Japan Open Science Summit 2023

メタデータ流通ガイドライン (研究データ編)

国立情報学研究所 学術基盤情報部 学術コンテンツ課 鹿沼智美

令和5年6月21日

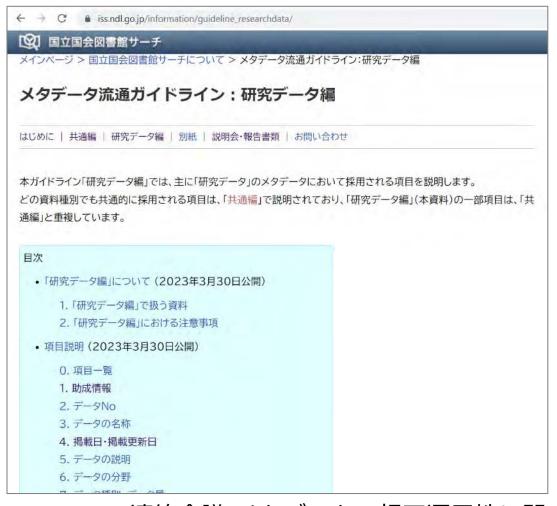




メタデータ流通ガイドライン 研究データ編 公開



2023年3月に国立国会図書館サイト上で公開



2022年3月 メタデータ流通ガイドライン(共通編): NDLはじめ各機関で集約した メタデータの流通を促進し、 発見可能性を高めるために公開

> (Japan Open Science Summit 2022 国立国会図書館スライドより)

2023年3月 メタデータ流通ガイドライン (研究データ編):を公開 主に研究データのメタデータで採用される 項目について説明

NII・JST・NDL連絡会議 メタデータの相互運用性に関する検討WG 課題10: 研究データに関する検討 にて作成



メタデータ流通ガイドライン(研究データ編)

「研究データ編」作成の背景



オープンサイエンスの潮流



国際的な動き (一部抜粋) https://rcos.nii.ac.jp/document/policy/

G7(2016年5月): G7茨城・つくば科学技術大臣会合 オープンサイエンスWGを設置

EU(2018年11月):欧州オープンサイエンスクラウド発足

OECD(2021年1月):公的資金による研究データアクセスに関する理事会勧告

UNESCO (2021年12月): オープンサイエンス勧告

米国(2022年8月): 論文と研究データの即時オープンアクセス方針を決定

現在 日本(2023年5月): G7科学技術大臣会合(仙台)

「G7は、公的資金による学術出版物及び科学データへの即時のオープンで公共的なアクセスを支援し、適切な科学的成果のより広範な共有のための学術出版における課題に対処する科学界の努力を支持する」

2025年度の新規公募分から学術論文等の即時オープン化へ



写真・内閣府「G7仙台科学技術大臣会合(概要)」より https://www8.cao.go.jp/cstp/kokusaiteki/g7_2023/ 2023.html

日本国内の動き



- 2018年 6月 国際的動向を踏まえたオープンサイエンスの推進に関する検討会「国立研究開発法人におけるデータポリシー策定のためのガイドライン」
- 2021年 3月 内閣府 「第6期科学技術・イノベーション基本計画」
- 2021年 4月 統合イノベーション推進会議「**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**」(メタデータ共通項目)
- 2021年 2月 GakuNin RDM本運用開始
- 2021年 4月 CiNii Research公開
- 2021年 9月 「ムーンショット型研究開発制度におけるメタデータ説明書 第2版」
- 2022年12月 日本学術会議 「回答・研究DXの推進 特にオープンサイエンス、データ利活用推進の 視点から - に関する審議について」
- 2022年12月 JPCOARスキーマ更新(メタデータ共通項目対応)
- 2023年 3月 「ムーンショット型研究開発制度におけるメタデータ説明書 第3版」
- 2023年 5月 G7科学技術大臣会合(仙台)
- 2025年度 (予定) 新規公募分から、学術論文等の即時オープン化の対象となる

政策的要請と期限



第6期科学技術・イノベーション基本計画 (2021年3月)



『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』

(2021年4月 統合イノベーション戦略推進会議)

「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る」
「公的資金により得られた研究データについて、産学官における幅広い利活用を図るため、2023 年度までに体系的なメタデータの付与を進め、研究データ基盤システム上でこれらのメタ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)。

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上において検索可能となるように登録する

「我が国の公的資金による研究開発で共通的に用いられるべきと考えられる**メタデータ共通項目**を別紙のとおり定める。また、研究開発を行う機関や資金配分機関において、それぞれの特性に応じてメタデータ項目を定める場合は、少なくともこの**メタデータ** 共通項目(うち必須項目)を含む必要がある。」

