Sophia: WEB Browsing System for Assembly Minutes

Ryosuke Fujioka(*1) and Toshihiko Watanabe(*2)
(*1) Kobe Sogo Sokki Co. Ltd., 4-3-8 Kitagase-dori, Chuo-ku, Kobe, Hyogo, 650-0012, Japan,
ryo@sogosokki.co.jp
(*2) Osaka Electro-Communication University, 18-8 Hatsucho, Neyagawa, Osaka, 572-8530, Japan,
t-wata@isc.osakac.ac.jp

Abstract- Motivated by the recent trend toward the public information disclosure, an increasing number of local governments make their assembly minutes openly accessible on the WEB. At present, there are about 1,800 local governments in Japan and half of them provide the minutes browsing services on the WEB. “Sophia” is one of the WEB-based minutes browsing systems that are commercially available in Japan. “Sophia” is already deployed in about 100 local governments. “Sophia” is basically a text information database system equipped with functions that are useful for the viewers. In this paper, we introduce the WEB-minutes browsing system named “Sophia” which contributes to the increase of viewers compared with the case where the minutes are just made open by printed matter or simple text on the WEB.

1. INTRODUCTION

In this paper, we introduce a WEB-based assembly minutes browsing system as our application domain. Before further explanation, we give a brief explanation of our application domain.

Motivated by the recent trend toward the public information disclosure as well as the prevalence of the internet-related technologies, an increasing number of local governments make their assembly minutes openly accessible on the WEB. At present, there are about 1,800 local governments in Japan [1], and more than half of them provide the minutes browsing services on the WEB. Strong needs towards the public information disclosure have motivated an increasing number of local governments to make their assembly minutes openly accessible on the WEB.

The assembly activity takes an important role as a legislative organ deciding the will of the basic inhabitants of the local autonomy. As for the main role of the assembly activity, there are the drafting of the necessary policy and regulations, the investigation of the business that an executive organ performs, the reflection to the measure of the demand of inhabitants. For a record of the assembly activity, the making of the assembly minutes and long-time preservation and exhibition are obliged to by the Local Autonomy Law which means the law that each local government establishes.

Now, it comes to be demanded information disclosure in not only a national assembly but also a local assembly. In such a reason, it may be said that the exhibition of the assembly minutes is a part of important administrative services. We can read the assembly minutes in the information corner of the government office, a public hall, each branch, etc. As a matter of fact, inhabitants have rare chance to read assembly minutes for a hand. It is important to disclose information in the form that is familiar with inhabitants in assembly minutes to accomplish a purpose to make civic participation-shaped administration.

On the other hand, the development of the computer technology of these days has been remarkable. For example, the speedup of the CPU, the increase of storage capacity, price reduction of the hardware, the spread of internet broadband environment made a great contribution. In such situation, it is expected that the inhabitants want to read the minutes on the WEB.

The minutes of local assemblies are basically text-only information. Detailed and correct description is more weighed than readability and understandability. Unfortunately, it is not so much readable. We want to make it more readable and understandable by using WEB system. That’s why we developed the minutes browsing and search system named “Sophia” to satisfy the hope of such inhabitants’ demand.
There already exist several software products to answer those demands. “Sophia” [2] is one of the WEB-based minutes browsing systems that are commercially available in Japan. “Sophia” is developed by the first author and is already deployed in about 100 local governments. “Sophia” is basically a text information database system equipped with functions such as full-text search, etc.

Due to the nature of the minutes as the formal and authoritative record of the assembly activities, the minutes should be precise and correct in the first place. This is quite natural because they are used as the evidence of the speeches and words delivered by the assembly members and administrative staffs. However, from the viewpoint of the public, they are sometimes redundant in expression and difficult to read. Efforts to make them more attractive and readable are worth pursuing to promote the information disclosure and the public interest in the local government. Several directions are already explored for this purpose. The works in [3] uses multi-media technologies, and, in our previous work [4-5], the application of automatic summarization method is discussed.

“Sophia” has many functions in searching, browsing and printing. Furthermore, for a trial of the future extensions, we think about an automatic summarization, the use of the multimedia, collaborative filtering and evaluation technique of the usability to make it easy to use this system [6-8]. The document database systems have become widely used in various fields. As one of them, we introduce a WEB based assembly minutes browsing system named “Sophia”.

2. BASIC CONCEPTS

Currently, most of the local assembly make their Web Page and open it. Inhabitants can read their Web Page using PC and the Internet from each home. In Fig.1, we show the structure that inhabitants can browse Web Page with PC and the Internet.

In Fig.2, we show the outline of our system. A user can browse and search assembly minutes text by selecting following menu. The basic operation in the information retrieval is the search, or the pattern matching.
(1) Input search condition in the function menu
(2) Select the kind of the meetings
(3) Click the “Retrieve” button
(4) Hit list is showed as an output of result

Input search condition

Analyzer
- Name of speaker
- Date of meeting
- Specific keywords
- The kind of the minutes
- Combination mentioned above

Record of the Assembly minutes

Thesaurus

Output (Hit list)

Fig.2 System configuration

After the 4th step, the minutes’ text for the selected meetings is shown in the web page. In Fig.3 and Fig.4 we show CRT snapshots of input and output page in this system, respectively.

Fig.1 The structure of web browsing system

Fig.3 Snapshot of input page
In Fig.5, we show CRT snapshot of minutes’ text. In the left frame of this page, names and roles of speakers (e.g., the governor, the chief of general affairs, the member of the assembly and so on) are listed in order of time so that a user can browse who spoke what and when in the meeting.

