

Manuscript Writing Support System for Researchers Based on Hypertext List of Their Achievements

S. Todoroki*

August 4, 2008

Abstract

A method is presented for shortening the time needed to prepare research papers and slides. The key is to enhance the searchability of related documents previously written by the author and stored in a PC. It is achieved by employing a hypertext list of the author's achievements, each item of which includes links to related files and web sites. Such a list was constructed with the aid of reference management software (BIBTEX) and a scripting language (Ruby). Automatic incorporation in the author's home page makes it easy to disseminate the author's achievements.

1 Introduction

When writing a manuscript on PC, you may often want to cut and paste something from a file that you made previously but cannot find. How can this problem be overcome? Some people may open all the folders and search them one after another, and others may use a full-text search software in their PCs such as Google Desktop.

Recently, I found a better way to deal with such a situation. The left window in Fig. 1 shows a list of my research papers displayed on a web browser. It is different from a conventional list because each item includes links to local files and web sites on the net. The clue to finding the target file is to remember that the file was used in a previous article or presentation. Although the idea is quite simple, I have found it to be an indispensable support system when publishing my research results. This paper outlines the approach.

Hypertext list of research achievements

- What is the benefit?
- How is it constructed?
- How does it promote information dissemination?

2 Useful functions

Promoting reuse of past documents

We can easily find target files by using our memory of our publications. This is because your memory is directly related to your list of research achievements and we researchers always focus on this list