内閣府「研究DX」サイトで確認できます



https://www8.cao.go.jp/cstp/kenkyudx.html



公的資金による研究データの管理・利活用

- ♪ 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
 - > 本文(PDF形式:84KB) · □
- > 概要(PDF形式:935KB) ■
- シメタデータの共通項目 最新版(PDF形式:516KB) 管
- 研究データ2022(PDF形式:3289KB)
 分割版1(PDF形式:1524KB)
 2(PDF形式:1724KB)
 3(PDF形式:1964KB
 4 (PDF形式:281KB)
- 研究データの管理・利活用に関する取組状況の評価体系への導入について(PDF形式:293KB) 管
- ♪ 国立研究開発法人におけるデータポリシー策定のためのガイドライン
- > 本文(PDF形式:755KB)
- → 解説(PDF形式:574KB) 🗎
- か研究データ基盤整備と国際展開ワーキング・グループ報告書
- > 第1フェーズ報告書(令和元年10月) <u>本文(PDF形式:771KB)</u> 図表集 <u>1(PDF形式:1119KB)</u> 図表集 <u>1(PDF形式:1945KB)</u> 図表集 1(PDF形式:1945KB) 図表 1(PDF形式:1945KB) 図表 1(PDF形式:1945KB) 図表 1(PDF形式:1945KB) 图表 1(PDF形式:194
- 第2フェーズ報告書(令和3年3月) 本文(PDF形式: 941KB) 国 図表集 1(PDF形式: 858KB) 国 2(PDF形式: 1695KB) 国

先進的データマネジメント導入事例:ムーンショット型研究 開発制度

- ♪ ムーンショット型研究開発制度ページ
- るムーンショット型研究開発制度におけるメタデータ説明書 第3.0版(PDF形式:1251KB) ■
- ◆ (英語版) Metadata Manual for the Moonshot Research and Development Program (2nd Version) (PDF形式:689KB)

「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る」
「公的資金により得られた研究データについて、産学官における幅広い利活用を図るため、2023年度までに体系的なメタデータの付与を進め、研究データ基盤システム上でこれらのメタデータを検索可能な体制を構築する。(2023年度まで)上

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)。

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上において検索可能となるように登録する

「我が国の公的資金による研究開発で共通的に用いられるべきと考えられるメタデータ共通項目を別紙のとおり定める。また、研究開発を行う機関や資金配分機関において、それぞれの特性に応じてメタデータ項目を定める場合は、少なくともこのメタデータ共通項目(うち必須項目)を含む必要がある。|

「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」 内閣府「研究DX(デジタル・トランスフォーメーション」より https://www8.cao.go.jp/cstp/kenkyudx.html

メタデータ共通項目を含む

メタデータを検索可能な体制の構築へ



メタデータ共通項目とメタデータ説明書



どの研究データにも最低限付与してほしい メタデータの共通項目を策定

「我が国の公的資金による研究開発で共通的に用いられるべきと考えられるメタデータ共通項目を別紙のとおり定める。また、研究開発を行う機関や資金配分機関において、それぞれの特性に応じてメタデータ項目を定める場合は、少なくともこのメタデータ共通項目(うち必須項目)を含む必要がある。」

「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」より

	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		
	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		
Li	プログラム名	任意	競争的研究費制度の名称		
	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		
3	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		
1	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の分からない名称は避ける		
3	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		
3	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。		
)	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		
0	概略データ量	任意	1GB未満、1GB以上10GB未満、10GB以上100GB未満、100GB以上等の区分により記載		
	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		
1	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		
	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		
2	リポジトリ情報 リポジトリ URL・ DOIリンク	必須 任意	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 情報があれば記載		
2	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		
3	データ作成者 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		
	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		
	データ管理機関コード	任意	データ管理機関のコード		
	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		
4			管理者のe-Radの研究者番号		
	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		
	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		
	備考	任意	2		

項目名(日本語)	管理対象データの利活用・提供方針			
項目名(英語)	Data utilization and provision policy 必須			
必須/任意				
記述言語	日本語・英語			
備考	対象となる管理対象データを提供する場合の条件など(無償/有償等の使用条件や、また論文等で引用する際の引用の仕方等)を入力 管理対象データを一定期間後に非共有・非公開から共有へ変更する予定のある場合等、特段の事情のあるデータについては、その方針を本項目に記載する。 (例) 一定期間後に事業の実施上有益なものに対して有償又は無償で提供を開始。但しデータのクレジット標記を条件とする。なおサンプルデータを公開している。			
項目名(日本語)	11. アクセス権			
項目名(英語)	Access rights 必須(選択式)			
必須/任意				
記述言語				
備考	 管理対象データのアクセス権について下記から選択 公開 (open access): 一般に公開されるデータ 共有 (restricted access): 限定された関係者での利活用を目的として共有されるデータ 非共有・非公開 (metadata only access): 公開も共有 			

「メタデータ共通項目」

「ムーンショット型研究開発制度におけるメタデータ説明書 第3.0版』より

JPCOARスキーマでの対応



メタデータ共通項目にJPCOARスキーマ2.0で対応して、メタデータ流通に乗せます

	共通項目	必須/任意	JPCOARスキーマ
1	資金配分機関情報	必須	jpcoar:fundingReference/jpcoar:funderldentifier
2	体系的番号におけるプログラム情報コード	任意	jpcoar:fundingReference/jpcoar:fundingStreamldentifier
	プログラム名	任意	jpcoar:fundingReference/jpcoar:fundingStream
2	—————————————————————————————————————	必須	jpcoar:fundingReference/jpcoar:awardNumber
3	プロジェクト名	必須	jpcoar:fundingReference/jpcoar:awardTitle
4	データNo.	必須	jpcoar:relatedIdentifier
5	データの名称	必須	dc:title
6	掲載日・掲載更新日	必須	datacite:date
7	データの説明	必須	datacite:description
8	データの分野	必須	jpcoar:subject
9	データ種別	必須	dc:type
10	概略データ量	任意	jpcoar:extent
	管理対象データの利活用・提供方針	必須	dc:rights
11	アクセス権	必須	dcterms:accessRights
	公開予定日	必須	datacite:date

