

研究データの可視化・検索向上を目指したメタ データマネジメントの実践

新堀淳樹¹、能勢正仁^{1,2}、三好由純¹、堀智昭¹、塩川和夫¹、大平司³、
端場純子³、直江千寿子³、我喜屋累³、岡本麻衣子³、
門倉昭⁴、阿部修司⁵、吉川 顕正⁵、

堀優子⁶、芦北卓也⁶、平野かおる⁶、岡村菜々子⁶、清水敏之⁶、相良毅⁷

¹名古屋大学宇宙地球環境研究所、²名古屋市立大学データサイエンス学部、
³名古屋大学附属図書館、
⁴情報・システム研究機構データサイエンス共同利用基盤施設
⁵九州大学国際宇宙惑星環境研究センター、⁶九州大学附属図書館、⁷情報試作室

1. はじめに

1.1 本共同研究に至った背景

文部科学省 → 国立大学法人

- 「公的資金による研究データの管理・利活用に関する基本的な考え方」に基づき、文部科学省から、2021年6月に国立大学法人へ通達。
- 「各大学における研究データの管理・利活用に係るデータポリシーの策定等を含め、適切に取り組んでいただくようお願いいたします。」

研究データ基盤システム
を用いる

データポリシーの策定
(2025年までに)

新規公募分についてDMP
の必須化
(2023年度までに)

