

Digital Document Network System for Organizing Individual Knowledge

Kenji Matsunaga and Kyoko Yoshida

School of Network and Information, Senshu University, Kanagawa, Japan
{matunaga, k-yoshida}@isc.senshu-u.ac.jp

Abstract. eBooks and many other documents are being kept in recent years in personal storage in a variety of digital formatted documents with the amount of digitized documents expected to increase further. Intellectual tasks to organize this knowledge and add new meaning based on mass digital documents that will replace paper ones is needed. Thus a system that utilizes and supports the benefits of digital based on traditional work methods is required. This paper describes and evaluates the system we developed for building and searching multiple digital document networks with marks, links and memos, while reviewing useful methods to organize knowledge and add new insights.

Keywords: eBooks, Digital Document, Knowledge Management.

1 Introduction

Individuals have become able to possess many books in recent years as digital documents due to the spread of eBooks. Many documents are already maintained in an individual's storage as digital documents in a variety of formats, with the volume of digitized literature expected to jump dramatically in the future. Paper media books and documents have traditionally been a useful knowledge base for organizing an individual's knowledge, while societies have tackled larger intellectual endeavors using paper-based books and documents. Intellectual pursuits that organize knowledge and add new meaning can be obtained based on massive digital documents instead of paper books. Thus a system may be needed to utilize and support digital benefits based on traditional work methods.

iBook and Kindle are typical readers for reading eBooks. These have useful functions for bookmarks, markers and other intellectual organizing, yet they are primarily for reading one (1) book, not assuming the work to create knowledge based on multiple books. Software such as Evernote can also search keywords across the documents and list notes where the same tags are kept for tagged documents. But one cannot jump to the location where it is written in the document using a direct link because the link is not added. There is a lot of value in adding a link to eBooks, documents and other information with many pages compared to the Web where the volume of a single piece of information is comparatively small.

We proposed a system to link multiple eBooks and switch between them while reading [1]. We also proposed a system for evaluating methods of organizing knowledge and those useful for adding new knowledge based on interviews with researchers who handle many books and documents to build a digital document network [2]. This paper describes the results of the development and evaluation of this system.

2 Analysis of Target Users

We conducted a survey of literary research methods specifically for studies where many books and documents are handled from among a wide range of research fields. We interviewed two (2) master’s level students who are researching Japanese literature and ancient Japanese history on research methods and the utilization method of books. We questioned the information collection and management methods, specifics on their work using documents, and methods they would like to use with digitized documents and obtained responses in the interviews.

The results of the interviews, suggested that the ability to choose what is needed from books, documents and other information collected, focusing on the relationships between them, thereby adding new interpretations and knowledge is needed in literary research. Researchers require a system to support work such as when they write memos to digitized books and documents, create a new network relating these memos and gaining new knowledge on network formats already obtained.

3 System Overview

This paper proposes a digital document network architecture system supporting work to organize an individual’s knowledge. A network can be created by adding a memo to a sentence (hereafter marks & notations) extracted from a digital document using this system. Memos can also be added to a link. They can be searched using marked or linked memos with the generated network (Fig. 1).

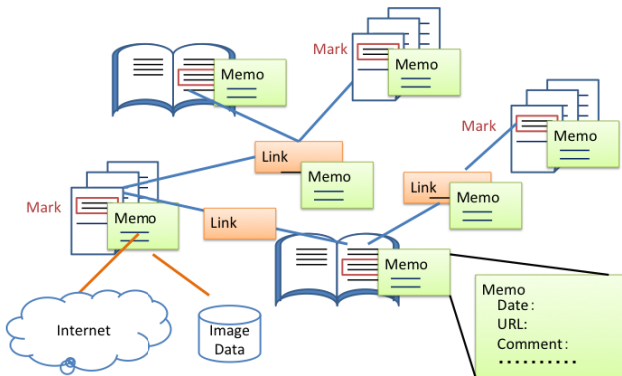


Fig. 1. Digital Document Network

3.1 System Functions

This system has “Create Network” and “Search Network” functions.

You can “Add Mark” and “Connect Links between Marks” when creating networks. Extract the marked digital document using a search keyword, etc. from an eBook or document, choose a sentence from a page to be marked from among them and place the mark there. You can add a memo to this mark. You can input these with a date, relevant items, people, comments and other “Key: Value” formats adding a memo. Choose two (2) or more items to link from among the marks, and set the links between them in order to relate links between marks. You can also add a memo to the link (Fig. 2).

