

# *Collection of Conference Proceedings and Improving Access to the Full Text of Proceedings*

Misa Hayakawa, Shun Nagaya, Mayuki Gonda, Takeyasu Fukazawa,  
Minoru Yonezawa, and Keizo Itabashi  
*Japan Atomic Energy Agency, Japan*

## **Abstract**

Conference Proceedings are “grey literature” due to the fact that they are not made commercially available frequently. While many Proceedings are published on the Internet, there are specific issues that can affect access, such as changes in the URLs. This paper introduces the case of the Japan Atomic Energy Agency (JAEA) library as an example of efforts to improve access of Proceedings using the Internet. JAEA Library uses the Internet to make available presentations by JAEA researchers. The paper notes that the conference secretariats tend to be temporary bodies and the links to conference websites are not permanent. The paper reports our investigation into these problems, and we introduce a new approach to provide access to these Proceedings.

## **1. Introduction**

### **1.1 Conference Proceedings, a type of “grey literature”**

Conference Proceedings are records of the papers and data which researchers submit for a specific theme at a conference or meeting. Proceedings contain information on the latest research trends, therefore, it is important information resources in the fields of science and technology.

These Proceedings papers are characterized by a shorter period between the writing and publication than the journal papers. In recent years, the period between writing and publication has been shortened even further by the Internet [1]. Due to this, the importance of Proceedings is growing.

Many researchers assume that the reliability of Proceedings is lower than journal articles. Nonetheless, many researchers use Proceedings as (1) points of reference in developing research topics, and (2) to investigate the latest research trends [2].

However, Proceedings are often distributed to the participants at a conference, and are often not commercially available. It is difficult for libraries to ensure a complete collection of Proceedings.

Proceedings have traditionally been published as conventional books. However, forms of publication have changed and many Proceedings are published as digital media, such as CD-ROM or Flash Memory, or are published only on the Internet.

Proceedings on the Internet may seem like an easy to collect, but there are in fact many challenges in terms of collection and use. For example, the websites on which Proceedings are usually published are owned and operated independently, and cross-searching these sites is difficult. Additionally, there is a problem of visibility as Proceedings on the Internet cannot be accessed when users do not know if the sites exist [3]. Furthermore, conference secretariats are often temporary bodies and the links to conference websites are not permanent. Changes to URLs happen frequently; a specific challenge relating to access to information on the Internet.

Proceedings are, therefore, still “grey literature” because they are difficult to search and use.

### **1.2 Purpose of this study**

This paper discusses ways to improve access of Proceedings especially on the Internet.

The forms of publishing Proceedings are changing, and number of Proceedings available on the Internet is increasing. They are important as a means of providing the full text of papers to researchers. However, Proceedings published on the Internet cannot be accessed if their existence is not known, and currently it is difficult to search and use Proceedings papers.

We would like to introduce the case of the JAEA library as an example of improving access of full text Proceedings on the Internet.

We also address how the many sources of information about a meeting such as the conference website and papers on institutional repositories can be used to improve the access of Proceedings.

We introduce a new approach in providing conference information and linking to the full text Proceedings on the Internet.

## 2. Publication forms of Proceedings and library collections

### 2.1 Survey of the forms of Proceedings publications

We begin with an survey of the current forms of publishing Proceedings.

Fukazawa conducted a survey on the various forms of publishing Proceedings in 1982 and 1992. The survey classified Proceedings by publication forms such as books, journals, reports, and others [4]. 64% of Proceedings were published in journals or books.

We conducted our own survey to ascertain whether there had been any changes in the publication forms. We examined the Proceedings of international conferences in which researchers from the Japan Atomic Energy Agency (JAEA) participated in FY2010 (212 conferences). JAEA researchers mainly participate in conferences in the field of nuclear science.

We classified Proceedings forms as books, journals, reports, and other publications (Figure 1). Then, we classified books by mediums such as book form and CD-ROM (Figure 2).

