Access log analysis of university repository

Ito, Eisuke
Research Institute for Information Technology, Kyushu University

Baba, Kensuke
Kyushu University Library

Hirokawa, Sachio
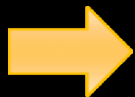Research Institute for Information Technology, Kyushu University

https://hdl.handle.net/2324/18333

ACCESS LOG ANALYSIS OF UNIVERSITY REPOSITORY

Eisuke Ito, Kensuke Baba and Sachio Hirokawa
Kyushu University
{ito.eisuke.523@m, baba@lib, hirokawa@cc} .kyushu-u.ac.jp

Outline

1. Introduction

- Scholarly papers should be accessible by anybody.
  - Because, they should be sharable by all human beings.
- Recently,
  - EJ (Electric Journal) becomes popular.
  - EJ is a closed service.
  - EJ subscription price becomes very high.
- To against EJ,
  - Open Access movement was arising.
- IR (Institutional Repository)
  - Provide open access to institutional research output by self-archiving it.

Problems of Institutional Repository

- Problems
    - No quality assurance of original papers
    - Difficult to increase papers
- Reasons
    - Based on self-uploading by volunteers
    - No incentive for researcher (Just overhead work)
- Solutions for IR problems
    - Automatic collection of published paper metadata
    - Prompt uploading with few work (easy uploading)

    } System Developing

    - *Show effectiveness of IR*

Objectives

- Analyze access log of university repository
  - To find new quality assurance method
  - To develop feedback system for researcher
  - To create something good search or recommendation system
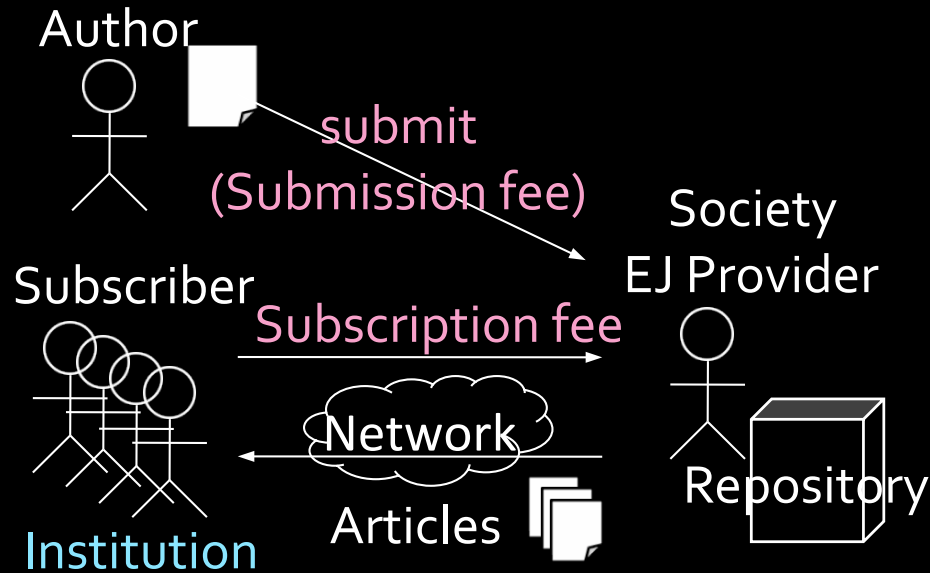 - To evaluate effectiveness of IR

Outline

1. Introduction
2. IR (Institutional Repository)
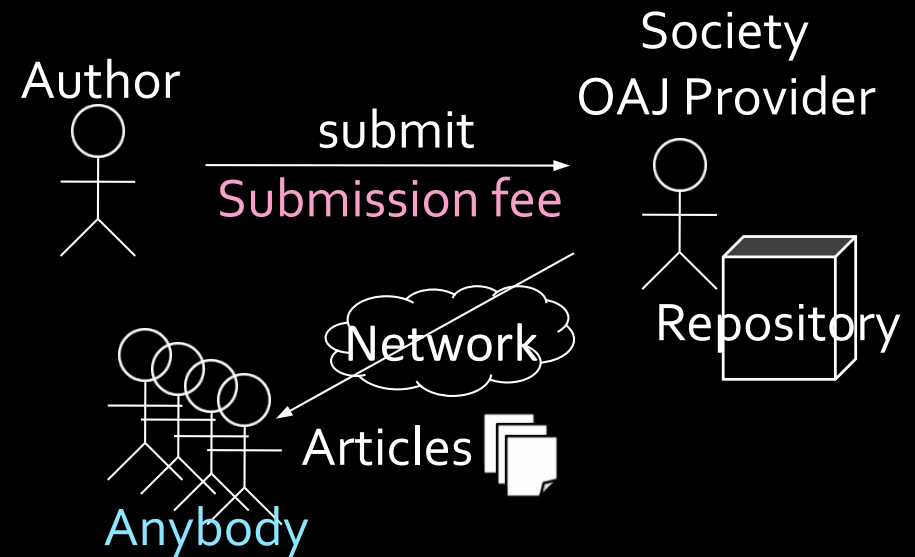3. QIR: Kyu(Q)shu University IR
4. Research topics
5. Conclusion

