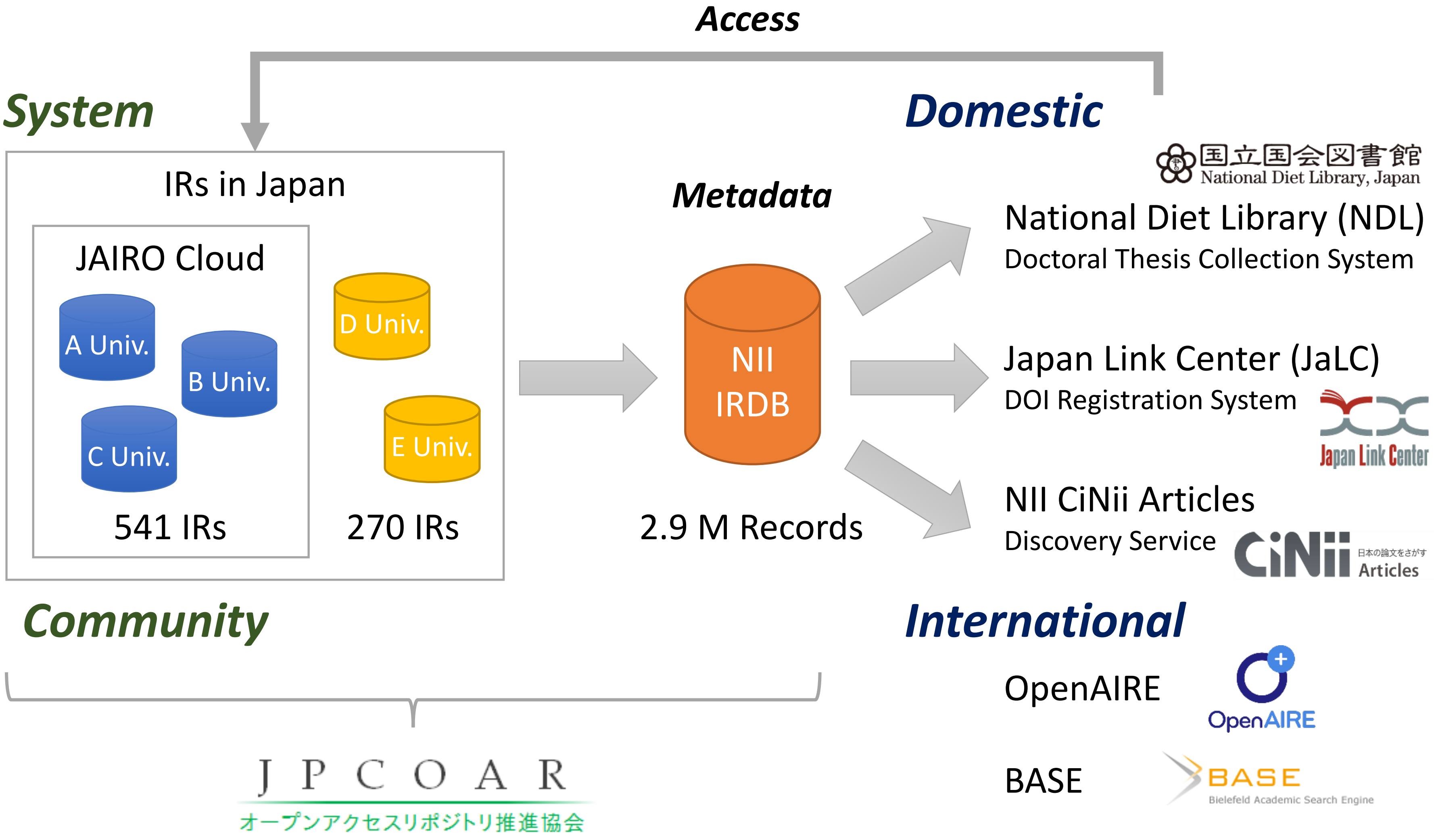
JP COAR and NII supporting local services and engaging internationally

Kazu YAMAJI, NII

Asia OS Annual Meeting 2019, Dhaka, Bangladesh
Date -6-7 March 2019

Japanese IR Network



Access

System

IRs in Japan

JAIRO Cloud

A Univ.
B Univ.
C Univ.

541 IRs

D Univ.
E Univ.

