

第1回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係:研究データの管理と流通から考える」

新たな学術情報流通において JPCOAR スキーマが果たす役割

片岡 朋子

(お茶の水女子大学/JPCOAR メタデータ普及タスクフォース)

講演要旨



近年、公的研究助成を受けた学術成果および関連する研究データの公開が助成団体や政府の方針として推進されている。公開された学術成果の発見や利活用を促進するためには、研究データ等の新たな学術資源や国際的な流通に対応した標準的なメタデータ交換フォーマットが必要となる。JPCOAR ではこうした社会的要請に対応するため、機関リポジトリのメタデータ交換フォーマットである junii2 を改訂し、JPCOAR スキーマを策定した。本発表では、新たな学術情報流通において JPCOAR スキーマが果たす役割について紹介する。



片岡 朋子

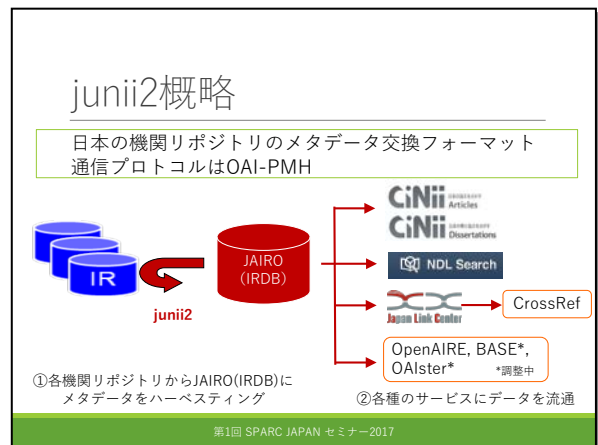
2015年よりお茶の水女子大学図書・情報課に勤務。情報基盤担当として大学全体の事務システムの運用支援を担当。2016年よりJPCOARスキーマの策定に携わっている。

本日は、私が所属している JPCOAR メタデータ普及タスクフォースが策定を進めている、新しいメタデータフォーマットである JPCOAR スキーマと、これからの学術情報流通において、この新しいメタデータフォーマットがどのように活用できるのかという事例についてご紹介します。

1.JPCOAR スキーマの概要

JPCOAR スキーマは、junii2 という日本の機関リポジトリのメタデータ交換フォーマットの代わりとなるものとして策定されました。junii2 は、図 1 にあるように、各大学の機関リポジトリから JAIRO (IRDB) にメタデータをハーベスティングする際に用いられて

いるメタデータの交換フォーマットです。この JAIRO に収集されたデータは、CiNii や NDL Search、JaLC といったサービスにも流通しており、それぞれのサービ



(図 1)

スや DOI の登録などにメタデータが活用されるという流れになっています。

近年では、こうした国内のサービスだけではなく、OpenAIRE や BASE、OAIster といった海外のサービスにもデータが流通する機会が増えており、この junii2 は日本独自のメタデータのフォーマットのため、こうした国際的なフォーマットとの間にギャップが生じていました。

また、本日の SPARC Japan セミナーのテーマでもある、研究データといった新しいコンテンツへの対応を進めるために、今回、junii2 を改訂し、JPCOAR スキーマを策定するという流れになっています。

図 2 は JPCOAR スキーマの検討メンバーです。昨年度からタスクフォースを設置して検討を進めていました。国立情報学研究所の研究者と職員に加え、各大学の図書館員と一緒に検討する形となっています。昨年度は新しいメタデータフォーマットの方向性に関する検討や国際的な動向調査も行い、JPCOAR スキーマ案としてまとめたものをパブリックコメントとして公開しました。

今年度は、そのパブリックコメントで頂いたご意見の反映や、国際的な連携先との最終的な調整を経て、現在この JPCOAR スキーマの確定版を策定する作業を進めています。この確定版の JPCOAR スキーマは来月初旬に公開する予定です。

JPCOARスキーマの検討メンバー

2016年度メタデータ検討タスクフォース
主査：高橋 菜奈子 (千葉大学)
協力員：佐々木 翼 (北海道大学)、前田 朗 (東京大学)、南山 泰之 (国立極地研究所)、香川 朋子 (お茶の水女子大学)、大園 隼彦 (岡山大学)、林 豊 (九州大学)
国立情報学研究所：片岡 真、田口 忠祐、大向 一輝、山地 一禎

2017年度メタデータ普及タスクフォース
主査：高橋 菜奈子 (千葉大学)
作業部会員：石田 唯 (東京大学)、片岡 朋子 (お茶の水女子大学)、松村 友花 (神戸大学)、林 豊 (九州大学)、前田 朗 (東京大学)
国立情報学研究所：片岡 真、田口 忠祐、大向 一輝、山地 一禎

第1回 SPARC JAPAN セミナー2017

(図 2)