2. IR (Institutional Repository)

- Open Access movement
  - Open access journal
  - Self archiving
- Institutional Repository
  - Provide open access to institutional research output by self-archiving it
- Why University Library operates IR system?
  - Showcase of Institution
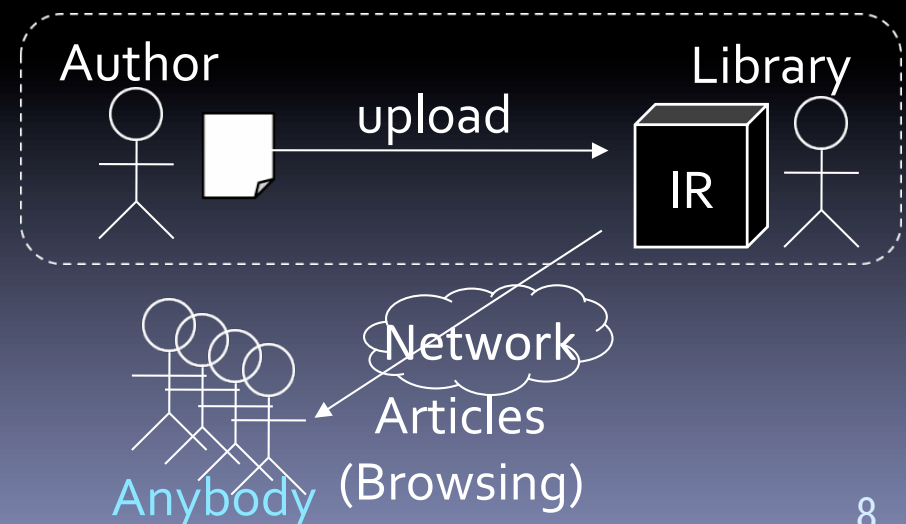 - Social contribution
  - To against EJ

E-Journal

Author
submit
(Submission fee)