270 IRs

Metadata

NII IRDB

2.9 M Records

Domestic

国立国会図書館
National Diet Library, Japan

National Diet Library (NDL)
Doctoral Thesis Collection System

Japan Link Center (JaLC)
DOI Registration System

Japan Link Center

NII CiNii Articles
Discovery Service

CiNii 日本の論文をさがす Articles

International

OpenAIRE

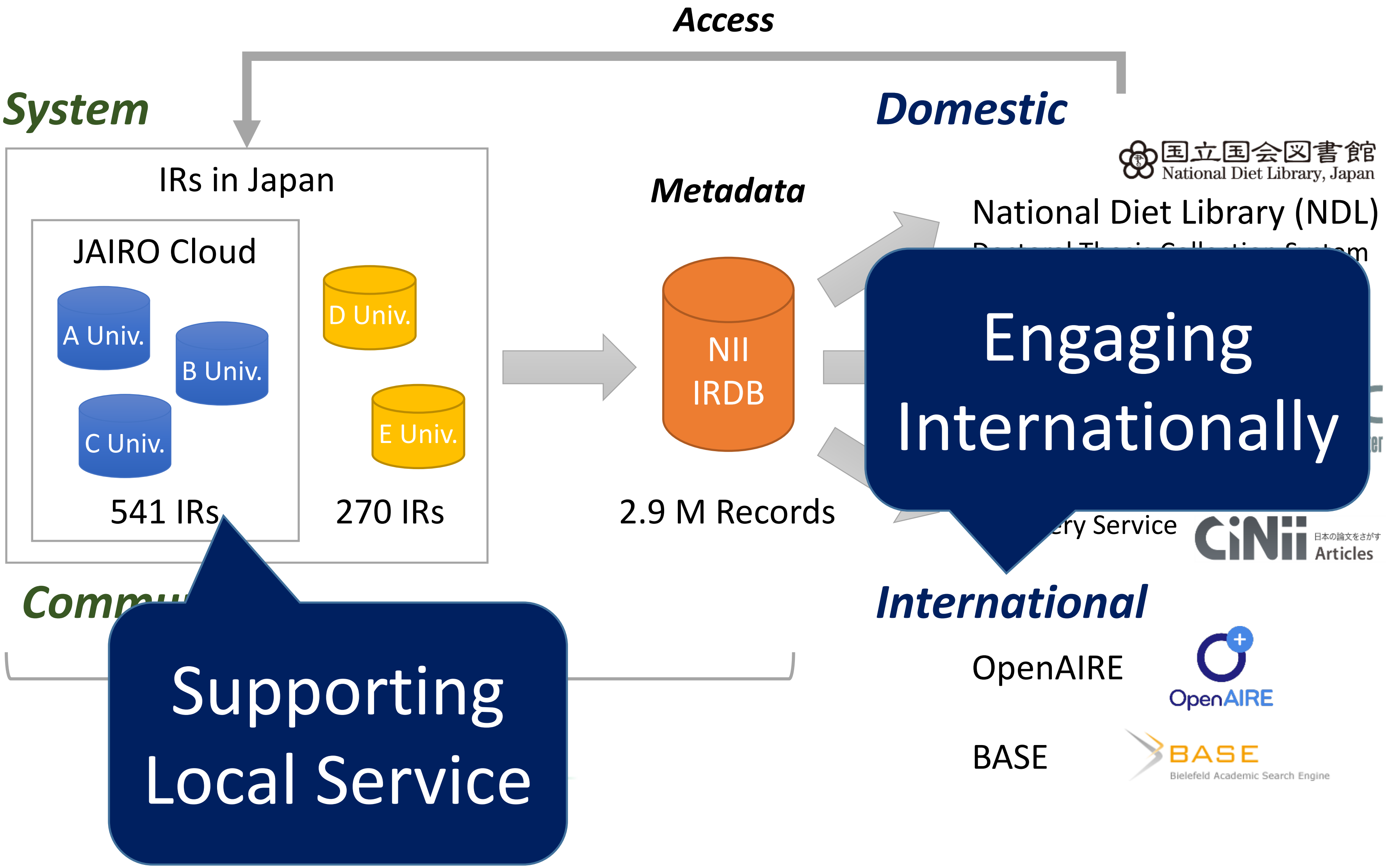
BASE

BASE Bielefeld Academic Search Engine

Community

JPCOAR
オープンアクセスリポジトリ推進協会

Japanese IR Network



Access

System

IRs in Japan

JAIRO Cloud

A Univ.

B Univ.

C Univ.

541 IRs

D Univ.

E Univ.

270 IRs

Metadata

NII
IRDB

2.9 M Records

Domestic

国立国会図書館
National Diet Library, Japan

National Diet Library (NDL)