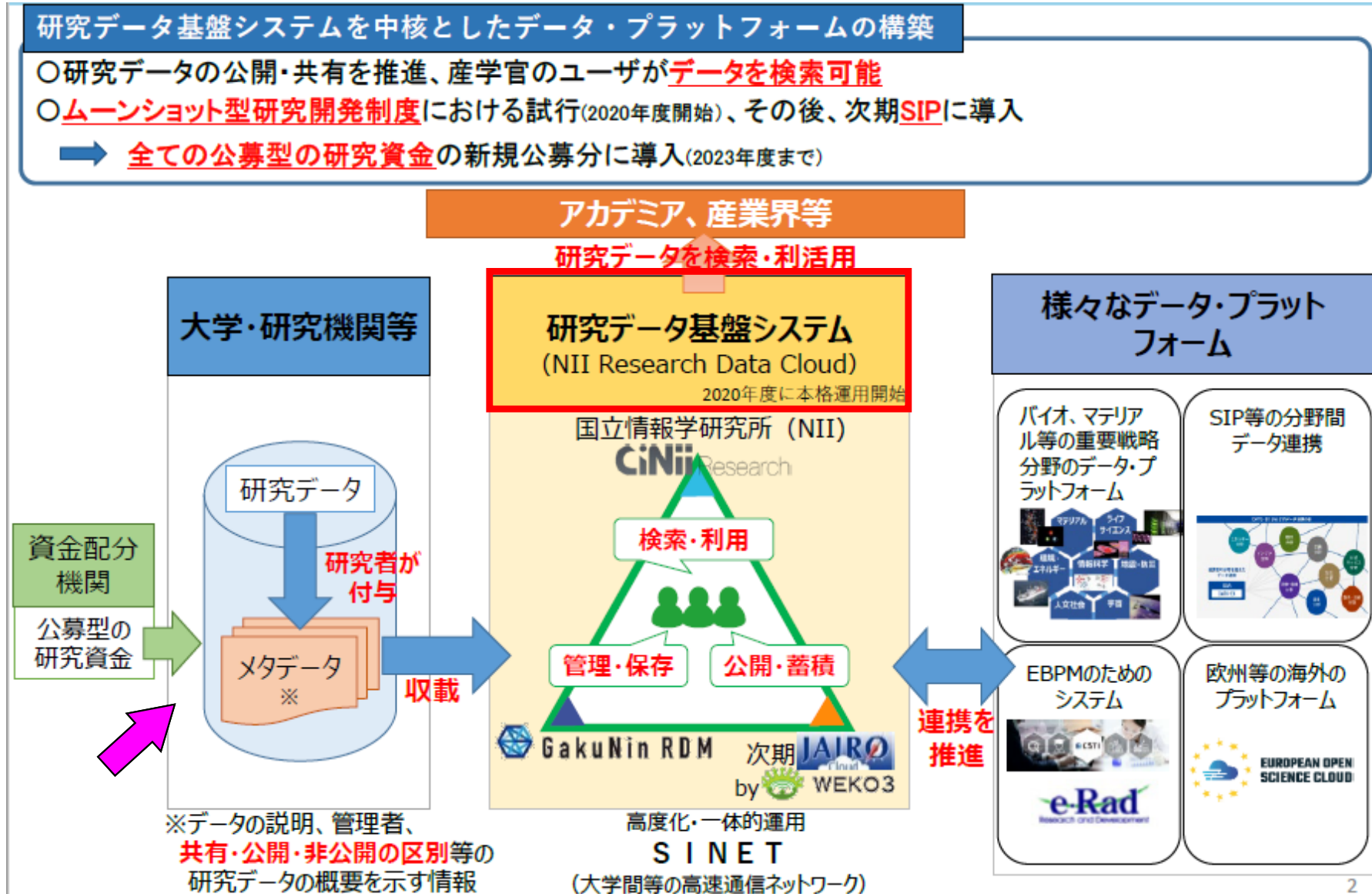
メタデータの作成と登録

<本考え方の主な内容>

- 研究データ基盤システム (NII Research Data Cloud) を中核的なプラットフォームとして位置付け、産学官における幅広い利活用を図るため、メタデータ (データを説明するための情報から構成されるデータ) を検索可能な体制を構築する。(2023年度まで)
- 研究開発を行う機関は、データポリシーを策定し、機関リポジトリへの研究データの収載を進める。(機関リポジトリを有する全ての大学・大学共同利用機関法人・国立研究開発法人においては、2025年までにデータポリシーを策定)
- 公募型の研究資金の全ての新規公募分について、メタデータを付与する仕組みを導入。(2023年度まで)
- 研究者は、所属機関のデータポリシーや公募型の研究資金における資金配分機関の基準等に基づき、管理対象データの範囲を定め、メタデータを付与し、研究データ基盤システムへ登録 (又は研究データ基盤システムで検索可能なプラットフォームへ登録) する。
- その他 (人材・支援体制の整備、取組状況の評価、他のデータ・プラットフォームとの連携等)

1. はじめに

1.2 公的資金による研究データの管理・利活用に関する基本的な考え方



2. 研究データの登録・管理に関する問題点

2.1 共通メタデータと研究分野メタデータの違い

- 政策的に期待される共通メタデータフォーマットと研究分野で利用されるメタデータフォーマットは大きく異なる。

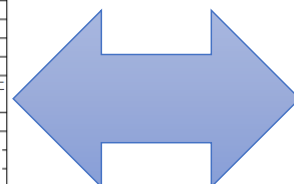
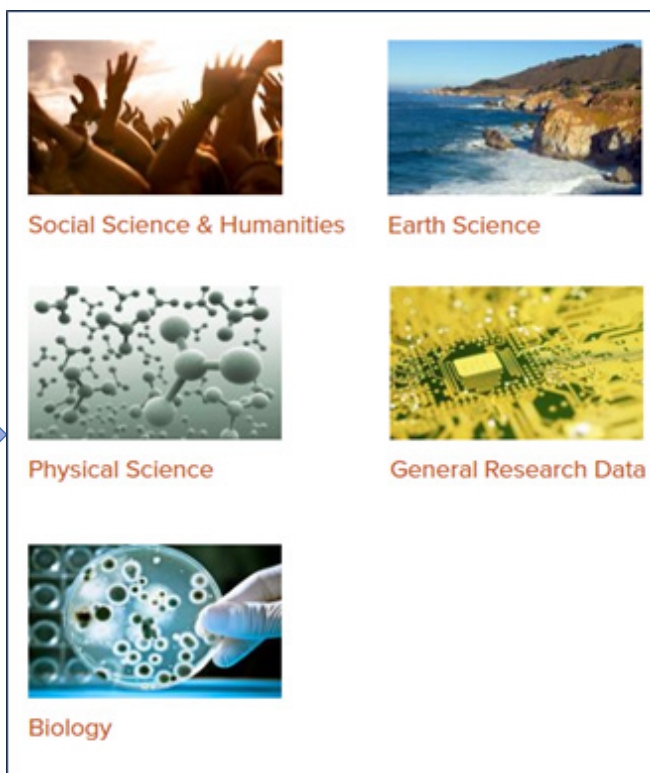
共通メタデータ

研究分野メタデータ

「公的資金による研究データの管理・利活用に関する基本的な考え方」におけるメタデータの共通項目
(2023年3月31日時点)