Subscriber
Subscription fee
Network
Articles
Institution

Society
EJ Provider
Repository

Open Access Journal

Author
submit
Submission fee

Society
OAJ Provider
Repository

Network
Articles
Anybody

Institutional Repository

Author
upload
Library
IR

Network
Articles
(Browsing)
Anybody

8

Outline

1. Introduction
2. IR (Institutional Repository)
3. QIR: Kyu(Q)shu University IR
4. Research topics
5. Conclusion

3. QIR: Kyu(Q)Shu University Repository

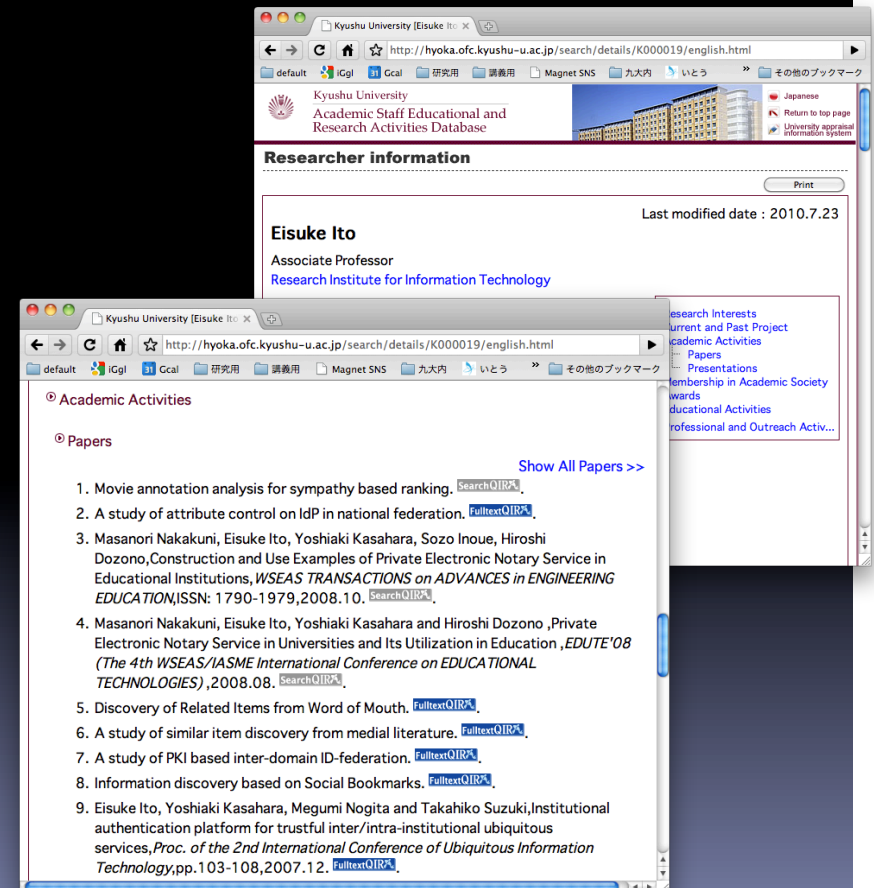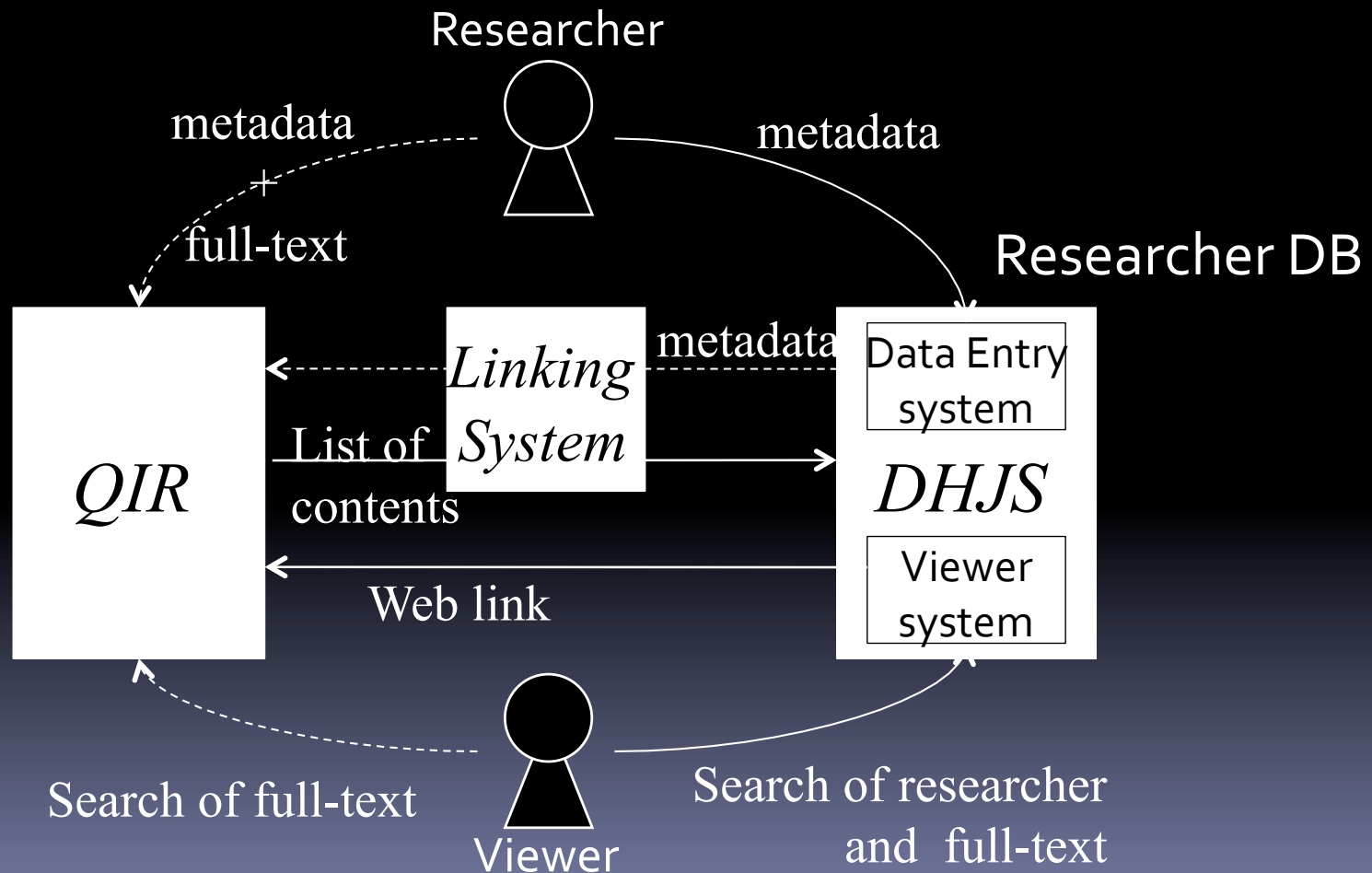*QIR*