Doctoral Thesis Collection System

ery Service

CiNii 日本の論文をさがす Articles

International

OpenAIRE

BASE

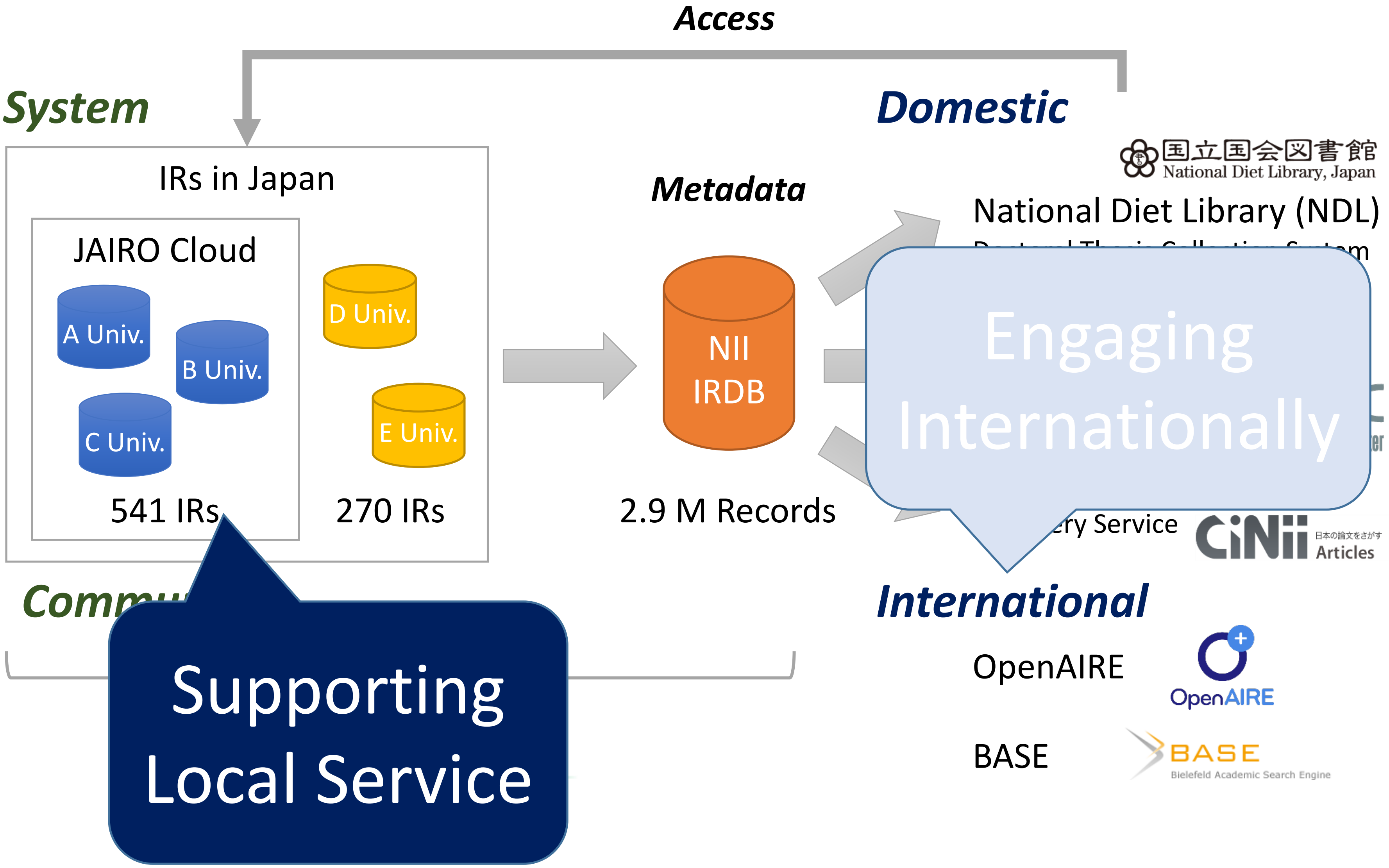
BASE Bielefeld Academic Search Engine

Commu

Engaging Internationally

Supporting Local Service

Japanese IR Network



Access

System

IRs in Japan

JAIRO Cloud

A Univ.

B Univ.

C Univ.

541 IRs

D Univ.

E Univ.

270 IRs

Metadata

NII
IRDB

2.9 M Records

Domestic

国立国会図書館
National Diet Library, Japan

National Diet Library (NDL)