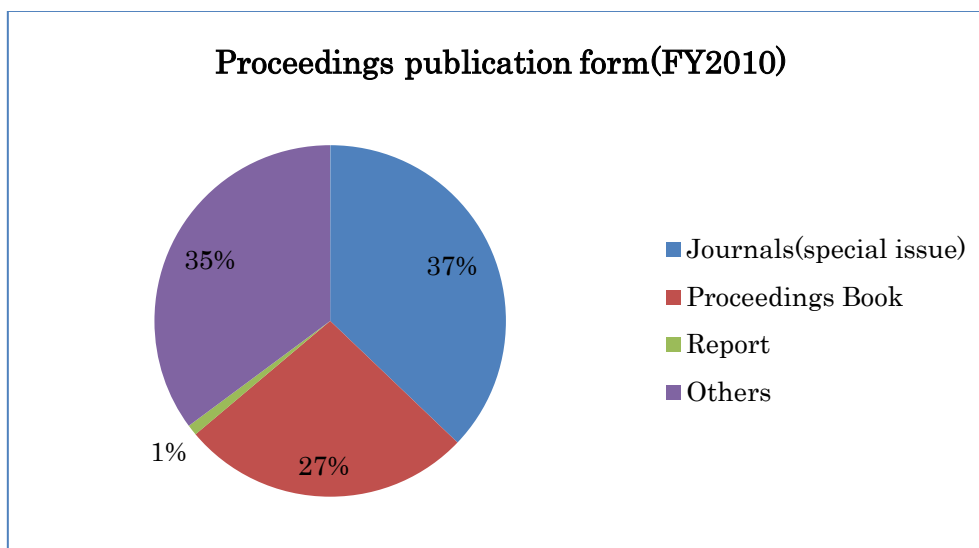


Figure 1: Proceedings publication form

We found that Proceedings are mainly published as journals (in special issues or as conference series). Next, Proceedings were commonly published in books. This is similar to the findings of Fukazawa's survey.

Incidentally, the “others” category includes cases where there was no form of publication of Proceedings (where, for example, only an abstract or program book were made available), and where there was an unknown publication form.

Proceedings published in journals are often commercially available, making them easy to access.

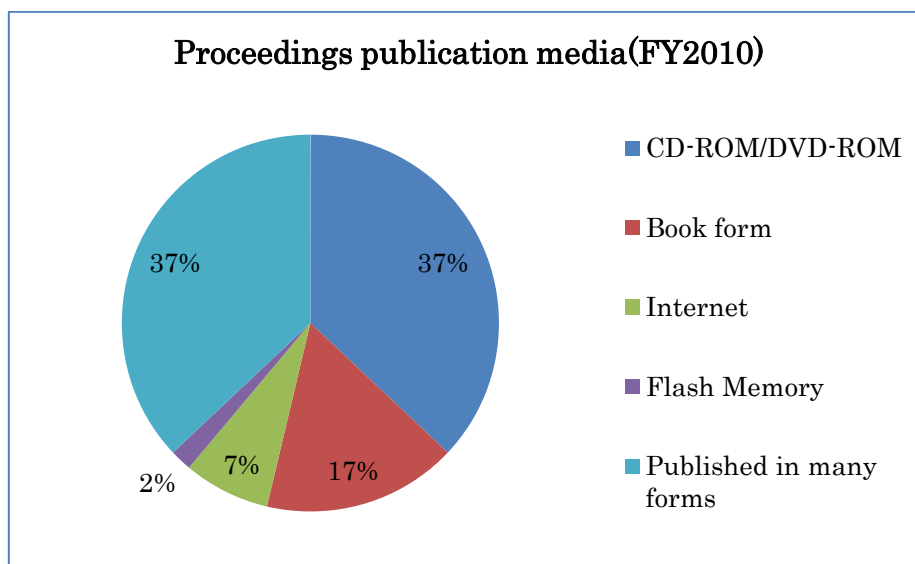


Figure 2: Proceedings publication media

We classified Proceedings by mediums, such as traditional books, CD-ROM/DVD-ROM, Internet sites, and Flash Memory. Proceedings published in two or more forms were classified as “Published in many forms” and duplication was not counted.

In all, 17% of the Proceedings were published as books, much less than the other forms. In comparison, 37% of the Proceedings were published as CD/DVD-ROMs.

Additionally, this survey shows many Proceedings are published more than one publication form. Many of these cases include Proceedings published on the Internet. Proceedings books or CDs were distributed to conference participants, and after the conference they were published on the Internet where they were available to buy. Published in a variety of forms in this way improves the availability of Proceedings.

Only 7% of the Proceedings were published solely on the Internet. In some cases, this publication form was open to the public, while some were only made available to the conference participants. This was, therefore, not necessarily a widely accessible form of providing access to the full text.

