

2020年10月26日（月）

学術コミュニケーション技術セミナー(JPCOAR Monday)

A.オープンアクセス及び機関リポジトリ周辺技術解説

リポジトリ周辺技術解説(2)

OAI-PMH, ResourceSync, IRDB連携

オープンアクセスリポジトリ推進協会（JPCOAR）

コンテンツ流通促進作業部会メタデータチーム

発表者：片岡 朋子（お茶の水女子大学）

スライド作成協力：田口 忠祐（東京大学）

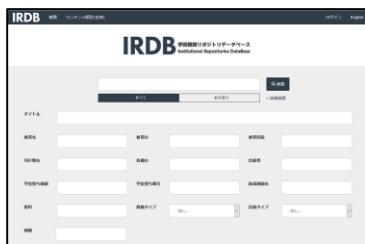
IRDBとは

IRDBとは

①収集機能

機関リポジトリ

学術機関リポジトリ
データベース (IRDB)



②提供機能

国内

海外

CiNii

医中誌Web
Japan Medical Abstracts Society

Japan Link Center

国立国会図書館サーチ

National Diet Library Digital Collections
国立国会図書館デジタルコレクション

CORE

OpenAIRE

BASE

The OALster® database

EBSCO
Discovery Service

Summon®
ウェブスケール・ディスカバリー

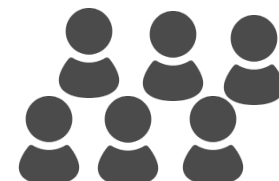
ExLibris Primo



大学内の学生・
研究者



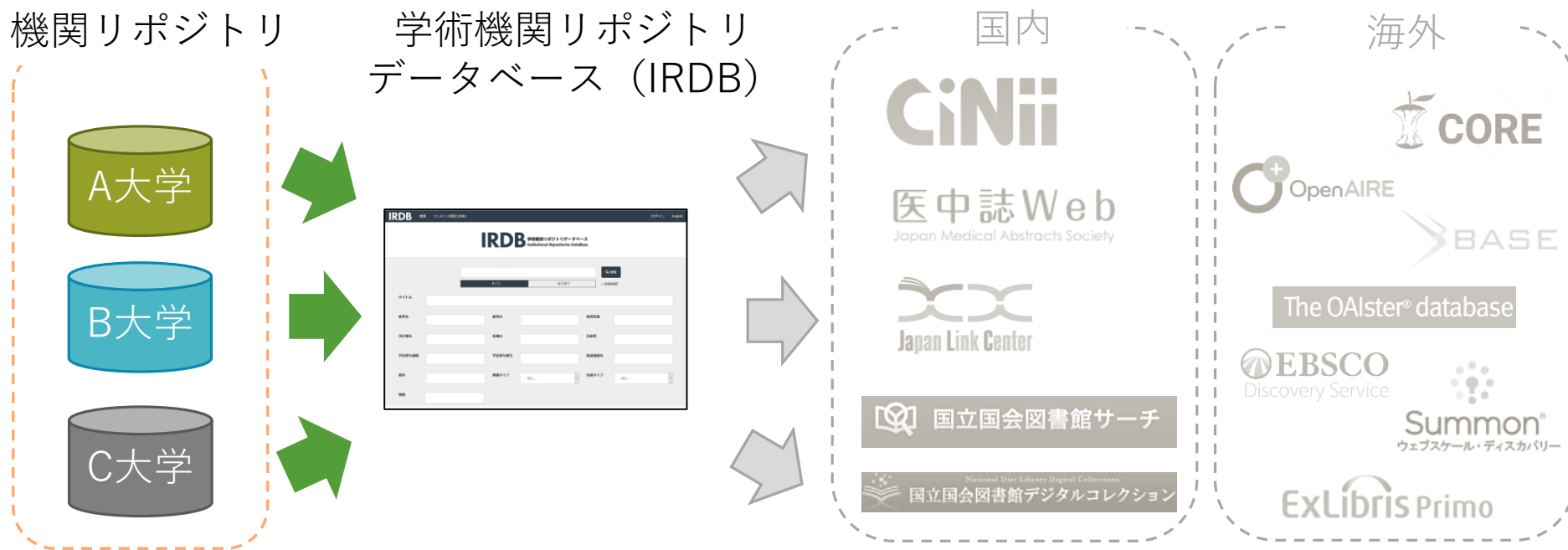
日本の学生・
研究者



世界の学生・
研究者

IRDBの機能

①収集機能



収集のプロセス

1. OAI-PMH プロトコルによりメタデータをハーベスト (JPCOARスキーマまたはjunii2に対応)
2. エラーチェック・正規化
3. junii2はJPCOARスキーマに変換して集約

IRDBの機能

②提供機能



- 日本の研究成果を集約して国内外の関係サービスに流通させる、学術情報流通のハブとしての機能
- 国立情報学研究所の取りまとめるJaLC準会員がIRDBを介してDOIを付与できる

異なるシステム間でのメタデータ交換



① メタデータスキーマの標準化

DCMI Metadata Terms、JPCOARスキーマなど

② 提供方式の標準化

OAI-PMH、ResourceSyncなど

OAI-PMHとは

OAI-PMHとは

Open Archives Initiative Protocol for
Metadata Harvesting (OAI-PMH)