項目	必須/任意	備考
1 資金配分機関情報	必須※	公募型の研究資金を配分した資金配分機関（府省含む）の英語略称 公募型の研究資金以外の場合は記入不要
2 体系的番号におけるプログラム情報コード	任意	公募型の研究資金の場合は、体系的番号のうち、「機関コード」および「施策・事業の特定コード」を表すコード 公募型の研究資金以外の場合は記入不要
プログラム名	任意	競争的研究費制度の名称
3 体系的番号	必須※	公募型の研究資金の場合は、研究費ごとに付与される体系的な番号 公募型の研究資金以外の場合は記入不要
プロジェクト名	必須※	プロジェクトの研究代表者が統括する研究開発の範囲の名称（e-Rad課題名称等） 公式な名称がない場合は、研究者の所属機関のルールに従って入力
4 データNo.	必須	管理対象データを一意に特定するための番号 公募型の研究資金を配分した資金配分機関が付与 公募型の研究資金以外の場合はデータ管理機関（項目14）が付与
5 データの名称	必須	学会資料、報告資料、測定結果などの中身に分らない名称は避ける
6 掲載日・掲載更新日	必須	メタデータの掲載日・掲載更新日
7 データの説明	必須	端的かつ中身の分かる内容を記載
8 データの分野	必須※	e-Radの研究分野（主分野）。e-Radとの連携により、自動入力される予定。
9 データ種別	必須	研究データ基盤システム上では、通常は「データセット」を標準とするが、データの特性に応じて「データセット」以外の種別を選択可能
10 概略データ量	任意	1GB未満、1GB以上10GB未満、10GB以上100GB未満、100GB以上等の区画により記載
11 管理対象データの利活用・提供方針	必須	無償/有償、ライセンス情報、その他条件（引用の仕方等）等を記載
12 アクセス権	必須	公開/共有/非共有・非公開/公開期間猶予から選択
13 公開予定日	必須	公開期間猶予を選択した場合、公開予定日を記載
14 リポジトリ情報	必須	現在のリポジトリ情報、あるいはプロジェクト後のリポジトリ情報
15 リポジトリURL・DOIリンク	任意	情報があれば記載
16 データ作成者	任意	管理対象データを生み出した研究者の名前
17 データ作成者のe-Rad研究者番号	任意	管理対象データ作成者のe-Radの研究者番号
18 データ管理機関	必須	各データを管理する研究開発を行う機関の法人名
19 データ管理機関コード	任意	データ管理機関のコード
20 データ管理者	必須	データ管理組織において各管理対象データを管理する担当者の名前
21 データ管理者のe-Rad研究者番号	任意	管理者のe-Radの研究者番号
22 データ管理者のe-Rad研究者番号	任意	e-Rad研究者番号がない管理者は記入不要、ある場合は番号研究者が番号を非公開にしたい場合を除き必須
23 データ管理者の連絡先	必須	データ管理者の所属機関の住所や電話番号、メールアドレス等
24 備考	任意	

