

2024年11月22日（金） デジタルアーカイブ×メタデータ勉強会 #4

② デジタルアーカイブ対応の 事例紹介



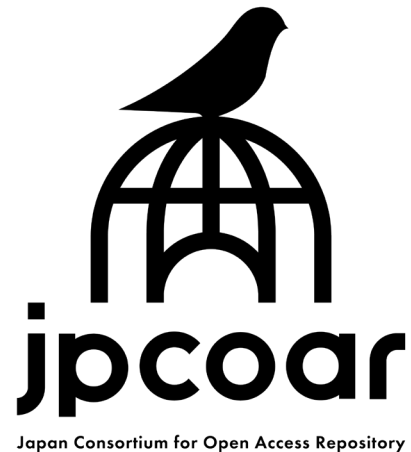
JPCOARスキーマ と 国書データベース の 未来



JPCOARコンテンツ流通促進作業部会メタデータチーム
堀野和子(国文学研究資料館)

本日のテーマ

- 1、国書データベースとは？
- 2、国書データベースのメタデータ
- 3、JPCOARスキーマとの出会い
- 4、JPCOARスキーマと国書DBの未来



1、国書データベースとは？

1、国書データベースとは？



大学共同利用機関法人 人間文化研究機構 国文学研究資料館
が公開しているデータベースです

「**国公立全ての大学の共同利用の研究所**」として、
個々の大学では整備・運営が困難な**最先端の大型装置や大量の学術データ、貴重な資料等を、全国の研究者に提供すること**を通じて大学の枠を越えた共同研究を推進し、**関連する研究分野の研究水準の向上を図る**ことを目的とする我が国独自の研究機関です。

わが国の学術研究体制は、多様な研究活動を行う大学と、それらを横断し、重点的に発展させる大学共同利用機関とを「**車の両輪**」として構築されており、お互いに切磋琢磨しあうことを通じて、我が国全体の研究水準の向上に貢献してきました。



1、国書データベースとは？



大学共同利用機関法人 人間文化研究機構 国文学研究資料館
で公開しているデータベースです

国文学研究資料館のめざすもの

日本文学の基盤的な総合研究機関

- ・ 国内各地の日本文学とその関連資料を大規模に集積
- ・ 日本文学をはじめとするさまざまな分野の研究者の利用に供する
- ・ 先進的な共同研究を推進する

創設以来50年にわたって培ってきた日本の古典籍に関する資料研究の蓄積を活かし、国内外の研究機関・研究者と連携し、日本の古典籍を豊かな知的資源として活用する、分野を横断した研究の創出に取り組めます。



東京都立川市

1、国書データベースとは？

国文研の歴史と国書データベース



- ・ 1972年 国文学研究資料館創設 **調査・収集事業開始**
- ・ 1977年 閲覧サービス開始
- ・ 1987年 「マイクロ資料目録データベース」「館蔵和古書目録データベース」のオンライン検索サービス開始
- ・ 2006年 「日本古典籍総合目録データベース」※1を公開
- ・ 2007年 デジタル収集画像を公開
- ・ 2011年 収集マイクロフィルムをデジタル化して公開
- ・ 2014年 大規模学術フロンティア促進事業
「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」開始
- ・ 2017年 「新日本古典籍総合データベース」※2公開
- ・ 2023年 ※1と※2を統合し「**国書データベース**」を公開
「**コーニツキー版 欧州所在日本古書総合目録**」統合
- ・ 2024年 大規模学術フロンティア促進事業
「データ駆動による課題解決型人文学の創成」開始
「**近代書誌・近代画像データベース**」統合
館蔵資料の二次利用条件をパブリックドメインへ変更