Open Archives Initiativeが策定した
メタデータの交換プロトコル（規約）

1999年 サンタフェ会議にてOAI発足

2001年 OAI-PMH Version 1.0公開

2002年 OAI-PMH Version 2.0公開

→図書館界におけるメタデータ交換の標準
プロトコルとして世界的に普及

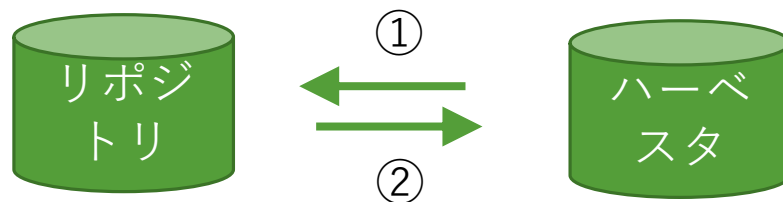
OAI-PMHの特徴

実装の容易さ “*low-barrier mechanism*”

- **HTTP**上のプロトコルを利用
- メタデータの形式は**XML形式**
- メタデータフォーマットは**oai_dc** (Simple Dublin Core[15要素]) への対応のみ必須

ハーベスト（データ収集）の流れ

①URL形式のリクエストを送信



②XML形式でメタデータの情報 that 返戻

OAI-PMHの仕組み

リクエストURL例

http://teapot.lib.ocha.ac.jp/oai?verb=ListRecords

①BaseURL（リポジトリ毎の基本URL） ②リクエストの種類

&metadataPrefix=junii2&from=2020-10-10

③パラメータ（条件）

&ResumptionToken=xxx

④フロー制御（ページめくり）

→お茶の水女子大学の機関リポジトリから（①）

2020年10月10日以降に登録・更新されたメタデータをjunii2形式で（③）

前回取得した件数以降分（④）取得する（②）

返戻値 (XML形式)

https://teapot.lib.ocha.ac.jp/oai?verb=ListRecords&metadataPrefix=junii2&from=2020-10-10

This XML file does not appear to have any style information associated with it.

http://teapot.lib.ocha.ac.jp/oai?verb=ListRecords&metadataPrefix=junii2&from=2020-10-10

1件目

```
<?xml version="1.0" encoding="UTF-8" ?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  <responseDate>2020-10-16T20:19:34Z</responseDate>
  <request verb="ListRecords" metadataPrefix="junii2" from="2020-10-10">https://teapot.lib.ocha.ac.jp/oai?verb=ListRecords&metadataPrefix=junii2&from=2020-10-10</request>
  <ListRecords>
    <record>
      <header>
        <identifier>oai:teapot.lib.ocha.ac.jp:00043009</identifier>
        <timestamp>2020-10-14T00:35:46Z</timestamp>
        <setSpec>00347:00362:02375</setSpec>
      </header>
      <metadata>
        <junii2 xsi:schemaLocation="http://irdb.nii.ac.jp/oai http://irdb.nii.ac.jp/oai/junii2-3-1.xsd">
          <title lang="ja"><書評>ゲイル・サラモン著, 藤高和輝訳 『身体を引き受ける トランスジェンダーと物質性のレトリック』</title>
          <language>jpn</language>
          <URI>http://hdl.handle.net/10083/00063804</URI>
          <selfDOI ra="JaLC">info:doi/10.24567/00063804</selfDOI>
          <NIItype>Departmental Bulletin Paper</NIItype>
          <creator lang="ja">中村, 美亜</creator>
          <jtitle>ジェンダー研究 : お茶の水女子大学ジェンダー研究所年報</jtitle>
          <issue>23</issue>
          <spage>229</spage>
          <epage>231</epage>
          <dateofissued>2020-07-31</dateofissued>
          <NCID>AA11391754</NCID>
          <textversion>publisher</textversion>
          <subject>367</subject>
          <fullTextURL>https://teapot.lib.ocha.ac.jp/?action=repository_action_common_download&item_id=43009&item_no=1&attribute_id=21&file_no=1</fullTextURL>
          <publisher>お茶の水女子大学ジェンダー研究所</publisher>
          <type>紀要論文</type>
          <relation>http://www2.igs.ocha.ac.jp/gender/gender-40/</relation>
          <date>2020-07-14</date>
        </junii2>
      </metadata>
    </record>
    <record>
      <header>
        <identifier>oai:teapot.lib.ocha.ac.jp:00043010</identifier>
        <timestamp>2020-10-14T00:38:59Z</timestamp>
        <setSpec>00347:00362:02375</setSpec>
      </header>
      <metadata>
        <junii2 xsi:schemaLocation="http://irdb.nii.ac.jp/oai http://irdb.nii.ac.jp/oai/junii2-3-1.xsd">
          <title lang="ja"><書評>石井香江著 『電話交換手はなぜ「女の仕事」になったのか : 技術とジェンダーの日独比較社会史』</title>
          <language>jpn</language>
          <URI>http://hdl.handle.net/10083/00063805</URI>
          <selfDOI ra="JaLC">info:doi/10.24567/00063805</selfDOI>
          <NIItype>Departmental Bulletin Paper</NIItype>
          <creator lang="ja">萩原, 久美子</creator>
          <jtitle>ジェンダー研究 : お茶の水女子大学ジェンダー研究所年報</jtitle>
          <issue>23</issue>
          <spage>232</spage>
          <epage>234</epage>
          <dateofissued>2020-07-31</dateofissued>
          <NCID>AA11391754</NCID>
          <textversion>publisher</textversion>
          <subject>367</subject>
          <fullTextURL>https://teapot.lib.ocha.ac.jp/?action=repository_action_common_download&item_id=43010&item_no=1&attribute_id=21&file_no=1</fullTextURL>
          <publisher>お茶の水女子大学ジェンダー研究所</publisher>
          <type>紀要論文</type>
        </junii2>
      </metadata>
    </record>
  </ListRecords>
</OAI-PMH>
```

