

2020年2月20日

デジタルアーカイブ学会JPS-SIG



ジャパンサーチの概要と検索の実際

国立国会図書館 電子情報部 電子情報企画課

はじめに：ジャパンサーチとは

- **さまざまな分野のデジタルアーカイブと連携し、我が国が保有する多様なコンテンツのメタデータ***をまとめて検索できる**国の分野横断統合ポータル**

*メタデータとは：

コンテンツの内容や所在等について記述したデータ。図書館の書誌データ、博物館・美術館の収蔵品の目録データなど。

目指す効果

1. コンテンツの所在等の明確化
2. 連携機関へのアクセス促進
3. データの利活用の促進
4. 連携機関への支援
5. 新規ビジネス・サービスの創出等

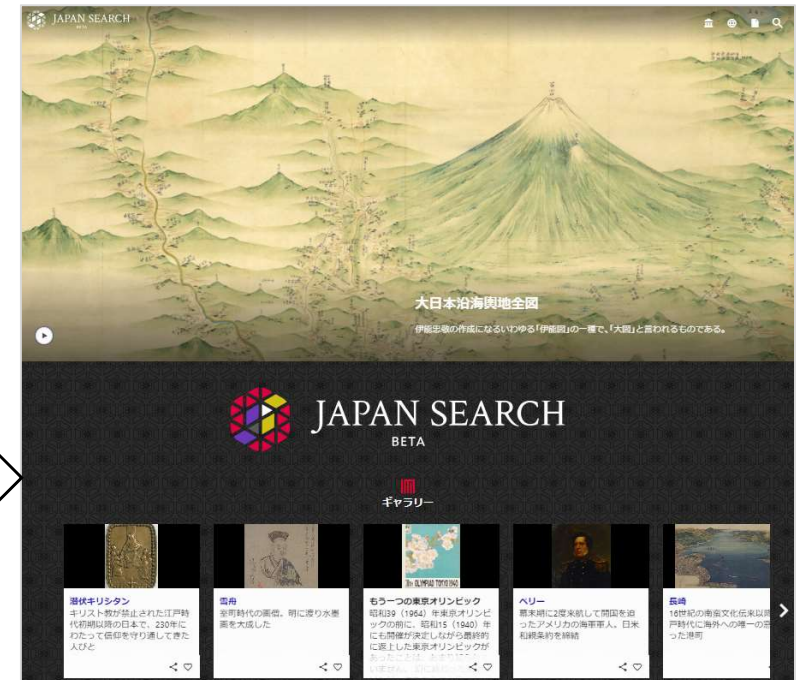
- **政府の「知的財産推進計画」等に掲げられている国の取組**

運用主体：デジタルアーカイブジャパン推進委員会・
実務者検討委員会
(事務局：内閣府知的財産戦略推進事務局)

