

WebAPI及びデータフォーマット (DC-NDL)の概要

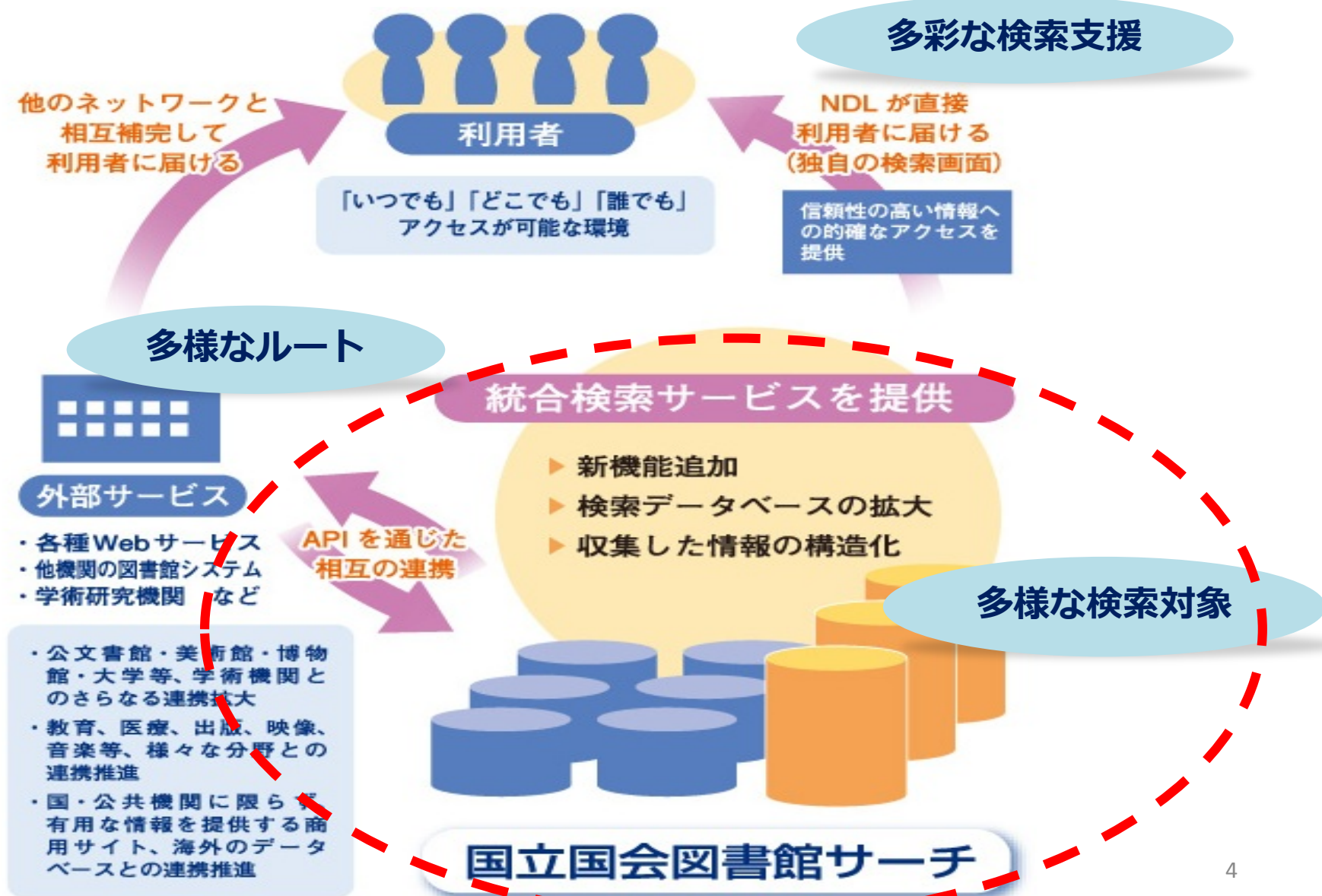
国立国会図書館電子情報部
電子情報サービス課

内容

- I. 国立国会図書館サーチとの連携方式
- II. 国立国会図書館サーチへ提供いただくメタデータ方式

1. 国立国会図書館サーチとの 連携方式

国立国会図書館サーチ (NDLサーチ) とは



外部提供インタフェース(API)とは

- あるプラットフォーム(OSやミドルウェア)向けのソフトウェアを開発する際に使用できる命令や関数の集合。また、それらを利用するためのプログラム上の手続きを定めた規約の集合。
- 個々の開発者は規約に従ってその機能を「呼び出す」だけで、その機能を容易に使うことができる。
- NDLサーチ自体もAPI提供をしており、他の図書館システムや各種の文献検索サービス、辞書サイト等で使われている。

※NDLサーチが提供する外部提供インタフェースについて
<http://iss.ndl.go.jp/information/api/>

NDLサーチへのデータ提供

1.メタデータを収集しての連携(ハーベスト)

連携先のシステムに、OAI-PMH形式かRSS形式のインタフェースを実装していただき、そのインタフェースを使って定期的にメタデータを収集する方法。

2.横断検索による連携

連携先のシステムに、SRU、SRW、OpenSearchのいずれかのインタフェースを実装していただき、そのインタフェースを使って、利用者からの検索の度に検索要求を投げ、結果を受け取る方法。

連携方式ごとの返戻メタデータ形式

連携方式	通信プロトコル	返戻メタデータ形式
メタデータ収集 (ハーベスト)	OAI-PMH	<ul style="list-style-type: none">DC-NDL(RDF)DC-NDL(Simple)SimpleDC(OAI-DC)
	RSS	<ul style="list-style-type: none">RSS1.0のデータ形式又はその派生形式RSS2.0のデータ形式又はその派生形式
横断検索	SRU/SRW	<ul style="list-style-type: none">DC-NDL(RDF)DC-NDL(Simple)SimpleDC(OAI-DC)
