). Using collaborative filtering to weave an information tapestry. *Commun. ACM* 35(12): 61-70.

Gómez, Vicenç, Andreas Kaltenbrunner and Vicente López (2008). Statistical analysis of the social network and discussion threads in slashdot. *Proceeding of the 17th international conference on World Wide Web*, Beijing, China, ACM, 645-654

Groot, Reuma De, Raul Drachman, Rakheli Hever, Baruch B. Schwarz, Ulrich Hoppe, Andreas Harrer, Maarten De Laat, Rupert Wegerif, Bruce M. McLaren and Benoit Baurens (2007). Computer supported moderation of e-discussions: the ARGUNAUT approach. Proceedings of the 8th iternational conference on Computer supported collaborative learning. New Brunswick, New Jersey, USA, International Society of the Learning Sciences**:** 168-170.

Grudin, J. (1987). Social Evaluation of the User Interface: Who Does the Work and Who Gets the BENEFIT? *IFIP INTERACT: Human-Computer Interaction*, 805-811

Grudin, Jonathan (1988). Why CSCW applications fail: problems in the design and evaluationof organizational interfaces. *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*, Portland, Oregon, United States, ACM, 85-93

Guan, Ziyu, Can Wang, Jiajun Bu, Chun Chen, Kun Yang, Deng Cai and Xiaofei He (2010). Document recommendation in social tagging services. *Proceedings of the 19th international conference on World wide web*, Raleigh, North Carolina, USA, ACM, 391-400

Halasz, Frank G., Thomas P. Moran and Randall H. Trigg (1987). Notecards in a nutshell. *Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*, Toronto, Ontario, Canada, ACM, 45-52

Hansen, D. L. (2007). Knowledge Sharing, Maintenance, and Use in Online Support Communities, University of Michigan. **Ph.D.**

Hansen, D. L., M. S. Ackerman, P. J. Resnick and S. Munson (2007). Virtual Community Maintenance with a Collaborative Repository. *ASIS&T*,

Harper, F. M., D. Raban, S. Faraeli and J. A. Konstant (2008). Predictors of Answer Quality in Online Q&A Sites. *CHI*, New York, NY, USA, 865-874

Heer, J. (2007, April 05, 2007). prefuse | interactive information visualization toolkit. Retrieved May 26, 2007, from http://prefuse.org.

Helic, D., H. Maurer and N. Scerbakov (2003). Reusing discussion forums as learning resources in WBT systems. *International Conference on Computers and Advanced Technology in Education*, 223-228

Hess, K. (2004). Applying Webb's Depth-of-Knowledge (DOK) Levels in Reading. Retrieved 07-02, 2010, from http://www.nciea.org/.

Hess, K. (2005). Applying Webb's Depth-of-Knowledge (DOK) Levels in Social Studies. Retrieved 07-02, 2010, from http://www.nciea.org/.

Hester, Andrea J. (2010). Increasing collaborative knowledge management in your organization: characteristics of wiki technology and wiki users. *Proceedings of the 2010 Special Interest Group on Management Information System's 48th annual conference on Computer personnel research on Computer personnel research*, Vancouver, BC, Canada, ACM, 158-164

Hewitt, J. (2001). Beyond Threaded Discourse. . *International Journal of Educational Telecommunications.* 7(3): 207-221.

Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion threads. *Journal of Educational Computing Research* 28(1): 31-45.

Hill, William C., James D. Hollan, Dave Wroblewski and Tim McCandless (1992). Edit wear and read wear. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Monterey, California, United States, ACM, 3-9

Hoadley, Christopher M. and Peter G. Kilner (2005). Using technology to transform communities of practice into knowledge-building communities. *SIGGROUP Bull.* 25(1): 31-40.

Hori, K., K. Nakakoji, Y. Yamamoto and J. Ostwald (2004). Organic perspecives of knowledge management: Knowledge evolution through a cycle of knowledge liquidization and crystallization. *Jounal of Universal Computer Science* 10(3): 252-261.

Ito, Kyoko, Eriko Mizuno and Shogo Nishida (2007). Design and development of computer-based discussion support tool for science and technology communication exercise. *Proceedings of the 2007 conference on Human interface: Part II*, Beijing, China, Springer-Verlag, 356-363

Jeon, J., W. B. Croft and J. H. Lee (2005). Finding Similar Questions in Large Question and Answer Archives. *CIKM*, 84-90

Jeon, Jiwoon, W. Bruce Croft, Joon Ho Lee and Soyeon Park (2006). A framework to predict the quality of answers with non-textual features. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, Seattle, Washington, USA, ACM, 228-235

Joyce, Michael (1991). Storyspace as a hypertext system for writers and readers of varying ability. *Proceedings of the third annual ACM conference on Hypertext*, San Antonio, Texas, United States, ACM, 381-387

Judge, Robert (2007). The Influence of Information Linkage on User Acceptance of a Knowledge Management System (KMS) in Small to Mid-size Enterprises (SME). *HICSS*, Waikoloa, HI 194a - 194a

Kim, G. T. (2007, March). KOREA: Naver's Knowledge Search Database Tops 60 Million. *Asia Media Media News Daily*. Retrieved June 2, 2008.

