

## A Synergistic System of Institutional Repository and Researcher Database

Kensuke Baba  
*Research and Development Division  
 Kyushu University Library  
 Fukuoka, 812-8581, Japan  
 Email: baba@lib.kyushu-u.ac.jp*

Masao Mori  
*Institutional Research Office  
 Kyushu University  
 Fukuoka, 812-8581, Japan  
 Email: mori@ir.kyushu-u.ac.jp*

Eisuke Ito  
*Research Institute for IT  
 Kyushu University  
 Fukuoka, 812-8581, Japan  
 Email: itou@cc.kyushu-u.ac.jp*

**Abstract**—The paper introduces a practical Web system which activates institutional repositories. Institutional repository is an important service of libraries in academic institutions. The authors of the paper developed a linking system between the institutional repository and the researcher database of their university. By the developed system papers registered in the institutional repository are linked from lists of papers in the researcher database, which improves the accessibility of the institutional repository. Additionally, the system reuses the metadata of the papers registered in the researcher database for registrations to the institutional repository, which reduces the efforts of researchers. As a result, this system is expected to encourage the registration of papers to the institutional repository. The paper describes the concepts and the details of the system. The essential idea can be applied to other academic institutions.

**Keywords**—institutional repository; Web database; Open Access; scholarly paper; library

### I. INTRODUCTION

The subscriptions for electronic journals occupy a large proportion of the budget of university libraries, which forces researchers to make funds for the price or restrict the number of journals to read. This is a quite unacceptable situation for academic societies. A solution of the problem is the idea of “Open Access [3][4]”, which is to open free-access information to the public. Publishing of scholarly papers on an institutional repository (IR) is one of activities based on the idea. An increase of the papers stored in IRs is expected to solve the problem of journal prices indirectly.

IR has been receiving increasing attention; actually 129 universities in Japan have their own IR (Aug. 12, 2010). However, the number of the papers in the IRs is small compared with the number of the papers actually written by the researchers in the universities. One of the obstacles to increase the number of papers is the fact that researchers are forced some efforts when they register their papers (and usually this operation is repeated when they upload the papers on their web-site, submit a list of the papers as a report to their institute, and so on!). Additionally, it is difficult to prove the effectiveness of IR, hence researchers do not think that registering papers to an IR is worth their time.

The aim of our research is to increase the number of the papers in IRs. Then, we set the following two purposes:

- To increase the number of accesses of papers on IRs,
- To reduce the efforts for registering papers.

The first purpose is to make clear an effectiveness of IRs. The number of citations to the registered papers from other papers is a measure of the effectiveness of an IR, and the number of the accesses of a paper can be considered as a rough upper-bound of the number of citations to the paper. The Ranking Web of World Repositories [11], which is provided by the Spanish Cybernetics Lab. in the Consejo Superior de Investigaciones Científicas (CSIC), takes account of the “visibility” in addition to the number of the contents. The visibility is the total number of links from external sites. The number of accesses of papers is expected to be improved by links from a reasonable list such as a database or a result of a search. As to the second purpose, the practical efforts of registration cannot be zero, therefore it seems to be effective that the information of a registration is shared and reused with other systems.

For the previous two purposes, we developed a linking system between the IR and the database of researcher activities of Kyushu University and are operating the system from Apr, 2010. For the first purpose, by the linking system, some papers in the IR are linked from a list in the researcher database, which improves the accessibility of the IR. For the second purpose, the information of the registered papers are shared between the IR and the researcher database by the linking system. In Kyushu University, any researcher has a duty to input his/her research and education results into the researcher database, therefore some efforts of a researcher can be reduced. It is not special for our university that researchers are required to register their research activity to a researcher database and an IR. For example, the National Academic Research and Collaborations Information System (NARCIS) [2] in the Netherlands has been developed in 2009 in order to run a central search engine of academic information with linking function to IR. NARCIS can be the one-stop national service by collecting all kind of lump academic information in the Netherlands. Our linking system help researchers to register their articles accurately

in researcher database and IR.