※公募型の研究資金による研究活動の場合



https://www8.cao.go.jp/cstp/common_metadata_elements.pdf

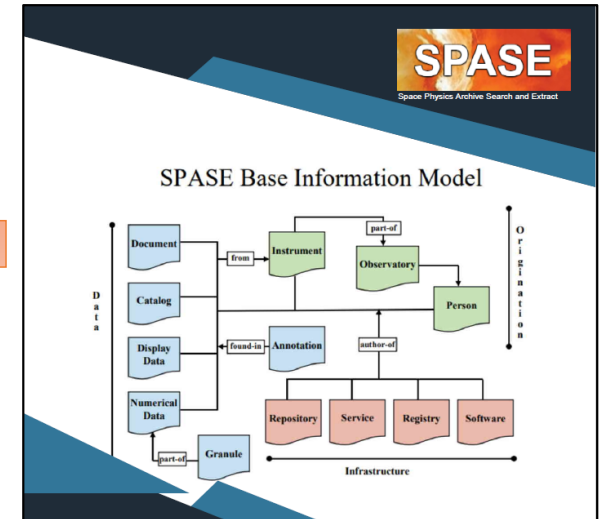
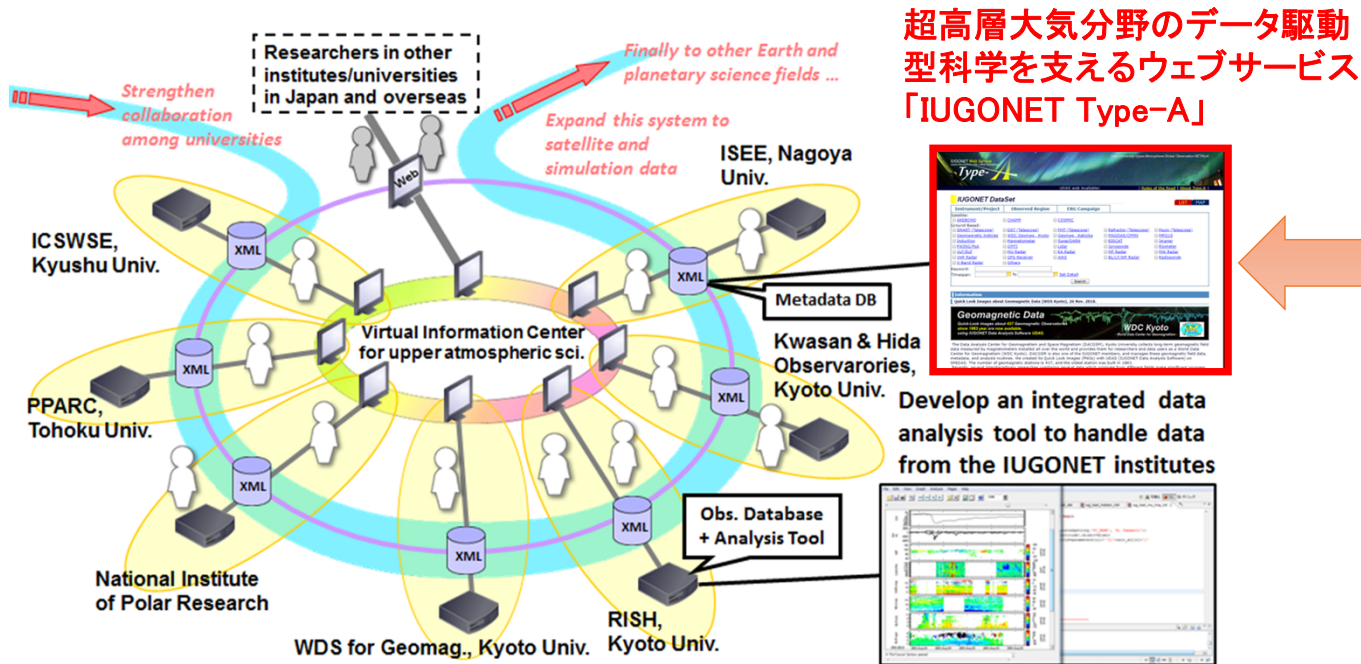
- 政策的に必要な情報
- RD基盤システムに登録するための一般的な情報

<https://www.dcc.ac.uk/guidance/standards/metadata>

- 研究遂行に必要な情報
- データを解釈するための詳細な情報

2. 研究データの登録・管理に関する問題点

2.2 宇宙地球科学分野におけるメタデータモデルと活用事例



- 宇宙地球科学分野においては、Space Physics Archive Search and Extract (SPASE)と呼ばれる国際コンソーシアムがメタデータモデルに従ったスキーマを策定している。
- SPASEメタデータスキーマは、主に科学人工衛星のデータを記述することに主眼を置いてきたが、最近では地上観測データやシミュレーションデータも記述できるように改良が進んでいる。
- この分野の「デファクトスタンダード」なメタデータモデルとして広く利用されている。
- 大学間連携プロジェクト「IUGONET」で作成されたメタデータ(データセット)が1200件以上、Type-Aに登録されており、これまでに180本以上の研究成果創出に貢献してきた。

2. 研究データの登録・管理に関する問題点

2.3 研究指向側のメタデータモデルの問題点

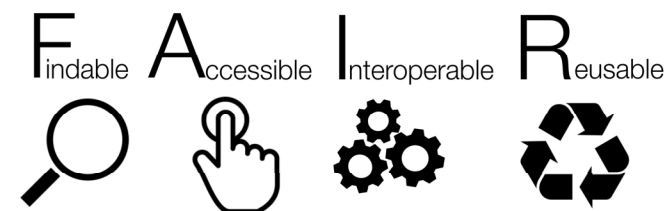
- 利用者は、様々なキーワードでメタデータ検索が可能。(例: 観測機器、物理量等)
→ SPASEメタデータモデルが、**科学研究に必要な情報**を記述できる項目を多く用意している