DHJS
Researcher DB

DSpace based IR system
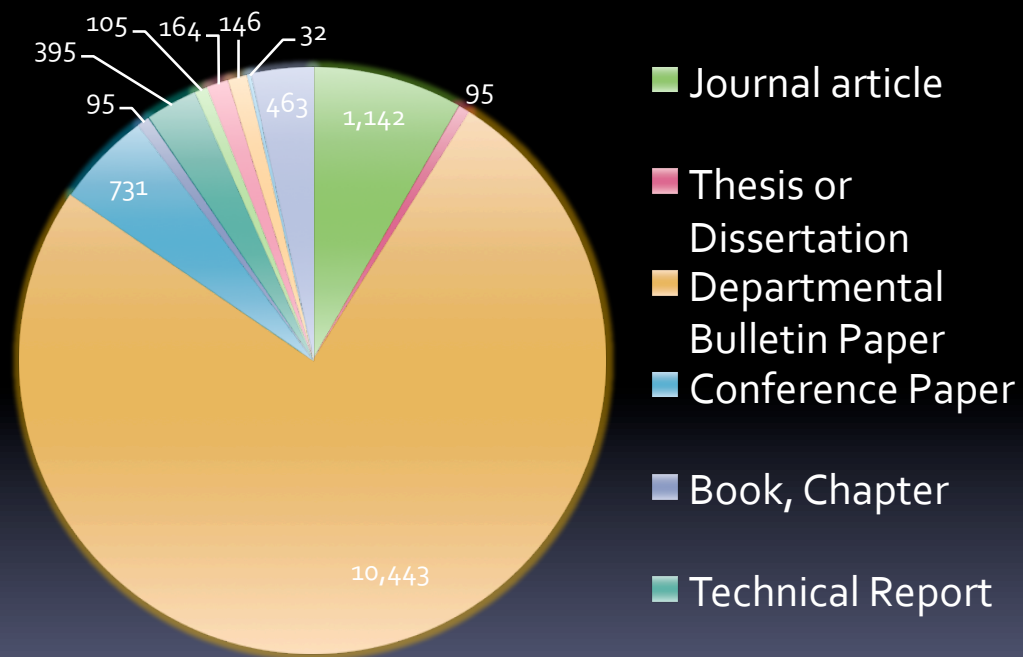by Kyushu Univ. Library
Since April 2006.

The outline of QIR, DHJS, and the linking system

Contents

- 13,948 contents (Mar. 31, 2010)

Type	contents	%
Journal article	1,142	8.27%
Thesis or Dissertation	95	0.69%
Departmental Bulletin Paper	10,443	75.61%
Conference Paper	731	5.29%
Book, Chapter	95	0.69%
Technical Report	395	2.86%
Research Paper	105	0.76%
Article	164	1.19%
Preprint	146	1.06%
Learning Material	32	0.23%
Others	463	3.35%
Total	13,811	100.00%

