

「メタデータ流通ガイドライン」の公開 ：メタデータ連携の現場から

メタデータ流通ガイドライン

<https://iss.ndl.go.jp/information/guideline>

情報組織化研究グループ月例研究会

2022年12月月例研究会 2022.12.17

国立国会図書館 電子情報部 電子情報サービス課

「メタデータの相互運用性に関する検討ワーキンググループ」 (メタデータWG)

- NDL、NII、JSTの三機関連絡会議のもとで立ち上げられる。
- WG活動の成果として「**メタデータ流通ガイドライン**」を公開。
- 「国立国会図書館サーチ（NDLサーチ）」は、図書館領域のメタデータを集約し提供する役割を担う。
メタデータWGには、NDLサーチ担当として参加。

NDLサーチが担うメタデータ連携の**現場**から、
「メタデータ流通ガイドライン」の**役割**をご説明します。

1. メタデータWGの活動
2. 「メタデータ流通ガイドライン」の作成
3. 「メタデータ流通ガイドライン」の特長
4. 公開後の動き
5. なぜ今、ガイドラインなのか
6. 「メタデータ流通ガイドライン」の使いどころ
7. 提供サービスと流通データの充実に向けて

1. メタデータWGの活動



- 三機関の業務連携のための恒常的な会議体。
- この活動の一環として、JSTより、メタデータの利活用促進に向けた、相互運用性改善を目的としたWGの設置について提案があった。
- あわせて、NIIより、デジタルアーカイブのメタデータ流通検討など、個別検討課題の提案。



メタデータWGの発足

メタデータWGの検討範囲

一 各種スキーマ改訂時に本WGでレビューを行うことで、メタデータを利用する側の観点を入れる。

二 異なるスキーマ間で項目のマッピングを検討。

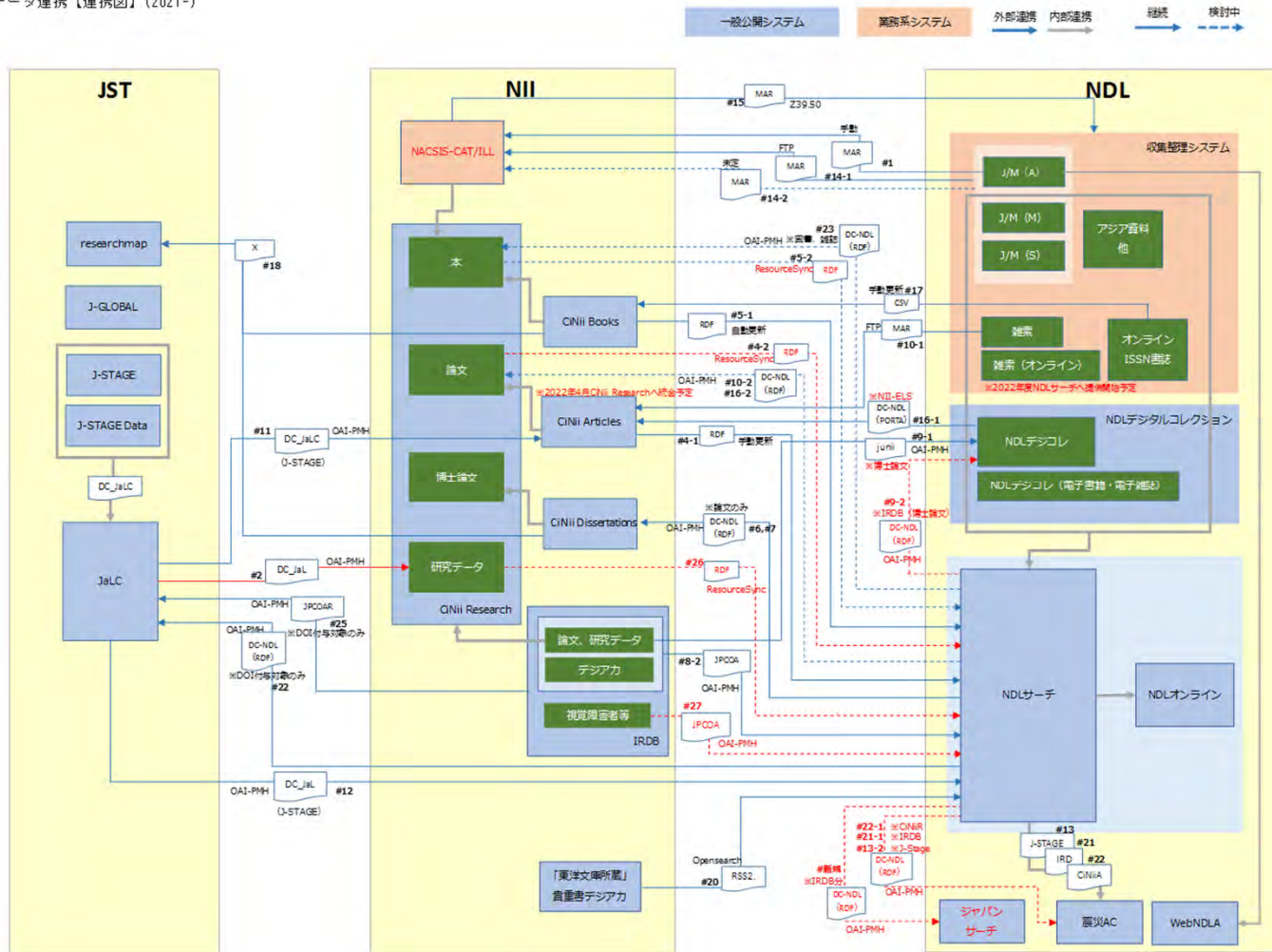
三 メタデータ交換の方法を検討し、相互運用性を改善するために求められることを見いだす。

四 データクレンジングに活用するために、メタデータの記述ルールを共有する

五 メタデータの利用条件を検討する。

三機関のメタデータ連携

データ連携【連携図】(2021-)



- 連携の種類、目的、運用状況などを整理する作業に着手。
- 現状を可視化することで、効率化と流通する情報の充実にもつなげる。



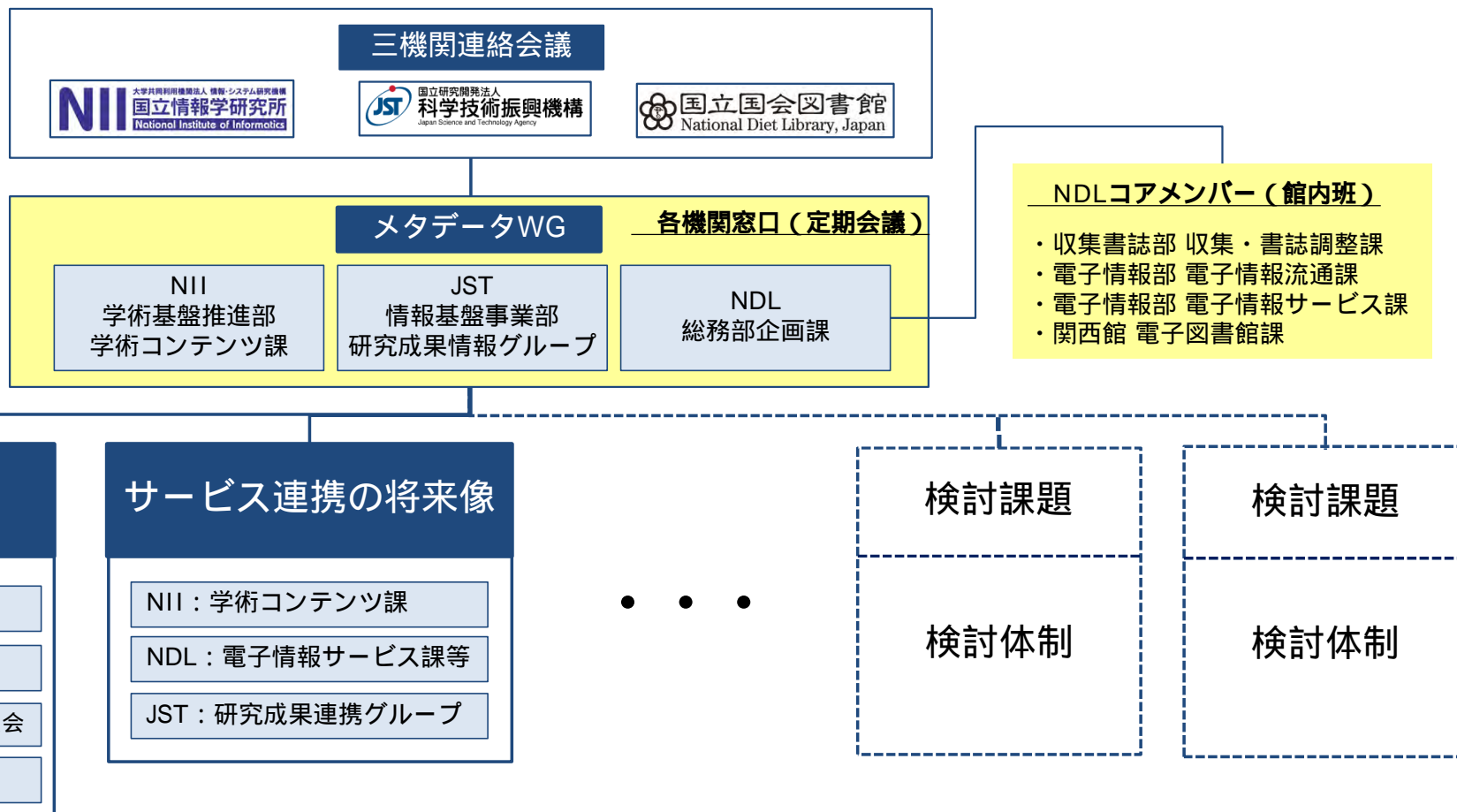
三機関間内の整理プロセスや検討体制を基盤とし、メタデータ連携（メタデータ流通）の改善検討をより広範囲に。

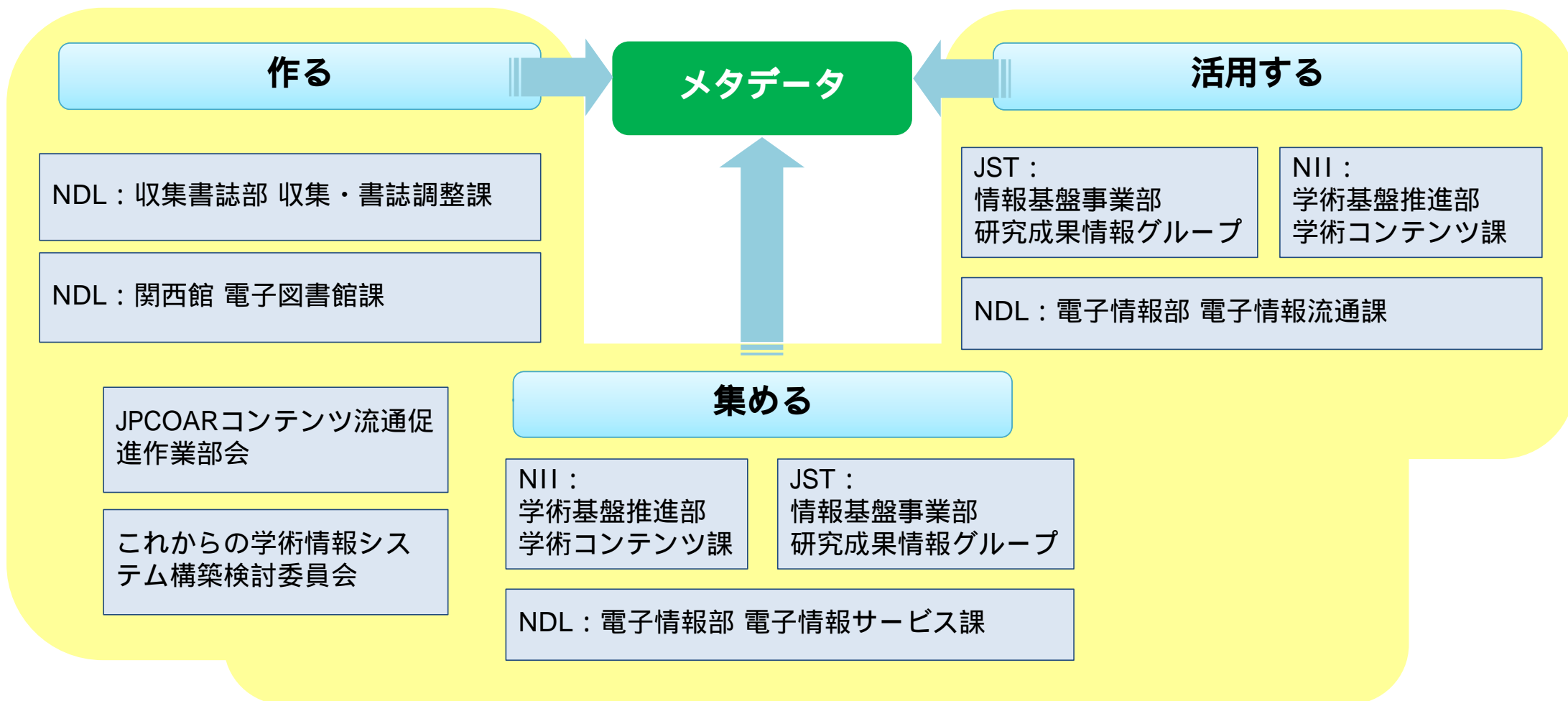
メタデータWGの検討体制

「メタデータの相互運用性に関する検討ワーキンググループ（メタデータWG）」実施体制

この検討基盤のもと
課題を柔軟に設定

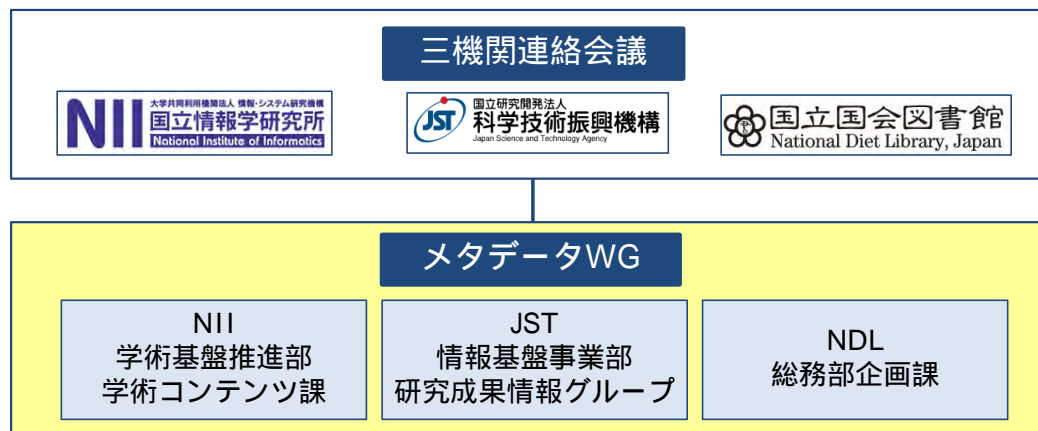
個別検討
課題に応じて検討チームを立ち上げる。
主担当機関を決め、窓口で状況報告





