

メタデータ

機関リポジトリ新任担当者研修
平成27年2月20日・東北学院大学

静岡大学附属図書館 杉山智章

機関リポジトリの“メタデータ”

メタデータとは…

「データに関するデータ」

ここでは…

メタデータ ≡ コンテンツの書誌情報

メタデータ入力画面

The screenshot shows a web browser window titled "SURE: Shizuoka University REpository: アイテムの記述 - Internet Explorer". The address bar shows "http://ir.lib.shizuoka.ac.jp/submit". The form contains the following fields and labels:

- タイトル (Title)
- その他のタイトル (Other titles) with a "フィールド追加" (Add field) button. Label: 別言語のタイトルやタイトルのヨミ
- 著者 (Author) with a "フィールド追加" button. Label: (和文なら日本語、欧文なら英語) 左欄:「姓」中欄:「名」右欄:「ID」姓 名 研究者ID
- 言語 (Language) dropdown menu. Label: コンテンツ本体の言語
- 資源タイプ (Resource type) dropdown menu. Label: 資源タイプについて(PDF)
- Relations (Relations) dropdown menu. Label: 博士論文→「NDLBIBID」, 特許文献→「KAKEN」, それ以外るとき→「NCID, PMID, DOI, NAID」をさがす それ以外にリンクするページがあるときは「URI」
- 掲載図書名 (Book title) with a label: 本タイトルのみ(アイテムが圖書の一部のときのみ使用)
- ISBN (ISBN) with a label: International Standard Book Number
- 掲載雑誌名 (Journal title) with a label: 本タイトルのみ(アイテムが雑誌の一部のときのみ使用)
- ISSN (ISSN) with a label: International Standard Serial Number
- E-ISSN (E-ISSN) with a label: (Online) International Standard Serial Number (ISSNと同じときは不要)
- 巻 (Volume)
- 号 (Issue)
- 開始ページ (Start page)

- リポジトリに論文等のコンテンツを登録する際には、ファイルとメタデータが必要
- 何のためにメタデータがあるのか
- どのようなことに気をつけて、メタデータを管理するのか

内部メタデータ

- 登録したメタデータは、アイテムページに、コンテンツの書誌情報として表示される。
- このように、リポジトリ内部で管理され使用されるメタデータを「内部メタデータ」と呼ぶ。

The screenshot shows a web browser window displaying the Shizuoka University REpository item page. The page title is "クラウドを使用した静岡大学附属図書館のシステム構築". The URL is <http://hdl.handle.net/10297/6885>. The page contains a table of files, a metadata section, and navigation links.

ファイル	記述	サイズ	フォーマット
121101001.pdf		372.74 kB	Adobe PDF

このアイテムの引用には次の識別子を使用してください:
<http://hdl.handle.net/10297/6885>

このアイテムのファイル:

タイトル: クラウドを使用した静岡大学附属図書館のシステム構築
その他のタイトル: Practical approaches to the cloud at Shizuoka University Library

著者: 杉山, 智章
森内, 文
高橋, 里江
森部, 圭亮

掲載誌名: 情報の科学と技術
出版者: 情報科学技術協会
巻: 62
号: 11
開始ページ: 478
終了ページ: 483
出版日付: 2012-11-01
NDC: 013
ISSN: 09133801 [静大OPAC](#)

NII書誌ID: AN10005857 [静大OPAC](#) [CiNii Books](#)