2件目

①BaseURL等の確認

IRDB 検索 コンテンツ統計(全体) マイコンテンツ コンテンツ統計(ユーザ) ユーザ情報 1 ログアウト English

2 ビュー 編集

3

機関	xxx大学
リポジトリ名	xxx大学リポジトリ
リポジトリURL	http://xxx.repo.nii.ac.jp/
週次ハーベスト頻度	月
データ取得種別	OAI-PMH
BaseURL	http://xxx.repo.nii.ac.jp/oai
メタデータPrefix	junii2
公開対象	公開
連携対象	CiNii
JaLC DOI Prefix	10.xxxxx
ハーベスログ送信先	xxx@xxx.ac.jp
次回ハーベスト起算日時	2019-04-04 15:00:05
前回ハーベスト開始日時	2019-04-04 15:00:05
前回ハーベスト終了日時	2019-04-04 16:03:31
前回ハーベスト対象データ件数	5,635
前回ハーベスト登録件数	65
前回ハーベスト更新件数	5,517
前回ハーベスト削除件数	53
前回ハーベストデータ取得エラー	0
前回ハーベステータス	正常
前回ハーベストエラーログ	前回ハーベスト開始日時: 2019-04-04 15:00:05 前回ハーベスト終了日時: 2019-04-04 16:03:31 ハーベスト対象データ件数: 5635 登録件数: 65 更新件数: 5517 削除件数: 53 データ取得エラー: 0 レコードエラー件数: 0 項目エラー件数: 16 ワーニング件数: 0 項目変換件数: 0 確認先URL: https://irdb.nii.ac.jp/usercontents

①BaseURL
(リポジトリ毎の基本URL)

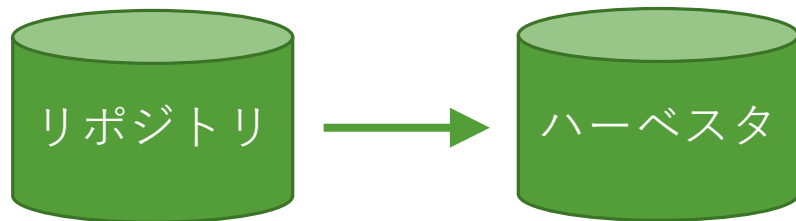
IRDBに収集されているメタデータ形式
次期JAIRO Cloud移行後は“jpcoar”
(JPCOARスキーマ形式) に

②リクエストの種類／③パラメータ¹³

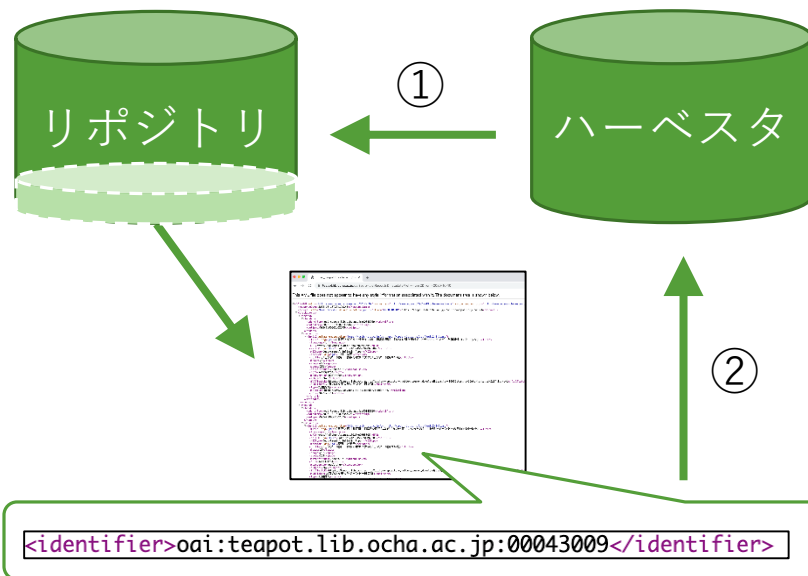
リクエストの種類	内容	パラメータ (*は必須)
Identify	リポジトリ名称、BaseURL、OAI-PMHのバージョン等の情報を取得する	指定不可
ListMetadataFormats	自機関リポジトリから出力可能なメタデータフォーマットの情報を取得する	ResumptionToken
ListSets	Set（データ集合）の一覧を取得する	ResumptionToken
ListIdentifiers	指定した抽出条件（From/Until/Set）に該当するメタデータのIDを取得する	From / Until Set MetadataPrefix* ResumptionToken
GetRecord	指定したID（Identifier）1件のメタデータを指定したメタデータフォーマット（MetadataPrefix）で取得する	Identifier* MetadataPrefix*
ListRecords	指定した抽出条件（From/Until/Set）に該当するメタデータを指定したフォーマット（MetadataPrefix）で取得する	From / Until Set MetadataPrefix* ResumptionToken