マイクロDB
+
館蔵DB
+
他機関目録DB



『十年の歩み』国文学研究資料館

1、国書データベースとは？

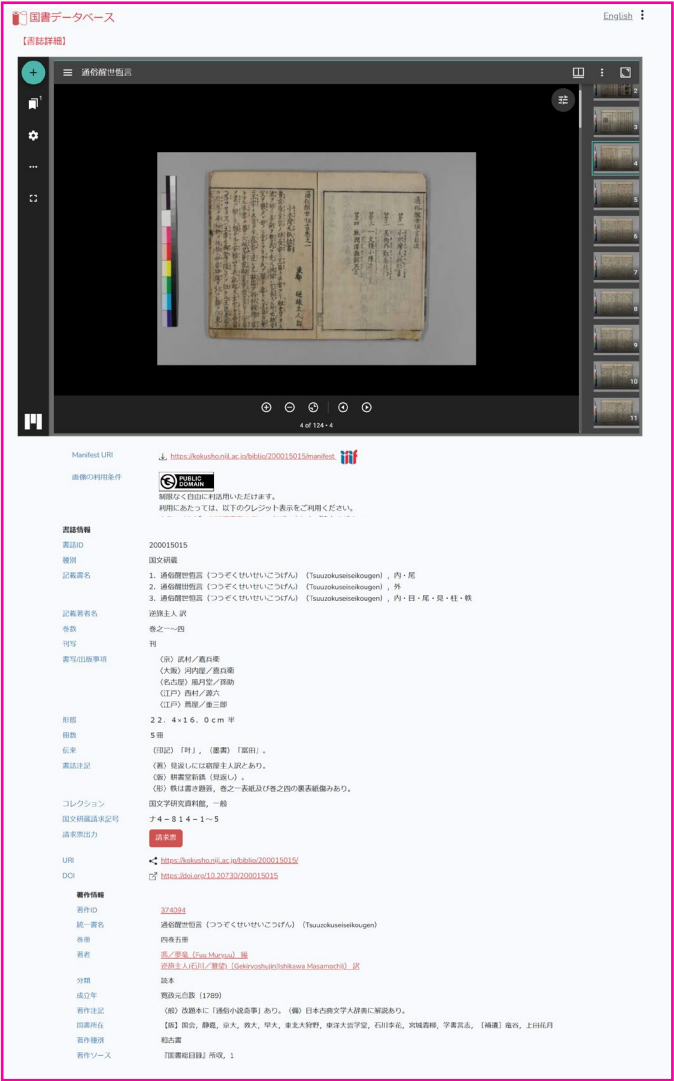
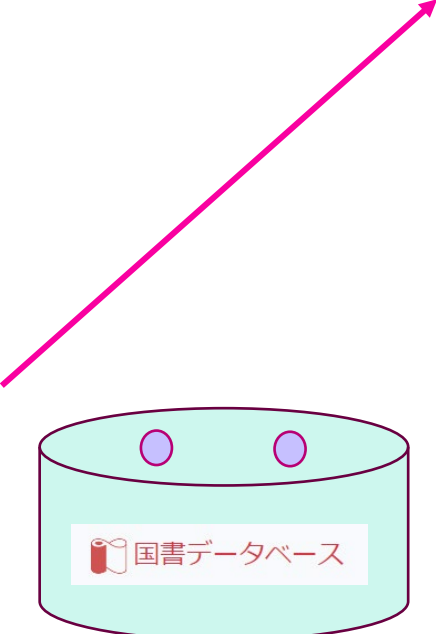
集積した成果を公開するデータベース

書誌詳細

トップページ



検索結果



2、国書データベースのメタデータ

2、国書データベースのメタデータ

書誌データを詳しく見てみると...