2. junii2 改訂の基本方針

junii2 の改訂に当たり、主に三つの基本方針に沿って検討を進めてきました。1 点目は、近年、学術機関に求められているオープンサイエンス・オープンアクセス方針に対応した要素の拡充です。2 点目は、識別子の拡充・メタデータ構造の修正です。3 点目は、近年、国際的な学術情報流通の機会が増えており、日本独自の定義と国際的に主流な定義との間にギャップが生じていたため、国際的によく使われているメタデータの定義を採用し、相互運用性の向上を図ることで

2-1. オープンサイエンス・オープンアクセス対応

ここからは、それぞれの基本方針に対応して、具体的にどのような改訂を行ったのかご紹介していきます。

1 点目については、近年では政府によるオープンサイエンスに関する検討や、助成機関や大学によるオープンアクセスの方針の公開が進むなど、オープンサイエンス・オープンアクセスに対応するという社会的な要請が高まっています。このような要請に応えるために、学術成果をより公開しやすくなる要素や、研究データを管理するための要素の拡充を行いました。

追加したものは、研究助成を受けた場合の助成機関情報を記述する要素です (図 3)。内容としては、助成機関そのものに関するものと、研究課題に関するものとで構成されており、それぞれ助成機関名と研究課題名という文字列による情報だけではなく、助成機関の識別子、また研究課題番号を関連させて記述ができ

助成機関情報の拡充

- 助成機関情報
 - 助成機関識別子
 - 助成機関名
 - 研究課題番号
 - 研究課題名

助成機関識別子の種類

- Crossref Funder
- GRID (Global Research Identifier Database)
- ISNI (国際標準名称識別子)
- other

```
<datacite:fundingReference>  
<datacite:funderIdentifier funderIdentifierType="Crossref Funder">  
  https://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求と高次生命現象</datacite:awardTitle>  
<datacite:awardTitle xml:lang="en">Post-transcriptional regulation associated with RNA modifications responsible for  
  higher order biological processes</datacite:awardTitle>  
</datacite:fundingReference>
```

(図 3)

るようにしています。こうしておくことで、このデータが他のデータベースやサービスに渡った際にも、助成情報に関する分析や活用が進むようになると考えています。

もう一つのポイントとして、研究データを単純に公開するだけではなく、適切なライセンスと一緒に公開することで、その研究データの共有や再利用がより促進されるのではないかと考えています(図4)。そのために権利情報と権利を持っている権利者に関する情報の二つを設定しており、それぞれ文字列で権利情報を書くだけではなく、権利情報の方は、例えばクリエイティブ・コモンズといった公開された標準的なライセンスも一緒に書けるようになっており、権利者の方は権利者の識別子を一緒に記述できるようにしています。

もう一つの側面としては、公的な研究助成を受けた学術成果は原則、公開を進めるようにという方向にな

ってきているため、オープンアクセスの状態を確認できるアクセス権という要素と、論文処理費用(APC)の支払い状況を確認できる要素の二つを追加しました(図5)。それぞれ、統制語彙と呼ばれる、決められた語彙で記述する形式で追加しています。

特に研究データは、細やかなデータのバージョン管理が求められる場合があるので、メタデータの単位と研究データの個別のファイル単位それぞれで、バージョンの情報を記述できるようにしました(図6)。また、その研究データを採取した場所や研究データが示している内容の位置情報を経度や緯度を使って記すことも、文字列で書くこともできるようにしています。

もう一つ、研究データに対応するために、寄与者の役割を属性として記述できるようにしました(図7)。研究データの場合は、データの作成・収集など、論文の執筆よりも多くの人に関係してくる場合が想定され

ライセンスの付与・表示

データの再利用を促進するためには、適切なライセンスの付与・明示が必要

- 権利情報


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

ライセンスのURI
- 権利者情報


```
<jpcoar:rightsHolder>
<jpcoar:nameIdentifier nameIdentifierScheme="ISNI" nameIdentifierURI="http://www.isni.org/isni/00000004043815">0000000404381592</jpcoar:nameIdentifier>
<jpcoar:rightsHolderName xml:lang="en">American Physical Society
</jpcoar:rightsHolderName>
</jpcoar:rightsHolder>
```

権利者識別子

第1回 SPARC JAPAN セミナー2017

(図4)

研究データに対応した要素

研究データのバージョン(メタデータ単位/ファイル単位)やデータを収集またはそのデータが対象とする位置情報を記述可能な要素を追加

- バージョン情報


```
<datacite:version>1.2</datacite:version>
```
- 位置情報
 - 位置情報(点)
 - 緯度
 - 経度
 - 位置情報(空間)
 - 西部経度
 - 東部経度
 - 南部緯度
 - 北部緯度
 - 位置情報(自由記述)