データ更新の仕組み

全件取得・・・登録されている全データを収集



差分更新・・・前回取得分以降に更新されたデータのみ収集



① リクエスト時に収集対象の期間を指定

2020/1/8のハーベスト

from=2020-01-01&until=2020-01-07

2020/1/15のハーベスト

from=2020-01-08&until=2020-01-14

⋮

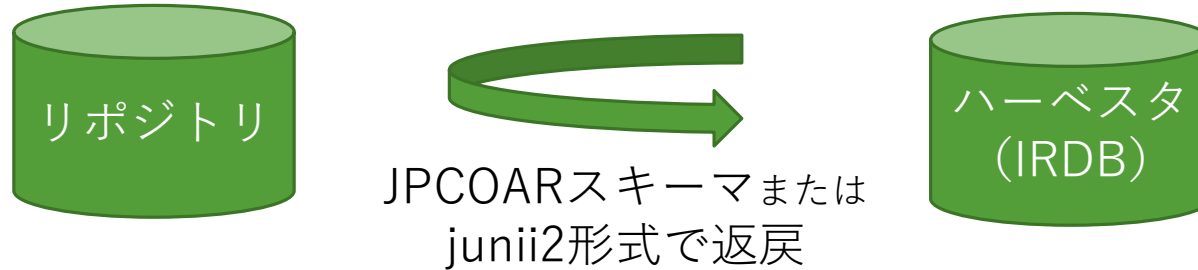
② header内のidentifierをキーに照合
identifierなし→新規登録
identifierあり

status="deleted"あり→削除

status="deleted"なし→更新

OAI-PMHの実装例

IRDBより週次でハーベスト



インデックス

紀要等 ▶ ジェンダー研究 ▶ 23

Permalink : <http://doi.org/10.24567/00063803>

幼稚園教諭賃金の地域格差に見る少子化時代の女性活躍推進のジレンマ——保育者の労働環境におけるジェンダー不平等の実証分析

[利用統計を見る](#)

File / Name	License
15,p207-228.pdf	
15,p207-228.pdf (1.34MB) [2 downloads]	

[OAI-PMH](#)
[BIBTeX](#)
[OWL/RWC](#)
[WFO](#)
[EXPORT](#)
[PRINT](#)

JaLC DOI	info:doi/10.24567/00063803
タイトル(英)	The Dilemma Posed by Measures to Promote the Social Advancement of Women Evaluated from the Regional Disparity of the Salaries of Kindergarten Teachers: An Empirical Analysis of Gender Inequality in the Working Conditions of ECEC Teachers
アイテムタイプ	紀要論文 / Departmental Bulletin Paper
言語	日本語
キーワード	幼稚園教諭, 賃金, 女性活躍推進, ジェンダー不平等, マルチレベル分析
著者	小林, 佳美
著者 (英)	KOBAYASHI, Yoshimi
抄録	本稿では、少子化対策の枠組みで進められた女性活躍推進が、幼稚園教諭の賃金にどのような影響を与えたのかを定量的に明らかにすることを目的に、1980年から35年間に、8時点の都道府県別平均賃金データを活用した実証分析を行った。女性一般労働者に比した幼稚園教諭の賃金水準は「男女雇用機会均等法」制定前夜の1980年代前半から、平均以下となり2015年に至るが、人口が少なく、もともと3歳未満児就園率の高い地域で低くなる傾向が明らかになった。さらに、2005年以降の幼稚園教諭の賃金水準は、女性が出産を経て就業を継続し、キャリアを重ねる傾向を強めている都道府県ほど低くなる傾向が、マルチレベル分析の結果から観察され、女性活躍推進により生成されたジェンダー内格差の内側で、幼児教育の領域にも不平等が生み出されていたことが示唆された。
雑誌名	ジェンダー研究：お茶の水女子大学ジェンダー研究所年報
号	23
ページ	207 - 228
発行年	2020-07-31

IRDB 検索 コンテンツ統計(全体)

[← 一覧に戻る](#)

タイトル	ja : 幼稚園教諭賃金の地域格差に見る少子化時代の女性活躍推進のジレンマ——保育者の労働環境におけるジェンダー不平等の実証分析
その他のタイトル	en : The Dilemma Posed by Measures to Promote the Social Advancement of Women Evaluated from the Regional Disparity of the Salaries of Kindergarten Teachers: An Empirical Analysis of Gender Inequality in the Working Conditions of ECEC Teachers
作成者	ja : 小林, 佳美
主題	Other : 幼稚園教諭 Other : 賃金 Other : 女性活躍推進 Other : ジェンダー不平等 Other : マルチレベル分析 Other : 367
内容注記	Other : type 紀要論文 Other : 本稿では、少子化対策の枠組みで進められた女性活躍推進が、幼稚園教諭の賃金にどのような影響を与えたのかを定量的に明らかにすることを目的に、1980年から35年間に、8時点の都道府県別平均賃金データを活用した実証分析を行った。女性一般労働者に比した幼稚園教諭の賃金水準は「男女雇用機会均等法」制定前夜の1980年代前半から、平均以下となり2015年に至るが、人口が少なく、もともと3歳未満児就園率の高い地域で低くなる傾向が明らかになった。さらに、2005年以降の幼稚園教諭の賃金水準は、女性が出産を経て就業を継続し、キャリアを重ねる傾向を強めている都道府県ほど低くなる傾向が、マルチレベル分析の結果から観察され、女性活躍推進により生成されたジェンダー内格差の内側で、幼児教育の領域にも不平等が生み出されていたことが示唆された。
出版者	お茶の水女子大学ジェンダー研究所
日付	Created : 2020-07-14, Issued : 2020-07-31
言語	jpn
資源タイプ	departmental bulletin paper

