2020.10.26　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(1)
JPCOARスキーマ

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
瀬川　結美（東京学芸大学）

JPCOARスキーマとは

- **メタデータを記入する際の規則や語彙を定義**したもの
- **機関リポジトリのメタデータ交換フォーマット**の一つ
- 2017年にVer.1.0（初版）リリース
- **新JAIRO Cloud（WEKO３）標準メタデータスキーマ**

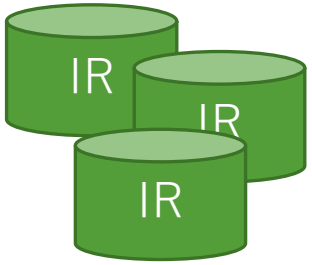
大園隼彦, 片岡朋子, 高橋菜奈子, 田口忠祐, 林豊, 南山泰之.
JPCOARスキーマの策定：日本の学術成果の円滑な国際的流通を目指して.
情報管理. 2018. 60(10), 719-729　　https://doi.org/10.1241/johokanri.60.719

スキーマ策定の経緯や
考え方はこちらをぜひ
ご参照ください！

メタデータ交換の流れ

②各リポジトリからIRDBに
メタデータをハーベスティング

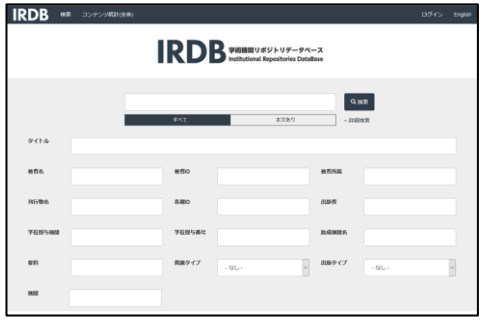
③国内外の各種サービスに
メタデータを流通



機関リポジトリ

JPCOAR
スキーマ

IRDB
学術機関リポジトリ
データベース



JPCOAR
スキーマ

CiNii 日本の論文をさがす Articles

CiNii 日本の博士論文をさがす Dissertations

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

医中誌Web
Japan Medical Abstracts Society

Japan Link Center

①リポジトリに
メタデータ等を登録

図書館員・研究者

連携先サービス例

JPCOARスキーマ　３つのポイント

1. **オープンアクセス、オープンサイエンスへの対応**

…アクセス権・APC等への対応、各種研究データへの対応

2. **識別子の拡充、メタデータ構造の修正**

…多様な識別子を採用、メタデータを3階層に構造化

3. **国際的な相互運用性の向上**

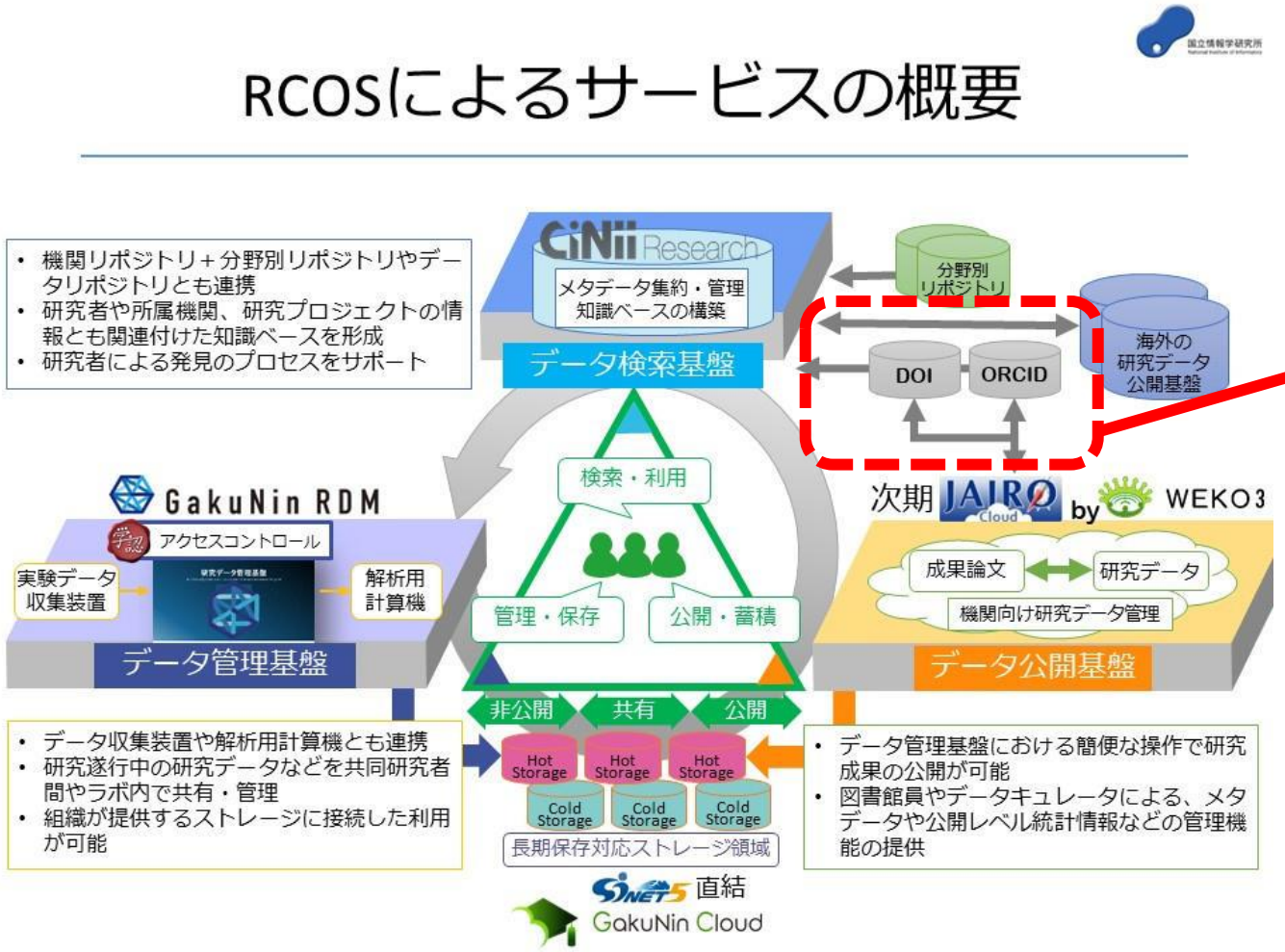
…国際標準メタデータスキーマ（Dublin Core、OpenAIRE、DataCite…）と互換性のある設計とし、国際的なデータ流通を目指す

他のスキーマを参照した要素

日本独自の事情への対応が必要な
要素は独自に定義…ex. 助成情報など