Kim, Han-Joon and Sang-Goo Lee (2000). A semi-supervised document clustering technique for information organization. *Proceedings of the ninth international conference on Information and knowledge management*, McLean, Virginia, United States, ACM, 30-37

Kintsch, E., D. Steinhard and G. Stahl (2000). Developing summarisation skills through the use of LSA-based feedback,

Kittur, Aniket, Bongwon Suh, Bryan A. Pendleton and Ed H. Chi (2007). He says, she says: conflict and coordination in Wikipedia. *Proceedings of the SIGCHI conference on Human factors in computing systems*, San Jose, California, USA, ACM, 453-462

Klein, D. and C. D. Manning (2001). Parsing with treebank grammars: empirical bounds, theoretical models, and the structure of the Penn Treebank. *Proceedings of the 39th Annual Meeting on Association For Computational Linguistics*, 338-345.

Kolla, Maheedhar and Olga Vechtomova (2007). Retrieval of discussions from enterprise mailing lists. *In Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, Amsterdam, The Netherlands, 881-882

Lampe, C. and P. Resnick (2004). Slash(dot) and burn: distributed moderation in a large online conversation space. *Proceedings of ACM CHI 2004 Conference on Human Factors in Computing Systems*, 543-550

Lee, J. (1990). SIBYL: A Tool for Managing Group Decision Rationale. *CSCW*, 79-92

Leuf, B. and W. Cunningham (2001). The wiki way.*(Ed.)*. Boston, MA, Addison-Wesley.

Lih, Andrew (2004). Wikipedia as Participatory journalism: reliable sources? metrics for evaluating collaborative media as a news resource. *In Proceedings of the 5th International Symposium on Online Journalism*, 16-17

Lin, Chin-Yew and E. Hovy (2003). Automatic Evaluation of Summaries Using N-Gram Co-occurrence Statistics. *Human Technology Conference*, Edmonton, Canada, 150-157

Liu, Yandong, Jiang Bian and Eugene Agichtein (2008a). Predicting information seeker satisfaction in community question answering. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, Singapore, Singapore, ACM, 483-490

Liu, Yuanjie, Shasha Li, Yunbo Cao, Chin-Yew Lin, Dingyi Han and Yong Yu (2008b). Understanding and summarizing answers in community-based question answering services. *Proceedings of the 22nd International Conference on Computational Linguistics - Volume 1*, Manchester, United Kingdom, Association for Computational Linguistics, 497-504

Lupfer, E. (2010). Expertise Location: The Killer App for Enterprise Social Computing. *ENTERPRISE SOCIAL TECHOLOGIES,SOCIAL MEDIA*.

MacLean, A., R. Young, V. Bellotti and T. Moran (1990). Questions, Options, and Criteria: Elements of a Design Rationale for User Interfaces. *EuroPARC/AMODEUS*,

Maltz, David and Kate Ehrlich (1995). Pointing the way: active collaborative filtering. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Denver, Colorado, United States, ACM Press/Addison-Wesley Publishing Co., 202-209

Marshall, C. C., F. G. Halasz, R. A. Rogers and W. C. Janssen (1991). Aquanet: a hypertext tool to hold your knowledge in place. *Proceedings of the ACM Conference on Hypertext*, 261-275

Marshall, C. C. and R. A. Rogers (1992). Two Years before the Mist: Experiences with Aquanet. *Proceedings of the ACM Conference on Hypertext (ECHT '92)*, Milano, Italy, 53-62

Marshall, C. C., F. M. Shipman and J. H. Coombs (1994). VIKI: Spatial Hypertext Supporting Emergent Structure. *Proceedings of the ACM European Conference on Hypermedia Technology (ECHT '94)*, Edinburgh, Scotland, 13-23

Masui, T. and S. Takabayashi (2003). Instant group communication with QuickML. *In Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work* Sanibel Island, Florida, 268-273

McDonald, David W. and Mark S. Ackerman (2000). Expertise recommender: a flexible recommendation system and architecture. *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, Philadelphia, Pennsylvania, United States, ACM, 231-240

McLaren, Bruce M., Oliver Scheuer, Maarten De Laat, Rakheli Hever, Reuma De Groot, Carolyn P. Ros\ and \#233 (2007). Using Machine Learning Techniques to Analyze and Support Mediation of Student E-Discussions. *Proceeding of the 2007 conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work*, IOS Press, 331-338

Miles, M. and H. Huberman. (1994). *Qualitative Data Analysis*, Sage Publications Inc.