A few conference Proceedings were published as flash memory, although we do not dwell on this form of publication in this paper. Flash memory Proceedings were often distributed to conference participants at the conference venue, and pose a particular challenge in terms of collecting Proceedings by libraries. Additionally, it is difficult to preserve flash memory readability because they deteriorate and standards change. Where Proceedings published as flash memory do not conform to a meta-data system, it is also necessary to consider how Proceedings are made available by libraries. JAEA Library currently does provide a flash memory. However, to collect and provide flash memory still remains a challenging issue.

Proceedings published as CD-ROMs present similar problems to those associated with flash memory. For example, a Proceeding’s CD-ROM published in 1996 created with Windows 95. CD-ROM also cannot be read when the operating systems or standards

have changed. It is necessary to convert CD-ROM / DVD-ROM to up-to-date forms or to print.

## 2.2 Collection of Proceedings at the Library

We surveyed the library's collection of papers presented by JAEA researchers in FY2010 to confirm access of the Proceedings (not including those published as journals or reports) in libraries. We inspected the collections at the JAEA Library, the National Diet Library (NDL) [5], and CiNii Books [6] (to look for information about books and journals held by other university libraries in Japan). Additionally, we also investigated the availability of Proceedings on the Internet (Figure 3).

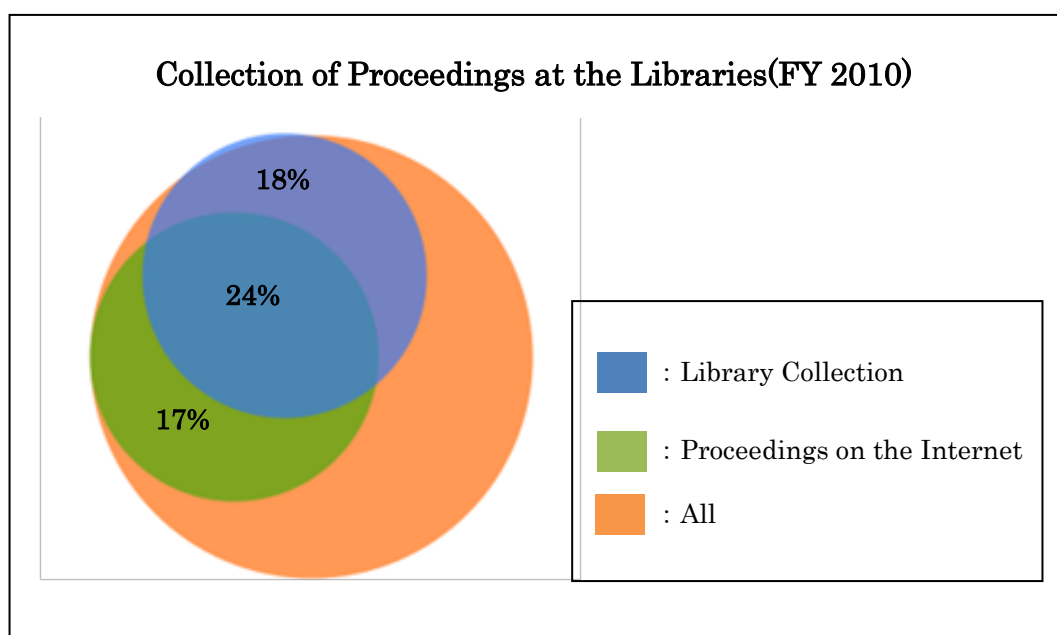


Figure 3: Collection of Proceedings at Libraries

The JAEA Library is one of the largest nuclear information centers in Japan, and we were collecting information on Proceedings specifically related to the field of nuclear science. We collected Proceedings from a variety of sources, including conference websites and from the researchers themselves.

However, we found that the JAEA Library together with the other libraries could provide only 42% of the Proceedings. As already mentioned, 17% of the Proceedings appear only on the Internet. This shows the difficulty of making Proceedings available in library collections. On the other hand, 41% of Proceedings were available as an Internet version. Providing links to the Proceedings on the Internet as well as the library's collections is important to improve access to Proceedings.

## 3. Management & providing links at JAEA

### 3.1 Providing links to Proceedings on the Internet within JAEA

As mentioned above, Proceedings on the Internet is increasing. Providing links to Internet Proceedings will improve access of Proceedings. We also advocated the management and provision of a system of the R&D results in JAEA as well as the links to the full text of the Proceedings.

We manage the publication and presentation of information by JAEA researchers at the JAEA Library. At JAEA, researchers are requested to submit bibliographic data (title, author, journal/meeting name, etc.) via web forms before they undertake

presentations or submit articles. Through this procedure, the library staff manage the information and produce an authority file of author names, conference names, and journal/book names.