スキーマ	項目名	要素名（第一階層）	
Dublin Core	タイトル	dc:title	世界的に最も汎用性のあるスキーマを採用し、データを流通させやすく…
	その他のタイトル	dcterms:alternative	
	アクセス権	dcterms:accessRights	
	権利情報	dc:rights	
	出版者	dc:publisher	
	言語	dc:language	
	資源タイプ	dc:type	
	時間的範囲	dcterms:temporal	
DataCite FIND, ACCESS, AND REUSE DATA	位置情報	datacite:geoLocation	研究データに対応
	内容記述	datacite:description	
	日付	datacite:date	
	バージョン情報	datacite:version	
国立国会図書館 National Diet Library, Japan DC-NDL	学位授与番号	dcndl:dissertationNumber	学位論文データのための連携
	学位名	dcndl:degreeName	
	学位授与年月日	dcndl:dateGranted	
OpenAIRE	出版タイプ	oaire:version	オープンアクセス対応
rioxx	APC(Article Processing Charge)	rioxxterms:apc	

NII研究データ基盤（NII Research Data Cloud）とのかかわり



機関リポジトリによる
研究成果の収集・発信に
JPCOARスキーマを
採用することによって、

多様なデータを
正確　かつ　**詳細に**
国内外に広く

流通させることが可能に…

国立情報学研究所オープンサイエンス基盤研究センター https://rcos.nii.ac.jp/service/

JPCOARスキーマの普及状況 (2020年6月)

対応年月	サービス/システム名
2017年12月	E-Cats Library Version 6
2019年3月	CiNii Articles, CiNii Dissertations
2019年3月	ジャパンリンクセンター（JaLC）
2019年3月	医中誌Web
2019年4月	IRDB（学術機関リポジトリデータベース）
2019年4月	AMLAD（上智大学学術情報リポジトリ）
2019年6月	東京大学学術資産等アーカイブズポータル(API出力)
2019年12月	DSpace（北海道大学学術成果コレクション）
2020年5月	国立国会図書館サーチ
2020年6月	琉球・沖縄関係貴重資料デジタルアーカイブ（琉球大学）
対応中	次期JAIRO Cloud（WEKO3）
対応中	CiNii Research, BASE, OpenAIRE