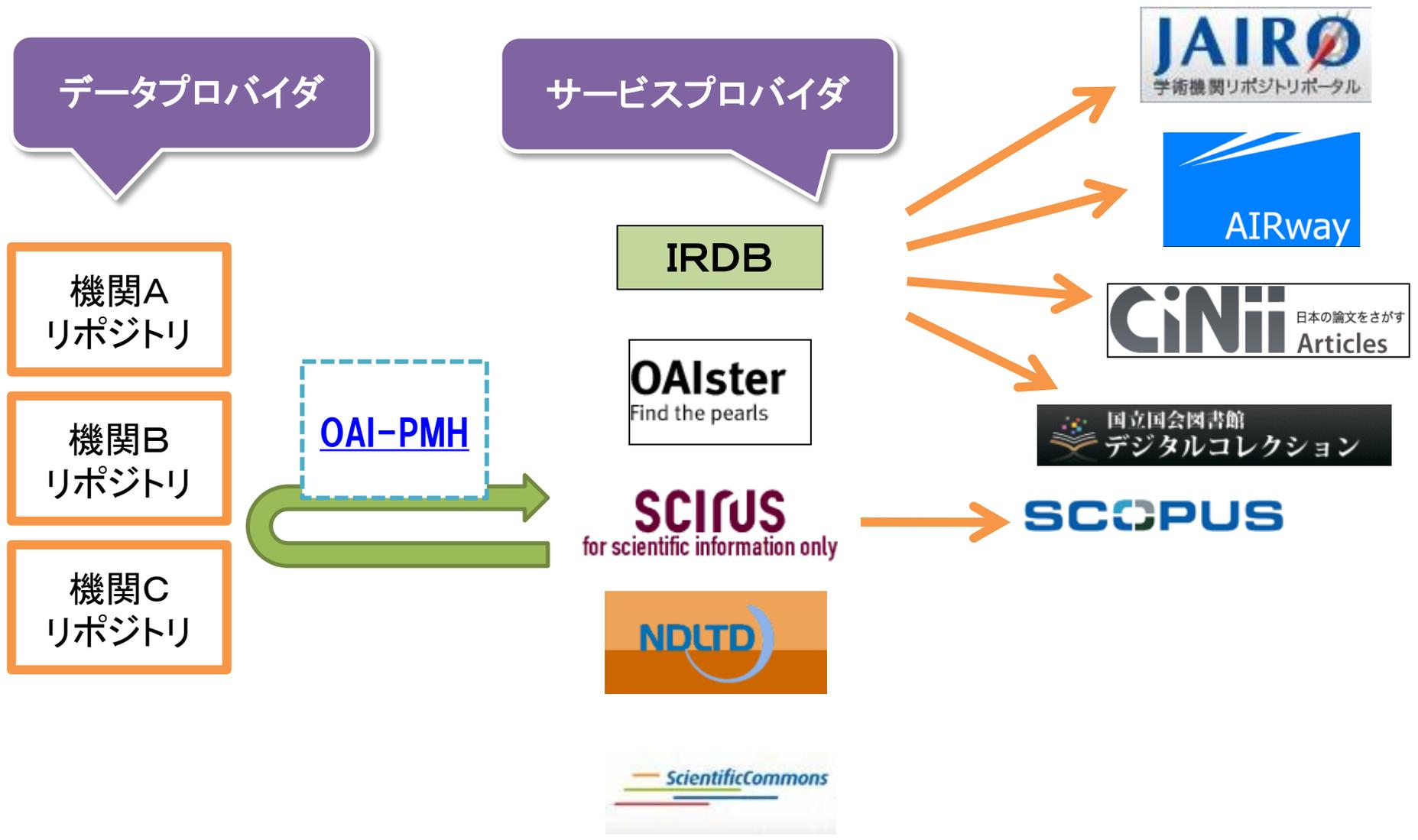
関連リンク: <http://www.infosta.or.jp/journal/journal.html>
バージョン: publisher
出現コレクション: 50. 雑誌論文・記事 (Journal Article, Article, Preprint)

[アイテムの詳細レコードを表示する](#)

外部メタデータ

- 登録されたコンテンツに、さまざまな手段を通してアクセスされやすくする、
＝可視性 (visibility: ビジビリティの向上)
つまり、さまざまなサービスと連携できることが求められる。
- そのために、リポジトリはOAI-PMHという仕組みでメタデータを提供できるようになっている。
- そのとき、外部のサービスに提供するメタデータを、「外部メタデータ」と呼ぶ。