This paper introduces the linking system we developed. First, the basic information of the IR and the researcher database are described, which makes the problems clear. Next, the two subsystems of the linking system are explained. The subsystems correspond to the previous two purposes, respectively. By the developed system the accessibility of the IR is improved and the effort of registration to the IR is reduced, which yields an increase of the papers stored in the IR. The essential idea of our system can be applied to other institutions which have their IR and researcher database.

The rest of this paper is constructed as follows. Section II describes the basic information and the problems in the IR and the researcher database in Kyushu University. Section III explains the concepts and the concrete functions of the developed system. Section IV concludes this paper and introduces our future work.

## II. PROBLEMS

This section describes the basic information of QIR (Kyu(Q)shu University Institutional Repository) [10] and DHJS (Academic Staff Educational and Research Activities Database in Kyushu University, “Daigaku Hyoka Joho System” in Japanese) [8], and then makes clear the problems we tackle.

### A. QIR

QIR is an IR based on DSpace [9] and operated by the Kyushu University Library since Apr, 2006. The total number of the contents in QIR is 13,948 (Mar. 31, 2010), which is extremely large compared with the IRs of the other Universities in Japan. However, in the Ranking Web of World Repositories [11], QIR is ranked at the 88th position (Jul, 2010). Table I shows the number of the contents in QIR. The largest content is “Departmental Bulletin Paper” and its ratio is about 75%. Most of QIR contents are original (that is, unpublished). Therefore, links on Web are expected to be an effective factor to increase the number of accesses to contents.

Generally IR stores the full-text of a paper in addition to the metadata of the paper such as the title, the authors, the name of the proceedings, and so on. Figure 1 is a web image of QIR. The page is the profile page of a researcher and the list is the result of a search of the name in the author fields. The third column is the title of each paper and linked to the site of detailed information of the paper with the full-text. The rightmost column is the number of the accesses to each paper. In addition to the search of author, it is possible to search by general keywords in the fields of the title, the abstract, and so on.

For researchers (at least in Kyushu University), one of the obstacles to register their papers to IR is the fact that the effectiveness of the registration is not clear. Watson [7] and

Table I  
THE NUMBER OF THE CONTENTS IN QIR (MAR. 31, 2010).

Type	Number of Contents
Journal Article	1,142
Thesis or Dissertation	95
Departmental Bulletin Paper	10,443
Conference Paper	731
Presentation	137
Book, Chapter	95
Technical Report	395
Research Paper	105
Article	164
Preprint	146
Learning Material	32
Others	463
Total	13,948

O’Leary [5] verified the validity of electronic journals by analyzing the relation between download history of papers and the citations in the papers. However, as to IRs in Japanese universities, no significant correlation is found as the relation [6]. The reason is considered that the number of the registered papers and the number of the accesses to the papers are not enough as samples to estimate some properties statistically.

### B. DHJS

Kyushu University developed a researcher database in 2005, and “DHJS” is the abbreviated name in Japanese of the database. DHJS has various kinds of data of any researcher in Kyushu University, for example, posts, research interests, and scholarly papers. The number of the researchers in Kyushu University is 2,197 at Oct. 1, 2009. DHJS consists of the two subsystems, the data-entry system and the viewer system. The data-entry system is developed from scratch by Kyushu University. The viewer system uses a commercial high-speed engine on XML as the backend search system. The main technology of the engine was developed in Kyushu University.

The data-entry system supports researchers to register their research activities to DHJS. In Kyushu University, any researcher has to register their research interests, research activities includes the metadata of scholarly papers, and so on. The viewer system shows the research activities registered in DHJS by the data-entry system. The data is separated with respect to each researcher and the registered metadata of scholarly papers of a researcher are described as a list.

Any researcher in Kyushu University has a duty to input his/her research and education results into DHJS. Therefore, DHJS has the list of (most of) the paper titles that were produced in Kyushu University in recent years. The number of the unique paper titles in DHJS is over 70,000. However, QIR has only 13,948 contents as mentioned in the previous subsection, that is, more than 56,000 papers are not uploaded

The screenshot shows the QIR (Kyushu University Institutional Repository) profile page. At the top, it says 'プロフィールページ' and '九州大学'. Below that, it displays the search results for 'Baba, kensuke' with 44 items. A table lists several papers with their titles, authors, and access counts.