ResourceSyncとは

ResourceSyncとは

- OAI-PMH の方向性を踏襲しつつ、より新しいWeb技術（Sitemapプロトコル）を採用
- メタデータだけでなくコンテンツ本文の収集や大容量データの収集を可能に

2011年 NISO（米国情報標準化機構）とOpen Archives Initiativeが検討開始

2014年 Version1.0 (ANSI/NISO Z39.99-2014)

2017年 Version1.1 (ANSI/NISO Z39.99-2017)

Sitemapプロトコルとは？

サーチエンジンのクローラーに適切にクロールしてもらうためにウェブサイト上のファイルの構成を伝えるためのファイル

XML形式で記述され、サイトのURL、各URLに対する追加のメタデータから構成されている

Google、Yahoo、Microsoftなど広く利用されている

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>http://www.example.com/</loc>
    <lastmod>2020-10-03T16:42:39+09:00</lastmod>
    <changefreq>monthly</changefreq>
    <priority>0.8</priority>
  </url>
</urlset>
```

sitemap.xsdを指定

クロールしてもらうURL（必須）

ファイルの最終更新日（オプション）

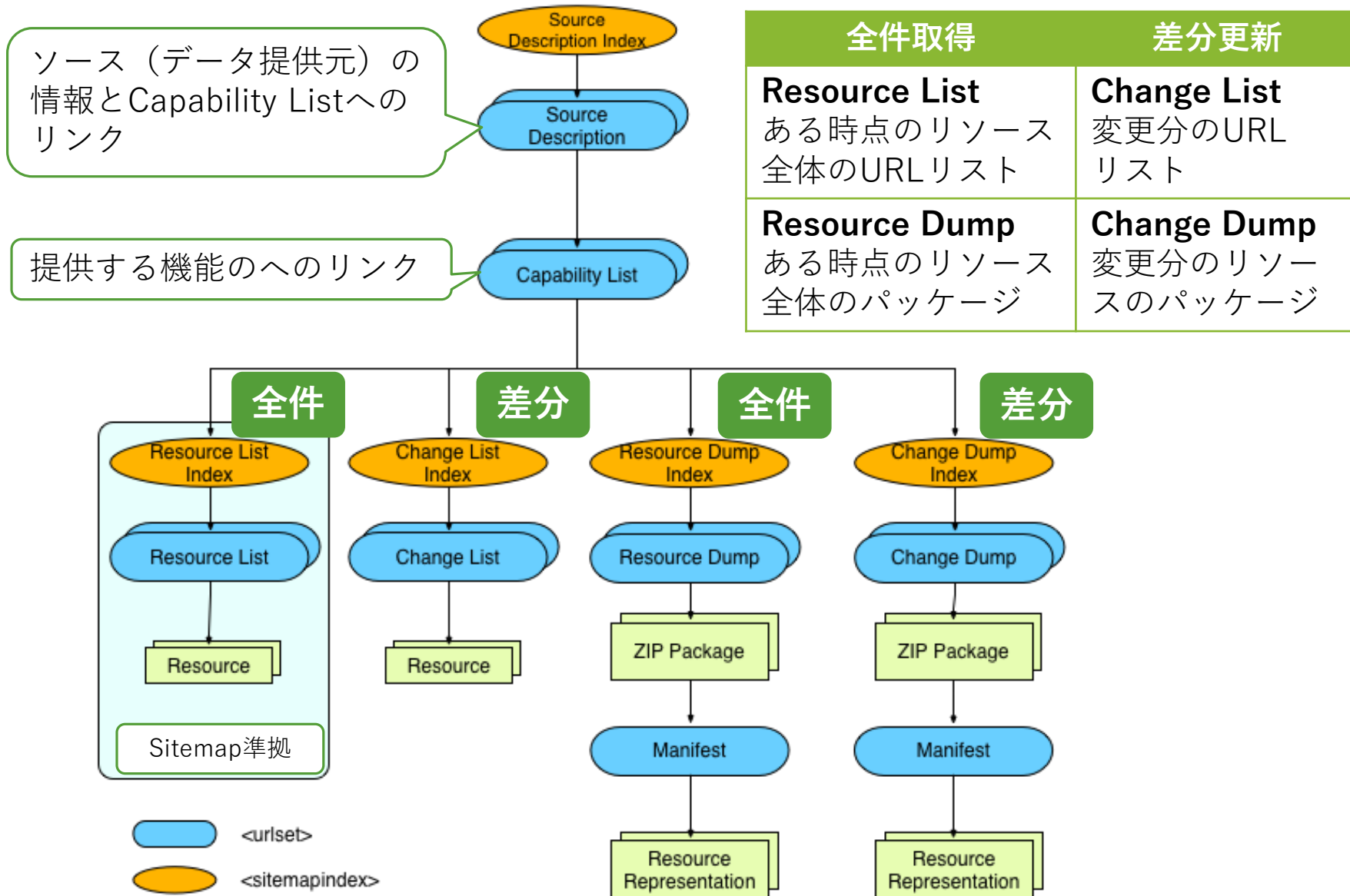
更新頻度（オプション）

優先度（オプション）

繰り返す

ResourceSyncの仕組み

19



ResourceSyncの具体例

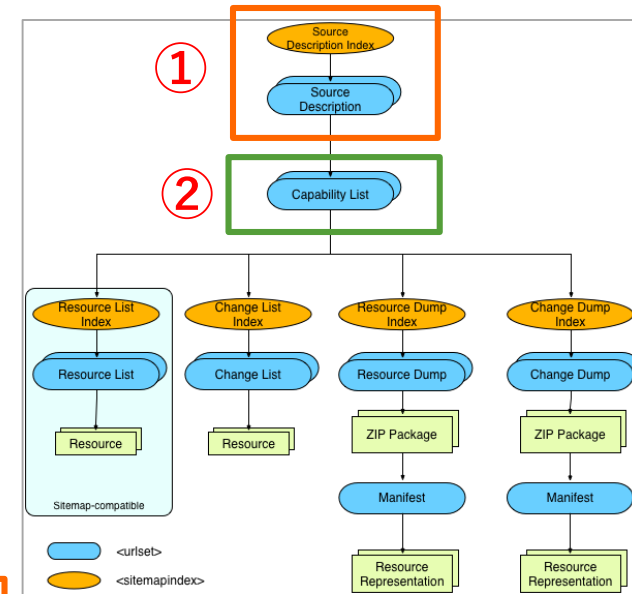
20

実際に想定されるResourceSyncのファイル例

最初にデスティネーション（データ収集を行う側）に
エン트리ポイント（入口）を示す

/robots.txt に記載する