	OpenSearch	<ul style="list-style-type: none">RSS1.0のデータ形式又はその派生形式RSS2.0のデータ形式又はその派生形式

OAI-PMH/DC-NDLでの連携

• OAI-PMHとは

⇒データの自動収集によってメタデータを交換するためのプロトコル。リクエスト種別や取得したいメタデータの条件、フォーマット等を指定したOAI-PMHリクエスト(URL)を送信すると、当該システムがそのリクエストを受け取り、リクエストされた条件やフォーマット等に合致するメタデータをXML形式で返戻する。

⇒差分更新が可能。またRSSと異なり、deleteレコードを配信し、受け取った側でそれを用いてデータを削除することも可能。

⇒OAI-PMHの基本仕様は以下のページにあり(NIIによる日本語訳)。

<http://www.nii.ac.jp/irp/archive/translation/oai-pmh2.0/>

• 「OAI-PMH/DC-NDL」でのデータ受領

⇒NDLサーチとしては、可能な限り、OAI-PMHでデータを受領したい。メタデータ交換のために最も効率的なプロトコルであり、連携後の運用コストがかからないため。また、OAI-PMHを実装しておくことで、当館以外との連携の可能性も広がることになる。

※DC-NDLでの受領については後述。

OpenSearch、SRU/SRWでの連携

- **OpenSearchとは**

⇒ URL で検索リクエストを発行し、結果をRSS形式で取得する横断検索用のプロトコル。

⇒ 基本仕様は次のページにあり。<http://www.opensearch.org/Home>

- **SRU/SRWとは**

⇒ URL (SRU)あるいはXML (SRW)で検索リクエストを発行し、結果をXML形式で取得する横断検索用のプロトコル。

⇒ 基本仕様は次のページにあり。<http://www.loc.gov/standards/sru/>

- **OpenSearch、SRU/SRWでの連携**

⇒これによりNDLサーチと連携したとしても、NDLサーチならではの書誌同定やグループ化の対象にはならない。また、現在のNDLサーチの仕様上、デフォルトの検索対象とはならない(「すべての連携先を検索する」にチェックを付す必要あり)。

