# Knowledge Activity Support System Based on Discussion Content

**TSUCHIDA, Takahiro**
Graduate School of
Information Science,
*tsuchida@nagoya-u.jp*

**OHIRA, Shigeki**
EcoTopia Science Institute
*ohira@nagoya-u.jp*

**NAGAO, Katashi**
Center for
Information Media Studies
*nagao@nuie.nagoya-u.ac.jp*

Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
011-81-52-789-5878

## Abstract

*We have developed a system to support knowledge activity. Knowledge activity is activity in which people continuously create ideas about themes and develop them into knowledge. First, the system generates minutes (linked to video and audio of discussions) and metadata (on significant discussions). Then, users tag the minutes in order to identify certain parts contemplate the discussion content and write notes derived from the minutes with the system for use as discussion content. The system converts ideas in notes, generated through ordinary tasks, into presentation material for subsequent discussion. It is possible to support knowledge activities by repeatedly using this system.*

## 1. Introduction

In knowledge activities, such as those in project planning in enterprises or research in universities, people continuously generate knowledge on various themes by performing common tasks such as information retrieval. We think knowledge activities will be improved by creating secondary content that is taken from knowledge generated in repeated cycles of knowledge activities.

Discussion plays an important role in knowledge activities. The knowledge generated by a closed community is dominated by the ideas held by that community. Since there are probably different points of view in other communities, people present and discuss knowledge generated by their own closed activities with people outside that community. We believe that it is possible to receive advice, to hear opinions and to review policy on knowledge activities through such discussion.

What we are concerned with here is that the content of an argument is forgotten with the passage of time. Consequentially, it is necessary to record the content in some way and to create hints on beneficial discussion. A lot of studies on systems that mechanically record the content of a discussion have been done in the past. Such studies include work on recording information not only as text information but also as audio-visual information and presentation material. By keeping minutes, the presenter can think about his or her own presentations, convey knowledge smoothly and efficiently, and improve the next presentation.

There has been little research on ways to utilize discussion content. However, there are close interrelationships between generating and utilizing it. Therefore, it is necessary to think not only of ways of generating discussion content but also ways of utilizing it.

In this paper, we first describe the features of knowledge activities and a knowledge activity support system (KASS) (Section 2). We then introduce discussion mining, a method that we developed as an environment in which discussion content can be created and browsed (Section 3). In Section 4, we describe the use of discussion contents created by discussion mining and a description of the discussion, reflection, investigation, preparation (DRIP) system. In Section 5, we describe some applications of the combination of discussion mining and DRIP systems. We then discuss the effectiveness of KASS.

## 2. Knowledge Activities

In our research, we focused on discussion, which is a knowledge activity. For example, a discussion is something like a seminar at a university, in which a presenter makes a presentation using knowledge that has been obtained from the presenter's own knowledge activity and participants ask questions about content they do not understand and express their opinions on the content they do understand. We think discussion has two roles: one is to disclose knowledge accumulated through previous knowledge activity to the public and the other is to share ideas with participants.

An important function of a discussion is to change the way participants conduct their knowledge activi-

ties. The participants may not have the same perspective as the presenter's on his or her knowledge activity. For this reason, opportunities for several people to meet and share a presenter's knowledge are important.

The participants in a seminar may have concerns and their own position, policy and practice on a certain knowledge activity. Therefore, it is very useful for the presenter to have a discussion with participants who have different ideas and knowledge because of the potential to get feedback via comments and opinions. For example, an investigation based on advice about relevant information can lead to further knowledge being acquired and confirmation of his or her activity's position in and benefit for society. Discussion thus influences normal research and knowledge accumulation. Furthermore, problems as well as new knowledge are discovered. More feedback can be acquired by further discussion. We believe that knowledge activities advance by combining discussion and normal research.

## 2.1. Predominant Features

Research into a system for creating minutes for the reuse of argument content has been conducted a number of times [1, 2]. Lee [3] proposed a method that records the participants' actions using cameras and microphones and then produces indexed minutes using auto-recognition technology. Chiu [4] integrated audio-visual information and information for presentation materials. We have developed a method to document, retrieve, and browse arguments that occur during a seminar as discussion contents. It is done by a method called discussion mining.

We conducted a questionnaire survey to 7 people from our laboratory. Each respondent made presentations about own research theme using the discussion mining for 10 months, then checked arguments in the presentations in order of date and answered questions about each arguments: whether he/she remembered the argument content, considered the argument important to perform common tasks and practically performed.