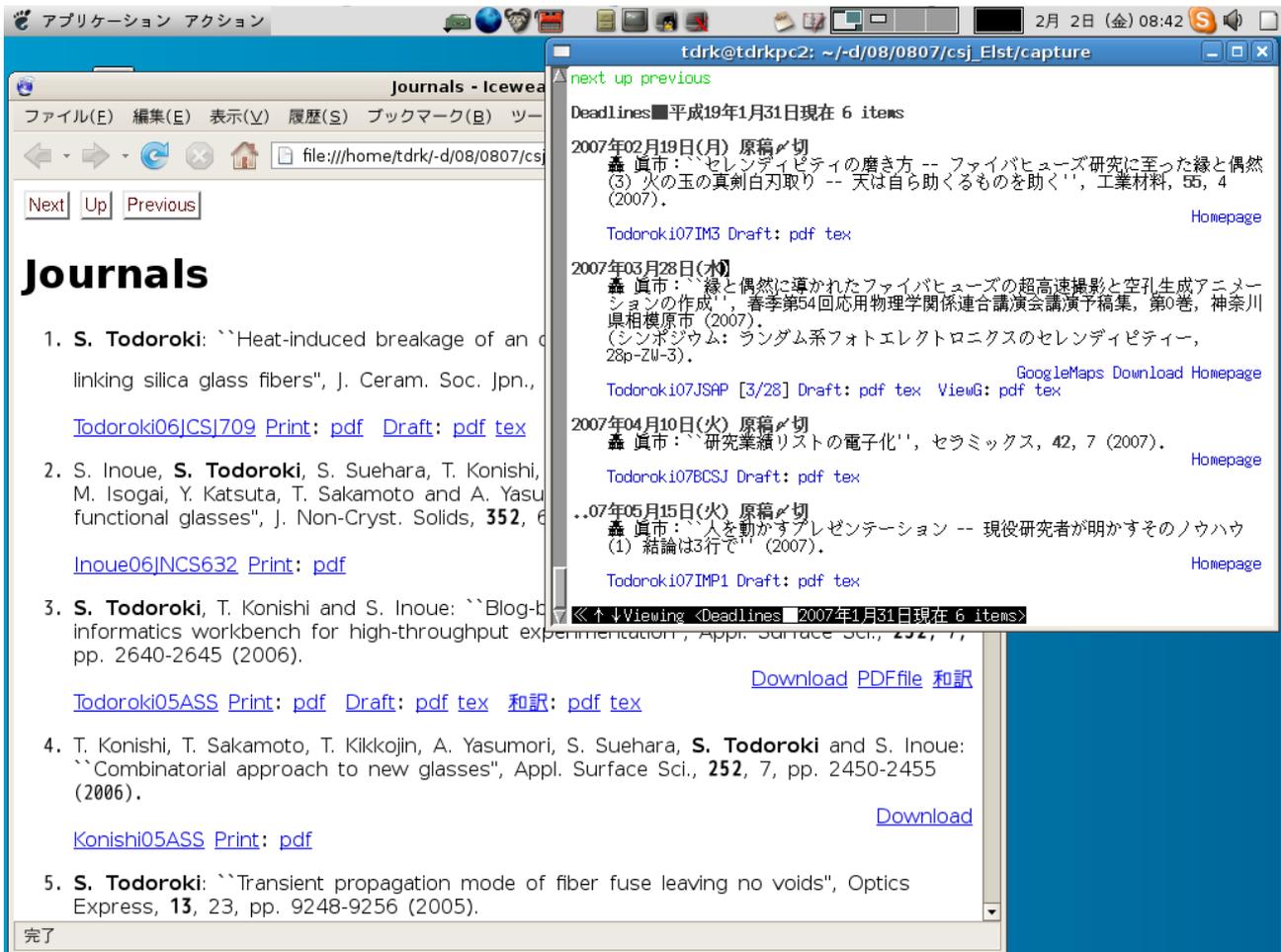


Figure 1: Typical pages of my hypertext list of research achievements. Each page is displayed on a different web browser. Left: List of research papers already published. Right: List of paper submission deadlines.

when applying for jobs, promotion and funding. Our primary mission is to increase the items in the list. Thus, most of the files we make are related to the list. We can remember our own list in detail.

Finding target files quickly means you have many chances to reuse your past documents. This reuse saves you time and provides extra time for polishing your document. To my surprise, this system also helps me find graphics files. In most cases, the clue to finding them is also my memory of my publications, i.e., "The figure I want was included in a certain paper or presentation". The list is also the best place to store links to photographs taken during conferences.

Impromptu presentation for visitors

Presentation slides are still useful after you have returned to your lab because they can be used to present your message visually to your guests. You can make a good impression on them if you can quickly find appropriate slides in your PC based on the visitor's interest. For this purpose, I prepared topic-specific hypertext lists including links to my presentation slides in order to be ready to make impromptu presentation in my lab.

Scheduling of document preparation

It is possible to make a list of future goals in a similar manner to that described above. The right window in Fig. 1 shows a list of deadlines in order of date on a certain day. I always open this page

when I start to write a manuscript, because it provides links to the files to be finished and related web pages. The latter is particularly useful when we prepare a photo-ready manuscript whose style is announced at a conference site, or make plans for a business trip by referring to an interactive map service such as Google Maps. Collecting related information in one place is a basic strategy for creative work.

Adding links to your list of achievements helps you

- to find files in your PC and to save time by reusing them
- to show appropriate slides promptly to your guests
- to manage manuscript and presentation deadlines

You may agree with me that such lists are useful, but who will make them for me given the time and effort that would involve? It is completely impractical unless the lists are compiled automatically. The next section outlines how to construct them.

3 Construction policy

Design

Three types of list are presented in the previous section: (1) lists organized by media (journals, international conferences, patents, etc...), (2) lists organized by subject, and (3) lists of things to be done (deadlines). It is impractical to construct these lists one by one manually. Thus, we must register all the items using a reference management software with unique identifiers, and retrieve the items needed to make one of these three types of list. All we need to do is to prepare recipes that contain a set of identifiers (see Fig. 2). To make a list of deadlines, we collect items containing future dates.

For example, the third item in the right window of Fig. 1 shows the deadline for this article (original version in Japanese[1]). This item is registered in the reference management software as shown in Fig. 3. [Todoroki07BCSJ](#) is its identifier under which its bibliographic data is stored. The **homepage** field contains the homepage address of the journal and the **draft** field the path of the manuscript file in my PC. These fields are the source information of the links in the hypertext list shown in Fig. 1.