書誌情報

書誌ID 200015015

種別 国文研蔵

記載書名 1. 通俗醒世恆言（つうぞくせいせいこうげん）（Tsuuzokuseiseikougen）, 内・尾
2. 通俗醒世恆言（つうぞくせいせいこうげん）（Tsuuzokuseiseikougen）, 外
3. 通俗醒世恆言（つうぞくせいせいこうげん）（Tsuuzokuseiseikougen）, 内・目・尾・見・柱・帙

記載著者名 逆旅主人 訳

巻数 巻之一～四

刊写 刊

②刊本か写本か

書写/出版事項 〈京〉武村／嘉兵衛
〈大阪〉河内屋／喜兵衛
〈名古屋〉風月堂／孫助
〈江戸〉西村／源六
〈江戸〉蔦屋／重三郎

③複数の出版者

形態 22. 4×16. 0 cm 半

冊数 5冊

伝来 (印記)「叶」, (墨書)「富田」。

④伝来

書誌注記 〈著〉見返しには宿屋主人訳とあり。
〈版〉耕書堂新鐫（見返し）。
〈形〉帙は書き題簽, 巻之一表紙及び巻之四の裏表紙傷みあり。

⑤多様な注記

コレクション 国文学研究資料館, 一般

⑥コレクション



国文研独自のスキーマ

古典籍特有の項目が
丁寧に表現されている

その他の項目として

- ・部編
- ・和暦
- ・残欠

…など

2、国書データベースのメタデータ

国書データベースの書誌データのルーツを探れ！

- ・ 1972年 国文学研究資料館創設
- ・ 1977年 閲覧サービス開始
- ・ 1977年 マイクロ目録（冊子体）

- ・ この段階で決定した項目がほぼそのまま使われてきていた！
- ・ さらに館蔵資料、他機関目録のDB作成時は少しずつ異なるスキーマを採用

国文学研究資料館は、人文科学系の資料センターとしては初めてコンピュータによる情報処理を行う、ということが、創設に際して企画されていた。

この目録はそのデータベースの一部となるべきものであり、機械可読の形式で作成されることになった。

（ただしこの段階ではコンピュータは導入されておらず、外部のものを使った）

『十年のあゆみ』国文学研究資料館

たしかに！



- ・ 冊子体出版時のみに使用する項目が残っている...
- ・ 1つのフィールドに押し込まれている情報...
- ・ 対象資料（マイクロ、館蔵資料、他機関目録）で項目が異なる...

- ・ 1987年 「マイクロ資料目録データベース」「館蔵和古書目録データベース」のオンライン検索サービス開始
- ・ 2006年 「日本古典籍総合目録データベース」を公開

...

マイクロDB
+
館蔵DB
+
他機関目録DB

2、国書データベースのメタデータ

書誌データの作成方法

※他機関目録をのぞく

- 1、原本を見て作成（館蔵資料）
- 2、画像を見て作成（他機関資料）



2014年～2023年

「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」

※10年間で30万点の画像公開＝書誌作成

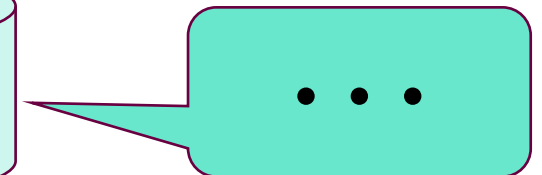
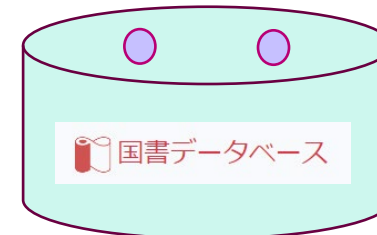