位置情報(空間)の例:

```
<datacite:geolocationBox>
<datacite:westBoundLongitude>-71.032</datacite:westBoundLongitude>
<datacite:eastBoundLongitude>68.211</datacite:eastBoundLongitude>
<datacite:southBoundLatitude>41.090</datacite:southBoundLatitude>
<datacite:northBoundLatitude>42.893</datacite:northBoundLatitude>
</datacite:geolocationBox>
```

位置情報(自由記述)の例:

```
<datacite:geolocationPlace xml:lang="en">Disko Bay
</datacite:geolocationPlace>
</datacite:geolocation>
```

第1回 SPARC JAPAN セミナー2017

(図6)

公的研究助成を受けた学術成果のオープン化促進

オープンアクセス(OA)の達成度を把握するため、OAの状態やAPCの支払い状況を記述できる要素を追加

- アクセス権


```
<coar:accessRight rdf:resource="http://purl.org/coar/access_right/c_abf2">
open access</coar:accessRight>
```

 - embargoed access: エンバゴ有
 - metadata only access: メタデータのみ
 - open access: オープンアクセス
 - restricted access: アクセス制限有
- APC


```
<rioxterms:apc>Paid</rioxterms:apc>
```

 - 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

第1回 SPARC JAPAN セミナー2017

(図5)

研究データに対応した要素

研究データの収集/分析等に貢献した寄与者の役割を示す属性を追加(contributorType)

ContactPerson	連絡担当者	ProjectManager	プロジェクト管理者
DataCollector	データ収集者	ProjectMember	プロジェクトメンバー
DataCurator	データキュレーター	RelatedPerson	関係者
DataManager	データ維持管理者	Researcher	調査者
Distributor	頒布者	ResearchGroup	調査グループ
Editor	編集者	Sponsor	スポンサー
HostingInstitution	提供機関	Supervisor	監督者
Producer	製作者	WorkPackageLeader	ワークパッケージ管理者
ProjectLeader	プロジェクトリーダー	other	その他

第1回 SPARC JAPAN セミナー2017

(図7)

ます。そうした人たちの役割を属性として記述しておくことで、その貢献度を明示したり、適切な評価に結び付けたりするといった活用方法もできると思います。こちらも統制語彙の中から選択する形で記述するようになっていきます。

2-2.識別子の拡充・メタデータ構造の修正

改訂方針の2点目については、情報をより正確に記述できるように、識別子の拡充とメタデータ構造の修正を行いました。例えば、作成者と寄与者が複数名いて、それぞれが識別子や所属機関の情報を持っていたときに、それを並列に書いてしまうと、どの情報がどの人に結び付いているか、判別するのが難しいケースが生じてしまいます。そうした課題に対応するために、JPCOAR スキーマでは、関連する情報同士をグループ化して、階層関係を持たせて記述できるようにしました。

図8の例では、第2階層のレベルで人の情報を書き、第3階層のレベルでその人が所属する機関の情報を記述しています。こうしておくことで、この方が複数の機関に所属している場合も、第3階層だけを繰り返して記述するという書き方ができるようになっています。こうした構造化は他の要素でも必要に応じて行っています。

もう一つの改善事項としては、識別子と関連性に関する記述を増やしています。識別子というのは情報を一意に特定し得るものです。図9はメタデータ単位に

付与するものの例です。メタデータ自体に付与されたDOI、HDL、URIは資源識別子、外部のデータベースやバージョン違いの研究データに対して付与される識別子は関連識別子という2種類を用意しています。

また、研究データの場合は、バージョン違いの研究データにリンクさせたり、その研究データを用いて執筆された本や論文にリンクさせたりというように、関連に関する記述が増えることが予想されるため、関連識別子の関連性の種類も若干増やしています。

2-3.国際的に相互運用性の高いデータ交換のためのスキーマ定義

改訂方針の3点目に関連して、新しいメタデータフォーマットを策定するに当たり、国際的な動向調査を行いました。ただ、現在のところ、一つのメタデータフォーマットに準拠すれば全ての課題が解決できるという状況にはなかったため、JPCOAR スキーマは複数の海外などの定義を参考にした、複合的なメタデータ定義となっています。

具体的には、図10の7種類のスキーマ定義からそれぞれ要素を採用しており、研究データについては、研究データの国際的なスタンダードとなっているDataCiteの要素を採用しています。その他にも、Dublin Core、COAR、OpenAIREの統制語彙を採用しました。The Bibliographic Ontologyからは雑誌の書誌記述に関する要素を採用しており、RIOXXからはAPCの要素を採用しています。また、国内では、国