「メタデータ流通ガイドライン」の公開

「メタデータの相互運用性に関する検討ワーキンググループ（メタデータWG）」実施体制



デジタルアーカイブのメタデータ流通

NII：学術コンテンツ課

JPCOARコンテンツ流通促進作業部会

これからの学術情報システム構築検討委員会

NDL：電子情報サービス課等

各大学図書館等でデジタル化された貴重資料等のコンテンツについて、ジャパンサーチ等の統合検索サービスを通じてアクセスを可能とするためのシステム、メタデータ、識別子付与のあり方を検討する。

検討成果として公開

メタデータ流通ガイドライン

2. 「メタデータ流通ガイドライン」の作成

- 対象は「**射程を幅広く**」、そして目的は「**明確に**」
- ガイドライン活用のため「**検討成果を形にする**」
- 継続的な維持・検討のため「**将来に活かす検討プロセスの構築**」

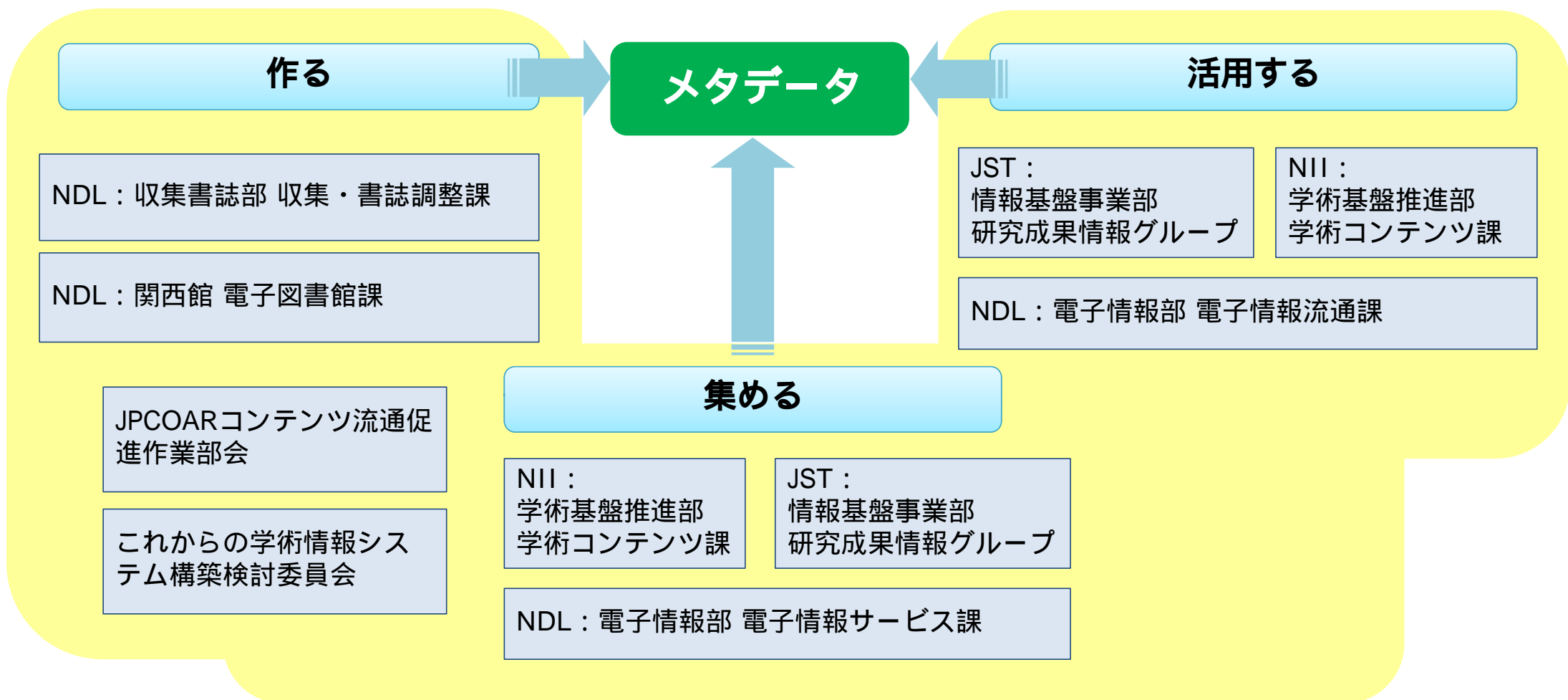
- 主な対象

デジタルアーカイブ、蔵書目録をはじめとする
各種データベースのメタデータ設計に関わる方

- 主な目的

図書館等が扱う多様な情報資源の発見可能性を高めるため、
持続可能性のあるメタデータ流通を実現する

(再掲) 様々な視点からの検討



検討成果を形にする、活用する

目的

多様な情報資源の
発見可能性の向上

持続可能性のあるメ
タデータ流通の実現
(システム間連携コ
ストの最小化)

R3までの検討範囲

デジタルアーカイブの流
通に必要なメタデータ要
素の検討

DC-NDL (RDF) と
JPCOARスキーマのデー
タ互換性の整理

学術機関(大学等)のデ
ジタルアーカイブのメタ
データの適切な流通経路
の整理

成果物

共通編

資料種別編

別紙：メタデータ流通経路

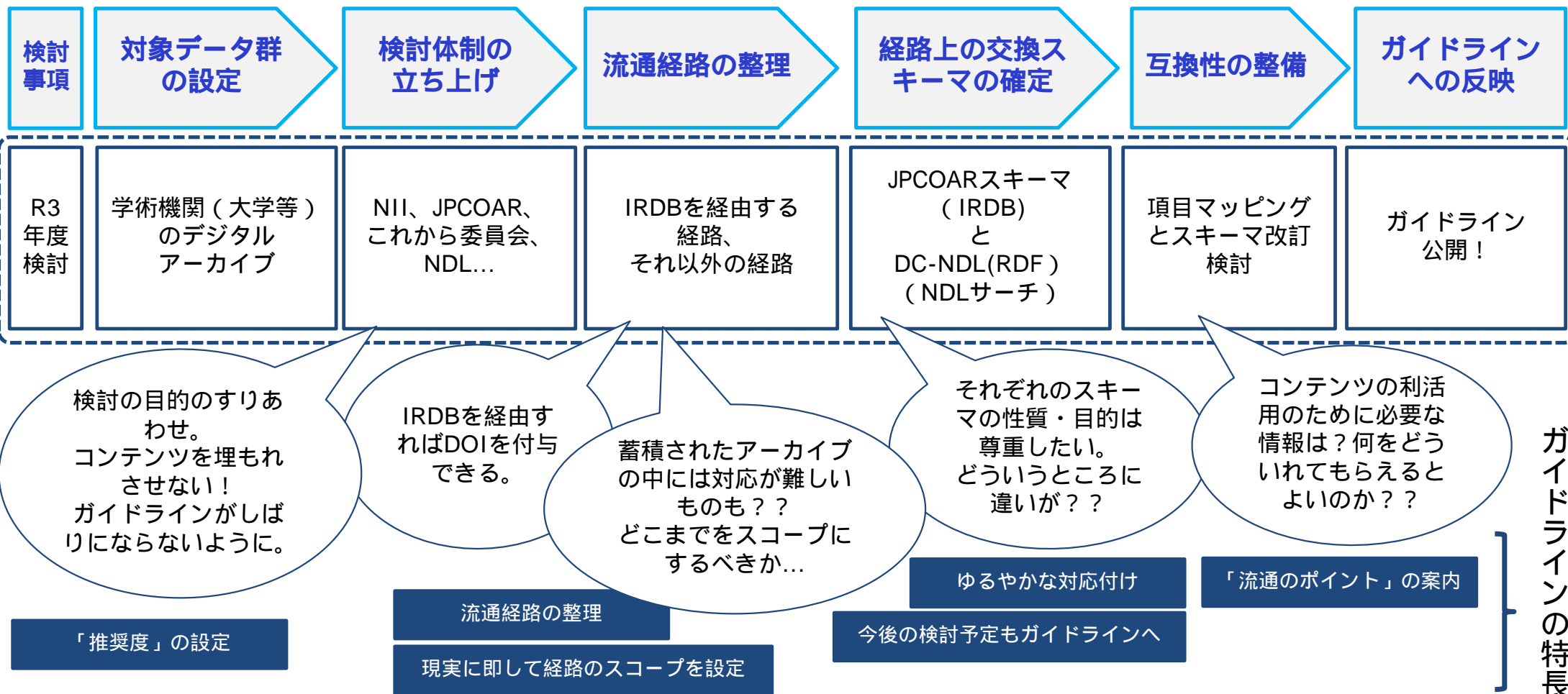
別紙：データ項目一覧

スキーマの改訂

システム対応

ガイドライン

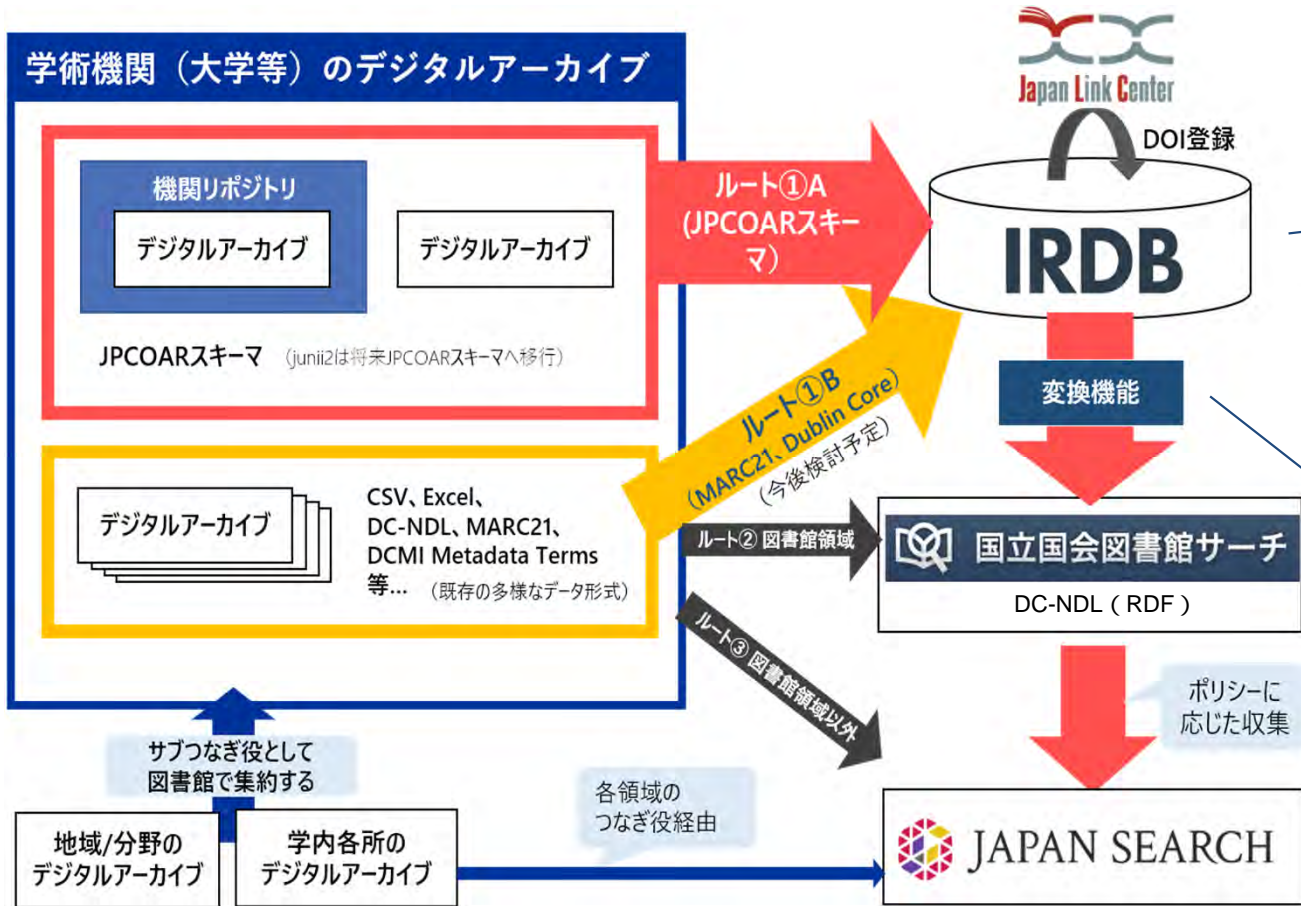
将来に活かす検討プロセスの構築



3. 「メタデータ流通ガイドライン」の特長

- **流通経路の整理、スキーマ間の互換性整備**
- **推奨度の設定**
- **「流通のポイント」を案内**
- **改善・拡充のプロセスの維持**

流通経路整理とスキーマ間の互換性整備



具体的な流通経路の整理

具体的なデータ群を想定した収集・集約の方向性の提案、またこれを基にした、IRDB、NDLサーチ、ジャパンサーチといったアグリゲータ同士のメタデータ流通の経路が整理されることで、より持続可能性が高く、広範囲で、効率的な流通を実現。