OAI-PMHイメージ



OAI-PMHとは？

- 複数のリポジトリのメタデータを収集し、それに基づいたサービスを提供するために開発されたデータ提供・収集用のプロトコル(通信手段)
- HTTP通信を使用しているため、特定のソフトウェアなどに依存せず、すべてのリポジトリに備わっている機能
- リポジトリを公開すること＝OAI-PMHで外部にメタデータ提供すること

クローリングとハーベスト

- クローリング

- GoogleやGoogle Scholarなどの検索エンジンが、HTTPで各ページの内容を読み取り、インデクシング(索引化)する。

- ハーベスト

- サービスプロバイダが、OAI-PMHで、専用のURLからメタデータを一括して刈り取る。

- サービスプロバイダは、さらに別のサービスにデータを提供し、広がっていくこともある。

NACSIS-CATとの比較

NACSIS-CAT



利用者



ダウン
ロード

入力



ローカルDB

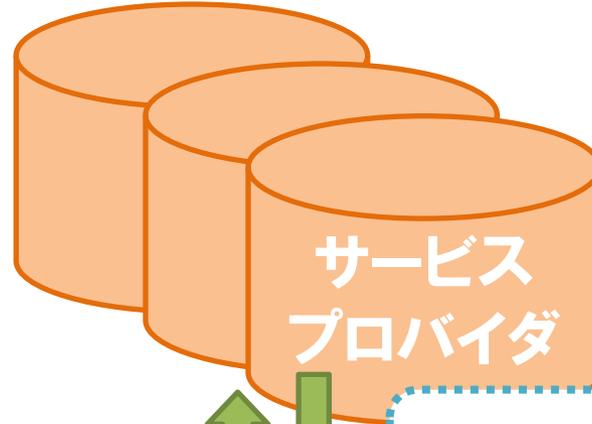


登録者

OAI-PMH



利用者



サービス
プロバイダ

これを実現するのが
OAI-PMH



リポジトリ

入力

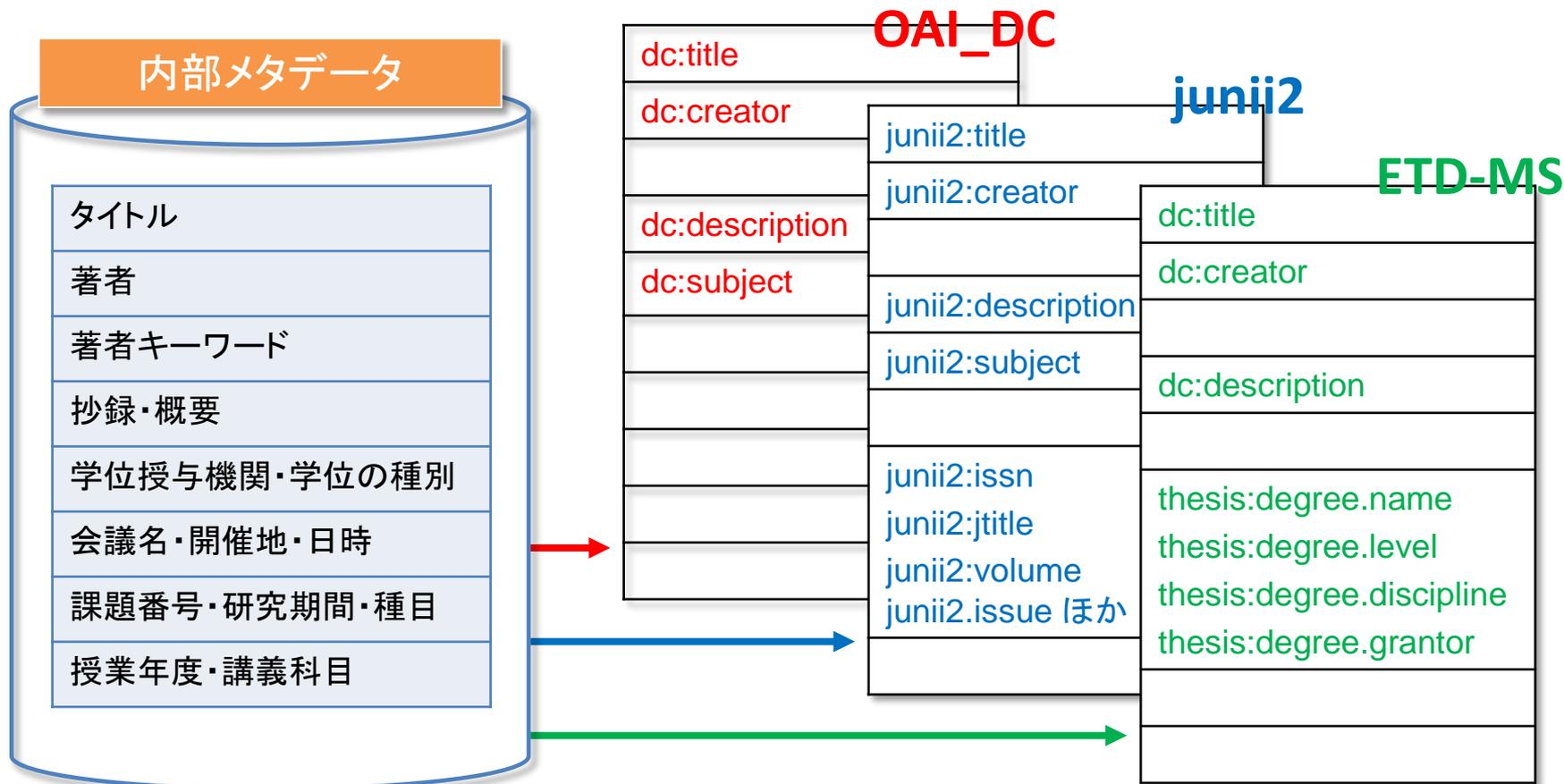


登録者

OAI-PMHメタデータフォーマット

- OAI_DC
 - ダブリン・コアの基本15要素項目
 - OAI-PMHで必須であり、世界標準
- junii2
 - NII等とのメタデータ連携で使用
 - 日本国内のほとんどのリポジトリが対応
- ETD-MS
 - NDLTD (Networked Digital Library of Theses and Dissertations) で使用
 - 学位論文用
- この他にも世界中でさまざまなフォーマットが使用されている

クロスウォーク



junii2バージョン

- ver. 1.0 (H18)
 - dc、dcterms、OpenURL (for Journal Articles)などの要素を取り入れている。textversion、fullTextURLなどが特徴的。