NEW

- 3、NACSIS-CATのデータがある資料は、
国書DB独自スキーマに変換して最低限の不足箇所を補充

+

1・2についても書誌の簡素化
→最低限必要な項目とは…？

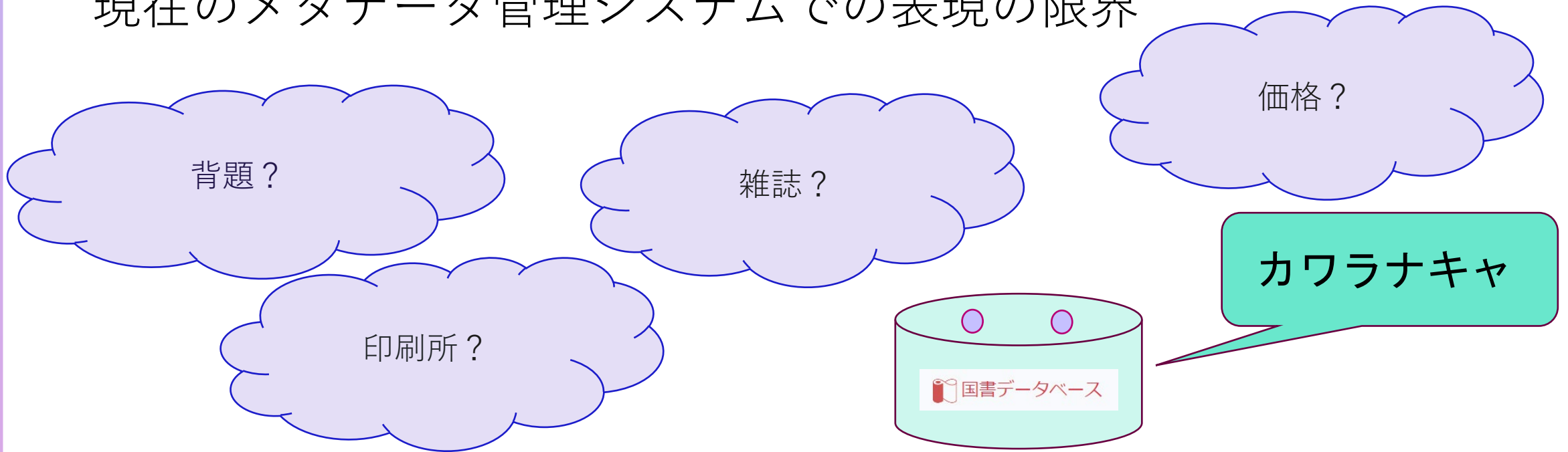


2、国書データベースのメタデータ

国文研のデータベース統合

- ・ 2023年 「コーニツキー版 欧州所在日本古書総合目録」 統合
- ・ 2024年 「近代書誌・近代画像データベース」 統合

→明治期資料と古典籍の違いをどう整理するか
現在のメタデータ管理システムでの表現の限界



3、JPCOARスキーマとの出会い

3、JPCOARスキーマとの出会い

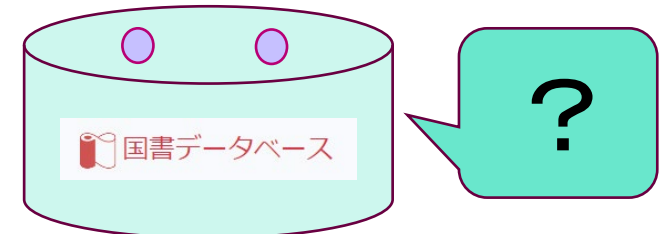
JPCOARスキーマ2.0策定にあたり、
JPCOARコンテンツ流通促進作業部会メタデータチームへ参加。

→流通に重点を置いた際の必要な情報とはなにか…

JPCOARスキーマVersion 2.0策定方針

基本方針

1. JPCOARスキーマが目指してきた標準的なスキーマへの準拠の方針を継承し、**独自項目の採用は最小限に留める。**
2. デジタルアーカイブの利活用可能なメタデータ流通に必要とされる**最小限のメタデータ項目のみ新規に採用する。**
3. デジタルアーカイブのメタデータ流通を鑑み、
アイテムの集合体としてのコレクション単位のメタデータ定義を導入する。
4. 研究データの管理・利活用を促進するため、統合イノベーション戦略推進会議により提示された、
国内の公的資金による研究開発において共通的に用いられる「メタデータの共通項目」に対応した拡張を行う。
5. 廃止となる項目や語彙については、影響を慎重に確認し、可能な限り下位互換性の確保に努める。



