JPCOAR
オープンアクセスリポジトリ推進協会

Challenges in adopting the vocabularies in national and regional repositories networks: Japan

May 8, 2017

Controlled Vocabularies Open Session

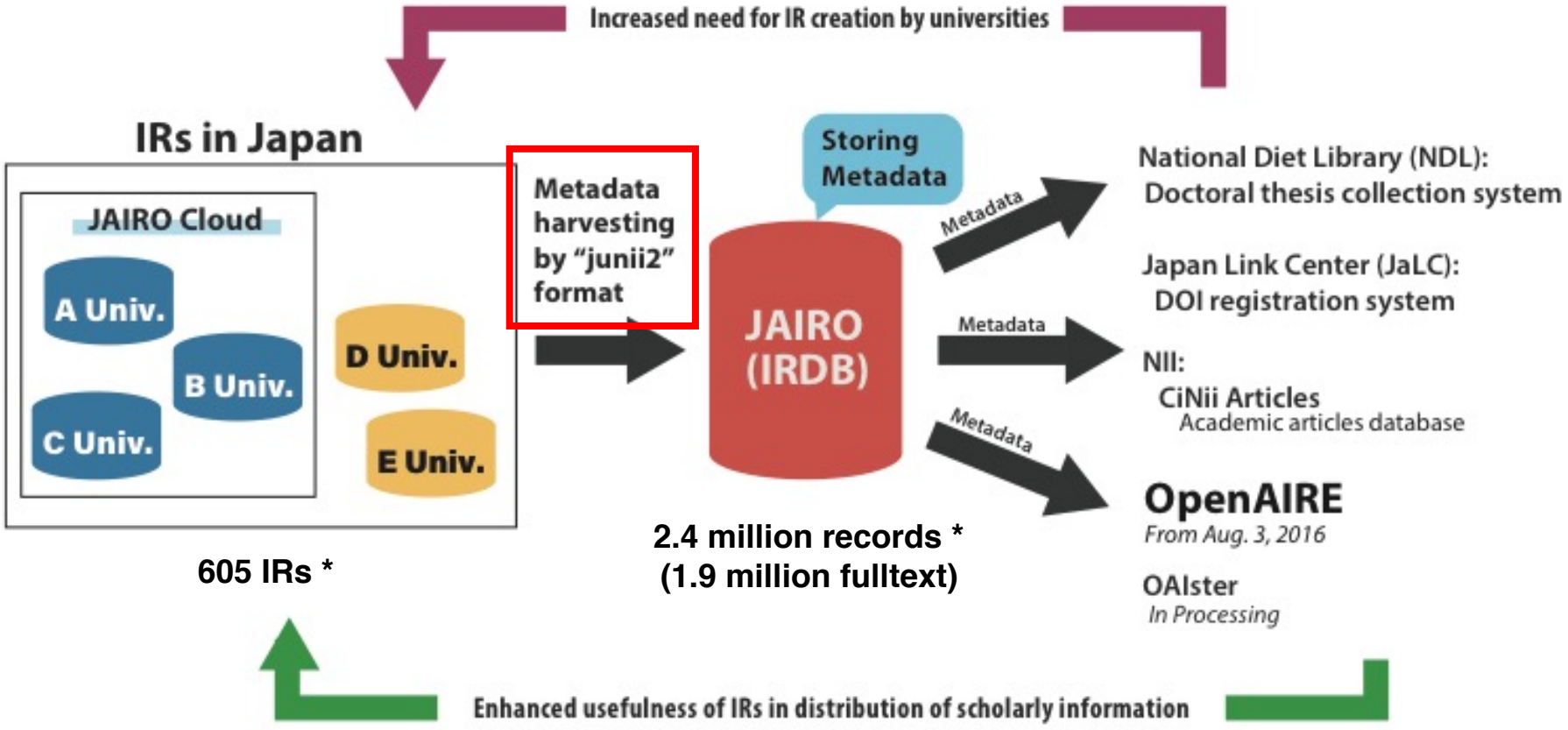
JPCOAR (Japan Consortium for Open Access Repository)