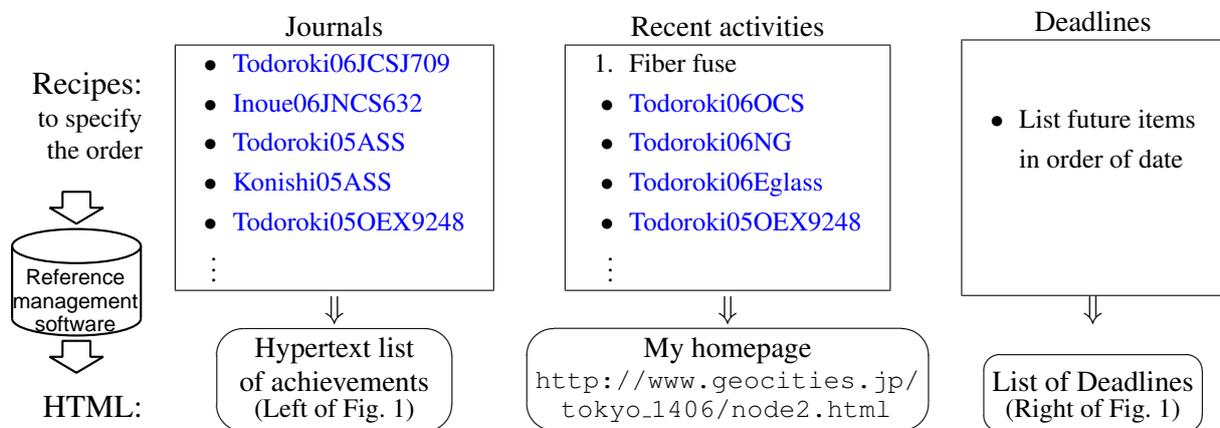


Figure 2: The procedure for making hypertext lists of achievements.

```
@Article{Todoroki07BCSJ,
  author = "S. Todoroki",
  title = "E-list of research papers",
  journal = "Ceramics",
  year = 2007,
  volume = 42,
  number = 7,
  homepage = "http://www.ceramic.or.jp/ihensyub/index_j.html",
  draft = HOME#"/-d/07/0704/csj_elst/csj_elst.pdf",
  annote = "070410: Deadline",
}
```

Figure 3: Example of an item registered in reference management software (The third item in the right window in Fig. 1.)

Required software

A search using any of the net search engine will show that many types of reference management software are available. What choices should you make when constructing this hypertext list of research achievements? My criteria are as follows.

Is the software compatible with your PC and software for manuscript writing? The software should be in harmony with your writing environment, not only with your operating system (OS; Windows, Mac, Linux etc...) but also your manuscript writing software (MS Office, OpenOffice, \LaTeX etc...), in order to avoid any reduction in performance.

Is it possible to export stored data for future migration? The lifetime of software is generally shorter than that of data, especially your research achievements. In the event that further development of the software is abandoned, you should have the skill needed to handle data stored in the software for future migration.

Is it possible to extend its functionality for customization? The style of the output list must be customizable. Specifically, you need to embed links in the output. If this is impossible, you have to modify the output by using other software. Such a cooperative process should be controlled by one of the programming languages. Some languages, called “glue language”, are very useful for this purpose, including Perl, Python, and Ruby.

Can you obtain the required output automatically? Since the final hypertext list is used for the daily tasks of searching files in your PC, the list should be always ready for updating at short notice. The updating process should be completed automatically just after you have typed the required data, such as bibliographic data about your article, the addresses of related web sites, and the deadline of your manuscript. Otherwise, you have to repeat this tiresome procedure every time and thus become impatient with the updating procedure.

A model case of mine

I always use $\text{\LaTeX}2\text{e}$ for writing manuscripts. Thus, $\text{BIB}\text{\TeX}$ is the most suitable reference management software. Since $\text{BIB}\text{\TeX}$'s output is not HTML, $\text{\LaTeX}2\text{HTML}$ is used to obtain the final HTML output. As shown in Fig. 4, these two types of software are combined by Ruby script, which reads a recipe, requests $\text{BIB}\text{\TeX}$ to obtain a required list, and converts it to HTML. However, you have to modify the format of $\text{BIB}\text{\TeX}$ output beforehand in order to embed links in it.

All these software are available free of charge. I believe \TeX -related software can be expected to survive for more than a decade since user and developer populations are sufficiently large. Since the only user of this customized hypertext list is the list's owner, I do not think that in the near future I need to abandon the present system and migrate to a new one. In fact, 20 years has passed since I

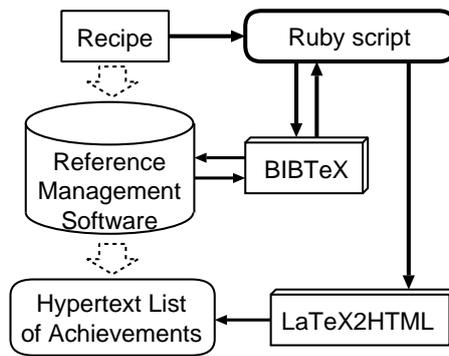


Figure 4: Data flow in the procedure for compiling a hypertext list of research achievements.

began to use \LaTeX . However, there may be a better choice for those who are willing to construct their own system.