システムの運用担当：国立国会図書館

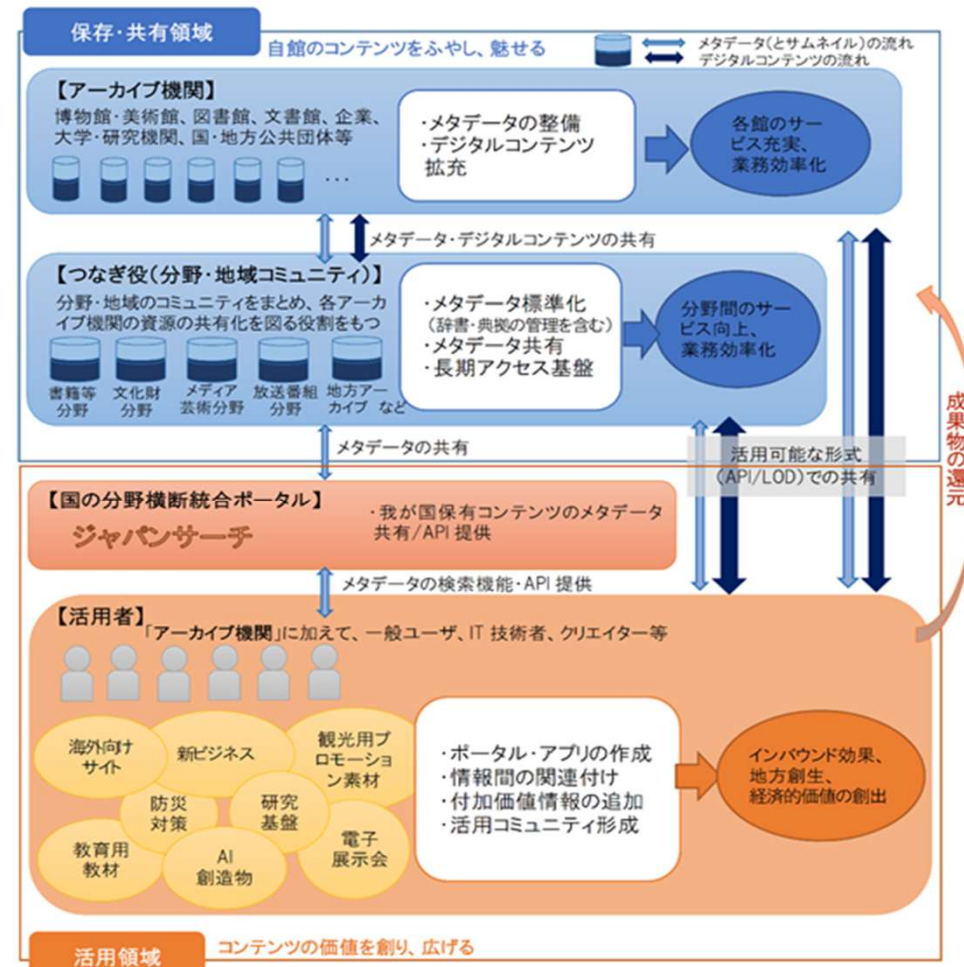
- **2019年2月に試験版を公開**
2020年夏までの正式版公開に向けて、
フィードバックを受付中

