Reviving past research data by utilizing institutional repositories

Eriko Amano
Kyoto University
Yoshida-honmachi
Sakyo-ku
Kyoto 606-8501
amano@kura.kyoto-u.ac.jp

Yasuyuki Minamiyama
National Institute of Polar Research
10-3 Midori-cho
Tachikawa-shi
Tokyo 190-8518
minamiyama@nipr.ac.jp

SUMMARY

A considerable number of databases created through Japanese public research funds are now under unstable management or use an outdated publication methodology; therefore, future access to the research results these databases contain cannot be guaranteed. Here, we examined the current status of databases listed in two major sources, Grant-in-Aid-funded databases and the National Diet Library Database Navigation Service (Dnavi), and reaffirmed this problematic situation. To ensure the sustainability of these databases, the Research Data Task Force of the Japan Consortium for Open Access Repository (JPCOAR) is currently testing a workflow for the transfer of contents from unstable databases to institutional repositories.

1. INTRODUCTION

Numerous databases in various fields have been published as the result of publicly funded research in Japan. From FY2009 to FY2011, the Grant-in-Aid for Publication of Scientific Research Results: Databases program, administered by the Japan Society for the Promotion of Science, supported the creation of 130 databases, 74 of which contained research results in the natural sciences and 56 of which contained research results in the humanities and social sciences (HSS).

2. PRELIMINARY SURVEY OF DATABASE SUSTAINABILITY

2.1 GRANT-IN-AID-FUNDED HSS DATABASES

To examine database sustainability, we conducted a preliminary survey in which we evaluated the sustainability of the 56 HSS Grant-in-Aid-funded databases with respect to their maintenance status and publication methodology. As of April 2015, 47 (84%) of the HSS databases were live, and their maintenance status was as follows: stable, 24 (51%); unstable, 20 (42%); inaccessible (including CD-ROM-based, fee-based, and registration-based databases), 3 (6%). Among the databases with unstable management, we found some that appeared to have not been updated in recent years, and for one of the databases the lead researcher had moved to another institute and the current owner of the database was not specified. Regarding publication methodology, some databases were provided in the legacy FileMaker format or as a single PDF file. Such methodologies do not ensure future accessibility to the database so newer technology should be applied to these databases to prevent loss of the research results they contain.

2.2 OTHER DATABASES

The National Diet Library Database Navigation Service (Dnavi), formaly managed by the National Diet Library of Japan and currently archived as a static dataset, is a searchable list of databases that includes not only databases published by academic institutions and individual researchers but also public use databases. Approximately 25% (5333) of the databases in Dnavi are managed by universities, laboratories, and technical colleges. A preliminary

survey conducted in 2016 revealed that the majority of databases in Dnavi created after 2000 were in most cases of discontinuation, and about half of these idle databases contained HSS research results. These databases will likely become inaccessible in the near future resulting in loss of the valuable data they contain. Therefore, to secure access to these databases and prevent data loss, the data should be migrated to a stable infrastructure such as an institutional repository.

3. CHALLENGES

3.1 DEVELOPING A WORKFLOW

To migrate unstable databases to institutional repositories, a workflow that is applicable to most databases and streamlines the process of converting and transferring the database metadata to the institutional repository is essential. In addition, a new metadata schema, the JPCOAR metadata schema, built specifically for the Japanese situation was developed recently by the Metadata Task Force of JPCOAR to replace the current standard schema for institutional repositories, junii2. The JPCOAR metadata schema assumes preservation of non-literature materials that were created in accordance to international standards.

The JPCOAR metadata schema was developed to ensure that the metadata contained in Japanese institutional repositories is compatible with the Open Science framework and is interoperable internationally. In the development of the JPCOAR metadata schema, the Metadata Task Force

- focused on interoperability with six established international metadata standards (i.e., RIOXX, BIBO, COAR vocabulary, OpenAIRE, DataCite, and Dublin Core),
- adopted major identifiers (Crossref Funder ID, ISNI, ORCID, etc.),
- monitored compliance with open access policies,

- hierarchized priority elements (creator, fundingReference, etc.), and
- enhanced elements related to research data.

In FY2016, the Research Data Task Force and Metadata Task Force carried out a trial migration of a Grant-in-Aid-funded database to an institutional repository; after subsequent review of the trial, they proposed the following workflow:

(1) Select a database in Dnavi that is under unstable management (live or discontinued databases are considered).
(2) Acquire permission from the database administrator(s) to migrate the contained data to a new database platform.
(3) Examine the database metadata and select and implement the most appropriate conversion method for registration of the metadata in an institutional repository.
(4) Map the database's unique metadata schema to the JPCOAR metadata schema.

3.2 FUTURE PLAN

In cooperation with several universities we are now testing our proposed workflow. By following this workflow, we intend to map the metadata from unstable research databases to the JPCOAR metadata schema and register the databases in real-world institutional repositories.

4. REFERENCES
Grant-in-Aid for Scientific Research (KAKENHI)
https://www.jsps.go.jp/english/e-grants/index.html

National Diet Library Database Navigation Service (Dnavi)
http://doi.org/10.11501/8427554

Japan Consortium for Open Access Repository (JPCOAR)
https://jpcoar.repo.nii.ac.jp/