識別子の拡充・メタデータ構造の修正

```

<jpcoar:contributor contributorType="Editor">
  <jpcoar:nameIdentifier nameIdentifierScheme="ORCID" nameIdentifierURI="https://orcid.org/0000-0001-0002-0003">0000-0001-0002-0003</jpcoar:nameIdentifier>
  <jpcoar:contributorName xml:lang="ja">山田, 一郎</jpcoar:contributorName>
  <jpcoar:contributorName xml:lang="en">Yamada, Ichiro</jpcoar:contributorName>
  <jpcoar:affiliation>
    <jpcoar:nameIdentifier nameIdentifierScheme="kakenhi">12601</jpcoar:nameIdentifier>
    <jpcoar:affiliationName xml:lang="en">University of Tokyo</jpcoar:affiliationName>
  </jpcoar:affiliation>
</jpcoar:contributor>
<jpcoar:contributor contributorType="Data Manager">
  <jpcoar:nameIdentifier nameIdentifierScheme="ORCID" nameIdentifierURI="https://orcid.org/0000-0004-0005-0006">0000-0004-0005-0006</jpcoar:nameIdentifier>
  <jpcoar:contributorName xml:lang="ja">田中, 太郎</jpcoar:contributorName>
  <jpcoar:contributorName xml:lang="en">Tanaka, Taro</jpcoar:contributorName>
  <jpcoar:affiliation>
    <jpcoar:nameIdentifier nameIdentifierScheme="kakenhi">12611</jpcoar:nameIdentifier>
    <jpcoar:affiliationName xml:lang="en">Ochanomizu University</jpcoar:affiliationName>
  </jpcoar:affiliation>
</jpcoar:contributor>
  
```

第1階層
第2階層
第3階層

識別子 (研究者/所属機関)

第1回 SPARC JAPAN セミナー2017

(図8)

識別子と関連性記述の拡充

識別子や関連性の統制語彙を拡充し、情報をより正確かつ機械的に処理可能に

- 資源識別子
 - DOI / HDL / URI
- 関連識別子
 - 識別子の種類:
 - ARK - arXiv - DOI - HDL - ICHUSHI - ISBN - J-GLOBAL - Local - PISSN - EISSN - NAID - PMID - PURL - SCOPUS - URI - WOS
 - 関連性の種類:
 - isVersionOf - hasVersion - isPartOf - hasPart - isReferencedBy - references - isFormatOf - hasFormat - isReplacedBy - replaces - isRequiredBy - requires - isSupplementTo - isSupplementedBy - isIdenticalTo - isDerivedFrom - isSourceOf

第1回 SPARC JAPAN セミナー2017

(図9)

立国会図書館の DC-NDL の学位に関する要素を取り入れています。

このように、できる限り外部の標準的な定義を取り入れる方向でまとめていますが、一部では、日本独自の事情や junii2 との互換性を保つために、どうしても独自に拡張せざるを得なかった要素も残っている状態です。

以上が JPCOAR スキーマの簡単な概要になります。

3.JPCOAR スキーマが果たす役割と可能性

ここからは、JPCOAR スキーマにどのような可能性があり、今後どのような役割を果たしていけるかというお話をします。

1 点目は、国際的な学術情報流通の促進です。現在の主要な連携先である OpenAIRE でも新しいガイドラインの改訂を行っており、お互いに情報共有をしな

がら改訂作業を進めています。かなり近い形にまとめられてきていますので、相互運用性は向上したと思います。また、研究データについては、DataCite という国際的なスタンダードから定義を採用しているため、もし今後、新たに研究データに関する連携先が増えた場合でも、ある程度の互換性は保てると思います。その他にも、図 11 のような主要な連携先とのマッピング表（対応表）をつくって公開する予定ですので、相互運用性の向上は図られていると思います。

2 点目は、国内サービスとの連携促進です。JPCOAR スキーマは機関リポジトリ向けのフォーマットとなっており、そのプラットフォームである JAIRO についても、今年度末から来年度にかけて、JPCOAR スキーマに対応した改修が行われる予定です。研究データを扱うようになることで、こうした従来のリポジトリの連携先であった CiNii Articles などだけではなく、より新たなサービスが連携先の候補として拡大するのではないかと考えています。

例えば、図 12 のように、KAKEN のような成果系のデータベースには助成関係の情報や研究課題、その研究課題によって生み出された本や論文の情報が登録されていますし、NACSIS-CAT、CiNii Books には、研究データから生み出された成果物としての本の情報が登録されることも考えられます。ERDB-JP という、国内の電子リソースの情報を集めているサービスにも、研究データからつくられた学術論文が収録されている電子ジャーナルの情報が収録されるという可能性も考

