

# OR2015 | 10th International Conference on Open Repositories

June 8-11, 2015, Indianapolis, Indiana, USA

## Title of Proposal

### System for cross-organizational big data analysis of Japanese institutional repositories

Takurou Kawamura, Hiroshima University, Japan, tkawa@hiroshima-u.ac.jp

Kenichi Igarashi, Keio University, Japan, igarashi@lib.keio.ac.jp

Hiroshi Kato, National Institute of Informatics, Japan, hkato@nii.ac.jp

Akira Maeda, National Institute of Informatics, Japan, maeda@nii.ac.jp

Toshihiro Aoyama, Suzuka National College of Technology, Japan, aoyama@info.suzuka-ct.ac.jp

Kazutsuna Yamaji, National Institute of Informatics, Japan, yamaji@nii.ac.jp

Sho Sato, Doshisha University, Japan, shsato@mail.doshisha.ac.jp

## Session Type (select one)

Poster

## Abstract

Institutional Repositories DataBase (IRDB) is the system which collects metadata from almost all Japanese institutional repositories (IR). As one of the application, IRDB content analysis system provides statistical information depending on content type and format. It allows users to compare the cross-organizational data. However, it doesn't have the data and functionality to compare the access statistics. This poster introduces the new feature of IRDB content analysis system with JAIRO Cloud, IR cloud service used by 250 institutions in Japan, which provides the cross-organizational access statistics of IRs.

The system can divide into two major components. One is the log repository which collects filtered access logs from Japanese IRs as data source. Unprocessed logs are filtered based on COUNTER-like processing such as robots exclusion. The filtering has to be done in each IR server since it includes elimination of the privacy data. The other component is the user interface. The system enables users, including repository managers, to analyze both content and access log to the content among Japanese IRs. We are planning to deploy the system to the institution which doesn't utilize JAIRO Cloud in order to make it the standardized log analysis system in Japan.

## Conference Themes

Select the conference theme(s) your proposal best addresses:

Exploring Metrics and Assessment

## Keywords

# OR2015 | 10th International Conference on Open Repositories

June 8-11, 2015, Indianapolis, Indiana, USA

User interface to show access and content analysis, cross-organizational access statistics, identical process for comparison of access, access log repository

## Audience

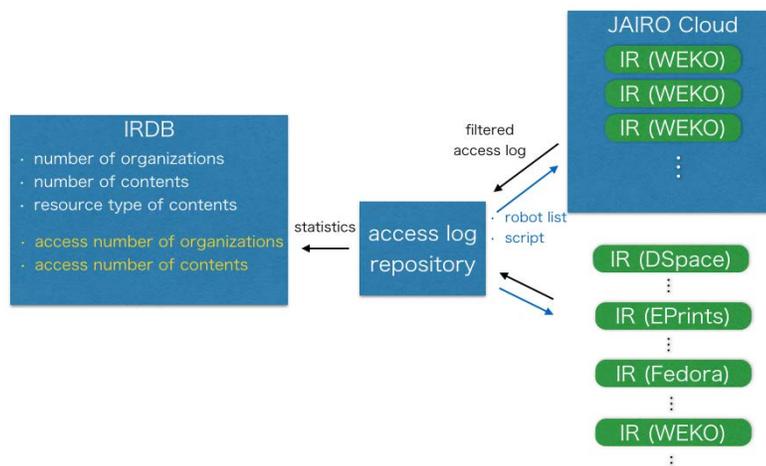
Repository managers, librarians

## Background

In order to show the repository value, it is important to know how the contents are used. Currently, the numbers of accesses are only counted by each institution using its own baseline since there is no way to compare the access among multiple repositories. However, it is interesting to analyze the cross-organizational access statistics of contents, promoting discussion for beyond current repository. This is the reason why the topic is suitable subject to the conference theme.

## Presentation content

- Background of the statistics of Japanese institutional repository
- Proposed system architecture



- Future perspective

## Conclusion

The system enables users, including repository managers, to analyze both the number of contents and the number of accesses from various angles such as content type among Japanese IRs. Some interesting facts of the IRs may be found by comparing the statistics among repositories, which will help to keep IRs moving forward.

## References

# **OR2015 | 10th International Conference on Open Repositories**

June 8-11, 2015, Indianapolis, Indiana, USA

[1] IRDB, NII Institutional Repository DataBase Contents Analysis,  
<http://irdb.nii.ac.jp/analysis/index.php>