<https://jpsearch.go.jp/>



連携方針

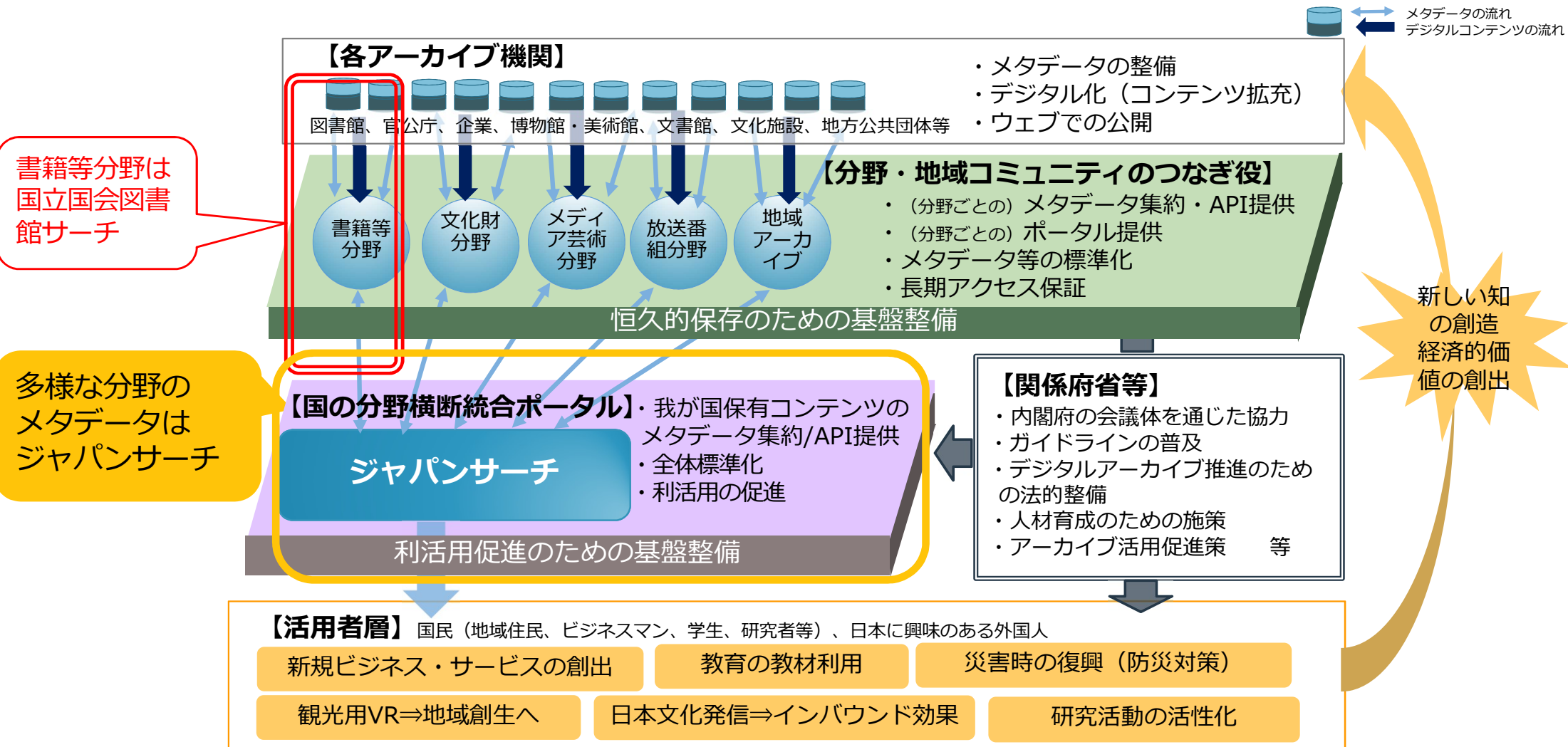
- 分野・地域の「つなぎ役」を通じた連携を原則とする
- ただし、「つなぎ役」が明確でない分野・地域では、以下の条件に当てはまるアーカイブ機関との直接連携を検討
 - 国の機関であり、当該分野におけるコンテンツを幅広くカバーしているアーカイブ機関
 - 公益に資する目的のため、当該分野におけるコンテンツを幅広くカバーしているアーカイブ機関
 - 唯一性・独自性の高いコンテンツ群を塊として扱う分野・地域を代表するアーカイブ機関
 - その他（実務者検討委員会において適当と認められるアーカイブ機関）



(出典) 「第二次中間取りまとめ」(実務者検討委員会, 平成31年4月) p. 33

https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/jitumusya/2018/torimatome2.pdf

ジャパンサーチ構築に向けて：役割の整理



連携状況（2020年2月14日現在）

16連携機関 64データベース メタデータ約1,970万件

分野	連携機関	データベース名
書籍等	国立国会図書館	「国立国会図書館サーチ」から、3件のデータベース
公文書	国立公文書館	「国立公文書館デジタルアーカイブ」
文化財	文化庁	「文化遺産オンライン」から、国指定文化財等データベース
	国立文化財機構	「ColBase 国立博物館所蔵品統合検索システム」
美術	国立美術館	「国立美術館所蔵作品総合目録検索システム」 「アートコモンズ」
	日本写真保存センター	「写真原板データベース」
メディア芸術	映像産業振興機構	「Japan Content Catalog」から、2件のデータベース
舞台芸術	早稲田大学坪内博士記念演劇博物館	「演劇情報総合データベース」から、2件のデータベース
自然史・理工学	国立科学博物館	「サイエンスミュージアムネット S-Net」
		「魚類写真資料データベース」
人文学	人間文化研究機構	「人間文化研究機構統合検索システム nihuINT」から、42件のデータベース
	立命館大学アート・リサーチセンター	「ARC浮世絵ポータルデータベース」「ARC古典籍ポータルデータベース」
	公益財団法人 渋沢栄一記念財団	「デジタル版「実験論語処世談」/渋沢栄一」
学術資産	東京大学	「東京大学学術資産等アーカイブズポータル」
放送番組	放送番組センター	「放送ライブラリー公開番組データベース（ドラマ）」
	日本放送協会	「動画で見るニッポンみちしる」
公共データ	総務省・内閣官房IT総合戦略室	「データカタログサイト」