スキーマ間の互換性整備

連携ルート上のメタデータ交換のスキーマ同士で互換性を整備。項目新設の検討、項目の解釈・運用の整理や、課題の共有などを行うことで流通促進と発見可能性向上の両方を実現。

ゆるやかな標準化

メタデータ流通に際して求められる各項目の推奨度を設定

メタデータの
詳細度が上がる

<p>強く推奨 (Strongly Recommended)</p>	<p>多くのサービスでの利用が想定され、メタデータやコンテンツの利用にあたって重要。ぜひ入力いただきたい情報。</p>
<p>推奨 (Recommended)</p>	<p>可能な限り付与することが望ましい情報。入力することで、類似資料との識別がしやすくなるなど、より活用の幅が広がる。</p>
<p>任意 (Optional)</p>	<p>メタデータをよりリッチにするための情報。細やかな検索や高度な利活用にも対応でき、専門的な研究にも資することができる。</p>



活用に重要な情報（強く推奨）と、よりリッチにしていくための情報を明確に

11.コンテンツ

項番	項目名	推奨度	DC-NDL (RDF)	JPCOARスキーマ	その他
11-1	IIIFマニフェスト	推奨	rdfs:seeAlso rdf:type="http://iiif.io/api/presentation/2#Manifest"/	[jpcoar:URI objectType = IIIFother]	seeAlso
11-2	サムネイル	強く推奨	資料のサムネイル画像URL : foaf:thumbnail	jpcoar:URI objectType = thumbnail	thumbnail
11-3	本文情報 (一次資料)	推奨	一次資料へのリンクURL : owl:sameAs	jpcoar:URI objectType = fulltext	sameAs

コンテンツに関する各種情報は、メタデータ流通の過程においてコンテンツの発見可能性を高めるため重要です。

「流通のポイント」を案内

9-1. ライセンス情報

項目名	推奨度	DC-NDL (RDF)	JPCOARスキーマ	その他
ライセンス情報	強く推奨	[license]	dc:rights	license

流通のポイント

ライセンス情報の記録は、汎用的な形式で出力いただくと利便性が高まります。

CC0	https://creativecommons.org/publicdomain/zero/1.0/deed.ja
PDM (パブリックドメイン)	https://creativecommons.org/publicdomain/mark/1.0/deed.ja
CC BY (表示)	https://creativecommons.org/licenses/by/4.0/deed.ja

権利区分の種類について以下のようなページも参照ください。

- 参考) ジャパンサーチ デジタルコンテンツの二次利用条件表示について
<https://jpsearch.go.jp/policy/available-rights-statements>

メタデータのライセンス情報は、流通上必須の項目であり、原則CC0と設定されることが推奨されます。

メタデータのオープンなライセンスは直接のメタデータ連携機関だけでなくその先のすべての利用者に対して利用の可能性を開くものです。ライセンスが付与されていない、または、ライセンスがオープンではない場合、メタデータの利用は直接のメタデータ連携機関に限られるものとなり、メタデータの利活用につながりません。仮にメタデータをオープンデータにできない場合でも、利用者のためには何らかのライセンス情報を付与することが推奨されます。

