National-wide Repository Network - Vision and Practice -

Prof. Hideaki Takeda, Dr. Eng.,
National Institute of Informatics (NII)
takeda@nii.ac.jp
What?

Repository Network

Network of Researchers

Network of Systems

Network of IR Managers
Repository Network

What?

• Network of Researchers
  – Mobility of researchers
  – Collaboration among researchers
• Network of Systems
  – Data-level interoperability
  – System-level interoperability
  – Policy/statistics-level interoperability
• Network of IR Managers
  – Mobility of IR managers
  – Collaboration among IR managers
Why National-wide network in Japan?

• Language
  – Language of contents
    • Over 80% of contents are written in Japanese
      – J:753,978, E:158,708
  – Language of IR managers

• Area of Mobility
  – Researchers/IR Managers

• Budget
Growth of Institutional Repositories in Japan

No. of Contents

<table>
<thead>
<tr>
<th>No. of Contents</th>
<th>No. of IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>26,095</td>
<td>33,862</td>
</tr>
<tr>
<td>44,252</td>
<td>52,640</td>
</tr>
<tr>
<td>70,176</td>
<td>146,277</td>
</tr>
<tr>
<td>234,375</td>
<td>384,449</td>
</tr>
<tr>
<td>482,296</td>
<td>552,909</td>
</tr>
<tr>
<td>49,536</td>
<td>170,090</td>
</tr>
<tr>
<td>209,833</td>
<td>482,964</td>
</tr>
<tr>
<td>251,744</td>
<td>290,325</td>
</tr>
<tr>
<td>2,32</td>
<td>552,909</td>
</tr>
</tbody>
</table>

Number of IRs
- National Univ.: 80
- Public Univ.: 24
- Private Univ.: 87
- Others: 41
- Total: 232

Types of Contents
- Journal articles: ca. 290,000
- Department Bulletin: ca. 550,000
- PhD Thesis: ca. 70,000
- Others: ca. 390,000
- Total: ca. 1,300,000

(March 31, 2012)

Types of Contents
- Journal Article (学術雑誌論文)
- Departmental Bulletin Paper (紀要論文)
- Thesis or Dissertation (学位論文)
- Conference Paper (発表論文)
- Presentation (発表用資料)
- Book (図書)
- Technical Report (テクニカルレポート)
- Research Paper (研究報告書)
- Article (一般雑誌論文)
- Preprint (プレプリント)
- Learning Material (教材)
- Data or Dataset (データベース)
- Software (ソフトウェア)
- Others (その他)
- 機関リポジトリ数
Repository Network
How?

• Network of Researchers
  – Aggregation of contents -> JAIRO
  – Interoperability of Researchers
    -> Researcher Name Name Resolver

• Network of Systems
  – Cloud Service -> JAIRO Cloud
  – Policy-level interoperability -> SPCJ
  – Statistics-level interoperability -> ROAT

• Network of IR managers -> DRF
Repository Network

Network of Researchers

Network of Systems

Network of IR Managers
Researcher Name Resolver (RNS)
- Switching Board for Researcher’s ID -
Researcher Name Resolver (RNR)

- Identity service of researchers in Japan
  - Basic information of researchers with unique ID
  - links to researcher’s web resources

- Basic Idea
  - Extend and use Kaken ID (given to researchers by MEXT)
  - Manual and Automatic mapping to ID systems in other services

http://rns.nii.ac.jp/nr/1000010295694

Researcher ID (consistent with Kaken ID)
Researcher name
(Kanji, Katakana, English character)
Researcher basic information

Direct links to the services outside
Search query formatted URL links to the services outside
KAKENHI project research fields
KAKENHI project keywords

URI for this page

198,861 researchers (27, Apr, 2012).
Name Disambiguation Approach for Institutional Repositories in Japan

- Problem: How to coordinate local ID (institutional ID) with global ID?
  - Some IRs contain global ID as author ID
  - Some IRs contain local ID as author ID
Name Disambiguation Framework for Federated Search Portal JAIRO

- This project aims at identifying authors in the federated search portal for Institutional Repositories in Japan which we call JAIRO (Japanese Institutional Repository Online).
The National Institute of Informatics (NII) has launched JAIRO Cloud service, from April 2012, to promote dissemination of research and educational outcomes and facilitate open access in Japanese universities and research institutions.

JAIRO Cloud provides a shared instance of institutional repository system to academic institutions. The system is implemented based on WEKO, a repository software developed by NII.