JPCOARスキーマを使用 しているシステム JAIRO Cloud 他 A大学 B大学 C機関 IR 国内外に 公開・流通 IR IR 大学・機関の リポジトリで 研究データメタデータ

の入力・流通を行う

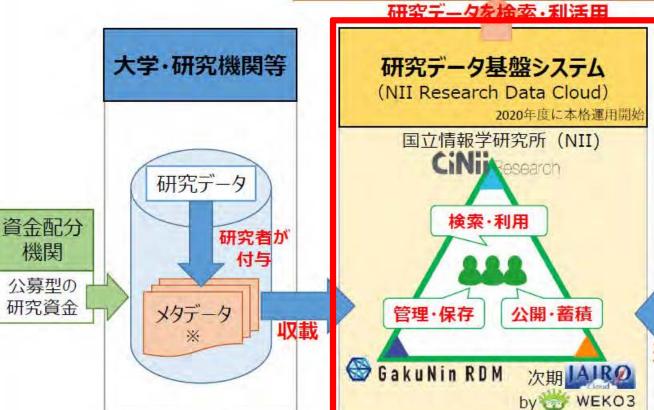
(項番12以下の表記は本スライド上では省略)

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ついて

アカデミア、産業界等

研究データ基盤システムを中核としたデータ・プラットフォームの構築

- ○研究データの公開·共有を推進、産学官のユーザがデータを検索可能
- 〇ムーンショット型研究開発制度における試行(2020年度開始)、その後、次期SIPに導入
 - ➡ 全ての公募型の研究資金の新規公募分に導入(2023年度まで)



※データの説明、管理者、

共有・公開・非公開の区別等の

研究データの概要を示す情報

様々なデータ・プラット フォーム

バイオ、マテリア ル等の重要戦略 分野のデータ・プ ラットフォーム



SIP等の分野間 データ連携



EBPMのための システム



欧州等の海外の プラットフォーム



高度化·一体的連用

SINET

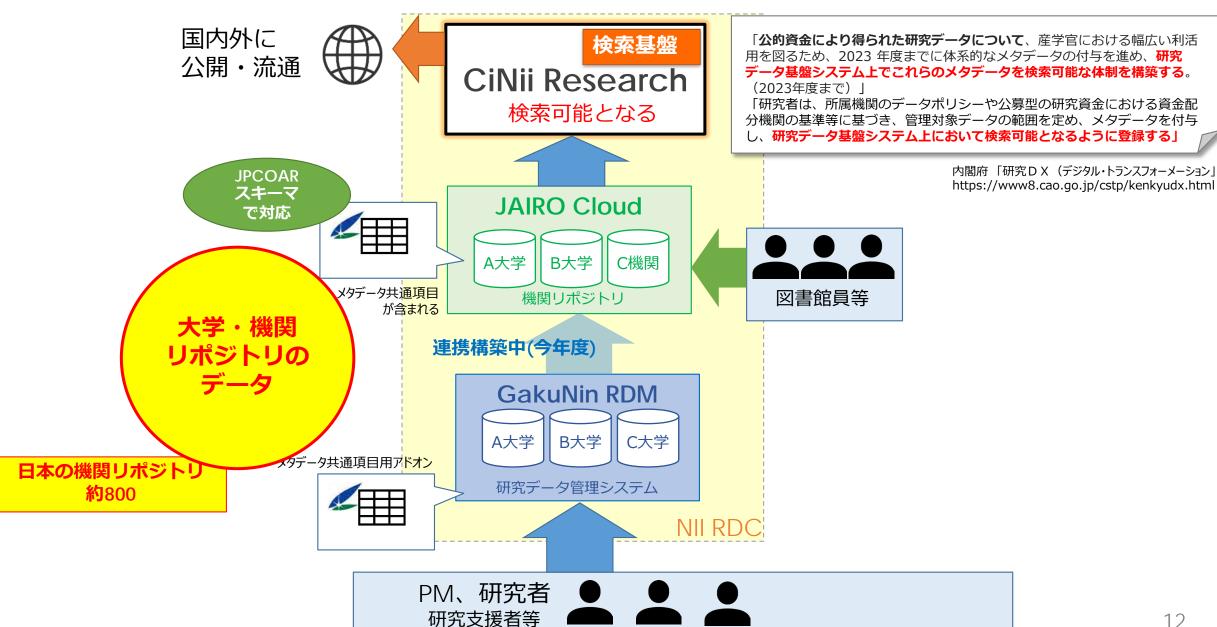
(大学間等の高速通信ネットワーク)

データの管理・利活用 に関する基本的な考え 方について」に追記 https://www8.cao.g o.jp/cstp/datapolicy _outline.pdf (内閣府サイト)