II. 国立国会図書館サーチへ提供いただくメタデータ方式

NDLサーチで採用している メタデータフォーマット

- NDLサーチでは、収集したメタデータを、全てDC-NDL(RDF)という形式にして格納し、出力している(正確には、内部DBでは、内部XMLという形式で保持しており、出力時にRDFに変換している)。
- NDLサーチでは、NDLの保有資源や外部連携先の情報資源を、紙・デジタルの媒体を問わず一括して検索することを可能とするため、国立国会図書館ダブリンコアメタデータ記述(DC-NDL)に基づくメタデータフォーマットを全資料に対して一元的に使用している。DC-NDLは、国際的なメタデータ標準である「Dublin Core(ダブリンコア)」をもとに日本語対応など独自に拡張したメタデータ記述語彙および記述規則の総称である。
- DC-NDLは、ダブリンコアと同様、セマンティックウェブ対応としてRDFのモデルを取り入れており、NDLサーチでもRDF/XML形式でメタデータを提供している。

Dublin Core (ダブリン・コア)

- 1995年頃からインターネット上の情報資源の発見を目的として開発が進められたメタデータ記述要素
- Dublin Coreの中核となる基本15要素を定義したのがDublin Core Metadata Element Set (DCMES)
- DCMESとは別の名前空間に、基本15要素を含む計55の記述要素(プロパティ)を定義したものが DCMI Metadata Terms

Dublin Core Metadata Element Set

基本記述要素	定義
Title	タイトル
Creator	作成者
Subject	キーワード
Description	内容記述
Publisher	公開者
Contributor	寄与者
Date	日付
Type	資源タイプ
Format	記録形式
Identifier	資源識別子
Source	出処
Language	言語
Relation	関係
Coverage	時空間範囲
Rights	権利管理

- 基本15要素を定義
- 国際標準 (ISO 15836)、国内標準 (JIS X 0836 ダブリンコアメタデータ基本記述要素集合)
- シンプルで記述の自由度が高い

DCMI Metadata Terms

- DCMESとは別の名前空間に、基本15要素を含む計55の記述要素(プロパティ)を定義
- 基本15要素を再定義し、DCMESの各記述要素に対する下位プロパティと位置づける
- 各プロパティに対して、定義域と値域を設定
- プロパティ以外にも、9個の語彙符号化スキーム、12個の構文符号化スキーム、22個のクラス、12個のDCMIタイプ語彙を定義

基本15要素	DCMI Metadata Terms	定義
title	alternative	タイトルの代替
description	tableOfContents	目次
	abstract	要約
date	created	作成日
	valid	有効期日もしくは期間
	available	利用可能日もしくは期間
	issued	正式発行日
	modified	更新日
	dateAccepted	(論文やジャーナル記事などの)受理日
	dateCopyrighted	著作権日
	dateSubmitted	(論文やジャーナル記事などの)提出日

⋮

DC-NDLとは

- DC-NDLは、国際的なメタデータ標準である「Dublin Core」をもとに、日本語対応など独自に拡張したメタデータ記述語彙および記述規則の総称
<http://www.ndl.go.jp/jp/aboutus/standards/meta.html>
- 語彙の意味定義は上記URLに書かれている「第一部 NDL Metadata Terms」を、語彙の用法についての規定は「第二部 Application Profile」を、RDFによる語彙の定義は「第三部 RDFスキーマ」をそれぞれ参照。

DC-NDLのメタデータフォーマット

- **DC-NDL (RDF)**
 - DC-NDLで記述したメタデータをRDF/XML形式で出力する際のフォーマット
 - RDFによる構造化表現を用いて、値とその読みなどをセットで表現
 - DC-NDL(Simple)に比べて情報量が多い
- **DC-NDL (Simple)**
 - DC-NDLで記述したメタデータをXML形式で出力する際のフォーマット
 - 構造化表現を用いず、文字列(リテラル)での記述を中心としたフラットな表現をとる
 - 利用頻度の高い要素のみを使用する

<参考> SimpleDC(oai_dc)

- Dublin Coreの中核となる基本15要素(Dublin Core Metadata Element Set)をXML形式で出力する際のフォーマット。情報量は最も少ない。¹⁷