Tomoko Kagawa

About JPCOAR

- Established in July 2016 as a comprehensive consortium for open access in Japan
  - took over the responsibilities of DRF (Digital Repository Federation) and IRPC (Institutional Repositories Promotion Committee)
- Promote open access repository and facilitate communication among member institutions
- Consist of 474 academic institutions
- Set up 3 working groups and 4 task forces

Current Japanese metadata schema “junii2”



*** as of March 31, 2017**

Background of challenge for revising “junii2”

- Japanese regional schema
  - “junii2”(formerly “junii”) has been used for 15 years as a widespread adopted schema in Japanese community, however, it has not enough interoperability with international scholarly communication
- Unstructured schema architecture
  - It has issues with appropriate and detailed metadata description such as describing association between creator name, creator ID, affiliation...
- Incapability for scholarly changes
  - It has not capability to innovative enhancement for change of scholarly environment such as research data management

Major concepts of new schema

1. Focusing on interoperability with established international standards/schemas

Dublin Core	COAR	OpenAIRE	DataCite	The Bibliographic Ontology	RIOXX
[dc] title language publisher rights [dcterms] alternative temporal	resourceType accessRight	versionType	fundingReference funderIdentifier funderName awardNumber awardTitle identifier version date geolocation description	journal volume issue identifier pageStart pageEnd	apc

Major concepts of new schema

Regional elements are still left corresponding to Japanese scholarly environment

DC-NDL (National Diet Library Dublin Core Metadata Description)	JPCOAR schema
dateGranted	creator
degreeGrantor	contributor
degreeName	file
dissertationNumber	relation
	rightsHolder
	subject
	identifierRegistration

Major concepts of new schema

2. Extending language attribute to deal with multilingual metadata

Example:

<dc:title xml:lang="ja">情報爆発時代の研究基盤構想</dc:title>

<dc:title xml:lang="ja-Kana">ジョウホウ バクハツ ジダイ ノ ケンキュウ キバン コウソウ</dc:title>

<dc:title xml:lang="en">Research Project on Cyber Infrastructure for Information-explosion Era</dc:title>

Major concepts of new schema

3. Hierarchizing schema architecture (for elements)

   Creator, Contributor, Rights Holder, Relation, Geo Location, Funding Reference, Degree Grantor, File

```
<jpcoar:creator>
    <jpcoar:nameIdentifier nameIdentifierScheme="ORCID" schemeURI="https://orcid.org/0000-0001-0002-003">
      0000-0001-0002-0003</jpcoar:nameIdentifier>
    <jpcoar:creatorName xml:lang="ja">夏目, 漱石</jpcoar:creatorName>
    <jpcoar:creatorName xml:lang="en">Natsume, Soseki</jpcoar:creatorName>
    <jpcoar:creatorName xml:lang="ja-Kana">ナツメ, ソウセキ</jpcoar:creatorName>
    <jpcoar:familyName xml:lang="ja">夏目</jpcoar:familyName>
    <jpcoar:givenName xml:lang="ja">漱石</jpcoar:givenName>
    <jpcoar:creatorAlternative xml:lang="ja">夏目, 金之助</jpcoar:creatorAlternative>
    <jpcoar:affiliation>
        <jpcoar:nameIdentifier nameIdentifierScheme="ISNI"  schemeURI=
          "http://www.isni.org/isni/0000000121691048">0000000121691048</jpcoar:nameIdentifier>
        <jpcoar:affiliationName xml:lang="en">University of Tokyo</jpcoar:affiliationName>
    </jpcoar:affiliation>
</jpcoar:creator>
```

Major concepts of new schema

4. Adoption of funding information elements

```
<datacite:fundingReference>
  <datacite:funderIdentifier funderIdentifierType =“Crossref Funder”>
    http://dx.doi.org/10.13039/501100001691</datacite:funderIdentifier>
  <datacite:funderName xml:lang=“ja”>日本学術振興会</datacite:funderName>
  <datacite:funderName xml:lang=“en”>Japan Society for the Promotion of
    Science</datacite:funderName>
  <datacite:awardNumber awardURI=“https://kaken.nii.ac.jp/grant/KAKENHI-PROJECT-
    22227006/”>22227006</datacite:awardNumber>
  <datacite:awardTitle xml:lang=“ja”>ＲＮＡ修飾が支配する遺伝子発現調節機構の探究と高次
    生命現象</datacite:awardTitle>
  <datacite:awardTitle xml:lang=“en”>Post-transcriptional regulation associated with RNA
    modifications responsible for higher order biological processes</datacite:awardTitle>
</datacite:fundingReference>
```