JPCOARスキーマ詳細解説

＊「スキーマ説明（JPCOARスキーマガイドライン）」に基づき、主な要素の入力方法について解説します

https://schema.irdb.nii.ac.jp/ja/schema

JPCOARスキーマの全体像

	JPCOARスキーマ ver1.0	junii2
構造	**3階層**	1階層
要素数	**79要素**（68種類）	64要素
移行・内容改訂	21要素	21要素
統合・廃止	7要素	43要素
新規追加・拡張	51要素	-
属性の種類	**15種類**	4種類
言語属性（使用可能な要素数）	**26要素**	6要素
言語属性以外（使用可能な要素数）	29要素	14要素
識別子の種類	**28種類**	12種類
統制語彙（使用する要素・属性数）	**22要素・属性**	5要素

＊**階層構造**を持たせ、データを**グルーピング**できる　ex.作成者・寄与者

＊記入できる要素がリッチに

＊**識別子**、**統制語彙**採用項目が増加し、データ互換性が向上

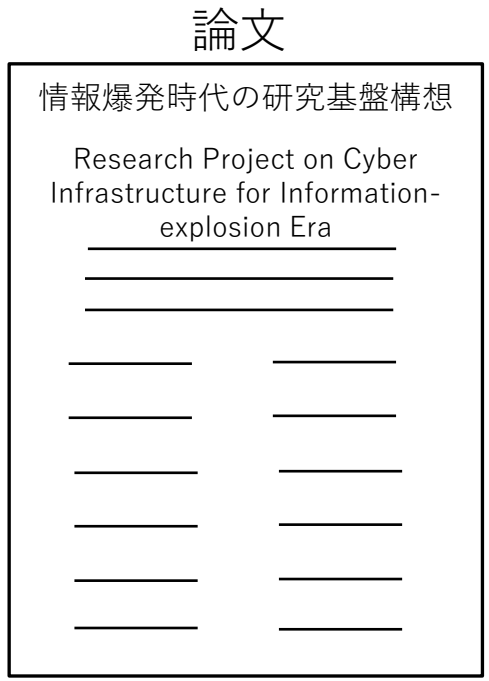
JPCOARスキーマ　新規追加要素

アクセス権	dcterms:accessRights
APC	rioxxterms:apc
権利者情報	jpcoar:rightsHolder
バージョン情報	datacite:version
位置情報	datacite:geolocation
助成情報	jpcoar:fundingReference
ページ数	jpcoar:numPages
会議記述	jpcoar:conference

JPCOARスキーマ基本構成、記入全般に関わる事項

要素と属性

要素は最小の情報単位。要素の付加的な情報は、**属性**で表現。
それぞれ、対象の値とペアで記入。



例：タイトル（日）
情報爆発時代の研究基盤
タイトル（英）
Research Project on Cyber Infrastructure for Information-explosion Era

データをスキーマにしたがって記入

属性
要素　属性名　属性の値　要素の値

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

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

JPCOARスキーマ基本構成、記入全般に関わる事項

記入レベルと繰返回数

記入レベル

M（Mandatory）	必須（必ず使用する）
MA（Mandatory when applicable）	該当する場合は必須 （該当する情報がある場合は、必ず使用する）
R（Recommended）	推奨（使用が望ましい）
O（Optional）	任意（使用は任意とする）

繰返回数

0-N	繰返可（必須以外）
1-N	繰返可（必須）
0-1	繰返不可（必須以外）
1	繰返不可（必須）

Check!