3、JPCOARスキーマとの出会い

「メタデータ流通ガイドライン：古典籍編」（2023年6月19日公開）への協力

デジタルアーカイブ、蔵書目録等のデータベースのメタデータ設計の関係者が、各機関の特性に応じた適切なメタデータの作成、流通経路の選択ができるように作成されました。2022年3月にドラフト版として＜共通編＞を、2023年3月30日に＜研究データ編＞を公開に続き、＜古典籍編＞を公開。

和古書（概ね江戸期以前に書写・刊行された日本語資料）のメタデータにおいて採用される項目について説明しているもの。

検討経緯、今後の発展性

（2023年6月19日公開）

古典籍に関するデータは、他の一般的な資料に比してその利用のニーズに応じ、マイクロ化・デジタル化など、様々な媒体への変換が行われてきました。また、近年では、翻刻等における最新技術により利活用の可能性も広がっています。

メタデータは、これら様々に取り巻く環境を、最大限必要な利用者に届けるための材料であり、何を必要な情報として格納するかまた流通させるかは、常にアップデートしていく必要もあります。こういった情勢をふまえ、今回「古典籍」に特化したガイドラインを作成しました。

本ガイドラインは、今後も次のような予定で内容を充実させていきます。

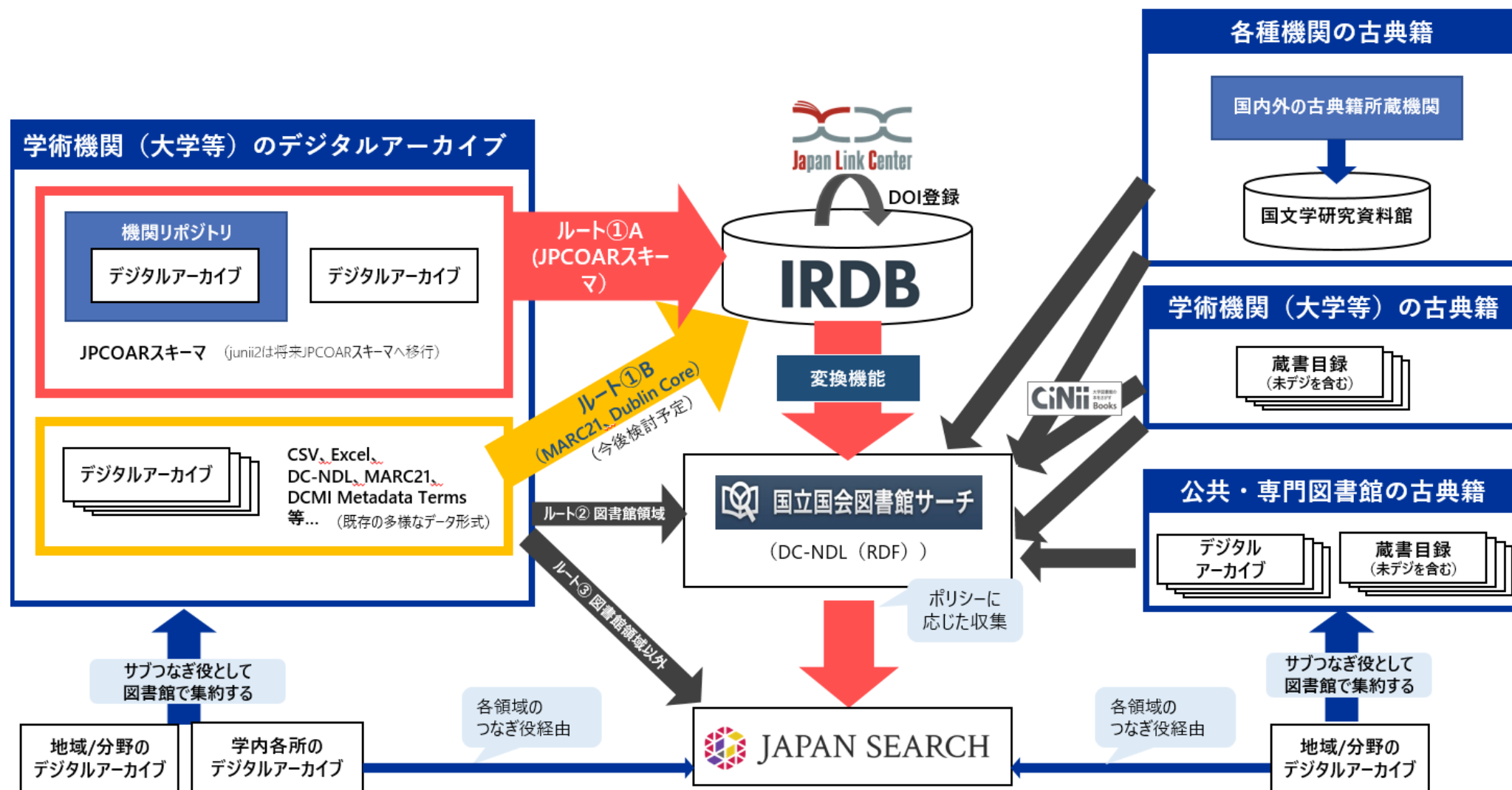
- 漢籍等、対象範囲の拡張
今回対象外とした「漢籍」「洋古書」等についても、情報の充実を目指します。
また、古典籍のうち、地図や絵図といった個別の資料に応じた記載も今後検討します。
- 古典籍に特有な項目に関する案内の充実
必要に応じて案内の追加を検討します。
また、各スキーマでは、必要な項目の新設などを継続的に検討します。