Doctoral Thesis Collection System

ery Service

CiNii 日本の論文をさがす Articles

International

OpenAIRE

OpenAIRE

BASE

BASE
Bielefeld Academic Search Engine

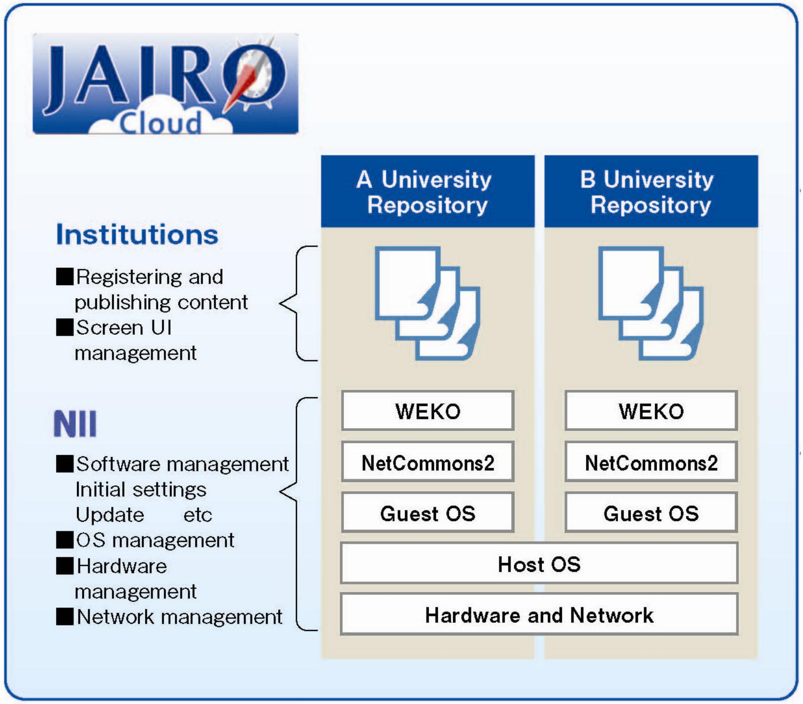
Commu

Supporting
Local Service

Engaging
Internationally

JAIRO Cloud

- Background
  - Limited resources and less technical knowledge hamper implementation of IR especially in small universities.
  - JAIRO Cloud provides a shared instance of IR system on the virtual server hosted by NII since April 2012.
- Service Architecture



JAIRO Cloud

Institutions

- Registering and publishing content
- Screen UI management

NII

- Software management
  Initial settings
  Update etc
- OS management
- Hardware management
- Network management

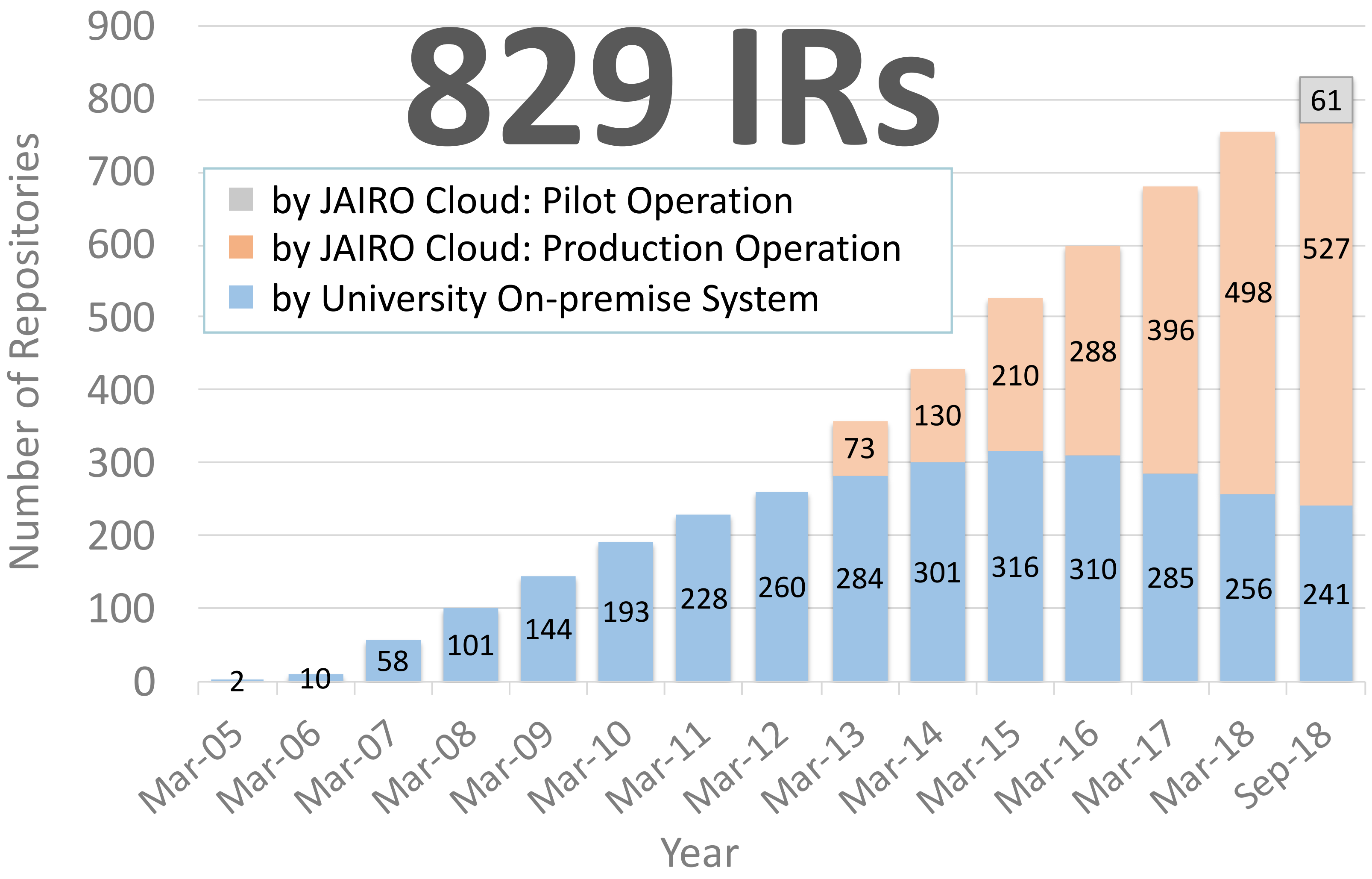
A University Repository	B University Repository
WEKO	WEKO
NetCommons2	NetCommons2
Guest OS	Guest OS