The ability of this form of unified management by the library is particularly important where conference names have various notations.

Once the information is collated, we disseminate it via the Internet using the JAEA Originated Papers Searching System (JOPSS) [7].

Additionally, since July 2011 we have provided hyperlinks on JOPSS for the full-text versions of articles submitted by JAEA researchers. Providing links via JOPSS allows direct access to the Proceedings on the Internet.

The registration of the DOI/URLs is performed every month. Information is sent in the form of a hyperlink from a JAEA R&D result management system. The search and input of the DOI are then carried out using the Web system (Figure 4).

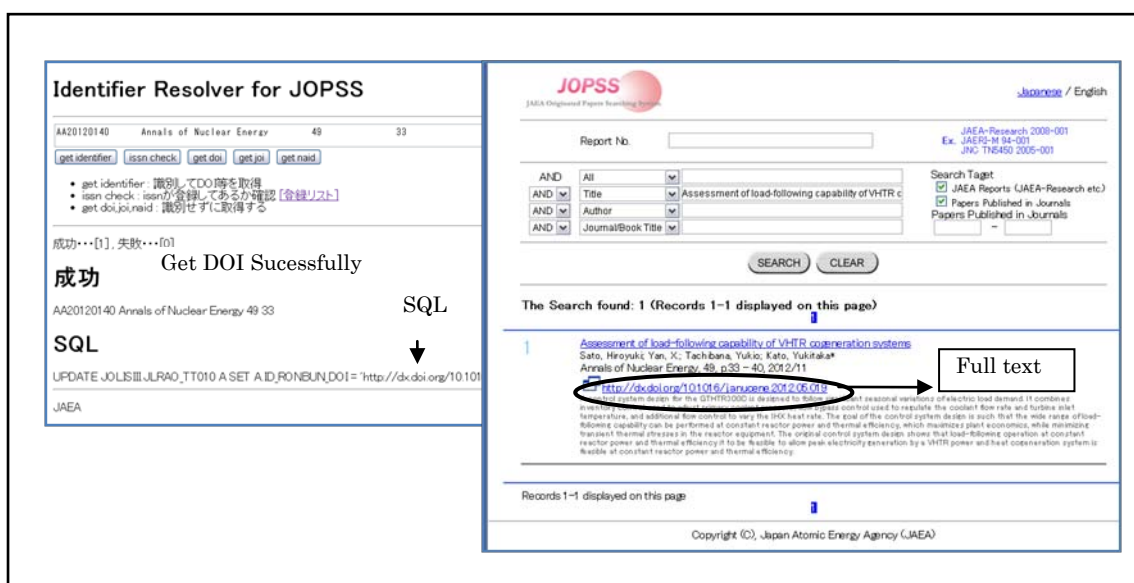


Figure 4: Providing hyperlinks to full text

We currently provide access to 14,459 papers from conference. Among them, 3,137 are hyperlinked to the full-text version or abstract (as of October 2012). Past paper/presentation data will also include DOIs/ URLs. Additionally, conference information (conference name, venue, and date) also provide from JOPSS since September 2012.

However, few Proceedings papers published on the Internet have DOIs (except for those published as journal papers). For example, Proceedings papers published on conference websites almost never have DOIs. In these cases, we provided a URL for the paper (such as a PDF etc.). Additionally, even when full-text versions could not be accessed from the Internet, many conference websites provided abstracts of conference papers. We therefore provided links to the abstracts of conference papers when full-text versions of the Proceedings could not be accessed from the Internet.

### 3.2 Surveys of Proceedings on Internet/Conference websites

In recent years, many conferences have had dedicated websites which provide conference information (dates, venues, programs, abstracts) to the public and the conference participants. However, this information faces a challenge that is specific to the information on the Internet, i.e., the possibility that the access links to the URLs

are either changed or the websites are not permanent.

We also, therefore, investigated access of the conference websites and the contents of websites for the conferences JAEA researchers participated in FY2010.

Access of conference websites with the URL registered before the meeting (Figure 5).