コンテンツのライセンス情報についても、情報資源の利用を促進する観点から付与することが推奨されます。

流通における考慮点の記載

メタデータスキーマの仕様とは異なり、メタデータを流通させる視点での留意点などを記載

他の基準や案内等への参照

新たな基準を作るのではなく、参照したほうがよい既存の基準や案内にそった記載。

ガイドラインが、メタデータ流通の促進を共通の目的とする、
複数のスキーマ、システム、機関等々が対話を行う土台となる

様々な関係者や利用者のフィードバック

情勢の変化に応じ、柔軟に検討・反映



継続的な改善

資料種別編の拡充



メタデータ流通ガイドライン

ガイドライン自体がコミュニケーションツールとして機能していく

課題の解決



目的の達成

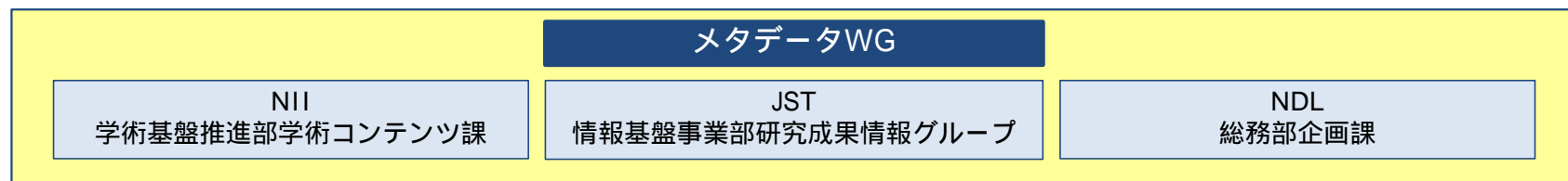
多様な情報資源の発見可能性の向上

持続可能性のあるメタデータ流通の実現（システム間連携コストの最小化）

4. 公開後の動き

R4年度メタデータWGの検討課題

「メタデータの相互運用性に関する検討ワーキンググループ（メタデータWG）」検討課題



デジタルアーカイブのメタデータ流通

NII：学術コンテンツ課

JPCOARコンテンツ流通促進作業部会

これからの学術情報システム構築検討委員会

NDL：電子情報サービス課等

アグリゲータで実施するデータクレンジング（正規化等）に関する情報共有

ガイドライン資料種別編の拡充

研究データ

NII：オープンサイエンス基盤研究センター、学術コンテンツ課

JST：情報基盤事業部研究成果連携グループ

NDL：電子情報サービス課等

ガイドライン資料種別編への反映

サービス連携の将来像

NII：学術コンテンツ課

NDL：電子情報サービス課等

JST：研究成果連携グループ

CiNii Researchリリース以降の新しい連携に向けた取組

メタデータのオープンライセンス化

JST：情報基盤事業部研究成果連携グループ

NII：学術コンテンツ課

JPCOARコンテンツ流通促進作業部会

これからの学術情報システム構築検討委員会

NDL：電子情報サービス課等

ガイドライン検討から独立して、課題の共有や抽出

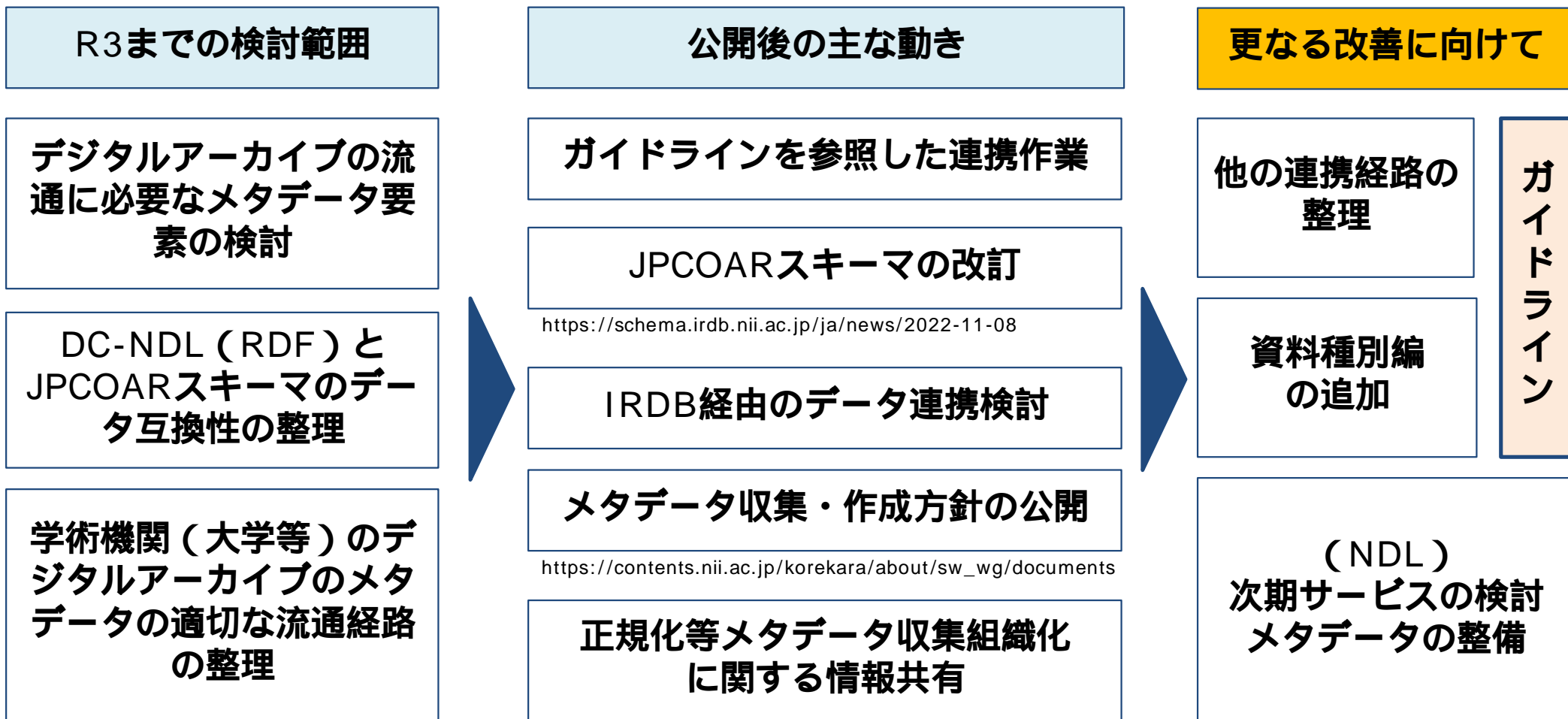
視覚障害者等用データに関する検討

NII：学術コンテンツ課

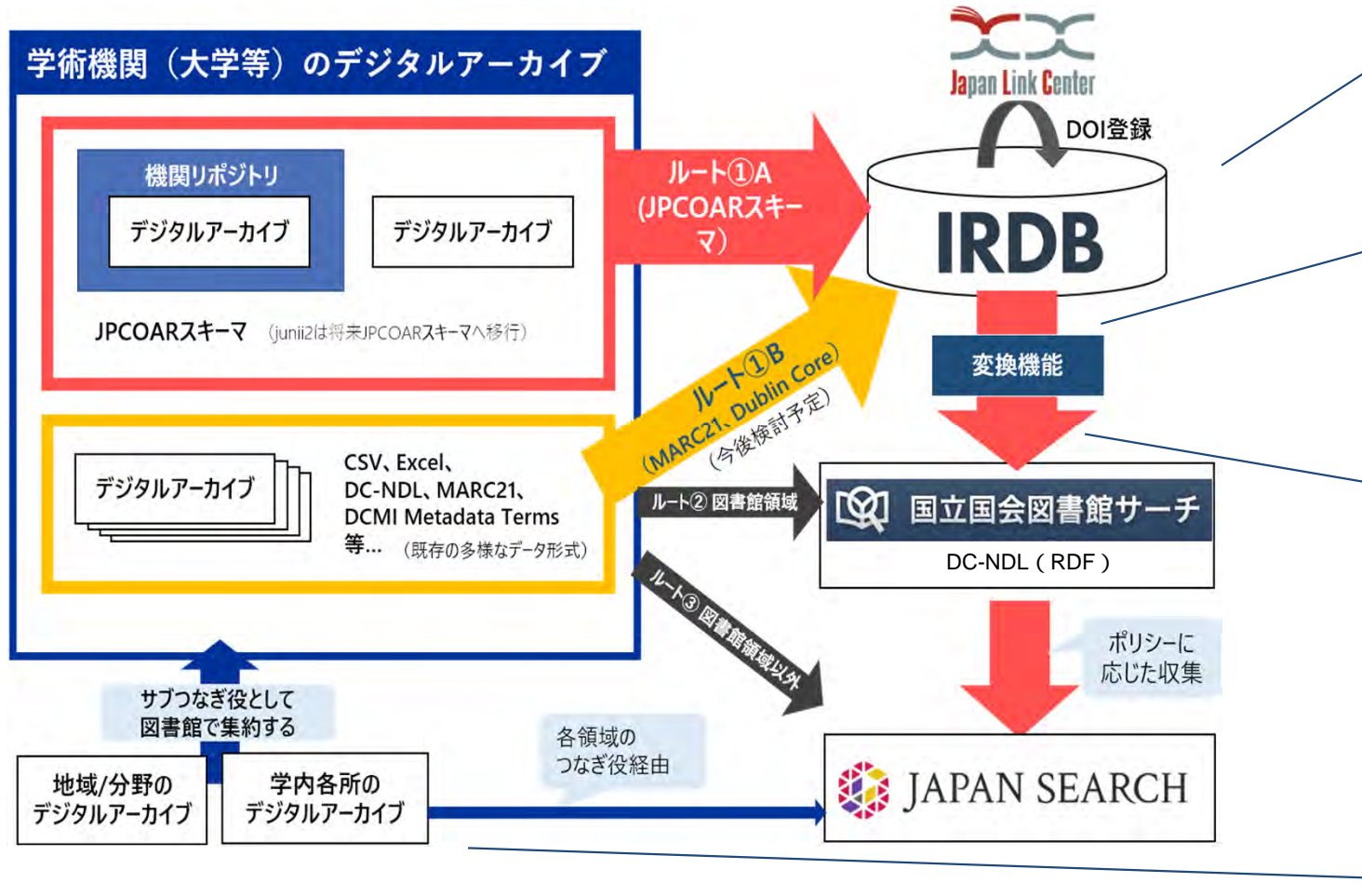
NDL：電子情報サービス課等

NDL：図書館協力課

大学等が保有するアクセシブルな資料のメタデータ共有システム構築と連携。また、アクセシブルな資料に関するメタデータ整備。



学術機関のデジタルアーカイブ検討



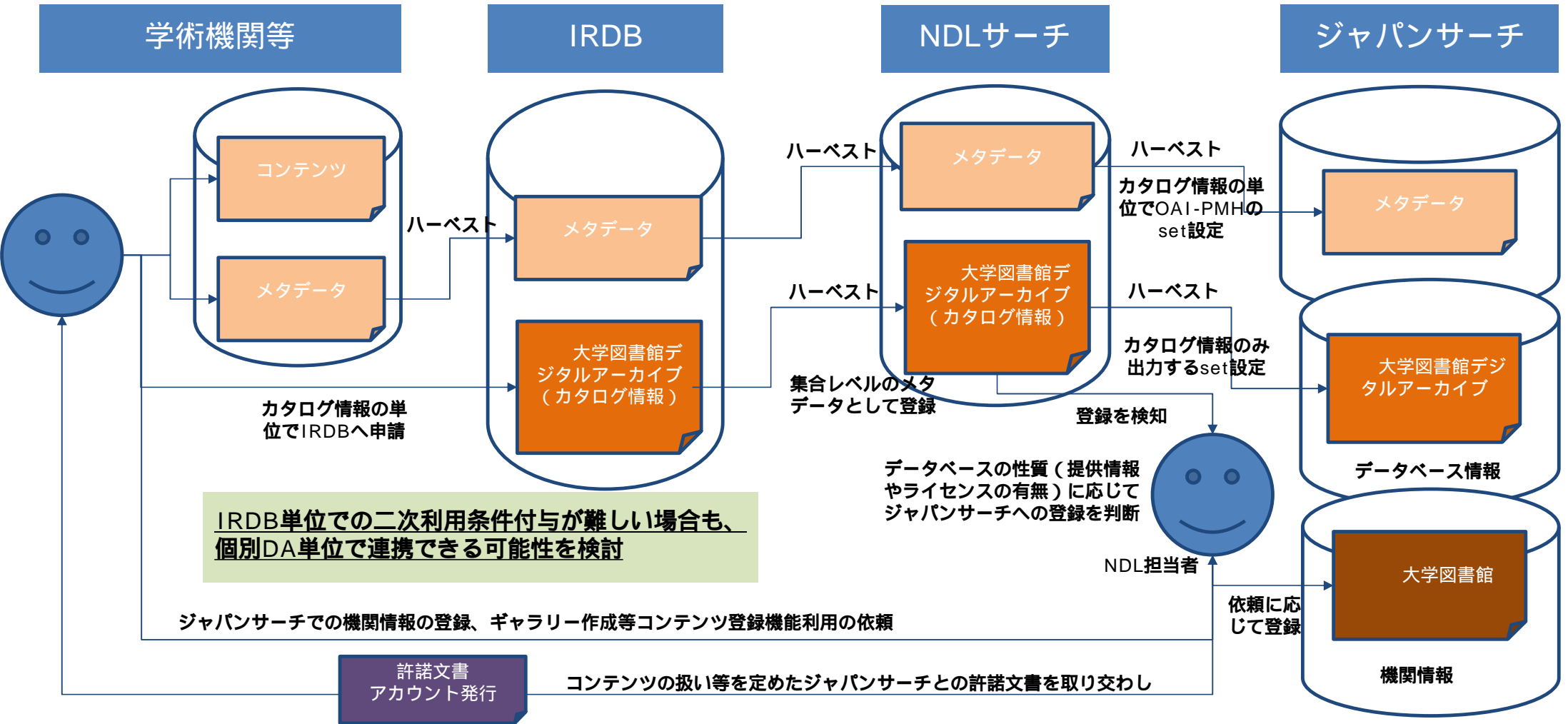
JPCOARスキーマの改訂
デジタルアーカイブに必要な要素をふまえた項目の追加等。

IRDB経由のデータ連携
個々のデジタルアーカイブの単位を維持した連携の検討。

メタデータのオープンライセンス化
IRDB経由のメタデータの条件整備、抄録・解説・サムネイルなど利用条件を示す項目の詳細度。個別課題に切り出し。

メタデータ収集・作成方針
学術機関のメタデータ作成と収集方針の公開。デジタルアーカイブの収集方針についても触れられている。

IRDB経由のデジタルアーカイブ連携



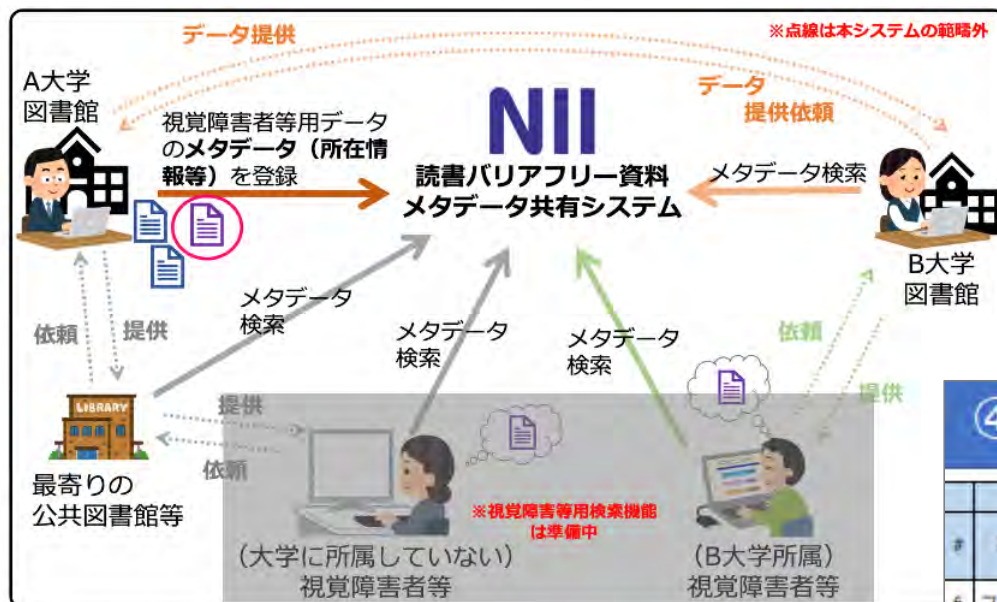
サービス連携の将来像

CiNii Researchとの連携検討

データ種別	収録データベース	NDL側での登録想定	提供方法（次期サービス）
本	CiNii Books 科学研究費助成事業データベース（図書）	連携対象	<ul style="list-style-type: none"> ・ ISBN等の識別子により、当館、公共図書館蔵書との同定実施 ・ 原資料が特定できる場合、デジタル化資料等の形態の異なる資料とも同定を実施検討
博士論文	CiNii Dissertations	連携対象外	
論文	Japan Link Center * IRDB * 雑誌記事索引データベース 国立国会図書館デジタルコレクション Crossref 医中誌Web 日経BP記事検索サービス CiNii Articles 国立国語研究所 * 情報学研究データリポジトリ * Dbpedia * 科学研究費助成事業データベース（論文）	連携対象 重複分については、登録時に除外するか、もしくは同定を検討	<ul style="list-style-type: none"> ・ DOIなどの識別子が一致する場合の同定を検討 <p style="text-align: right;">下線 = NDLサーチ既存連携と重複</p> <p style="text-align: right;">* = 研究データの提供もあり</p>
研究データ	DataCite JDCat Rikkyo University Data Archive Integbioデータベースカタログ	連携対象	<ul style="list-style-type: none"> ・ デフォルトの検索対象とはせず、ヒットした論文に（DOI等で）紐づく場合に書誌詳細の関連情報として表示予定 「研究データ」として区別できる場合に限る
プロジェクト・人物	科学研究費助成事業データベース	連携対象外	

視覚障害者等用データに関する検討

国立情報学研究所読書バリアフリー資料メタデータ共有システム <https://a11y.pub.nii.ac.jp/>



- 2022年10月4日本格運用開始
- IRDBを通じてメタデータが提供される
- 今後、連携に向けて調整開始

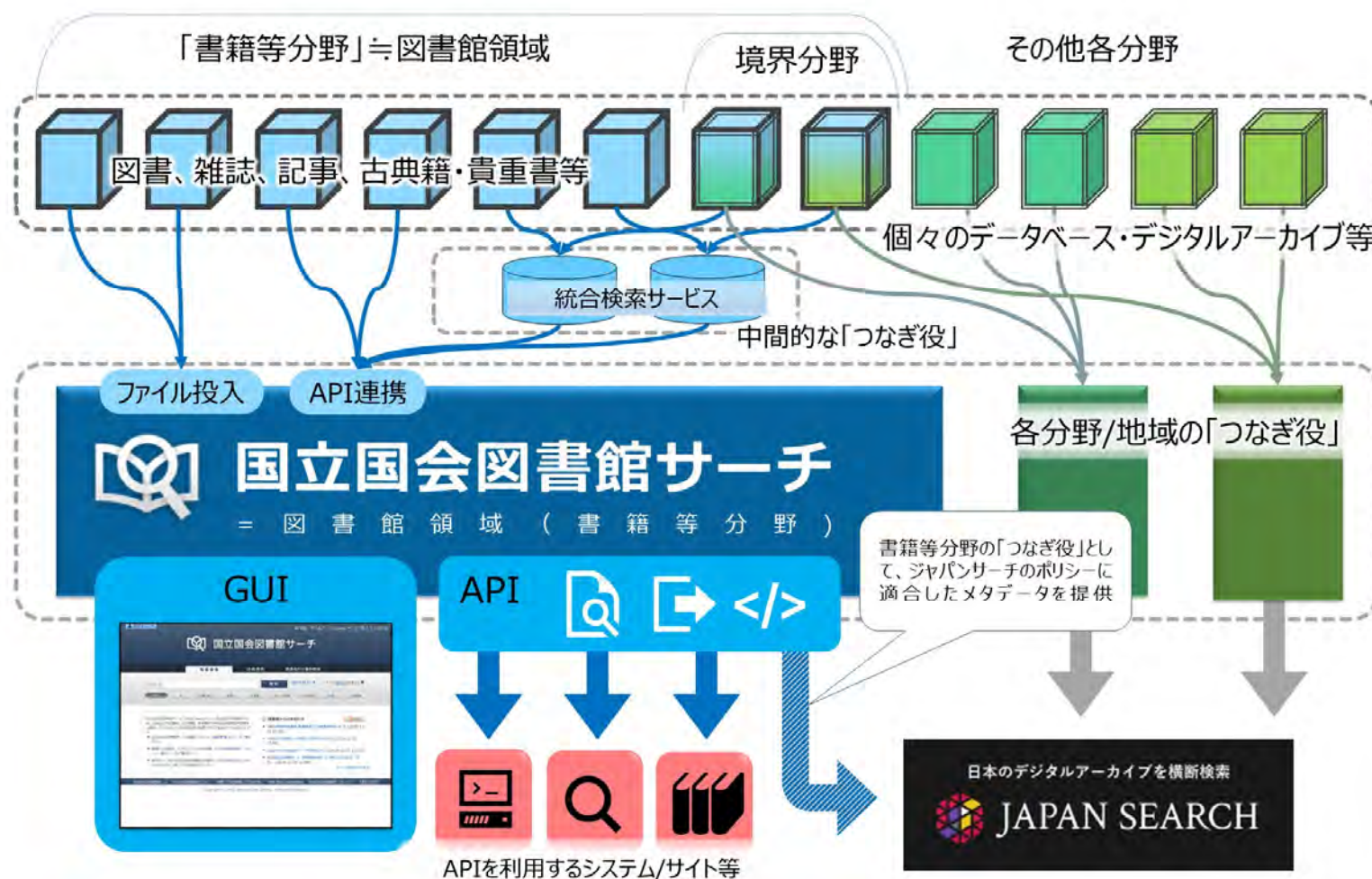
④メタデータ項目

#	項目名		プロパティ		オプション		入力マニュアル		
	項目名(日)	項目名(英)	プロパティ	サブプロパティ	必須	繰返	図書(の一部)	雑誌論文・記事	
6	ファイル形式	Format	内容記述						
						○			自由記述 (PDFファイル、テキストファイル/ テキストDAISY/音声DAISY/マルチメディア DAISY/点字データ、等)
						○		Otherを選択する	
7	校正レベル	Proofreading Level	内容記述						
									自由記述 (校正済/未校正、等)
									Otherを選択する