A questionnaire result shows that about 74% of arguments made 10 months ago are forgotten while most recent arguments are little forgotten; the content of an argument is forgotten with the passage of time. It also shows that about 21% of arguments well before 6 months, which respondents are felt important, are not reflected on common tasks. If not only arguments respondents remembered when the questionnaire was made but also arguments forgotten are included, the value may become bigger. We think they are forgotten because certain discussions are postponed because they are deemed future works or because the actual discussion topic is not brought to the attention of relevant people. The questionnaire also revealed that it is not sufficient to merely create minutes in order for knowledge activities to be properly performed.

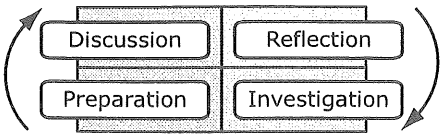


**Figure 1. DRIP Cycle**

Based on this point of view, we regard knowledge activities process as the cycle of four phases listed below.

- Discussion Phase
- Reflection Phase
- Investigation Phase
- Preparation Phase

In the discussion phase, we present the ideas and products presentation with material and discuss with participants. Arguments give birth to new advice and opinion because of understanding from several points of view. So as not to waste these desirable arguments, we then arrange arguments in the reflection phase. In the investigation phase, we create and record new ideas and products based on the past arguments. We create presentation material for next discussion with conversion from the recorded ideas. We get feedback again via discussion in a presentation using created material.

We consider that knowledge activities advance through circulating these four phases shown in Figure 1. Once this cycle rides the entire loop, it is not a cause for concern where the starting point of this cycle is. We define knowledge activities' process with this cycle "DRIP (Discussion-Reflection-Investigation-Preparation) cycle".

We think that it is important to associate the knowledge generated through various works with discussion as a trigger for the generation of more knowledge in the drip cycle. As a result, a person can understand what knowledge is generated by discussion, moreover, decrease the amount of discussion that is neglected by warning of the existence of discussion that is not related to the knowledge being generated.

Linking information connecting a discussion and the knowledge generated from the discussion can also lead to better presentations. A presenter can arrange the content of a presentation and the participants can better understand the context of the presentation. The participants are especially motivated to discuss the content of a presentation because they can make sure that their own remarks are reflected in the presenter's knowledge activity.

## 2.2. Knowledge Activity Support System

The purpose of this research is to develop a system to record discussion arguments as discussion con-

tent and to activate user knowledge activity using this discussion content. To accomplish this, we have developed a system called a knowledge activity support system. KASS contains two environments (Figure 2): the discussion mining environment in which discussion content is generated and browsed and the DRIP environment in which generated discussion content is used.

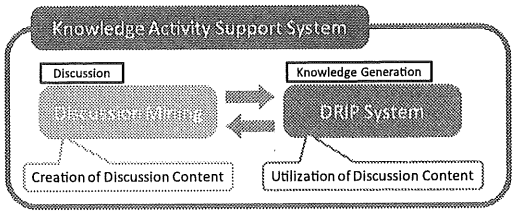


**Figure 2. Knowledge Activity Support System**

In the discussion mining environment correspond to the discussion phase in the drip cycle , reusable discussion content is generated by gathering real-world information from text, video, metadata, etc. on discussions that take place in face-to-face meetings and semi-automatically structuring this information. Furthermore, there is an interface for efficiently browsing discussion contents.

The DRIP system provides functions to arrange discussion content from their own viewpoints in the reflection phase, to record any newly generated ideas as notes and relationships with the discussion content and notes in the investigation phase. The system can also provide accumulated information about discussion content and notes for making presentation materials in the preparation phase.

## 3. Discussion Mining

We have developed a method called discussion mining [5], which semi-automatically generates reusable discussion content. Real-world information, such as text information, audio-visual information, and metadata, is recorded and structured as discussion content. Discussion mining consists of a discussion recorder, which is an environment in which discussion content is generated, and a discussion browser, which is an environment in which discussion content can be effectively browsed.