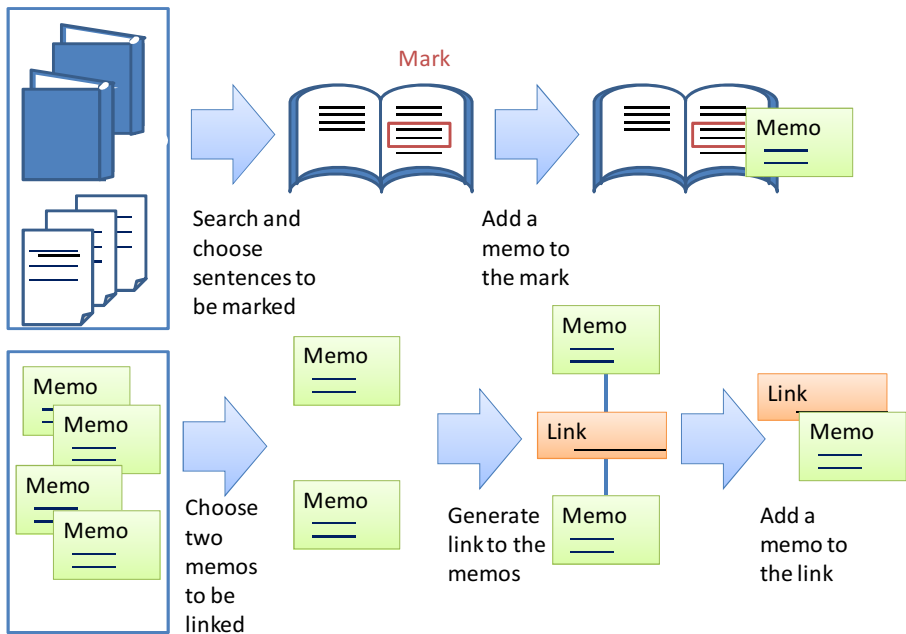


Fig. 2. Create Network Function

You can search the network created for eBooks and documents based on memo information with network search. First, choose the memo using a search keyword, etc. from the contents added to multiple memos created. The sentence marked with the memo added is displayed for marked memos. Two (2) or more sentences with links are displayed for linked memos. In terms of displayed items, you can view the location of marked sentences, as well as adjoining sentences and pages. You can also view the entire network image for all books and documents linked to selected memos (Fig. 3).