JPCOARスキーマの各要素・属性のそれぞれに**記入レベル**、**繰返回数**の指定があります

JPCOARスキーマ基本構成、記入全般に関わる事項

言語属性

JPCOARスキーマでは、多くの要素で、言語属性(xml:lang)が記入できる。
言語属性では、**要素の値**の言語を指定する。

言語属性　　**日本語**

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

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

英語

- **登録コンテンツ本文の言語を指定するわけではない**
- 基本的にはISO 639-1の2桁の言語コードを使用
  日本語のヨミは “ja-Kana”を使用する
  ヨミを入力する場合は、日本語（xml:lang=“ja”）による要素入力を共に行うことが必須

JPCOARスキーマ基本構成、記入全般に関わる事項

統制語彙 controlled vocabularies

要素・属性には予め定められた用語（統制語彙）を選択記入する項目がある。
統制語彙に対応したURIを属性(rdf:resource)に入力しなければならない要素もある（ex. 出版タイプ）。

統制語彙の例：APC

- **Paid** - 支払済
- **Partially waived** - 一部免除
- **Fully waived** - 全て免除
- **Not charged** - 無料
- **Not required** - 不要
- **Unknown** - 不明

COAR、DataCiteなど国際的に広く使用されている統制語彙を採用。
日本独自の事情を持つ要素・属性は独自に設定。
統制語彙を用いることで入力効率を上げ、データ互換性を高める。

＊統制語彙を採用する要素は、JPCOARスキーマガイドラインに指示 統制語彙 と説明がある

JPCOARスキーマ要素

タイトル、その他のタイトル

タイトル<dc:title>

- 本タイトルを記入
- コンテンツ本文と同じ言語のタイトル情報は必ず記入
- タイトル情報が複数言語ある場合は、繰り返して記入
- 各言語の出現回数は**1回まで**
- 言語属性（xml:lang）の指定がない入力は推奨しない

その他のタイトル<dcterms:alternative>

- 目次タイトル等、本タイトル以外を記入
- 各言語の出現回数に制限はなく、複数回繰り返して記入できる

JPCOARスキーマ要素

作成者、寄与者

人物の情報、所属情報について識別子による記入が可能。

名称を記入する場合は言語属性(xml:lang)を指定。

寄与者は研究データに対応、役割属性(contributorType)を統制語彙で指定。

第1階層 / 第2階層 / 第3階層

```
<jpcoar:creator>
    <jpcoar:nameIdentifier nameIdentifierScheme="ORCID" nameIdentifierURI="https://orcid.org/0000-0004-0005-0006">
        0000-0004-0005-0006</jpcoar:nameIdentifier>
    <jpcoar:creatorName xml:lang="ja">田中, 太郎</jpcoar:creatorName>
    <jpcoar:creatorName xml:lang="en">Tanaka, Taro</jpcoar:creatorName>
    <jpcoar:creatorName xml:lang="ja-Kana">タナカ, タロウ</jpcoar:creatorName>
    <jpcoar:affiliation>
        <jpcoar:nameIdentifier nameIdentifierScheme="kakenhi">12604</jpcoar:nameIdentifier>
        <jpcoar:affiliationName xml:lang="en">Tokyo Gakugei University</jpcoar:affiliationName>
    </jpcoar:affiliation>
</jpcoar:creator>
```

属性：①識別子スキーマ

②識別子IDをURIで表現したもの

…科学研究費助成事業の機関番号

同一人物の著者名、英語表記、ヨミ、識別子、所属機関情報をグループ化して流通させることができる

JPCOARスキーマ要素

作成者、寄与者に関わる識別子

個人を示す識別子

統制語彙	解説
e-Rad	府省共通研究開発管理システム（e-Rad）の研究者番号
NRID	KAKEN研究者ページのID
ORCID	ORCID
ISNI	国際標準名称識別子
VIAF	バーチャル国際典拠ファイル識別子
AID	NACSIS-CAT著者名典拠レコードID
kakenhi	科学研究費助成事業の機関番号
Ringgold	Ringgold Identifier （Ringgold社の学術機関識別システムの識別子）
GRID	GRID（Global Research Identifier Database）の識別子

所属機関を示す識別子

統制語彙	解説
kakenhi	科学研究費助成事業の機関番号
ISNI	国際標準名称識別子
Ringgold	Ringgold社の学術機関識別システムの識別子
GRID	Global Research Identifier Databaseの識別子

JPCOARスキーマ要素

寄与者の役割を示す属性

役割属性(ContributorType)はDataCite (※) の統制語彙を採用

該当する語彙が無い場合は、”Other”を選択する

統制語彙	解説
ContactPerson	連絡担当者
DataCollector	データ収集者
DataCurator	データキュレーター
DataManager	データ維持管理者
Distributor	頒布者
Editor	編集者
HostingInstitution	提供機関
Producer	制作者
ProjectLeader	プロジェクトリーダー