### 3.1. Discussion Recorder

We analyze meetings not only with natural language processing to support the comprehension of arguments in a discussion but also form diversifed

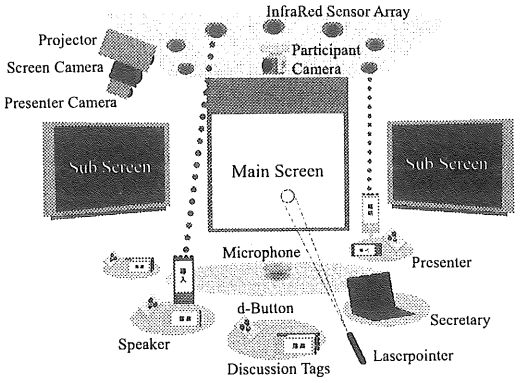


**Figure 3. Discussion Recorder**

perspectives using auditory and visual information in slides and other presentation content, and we use metadata to deal with the discussion content. The discussion recorder supports the creation of discussion content for face-to-face meetings, records the meeting environment with cameras and microphones, and writes meta-information that relates content in an XML database.

Most studies that provide technology for discussions and generating minutes have focused on automatic recognition techniques for auditory and visual data, such as meeting browsers [1]. We have developed a discussion recorder for generating the content from meetings in a more semantic manner. In discussion mining, human activity in the real world is recorded using two or more cameras and microphones. We target meetings that include a presenter, a secretary, and participants, and the presenter presents her or his agenda by using Microsoft PowerPoint. Figure 3 shows an image of the discussion room.

The presenter uses a browser-based interface to display slides and to change them. The information is recorded automatically. Participants in the meeting transmit their IDs and comment types using tag devices called discussion tags in order to properly structure the discussion. A secretary records any arguments using a browser-based interface. A record of the arguments in XML and MPEG-4 format is saved as discussion content in an XML database.

### 3.2. Discussion Browser

The information accumulated by the discussion recorder is presented as discussion content in the discussion browser[1]. Figure 4 shows the construction of the discussion browser.

This browser screen consists of (1) a video view, (2) an in-depth view, (3) a search table, and (4) a lay-

[1]The discussion browser is exhibited at http://dm.nagao.nuie.nagoya-u.ac.jp/

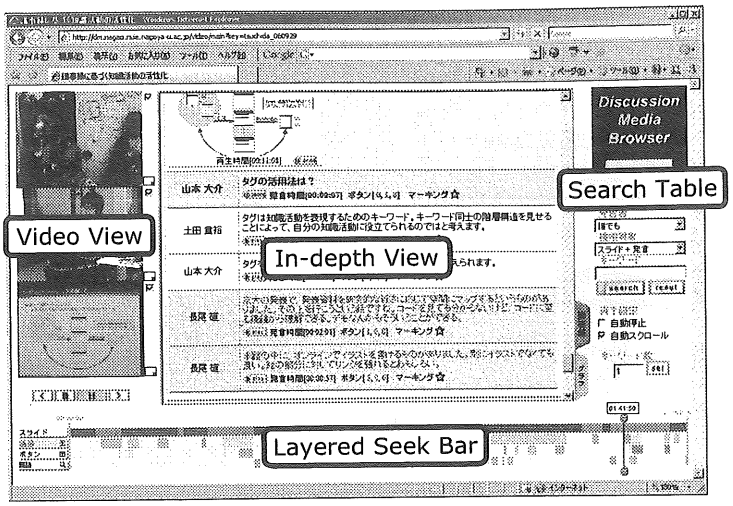


**Figure 4. Discussion Browser**

ered seek bar. The video view provides videos of the discussion, including the participants, presenter, and screen. The in-depth view consists of a text-based view, which display any discussion content that is text information, and a graph view, which displays the structure of the discussion. Users can toggle between these views. In the search table, three types of search query are available: speaker ID, the target of the search (either the contents of the slide or statement, or both), and keywords. The elements which compose discussion content are displayed in the layered seek bar. The discussion browser enables searching and browsing of the details of the discussion that correspond to user queries. For example, when a participant of a discussion wants to refer to certain important previous discussions, the participant will search for statements by using keywords or speakers' names and then browse the details of the statements in the search results.

## 4. DRIP System - Utilization of Discussion Content

The DRIP system is a server/client system. The client application has three main functions: arranging discussion content created by discussion mining, creating a memorandum of generated knowledge, and creating presentation materials using the information accumulated. The server manages data transferred from the client and conducts periodic auditing and notification of users' knowledge activity.

### 4.1. Tagging Discussion Content

When past discussion is not been dealt with, they are forgotten. As a result, the amount of forgotten discussion content increases. The DRIP system enables users to arrange discussion content so that it is less easily forgotten.

Figure 5 shows an example of the interface used when arranging discussion content with the DRIP system. There are various kinds of discussion. For example, important ones that affect a presenter's knowledge activity and less important ones that involve questions and answers. Users thus can mark statements evaluated important in the discussion content using this interface. The DRIP system can be used during and after discussion.