利用したいコンテンツを、
より特定しやすいメタデータへ

5. なぜ今、ガイドラインなのか ～ NDLサーチのメタデータ連携から～

NDLサーチの役割



- 図書館等が扱う情報資源のメタデータの統合的検索機能の提供。
- メタデータの標準化・オープンライセンス化の促進を図るなど、メタデータ流通を維持・向上するための取り組み。



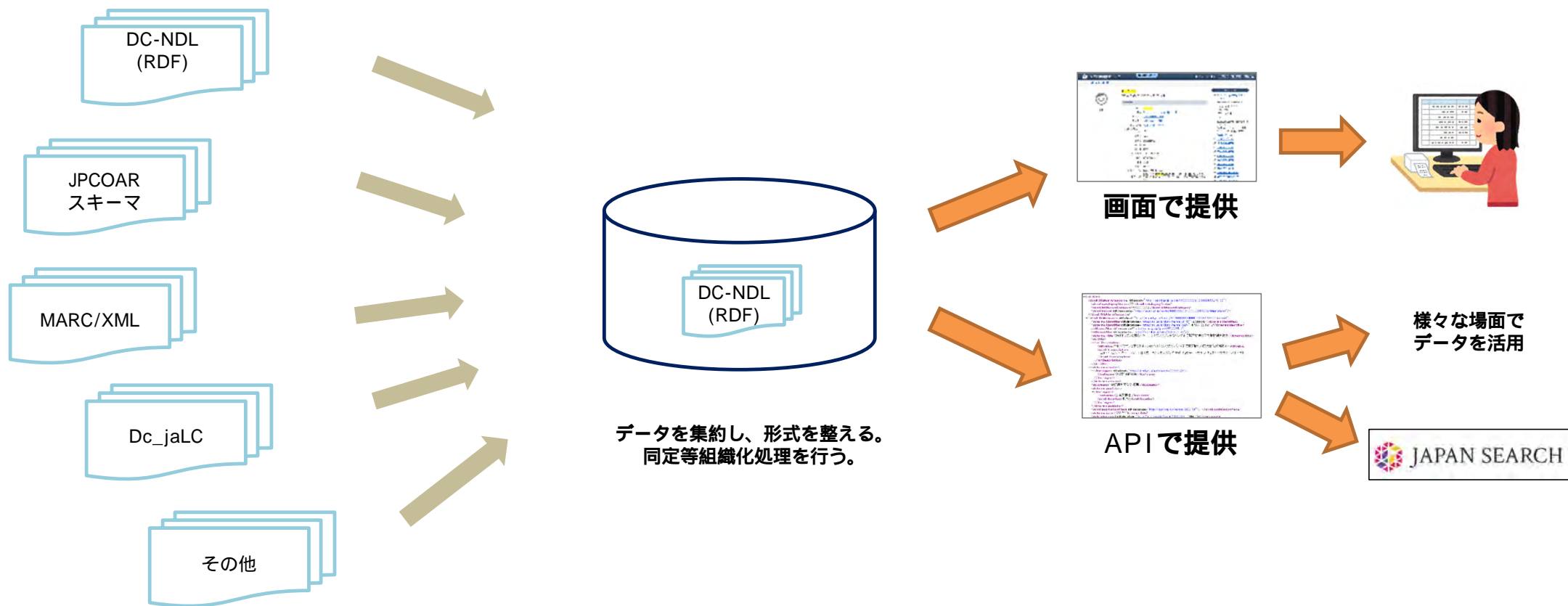
集約したメタデータの流通促進により、連携機関のコンテンツと利用者をつなぐ。

NDLサーチ上でのメタデータの流れ

作る

集める

活用する



メタデータ連携の種類 (1) API

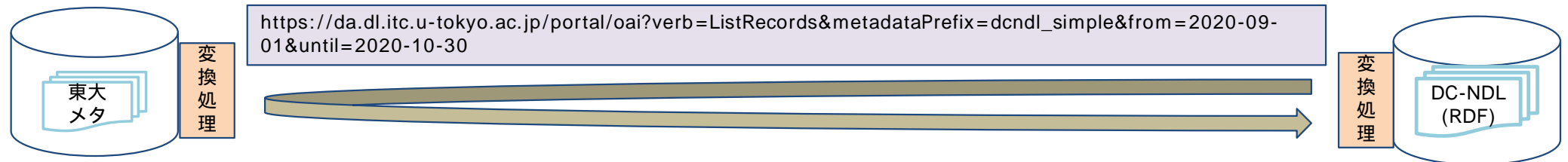
<https://da.dl.itc.u-tokyo.ac.jp/portal/help/api>



- 連携先の更新頻度に応じて、NDLサーチからリクエストを行う。
- 比較的更新の多いDBで連携実績が多い。

ハーベスト 用API	OAI-PMH	URL	XML	https://da.dl.itc.u-tokyo.ac.jp/portal/oai
---------------	---------	-----	-----	---

NDLサーチから、東大ポータルAPIを使ってリクエスト



東大ポータルから、リクエストに応じたデータが返戻される。

```
<dcndl_simple:dc xmlns:dcndl_simple="http://ndl.go.jp/dcndl/dcndl_simp
<dc:identifier>a135b4e4-a0b3-9007-2fd5-4922d2a40471</dc:identifier>
<rdfs:seeAlso rdf:resource="https://da.dl.itc.u-tokyo.ac.jp/portal/as
<dcterms:source rdf:resource="https://iiif.dl.itc.u-tokyo.ac.jp/repo/
<foaf:thumbnail rdf:resource="https://da.dl.itc.u-tokyo.ac.jp/portal/
<rdfs:seeAlso rdf:resource="https://iiif.dl.itc.u-tokyo.ac.jp/repo/ii
<dc:title>大日本海志編纂資料目録</dc:title>
```

返戻されたデータ例 (抜粋)

DC-NDL (RDF) に変換してNDLサーチに格納する。

```
<dcndl:BibResource rdf:about="http://iss.ndl.go.jp/books
<rdfs:seeAlso rdf:resource="https://da.dl.itc.u-tokyo.
<dcterms:title>大日本海志編纂資料目録</dcterms:title>
<dc:title>
  <rdf:Description>
    <rdf:value>大日本海志編纂資料目録</rdf:value>
```

変換した後のデータ例 (抜粋)

メタデータ連携の種類（2）ファイル提供

https://www.plib.pref.aomori.lg.jp/top/digital/about.html

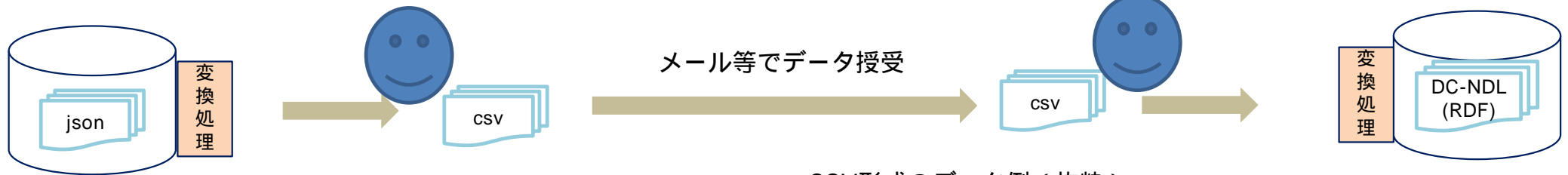
青森県立図書館 デジタルアーカイブ

- ・ 連携先で用意されているCSV等のファイルを取得して格納。
- ・ 比較的更新頻度の少ない連携の場合に採用。

「青森県立図書館デジタルアーカイブ」書誌情報オープンデータセット



(csv形式; 254KB; UTF-8)



CSV形式のデータ例（抜粋）

変換した後のデータ例（抜粋）

```
# 青森県立図書館デジタルアーカイブ書誌情報オープンデータセット (2020-04-14),,,,,,  
,リンク,資料番号,タイトル,タイトルカナ,巻号,著者等,著者等カナ,シリーズ名,シリーズ名カナ,出版者,出版地,  
link,dcterms:identifier,dc:title,dcndl:titleTranscription,dcndl:volume,dc:creator,dcndl:creatorTransc  
1,https://www.plib.pref.aomori.lg.jp/top/digital/cont/10203132933.html,10203132933,数学端記,スウガクタ  
2,https://www.plib.pref.aomori.lg.jp/top/digital/cont/10203132944.html,10203132944,数学端記,スウガクタ
```

```
<rdf:seeAlso rdf:resource="https://www.plib.pref.aomori.lg.jp/top/digital/cont/10203132933.html"/>  
<dcterms:title>数学端記 巻之1</dcterms:title>  
<dc:title>  
<rdf:Description>  
<rdf:value>数学端記</rdf:value>  
<dcndl:transcription>スウガクタンキ</dcndl:transcription>
```

メタデータの流れとシステム連携

連携先図書館

NDLサーチ

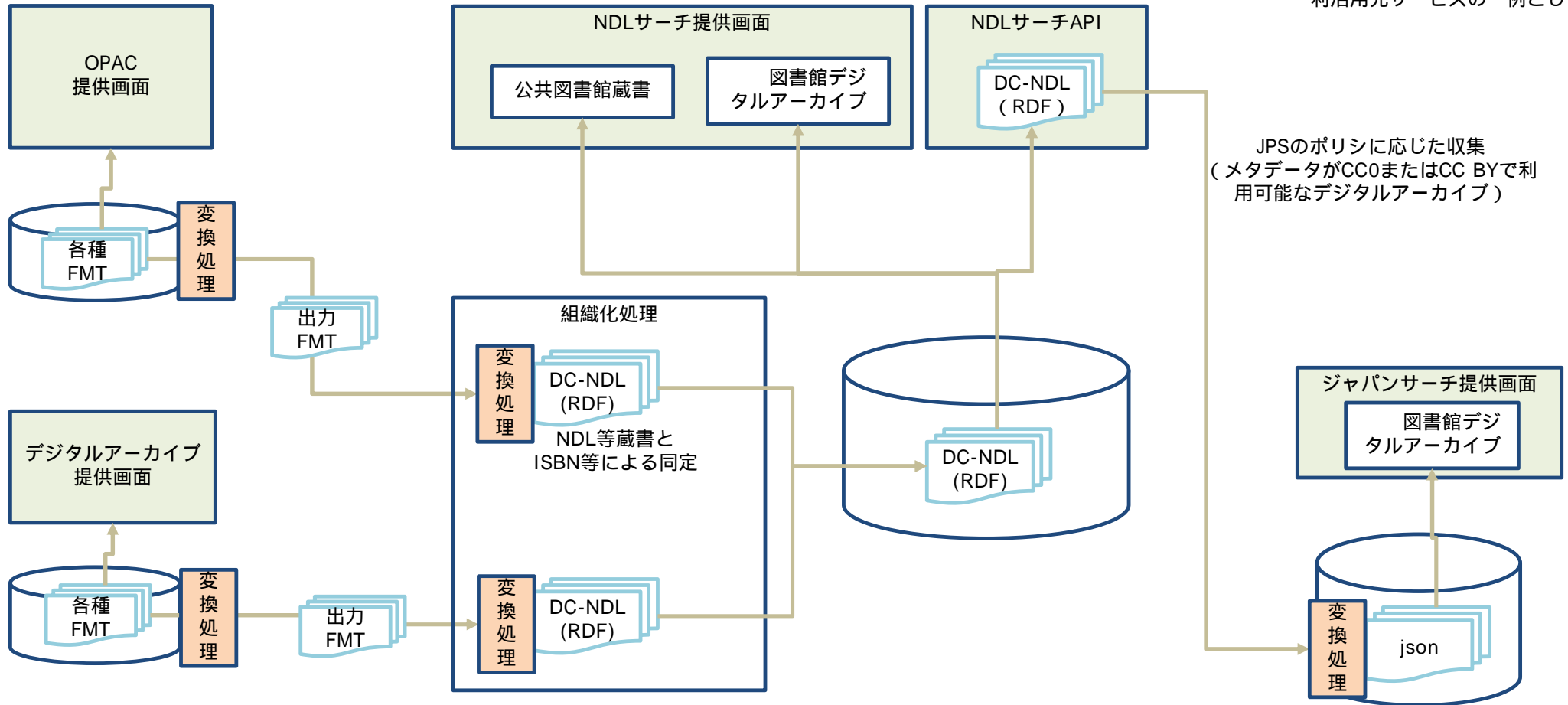
ジャパンサーチ

利活用先サービスの一例として

JPSのポリシーに応じた収集
(メタデータがCC0またはCC BYで利用可能なデジタルアーカイブ)