Besides, a user can search specific remarks of assembly minutes. Several kinds of search conditions, e.g., the period of time (from and to) speakers, chairpersons, free keywords. In the case of keyword-based search, a user is expected to input keywords and press the “Retrieve” button. The user can browse the list for the search and select items in the list to browse the content text of the hit item so-called “snippets” of the hit items. Snippets are a piece of text inserted in web pages. It can include code to run at insertion variables (like selected text). In our system, snippets are the link to the pages (e.g. search result) using the highlight text. The sentence in which one or more keywords are included is partially shown, so that a user can find remarks of interest without switching from/to the hit list and hit items.

### 3. MAIN FUNCTIONS

There are three major functions in this system. One of them is a function to input various search conditions into this system. The second and third one is a display function and an output function. We show those functions in Table 1, 2 and 3, respectively.

In Table 1, we show search conditions. User can retrieve specific remarks to input this search condition or its combination into this system. For example, when viewer wants to search the remark of a certain political party (the group of assembly members), viewer can use the function of “Retrieval by group.” Using this function, viewer can confirm that they do not have a different political opinion among the group. This function is quite a unique but very important for minutes retrieving system. The function of “neighborhood retrieval” is to search texts by two keywords in the constraint where two keywords are within ten characters of one another, in either order in the sentence, such as “NEAR” operator adopted in Alta Vista Advanced Search engine [9].

<table>
<thead>
<tr>
<th>Table 1 Search conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functions</strong></td>
</tr>
<tr>
<td>Retrieval by keyword</td>
</tr>
<tr>
<td>Retrieval by speaker’s name</td>
</tr>
<tr>
<td>Retrieval by date</td>
</tr>
<tr>
<td>Retrieval by the meetings</td>
</tr>
<tr>
<td>Retrieval by group</td>
</tr>
<tr>
<td>Neighborhood retrieval</td>
</tr>
<tr>
<td>Fuzzy retrieval</td>
</tr>
</tbody>
</table>

In Table 2, we show display functions. In Fig.6, we show a CRT snapshot of the first line display. When viewer clicks the button of “the first line display” in the left frame, the first sentence is shown under the name of speaker. Viewer can easily grasp the outline of the meeting at the first glance. In Fig.7, we show a CRT snapshot of the comparative display of two speakers. In
the upper frame, a certain question by assembly member is shown, and in the lower frame, the answer to the question by a certain administrative staff. Using this function, we can easily understand a pair of question and answer in the minutes.

Table 2 Display functions

<table>
<thead>
<tr>
<th>Functions</th>
<th>Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>First line display</td>
<td>Only the first sentence of the speech is extracted by clicking a speaker's name.</td>
</tr>
<tr>
<td>Comparative display of two speakers</td>
<td>Using two frames, the words by an assembly member and an administrative staff are displayed comparatively.</td>
</tr>
<tr>
<td>Multimedia display</td>
<td>Display not only texts but also streaming data</td>
</tr>
</tbody>
</table>

In Table 3, we show output functions. For example, when viewer wants to know the whole activity of a certain assembly member in his/her active period, viewer can use the function of “Extracting Q and A.” Using this function, viewer can easily get a kind of book which question of the assembly member and its answer by the administrative stuff is written. When we set a search period with ten years, this book can be sometimes more than 100 pages.

“Sophia” is basically a text information database system equipped with several useful functions for the viewers. In addition to the basic search functions such as keyword search that are common to any text database systems, there are functions that are unique to the minutes browsing such as the first line extraction and the comparative display of speech texts of the assembly members.

These unique features are expected to help viewers quickly grasp the outline of the minutes. In general, the texts in the minutes tend to be of lengthy and redundant nature because the minutes are basically the transcript of the spoken words, and also because the precision and correctness precede the readability. Thus, such functions as mentioned above are quite useful and necessary for the viewers.

4. FUTURE WORKS

One of the purposes of the minutes on the WEB is to the arousal of interest and the promotion of the participation of the citizens into the local government activities. As stated above, though the first priority of the minutes is the archive of the facts, at the same time, it should be an appealing information source beyond the mere archive of facts to achieve the above-mentioned purpose. In this paper we introduced “Sophia” and its functions in detail. Reminder of this paper we describe functions of “Sophia” in our future works.

To make the minutes more appealing information source, the works in [6] use multi-media technologies. The use of visual images along the related texts makes the information more intuitively understandable. Such use of multimedia technologies is quite worth noting, consider-
ing the prevalence of the broadband internet services in these days. The example of streaming contents output from our developed system is shown in Fig.8. We realize the other function such as the synchronous show of speaker career and speech title.

![Fig.8 Streaming contents example of system output](image)

Furthermore, we focus on another approach that is mainly concerned with text information processing [7]. In our previous work [4-5], the use of automatic summarization method is presented to make the minutes more compact and condensed information so that the readers easily understand the main topics of the discussion taken place in the assembly. In short, the automatic summarization extracts interesting sentences from the minutes using some built-in preference evaluation functions.

At the last, we propose a method that detects inconsistencies between user interaction logs of a task and desired sequences for the task based on mouse click coordinate logs [8]. The method models two successive clicks as a vector and thus a sequence of operation in a user log as a sequence of vectors. To detect inconsistencies in user interactions and desired sequences, each vector form user logs is compared with user each vector form desired logs. The authors have developed a computer tool for logging and analyzing user interactions and desired sequences by the proposed method. The tool can be applied to evaluation of “Sophia” to make it easy using this system.

5. CONCLUSION

In this paper, we introduce the WEB-minutes browsing system named “Sophia.” “Sophia” is basically a text information database system equipped with functions. In general, the texts in the minutes tend to be of lengthy and redundant nature because the minutes are basically the transcript of the spoken words, and also because the precision and correctness precede the readability. “Sophia” has unique features to help viewers quickly grasp the outline of the minutes.

REFERENCES