Outline

1. Introduction
2. IR (Institutional Repository)
3. QIR: Kyu(Q)shu University
4. Research topics
5. Conclusion

4. Access log analysis of QIR

- Logs
  - QIR Log (Apache style)
  - DSpace's stat data
  - Google Analytics
  - Log of Researcher activity DB
  - Log of Link Database (between QIR and Researcher DB)
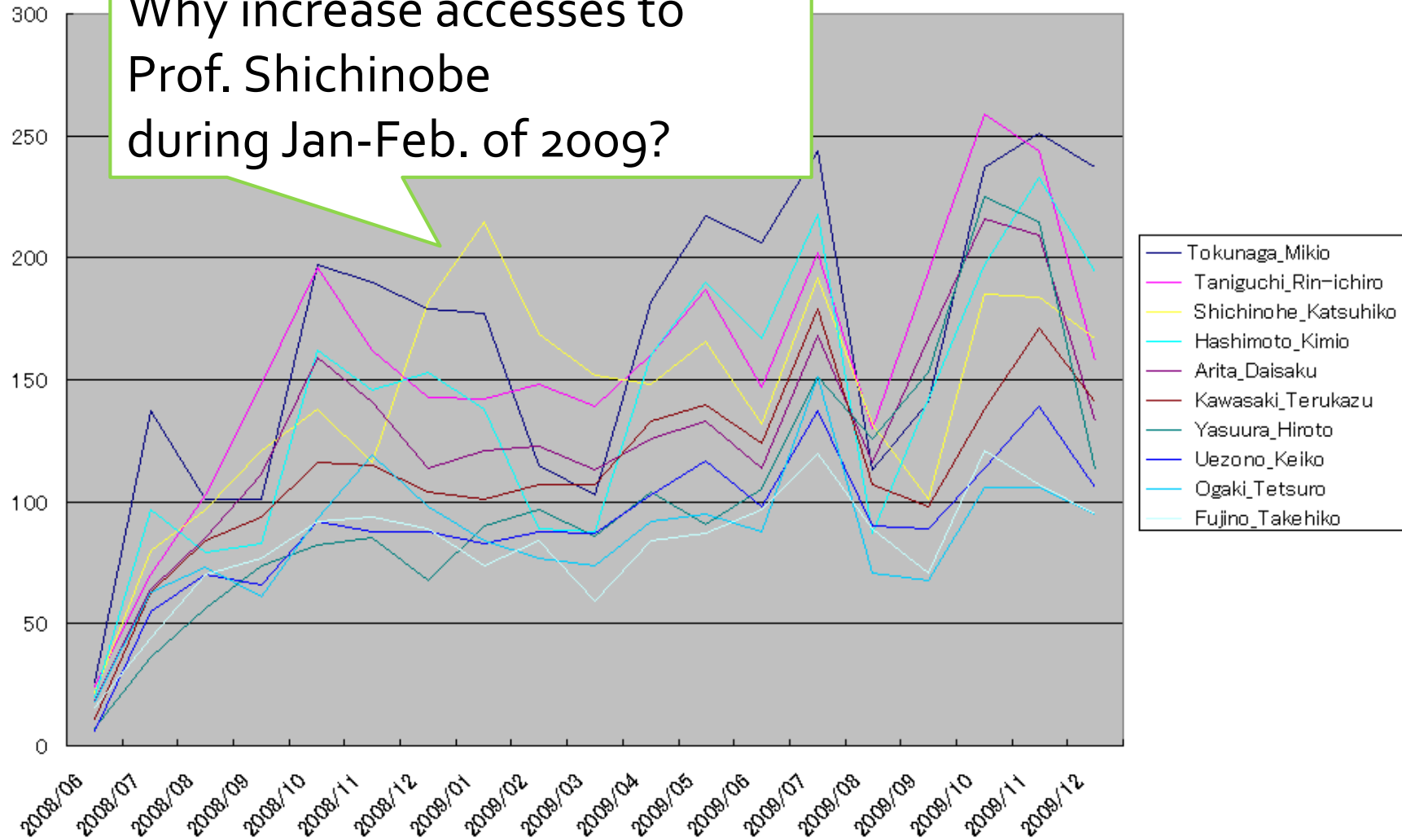
4.1 Quality assurance method

- Quality assurance of Published papers
  - Peer review for brand-new paper
  - Journal Impact factor, or Number of citation
- Many original papers on IR
  - University bulletins
- How to estimate quality of original paper?
  - Number of page view
  - Number of visited unique users   } Access log analysis
  - Author's point

4.2　Feedback to researcher

- Present IR gives few incentive for researcher
- No feedback from IR
- Develop feedback system from IR
  - Monthly report of page view, download number
  - Nations of access
  - Number of researcher's access
    - Who has interest to your research
  - Number of student's access
    - Who want to belong your laboratory
  - Relation with access spike and news topic.
    - Ex. Volcano news → Spike of access to