なお、本ガイドラインの編纂にあたり、[国文学研究資料館](#)にご協力いただきました。

事務職員だけでなく、
教員も参加しました！

3、JPCOARスキームとの出会い

国書
データベース



図：古典籍のメタデータ流通経路

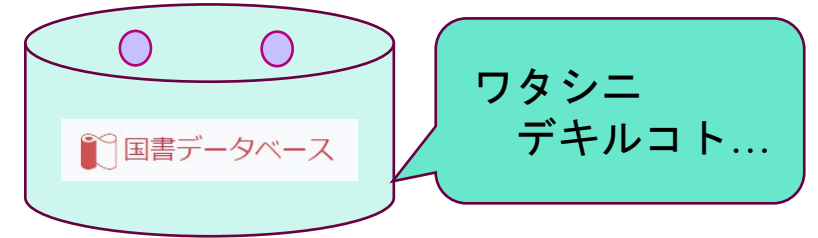
4、JPCOARスキーマと国書DBの未来

4、JPCOARスキーマと国書DBの未来



古典籍画像の発見可能性を高めるために当館でできること

→ 効果的にメタデータを流通させる！



STEP 1 国書データベースのメタデータ管理システムを再構築する

- ❌ 古すぎる : 未使用項目・用途不明項目が多数あり
- ❌ 複雑すぎる : システムは教員作成後、様々な業者による機能追加
- ❌ 敷居が高すぎる : システムを熟知した職員のみしか対応できない

→ データを分析し、メタデータ設計を改めて行うことで柔軟な対応が可能に！

(校正テキスト、解題等、多様な付加データへの対応...など)