Millen, D. R. and M. A. Fontaine (2003). Multi-team facilitation of very large-scale distributed meetings. *European Conference on Computer Supported Cooperative Work*, 259-275

Moran, T. P. and J. M. Carroll (1996). Design Rationale: Concepts, Techniques, and Use.*(Ed.)*. Mahway, NJ, Lawrence Erlbaum Associates.

Nakanishi, Takafumi, Koji Zettsu, Yutaka Kidawara and Yasushi Kiyoki (2009). SAVVY Wiki: a context-oriented collaborative knowledge management system.

Proceedings of the 5th International Symposium on Wikis and Open Collaboration, Orlando, Florida, ACM, 1-8

Nam, K. K., M. S. Ackerman and L. A. Adamic (2009). Questions in, knowledge in?: a study of naver's question answering community. *In Proceedings of the 27th international conference on Human factors in computing systems*, Boston, MA, USA, 779-788

Nam, Kevin K. and Mark S. Ackerman (2007). Arkose: reusing informal information from online discussions. *In Proceedings of the 2007 international ACM conference on Supporting group work*, Sanibel Island, Florida, USA, 137-146

O'Conner, John (2007). Using the Java Persistence API in Desktop Applications. Retrieved 04/09, 2009, from http://java.sun.com/developer/technicalArticles/J2SE/Desktop/persistenceapi/.

Ozaku, Hiromi, Kiyotaka Uchimoto, Masaki Murata and Hitoshi Isahara (2000). Topic search for intelligent network news reader HISHO. *Proceedings of the 2000 ACM symposium on Applied computing - Volume 1*, Como, Italy, ACM, 28-33

Passonneau, R. J. and D. J. Litman (1997). Discourse segmentation by human and automated means. *Comput. Linguist* 23(1): 103-139.

Porter, C. (2004). A Typology of Virtual Communities: A Multi-Disciplinary Foundation for Future Research. *Journal of Computer Mediated Communication* 10(1).

Porter, M. F. (1980). An algorithm for suffix stripping. *Program* 14(3): 130-137.

Preece, J. (2000). Online Communities: designing usability, supporting sociability.*(Ed.)*. New York, Wiley.

Preece, J., B. Nonnecke and D. Andrews (2004). The Top Five Reasons for Lurking: Improving Community Experiences for Everyone. *Computers in Human Behavior* 20(2): 201-223.

Radev, D. R. and E. Hovy (1999). Intelligent Text Summarization - Report on the AAAI Spring Symposium*. AI Magazine*. 20(3).

Resnick, Paul, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom and John Riedl (1994). GroupLens: an open architecture for collaborative filtering of netnews. *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, Chapel Hill, North Carolina, United States, ACM, 175-186

Ridings, C. M. (2004). Virtual community attraction: Why people hang out online. *Journal of Computer Mediated Communication* 10(1).

Sack, Warren (2000a). Conversation map: a content-based Usenet newsgroup browser. *Proceedings of the 5th international conference on Intelligent user interfaces*, New Orleans, Louisiana, United States, ACM, 233-240

Sack, Warren (2000b). Discourse Diagrams: Interface Design for Very Large-Scale Conversations. Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 3 - Volume 3, IEEE Computer Society**:** 3034.

Salton, G. and C. Buckley (1988). Term-Weighting Approaches in Automatic Text Retrieval. *Processing and Management* 24(5): 513-523.

Salton, Gerard, Edward A. Fox and Harry Wu (1983). Extended Boolean information retrieval. *Commun. ACM* 26(11): 1022-1036.

Shipman, F. M. (1999). Spatial Hypertext: An Alternative to Natigational and Semantic Links. *ACM Computing Surveys* 31(4).

Shipman, F. M. and C. C. Marshall (1999). Formality Considered Harmful: Experiences, Emerging Themes, and Directions on the Use of Formal Representations in Interactive Systems. *Computer Supported Cooperative Work Journal* 8(4): 333-352.

Shipman, F. M. and R. McCall (1994). Supporting knowledge-base evolution with incremental formalization. *Proceedings of the SIGCHI Conference on Human Factors in Computing System*, 285-291

Smith, Marc (1997). Netscan: Measuring and Mapping the Social Structure of Usenet. *Social Network Conference*, San Diego, California,

Smith, Marc A. and Andrew T. Fiore (2001). Visualization components for persistent conversations. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Seattle, Washington, United States, ACM, 136-143

Stahl, G. (1999). Reflections on WebGuide: seven issues for the next generation of collaborative knowledge-building environments. . *In Proceedings of the 1999 Conference on Computer Support For Collaborative Learning*, Palo Alto, California,

Streeter, L. A. and K. E. Lochbaum (1988). Who Knows: a system based on automatic representation of semantic structure. *RIAO*, 380-388

Suthers, Daniel D, Ravi Vatrapu, Richard Medina, Samuel Joseph and Nathan Dwyer (2008). Beyond threaded discussion: Representational guidance in asynchronous collaborative learning environments. *Comput. Educ.* 50(4): 1103-1127.