Number of institutions applied

69

(as of May 2012)
Service Architecture

USER

Univ.

- Deposit of contents
- UI Management

NII

- SW management
  - Initial setting
  - Update etc.
- OS management
- SW management
- NW management

HW & Network

Host OS

NetCommons

JAIRO

A Univ. IR

B Inst. IR

C Univ. IR

WEKO

NetCommons 2

GuestOS

CiNii

Deposit of contents

SW management

Initial setting

Update etc.

OS management

NUXT

UI Management

SW management

NW management

HW & Network

Univ.

access
Check academic societies' copyright policies

<table>
<thead>
<tr>
<th>Color</th>
<th>Policy</th>
<th>Society</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>can archive pre-print and post-print</td>
<td>139</td>
<td>5.4%</td>
</tr>
<tr>
<td>Blue</td>
<td>can archive post-print</td>
<td>668</td>
<td>25.8%</td>
</tr>
<tr>
<td>Yellow</td>
<td>can archive pre-print</td>
<td>11</td>
<td>0.4%</td>
</tr>
<tr>
<td>White</td>
<td>archiving not formally supported</td>
<td>223</td>
<td>8.7%</td>
</tr>
<tr>
<td>Gray</td>
<td>Under review / Not published / No response / Other</td>
<td>1511</td>
<td>59.4%</td>
</tr>
</tbody>
</table>

No. of registered societies 2542

SCPJ Policy Colours
- Green: 5.4%
- Blue: 25.8%
- Yellow: 8.7%
- Gray: 59.4%
- White: 0.4%

Digital Repository Federation
COAR General Assembly 2012 in Sweden
SCPJ project
(Societies Copyright Policies in Japan)

• Outline
  – Investigate Japanese scholarly societies’ open access policies and publicize this information through SCPJ (Societies Copyright Policies in Japan) database.
  – http://scpj.tulips.tsukuba.ac.jp/

• Two Features of SCPJ database
  – The only database that covers OA policies of almost all Japanese academic societies (about 1,800)
  – No policy is a policy: to regard responses of “under consideration” or “No response” as one of the policies
SCPJ project
(Societies Copyright Policies in Japan)

- Color coding of OA policy in SCPJ database

<table>
<thead>
<tr>
<th>Color</th>
<th>OA policy</th>
<th>Num</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>Author can archive both Pre-prints &amp; Post-prints</td>
<td>44</td>
</tr>
<tr>
<td>Blue</td>
<td>Author can archive only Post-prints</td>
<td>252</td>
</tr>
<tr>
<td>Yellow</td>
<td>Author can archive only Pre-prints</td>
<td>4</td>
</tr>
<tr>
<td>White</td>
<td>Author can not archive</td>
<td>157</td>
</tr>
<tr>
<td>Gray</td>
<td>Under consideration</td>
<td>1,379</td>
</tr>
</tbody>
</table>
HTTP log files uploaded from IRs

User interface for uploading, downloading and maintenance

- Resolve/Merge/Sort
- Filter a la COUNTER
- Cut off bot’s accesses
- Show the results of analysis vis à vis the metadata of articles

Bots database, maintained collaboratively

ROAT: Repository Output Assessment Tool

- Shared online platform for assessing the impact (= downloads) of repositories
- Standardized counting through various repository softwares
- Metadata based integrated reporting
- Serving 30 repositories currently

Stats summary

Output stats for each IR

JAIRO

Metadata ingestion

HTTP log files uploaded from IRs
Repository Network
What?

Network of Researchers

Network of Systems

Network of IR Managers
DRF 2011

Executive Board

Advisers

Planning and Coordinating WG

International Relations WG

Technical Support WG

127 members

Activities
- 3 Trainings for beginner
- 1 Training for senior
- 1 Workshop
- 1 International meeting

- Poster presentation at Berlin 9
- Dispatch of 11 instructors for 8 lecture meetings
- 1 Technical workshop
- MoU: DRF, RSP and UKCoRR
DRF Activity Challenges: 2010 - 2012

1. Survey and Research to Enhance Content
2. Development of Institutional Repository Technical Infrastructure
3. Information Sharing, Staff training, Regional support
4. Promotion of International Cooperation

133 members (As of May, 2012)
Summary

• *The Network Effect* is working!
  – Focusing on national-wide network sharpens it
  – Variety of supports prompts newcomers to join

• Network of contents should be more functional
  – more collaboration in ID (e.g., ORCID)
  – Collaboration with other data

• More expectations on network of IR managers