オープンサイエンスとデータ

武田英明

国立情報学研究所
Scholarly Information Infrastructure

- **CiNii Articles**
  - Metadata and links of Japanese journal articles
  - 19 M records

- **JAIRO**
  - Metadata and links of Japanese institutional repositories
  - 2.5 M records

- **CiNii Books**
  - Catalog of materials held by universities
  - Bibliographic info 11 M records
  - Holding information 137 M records

- **KAKEN**
  - Project reports of MEXT supported scientific researches
  - 820 K record

- **NACSIS-CAT**
  - Shared Repositories
  - Compilation

- **JSPS**
  - Linkage to other DB services

- **MEXT**
  - Universities and Research Institutions

- **J-Stage (JST)**
  - Linkage to other DB services
  - More than 800 repositories

- **NDL**
  - More than 1,300 libraries

Note: The record numbers are as of March 2017.
Framework of the Open Science in Japan

Current NII Services and Research Workflow

- **Discovery Service**
- **Aggregation**
- **Publication Platform**

Research data life cycle:
- Discover, reuse and cite
- Plan and design
- Collect and capture
- Collaborate and analyze
- Manage, store and preserve
- Share and publish
- Data creation and deposit
- Data catalogs and registries
- Managing active data
- Data repositories and archives

Stages:
- Project Start (Application)
- Member Management
- Initial Setting
- Data Acquisition
- Data Management
- Data Analysis
- Paper writing
- Deposit with Supplemental Data

Statuses:
- Open
- Closed
- Not yet supported
Research Data Infrastructure for Open Science

**Discovery Platform**
- Linking Func between Article and Data
- Researcher and Research Project Identification and Management Func
- Data Exchange with International Discovery Service

**Discovery Service**
- **Metadata Management**
  - DOI
  - International Metadata Aggregator

**Data Depositor**
- Article
- Supplemental Data

**Data User**
- Exp/Data
- Exp/Store

**Journal Article**
- Subject Repository

**Research Data Repository**
- Private
- Shared
- Public

**Research Data Management System**
- Research Data Mng
- User Interface
- Access Control
- Metadata Mng

**RDM Platform**
- High Speed Access using SINET5
- Data Sharing Func using Virtual NW and ID Federation
- Effective Data Storage Switcher

**Publication Platform**
- Data oriented Self-Archiving Func
- Versioning and auto-Packaging Func
- User Dependent Personal Data Pseudonym Func

**GakuNin RDM**
- Storage Area for Long-term Preservation
Data Publication Platform

Basic Idea of New System

• Separate RDM Platform (Researcher) and Publication Platform (Librarian)
• Employ Micro Service Architecture which ensures Extensibility and Flexibility in Development and Operation

What will become of?

• Easy to Deposit Research Output through RDM Platform
• Possible to handle Large-scale Research Data with appropriate Interface
• Expandable to Subject Repository in addition to Institutional Repository
Data Discovery Platform

Crucial Backend Technology for integrating different types of entities and providing high UX as CiNii Research

Integrated Knowledge Base
Article • Data • Researcher • Project

Book (Disambiguation)
Dissertation (Disambiguation)
Journal Article (Disambiguation)

Publication Platform
CJP
DB Magazine
J-STAGE
Institutional Repository
How to realize our RDM Platform

- Establish corroborative development with the Western projects
- Our Functional Requirements
  - Connection to Institutional Storage Service
  - Institutional Level Control Panel (Storage Configuration, Plugin Configuration)
  - SAML Authentication and connect with VO Platform
  - Metadata Management Functionality
  - Easy deposit function to JAIRO Cloud
  - Mash-up with other scholarly information services in Japan

Advantage of OSF is its “Flexible” and “Extensible” architecture
New Service GakuNin RDM

Manage Research Data by Research Project

Connect Cloud Storage from Various Plugin

Share Research Data within Collaborators authn by ID Fed

NII: Frontend Service

University: Backend Storage

RDM Platform

Cloud Storage

Customize selectable Plugin depending on University Environment and Policy

Public Cloud (Provider DC)

Private Cloud (On-premise)

Public Cloud (Provider DC)

Default (minimum?) Storage by NII
RDM Platform v.s. Publication Platform

**Researcher**
- File Management
- Timestamp Proofing
- Long-term Preservation
- ...

**Librarian, Research Office**
- Metadata Management
- Data Publication
- DOI Registration
- ...

Develop Additional Functionalities in consistent with University Workflow