参照したスキーマ定義	
Dublin Core	Dublin Core Metadata Initiativeが維持管理する基本的な要素セット
COAR	オープンアクセスリポジトリ連合 (COAR) が策定するセマンティックWebを意識した語彙セット
OpenAIRE	EU Horizon 2020政策下で国際的に研究成果を収集するプロジェクトが使用する要素
DataCite	研究データにDOIを付与し、正確な特定・引用を目指す国際組織が定義するスキーマ定義
The Bibliographic Ontology	書誌事項をセマンティックWebに対応した形式 (RDF) で記述可能とするスキーマ定義
RIOXX	英国Jiscの助成を受けて策定されたスキーマ定義
DC-NDL (国立国会図書館)	国立国会図書館が提供するThe Dublin Core Metadata Element Setを拡張したメタデータ記述語彙および記述規則
JPCOARスキーマ (独自定義)	標準的なスキーマ定義では不足する部分をJPCOARが独自に拡張したもの

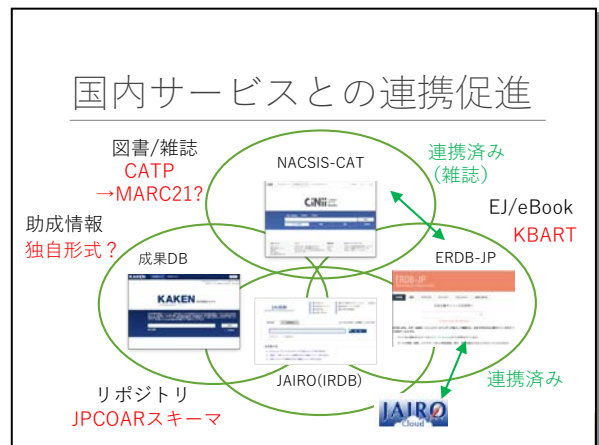
第1回 SPARC JAPAN セミナー2017

(図 10)

国際的な学術情報流通の促進	
主要連携先とのマッピングの提供 (予定)	
oai_dc	
OpenAIRE Guidelines	<ul style="list-style-type: none"> Literature Repositories Data Archives 3.0
DC-NDL	<ul style="list-style-type: none"> RDF Simple
JaLC	<ul style="list-style-type: none"> 汎用データ ジャーナルアーティクル 書籍 e-learning 研究データ
junii2	<ul style="list-style-type: none"> junii2→ JPCOARスキーマ JPCOARスキーマ→junii2

第1回 SPARC JAPAN セミナー2017

(図 11)



(図 12)

えられます。このように、従来からこのようなサービスは割と近い関係にあり、一部ではシステム連携も進んできていますが、そこに研究データが加わることによって、さらに関連性が高まっていくのではないかと思います。

ただ、それぞれのサービスで使われているメタデータのフォーマットは全く異なっており、連携を進めるに当たってはマッピングや対応表をつくる作業が必要になってくると思うので、今回の JPCOAR スキーマの策定の過程で得られた経験や課題を共有することによって、そうしたサービスの連携の促進につながればと考えています。

3 点目は、新しいサービスの展開です。研究データについては、データを管理をするだけではなく、それを広く一般に公開し、発見したりデータを再利用したりという仕組みが必要になってくるかと思えます。図 13 は現在検討中のデザインのイメージ図を借りてきたもので、例えば従来の CiNii にはなかった「すべて」という検索メニューや、「研究」「研究データ」「人物」といったメニューが増えていることがお分かりいただけると思います。

「すべて」というメニューで検索した場合、キーワードに関連した研究データ、論文、本、人物、研究の課題といったあらゆる情報が一度に検索される予定です。従来の CiNii では、本は本、論文は論文という形で個別に検索するサービスでしたが、そこに研究データが加わることによって、本や論文やデータが関連性

をより高めてくると思いますので、このような網羅的な検索に対する需要も高まってくると予想しています。

「研究」というメニューで検索した場合、その研究に関連する研究データへのリンクや、関連するサービスである KAKEN へのリンク、またその研究によって生み出された文献・研究データの情報が掲載される予定です。さらに関連書籍、先行研究、被引用文献といった情報や、その他にも関連する論文、図書、研究というように、従来のサービスに比べて、かなり充実した内容の情報が掲載されます。

今後、研究データを扱うようになるなど、学術成果の情報量は増えていく傾向にある中で、関連する情報まできっちりサービスを提供することはなかなか難しくなってくると思いますが、例えばメタデータの中のある文字列からこうしたリンクを形成することよりも、メタデータの中に識別子や URI を含めておくことで、より正確な関連性の情報を提供できるのではないかと考えています。

「研究データ」で検索した場合、研究データへのリンクやクリエイティブ・コモンズといったライセンスへのリンクが表示される予定です。ライセンスの情報を文字列ではなく URI の形式で記述しておくことによって、サービスでもこのように正確なライセンスへのリンクが表示されるようになると思います。また、その研究データに基づく研究や論文、被引用文献、関連する情報も併せて表示される予定です。