蔵書目録

デジタルアーカイブ



1つのメタデータには複数の役割がある

連携先図書館

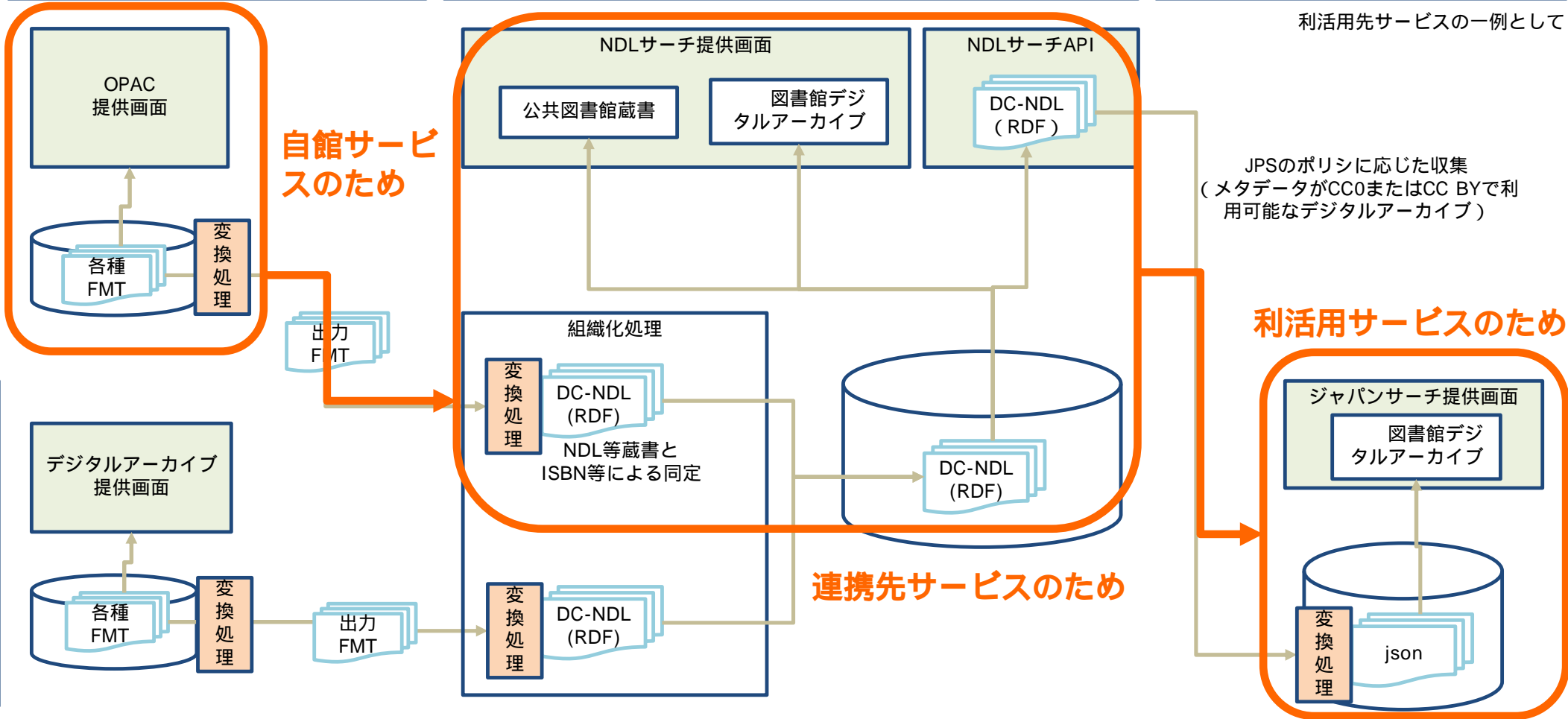
NDLサーチ

ジャパンサーチ

利活用先サービスの一例として

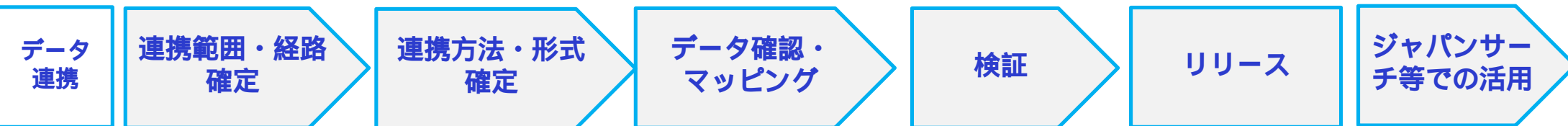
蔵書目録

デジタルアーカイブ

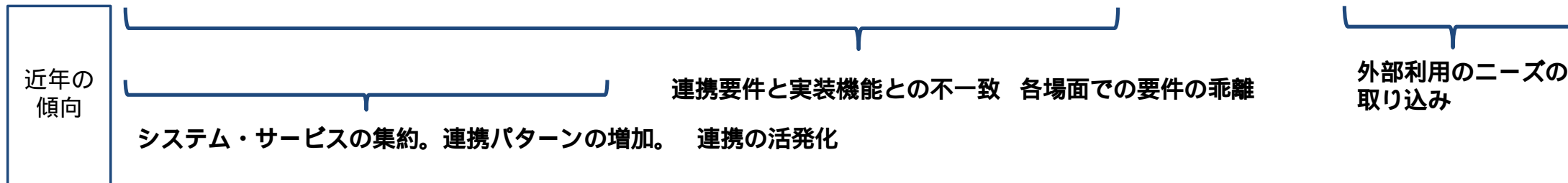


6. 「メタデータ流通ガイドライン」の使いどころ

連携開始に至るまでの手順



主な実施事項	<p>ヒアリングシート、打ち合わせ等により先方の要件を確認し、連携方針等とすり合わせたうえで仕様を確定していく。</p> <p>先方のベンダや、デジタルアーカイブの場合ジャパンサーチ担当も交えて関係者調整 メタデータの二次利用条件の確認も</p>	<p>サンプルデータ等を受領してNDLサーチのデータ形式にマッピング。</p> <p>出力が不足しているデータ項目の追加依頼も。</p>	<p>データ仮登録、性能確認を含む連携テスト。</p> <p>許諾取り交わし、全件データの取得なども。</p>	<p>本番環境へのデータ登録。 リリース後、定常更新も開始。</p>	<p>APIが公開され、他サービスからもデータ利用可能に。</p>
--------	---	--	---	--	-----------------------------------