国立国会図書館サーチ経由での連携状況（2020年2月13日現在）

9連携機関 10データベース メタデータ約868万件

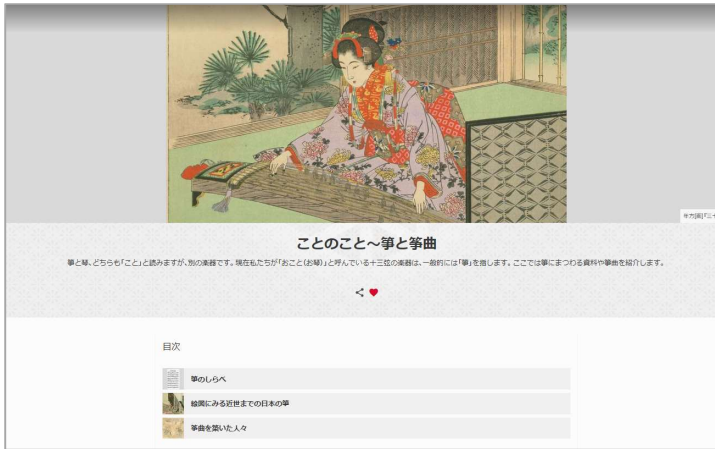
データ提供機関	データベース名
国立国会図書館	全国書誌
	国立国会図書館デジタルコレクション
TRC (ADEAC)	デジタルアーカイブシステムADEAC
青森県立図書館	青森県立図書館デジタルアーカイブ
千葉県立中央図書館	千葉県デジタルアーカイブ
奈良県立図書館	まほろばデジタルライブラリー
慶應義塾大学メディアセンター	慶應義塾大学メディアセンターデジタルコレクション
岩手県立図書館	イーハトーブ岩手電子図書館
宮城県図書館	叡智の杜Web
札幌市中央図書館	札幌市中央図書館デジタルライブラリー

連携に当たっての技術的な要件・手続等は、
国立国会図書館サーチの「連携を希望される機関の方へ」を参照
<http://iss.ndl.go.jp/information/renkei/>

ジャパンサーチ試験版の機能

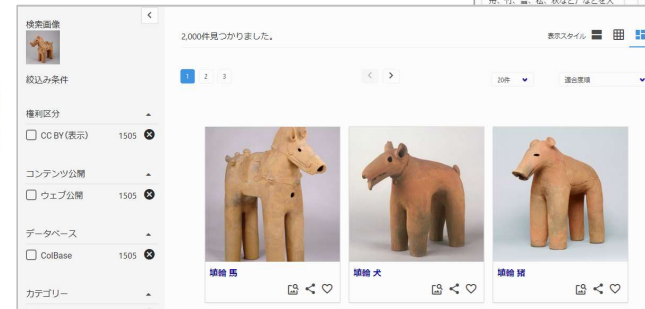
クリックするだけで楽しめる 多様なコンテンツの紹介

- 様々なテーマ*に関する「ギャラリー」を提供
(*テーマ例：長崎、富士山、刀剣、和食、北斎、伊能忠敬 etc.)



分野の特性を活かした検索機能

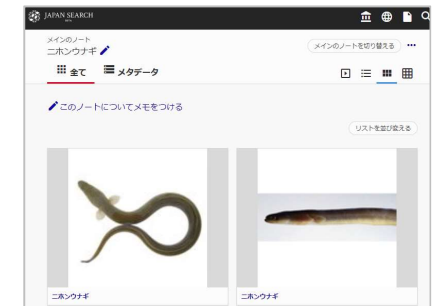
横断検索



テーマ別検索



画像検索



マイノート画面

利活用促進の基盤となる利活用機能

- ジャパンサーチが集約したメタデータをAPI（外部提供インタフェース）で提供
- お気に入りのメタデータ・検索結果・ギャラリーをマイノートに登録可能
- マイノートを外部のWebサイトに貼付できるウェブパーツ機能