「人物」で検索した場合、その人物が執筆した論文・本、関連する人物、関連する情報が表示される予定です。特に日本人の場合は、日本語や英語など複数の言語で著者名が表記される場合が多いと思いますので、メタデータにきちんとした識別子が含まれるかどうかによって正確な情報の集約が可能になると思います。

JPCOAR スキーマでは、国内的な普及状況を考えて、ORCID や ISNI といった国際的な識別子に加えて、日本で流通している e-Rad (科研費など日本の公的な助成事業で使われている研究者番号) や研究者リゾルバ



(図 13)

一の ID も記述できるようにしています。

ただ今後、国際的な学術情報流通を考えた場合には、やはり ORCID のような国際的な識別子を付与しておくことの方がより望ましいため、今後の国内における識別子の普及状況にも注意を払う必要があると考えています。

4 点目は、研究データ管理の基盤としての波及効果です。junii2 は約 15 年間にわたって機関リポジトリのメタデータ交換フォーマットとして、かなり幅広く活用されてきました。現在のところ、約 645 の機関リポジトリが junii2 の形式に対応しているといわれており、JPCOAR スキーマはこの junii2 の基盤を継承する形となります。しばらくは並行運用の期間を設けますが、最終的には JPCOAR スキーマに移行していただく予定になっており、移行するときにまず、この機関リポジトリの管理者である図書館員がどうしたら JPCOAR スキーマに対応できるのかについて考える一つの機会になると思っています。

例えば、助成情報をどうしたら効率的に集められるのかといったことや、研究データにはライセンスを適切に付与した方が良く、メタデータに識別子・URI などを含んでおいた方が望ましいといった一つ一つの考え方の変更点が、研究データの管理や流通に関する意識の喚起にもつながればと思っています。

また、図書館員は研究データを持っているわけではないので、図書館員だけでは研究データ管理ができません。実際にこの研究データを管理している研究者の実情について理解を深めたり、研究者とコミュニケーションを取ったりする一つのきっかけにいただければと考えています。

最後になりましたが、10 月 10 日に国立情報学研究所の会場をお借りして、JPCOAR スキーマの説明会を開催する予定です。ご興味のある方がいらっしゃいましたら、ぜひご参加をご検討ください。

●フロア 1 JST バイオサイエンスデータベースセンターの職員です。機関リポジトリ以外のサイトで JPCOAR のメタデータを付けて、それを例えば CiNii で検索できるようにすることは想定されているのでしょうか。

●片岡 現在のところは機関リポジトリがメインのターゲットになると思っていますが、やはり検討の過程で、例えばデジタルコレクションやデジタル化された図書についても、スキーマを考えていかなければいけないという課題は出ています。現段階ではリポジトリがメインですが、次の段階ではそうしたデジタルアーカイブや電子化された資料についても取り込めるようなことを考えています。

●フロア 2 専修大学の図書館員です。今日のお話はとても興味深いと思って聞いていたのですが、少し気になることがあります。今回、この junii2 を拡張した JPCOAR スキーマにいろいろ変更点が出てきていますが、それは機関リポジトリにデータを載せることを考えたから、junii2 の拡張のような形で考えられたのですか。

参考にされていた国際的な規格なども、データサイエンスに対応しているということによろしいですか。「機関リポジトリにこういうものが出ました」という話になって、「データサイエンスやデータも載せられるようになりました」と言うと、やはり「載せてください」となってしまうので、そのときに、別のサーバー、システムの方で立ち上げた方がいいのか、それともリポジトリの中に論文とそのデータが混在しているような形を想定されていたのかを聞きたいです。

今回、CiNii の方が、Research と Books と Articles が分かれているので、そういうことを考えたときに、各大学のリポジトリは共存してしまっているのか気になったので、そのあたりを教えてください。

●片岡 一番参考にした OpenAIRE のガイドライン

でも、Literature Repositories という文献のリポジトリに、DataCite といった研究データの要素を追加して、研究データを載せられるようにするという流れがあったため、従来のリポジトリに研究データを載せるという方向性が一番自然なのではないかと考えて、リポジトリのスキーマである junii2 を拡張して、JPCOAR スキーマを策定しています。

CiNii の方はまだ検討中ということですが、資料種別ごとに個別に検索したい方は個別の検索をしていただけますし、網羅的に検索をするということも考えられているようです。これからの利用者のニーズがどのようになっていくかにもよるかと思いますが、私の個人的な感想としては、この「すべて」という検索がより使われるようになると期待しているところです。

●**服部** SPARC Japan 事務局の服部です。CiNii の今の担当が出張に行っておりまして、前の前の担当としての発言をいたします。

CiNii Research では「すべて」だけではなく、論文 (Articles)、本 (Books) というくくりがまだ続くのかということですが、分けて利用したいというニーズがあるのだと思います。「論文を探したい」「本を探したい」ということもあると思うので、全て混ぜてしまうのはベストではないと思います。