```
Disallow: /index.php/user/register/  
Disallow: /index.php/user/login/  
Disallow: /index.php/user/logout/  
ResourceSync: http://irdb.nii.ac.jp/.well-known/resourcesync
```



/.well-known/resourcesyncにSource Description ① を作成する

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"  
  xmlns:rs="http://www.openarchives.org/rs/terms/"  
  <rs:ln rel="describedby"  
    href="http://example.com/info-about-source.xml"/>  
  <rs:md capability="description"/>  
</urlset>  
  
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"  
  xmlns:rs="http://www.openarchives.org/rs/terms/"  
  <rs:ln rel="describedby"  
    href="http://example.com/info-about-set1-of-resources.xml"/>  
  <rs:md capability="capabilitylist"/>  
</urlset>
```

②へのリンク

以降、xml宣言は省略

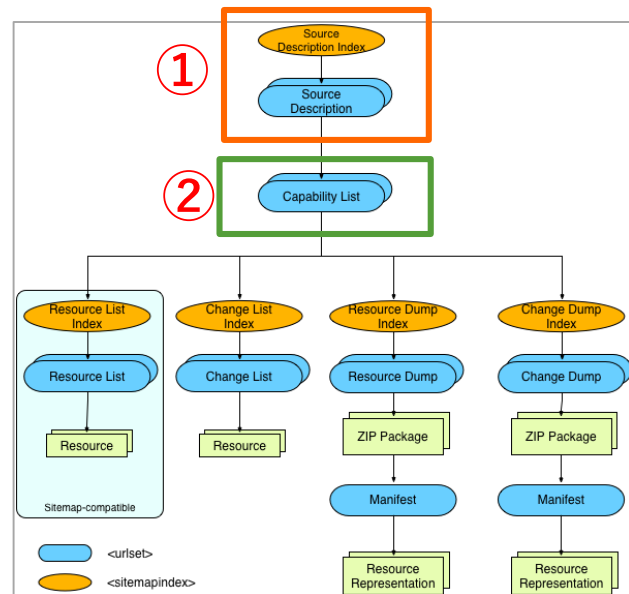
Source Descriptionの例

21

description.xml ①

→ 提供するcapabilitylist.xml ② を記載

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:rs="http://www.openarchives.org/rs/terms/"
  <rs:ln rel="describedby"
    href="http://example.com/info_about_source.xml"/>
  <rs:md capability="description"/>
  <url>
    <loc>http://example.com/capabilitylist1.xml</loc>
    <rs:md capability="capabilitylist"/>
    <rs:ln rel="describedby"
      href="http://example.com/info_about_set1_of_resources.xml"/>
  </url>
  <url>
    <loc>http://example.com/capabilitylist2.xml</loc>
    <rs:md capability="capabilitylist"/>
    <rs:ln rel="describedby"
      href="http://example.com/info_about_set2_of_resources.xml"/>
  </url>
</urlset>
```



②へのリンク

Capability Listを
2つ提供する場合は
繰り返して指定

②へのリンク

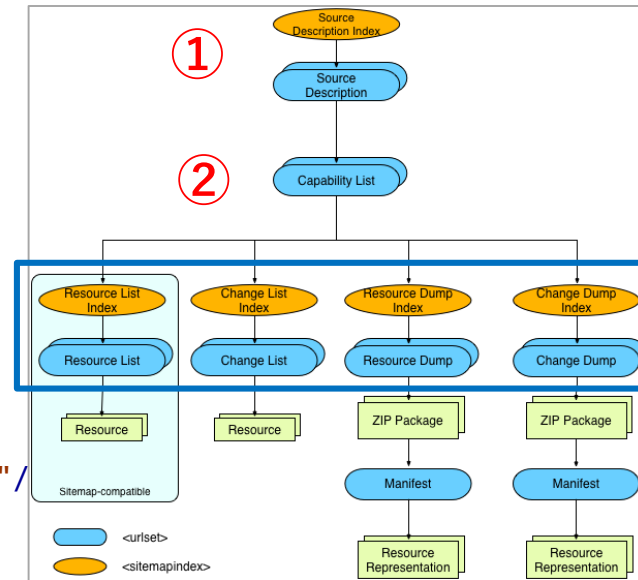
Capability Listの例

22

capabilitylist.xml ②

→ 提供している機能リストを記載

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:rs="http://www.openarchives.org/rs/terms/"
  <rs:ln rel="describedby"
    href="http://example.com/info_about_set1_of_resources.xml"/>
  <rs:ln rel="up"
    href="http://example.com/resourcesync_description.xml"/>
  <rs:md capability="capabilitylist"/>
  <url>
    <loc>http://example.com/dataset1/resourcelist.xml</loc>
    <rs:md capability="resourcelist"/>
  </url>
  <url>
    <loc>http://example.com/dataset1/changelist.xml</loc>
    <rs:md capability="changelist"/>
  </url>
</urlset>
```