Major concepts of new schema

5. Enhancement of elements for research data

- various identifiers

  datacite:identfier, jpcoar:relatedIdentifier,

  jpcoar:nameIdentifier(creator/affiliation)

- enhancement for rights description

  dc:rights, jpcoar:accessRights, jpcoar:rightsHolder, rioxxterms:apc

- additional elements/attribute

  datacite:version, datacite:geolocation, contributorType

Adopting vocabularies from established standards

Elements

- coar:resourceType
  (selected 38 vocabularies)
- coar:accessRight (now establishing)
  open access/metadata only access/embargoed access/restricted access
- openaire:versionType
  accepted/published/draft/submitted/updated
- rioxxterms:apc
  Paid/Fully waived/Not required/Partially waived/Not charged/ Unknown

Adopting vocabularies from established standards

Attributes

- Name Identifier Scheme of Creator/Rights Holder/ Affiliation/Degree Grantor

  Example) Creator/Rights Holder: e-Rad/ORCID/ISNI/VIAF

  Affiliation: kakenhi/ISNI/Ringgold/GRID

- Identifier Type of Resource Identifier/Related Identifier/ Journal Identifier/Registration Identifier/Funder Identifier

  Example) Resource Identifier: DOI/HDL/URI

- Other attributes

  Contributor Type/Description Type/Date Type/Subject Scheme/ Relation Type/File Object Type

Compliant with COAR "Resource Type" vocabularies

Metadata Field Set	Vocabulary
Article	conference paper
Article	data paper
Article	departmental bulletin paper
Article	editorial
Article	journal article
Article	article
Book	book
Book	book part
Cartographic Material	cartographic material
Cartographic Material	map
Conference Object	conference object
Conference Object	conference poster
Dataset	dataset
Image	image
Image	still image
Image	moving image
Image	video
Lecture	lecture
Patent	patent

Metadata Field Set	Vocabulary
Report	report
Report	research report
Report	technical report
Report	policy report
Report	report part
Report	working paper
Sound	sound
Thesis	thesis
Thesis	bachelor thesis
Thesis	master thesis
Thesis	doctoral thesis
Multiple	interactive resource
Multiple	learning material
Multiple	musical notation
Multiple	research proposal
Multiple	software
Multiple	technical documentation
Multiple	workflow
Multiple	other

Why we comply with COAR “Resource Type” vocabularies?

- We put priority on complying with established standards to improve interoperability with international networks.
- COAR “Resource Type” is one of the most influential vocabulary and has strong potential to improve Japanese resources sharing to international frameworks.
- However, we had to maintain compatibility with current 14 vocabularies and considered the impact of repository systems so we selected 38 vocabularies and categorized flat structure.
- Also, JPCOAR schema still contains an alternative label of COAR vocabulary and regional vocabulary.

Exception case

Concept URI
(“journal article” URI)

Case1: journal article

<coar:resourceType rdf:resource=“http://purl.org/coar/resource_type/c_6501”>
journal article</coar:resourceType>

<coar:resourceType rdf:resource=“http://purl.org/coar/resource_type/c_6501”>
departmental bulletin paper</coar:resourceType>

<coar:resourceType rdf:resource=“http://purl.org/coar/resource_type/c_6501”>
article</coar:resourceType>

Terms
(including alternative label or regional vocabularies)

Concept URI
(other” URI)

Case2: other

<coar:resourceType rdf:resource=“http://purl.org/coar/resource_type/c_1843”>
learning material </coar:resourceType>

Next steps

Implementation schedule

- Finalize JPCOAR schema and disseminate to JAIRO, JAIRO Cloud and other major Japanese repository systems
  - Planned: FY2017 JAIRO, FY2018- repository systems
  - Japanese centralized sharing repository system “JAIRO Cloud”, implemented 485 institutions, is enabled us to achieve this innovative schema revision

Additional objectives

- Digital collections management (e.g. digitalized handwritten classical manuscripts)
- Learning materials