Host OS

Hardware and Network

みんぱくリポジトリ
IDE-JETRO
ARRIDE
つくばリポ
Tulips
リポジトリ
REPOSITORY of SHINSHU

Number of Institutional Repository in Japan



From Open Access
To **O**pen **S**cience

- **Current System WEKO2**
  - Journal Article Repository
  - Add Functions more and more

→ ✗ Research Data

Trend: Software Development

Name	Version	License	Version control	Issue tracker	CI	Mainling list	Chat	Vagrant	Docker	cloud deploy
DSpace6	6.2	BSD 3-Clause	GitHub	JIRA	Travis CI	Yes	IRC, Slack	Yes	-	-
DSpace7	-	BSD 3-Clause	GitHub	JIRA	Travis CI	Yes	IRC, Slack	-	-	-
EPrints3.3	3.3.16	GPL v.3	GitHub	GitHub	-	Yes	Gitter	Yes	-	-
Eprints3.4	3.4.0	GPL v.3	GitHub	GitHub	-	Yes	Gitter	-	-	-
Fedora 4	4.7.5	Apache 2.0	GitHub	JIRA, GitHub	Travis CI	Yes	IRC, Slack	Yes	-	-
HUBzero	2.1.15	GPL v.2	GitHub	Hubzero	Travis CI	-	-	Yes	-	-
Invenio3	3.0.0rc1	MIT	GitHub	GitHub	Travis CI	-	Gitter	Yes	Yes	OpenShift template
Zenodo	-	GPL v.2	GitHub	GitHub	Travis CI	-	Gitter	-	Yes	-
Hyrax	2.1.0.rc1	Apache 2.0	GitHub	GitHub	Travis CI	Yes	Slack	Yes	-	-
Hyku	v1.0.0.beta2	Apache 2.0	GitHub	GitHub	Travis CI	Yes	Slack	Yes	Yes	AWS template
Samvera	-	Apache 2.0	GitHub	GitHub	Travis CI	Yes	Slack	Yes	-	-
Islandra	7.x-1.10	GPL v.3	GitHub	JIRA	Travis CI	Yes	IRC	Yes	-	-
Islandora CLAW	-	GPL v.2	GitHub	GitHub	Travis CI	Yes	IRC	Yes	-	-
Dataverse4	4.8.6	Apache 2.0	GitHub	GitHub	Travis CI	Yes	IRC	Yes	Yes	OpenShift template
WEKO2	2.4.2	BSD 2-Clause	Bitbuket	-	-	Yes	-	Yes	Yes	-

Use of GitHub,Travis CI and Vagrant are de facto standards

Use of Slack, Gitter are spreading

Invenio, Hyku and Dataverse are providing cloud deploy template using AWS or OpenShift

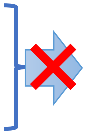
Trend: System Architecture

Name	Lang	Framework	Base	DataBase	O/Rmapping	Search	Full Text	REST API	Messaging	Hosting
DSpace6	Java	Spring	-	PostgreSQL, Oracle	Hibernate	Solr	Yes	Yes	-	Yes
DSpace7	Java	Spring	-	PostgreSQL, Oracle	Hibernate	Solr		Yes	-	-
EPrints3.3	Perl	-	-	MySQL		Internal	Yes	Yes	-	Yes
Eprints3.4	Perl	-	-	MySQL		Internal	Yes	Yes	-	Yes
Fedora 4	Java	Spring	ModeShape	ModeShape		Internal	-	Yes	-	-
HUBzero	PHP	Joomla!	-	MySQL		Internal	-	-	-	Yes
Zenodo	Python	Flask	Invenio3	SQLite, PostgreSQL, MySQL	SQLAlchemy	Elasticsearch	-	Yes	Celery/Rabbitmq	-
Invenio3	Python	Flask	-	SQLite, PostgreSQL, MySQL	SQLAlchemy	Elasticsearch	-	Yes	Celery/Rabbitmq	-
Hyrax	Ruby	Rails	Samvera	SQLite, PostgreSQL, MySQL	ActiveFedora	Solr	Yes	Yes	Sidekiq/Redis	-
Hyku	Ruby	Rails	Hyrax	SQLite, PostgreSQL, MySQL	ActiveFedora	Solr	Yes	Yes	Sidekiq/Redis	Yes
Samvera	Ruby	Rails	Fedora4	-	ActiveRecord	Solr	-	Yes	-	-
Islandra	PHP	Drupal7	Fedora3	MySQL	Doctrine	Solr	Yes	Yes	-	-
Islandora CLAW	PHP	Drupal8	Fedora4	MySQL	Doctrine	Solr	-	Yes	API-X(Karaf), ActiveMQ	-
Dataverse4	Java	primefaces	-	PostgreSQL	Hibernate	Solr	Yes	Yes	-	-
WEKO2	PHP	NetCommons2	-	MariaDB		Mroonga	Yes	-	-	Yes