⇔ 利用者はある程度、以下の知識が必要

- 宇宙地球科学、IUGONET(プロジェクトの概要、メタデータ検索機能等の情報)

こうした地上観測データは

- 宇宙地球科学の研究者にとっては、**見つけられる**。
- 他分野の研究者にとっては、**見つけづらい**。
- 一般利用者にとっては、**見つけられない**。



このような状況は、FAIR原則が掲げる“Findable”からは程遠い

そこでこの状況を改善すべく、次のような取り組みを行った。

1. SPASEメタデータスキーマからもっと一般的なメタデータスキーマへの項目対応表を作成する。
2. この対応表に従って、SPASEスキーマに基づくメタデータを一般的なスキーマに基づくメタデータに変換する。
3. 一般的なスキーマに基づくメタデータを機関リポジトリに登録する。

3. 解決方法と結果

3.1 JPCOARメタデータスキーマ

- 一般的なメタデータスキーマとして、**JPCOARメタデータスキーマ**を採用
- このメタデータスキーマは、Japan Consortium for Open Access Repository (JPCOAR)によって、日本国内の機関リポジトリの運用のために開発されている。
- このスキーマは、44個の比較的一般的な項目から構成されており、そのほとんどは学術雑誌や図書を記述するために用意されている。
- 研究データについては、10-20個程度の項目(例: タイトル、作成者、寄与者、内容記述、日付など)が対応する。



3. 解決方法と結果

3.2 メタデータ項目対応表の作成

- 2022年7月から、SPASEスキーマからJPCOARスキーマへの項目対応表を作成している。
この対応表の作成には、次の3者協働が不可欠であった。
 - 宇宙地球科学分野の研究者・・・ SPASE・IUGONETに関する知識
 - 図書館専門職員・・・ JPCOAR・リポジトリに関する知識
 - 情報科学エンジニア・・・メタデータ変換プログラム作成

JPCOAR						SPASE				
No	項目名 (日本語)	要素名	属性	言語 属性 xml: lang	記 入 レ ベ ル	繰 返 回 数	SPASE 2.4.0	注釈	説明	図書館コメント
1	タイトル	dc:title		TRUE	M	1-N	<*> +<ResourceHeader> ++<ResourceName>	データのタイトル SPASEの場合はすべて英語。 国際的に流通しているの。 JPCOARのxmlを作成するとき に、<***:lang="en">を入れる ことになる。	コンテンツのタイトル。論文の場 合、論題である。コンテンツ本文 と同じ言語のタイトル情報は必ず 記入する。タイトルの言語情報は xml:langに記入する。	言語属性を追加する必要あり。 すべてenで良いか、要確認。
2	その他の タイトル	dcterms:alternative		TRUE	MA	0-N	<*> +<ResourceHeader> ++<AlternateName>	データの別名 JPCOARのxmlを作成するとき に、<***:lang="en">を入れる ことになる。	目次タイトル、奥付タイトル等の 本タイトル以外のタイトルを記入 する。	言語属性を追加する必要あり。 すべてenで良いか、要確認。

3. 解決方法と結果

3.3 メタデータ変換ツールの作成

- メタデータ項目対応表をXSLT (XSL Transformations)で実装した。
- このXSLTファイルを用いて、名古屋大学で作成しているSPASEメタデータスキーマに基づくメタデータ(277個)を、JPCOARメタデータスキーマに基づくメタデータに変換した。

SPASEからJPCOARへの変換を行うXSLTファイル

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:spase="http://www.spase-group.org/data/schema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:orcid="http://www.orcid.org/ns/orcid"
  xmlns:jpcoar="https://github.com/JPCOAR/schema/blob/master/1.0/"
  xmlns:datacite="https://schema.datacite.org/meta/kernel-4/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">

<xsl:template match="/spase:Spase">

  <jpcoar:jpcoar
    xmlns:jpcoar="https://github.com/JPCOAR/schema/blob/master/1.0/"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:dcterms="http://purl.org/dc/terms/"
    xmlns:datacite="https://schema.datacite.org/meta/kernel-4/"
    xmlns:oaire="http://namespace.openaire.eu/schema/oaire/"
    xmlns:dcndl="http://ndl.go.jp/dcndl/terms/"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="https://github.com/JPCOAR/schema/blob/master/1.0/ jpcoar_scm.xsd">
    <!-- 7. 研究データ (オープンアクセス, JaLC DOI登録あり) / Research Data (Open Access, JaLC DOI registered) -->