③へのリンク

機能がResource Listであることを指定

③へのリンク

機能がChange Listであることを指定

Resource Listの例

23

resourcelist.xml ③

→ リソース全体(メタデータやコンテンツ)のURL一覧

機能がResource Listであることを指定

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
xmlns:rs="http://www.openarchives.org/rs/terms/"
<rs:ln rel="up"
href="http://example.com/dataset1/capabilitylist.xml"/>
<rs:md capability="resourcelist"
at="2013-01-03T09:00:00Z"
completed="2013-01-03T09:01:00Z"/>
```

URL

<url>

```
<loc>http://example.com/res1</loc>
```

```
<lastmod>2013-01-02T13:00:00Z</lastmod>
```

最終更新年月日

```
<rs:md hash="md5:1584abdf8ebdc9802ac0c6a7402c03b6"
length="8876"
type="text/html"/>
```

④への
リンク

</url>

リソースのメディアタイプ

<url>

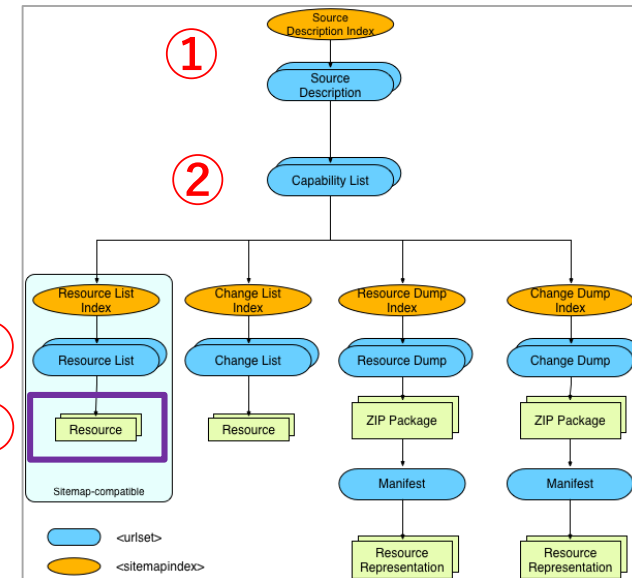
```
<loc>http://example.com/res2</loc>
```

```
<lastmod>2013-01-02T14:00:00Z</lastmod>
```

```
<rs:md hash="md5:1e0d5cb8ef6ba40c99b14c0237be735e
sha-256:854f61290e2e197a11bc91063afce22e43f8ccc655237050ace766adc68dc784"
length="14599"
type="application/pdf"/>
```

</url>

</urlset>



JPCOARスキーマ形式での メタデータ出力例

resourcelist.xml

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:rs="http://www.openarchives.org/rs/terms/"
  <rs:ln rel="up"
    href="http://example.com/dataset1/capabilitylist.xml"/>
  <rs:md capability="resourcelist"
    at="2013-01-03T09:00:00Z"
    completed="2013-01-03T09:01:00Z"/>
  <url>
    <loc>http://example.com/item-0001.xml</loc>
    <lastmod>2013-01-02T13:00:00Z</lastmod>
  </url>
</urlset>
```

← リソースのURLを指定
← 最終更新年月日



item-0001.xml

```
<dc:title xml:lang="ja">情報爆発時代の研究基盤構想</dc:title>
<jpcoar:creator>
  <jpcoar:nameIdentifier nameIdentifierScheme="ORCID"
    nameIdentifierURI="https://orcid.org/0000-0001-0002-0003">
    0000-0001-0002-0003
  </jpcoar:nameIdentifier>
  <jpcoar:creatorName xml:lang="ja">夏目， 漱石</jpcoar:creatorName>
</jpcoar:creator>
```

← JPCOARスキーマ形式で出力

Resource List Indexの例

Resource List Index → 複数のresourcelist.xmlを指定

 sitemapindex形式を利用する

```
<sitemapindex xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:rs="http://www.openarchives.org/rs/terms/">
  <rs:md capability="resourcelist"
    at="2013-01-03T09:00:00Z"/>
  <sitemap>
    <loc>http://example.com/resourcelist-part1.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://example.com/resourcelist-part2.xml</loc>
  </sitemap>
</sitemapindex>
```