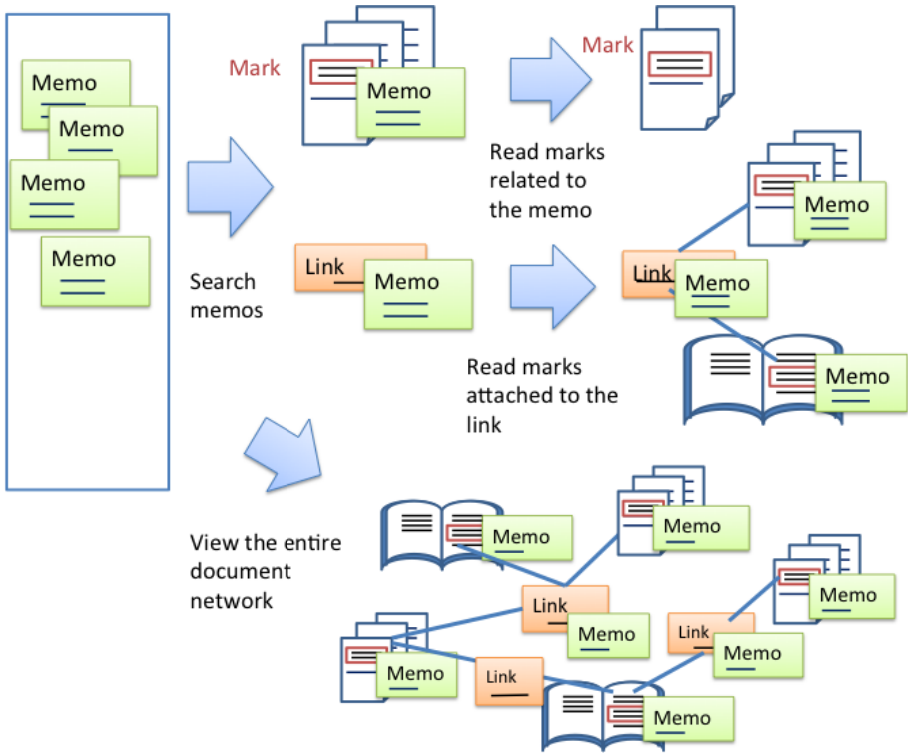


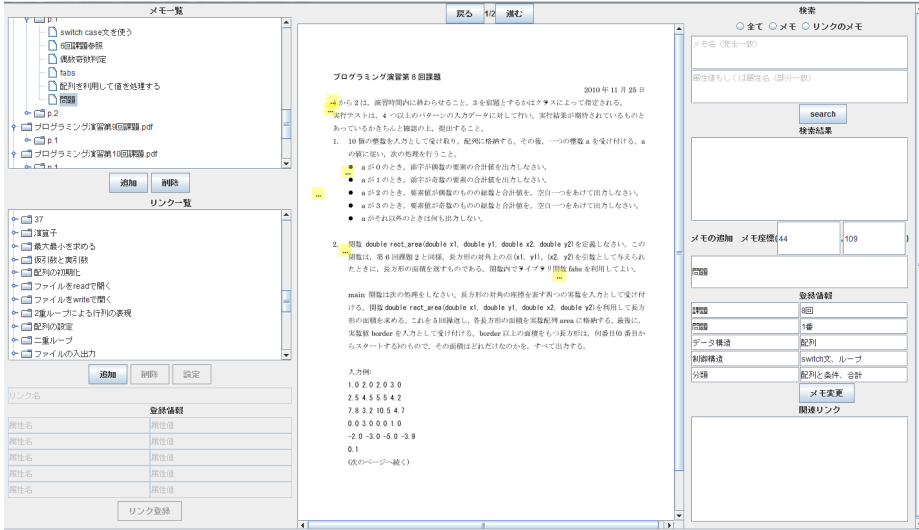
Fig. 3. Search Network Function

3.2 System Screen

Screen configurations were designed so that users could easily use system functions (Fig. 4).

A window where documents, pages and marks are arranged hierarchically is located at the left where pages can be followed while checking the marks. The document with marks embedded is shown in the center. Created, selected and searched marks and link memo contents are shown on the right. There is also a form included for searching memos. Information can be accessed from the route where the document is the origin and from the route where the mark and link are the origin via these interfaces.

Figure 4(a) shows the utility screen for choosing and reading a document, selecting the range to be marked and noting a memo for them. Figure 4(b) shows a utility screen for choosing two (2) marks, creating a link for them and noting a memo for them while a number of marks already exist.



(a)



(b)

Fig. 4. System Screen

3.3 System Configuration

Figure 5 shows the system configuration. The system input files handle eBook and PDF formats here, and are intended for paper media documents to be used by converting them to PDF via scanning, etc. The internal system uses unified PDFs. Marks, links, memos and other data users created with the system edit function are stored in the database. The search function searches information saved to the database.

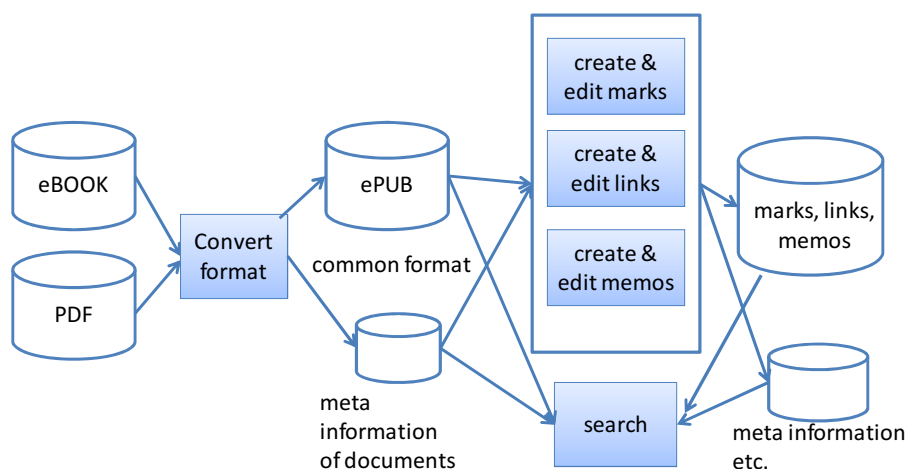


Fig. 5. The system configuration

4 System Evaluation

We conducted tests whereby we created a network for files saved to the system and searched information using the created network to evaluate it. The tests to create the network and to search information using the created network were not continuous, but conducted on separate days. Five (5) undergraduate seniors created one (1) network in 5 hours. Five (5) other undergraduate seniors conducted tests on searching information in 2 hours. A total of 11 files, five (5) being programming reference materials and six (6) being related practice problems, were the documents used. We observed and evaluated the behavior of subjects in the work process of this test (Fig. 6).

The task for the test to create the network is creating links between the 11 files for reference materials and practice problems as examples for these. Reference materials and practice problems were read, and marks and links were added to locations containing relevant items and other items to be referred. The result of this task was that around 150 marks and 70 links were created, and memos were created for each mark and link. Marks were drawn and memos were written smoothly here. However, only 20 links for these marks were drawn at first. Thus a specific example to adding links was instructing to search locations of references and examples deemed to be useful with practice problems. The tasks for tests on searching information was searching reference materials and practice problems to solve four(4) programming test problems with the created network. The knowledge required to solve these problems, not merely solving them, was extracted from reference materials and practice problems, and their locations were noted. Tests were conducted with groupware due to the long pre-trial time requirements for one (1) person to create a network for this system.