「公的資金による研究

NII研究データ基盤 (NII RDC)による基本のルート







「公的資金により得られた研究データについて、産学官における幅広い利活用を図るため、2023年度までに体系的なメタデータの付与を進め、研究データ基盤システム上でこれらのメタデータを検索可能な体制を構築する。(2023年度まで)」

JaLCやNDLサーチ経由のメタデータ

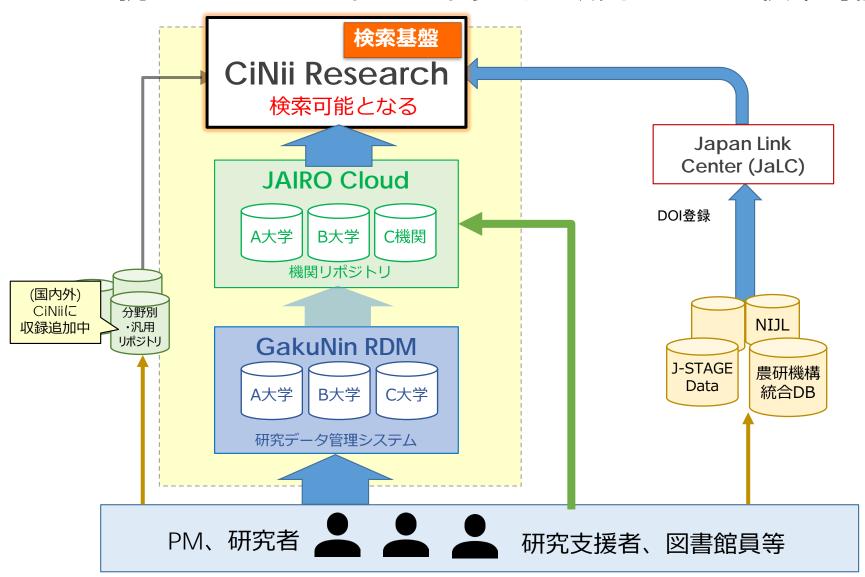
メタデータを検索可能な体制の拡張へ



メタデータ検索可能な体制の拡張 他のシステムとの連携



他システムとも連携することにより、より多くの研究データが検索可能になります



DC_JaLCでの対応 ※JaLC (Japan Link Center) とは: DOIを付与する機関

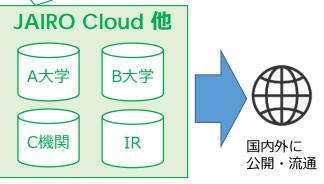


メタデータ共通項目にJPCOARスキーマ+DC_JaLCで対応して、メタデータ流通に

乗せます

	共通項目	必須/任 意	JPCOARスキーマ	DC_JaLC
1	資金配分機関情報	必須	jpcoar:fundingReference/jpcoar:funderldent ifier	fund_list/fund/funder_identifier
2	体系的番号におけるプログラム情 報コード		ipcoar:fundingReference/jpcoar:fundingStre amldentifier	
	プログラム名		jpcoar:fundingReference/jpcoar:fundingStre am	
3	体系的番号		jpcoar:fundingReference/jpcoar:awardNum ber	fund_list/fund/award_number_group/awar d_number
	プロジェクト名	必須	jpcoar:fundingReference/jpcoar:awardTitle	
4	データNo.	必須	jpcoar:relatedIdentifier	alternate_identifier_list/alternate_identifier
5	データの名称	必須	dc:title	title_list/titles/title
6	掲載日・掲載更新日	必須	datacite:date	publication_date date_list/date@type="Updated"
7	データの説明	必須	datacite:description	description_list/description
8	データの分野	必須	jpcoar:subject	subject_list/subject
9	データ種別	必須	dc:type	resource_type
10	概略データ量	任意	jpcoar:extent	size_list/size
	管理対象データの利活用・提供方 針	必須	dc:rights	rights_list/rights
11	アクセス権	必須	dcterms:accessRights	
	公開予定日	必須	datacite:date	date_list/date@type="Available"

JPCOARスキーマを使用 しているシステム



大学・機関のリポジトリ で研究データメタデータ の入力・流通



(項番12以下の表記は本スライド上では省略)

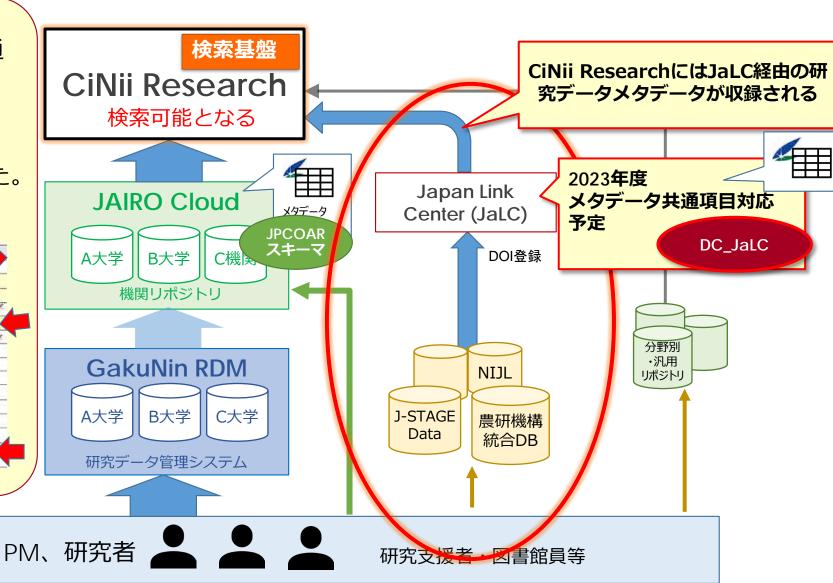
JaLCとの連携

※JaLC (Japan Link Center) とは: DOIを付与する機関



JaLCは2023年度にメタデータ共通 項目対応を行う予定です 「メタデータ流通ガイドライン 研 究データ編」を作成していく中で、 対応すべき内容が明確になりました。