- ver. 2.0 (H23)
 - creator等に著者IDを追加(実証実験)
- ver. 3.0 (H25)
 - 学位規則改正(H25.4.1施行)、JaLC DOI等対応の拡張
- ver. 3.1 (H26)
 - 学位授与番号の改訂

junii2ガイドライン

junii2の各要素の説明と使用例が書いてある

<http://www.nii.ac.jp/irp/archive/system/junii2.html>

メタデータは変換可能なことが重要

- NIIサービスにとって扱いやすい
- OpenURLに変換できる
- OAI_DC(DRIVERガイドライン)にダムダウン(簡素化)できる
- 書誌情報(SIST等)として適正

例) 著者(creator)

junii2ガイドライン3.1 (抜粋)

個人名について、「姓,△名」とする。肩書き等は記入しない。

1作成者について、出力する表記形は1つのみで、同一作成者の異表記形を出力しない。

DRIVERガイドライン2.0 (抜粋)

逆の語順の名前を使用し、構文は「姓,△頭文字(名)△接頭辞」とする。

たとえばJan Hubert de Smit は次のようになる:

```
<dc:creator>Smit, J.H.(John) de</dc:creator>
```

Using Dublin Core - The Elements (抜粋)

Creators should be listed separately, preferably in the same order that they appear in the publication. Personal names should be listed surname or family name first, followed by forename or given name. When in doubt, give the name as it appears, and do not invert.