Software development has become complicated
However, automating build/test/deployment with CI/CD tool can help developers
Container orchestration software will improve maintenance and operability

WEKO3

- **Current System WEKO2**
  - Journal Article Repository
  - Add Functions more and more



Research Data

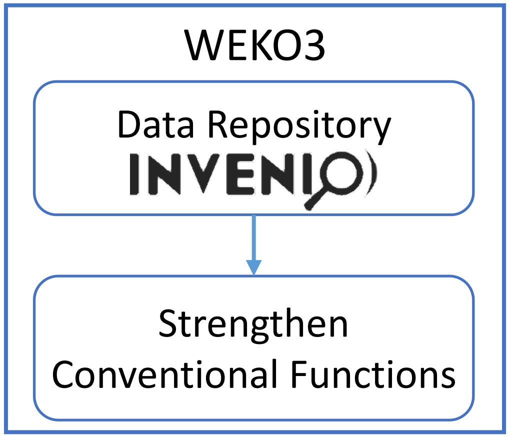
WEKO3

- **Current System WEKO2**
  - Journal Article Repository
  - Add Functions more and more

✕ Research Data

○

- **New System WEKO3**
  - Based on Invenio3 which is originally focused as Data Repository
  - Integrate WEKO2 Functions into Invenio3


WEKO3
Data Repository
INVENIO
Strengthen Conventional Functions

Realize New Publication Platform based on sophisticated Invenio3 Architecture
(Invenio3 = our RDM Platform in Architecture)

- Effective Development and Operation
- Domain Use-case by Extensibility

An Overview of INVENIO

- **Flexible** – Invenio is a framework for building large scale digital repositories: RDM, Media Archive, Institutional Repository, Search Engine, ILS, Brokers.
- **Scalable** – Handles 100M records and petabytes of files and integrates with any storage system.
- **Powerful search** – with the full capabilities of Elasticsearch such as geospatial queries.
- **Born for the web** – REST APIs everywhere.
- **Data model agnostic** – Natively based on JSON from storage to indexing.
- **Best of Open Source** - Built on top of the best: Elasticsearch, PostgreSQL/MySQL, Flask & Python.