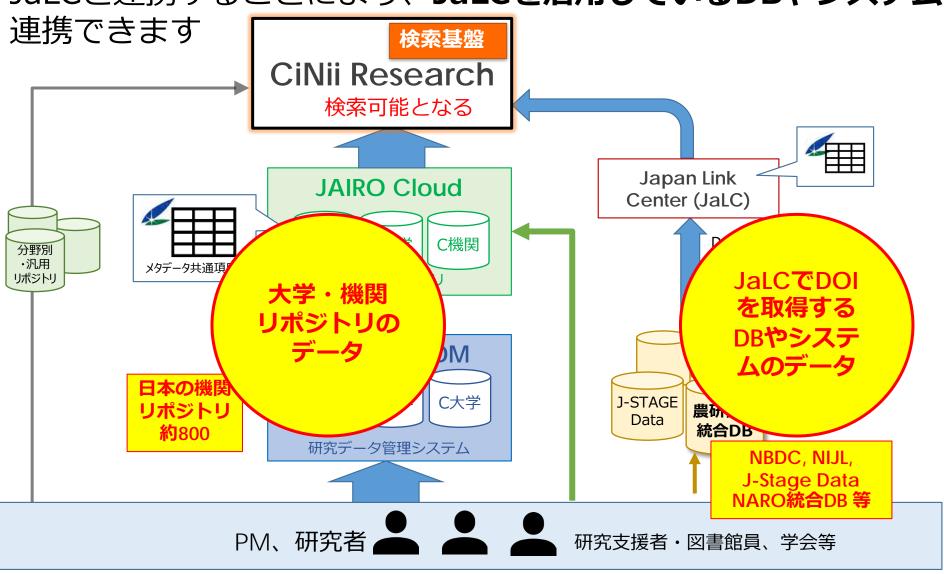




JaLC経由の連携



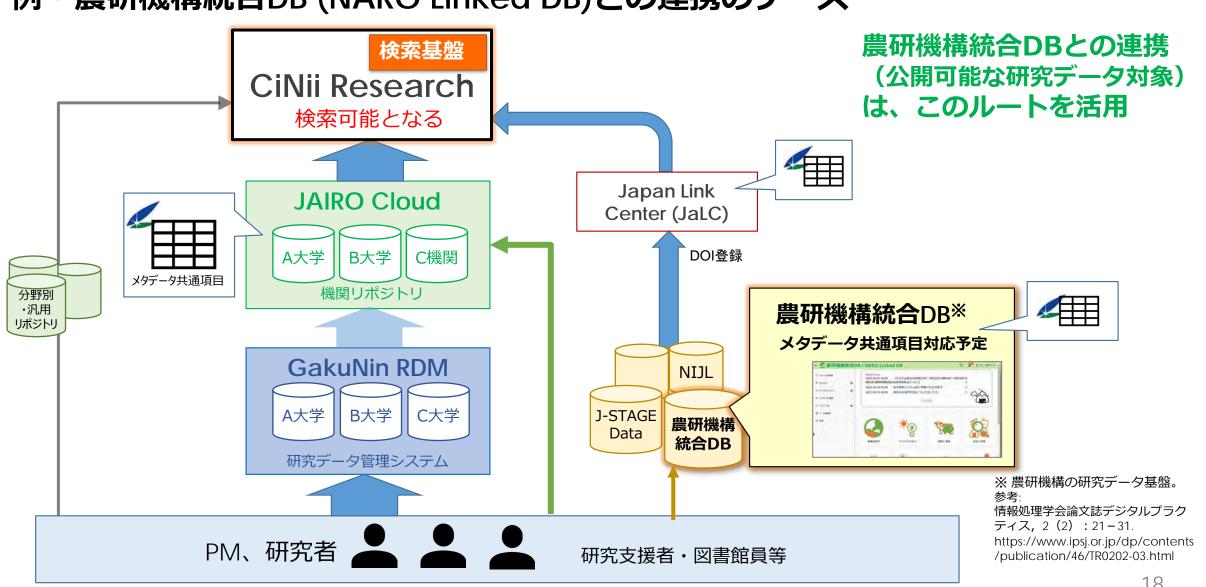
JaLCと連携することにより、JaLCを活用しているDBやシステムの研究データと