全選択	発行	タイトル	著者	アクセス数
<input type="checkbox"/>	2010-04-06	An Identifiable Yet Unlinkable Authentication System with Smart Cards for Multiple Services	Nakamura, Toru Inenaga, Shunsuke Ikeda, Daisuke Baba, kensuke Yasuura, Hiroto	120
<input type="checkbox"/>	2010-03	String Matching with Mismatches by Real-valued FFT	Baba, Kensuke	193
<input type="checkbox"/>	2010-02-28	機関リポジトリの有効性分析	馬場, 謙介 伊東, 栄典 吉松, 直美 星子, 奈美	603
<input type="checkbox"/>	2009-12	A Model of Publication of Scholarly Papers on Institutional Repositories	Baba, Kensuke Ito, Eisuke Yoshimatsu, Naomi Hoshiko, Nami Murakami, Kazuaki	125
<input type="checkbox"/>	2009-10-27	PIRに基づく匿名認証とその応用	中村, 徹 稲永, 俊介 池田, 大輔 馬場, 謙介 安浦, 寛人	212
<input type="checkbox"/>	2009-07	Anonymous Authentication Systems Based on Private Information Retrieval	Nakamura, Toru Inenaga, Shunsuke Ikeda, Daisuke Baba, Kensuke Yasuura, Hiroto	359

Figure 1. The web image of a list of papers stored in QIR. This example is the result of a search of “Kensuke Baba” in the author fields.

to QIR. Therefore, there is yet room for improvement of the number of the contents in QIR.

### III. SYNERGY OF QIR AND DHJS

We developed a linking system between QIR and DHJS, and are operating the system from Apr, 2010. This section explains the concepts of the system and concrete functions of the system in the two points, that is, the connections with the data-entry system and the viewer system in DHJS.

#### A. Overview of System

We implemented two concepts to encourage researchers to register their papers to QIR. One is a support for registration to QIR by reducing the efforts of the input of the metadata of papers. The other is an improvement of the accessibility of the papers in QIR by linking from lists of papers in DHJS.

Figure 2 is the outline of the systems. For the former concept, researchers can register their paper to QIR by submitting the full-text instead of the metadata and the full-text. Additionally, researchers can know whether each paper is stored in QIR. As to the latter concept, users can search the full-text of a paper at the same time they search a researcher in DHJS in addition to search in QIR. This function should increase the number of the accesses to papers in QIR, which encourages researchers to register their paper.

#### B. Connection with Data-entry System

The linking system realizes two functions on the data-entry system. Figure 3 is a web image of the interface of

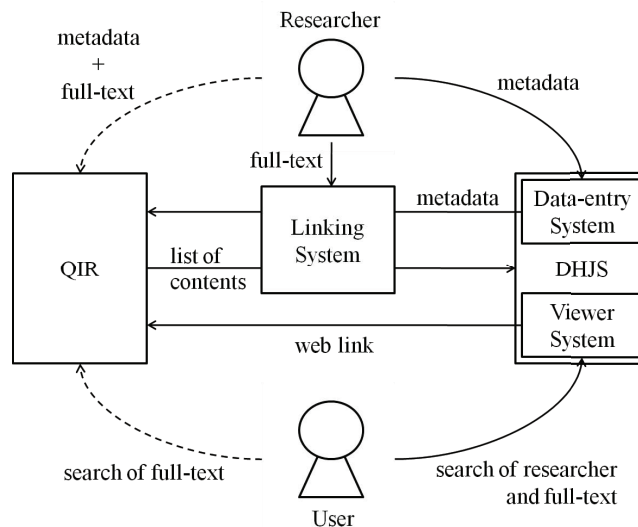


Figure 2. The outline of QIR, DHJS, and the linking system.

the data-entry system, especially for registration of scholarly papers.