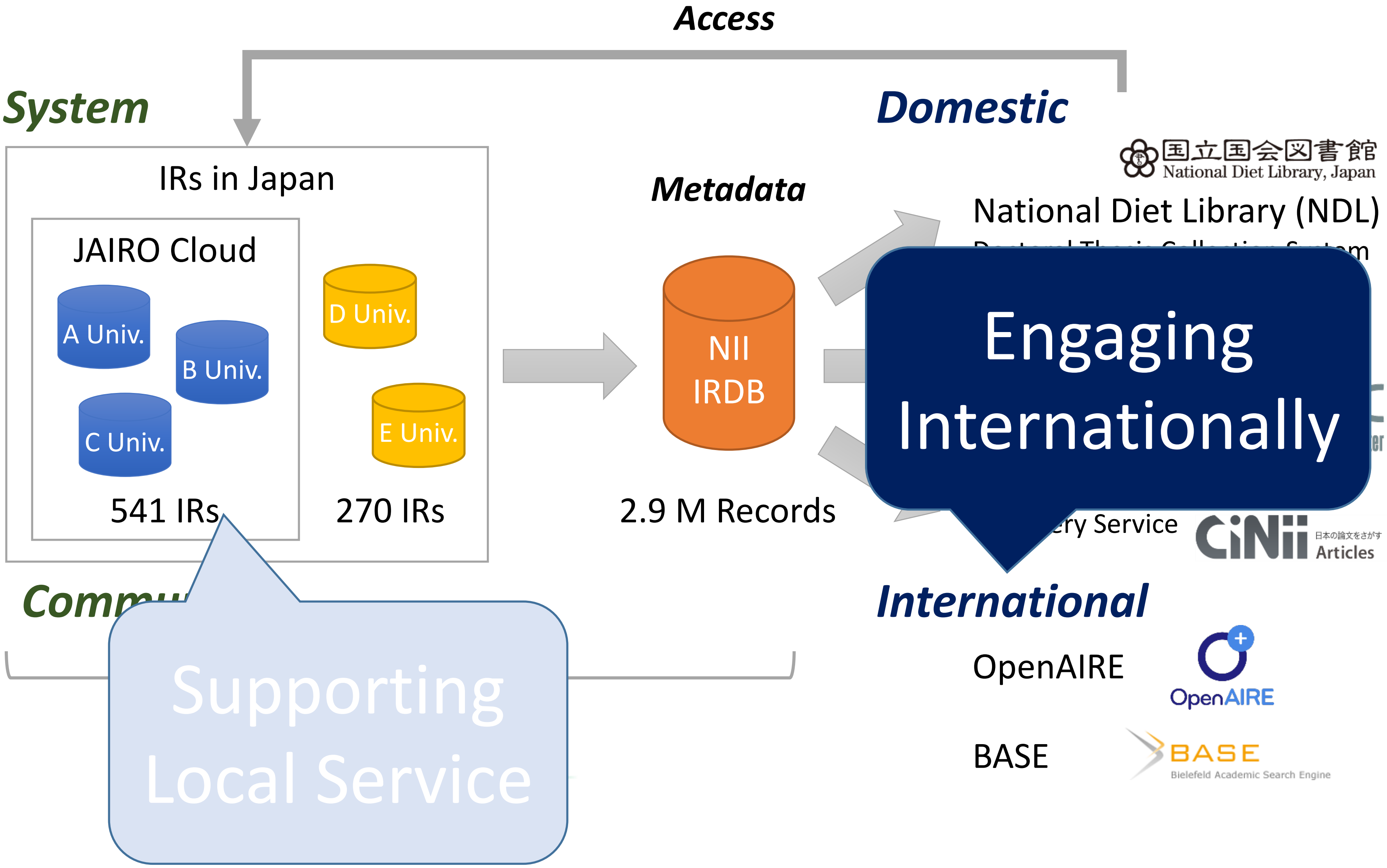
WEKO3

- A flavor of Invenio v3 for Institutional Repository
- Main features of WEKO3 (including future plan)
  - Integrate WEKO2 functions for IR based on Practical Requests from JAIRO Cloud Users
  - Flexible **metadata editor** and OAI-PMH provider feature
  - Flexible **workflow** feature for ingest workflow
  - **Cloud storage integration** for storing research data
  - **Multitenancy** features for providing perspective on needs of institution and the JAIRO Cloud service

Plan for NGR Implementation

	Technologies	Release plan
CERN is developing	ETag	June 2018
	ORCID	June 2018
	Sitemap	June 2018
	IIIF (Image API)	November 2018
CERN will develop	Signposting	Early in 2019
	COUNTER	Middle in 2019
	SUSHI	Middle in 2019
	WebSub	Later in 2019
NII will develop	ResourceSync	December 2018
	SWORD v3	December 2018

Japanese IR Network



Access

System

IRs in Japan

JAIRO Cloud

A Univ.

B Univ.

C Univ.

541 IRs

D Univ.

E Univ.

270 IRs

Metadata

NII IRDB

2.9 M Records

Domestic

国立国会図書館 National Diet Library, Japan

National Diet Library (NDL)

Doctoral Thesis Collection System

ery Service

CiNii 日本の論文をさがす Articles

International

OpenAIRE

OpenAIRE

BASE

BASE Bielefeld Academic Search Engine

Commu

Engaging Internationally

Supporting Local Service

Metadata

- junii2
  - Current Schema
    - Since 2010
    - Current Version 3.1
  - **Not designed for**
    - Layered Elements
    - Research Data
    - **Interoperability with International Standards**
- JPCOAR
  - Focusing on interoperability with established international metadata standards
  - Adopting major identifiers (Crossref Funder id, ISNI, ORCID etc)
  - Monitoring compliance with Open Access policies
  - Hierarchizing some elements (creator, fundingReference etc)
  - Enhancement of elements for research data