資料種別(NIIttype)

各要素の使用方法は資料種別ごとに考える

- ジャーナルアーティクル系
 - Journal Article, Departmental Bulletin Paper, Article, Preprint
- 学位論文
 - Thesis or Dissertation
- 書籍系
 - Conference Paper, Book, Technical Report, Research Paper
- その他
 - Presentation, Learning Material, Data or Dataset, Software, Others

CiNii Article連携

visibility: ビジビリティの向上



- Journal Article, Departmental Bulletin Paper, Article
- fullTextURLがある
- URIを変更しない

ジャーナルアーティクル系

- 掲載誌情報
 - 参考) OpenURL
- ID (pmid、doi、NAID、ichushi)
 - 論文を特定しやすく
- 著者版フラグ
 - “author” or “publisher”
はっきりと示す

OpenURL

```
http://xxx.xx.xx/xxx?.....  
&rft_id=info:doi/10.1126/science.275.5304.1320  
&rft_id=info:pmid/9036860  
&rft.issn=00368075  
&rft.eissn=10959203  
&rft.atitle=Isolation of a Common Receptor  
&rft.jtitle=Science  
&rft.aulast=Bergelson  
&rft.aufirst=Jeffrey  
&rft.date=1997  
&rft.volume=245  
&rft.issue=5304  
&rft.spage=1320  
&rft.epage=1323
```

国立国会図書館による 博士論文本文の自動収集



- NIIttypeが”Thesis or Dissertation”
- textversionが”ETD”でfulltextURIがある
- dateofgrantedが”2013-04-01”以降

学位論文

- 博士論文情報
 - 博士論文番号
科研費機関番号(5桁)+[甲|乙|*]+第*+報告番号+号
- textversion
 - 博士論文全文は”ETD”
修士、卒業論文には”ETD”は使用しない
- creator
 - 指導教員はcreatorにしない(使用例は少ないが contributorを使用できる)

ETD-MS

- NDLTDで策定された学位論文のためのメタデータフォーマット
- 基本はDC(ダブリン・コア)と同じだが拡張要素あり

主な拡張要素	説明
thesis.degree.name	Junii2:degreeNameに相当
thesis.degree.level	Doctoral Masters Undergraduate
thesis.degree.discipline	the name of a program or department
thesis.degree.grantor	Junii2:grantorに相当

- いつでも対応できるように、内部メタデータにデータを保持するとよい

書籍系 & その他

- 登録件数が少なく細かい標準化はこれから
- 書籍系
 - 図書1冊か章(Chapter)単位なのか注意
 - ISBN
- Data or Dataset、Software
 - 研究データのjunii2使用方法が検討中

DOI

- 電子データなどに登録される国際的な識別子で、論文などの学術情報資源の識別子として広く使用されている。

プレフィクス
(出版社や機関)

サフィックス
(コンテンツ)