“国の分野横断統合ポータルサイト”「検索機能」

横断検索

NDLサーチと同じ方式

- 共通項目による横断的な検索、絞込みも可能
(でも全てを拾いきれない? その場合は

テーマ別検索

NDLサーチと異なるところ

- 各分野の特性を生かして、連携機関の元の項目を最大限に活用した検索が可能
- 「ギャラリー」にも活用

- 検索結果の詳細表示画面では、連携元の生データの情報もそのまま表示

- 分野横断的な共通項目を予めマッピング

マッピング自動機能
(メタデータアナライザー)

手動補正

NDLサーチと異なるところ

- ユーザ自ら、特定のDB・項目間のマッピングをカスタマイズ
- 検索式の共有
- テーマ別検索を共有するためのページも用意

個別に検索式を作成する機能
(UIを別途用意)

連携機関

- 連携機関の生データをそのまま保持

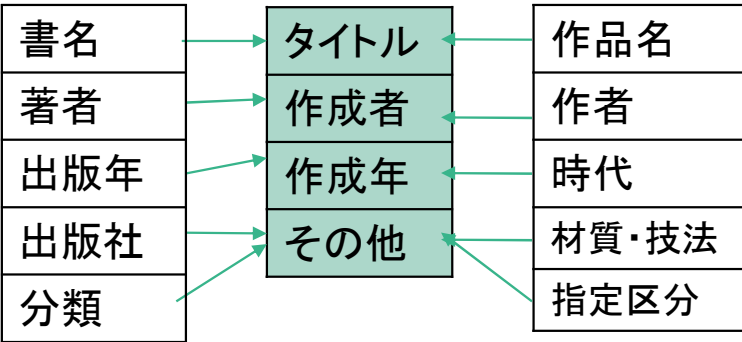
システム連携 (ファイル連携 or OAI-PMH) or 管理画面上でのアップロード

共通項目とマッピングについて

ジャパンサーチでのマッピングの前提

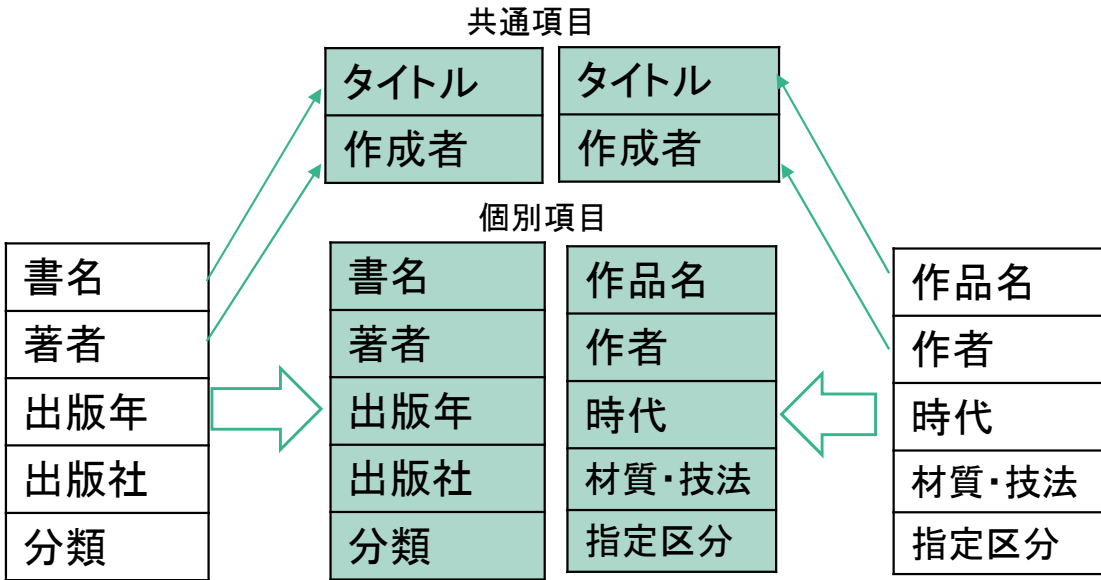
- 連携先追加の容易性を確保するため、マッピングのコストを最小限にする。
- マッピングした共通項目を連携機関の生データに追加して保持する。
- マッピングはメタデータアナライザーで自動推定を行う。