まだ CiNii Research は検討の段階ですが、ここからは宣伝になります。11月7日火曜日の図書館総合展の10時からのフォーラムでは『次世代 CiNii』の展望」という企画を行います。9月20日ぐらいから参加申し込みが可能になりますので、この話の続きが気になるという人は、ぜひ申し込みをお願いいたします。

●**能勢** 宣伝をどうもありがとうございました。

私から一つ伺ってよろしいですか。この JPCOAR スキーマは先ほどおっしゃったように元の junii2 を拡張したという形で、例えば研究者として、データをオープンにして登録したいと考えて、各大学の研究データリポジトリにデータを載せる場合、提供するデータ

も、JPCOAR スキーマに即したものを登録していくという形になっていくのですか。

●**片岡** 機関リポジトリが JPCOAR スキーマに対応している場合は、やはり JPCOAR スキーマに近い形で渡していただくと図書館員もそのまま取り込めて助かる場所があると思います。ただ、JPCOAR スキーマはどちらかというと流通用に割と項目を限定して、最小限の項目にしてつくったものですので、管理用にもっとリッチな情報が必要な機関もあるかもしれません。やはりそこは各機関でどういう項目でデータを集めるかを決めていただければと思います。

●**能勢** そうなると、また先ほどの大澤先生の話にあった「いっぱい入力して、全部はねられて」という繰り返しになってしまうかもしれません。研究者側としては、やはり新しいスキーマなので、どういう情報が必要になってくるのかというのが知りたいところです。拡張したものは、助成機関の情報やオープンアクセスの状態、観測の位置などの情報が新たに必要になってきて、それを提供するということですね。

私は宇宙科学分野においてメタデータを自分で作成したことがあります。それはスペーススキーマという主に NASA でつくられたスキーマに従って書くのですが、この情報は必要でここに入力をしなければいけない、これは必須のマandatreeで、これはオプションとか、やはりメタデータのスキーマを理解していくことが非常に大変で、もう二度とやりたくないのです。

それでいったん理解すると、今度はマイナーバージョンアップして、研究者としてデータをオープンにしたいのですが、とてもではないけれど、そうするためのメタデータを作成する気力が残っていませんでした。

先の話なので分からないかもしれませんが、実際に運用して JPCOAR スキーマでデータを流通していった場合、各大学においてどういう形で研究者がメタデータの情報を提供していくとうまくいきそうかという

見込みはありますか。例えば、Excel シートでフォーマットに入力していけばいいぐらいであれば、それほど大変ではないのですが、「xml を書け」「xml エディタをつくって研究者でやってください」と言われても、まずやりたくないですね。

●片岡 そうですね。このデータを研究者の方から頂きたいと思っているわけではなく、恐らく、お使いの機関リポジトリのシステムによって収集の仕方は変わると思っています。例えば識別子だけ持っていれば、その他の情報は補完してくれるというサービスもあるかと思えます。それはリポジトリのシステム側の機能として私も少し期待したいと思っている部分があります。その辺の情報交換も含めて、10月10日に図書館員だけではなく、システムベンダーさんや研究者の方も含めた説明会と質疑応答する機会を設けておりますので、ぜひご参加いただければと思います。

●能勢 ご参加いただければと思います。各機関の研究データリポジトリが一般的になっていくと、junii2の継承もあって、JPCOAR がメインストリームになっていくだろうと考えていてよろしいですか。この先どういった情報を提供すれば代わりにメタデータをつくっていただけるかという観点から質問をしているのですが。

●片岡 JAIRO で日本の機関リポジトリのデータを全て収集しているため、そのフォーマットが JPCOAR スキーマになるということで、今後は JPCOAR スキーマを目指して対応いただけるとありがたいです。

●武田 国立情報学研究所の武田です。個人的な立場から発言します。

今おっしゃった、このメタデータは誰が作るのかというときに、システムベンダーを入れてという話ですが、そこは結構重要なところですね。junii2 ぐらいのレベルであれば、今まで図書館員が自分で手で書いて

いた、書き写していたといったことが多かったのではないかと思います。junii2 はヒューマンリーダブルな世界でつくっていたので、構造もフラットでした。

しかし、この junii2 から JPCOAR になったら、もうかなり発想が変わっています。JPCOAR になった途端に、人間は普通ではほとんど読めない形式になるわけです。読むようにトレーニングすると読めるようになるのですが、普通には読めない。そここのところに大きな差があって、もうこれは機械処理を前提とするメタデータです。逆に言うと、そうだからいろいろ活用できるというメリットがあります。そういう意味では、つくる方も、人間が手で作るというよりは、システム連携でデータを入れていくのが基本になるでしょう。そこで大きなステップに踏み込んだと私は理解しています。だから、必要な情報はシステムからうまく補完されます。ID を入れれば、もうその人の名前が補完されます。逆もあるかもしれません。