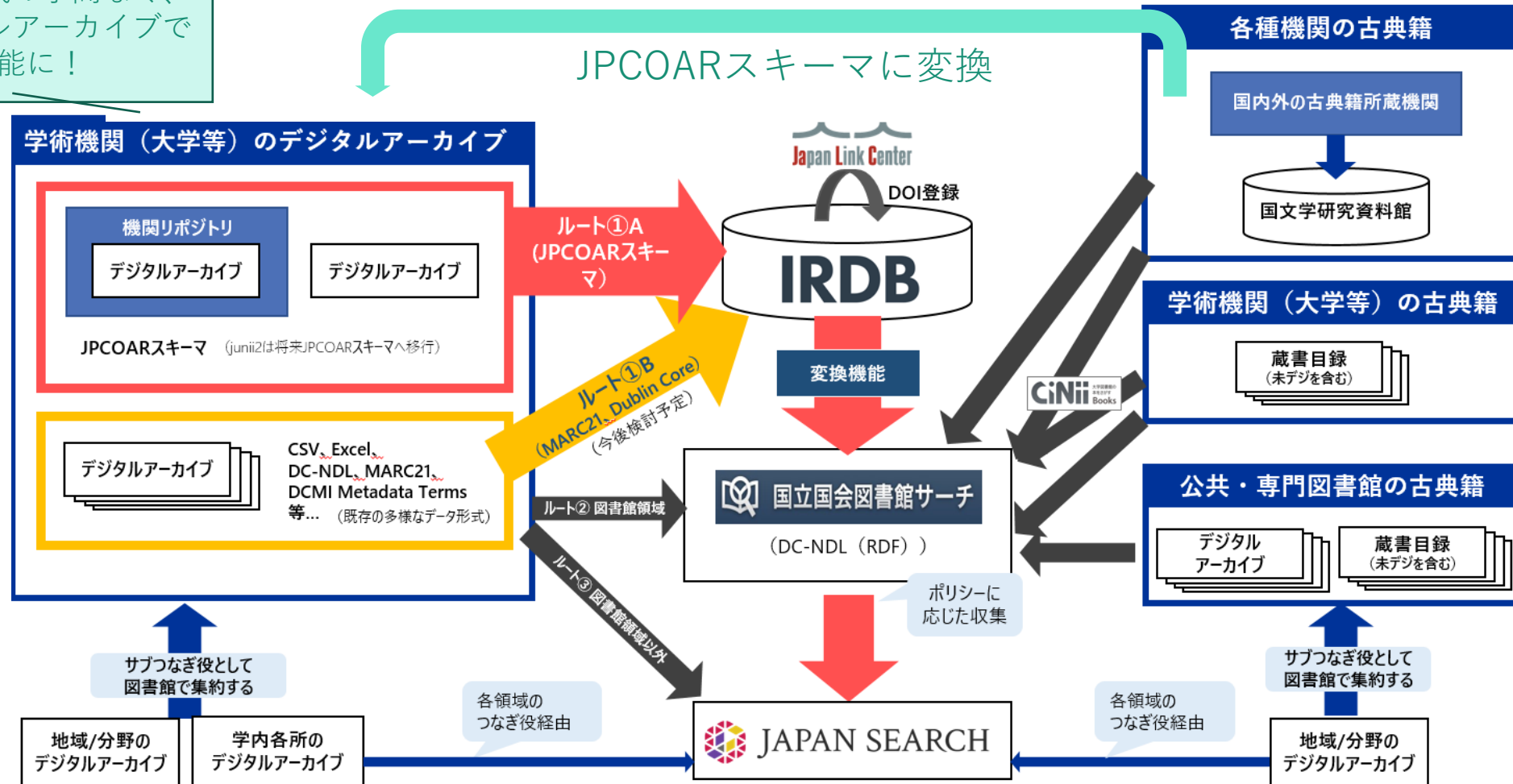
→ 汎用性が高く、自由なスキーマ設計に対応した汎用プラットフォームへ！

(シー・エム・エス社のe-Catsメタデータ管理オプションで実現)