従来 방식



※全ての項目をマッピング

ジャパンサーチの方式



※個別項目に共通項目を追加した形でデータをもつ

「ギャラリー」と「マイノート」

・ ギャラリー

検索機能を利用しない場合でも見て楽しめるページ

- ・ エディタで作成可能（現在は連携機能のみ）
- ・ 画像はIIIFに変換して掲載
- ・ SNSで簡単に拡散可能な仕組みも用意



ギャラリー「長崎」

<https://jpsearch.go.jp/curation/ndl-j6b49EaX1nFwB9>

・ マイノート

お気に入りのメタデータ/ギャラリー/検索結果を登録

- ・ ブラウザのキャッシュで保存
- ・ 注記の追加、データのエクスポート等が可能
- ・ 誰でも利用可能（ユーザ登録不要）



マイノートの例

マイノート機能を使ったキュレーション体験イベント 「ジャパンサーチ×エディタソンー新しいキュレーションを模索する」

<https://lab.ndl.go.jp/cms/>

- 2019年11月30日(土)13:00～17:00 国立国会図書館東京本館で開催
- 講師2名から講義:キュレーションの手法、デジタルデータを用いたキュレーションの可能性等について
参加者は、5チームに分かれ、マイノート機能を用いて約2時間のキュレーション作業を体験
- 参加者:16名(会社員、学生、研究者、図書館職員等)
- アンケート結果:
 - 満足度:100%(「満足」「どちらかと言えば満足」の合計)
 - 主な意見:
 - ・「マイノート機能は自分の興味関心をまとめられるよいツールだと感じた。」
 - ・「ジャパンサーチで利用可能なサムネイル画像をもっと増やしてほしい。」
 - ・「次回もジャパンサーチ×エディタソンの開催を期待する。」
 - ・「ギャラリーを使ったコンテストを開催すると面白い。」



キュレーション作業の様子



成果物の発表・講師講評

参加者作品一覧

干支の動物たち(1班)
日本の初期の物理学(2班)
物理学...物理?(2班)
ラクガキ素材～自然科学編(2班)
地図から消えた因島市～「尾道市因島」への軌跡～(3班)
教育双六(3班)
旅するイメージ「下る男と上る男」(3班)
箱根駅伝(4班)
描かれた妖怪たち(5班)
書跡の題字と肖像(講師 田良島哲氏)
GORAKU～♪(講師 大井将生氏)
9ga tukutoki(講師 大井将生氏)

エディタソン参加者の作品「干支の動物たち」




えとのどうぶつたち
干支の動物たち
年賀状にも使える十二支セレクション

このページは、2019年11月20日に限り国立国会図書館が主催したイベント「シヤパンサーチ&エディタソン 新しいエディタソンを体験する」の参加者が作成したものである。


十二支とは、本来は、木星が十二年で天を一周することから、中国の天文学で毎年度における木星の位置を示すために占を十二分し、それぞれを動物の名前に当てた呼称である。日本では、ね(子)・うし(丑)・う(寅)・たつ(辰)・み(巳)・うま(午)・ひつじ(未)・さる(申)・とり(酉)・いぬ(戌)・い(亥)とよむ。十二年ごとに一進する年周りを表すほか、編制や方角などを示すのに用いられる。

十二支の動物たち


ねずみ (子)



うし (丑)



いのしし (亥)



十二支に関する資料

十二経巻物
十二支の動物たちが活躍するお伽話を描いた絵巻。これは、十五夜に、十二支の動物たちが歌合(うたあわせ)和歌を読み、その優秀を競う遊(あそび)をしている場面。

十二経絵巻写
十二経絵巻を墨で写したものと思われる。京都国立博物館蔵。重要文化財。

十二支図三所物
縁起図集(4冊)
東京国立博物館 / CoBase

十二支木彫模付
縁起図(7巻)
東京国立博物館 / CoBase

四神八卦十二支図
京都国立博物館 / CoBase

もっと見る

参考文献

1. 日本国語大辞典(データベース「Japanshoni.kijg.jp」)