JaLC経由の連携



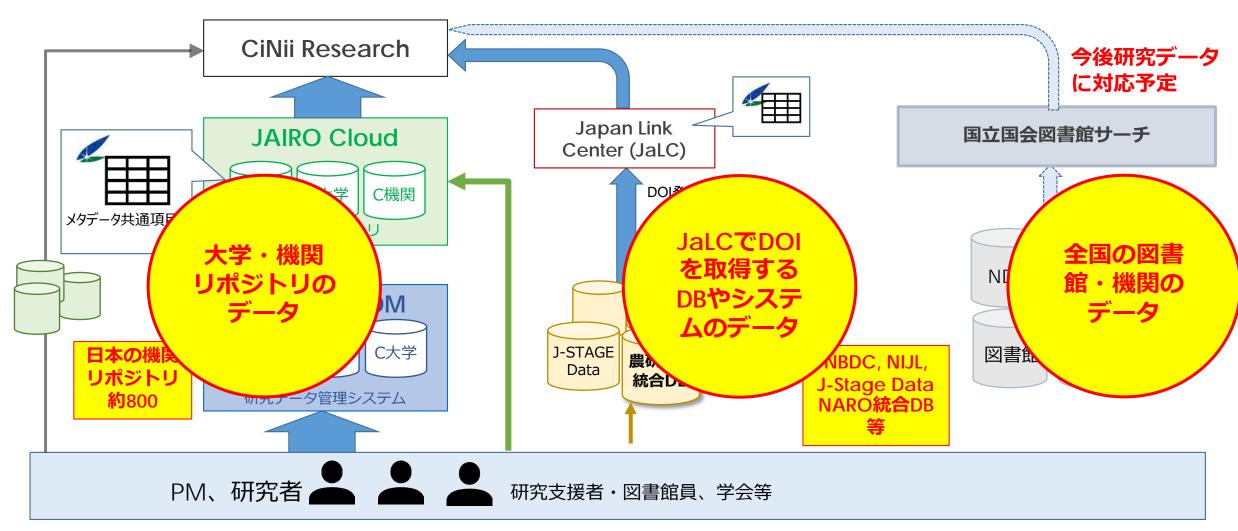
例・農研機構統合DB (NARO Linked DB)との連携のケース



国立国会図書館サーチ経由の連携



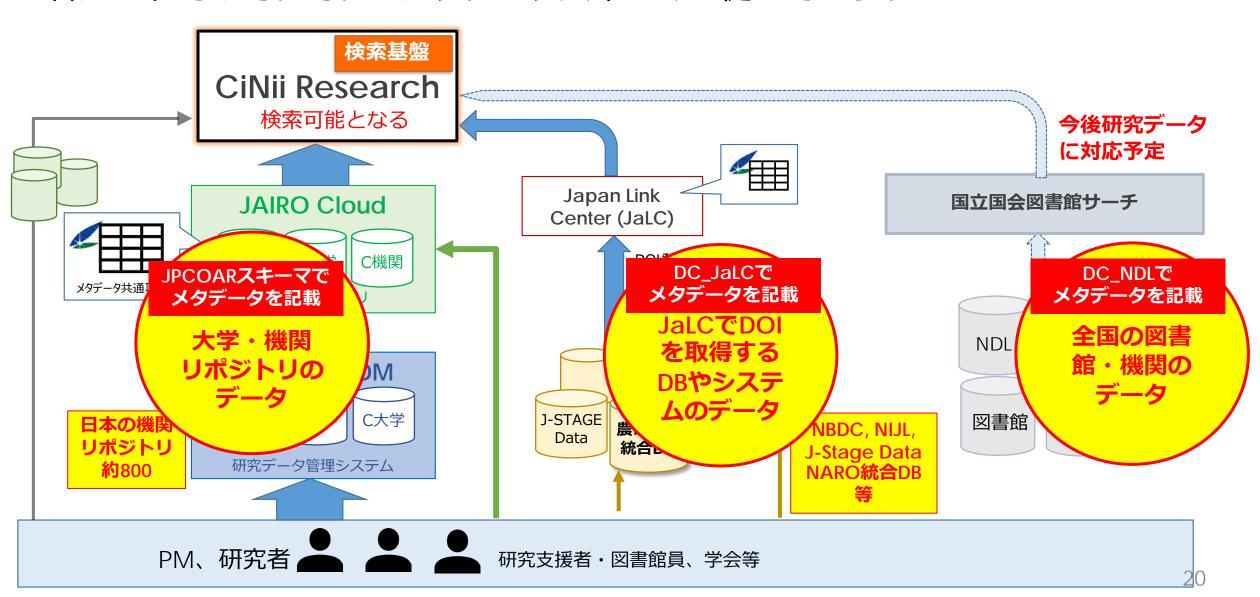
国立国会図書館サーチとも連携する予定



国立国会図書館サーチ経由の連携



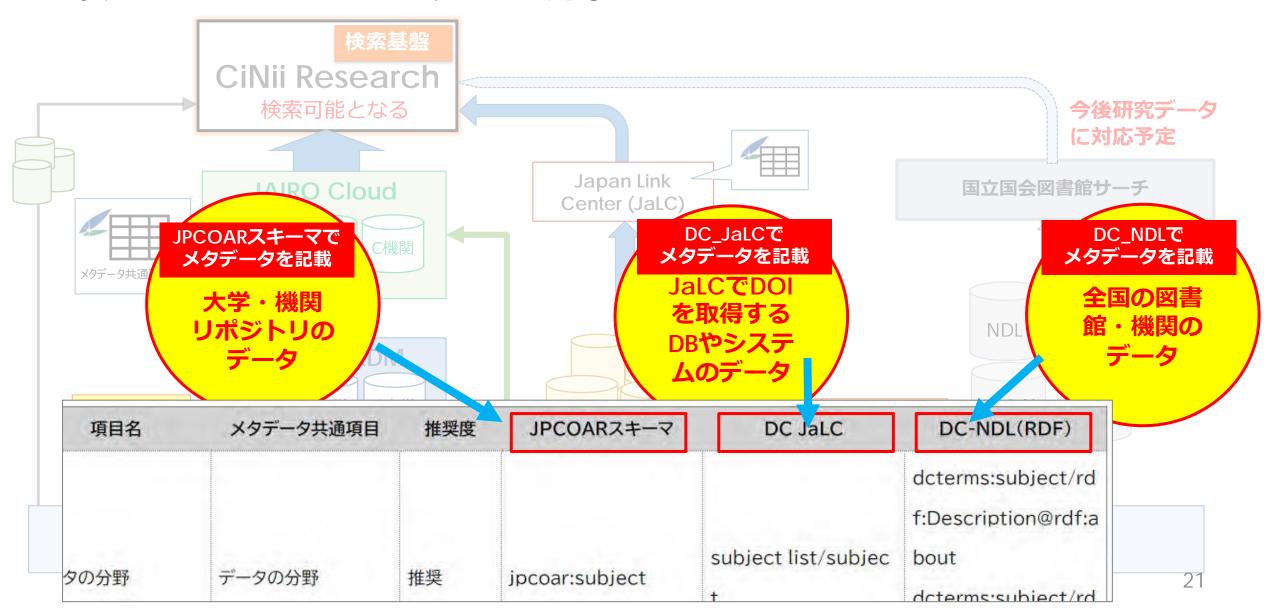
各ルートではそれぞれのメタデータスキーマに従っています



メタデータ流通ガイドライン 研究データ編



主要な3つのメタデータスキーマを対象としています





中身をご紹介します

メタデータ流通ガイドライン (研究データ編) の内容



研究データ編の内容



研究データ編について

- 1. 「研究データ編」で扱う資料
- 2. 「研究データ編」における注意事項

項目説明

- 0. 項目一覧
- 1. 助成情報 ~ 13. メタデータのライセンス

研究データのメタデータ流通経路

- 1. 対象範囲
- 2. 対象範囲のメタデータスキーマ

検討経緯、今後の発展性



https://iss.ndl.go.jp/information/guideline_researchdata/

メタデータ項目の説明

NII

- ・「メタデータ共通項目」とされた各項目について JPCOARスキーマ、DC_JaLC、DC-NDLでのマッピング表
- ・基本事項、各スキーマでの記載方法(データ例・適用状況)など

1. 助成情報

項番 項目名		メタデータ共通項目	推奨度	JPCOARスキーマ	DC JaLC	DC-NDL(RDF)
1-0	〔助成情報〕		推奨	jpcoar:fundingRe ference	fund list/fund	
1-1	資金配分機関情報	資金配分機関情報	推奨	jpcoar:funderIde ntifier jpcoar:funderNam e	fund list/fund/fu nder identifier fund list/fund/fu nder name	dcterms:descripti on