4、JPCOARスキーマと国書DBの未来

STEP2 国書データベース独自スキーマから各種スキーマへの変換

メタデータ作成の手間なく、
自館のデジタルアーカイブで
公開可能に！



4、JPCOARスキーマと国書DBの未来

国書スキーマとJPCOARスキーママッピング例				
※国書データベースにはほかに「著作」と「著者」のスキーマもあるが、ここでは記述内容の多い書誌を取り上げた。				
国書データベース「使い方」ページでの書誌事項説明		JPCOARスキーマver2.0		
国書DB書誌事項	書誌事項説明	要素名	属性値	子要素
書誌ID	書誌レコード番号	jpcoar:identifier	identifierType="URI"	
種別	書誌種別	(マッピング対象外)		
標目書名	本としての資料の代表書名（マイクロ資料の場合はターゲット書名）、書名の種別	dc:title	xml:lang="ja ja-Hira ja-Latn"	
記載書名	資料に記載されている書名、書名の種別	dc:terms:alternative	xml:lang="ja ja-Hira ja-Latn"	
記載著者名	資料に記載されている著者名 他等 役割/伝 (部編等の注記)	jpcoar:creator	creatorType	jpcoar:creatorName
巻数	資料の巻数	datacite:description	descriptionType="Other", label="巻数"	
刊写	刊：印刷による、写：書写による、混：刊写入り混じり	datacite:description	descriptionType="Other", label="刊写"	
書写／出版事項	出版者、出版年、書写者、書写年（部編等の注記）	jpcoar:publisher		jpcoar:publisherName jpcoar:publisherDescription dcndl:location
形態	丁数、大きさ、紙型などの形態的事項	dc:terms:extent		
冊数	冊数			
残欠	完本でない場合の残存欠落に関する表示	datacite:description	descriptionType="Other", label="残存"	
書誌構造	原書、合綴、合写・合刻等の表示、原書番号	datacite:description	descriptionType="Other", label="書誌構造"	
親書誌	原書、合綴、合写・合刻等の全体の書名	jpcoar:relation	relationType="isPartOf"	
子書誌	原書、合綴、合写・合刻等の細目の書名	jpcoar:relation	relationType="hasPart"	
書誌注記	注記種別*3ごとの注記	datacite:description	descriptionType="Other", label="系統注記 著者注記 出版注記..."	
コレクション	所蔵者名、コレクション名、所蔵者請求記号	jpcoar:holdingAgent		jpcoar:holdingAgentName
目録分類・データソース	所蔵者目録での分類、データソースID、データソース内連番	jpcoar:relation	relationType="inSeries"	
和古書請求記号	国文研蔵の場合の原資料の請求記号	datacite:description	descriptionType="Other", label="目録分類"	
マイクロ請求記号	国文研所蔵マイクロフィルムの場合のフィルム請求記号（下にコマ数、サービス区分*4）	(マッピング対象外)		
紙焼写真請求記号	国文研所蔵紙焼写真本請求記号（下にコマ数、サービス区分*4）	(マッピング対象外)		
デジタル請求記号	デジタル画像請求記号（下にコマ数、サービス区分*4）	(マッピング対象外)		
請求票出力	請求票出力ボタン（国文研で資料請求をする場合）	(マッピング対象外)		
URL	当該ページのURL	書誌IDと同じ		
DOI	当該ページのデジタルオブジェクト識別子 ※詳しくはこちら（ダウンロードの必要あり）	jpcoar:identifier	identifierType="DOI"	
外部リンク	CiniiBooks（書名検索）、人文学オープンデータ共同利用センター（データセットのダウンロ	jpcoar:file		jpcoar:URI

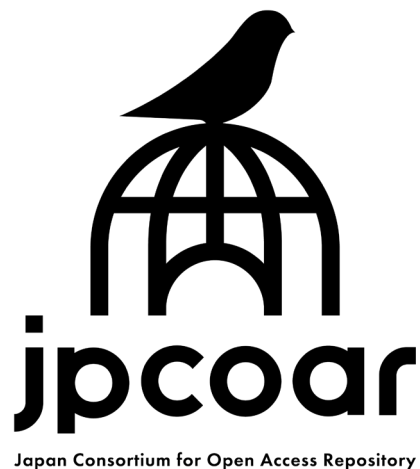
課題

- 資料ごとのメタデータが一律的でなく多様性に富んでいる古典籍資料をどう記述するか
（[datacite:description](#)かつdescriptionType="Other" が多くなる？）

メタデータ流通に必要な最小限のメタデータ ⇔ 多様性に富んでいる古典籍資料



どうもありがとうございました



国書データベース