Procedure for making hypertext list of research achievements

- Register all your research achievements using reference management software.
- Configure the software to obtain lists with links to related resources.
- Generate lists according to recipes containing the order of items.

My hypertext list originated from a set of scripts to simplify the procedure for uploading the list of research achievements to my homepage. After I came up with the idea of embedding links in each item, the list grew into the present style and has benefited my homepage. Some of these benefits are presented below.

4 Merits of synchronizing with web site

Making search engines your spokesman

Once you have completed a system for making a list of your latest achievements in HTML, it is easy to publish it on your homepage. Then, it is easy for people on the net to access your latest information through search engines. However, you should note that the frequency with which a search engine will visit your site is proportional to the frequency with which you update it. Thus, it is better to update your site as soon as you find something to be renewed. Otherwise, it takes long time for search engines to retrieve your latest information from your site.

Here I report an episode that showed me the prowess of search engines. Just after I had an article published in an academic journal in English[2], I uploaded a Japanese translation on my home page for the convenience of domestic readers. I did not expect a quick reaction, but two weeks later, I found a blog article introducing the Japanese document written by a graduate student I had never met. At that time, I had synchronized the list and my homepage with a cycle of couple of days. Therefore, the frequency with which the search engine visited my site must have been at least every 10 days. In addition, the blog article attracted further readers to my article.

Putting forward recommending articles

We often see lists of research achievements on researchers' homepages, but most of them provide only bibliographic data. They may be adequate as a target for search engines, but completely inadequate for visitors who are guided by search engines. It is surely useful to provide them with links to the

contents, such as specific pages of online journals and manuscript files on your site*. In addition, they appreciate seeing the author's choice of articles organized by subject.

I opened some web pages edited using this approach, and the address is shown at the bottom center of Fig. 2. Sometimes I hear comments stating that this page is interesting. However, it requires minimal labor, for example editing recipes, to update the page. The rest of the works is done automatically including the generation of pages in English and Japanese.

Publishing paper list on your area of interest

Nowadays, most of the research articles we get are PDF files rather than papers. To find required documents among them, I made another hypertext list of these articles in the same way. I have also published this list on the net (see http://www.geocities.jp/tokyo_1406/node3.html) for the convenience of researchers in my field.

Merits of synchronizing hypertext list and homepage

- Enhanced access via search engines if your site is updated frequently
- Attractive contents with minimal management
- Providing expert knowledge source for colleagues

5 Conclusion

You may feel that this article is too geeky for ordinary researchers, but the ultimate purpose of my system is to enhance the quality of research publications, as pointed out in Section 2. The thoughtful delivery of information attracts many people even after it is out of your control. Furthermore, it sometimes has serendipitous consequences[3]. I hope this article will enable readers to recognize my motivation and try to find their own approach to qualified presentation.

For readers in their twenties and thirties, I recommend mastering a programming language that assists daily writing tasks. At present, you may not need to manage your research achievements, but the quantity will increase steadily and move out of your control before you notice it. It is important to learn one programming language thoroughly rather than many languages incompletely. Then, you will find many daily tasks can be automated. My recommendation? Choose Ruby[4].

Never be afraid of a fighter who knows thousands of arts. Be afraid of the fighter who is skilled in one.

—A proverb from Chinese martial arts[5]

References

- [1] S. Todoroki: “E-list of research papers — a manuscript writing support system”, *Ceramics Japan*, **42**, 7, pp. 520–524 (2007) in Japanese. <http://pubman.mpdl.mpg.de/pubman/item/escidoc:33096>
- [2] S. Todoroki, T. Konishi and S. Inoue: “Blog-based research notebook: personal informatics workbench for high-throughput experimentation”, *Appl. Surface Sci.*, **252**, 7, pp. 2640–2645 (2006).
- [3] S. Todoroki: “Two serendipitous episodes — How I embarked on fiber fuse research” (2007) translated from articles in Japanese. <http://pubman.mpdl.mpg.de/pubman/item/escidoc:33124>

*It should be noted that few publishers allow the author to publish the final version of pdf files on his/her homepage.

- [4] C. Pine: “Learn to Program (Pragmatic Programmers)”, Pragmatic Bookshelf (2006).
- [5] R. Matsuda and Y. Fujiwara: “Kenji (manga)”, Vol. 21, Shogakkan (1992). (in Japanese, see Wikipedia for details).