1-2	プログラム情報	体系的番号における プログラム情報コー ド	任意	jpcoar:fundingStr eamIdentifier	=	
		プログラム名	任意	jpcoar:fundingStr		
1-3	課題情報(課題番号·課題名)	体系的番号	推奨	jpcoar:awardNum ber	fund list/fund/aw ard number grou p/award number	
		プロジェクト名	推奨	jpcoar:awardTitle	_	

基本事項

研究データ作成にあたり、公募型の研究資金の提供を受けた場合は、助成情報を記入します。公募型の研究資金以外の場合は記入不要です。

流通のポイント

• 「助成情報」には「資金配分機関情報」「プログラム情報」「課題情報」が含まれ、研究資金の成果集計等に用いられます。

基本事項

。研究課題番号と、研究課題名を記載します。

メタデータ共通項目:「3.体系的番号」

- 。 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番号。公募型の研究資金以外の場合は記入不要
- 。「メタデータ共通項目」上は、公募型の研究資金による研究活動の場合必須項目となっています。

メタデータ共通項目: 「3.プロジェクト名」

- 。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の課題名称があればそれを入力します。
- 。「メタデータ共通項目」上は、公募型の研究資金による研究活動の場合必須項目となっています。

JPCOARスキーマ

- データ例:
 - 研究課題番号
 - 研究課題名
- 。 適用状況:
 - 体系的番号を記入(e-Rad課題番号とは異なるので注意) 「体系的番号」の参照先は科学技術・学術政策研究所(NISTEP)のウェブサイトを参照のこと

JaLCメタデータ

- 。 データ例:
 - ▶ [award_number]
- 。 適用状況:
 - 課題名については来年度の改定での対応を検討中

DC-NDL(RDF)

- データ例:
 - 注記
- 。 適用状況:

メタデータ項目の説明 具体例



例えば、研究データの「権利情報」について、各スキーマではそれぞれ以下のように表すことがわかります



研究データのメタデータ流通経路



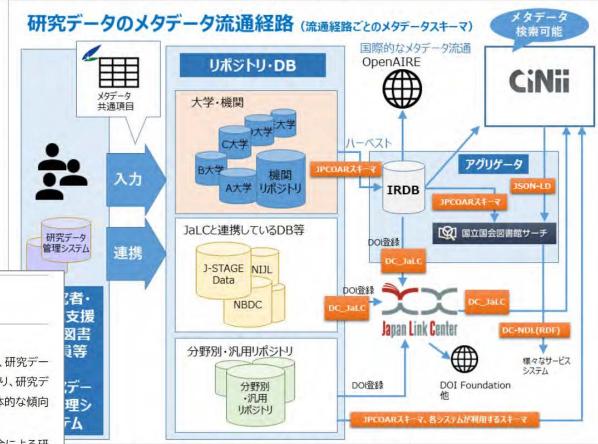
2. 対象範囲のメタデータスキーマ

本ガイドラインは、公的資金による研究開発によって生み出された研究データを対象とした「メー 参照)への準拠を基本としつつ、学術機関リポジトリデータベース(IRDB)と、Japan Link C NDLサーチを経由するメタデータ流通にスコープを絞り作成しています。

システム	メタデータスキーマ
学術機関リポジトリデータベース(IRDB)	JPCOARスキーマ
Japan Link Center (JaLC)	DC_JaLC
国立国会図書館(NDL)サーチ	DC-NDL(RDF)

- 「メタデータ共通項目」については、研究データ編について < 2. 研究データ編における注意い。
- メタデータスキーマについての全般的な解説、および、「JPCOARスキーマ」と「DC-NDL() 編<別紙4:ガイドラインの対象スキーマ>を参照してください。

本ガイドラインが対象としている流通経路では、経路ごとに異なるメタデータスキーマ(以下図中ではJPCOARスキーマ、DC_JaLC、DC NDL、各システムが利用するスキーマ)を用いています。



ータのメタデータ流通経路(流通経路ごとのメタデータスキーマ)

検討経緯、今後の発展性

研究データは、その学問領域ごとに様々な流通経路が存在します。いくつかの領域では既に流通体制が確立し、研究データの共有や公開が当然となっています。一方、データ共有・公開にまだ馴染みがない分野もあります。近年になり、研究データの利活用及び研究公正の観点、政策的なニーズも要因として、データの保管から共有、そして公開が、全体的な傾向として進められつつあります。

本ガイドラインでは、このような現状を反映し、第6期科学技術・イノベーション基本計画を背景とした「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」および「メタデータ共通項目」「ムーンショット型研究開発制度における

ソクニーク説明事(※の近日を中か) 問事オフソクニークの次名級吹を主しています



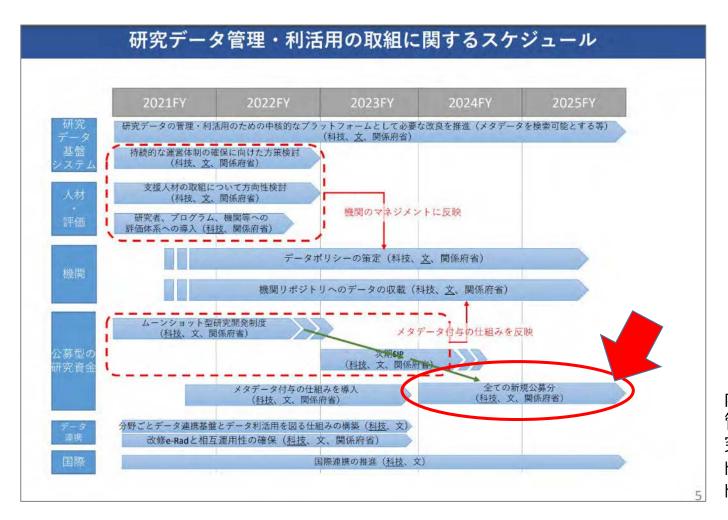
今後の課題



今後の連携や課題



2024年度から、全ての新規公募分にメタデータ付与の仕組みを導入することになります 既存のスキーマ・システムでの対応と、新たな相互連携を進めていく必要があります



内閣府「公的資金による研究データの管理・利活用に関する進捗と事例 〜研究データ2022〜」より https://www8.cao.go.jp/cstp/researc hdata2022.pdf

今後に向けて



まず「メタデータ共通項目」にのみ準じた最低限の内容を対象として公開しました。 本ガイドラインの擦り合わせの過程で各スキーマやシステムでの対応も進んでいます。 これらは追って「ガイドライン」に反映し、当ガイドラインを更新する予定です。

また、現在の当ガイドラインの記載項目は助成情報の比率が高い状況ですが、 今後はより研究に役立つ研究データメタデータの流通の為に、

- ・研究分野に特有の情報の流通を可能にする
- ・識別子の国際性・普遍性の担保など様々な側面の拡充・更新が必要と思われます。

「メタデータ流通ガイドライン 研究データ編」は今後も拡張・更新を続けてゆきます。 今後ともご協力をよろしくお願いします。