統制語彙	解説
ProjectManager	プロジェクト管理者
ProjectMember	プロジェクトメンバー
RelatedPerson	関係者
Researcher	調査者
ResearchGroup	調査グループ
Sponsor	スポンサー
Supervisor	監督者
WorkPackageLeader	ワークパッケージ管理者
Other	その他

DataCite: https://datacite.org 研究データの正確な特性・引用を目指す国際機関

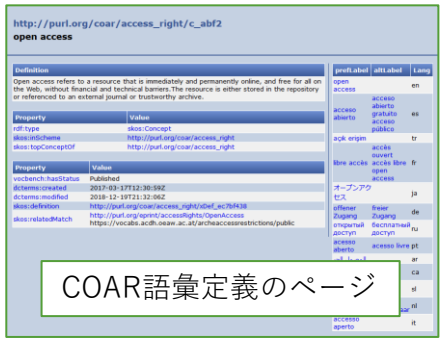
JPCOARスキーマ要素

アクセス権、資源タイプ

統制語彙を値に記入する要素（COAR※の統制語彙を採用）。

属性(rdf:resource)には、選択した統制語彙に対応する**URI**を記入する。

※オープンアクセスリポジトリ連合（COAR） https://www.coar-repositories.org



例：アクセス権

属性　　統制語彙に対応するCOAR Access RightsのURI

<dcterms:accessRights **rdf:resource="http://purl.org/coar/access_right/c_abf2"**>

open access</dcterms:accessRights>

統制語彙：4種類から選択して記入

- open access
- metadata only access
- embargoed access
- restricted access

例：資源タイプ

<dc:type **rdf:resource="http://purl.org/coar/resource_type/c_6501"**>

属性　　統制語彙に対応するCOAR Resource TypeのURI　**journal article**</dcterms:accessRights>

統制語彙：**資源タイプ語彙別表**から選択して記入

JPCOARスキーマ要素

APC Article Processing Charge

オリジナルのコンテンツに関するAPCの状態を記入。

RIOXXのスキーマを採用。

要素の値は統制語彙から選択。

統制語彙	解説
Paid	支払済み
Fully waived	全て免除
Not required	不要
Partially waived	一部免除
Not charged	無料
Unknown	不明

（例）

統制語彙

<rioxxterms:apc>Paid</rioxxterms:apc>

RIOXX： http://rioxx.net/guidelines
英国研究会議、英国高等教育助成会議のオープンアクセス方針に準拠したメタデータ仕様

JPCOARスキーマ要素

権利情報

コンテンツの、利用に関する権利情報を属性(xmi:lang)と共に記入。

知的所有権や著作権等に関する情報を含む。

ライセンスにURIがある場合、属性(rdf:resource)に記入。

例：クリエイティブコモンズライセンスの記入

<dc:rights xml:lang="en"
rdf:resource="**https://creativecommons.org/licenses/by/4.0/deed.en**">
Creative Commons Attribution 4.0 International</dc:rights>

URI

ライセンスの正式名とバージョン情報を示す
※省略形で記入しない

公開サイトでのアイコン表示が可能となり、言語に依存しない権利情報の表現が可能となる

JPCOARスキーマ要素

日付<datacite:date>

コンテンツに関する日付、コンテンツの個別ファイルに関する日付。
（DataCiteの統制語彙を採用）

属性(dateType)の記入が必須。
属性の値は右表の統制語彙から選択。

日付は、ISO-8601の3形式で記入。
- YYYY-MM-DD
- YYYY-MM
- YYYY

＊権利情報として”embargoed access”を指定した場合は
属性に”Available”を指定し、利用開始日を記入
＊学位授与年月日は別要素

統制語彙	解説
Accepted	受理日
Available	利用開始日
Collected	収集日
Copyrighted	著作権発効日
Created	作成日
Issued	発行日
Submitted	提出日
Updated	最終更新日
Valid	有効期日

JPCOARスキーマ要素

出版タイプ

出版物の版に関する情報を統制語彙から選択して記入。

記入の際は、属性(rdf:resource)の値として、統制語彙に対応する。

COAR VersionのURIを併せて記入することが必要。

例：著者最終稿（査読後、最終的に出版社に受理された原稿）の場合

<oaire:version **rdf:resource**="**http://purl.org/coar/version/c_ab4af688f83e57aa**">**AM**</oaire:version>