    <!-- title / alternative -->
    <dc:title xml:lang="en">
      <xsl:value-of select="//spase:ResourceHeader/spase:ResourceName"/>
    </dc:title>

    <xsl:if test="//spase:ResourceHeader/spase:AlternateName">
      <dcterms:alternative xml:lang="en">
        <xsl:value-of select="//spase:ResourceHeader/spase:AlternateName"/>
      </dcterms:alternative>
    </xsl:if>

    <!-- creator / contributor -->
    <xsl:apply-templates select="//spase:ResourceHeader/spase:Contact"/>

    <!-- accessRights -->
    <xsl:apply-templates select="//spase:AccessInformation/spase:AccessRights"/>

    <!-- rights -->
    <xsl:apply-templates select="//spase:ResourceHeader/spase:Acknowledgement"/>

    <!-- subject -->
    <xsl:apply-templates select="//spase:Keyword"/>

    <!-- description -->
```

3. 解決方法と結果

3.4 JPCOARメタデータの機関リポジトリへの登録

- JPCOARメタデータスキーマに基づくメタデータを、**名古屋大学の機関リポジトリ(NAGOYA Repository)**に登録した。登録件数は、284件(研究代表者が名古屋大学所属者のもの)
- 一般利用者にとって、データをより発見しやすく(Findable)になった。

登録したメタデータの一覧

名古屋大学学術機関リポジトリ
NAGOYA Repository

トップ ランキング

入力後、Enterキーを押下し検索してください 🔍 検索 ▼ 詳細検索

● 全文 ○ キーワード

インデックスツリー

- ▶ A100 文学部/人文学研究科・文
- ▶ A200 教育学部/教育発達科学研
- ▶ A300 法学部/法学研究科
- ▶ A400 経済学部/経済学研究科
- ▶ A450 国際経済政策研究センター
- ▶ A500 情報学部/情報学研究科・
- ▶ B100 理学部/理学研究科
- ▶ B200 工学部/工学研究科
- ▶ B300 農学部/生命農学研究科
- ▶ C100 医学部/医学系研究科
- ▶ D100 大学院国際開発研究科