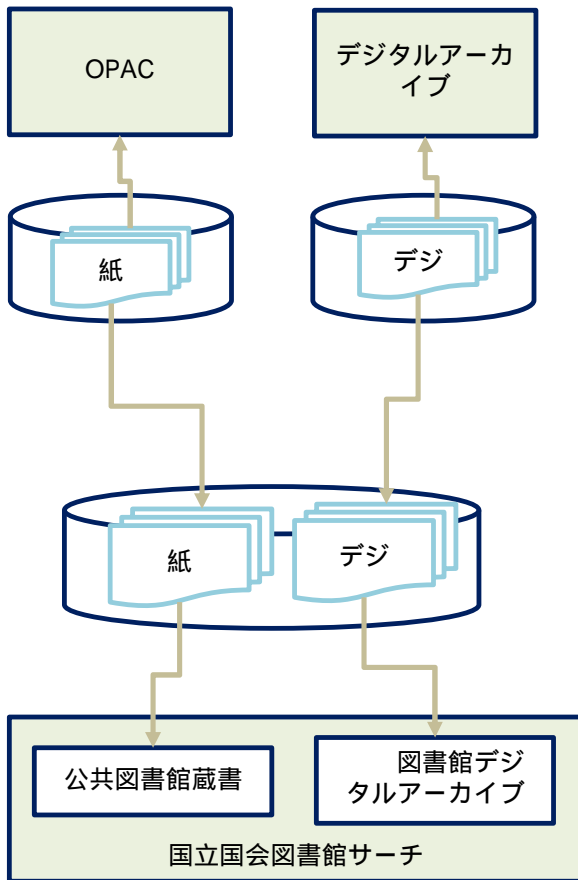


サービス・システムの集約

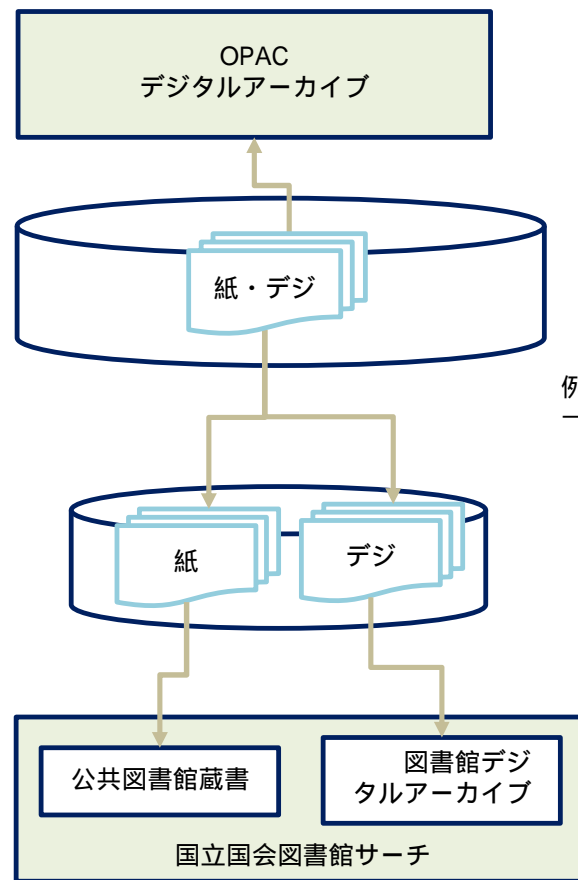
連携先図書館

国立国会図書館

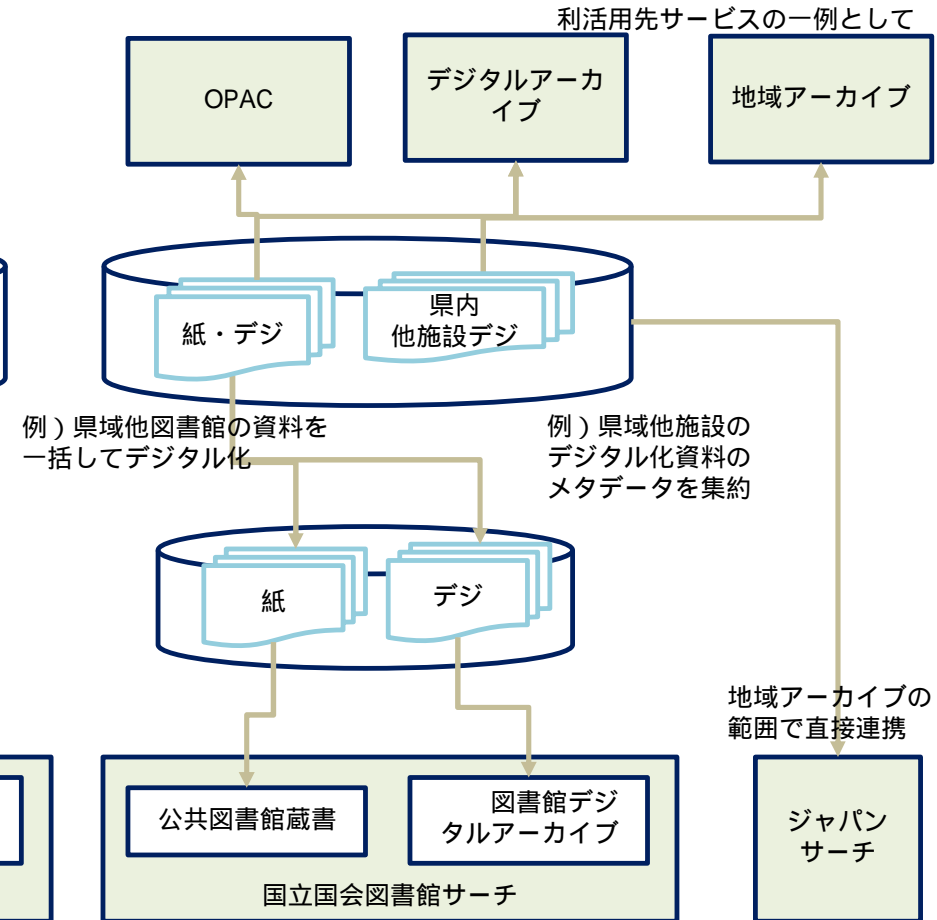
OPAC・デジアカ別々のシステム



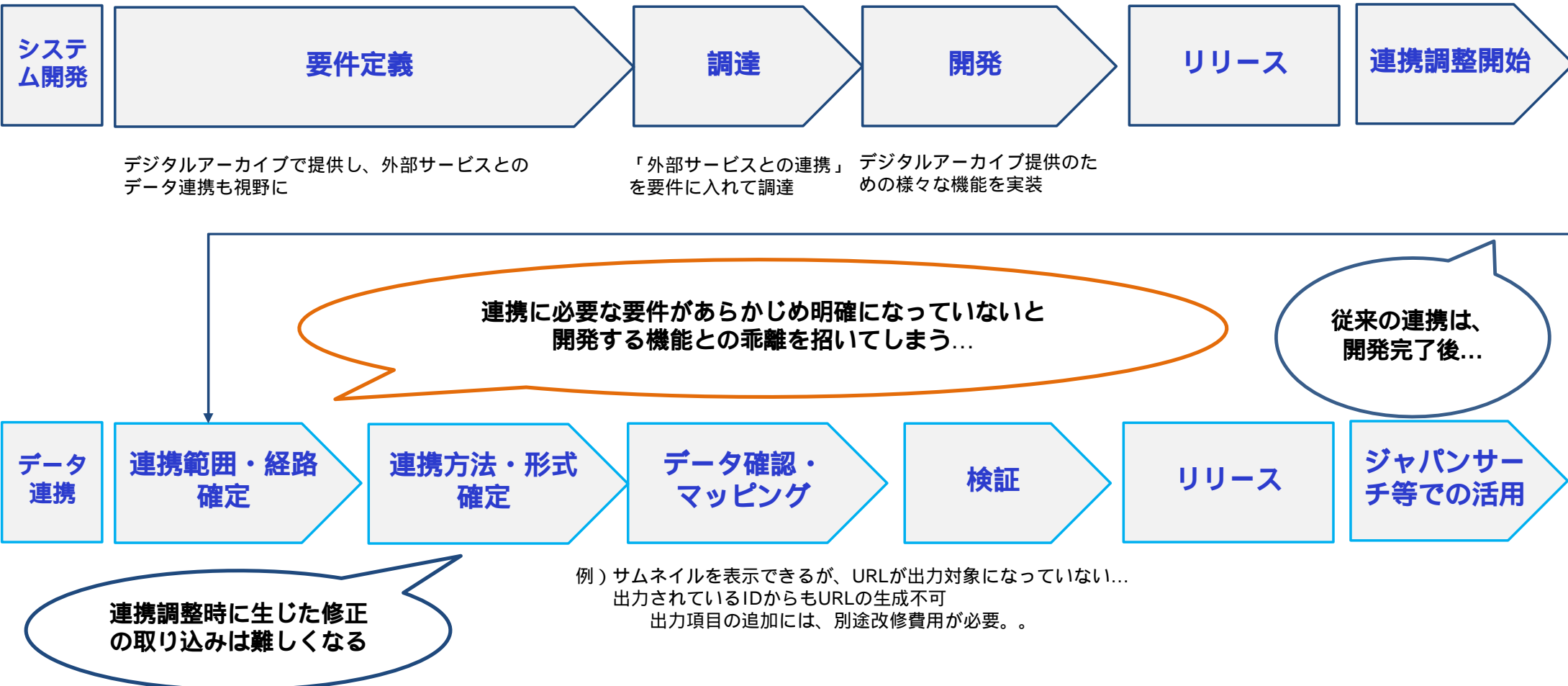
OPAC・デジアカ同じシステム



さらに、様々なデータを集約



連携要件と実装機能との乖離



乖離を防ぐためのコミュニケーション

データ活用側の視点を含む課題の検討

課題検討 ~ ガイドラインへの反映



自館のデータをどこに連携すれば効果的か？

どのスキーマを実装すればよいか？

メタデータの品質。どのレベルまでを出力項目とする？

他スキーマに変換する場合の仕様は？

従来の連携は、開発完了後...
これによる手戻りを防ぐには...

連携先のシステム開発



実装するIFを確定

調達時にガイドラインを参照してもらってもよい

ガイドラインを通じて、
連携開始前からのコミュニケーションを可能に

データ連携



ガイドラインを参照した連携の実現

ガイドラインの特長

流通経路の整理
スキーマ間の互換性整備

推奨度の設定

流通のポイント

資料種別編

R4年度作成中。今後公開予定。

ガイドラインを参照すると...

適切な流通経路を選択

連携機能の実装範囲
を絞り込む

連携メタデータの品質を
状況に応じてコントロール

自館に適したより効果的な
メタデータ連携を選択できる

たとえば...

IRDBを経由させれば、DOIも付与できて、ジャパンサーチまでデータを連携できそうだ...

OAI-PMHとJPCOARスキーマでの出力に対応すればよさそうだ...
他の出力機能は不要。

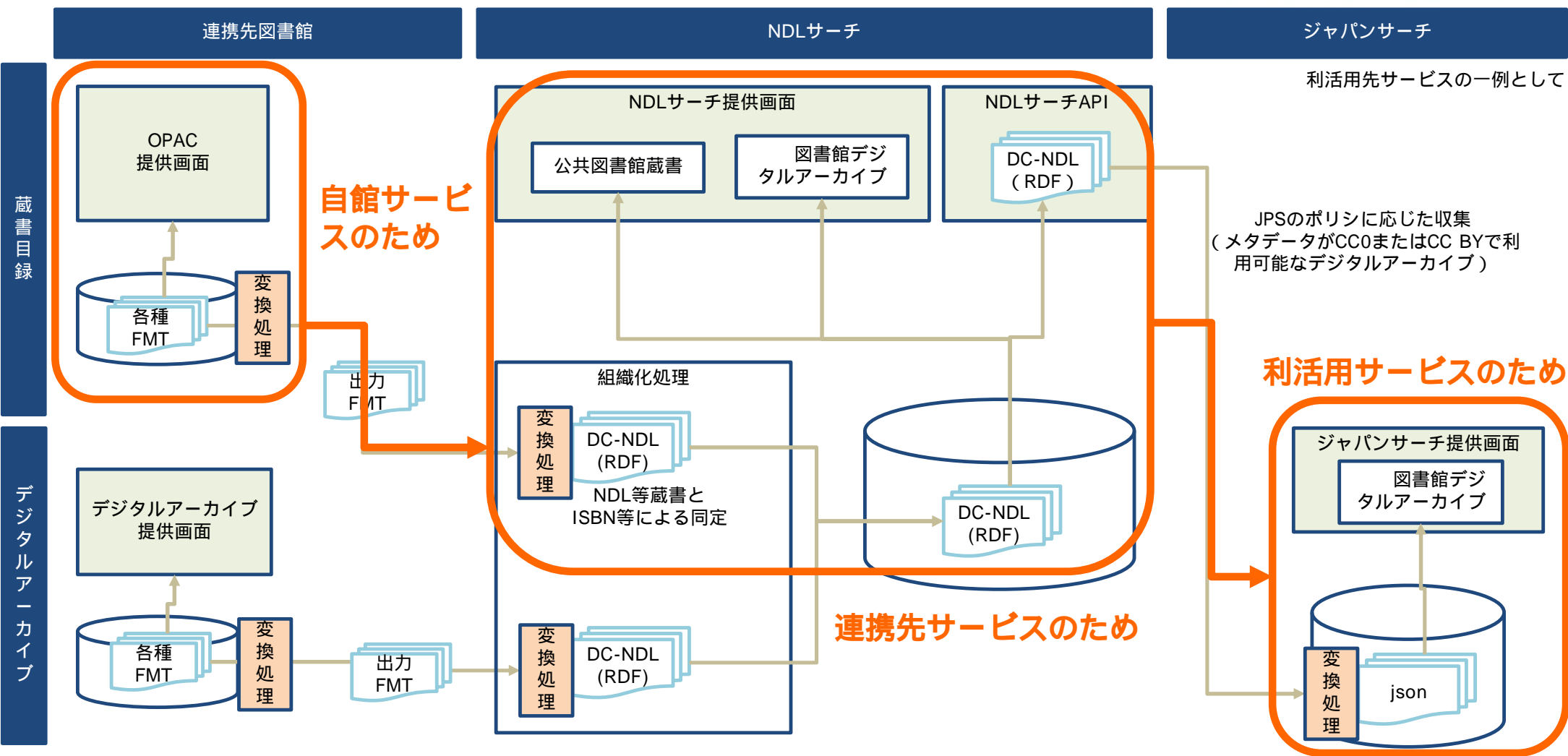
「流通のポイント」を参照することで、外部利用のニーズにも対応できる。
「強く推奨」レベルのポイントに留意して出力仕様を実装。
ただ、自館のデータは古典籍が主体。もう少し古典籍に特化した説明が欲しい...

フィードバックを受けた改善

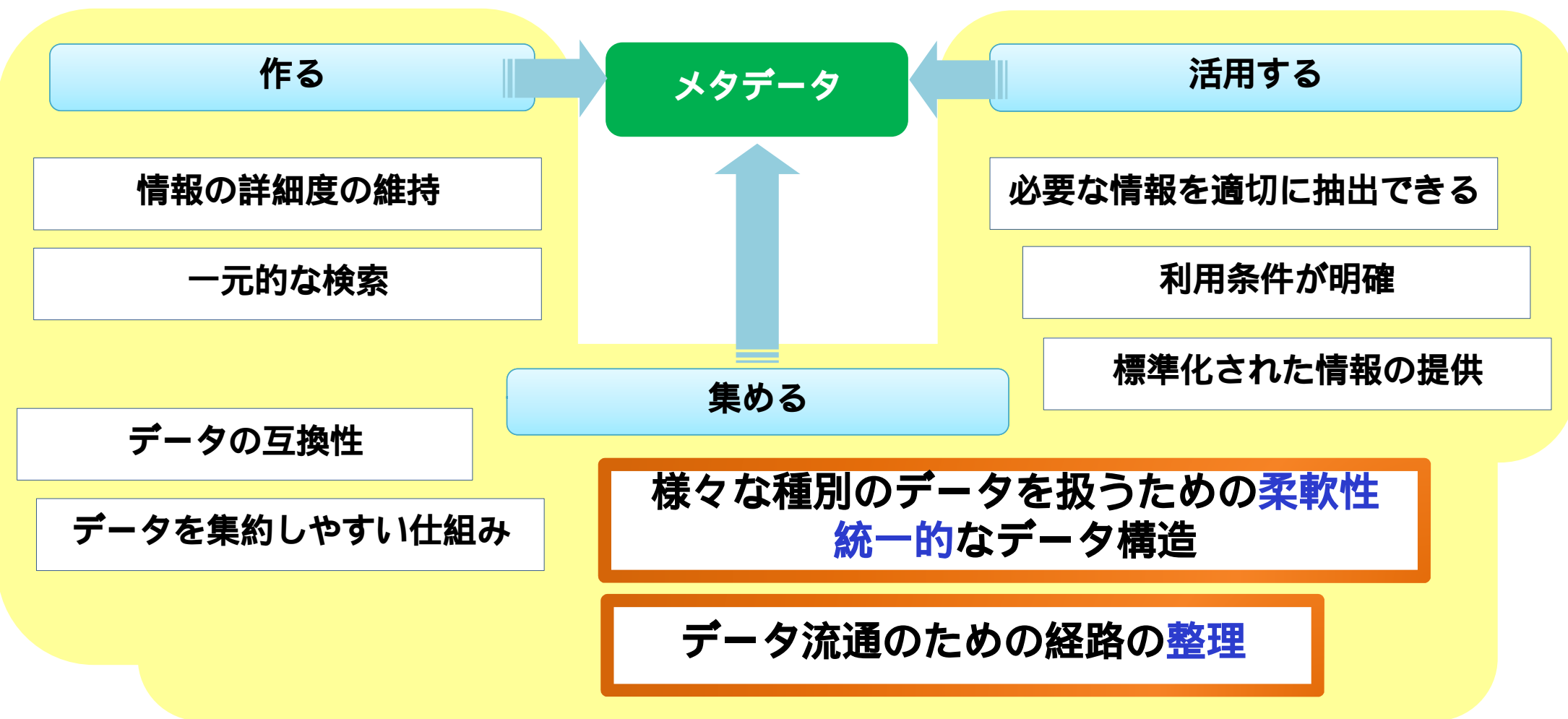
ガイドラインはあくまでも選択肢の提示。
そして、フィードバックの受け口としても機能。

7. 提供サービスと流通データの充実に向けて

NDLサーチのメタデータ...