統制語彙

COAR Version URI

統制語彙の一覧・説明と各COAR Version URIは
JPCOARスキーマガイドラインページの「出版タイプ」参照

＊論文の場合は記入が必須
＊メタデータのみの場合は、出版タイプを記入せずにアクセス権(dcterms:accessRights)に記入
＊データのバージョン情報については「バージョン情報」要素(datacite:version)を使用する

JPCOARスキーマ要素

識別子

コンテンツ自身のユニークなIDを記入する要素。

識別子の種類は、属性（identifierType）に統制語彙を選択して記入する。

要素の値は、識別子の種類に応じた形式で記入。

統制語彙	解説
DOI	デジタルオブジェクト識別子（DOI：Digital Object Identifier）
HDL	ハンドルシステム識別子（Handle URL）
URI	統一資源識別子（URI：Uniform Resource Identifier）

※JaLC DOIを登録する場合は「ID登録」の要素にも、登録DOIを「prefix/suffix」の形式で記入

例：DOI（デジタルオブジェクト識別子）

<jpcoar:identifier **identifierType="DOI"**> **https://doi.org/10.18926/AMO/54590** </jpcoar:identifier>

統制語彙　　URI

JPCOARスキーマ要素

関連情報（関連識別子・関連名称）

登録するコンテンツと関連するコンテンツ間の関連性を記入。

属性(relationType)の値には、関連性を統制語彙から選択して記入。

関連コンテンツは識別子（種類を属性(identiferType)で指定）での記入が望ましい

例：

～の異版である

```
<jpcoar:relation relationType="isVersionOf">
    <jpcoar:relatedIdentifier identifierType="DOI">
      https://doi.org/10.1371/journal.pone.0170224
    </jpcoar:relatedIdentifier>
</jpcoar:relation>
```

関連コンテンツの情報

登録コンテンツと関連コンテンツの関連性

※関連コンテンツは識別子で記入することが望ましいが、
識別子が無い場合は言語属性を指定して名称を記入できる。

JPCOARスキーマ要素

助成情報

コンテンツ作成に研究助成を受けた場合、利用した助成情報を記入。

助成機関の情報（識別子、機関名称）、研究課題（課題番号、課題名）を記入。

属性(funderIdentifireType)の統制語彙

統制語彙	解説
Crossref Funder	Crossref Funder Registryの助成機関識別子
GRID	GRID（Global Research Identifier Database）の識別子
ISNI	国際標準名称識別子（ISNI：International Standard Name Idenitifier）
Other	その他

助成機関の識別子は、属性(funderIdentifierType)で識別子の種類を統制語彙で指定し、識別子のURIを記入する。

研究課題番号は、属性(awardURI)に助成団体が提供する当該助成に関する詳細情報のURIを指定した上で、記入する。
※科研費の場合はKAKENデータベースの当該研究のページのURLを指定。

助成機関の名称や研究課題名は、言語属性(xml:lang)を指定して記入。

JPCOARスキーマ要素

ファイル情報

コンテンツ本体に関する情報を記入。

```
<jpcoar:file>
本文URL… <jpcoar:URI objectType=“fulltext” label=“abcd.pdf”>
                https://xxx/xxxx/xxxx/abcd.pdf</jpcoar:URI>
フォーマット…<jpcoar:mimeType>application/pdf</jpcoar:mimeType>
サイズ…<jpcoar:extent>3M</jpcoar:extent>
日付…<datacite:date dateType="Issued">2015-10-01</datacite:date>
バージョン情報…<datacite:version>1.2</datacite:version>
</jpcoar:file>
```

本文URL
- 属性(objectType、label)と共に記入
- objectTypeの値は以下の統制語彙を選択
  abstract / summary / fulltext / thumbnail / other
- labelはファイル名がある場合、記入する