- ▶ D200 大学院多元数理科学研究科

Index List

F300 宇宙地球環境研究所 / F300i データ / IUGONET

アイテムリスト 1 - 20 of 284 results.

📄 エクスポート 表示順: ID asc 表示数 20

📄 Keogram data derived from auroral images taken by the all-sky EMCCD imager with a RG665 filter at Kevo, Finland

Kazuo Shiokawa, Yoshizumi Miyoshi, ISEE IUGONET Metadata Management Group
発行年 2017-04-01

URL

メタデータ表示の一例

F300 宇宙地球環境研究所 / F300i データ / IUGONET

Keogram data derived from auroral images taken by the all-sky EMCCD imager with a RG665 filter at Kevo, Finland

<http://hdl.handle.net/2237/0002004390>

名前 / ファイル	ライセンス	アクション
https://ergsc.isee.nagoya-u.ac.jp/psa-gnd/bin/psa.cgi		Information

121 views

total

See details

Versions

Ver.	2023-01-30 09:20:57.251467
Ver.1	2023-01-30 09:20:57.251467

Show All versions

Share

ProjectMember
姓名 Kazuo Shiokawa
言語 en
所属機関名 Institute for Space and Earth Environmental Research, Nagoya University

https://nagoya.repo.nii.ac.jp/search?page=1&size=20&sort=controlnumber&search_type=2&q=1668993801802

3. 解決方法と結果

3.5 機関リポジトリ登録の意義、発展性

●メタデータのハーベストによる被検索可能性向上

機関リポジトリ → (NII) IRDB → データカタログ横断検索システム → Google Dataset Search

●DOI取得可能

●大学や機関の知の資産に

●NII Research Data Cloudへの情報収蔵

研究データが、非常に広い範囲において、より見つけられやすくなった



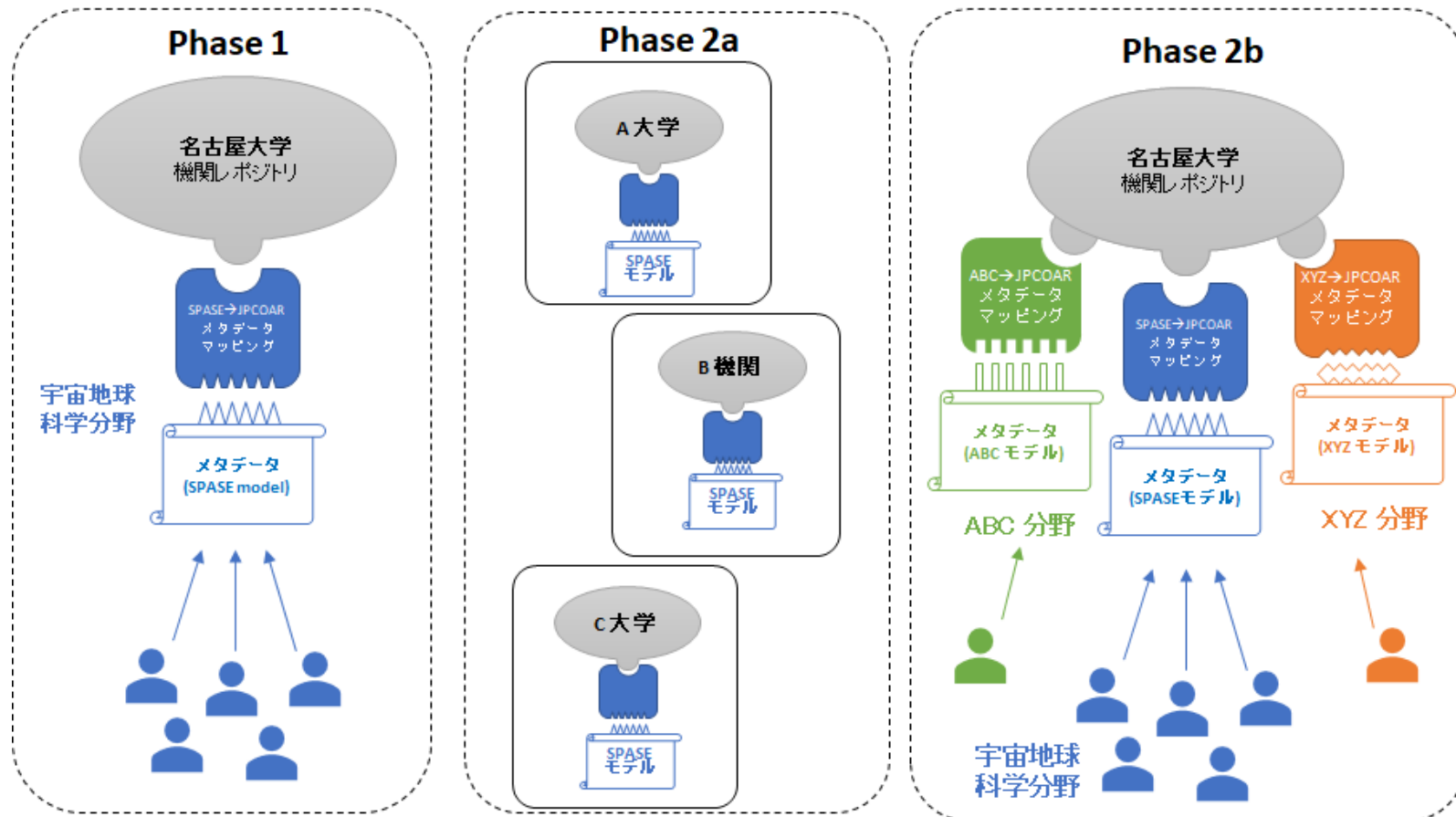
共同研究、市民科学、データそのものの評価向上、データの作成者・管理者の評価向上につながる



3. 解決方法と結果