第一部
NDL Metadata Terms

NDL独自語彙の定義

第二部
Application Profile

NDL独自語彙の用法
(値の形式・入力レベル等)

Dublin Core Metadata Terms 等
の語彙の用法
(値の形式・入力レベル等)

第三部 RDFスキーマ

NDL独自語彙の定義
(RDF形式)

DC-NDL

DC-NDL
(RDF)

DC-NDL
(Simple)

メタデータ
フォーマット

DC-NDL(RDF)の語彙

DC-NDL(RDF)では、以下の語彙をミックスして用いている。

<ダブリン・コアの語彙>

- dc
- dcterms

<当館独自の語彙>

- dcndl

<RDFの語彙>

- rdf
- rdfs
- foaf
- owl

DC-NDL(RDF)の構造

以下の3つの構造から成る
(XML宣言等は除く)。

- ①管理情報...書誌レコードのステータス(「作成中」、「校了済み」等)、リポジトリ番号(当該書誌データの提供元を表す番号)等
- ②書誌情報...情報資源に対する記述情報
- ③個体情報...アイテムレベルの情報(所蔵館)。

三層構造

管理情報

BibAdminResource

メタデータに関する情報
新規・更新等メタデータの作成状態等

書誌情報

BibResource

情報資源に対するいわゆる記述情報

個体情報

Item

各機関で所蔵する資料に関する情報
請求記号や欠号情報等

-<rdf:RDF>

```
-<dcndl:BibAdminResource rdf:about="http://iss.ndl.go.jp/books/R100000002-1000011254088-00">  
  <dcndl:catalogingStatus>C7</dcndl:catalogingStatus>  
  <dcndl:bibRecordCategory>R100000002</dcndl:bibRecordCategory>  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000002-1000011254088-00#material" />  
</dcndl:BibAdminResource>
```

管理情報
BibAdminResource

```
-<dcndl:BibResource rdf:about="http://iss.ndl.go.jp/books/R100000002-1000011254088-00#material">  
  <dcterms:identifier rdf:datatype="http://ndl.go.jp/dcndl/terms/JPN0">21969609</dcterms:identifier>  
  <dcterms:identifier rdf:datatype="http://ndl.go.jp/dcndl/terms/ISBN">978-4-87582-717-7</dcterms:identifier>  
  <rdfs:seeAlso rdf:resource="http://id.ndl.go.jp/jpno/21969609" />  
  <rdfs:seeAlso rdf:resource="http://iss.ndl.go.jp/isbn/9784875827177" />  
  <dcterms:title>国立国会図書館資料デジタル化の手引 2011年版</dcterms:title>  
-<dc:title>  
  -<rdf:Description>  
    <rdf:value>国立国会図書館資料デジタル化の手引</rdf:value>  
    <dcndl:transcription>コクリツ コッカイトショカン シリョウ デジタルカノ テビキ</dcndl:transcription>  
  </rdf:Description>  
</dc:title>  
-<dcndl:volume>  
  -<rdf:Description>  
    <rdf:value>2011年版</rdf:value>  
    <dcndl:transcription>2011ネンバン</dcndl:transcription>  
  </rdf:Description>  
</dcndl:volume>
```

書誌情報
BibResource

```
-<dcndl:BibResource rdf:about="http://iss.ndl.go.jp/books/R100000002-1000011254088-00#material">  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000002-1000011254088-00#item" />  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000001-1040988810-00#item" />  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000001-1041251690-00#item" />  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000001-1042544334-00#item" />  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000001-1042549039-00#item" />  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000001-1042610332-00#item" />  
  <dcndl:record rdf:resource="http://iss.ndl.go.jp/books/R100000001-1043238745-00#item" />  
</dcndl:BibResource>
```

```
-<dcndl:Item rdf:about="http://iss.ndl.go.jp/books/R100000002-1000011254088-00#item">  
  -<dcndl:holdingAgent>  
    -<foaf:Agent>  
      <foaf:name>国立国会図書館</foaf:name>  
      <dcterms:identifier rdf:datatype="http://ndl.go.jp/dcndl/terms/NDLLibCode">0000</dcterms:identifier>  
    </foaf:Agent>  
  </dcndl:holdingAgent>  
  <rdfs:seeAlso rdf:resource="http://id.ndl.go.jp/bib/000011254088" />  
  <dcterms:identifier rdf:datatype="http://ndl.go.jp/dcndl/terms/NDLBibID">000011254088</dcterms:identifier>  
  <dcndl:callNumber>UL755-J20</dcndl:callNumber>  
  <dcndl:localCallNumber>棚55b/当館刊行物</dcndl:localCallNumber>  
</dcndl:Item>
```