バージョン情報
- 研究データの場合のみ使用
- バージョンの変更がある場合は、記入が望ましい

JPCOARスキーマ要素

博士論文に関する諸要素

- **アクセス権** <dcterms:accessRights> ＝ 統制語彙 **"open access“**
- **資源タイプ** <dc:type> ＝統制語彙 **“doctoral thesis”**
- **学位授与番号** <dcndl:dissertationNumber> → **学位授与番号**を値に記入
- **学位名** <dcndl:degreeName> → **学位規則による学位の種別名と分野名**を値に記入
- **学位授与年月日** <dcndl:dateGranted> → ISO-8601 で規定する3 形式で値に記入
- **学位授与機関** <jpcoar:degreeGrantor> → **識別子**と**名称**を値に記入
- **本文URL** <jpcoar:URI>→ **URL**を記入、

  objectType属性 →収録本文によって統制語彙 **“fulltext”(本文あり)、**

  “summary”(要約）、“abstract”(要旨) を選択記入

学位名と**学位授与機関**は
言語属性の指定を推奨

属性の記入は、収録本文の
見分けやすさにつながる

JPCOARスキーマをもっと知りたい、確認したいときに…

JPCOARスキーマガイドライン

スキーマ説明
JPCOARスキーマについて、
詳細な入力ルールを確認できる。

FAQ・問い合わせ
不明点や疑問点はこちらで…
解決しない場合はお問い合わせください。

関連サイト
XMLスキーマ定義（GitHub）、
連携する外部スキーマ情報は
こちらから…

https://schema.irdb.nii.ac.jp/ja

さいごに
JPCOARスキーマのこれから

国内外の学術情報流通の変化に応じ、年1回程度改訂

2017年10月　Version 1.0（初版）公開

2018年 8月　Version 1.0.1公開
- OpenAIREのガイドライン改訂に伴う出版タイプの要素名変更等

2019年11月　Version 1.0.2公開
- COAR Resource Type Vocabulary Version 2へ改訂に伴う資源タイプ語彙追加等

＊要素拡充・改訂を継続
＊デジタルアーカイブ対応の検討、
識別子・ライセンスに関する検討が今後の課題

補足：10月26日にいただきましたご質問への回答

※後日回答させていただくこととなっておりましたご質問について、補足します。

ご質問１：
現在、junii2に基づきメタデータを作成していますが、早くJPCOARスキーマに修正したほうが良いでしょうか？
著者版フラグ「ETD」と入力している項目が、IRDB上では自動的に「VoR」（出版タイプ）>と表示されています。
本来「AM」とすべきものも一律「VoR」に変換されてしまうので、こちらのメタデータを>JPCOARスキーマに合わせて修正したほうが良いでしょうか。

回答：
いくつか項目を整理して回答させていただきます。

(1)IRDB上での変換内容について
junii2で著者版フラグ「ETD」と入力した場合、IRDB上で自動的に「VoR」（出版タイプ）と表示されることは、正常な変換内容となります。

※junii2に基づき作成いただいたメタデータの、IRDB上での変換内容は、「IRDB 学術機関リポジトリデータベース サポート」ウェブサイトの「エラーチェック・正規化仕様」のページhttps://support.irdb.nii.ac.jp/ja/harvest/junii2/mappingに掲載の「マッピング・エラーチェック・正規化仕様_junii2.pdf」で確認が可能です。
「連番60　著者版フラグ」の列の「JPCOARへの変換時対応（マッピング）」をご参照ください。

(2)junii2における著者版フラグの指定内容について
junii2で著者版フラグを指定される場合、
・博士論文の本文を登録される場合は「ETD」を入力します。
・博士論文に収録された投稿論文を登録される場合、著者版原稿でしたら「author」、出版社版原稿でしたら「publisher」を入力します。
登録対象のコンテンツがどういったものかについて、まずはjunii2の枠組みでご確認ください。

(3)JPCOARスキーマの適用を急ぐべきかどうかについて
まず、登録する対象コンテンツをどのようなものとして登録なさりたいかについて、上記(1)、(2)の内容を再度ご確認ください。
また、貴館様のリポジトリ全体における登録コンテンツの状況も併せてご考慮いただき、JPCOARスキーマの適用を急ぐかどうかを、ご検討いただければと思います。

ご質問１：
Crossref Funder Registryで助成機関の識別子を調べる方法を解説いただけないでしょうか。

回答：
Crossref Founder IDについては、CrossrefからAPIおよびCSVファイルが提供されております。このため、Crossrefのページから手動で検索するのではなく、論文投稿システムやリポジトリシステム等からAPIまたはCSVで連携して検索するという形となります。

※Crossref FounderデータのAPI、CSVファイルの提供について
https://www.crossref.org/education/funder-registry/funder-data-via-the-api/

※連携システムの一例　"CHORUS"
https://current.ndl.go.jp/node/39562

ご清聴ありがとうございました