責任表示
国立国会図書館 作成 2019年11月20日 「シヤパンサーチ&エディタソン」 1巻 (日岡暁、荒木千穂、深瀬西穂香、原崎良平) 巻中
最終更新日
2020.03.03
二次利用について
CC BY (表示) ただし、画像は提供元の権利表示による

<https://jpsearch.go.jp/curation/ndl-6M2QogqA47Q>

ジャパンサーチにおける二次利用条件表示

● メタデータ・サムネイル画像の二次利用条件表示

- 登録データベース単位で設定
 - 自由記述（外部サイトへのリンク可）

→ データベース紹介ページで表示

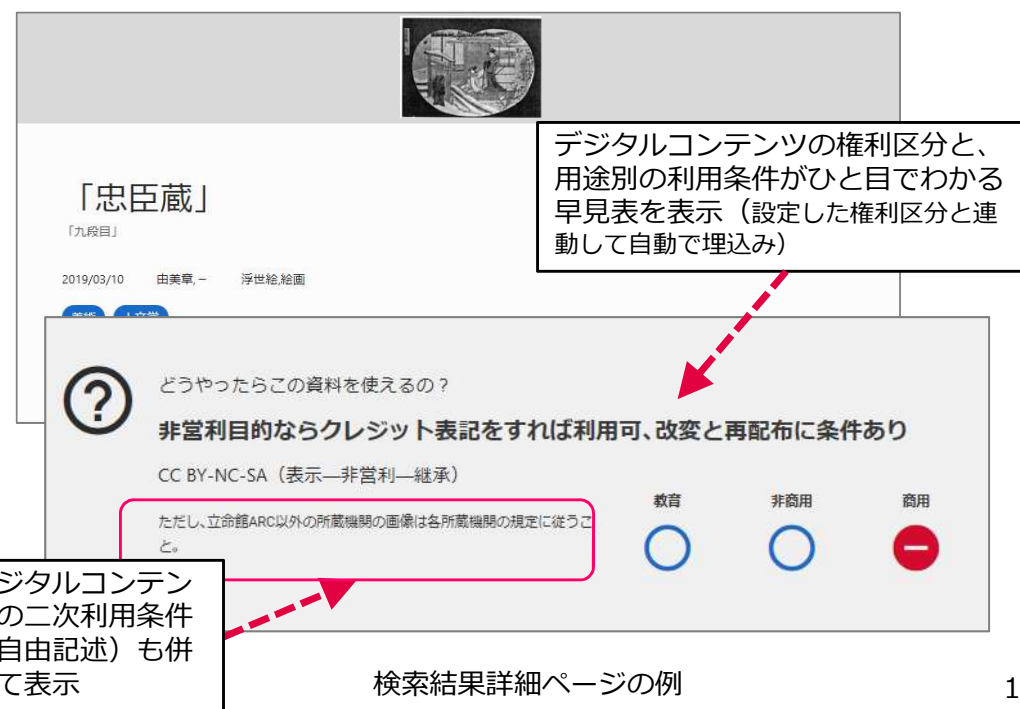


データベース紹介ページの例（ARC浮世絵ポータルデータベース）

● デジタルコンテンツの二次利用条件表示

- 登録データベース単位で設定
 - 自由記述（外部サイトへのリンク可）
 - 15種類の“権利区分”から1つ選択※（検索結果の絞込み用）※コンテンツ単位で二次利用条件が異なる場合は、コンテンツ単位でも設定可能

→ データベース紹介ページ、検索結果詳細ページで表示



検索結果詳細ページの例

実務者検討委員会
「デジタルアーカイブ
における望ましい二次
利用条件表示の在り方
について(2019年版)」
に基づき設定

デジタルコンテンツの権利区分と早見表対応表

		教育利用	非商用利用	商用利用
ツール	CC0	○	○	○
	PDM (パブリックドメインマーク)	○	○	○
クリエイティブ ライセンス commons	CC BY (表示)	○	○	○
	CC BY-SA (表示-継承)	○	○	○
	CC BY-ND (表示-改変禁止)	△	△	△
	CC BY-NC (表示-非営利)	○	○	×
	CC BY-NC-SA (表示-非営利-継承)	○	○	×
	CC BY-NC-ND (表示-非営利-改変禁止)	△	△	×
Rights Statements	著作権あり	×	×	×
	著作権あり – 教育目的の利用可	○	△	×
	著作権なし – 契約による制限あり	△	△	△
	著作権なし – 他の法的制限あり	△	△	△
	著作権未評価	×	×	×
その他	著作権未決定 – 裁定制度利用著作物	×	×	×
	その他	—	—	—