個体情報
Item

```
-<dcndl:Item rdf:about="http://iss.ndl.go.jp/books/R100000001-1040988810-00#item">  
  -<dcndl:holdingAgent>  
    -<foaf:Agent>  
      <foaf:name>福井県立図書館</foaf:name>  
      <dcterms:identifier rdf:datatype="http://ndl.go.jp/dcndl/terms/NDLLibCode">1811</dcterms:identifier>  
    </foaf:Agent>
```

DC-NDL(RDF)の特徴

- 「タイトル」と「タイトルよみ」、「著者」と「著者よみ」、「出版者」と「出版地」などの対応関係の維持等、構造化すべき項目は構造化して表現する。

例) 以下のように構造化・非構造化の表現を重複してもつ場合がある。

【タイトル】

dc:title : タイトルと読みをセットで記述する(構造化する)場合に使用

dcterms:title : タイトルのみを文字列で記述する(構造化しない)場合に使用

【作成者】

- dcterms:creator : 主に「著者標目」の表現に使用
 - 作成者名とその読みをセットで記述する(構造化する)場合に使用
 - URIを使う場合に使用
- dc:creator : 主に「責任表示」の表現に使用
 - 文字列(リテラル)で記述する(構造化しない)場合に使用

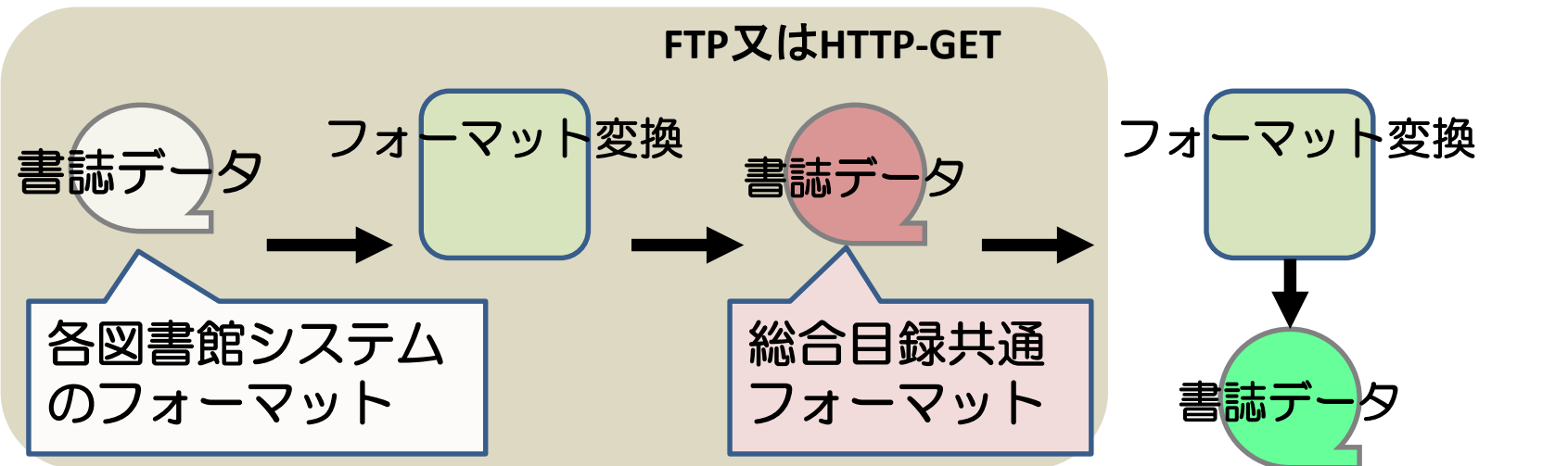
DC-NDL(RDF)の特徴

- 標準番号のうち代表的なもの（ISBN、ISSN等）、全国書誌番号（JP番号）は、URI形式とリテラルの両方で保持している。
- NDLサーチでは、書誌同定処理（前項参照）により、ひとつの書誌データに複数の個体データが紐づく場合がある。