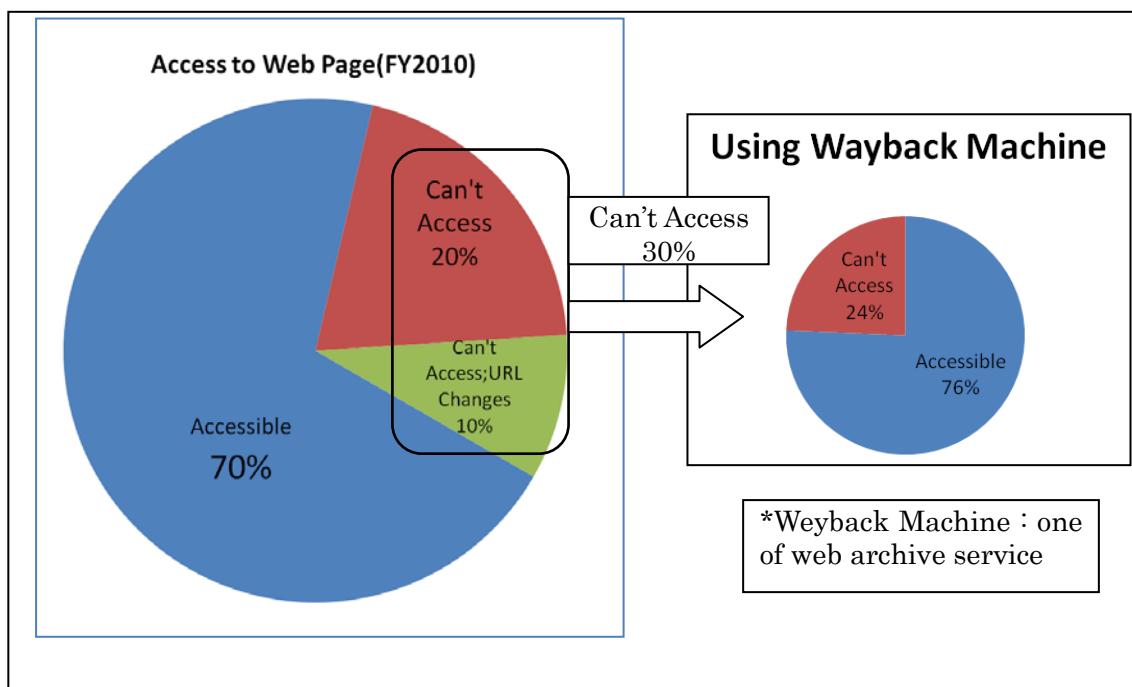


Figure 5: Access to conference webpage

Overall, 30% of the conference websites can no longer be accessed within 3 years. In all, 20% of the websites are no longer accessible because the website no longer exists, while 10% of the websites are not accessible because the URL has been changed. But 76% of these websites can access when use the web archive service.

Our investigation of website content showed that 10% of the websites provided the full text of the Proceedings, while 23% provided the abstracts. Some websites also provided presentation files and movie files, etc.

Some Proceedings available on the Internet were published on conference websites, while some were published on websites dealing with Proceedings in specific fields such as JACoW [8]. Moreover, some Proceedings available on the Internet were published on publisher websites.

Proceedings available on conference websites were published in a variety of formats, such as PDFs of each paper or PDFs of the whole Proceedings combined. Most of them were open to the public for no charge. Some, however, which were accessible only to conference participants. It required the ID and password to access details. These accounted for very few of the Proceedings, but did make it difficult to obtain full-text versions of the papers.

Proceedings made available by publishers generally have DOIs, making them more accessible than other forms.

### 3.3 A new approach to providing Proceedings information

A large amount of conference information and a great number of Proceedings are available to the public on the Internet. The links to Proceedings and conference information on the Internet as well as in the library's collection is useful to improve

access to the Proceedings. We, therefore, developed a prototype webpage that provides conference information, library collections of Proceedings, and Proceedings availability on the Internet using Google Calendar. Google Calendar is a free online calendar, and it is useful for information sharing and possible to integrate with other systems (Figure 6).