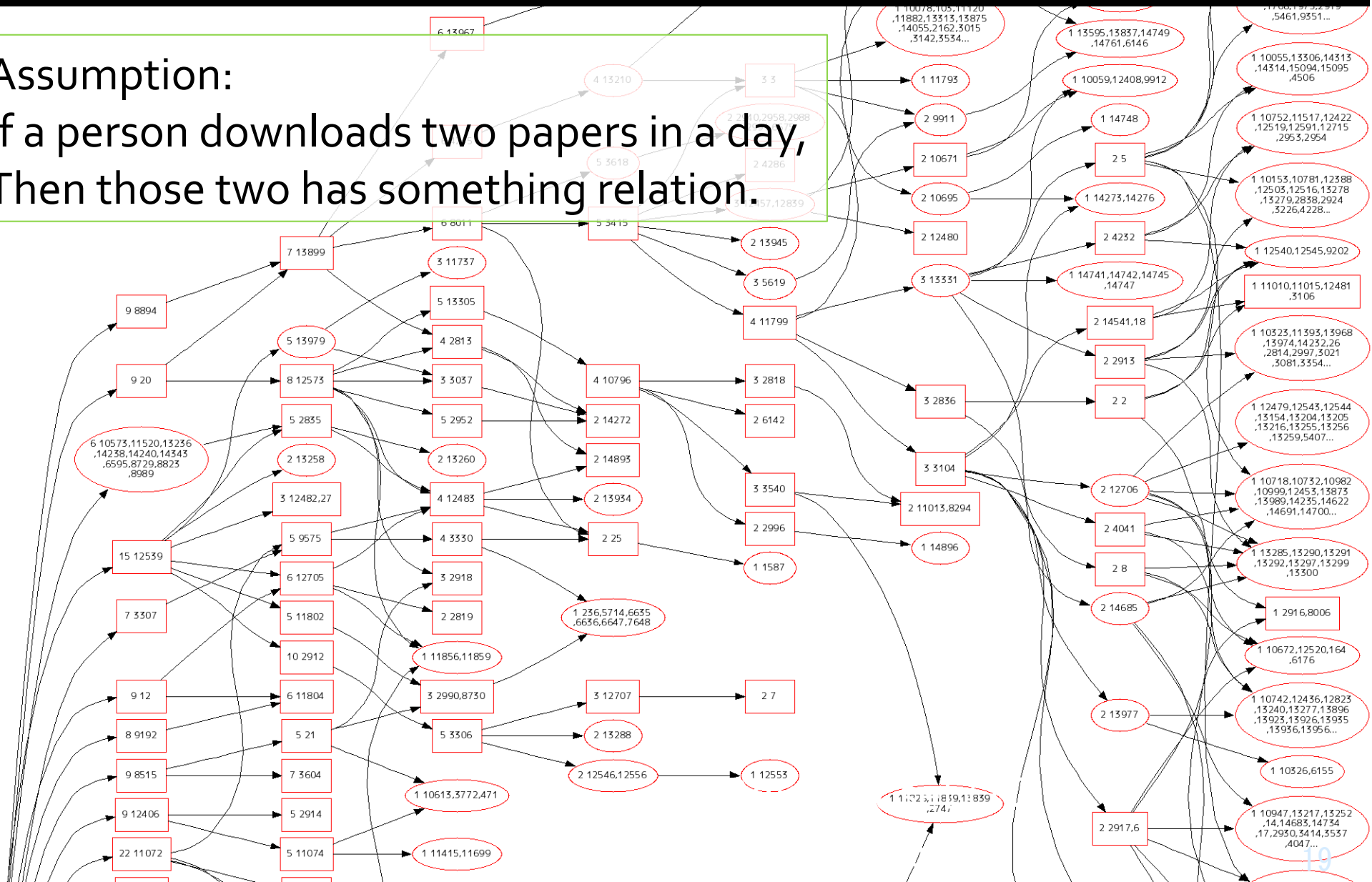Different access pattern, maybe.

Access trend

4.3 Recommendation system

- Paper recommendation
  - Like amazon.com
    - *Frequently download together*
    - *Users who download this paper also download*
  - Query Keyword
    - Collect Query words of search engine for a paper
    - If a user arrived a paper using search engine with query words,
    - Then recommend other frequently accessed paper via the similar query words

Assumption:

If a person downloads two papers in a day,

Then those two has something relation.

Outline

1. Introduction
2. IR (Institutional Repository)
3. QIR: Kyu(Q)shu University
4. Research topics
5. Conclusion

5. Conclusion

- IR may be important service for academic community.
- Access log analysis of university repository
  - To find new quality assurance method
  - To develop feedback system for researcher
  - To create something search or recommendation system
  - To evaluate effectiveness of IR
- Future works
  - To be continued

Thank you for your attention.