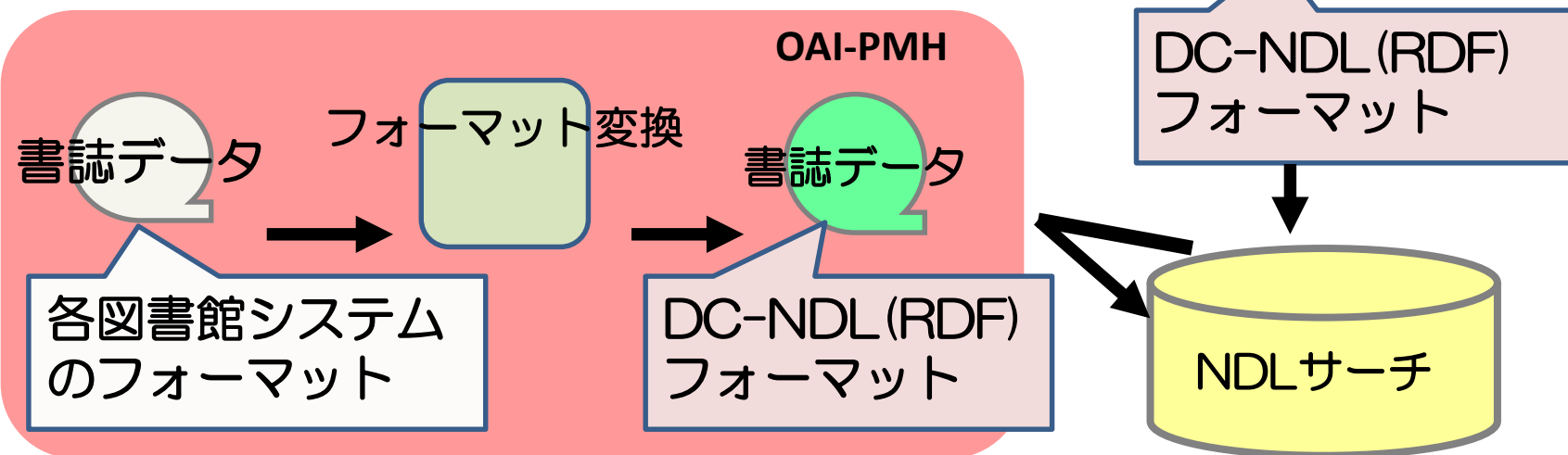
※複数の個体データを持っており、それを出力可能であることは、書誌データを集約しているNDLサーチならではの特徴。

データ提供の流れ

現状の流れ



新たな流れ



参考

- ・連携を希望される機関の方へ

(『WebAPIによるシステム連携ガイドライン』の掲載ページ)

<http://iss.ndl.go.jp/information/renkei/>

- ・メタデータフォーマット仕様について

(NDLサーチが提供するメタデータの形式のページ)

<http://iss.ndl.go.jp/information/metadata/>

- ・API仕様書について

(NDLサーチが提供する外部提供インタフェース(API)のページ)

<http://iss.ndl.go.jp/information/api/>

- ・国立国会図書館ダブリンコアメタデータ記述(DC-NDL)のページ

<http://www.ndl.go.jp/jp/aboutus/standards/meta.html>

- ・国立国会図書館ダブリンコアメタデータ記述に関する解説

http://www.ndl.go.jp/jp/aboutus/standards/meta/about_dcndl.html

- ・国立国会図書館ダブリンコアメタデータ記述に関する事例集

http://www.ndl.go.jp/jp/aboutus/standards/meta/dcndl_examples.html