Firstly, by a link to the registration form of QIR, researchers can register their papers on the data-entry system. Additionally, when a researcher wants to register a paper, the metadata of the paper is used to fill the registration form. In Figure 3, one of the icons in the rightmost row is the link to



Figure 3. The data-entry system of DHJS connected with QIR. This example is the interface for registration of (the metadata of) scholarly papers for the user “Kensuke Baba”.

the registration form of QIR. The interface will be improved by Nov, 2010 so that users can upload a full-text of their paper directly in DHJS.

Secondly, the linking system obtains from QIR

- Whether the papers of the current user (researcher) are registered to QIR, respectively and
- The number of the registered papers of the user in QIR

and describes them on the data-entry system. In Figure 3, the column named “QIR” means whether the user want to link to QIR for each paper, and the next column means whether the paper is stored in QIR (“o” if stored). The three numbers (with some Japanese words) on the table are, respectively, the number of the papers registered in DHJS, the number of the papers checked at the “QIR” column, and the number of the papers stored in QIR. This function is expected to encourage researchers to register by showing the current state of their registration.

### C. Connection with Viewer System

The papers which checked on the “QIR” column at the data-entry system are linked to the full-text of the corresponding paper in QIR by icons on the viewer system. There exist two kinds of icons which distinguishes whether the paper is registered in QIR. Figure 4 is the list of the papers registered in DHJS. For each paper in the list, a dark-colored icon “fulltextQIR” is added, a light-colored icon “searchQIR” is added, or there is no icon. The first

case means that there is the full-text of the paper in QIR, the second that there is no full-text (although the researcher wants to register), and the other that the researcher does not want to link. In the second case, the linking system returns the result of the search by the author name in QIR.

The numbers indexing the papers in the table in Figure 4 are the numbers in Figure 3. (The paper 3 is written in Japanese.) You can see that the situation of the “QIR” column and the next column in the data-entry system is reflected to the icons in the viewer system.

## IV. CONCLUSION AND FUTURE WORK

In this paper, the linking system between the IR and the researcher database in Kyushu University were introduced. The system provides links from the metadata of papers in the researcher database to the corresponding full-text in the IR. Additionally, the system reuses the metadata of the papers in the researcher database for registrations to the IR. By the previous functions, the linking system is expected to enhance the registration of papers to the IR. The concepts of the developed system were explained generally, which leads the idea of the system to be applicable to other academic institutions.

It is our future work to analyze the number of the accesses to the papers in the IR. We already analyzed the access log of the IR from Apr, 2009 to Mar, 2010 [1], and are taking the access log from Apr, 2010. Also we are going to verify