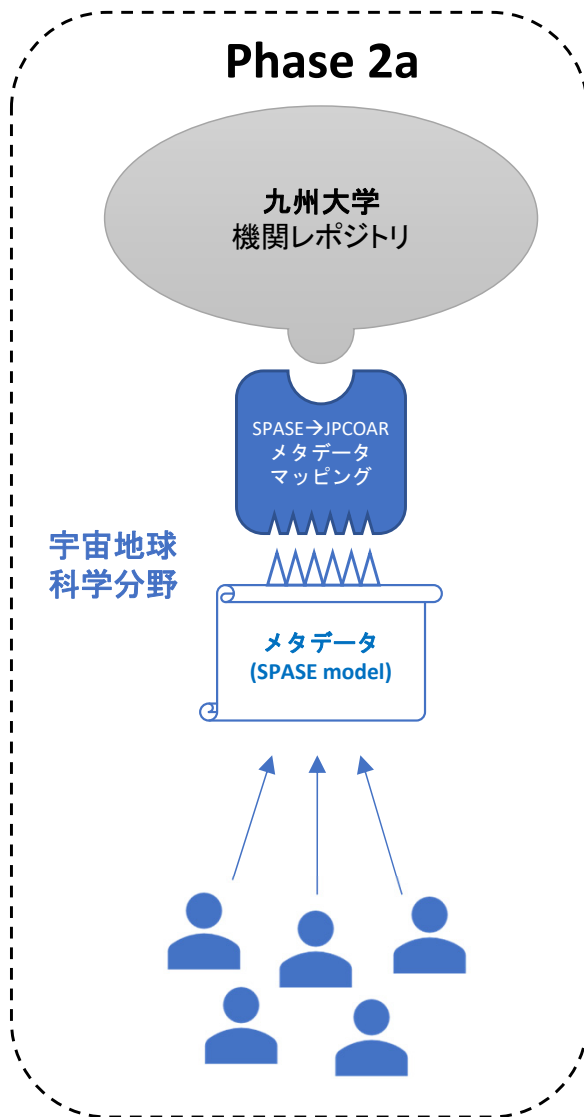
3.6 他大学・他研究機関への展開

- 他大学でSPASEメタデータスキームを利用しているところがあれば、今回の取り組みで開発した手順・成果物をそのまま適用可能。(Phase 2a)
- 他の研究分野で、デファクトスタンダードなメタデータを利用している機関があれば、今回の取り組みで得た経験・知識・教訓を活かせる。(Phase 2b)



3. 解決方法と結果

3.7 九州大学の宇宙地球科学分野への展開



- 基本的には名古屋大学で構築したメタデータマッピング手法を研究データの九州大学機関レポジトリへの登録に適用
- ※ 九州大学における宇宙地球科学分野の一部のメンバーが IUGONETメンバーでもあり、本取り組みが実施し易いと考えた
- 本取り組みを実施するにあたって、大学・研究機関によって機関レポジトリへの登録システムが異なることが判明

IUGONETメタデータ

SPASE→JPCOARへ
マッピングして登録

九州大学機関レポジトリ

- ・名古屋大の手法をふまえながら試行中
- ・GakuNinRDMのデータ解析機能を活用
(PythonやRのプログラムの実行・開発環境)

- ・JPCOARスキーマ使用
- ・レポジトリのシステムは名古屋大とは異なる (≠JAIROCloud)
- ・登録に向けたパイプラインの構築が必要

4. まとめ

- 「公的資金による研究データの管理・利活用に関する基本的な考え方」により、研究者は研究データにメタデータ(共通メタデータ)を付与し、NII研究データ基盤システムへの登録が期待されている。
- 宇宙地球科学分野では、**SPASEメタデータスキーマがデファクトスタンダード**として使用され、IUGONETプロジェクトでは、1200以上のデータセットに対するメタデータがSPASEメタデータモデルに即して作成し、IUGONETの横断検索プラットフォームでデータ検索が可能となっている。
- 専門知識を持たない一般利用者にとっては、これらのメタデータデータベースを利用して研究データを見つけることは容易ではない。

- FAIR原則が掲げる“Findable”を推進するために、**SPASEスキーマによるメタデータをJPCOARスキーマによるメタデータに変換することにした。**
- JPCOARスキーマによるメタデータは、**名古屋大学の機関リポジトリに登録し、研究データが、広い範囲でより見つけられやすくなった。**
- 機関リポジトリに登録したメタデータは、**他のメタデータデータベースによりハーベストされ、国立情報学研究所学術機関リポジトリデータベースやGoogle Dataset Searchなどでも検索可能になる。**

- このような取り組みを**九州大学の宇宙地球科学分野に適用し、実際に九州大学の機関リポジトリへの登録の作業が進んでいる。**
- 今後、他機関・他分野に本取り組みを展開することによって、分野をまたがる研究データを更に“Findable”にしていく。