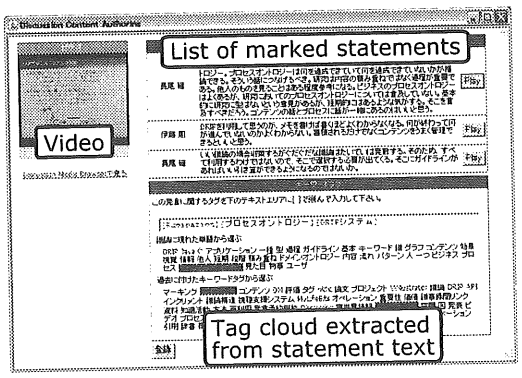


**Figure 5. Discussion Content Tagger**

The user can also tag the discussion content for searching. This is because it is not possible to tell from context alone why certain marked statements are important. Determining why something is important is difficult using automatic recognition techniques on audio and visual data. The DRIP system has two tagging methods: one for tagging new data and one for tagging a tag cloud containing past tags. The DRIP system prevents users overlooking discussion content

by warning the user that certain discussion content has not been arranged.

Marking and tagging discussion content enables the user to efficiently search for relevant arguments. A result on discussion content retrieved using the DRIP system is shown as a list of selected statements. The system redirects users to the discussion browser so that they can browse details of specific discussion content.

### 4.2. Note Writing - Derivations of Discussion Content

Users can note knowledge generated by work based on discussion content with the DRIP system. Figure 6 shows the interface for writing notes. The interface includes a discussion content viewer, a note editor, and linking information. When an idea occurs to the user while browsing discussion content, the user inputs the idea using the note editor, which is opened by clicking an icon in the discussion content viewer. Linking information, which explains the relationship between discussion content and notes, appears as an arrow between the discussion content viewer and the note editor. The user can also link discussion content and notes by dragging the note editor to other discussion content viewers.

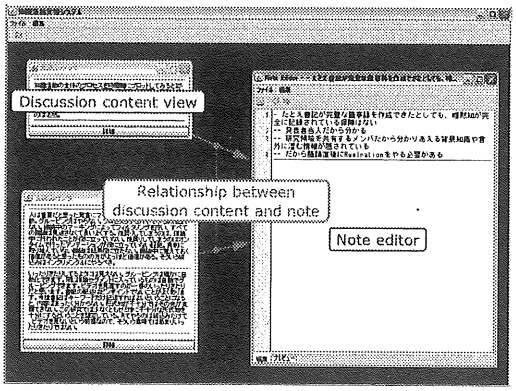


**Figure 6. Annotating Discussion Content with Personal Notes**

The server in the DRIP system manages all data displayed and input via the note editor. Thus, the DRIP system creates a synchronized environment across different PCs. The server regularly checks for discussion content that has not been annotated or viewed for a certain period of time. The server notifies the user of such content in order to decrease the amount of forgotten discussion.

### 4.3. Creating Presentation Materials for Next Discussion

Knowledge generated in work based on past discussion content is very useful for further discussion. This knowledge should be included in presentation material for future discussion. The DRIP system helps the user make presentation materials with an interface called presentation material creator, as shown in Figure 7. The presentation material creator has the following functions:

- Creation, deletion, reordering of slides
- Importing of accumulated notes into slides
- Conversion of slides into a PowerPoint document

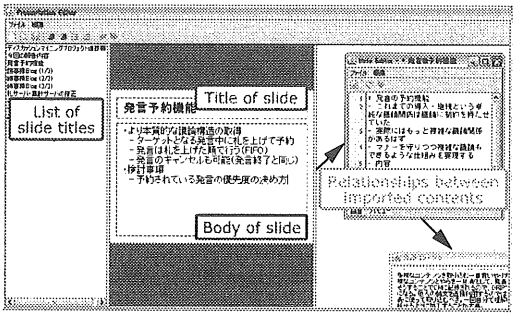


**Figure 7. Presentation Material Creator**

The user makes an outline of his or her presentation material by creating, deleting and reordering slides. The slides will be in a format such as Microsoft PowerPoint. Each slide, which contains a title, text body and a unique ID, is a derivative document of notes.

To include the accumulated knowledge in the presentation material, the user conducts an operation called "Import". For the import, the system retrieve notes based on keywords or dates. When the user searches using dates, a history of note edits starting from the query date is returned. It is thus possible to include a difference value in the slides of the user's knowledge activity from previous discussions. The contents of the retrieval results are inserted automatically into the text body editor by dragging and dropping. This function is used to make complete slides.