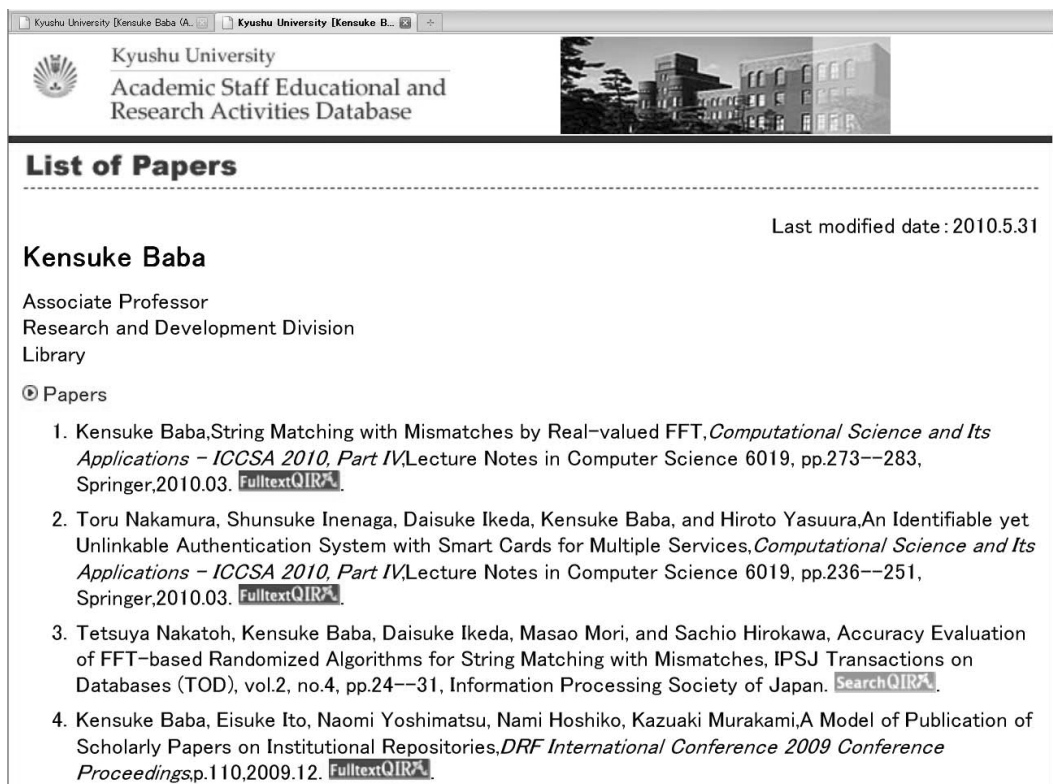


Figure 4. The viewer system of DHJS connected with QIR. This example is the list of the papers of “Kensuke Baba” in DHJS.

an effectiveness of the developed system by analyzing the number of the registrations and the ratio of the registrations by the system. Additionally, we are planning to develop an “embargo” system in the IR, that is, a system to manage the exhibition of the papers on the basis of the copyrights and the policies of publishers. The embargo system can make more efficient the operations for registrations to IRs.

#### ACKNOWLEDGMENTS

We thank the staff of the Kyushu University Library, Naomi Yoshimatsu, Nami Hoshiko, and Eriko Kudo, for their comments about QIR and general institutional repositories. We also thank the anonymous referees for their helpful comments to improve this work.

#### REFERENCES

- [1] Baba, K., Ito, E., Yoshimatsu, N., and Hoshiko, N.: “An Analysis of Validity of Institutional Repository” (in Japanese), Proc. The Second Forum on Data Engineering and Information Management (DEIM 2010), F7-3, 2010.
- [2] Dijk, E., Baars, C., Hogenaar, A., and van Meel, M.: “NAR-CIS: The Gateway to Dutch Scientific Information”, Proc. ELPUB2006 Conference on Electronic Publishing, 2006.
- [3] Harnad, S.: “The Green Road to Open Access: A Leveraged Transition”, The Culture of Periodicals from the Perspective of the Electronic Age, L’Harmattan, pp. 99–105, 2007.
- [4] Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., Stamerjohanns, H., and Hilf, E.: “The Access/Impact Problem and the Green and Gold Roads to Open Access”, Serials Review 30, 2004.
- [5] O’Leary, D. E.: “The relationship between citations and number of downloads in Decision Support Systems”, Decision Support Systems, vol. 45, no. 4, pp. 972–980, 2008.
- [6] Sato, S., Tomimoto, H., and Itsumura, H.: “The Relationship between Citations and Number of Downloads in Institutional Repositories” (in Japanese), <http://www.tulips.tsukuba.ac.jp/dspace/handle/2241/104229> (Aug. 12, 2010).
- [7] Watson, B. A.: “Comparing citations and downloads for individual articles”, A Journal of scientific research on biological vision, vol. 9, no. 4, pp. 1–4, 2009.
- [8] DHJS: Academic Staff Educational and Research Activities Database (“Daigaku Hyoka Joho System” in Japanese) in Kyushu University, [http://hyoka.ofc.kyushu-u.ac.jp/search/index\\_e.html](http://hyoka.ofc.kyushu-u.ac.jp/search/index_e.html) (Aug. 12, 2010).
- [9] DSpace, <http://www.dspace.org/> (Aug. 12, 2010).
- [10] QIR: Kyu(Q)shu University Institutional Repository, <https://qir.kyushu-u.ac.jp/dspace/> (Aug. 12, 2010).
- [11] Ranking Web of World Repositories, <http://repositories.webometrics.info/> (Aug. 12, 2010).