In the tests on creating the network, linking process was promoted by indicating tasks assigned. This may be because many marks and memos are often created with

paper documents, but adding links is not an action done daily. Other qualitative differences were also observed in the contents of memos by subjects because there are no clear standards for writing their contents. There were instances of writing styles used where their use in later searches was not insightful such as “Important” or “Will be on test”, and where the same memo contents were written many times over. A perfect example was also the usefulness in organizing document knowledge such as adding supplemental information not written in the document, excerpting important items and recording important relevant words. In addition, tests were conducted while students were talking to each other, and tests for creating networks had work to outline links in the head, so collaborative work may be more efficient in creating them than working alone.

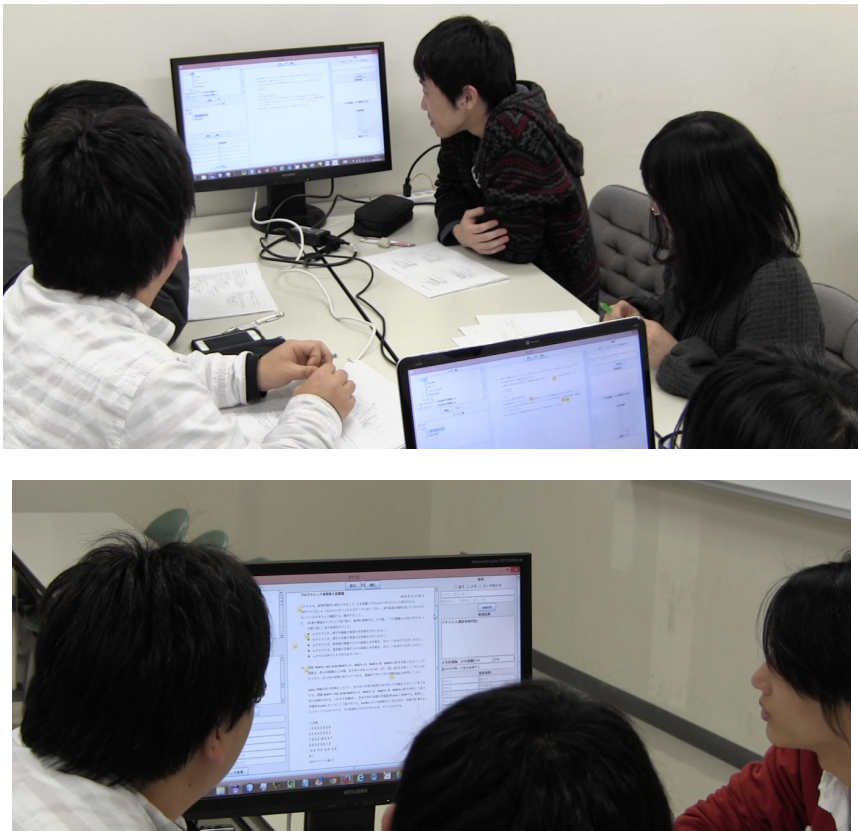


Fig. 6. User Test

Looking for the initial starting point to search was difficult in the tests for searching information. It was easy to search important information in the documents because they looked unconsciously for which documents had something written as there was a link added from the initial starting point. They located information by searching with memo contents used as tags when there was no link. For creating

networks with faster searches, differences appeared based on networks being created assuming what searches would come later when creating links and memos. One may be able to learn how to add higher value links by accumulating experience in searching.

Comments on the system made by the subjects were, for example, “Links and memos were effective in searching relevant information between multiple documents together” and “It was convenient for arriving at relevant information using memo contents”.

5 Conclusion

This paper described results in the development and evaluation of a system for creating a digital document network.

We will add functions for viewing the entire network structure for all books and documents relevant to selected memos and effective functions for group working. Some support may also be needed to create effective networks. The issues for this will be support for outlining links to various knowledge and preparing standards for creating networks.

References

1. Yoshikoshi, M., Matsunaga, K., Yoshida, K.: A Personal Document Network Building System for Digital Document Searches. In: Stephanidis, C. (ed.) Posters, HCII 2013, Part II. CCIS, vol. 374, pp. 458–461. Springer, Heidelberg (2013)
2. Matsunaga, K., Yoshida, K.: A Personal Document Network Building System for Organizing Knowledge. IPSJ SIG Technical Report 2013-IS-125(10), 1–5 (2013) (in Japanese)