Uchimoto, Kiyotaka, Hiromi Ozaku and Hitoshi Isahara (1997). A Method for Identifying Topic-changing Articles in Discussion-type Newsgroups within the

Intelligent Network News Reader HISHO. *Natural Language Processing Pacific Rim Symposium*,

Viégas, F., M. Wattenberg and K. Dave (2004). Studying Cooperation and Conflict Between Authors with History Flow Visualizations. *CHI*, 575 - 582

Vries, Sjoerd de and Piet Kommers (2004). Online knowledge communities\&\#58; future trends and research issues. *Int. J. Web Based Communities* 1(1): 115-123.

Wang, Y., M. Joshi and C. P. Rosé (2008). Investigating the effect of discussion forum interface affordances on patterns of conversational interactions. *In Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work* San Diego, CA, USA, 555-558

Watanabe, Y., K. Sono, K. Yokomizo and Y. Okada (2004). A question answer system using mails posted to a mailing list. . *In Proceedings of the 2004 ACM symposium on Document engineering*, Milwaukee, Wisconsin, USA, 67-73

WCER (2007). Aligning State-Level Standards and Assessments. Retrieved 07-02, 2010, from http://www.wcer.wisc.edu/news/coverStories/aligning_state_standards.php.

Webb, Norman L. (2005). Depth-of-Knowledge Level Definitions. Retrieved 11-01, 2009, from http://facstaff.wcer.wisc.edu/normw/state%20alignment%20page%20one.htm.

Wegerif, Rupert, Bruce M. McLaren, Marian Chamrada, Oliver Scheuer, Nasser Mansour, Jan Miks\, \#225 and tko (2009). Recognizing creative thinking in graphical e-discussions using artificial intelligence graph-matching techniques. *Proceedings of the 9th international conference on Computer supported collaborative learning - Volume 1*, Rhodes, Greece, International Society of the Learning Sciences, 108-112

Welser, H. T., E. Gleave, D. Fisher and M. Smith (2007). Visualizing the Signatures of Social Roles in Online Discussion Groups. *Journal of Social Structure* 8(2).

Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity.* New York, Cambridge University Press.

Knowledge Search. (2004). In *Wikipedia, the Free Encyclopedia*. Retrieved 1-June-2008, from http://en.wikipedia.org/wiki/Knowledge_Search.

Internet forum. In *Wikipedia, the free encyclopedia*. Retrieved 11/01/2009, from http://en.wikipedia.org/wiki/Internet_forum.

Knowledge Search. (2009). In *Wikipedia, the Free Encyclopedia*. Retrieved 1-June-2008, from http://en.wikipedia.org/wiki/Knowledge_Search.

Mailing list. In *Wikipedia, the free encyclopedia*. Retrieved 11/01/2009, from http://en.wikipedia.org/wiki/Mailing_list.

Metcalfe's law. In *Wikipedia, the free encyclopedia*. Retrieved 07/07/2010, from http://en.wikipedia.org/wiki/Metcalfe's_law.

Newsgroup. In *Wikipedia, the free encyclopedia*. Retrieved 11/01/2009, from http://en.wikipedia.org/wiki/Newsgroup.

Yang, J., L. A. Adamic and M. S. Ackerman (2008a). Crowdsourcing and Knowledge Sharing: Strategic User Behavior. *EC*, 246–255

Yang, K.-H., C.-Y. Chen, H.-M. Lee and J.-M. Ho. (2008b). Efs:expert finging system based on wikipedia link pattern analysis. *In the 2008 IEEE International Conference on Systems, Man and Cybernetics (SMC 2008)*.

Yu, Bin and Munindar P. Singh (2003). Searching social networks. *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, Melbourne, Australia, ACM, 65-72

Yu, J., Z. Jiang and H. C. Chan (2007). Knowledge contribution in problem solving virtual communities: the mediating role of individual motivations. *SIGMIS-CPR*, 144-152

Zakon, R. H. (2010). Hobbes' Internet Timeline. Retrieved 07-01, 2010, from http://www.zakon.org/robert/internet/timeline/.

Zhang, J., M. S. Ackerman and L. A. Adamic (2007a). Expertise Networks in Online Communities: Structure and Algorithms. *WWW*, 221-230

Zhang, J., M. S. Ackerman, L. Adamic and K. K. Nam (2007b). QuME: a mechanism to support expertise finding in online help-seeking communities. *In Proceedings of the 20th annual ACM symposium on User interface software and technology*, Newport, Rhode Island, USA, 111-114