JPCOAR Schema Properties

1. dc:title

2. jpcoar:creator

3. jpcoar:contributor

4. jpcoar:accessRights

5. rioxxterms:apc

6. dc:rights

7. jpcoar:rightsHolder

8. jpcoar:subject

9. datacite:description

10.dc:publisher

11.datacite:date

12.dc:language

13.coar:resourceType

14.datacite:version

15.openaire:versionType

16.datacite:identifier

17.jpcoar:identifierRegstration

18.jpcoar:relation

19.dcterms:temporal

20.datacite:geolocation

21.datacite:fundingReference

22.bibo:identifier

23.bibo:journal

24.bibo:volume

25.bibo:issue

26.bibo:pageStart

27.bibo:pageEnd

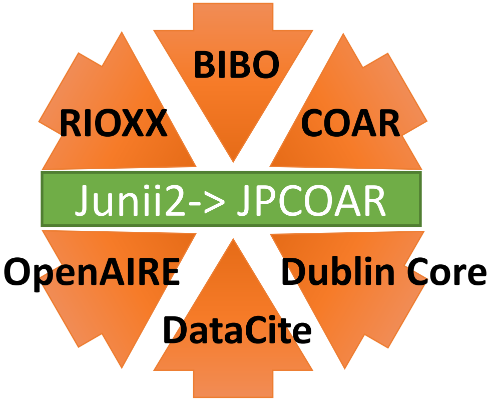
28.jpcoar:dissertationNumber

29.jpcoar:degreeName

30.jpcoar:dateGranted

31.jpcoar:degreeGrantor

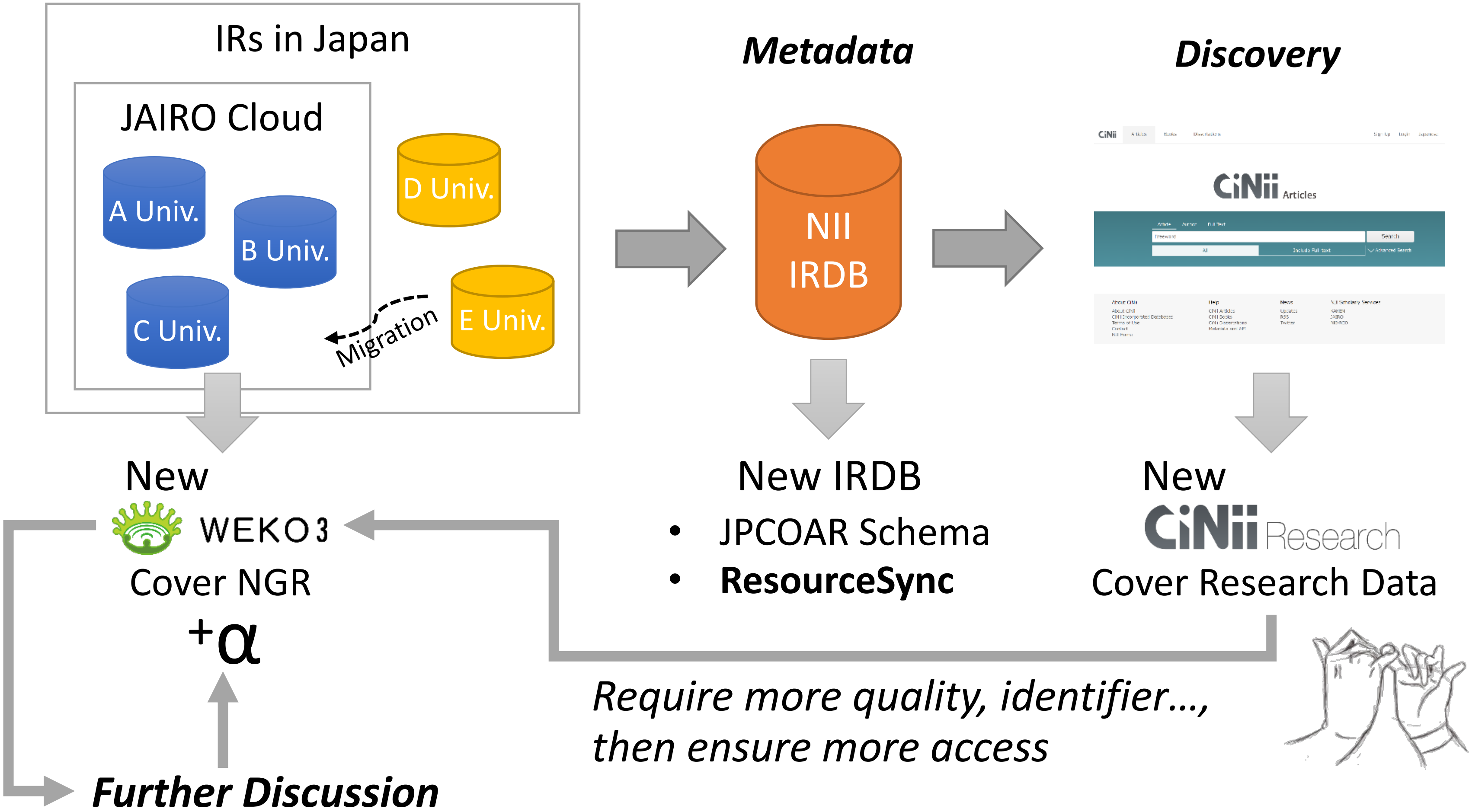
32.jpcoar:file



Example of hierarchized elements

```
<jpcoar:creator>
 <jpcoar:nameIdentifier nameIdentifierScheme= “ORCID”
  schemeURI= “http://orcid.org/0000-0003-2384-2166” >
  0000-0003-2384-2166</jpcoar:nameIdentifier>
 <jpcoar:creatorName xml:lang= “ja” > 大隅 , 良典
 </jpcoar:creatorName>
 <jpcoar:creatorName xml:lang= “en” >Ohsumi, Yoshinori
 </jpcoar:creatorName>
 <jpcoar:affiliation>
 <jpcoar:nameIdentifier nameIdentifierScheme= “ISNI”
  schemeURI= “http://www.isni.org/0000000121792105” >
  0000000121792105</jpcoar:nameIdentifier>
 <jpcoar:affiliationName xml:lang= “en” >
  Tokyo Institute of Technology
 <jpcoar:affliationName>
 </jpcoar:affiliation>
<jpcoar:creator>
```

Produce Positive Feedback



IRs in Japan

JAIRO Cloud

A Univ.

B Univ.

C Univ.

D Univ.

E Univ.

Migration

Metadata

NII IRDB

Discovery

New WEKO3

Cover NGR

+α

New IRDB

- JPCOAR Schema
- **ResourceSync**

New CiNii Research

Cover Research Data

Require more quality, identifier…,
then ensure more access

Further Discussion

- How to think about ORCID National Consortium?
- How to collect Evidence Data in accordance with Institutional Policy?
- …