<http://doi.org/10.1241/johokanri.57.591>

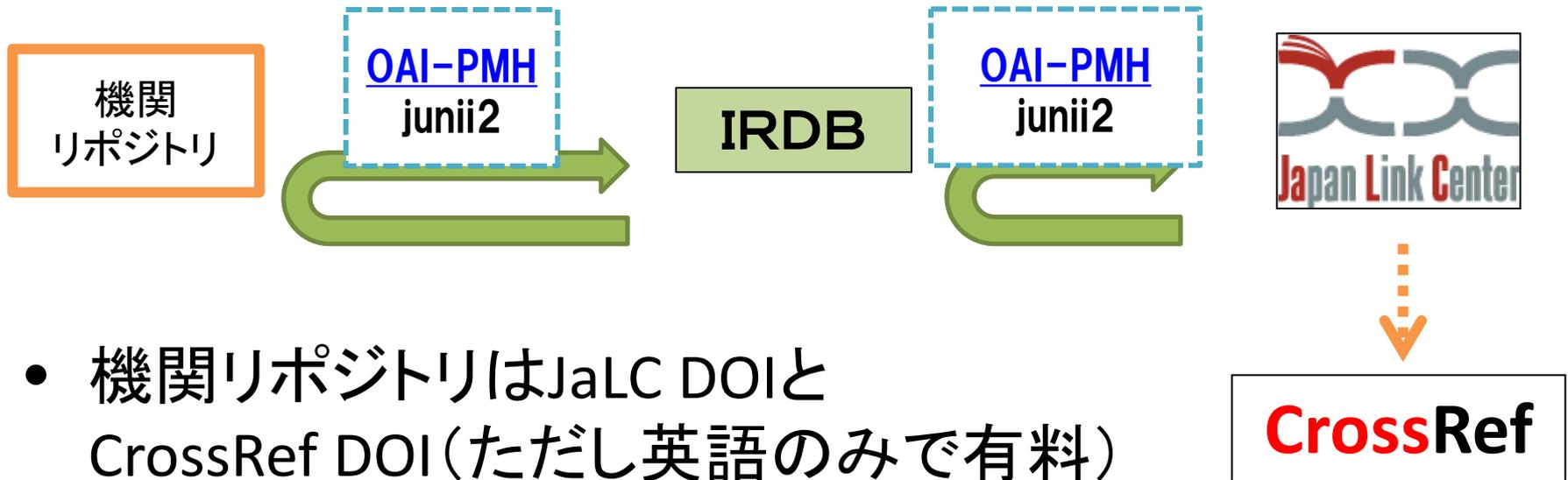


https://www.jstage.jst.go.jp/article/johokanri/57/8/57_591/_article

- <http://doi.org/>に続けてDOIを入力すると論文のあるURLに自動転送される。

ジャパンリンクセンター (JaLC)

- 日本初のDOI登録機関 (RA:Registration Agency)で、海外学術出版社のRAであるCrossRefとも連携



- 機関リポジトリはJaLC DOIと CrossRef DOI (ただし英語のみで有料) どちらを登録するか選択できる

JaLCガイドライン

- 正しくDOIを登録するためにNIIの「JaLCガイドライン」(<http://www.nii.ac.jp/irp/archive/system/jalc>)を参照
- 登録できるコンテンツ
 - 自機関で出版(学位授与)、または出版者から管理委託されたコンテンツ
 - 著者版コンテンツ
 - 1990年代の博士論文で、NDLですでにJaLC DOIを登録している論文は、リポジトリでもそれと同じDOIを登録する(マルチプルレゾリューション) 例doi:[10.11501/3111319](https://doi.org/10.11501/3111319)

ハーベスト確認

あなたの機関のBaseURLは？

<http://ir.lib.shizuoka.ac.jp/dspace-oai/request>

?verb=ListRecords

データを全部ください

&metadataPrefix=junii2

junii2 形式で (OAI_DCなら"oai_dc")

&from=2015-02-09

2015年2月 9日 から

&until=2015-02-16

2015年2月16日まで

登録・更新
された

&set=hdl_10297_20

コレクションを特定する場合のみ

- 各メタデータフォーマットでの出力状況を確認し、NIIからのIRDBハーベストログメールも毎回チェック

さいごに

- 内部メタデータと、外部メタデータのOAI_DCやjunii2などとのクロスウォークを確認する。
- 各サービスプロバイダに、正しいメタデータを提供するために、junii2ガイドラインやJaLCガイドラインなどに従う。
- 機関内だけで悩まないで、コミュニティに相談しましょう。みんな同じような経験を積んでいます。