The presentation material creator converts the completed slides into a PowerPoint document and transfers information about importing the contents into slides to the server. The information is made available in the discussion recorder. It is useful to display such information on sub screens in order to help participants understand presentations.

## 5. Application of KASS

KASS provides functions enabling users to relate discussion content and notes and to create presentation

materials using discussion content and/or notes. Each time the system is used, the relationships between discussion content and notes will increase (Figure 8). We regard the relationships as contextual information on knowledge activities.

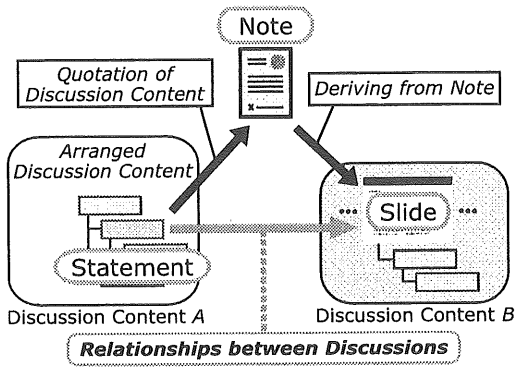


**Figure 8. Relationships between Discussions via Notes**

We believe it is possible to improve the understanding of the background to a user's knowledge activity by using an interface to efficiently browse the contextual information on the discussion browser. When the user browses information about his or her own activity, this knowledge activity will become more meaningful by determining its significance or assigning it to a particular activity.

Moreover, knowledge management can be supported by sharing the contextual information within a group. For instance, a person newly assigned to a project has little knowledge concerning previous approaches used in the project. In this case, contextual information on the knowledge activities of others may help the person to better understand these previous approaches.

Studies on systems that store and collate personal information quickly and flexibly have been done [6, 7]. Most studies have focused on storing content metadata, such as a title of an article, a URL, and creation dates. However, we believe that information on relationships between content is necessary for retrieving useful information.

There are two methods of acquiring information relating contents: one is an automatic recognition technique, such as natural language processing and image data processing, and the other is manual input. Because of the difficulty in interpreting semantic content by machine, the accuracy of the former method is not adequate. The problem with the manual input method is that it is expensive. However, we believe that relationships between content can be acquired easily using KASS due to tacit acquisitions based on users' natural activity.

## 6. Conclusions and Future Work

This paper describes KASS. Discussion mining, which is part of the system, generates reusable content linked to video and audio data of discussions and metadata. Users can then arrange and use the discussion content with the DRIP system. It is possible to perform knowledge activities by repeatedly using KASS.

Future research will concentrate on the items listed below.

- Evaluation of KASS
- Activation of discussion by using information accumulated when creating presentation material
- Implementation of a system that supports not only individual knowledge activities but also cooperative knowledge activities within a group

## References

[1] Schultz, T., Waibel, A., Bett, M., Metze, F., Pan, Y., Ries, K., Schaaf, T., Soltau, H, Martin, W., Yu, H. and Zechner, K: The ISL Meeting Room System, Proc. of the Workshop on Hands-Free Speech Communication (HSC-2001), 2001.

[2] Rowe, L.A. and Casalaina, V: Capturing Conference Presentations, IEEE MultiMedia, vol.13, no.4, pp.76–84, 2006.

[3] Lee, D., Erol, B., Graham, J., Hull, J., and Murata, N: Portable Meeting Recorder, ACM Multimedia 2002, pp.493–502, 2002.

[4] Chiu, P., Boreczky, J., Girgensohn, A., and Kimber, D: LiteMinutes: An Internet-based system for multimedia minutes, Proceedings of Tenth World Wide Web Conference (2001), pp.140–149, 2001.

[5] Nagao, K., Kaji, K., Yamamoto, D. and Tomobe, H: Discussion Mining: Annotation-Based Knowledge Discovery from Real World Activities, Proc. of the Fifth Pacific-Rim Conference on Multimedia (PCM 2004), pp.522–531, 2004.

[6] Fertig, S., Freeman, E., and Gelernter, D: Lifestreams: an alternative to the desktop metaphor, published in the ACM SIGCHI Conference on Human Factors in Computing Systems Conference Companion (CHI '96), pp.410–411, ACM Press, 1996.

[7] Gemmell, J., Bell, G., Lueder, R., Drucker and S., Wong, C.: MyLifeBits: fulfilling the Memex vision, ACM Multimedia, pp.235–238. ACM, 2002.