いろいろな情報を全て埋めないと駄目で、大澤さんが言ったように「全部フォーマットを埋めないと駄目です」と返されるということとはなくなり、そこはシステムが頭良く働いてきちんと埋めてくれるようになるのです。

しかも、これは大学のファカルティ・マネジメント・システムと連携して、著者情報はそこから取ってくるといったシステム連携を前提として運用されることが期待されているのではないかと思います。最後の部分はそれをおっしゃったと思うのです。だからそのことを前提に皆さんがこれを使い倒していただけるようになるとうれしいなと思っています。

●片岡 言いたいことを詳しく説明していただいて、ありがとうございました。

●逸村 筑波大学の逸村です。さらにそれに乗るような感じなのですが、まず能勢先生が言われたマイナーバージョンアップについてです。これは研究者にとっては大変重要なのです。いわゆるアーリーアダプター

は、一生懸命こういうものが重要だと思って、研究者としていろいろつくります。しかし、それが知らないうちに変わっていたら、後で大変なショックを受けるのです。やはりそこは何らかのフォローが欲しいです。

そのためには、今の武田先生のお話のとおりで、極端な話、ORCID を入れれば大抵のものは全部自動的にあちこちから集められるというぐらいにならないと、やはり研究者としてはちょっと付き合いきれません。特に若手は今、異動が激しいので、それごとにシステムが違うと「何だこれは」となると思います。老婆心でしょうか。でもやはり、そのためにぜひ頑張っていたきたいと思います。

●片岡 ありがとうございます。

●能勢 私が述べたマイナーバージョンアップについて、同じ思いを共有していただいて非常にありがたいです。バージョンが変わるとショックを受けて、どこが変わったかまた理解しなくてははいけません。二度とやりたくないですね。

もう一つ私からよろしいですか。細かい話なのですが、図 5 の APC というものがちょっとよく分からなくて、どういうものですか。データについて、paid や waived というのがあまりよく理解できませんでした。

●片岡 論文を投稿するときに、APC (Article Processing Charges) という論文処理費用を支払えばオー

公的研究助成を受けた学術
成果のオープン化促進

オープンアクセス (OA) の達成度を把握するため、OA の状態やAPCの支払い状況を記述できる要素を追加

- アクセス権
<coar:accessRight rdf:resource="http://purl.org/coar/access_right/c_abf2">
open access</coar:accessRight>
 - embargoed access: エンバゴ有
 - metadata only access: メタデータのみ
 - open access: オープンアクセス
 - restricted access: アクセス制限有
- APC
<rioxterms:apc>Paid</rioxterms:apc>
 - 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

第1回 SPARC JAPAN セミナー2017

(図 5)

ブンになるというものです。

●能勢 ああ、一般的な APC ですね。そういう情報がなぜデータのオープンに必要なってくるか、あまり理解できなくて。別に APC は APC で、書誌情報でいいのではないかと、データは関係ないのではないかと私は思ってしまったのですが、こういう情報も必要になってくるのですか。

●片岡 こちらもオープンアクセスがどの程度達成できているかを把握するにはどういう要素があればいいかを検討する中で出てきたものです。確かにメタデータにまで APC は要らないのではないかとのご意見と、メタデータに APC も含めておいて、それを把握したり開示したりしたいというご意見の両方があり、今回は一応含めておく方向となっています。

●能勢 こういうメタデータの中身の説明は、Research Data Management (RDM) のマニュアルと何か連携はしているのですか。「この項目は RDM の資料の何ページに書かれている」などです。なぜかという、「クリエイティブ・コモンズ (CC)」はここでは当然のように使っていますが、一般の研究者は、CC とは何か分からない、六つあるのも知らない、どれに該当するかも分からないという方も多いでしょうし、ISNI と言われても知らない方も多いと思います。

一般に広く研究者が使うには説明が必要です。メタデータを理解する、スキーマを理解するために、やはり何か資料が欲しいのです。APC に関しても、今私が思ったように「これは何だろう」と思いますので、そういうマニュアルや簡単な説明をつくられる予定などはございますか。

●片岡 一応、今、ガイドラインというものを作成しているのですが、それは主にリポジトリを管理する図書館員向けの内容になっています。研究者の方にこんなに関心を持って見ていただけるとあまり思っておら

ず、申し訳ありません。

●能勢 この部屋の中でだけかもしれませんが。

●片岡 研究者向けの教材との連携はまだできていませんが、確かにそういうものもあると良いかと思いましたが、今後の検討の課題とさせていただきます。

●能勢 この部屋にいる研究者は「メタデータのスキーマ」と言っても興味を持つのかもしれませんが。

●片岡 ありがとうございます。