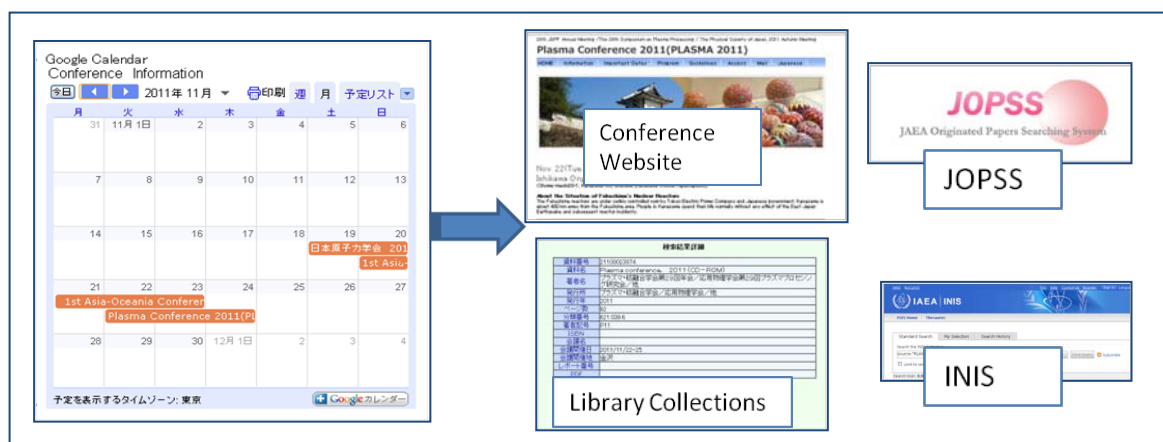


Figure 6: Google Calendar to Internet sources

One of our initial difficulties was in providing conference information where the websites had been deleted within a few years of the conference. We found that it was possible to solve this problem by using the web archives that has developed in recent years. The Web archives collect Internet web pages, archives them, and makes them available to the public. We used the Internet Archive's Wayback Machine [9] to provide conference information from the deleted conference websites. Wayback Machine is a major Internet archive, offering access to Web pages from 1996. As mentioned above, web archives will improve access to deleted conference web pages.

Additionally, we use library collections information from the JAEA Library/NDL/CiNii Books. This covers most of the information on Proceedings in Japan.

Some of the Proceedings papers were included in institutional repositories by the authors. In recent years, institutional repositories have become popular, and some institutional repositories include Proceedings not available anywhere else. For this reason we also provided the hyperlink to the search results of Japanese Institutional Repositories Online(JAIRO) as a way of increasing the availability of conference Proceedings papers. JAIRO includes 11,888 of presentation files (2010/10).

We also provide hyperlinks to the search results of the International Nuclear Information System (INIS), which includes information from conference Proceedings in the fields of nuclear science and technology. INIS hosts the world's largest collection of published information on nuclear science and technology [10], and is appropriate for conference information relating to the participation of JAEA researchers.

We have registered conference information from FY2010.

This conference information on Google calendar is not yet open to the public, but we want to improve the resource and make it publicly available.

#### 4. Summary

There are a variety of forms of publishing Proceedings and the number of Proceedings available on the Internet is increasing. Using links to Proceedings on the Internet as well as library collections can improve the access of the Proceedings. However, websites are not permanent and changes in URLs are frequent. The



URLs of Proceedings on the Internet, especially those on conference websites, are particularly susceptible to change. When we provide links to Proceedings, we need to, therefore, undertake periodic checks on the links. Web archive also improve access to web pages when URL changes or deleted pages.

Additionally, some of the Proceedings available on the Internet are not open to the public, making it difficult to access the full text. These accounted for very few of the Proceedings, but we have to consider how to provide these Proceedings.

In recent years, some of the Proceedings papers were included in institutional repositories by the authors. Institutional repositories improve access to full-text Proceedings that are not available anywhere else.

There are a large number of information sources on the Internet. We aim to improve the access of the Proceedings and aim to provide access to this information.

### References

- [1] Tsuda, Yoshiomi. Conference Proceedings and materials. The present state of conference Proceedings and materials. *Jyoho no Kagaku to Gijyutu*. 2005, 55(5), p.208-213. (Japanese)
- [2] Toyoda, Yuji. Possibilities for effective use and distribution of preprints. *Jyoho no Kagaku to Gijyutu*. 1998, 48(6), p.336-343. (Japanese)
- [3] Ikeda, Kiyoshi. Grey literature. Overview: rethinking the grey literature's definition. *Jyoho no Kagaku to Gijyutu*. 2012, 62(2), p.50-54. (Japanese)
- [4] Fukazawa, Takeyasu. Analysis of the published Proceedings of international conference-Scientific conference held in Japan. *Jyoho no Kagaku to Gijyutu*. 1993, 43(10), p.913-915. (Japanese)
- [5]NDL OPAC. <https://ndlopac.ndl.go.jp/>
- [6]CiNii Books. <http://ci.nii.ac.jp/books/>
- [7]JOPSS. <http://jolissrch-inter.tokai-sc.jaea.go.jp/search/servlet/interSearch>
- [8]JACoW. <http://www.jacow.org/>
- [9] Wayback Machine. <http://archive.org/web/web.php>
- [10]INIS. <http://www.iaea.org/inis/>