 機能がResource Listであることを指定

 複数のファイルを指定

 複数のファイルを指定

sitemapプロトコルの制限を超える場合は、sitemapindex形式を利用したResource List Indexを作成してファイルを分割するか、gzipを利用して圧縮する

sitemapプロトコルの制限
URL: 50,000件以下
サイズ: 50MB以下

差分更新の仕組み

changelist.xml → 変更が発生したリソースのURLの一覧

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:rs="http://www.openarchives.org/rs/terms/"
  <rs:ln rel="up"
    href="http://example.com/dataset1/capabilitylist.xml"/>
  <rs:md capability="changelist"
    from="2013-01-03T00:00:00Z"/>
  <url>
    <loc>http://example.com/res1.html</loc>
    <lastmod>2000-01-01T01:01:00Z</lastmod>
    <rs:md change="created" datetime="2013-01-03T11:00:00Z"/>
  </url>
  <url>
    <loc>http://example.com/res2.pdf</loc>
    <lastmod>2013-01-03T13:00:00Z</lastmod>
    <rs:md change="updated" datetime="2013-01-03T13:00:00Z"/>
  </url>
  <url>
    <loc>http://example.com/res3.tiff</loc>
    <rs:md change="deleted" datetime="2013-01-03T18:00:00Z"/>
  </url>
</urlset>
```

機能がChange Listであることを指定

この日時以降に変更が発生

change属性で、作成、更新、削除
なのかを指定する

Change Listもsitemapプロトコルの制限を超える場合には、Resource List Indexと同様にChange List Indexを利用する

ResourceSyncの実装例

27

IRDB 検索 コンテンツ統計(全体) マイコンテンツ コンテンツ統計(ユーザ) ユーザ情報 English

Unmasquerade ログアウト

ビュー 編集

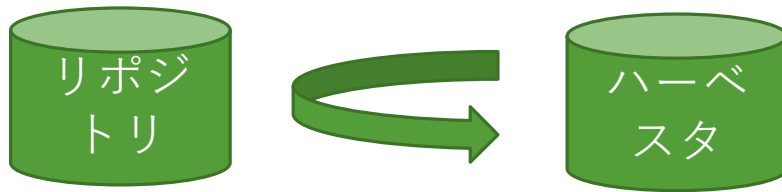
機関	
リポジトリ名	
リポジトリURL	
月次ハーベスト頻度	2
データ取得種別	ResourceSync
BaseURL	
メタデータPrefix	jpcoar
公開対象	非公開
連携対象	CiNii
ハーベスログ送信先	
次回ハーベスト起算日時	2020-10-02 15:00:03
前回ハーベスト開始日時	2020-10-02 15:00:03
前回ハーベスト終了日時	2020-10-02 15:01:29
前回ハーベスト対象データ件数	177
前回ハーベスト登録件数	162
前回ハーベスト更新件数	15
前回ハーベスト削除件数	0
前回ハーベストデータ取得エラー	0
前回ハーベステータス	正常
前回ハーベストエラーログ	前回ハーベスト開始日時: 2020-10-02 15:00:03 前回ハーベスト終了日時: 2020-10-02 15:01:29 ハーベスト対象データ件数: 177 登録件数: 162 更新件数: 15 削除件数: 0 データ取得エラー: 0 レコードエラー件数: 0 項目エラー件数: 0 ワーニング件数: 177 項目変換件数: 0 確認先URL: https://irdb.nii.ac.jp/usercontents

対応状況の例

- CiNii Dissertations (提供機能)
- KAKEN (提供機能)
- IRDB (収集機能、提供機能)
- JSTプロジェクトデータベース (提供機能)
- Materials Data Repository (提供機能)

OAI-PMHとResourceSyncの違い

OAI-PMH



- ハーベスタ側が必要な期間のデータを必要なフォーマットで取得
- 収集対象はメタデータ
- リポジトリ側からハーベスタに変更を通知する仕組みはない
- 全件取得時やハーベスト件数が多い場合にリポジトリ側のシステム負荷が高い (システムダウンなど)

ResourceSync



- リポジトリ側が予め作成したxmlファイルをハーベスタが取得
- 収集対象はメタデータやコンテンツ本文などWeb上のURIを持つリソース全般
- 変更発生時にリポジトリ側より変更通知用URLにプッシュ型で通知する機能あり
- データを速く収集でき、複数のハーベスタが同時にハーベストしてもリポジトリ側の負荷は低い

参考資料

- 学術機関リポジトリデータベースサポート
<https://support.irdb.nii.ac.jp/>
- Open Archives Initiative
<https://www.openarchives.org/>
- OAI-PMH2.0日本語訳
<https://www.nii.ac.jp/irp/archive/translation/oai-pmh2.0/>
- 国立国会図書館サーチが提供するOAI-PMH > 「OAI-PMHの要点」
https://iss.ndl.go.jp/information/api/api-lists/oai-pmh_info/
- ResourceSync Framework Specification (ANSI/NISO Z39.992017)
<http://www.openarchives.org/rs/1.1/resourcesync>
- 林豊. ResourceSync : OAI-PMH の後継規格. カレントアウェアネス. 2015, (323), CA1845, p. 17-21.
<https://doi.org/10.11501/9107589>
- Klein, Martin. (2018, September). ResourceSync: A Modular Framework for Web-Based Resource Synchronization. Zenodo.
<https://doi.org/10.5281/zenodo.1433806>

ご清聴ありがとうございました