その役割、求められる要件とは



• 内部データの仕様

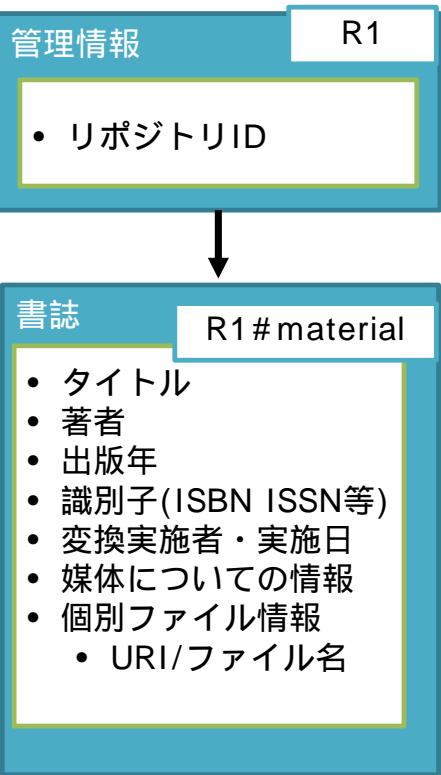
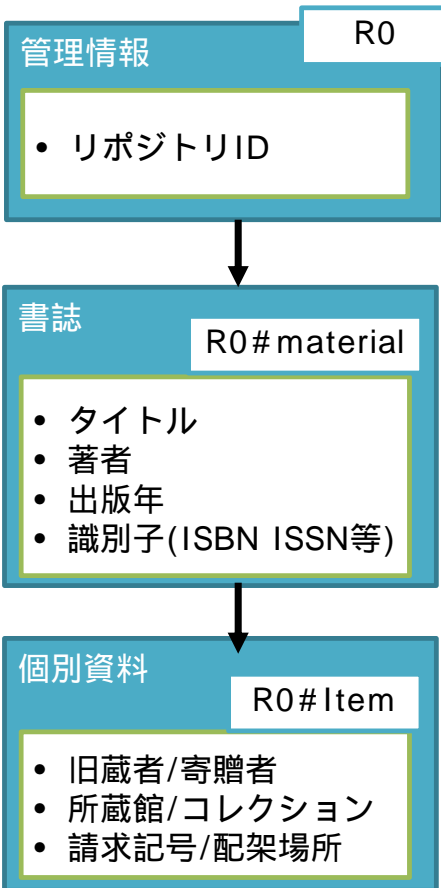
- 特定のスキーマに依存しない仕様とし、**柔軟性**を確保
出力仕様は別定義とし、サービス要件を切り離しやすく
- 資料種別や形態に応じデータ項目を**統一的**に管理
求められる「一元的な検索」の担保

• メタデータの構造

- 紙やデジタルといった形態を問わず、すべて三階層に**統一**
形態に応じた利用方法・条件の特定を容易に
データの関連、由来に関する情報の充実も

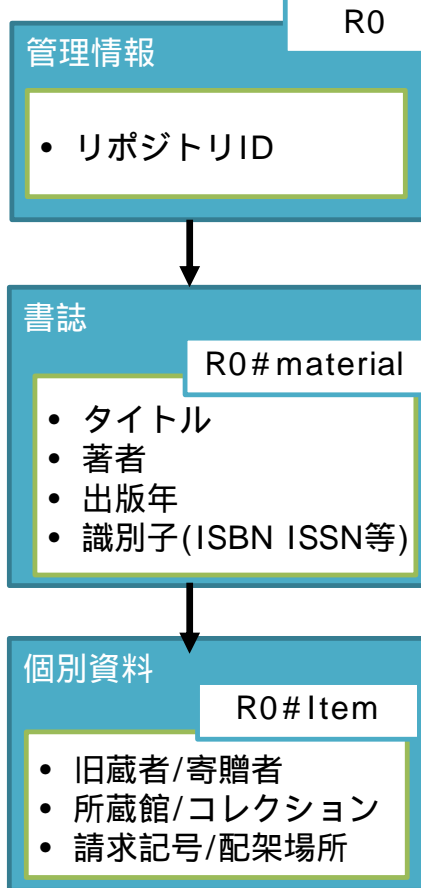
NDLサーチのメタデータ構造と改善点

現在

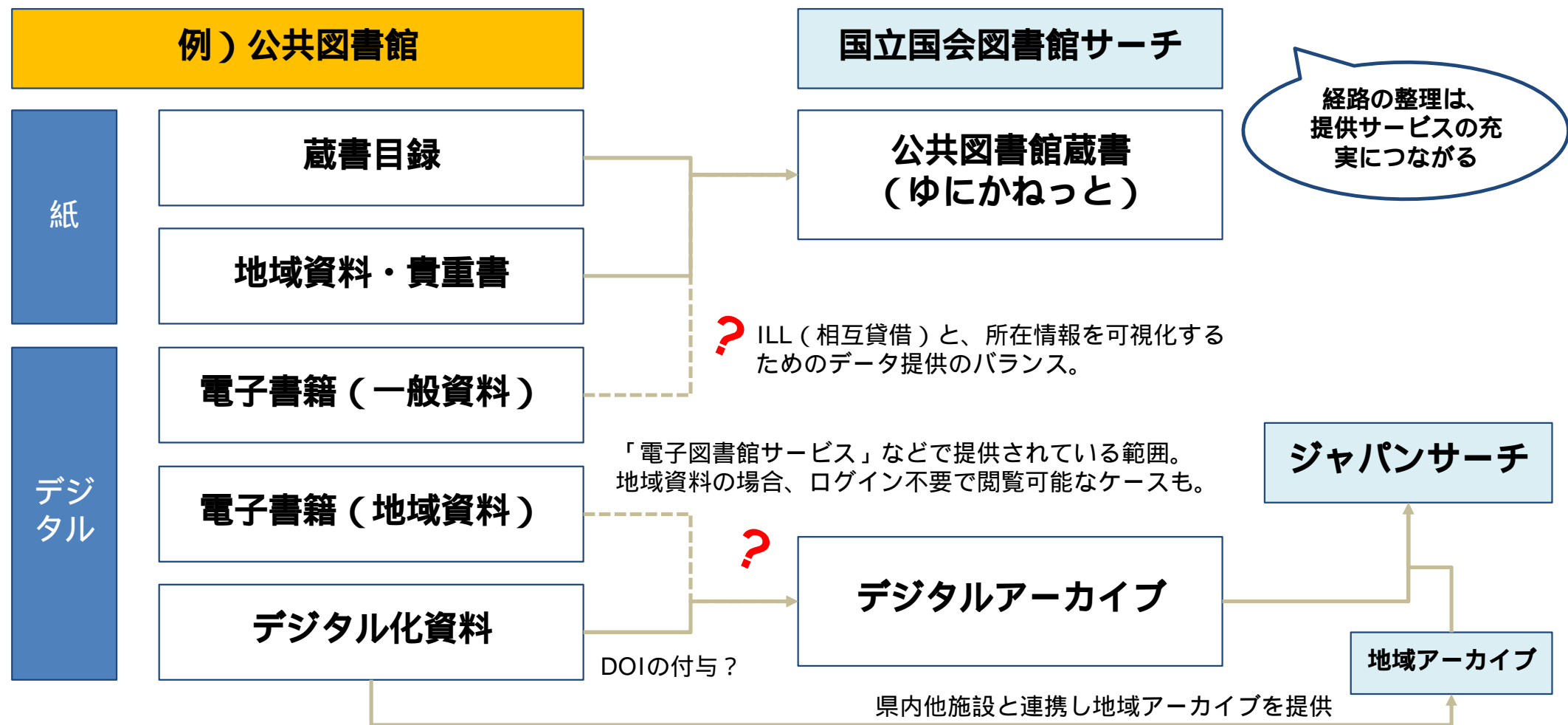


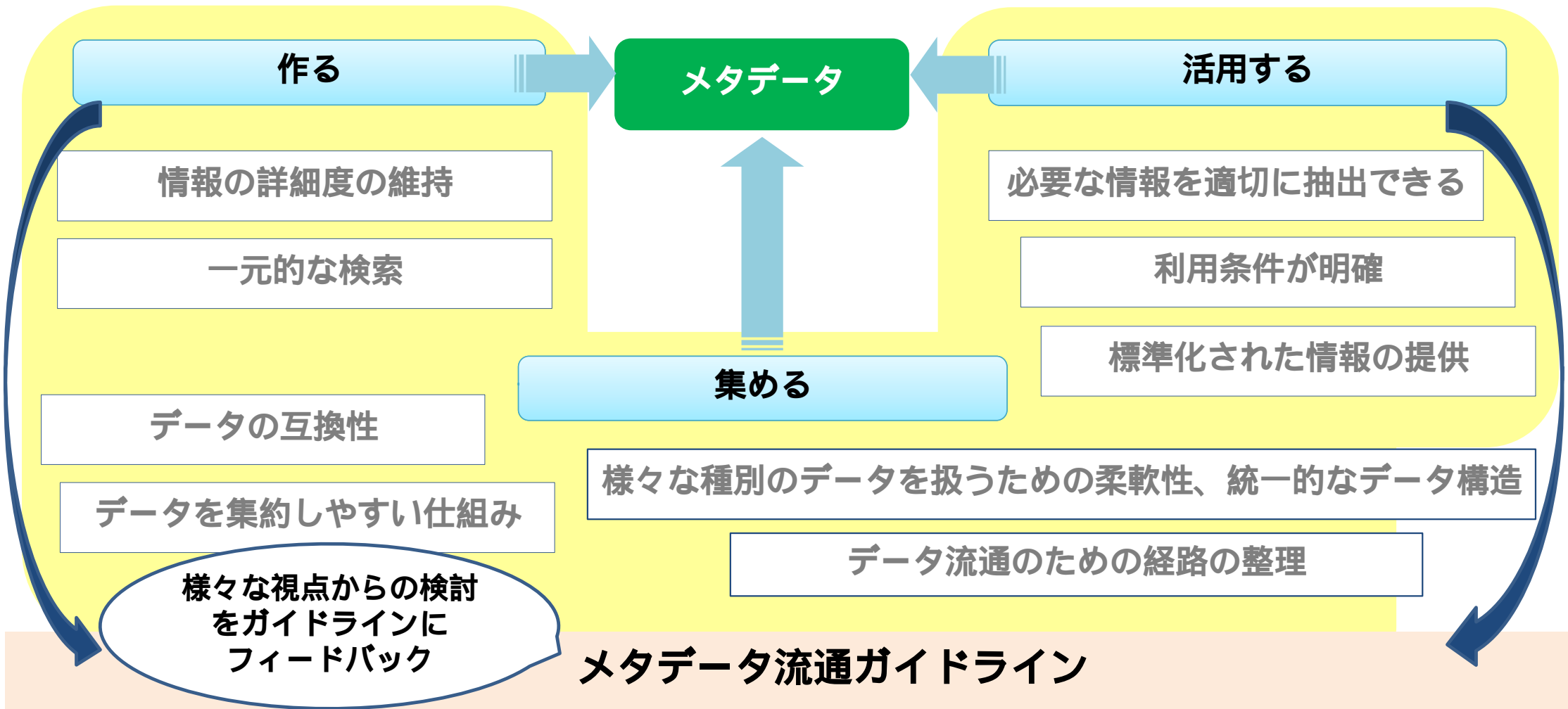
蔵書目録（書誌同定対象）のみ
個別資料（item）を生成

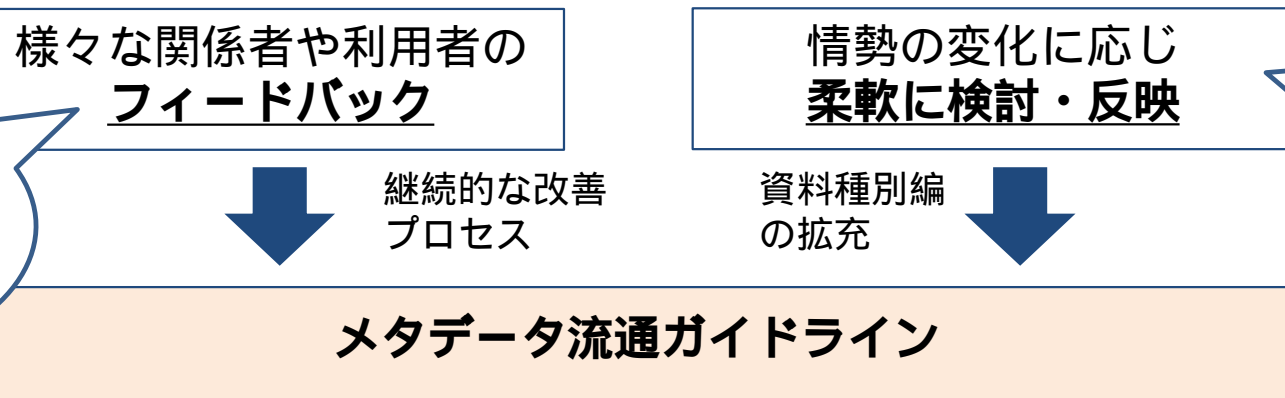
次期サービス



データ流通のための経路の整理







利用者からみると、提供されるデータはフラットな状態。様々な視点からの意見を取り込んでいく。

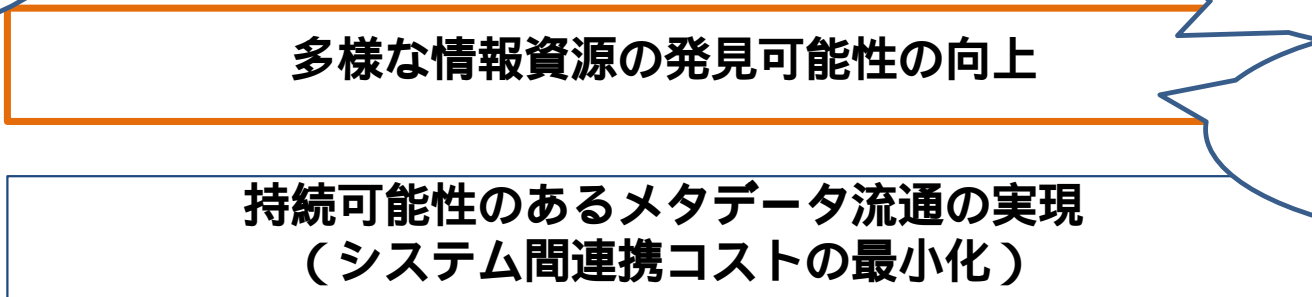
収集対象になる情報資源の拡充や詳細化、採用される技術仕様の変化、等

ガイドライン自体がコミュニケーションツールとして機能していく

同時に、誰が何を担うのかを可視化していく作業でもある

課題の解決 ↓ 目的の達成

メタデータはあくまで利用者と情報資源を媒介するもの。



資源の所在を明確にし、利用者と情報資源を結びつけるためのメタデータ流通を目指して。