早見表の例

【参照】
 「デジタルアーカイブにおける望ましい二次利用条件表示の在り方について (2019 年半)」 (『第二次中間取りまとめ』補足資料)
[https://www.kantei.go.jp/jp/singi/titeki2/digital archive_suisiniinkai/jitumusya/2018/nijiriyu2019.pdf](https://www.kantei.go.jp/jp/singi/titeki2/digital_archive_suisiniinkai/jitumusya/2018/nijiriyu2019.pdf)

○ = 利用可、△ = 条件付き利用可、× 無断利用不可 (要許諾)

JPS利活用スキーマとAPI

ジャパンサーチにおけるデータ変遷 (イメージ)

連携機関

(オリジナル) ソースデータ

名称 : 刀
サイズ : 68.1弱
和暦 : 長禄2
分野 : 工芸
作者名 : 長光
所蔵 : X博物館



登録・共通項目ラベル付与

共通項目ラベルの付与

共通項目ラベル名

名称/タイトル

名称 : 刀

サイズ : 68.1弱

時間/時代

和暦 : 長禄2

分野 : 工芸

人物/団体

作者名 : 長光

提供者

所蔵 : X博物館

DB単位の情報

データベース : Y DB

タイプ : 文化財

マッピング

詳細モデル化

ラベル : 刀剣
(<https://id.ndl.go.jp/auth/ndlsh/00573180>)

コンテンツ記述 : 68.1弱

時間 : 1459

主題 : 工芸

作者 : 長光

提供者 : X博物館

アグリゲーター : Y DB

タイプ : 文化財

名称 : 刀

サイズ : 68.1弱

和暦 : 長禄2

分野 : 工芸

作者名 : 長光

所蔵 : X博物館

ソースデータ

- 共通項目ラベルは自動で付与後、連携機関で確認・修正が可能
- 最小限度の項目のみ実施

魅力的表示



検索機能



- 一部項目を抽出・コピー/変換 (正規化)
- 値はできるだけURI化
- ソースデータも保持

検索機能の
精緻・高度化



「利活用フォーマット」
で出力



「JPS利活用スキーマ」の概要

目的

- ✓ ジャパンサーチに登録された多種多様なコンテンツのメタデータを共通の形式に変換して提供することにより、より精緻な検索を可能にすること。また、分野横断的な利活用（付加価値をもたらす二次利用）を促進すること。

※検討に当たっては、Europeana等の海外事例も踏まえつつ、シンプルで使いやすいものであると同時に必要十分な情報を提供するためにはどのようなモデルがよいかといった視点に留意した

データモデル

- ✓ 提供元からのソースデータを来歴情報明記の上でそのまま保持
- ✓ ソースデータのうち、利用者の4つのタスク（発見・識別・選択・取得）に特に有益な項目を共通の形式に変換した共通情報を新たに生成し保持

提供

- ✓ 必要十分な情報を、平易な構造で、汎用性のある形式でファイル出力
- ✓ 詳細画面で共通情報とソースデータの両方を表示

詳細は「利活用スキーマ概説」を参照 <https://jpsearch.go.jp/api/introduction/>

JPS利活用スキーマ 共通情報の基本項目

「いつ」、「どこで」、「だれが」、「何を」を基本に項目設定。

基本項目	内容
タイプ	コンテンツの基本区分（書籍、文化財など大きく情報を区分する枠）
名称	タイトル、別名、読みなど検索対象とする名前
寄与（者）関係	コンテンツに寄与した人／組織（作者、発行者、出演者など）
場所関係	場所に関する情報（発行地、制作地など）
時間関係	時間に関する情報（制作年、対象時期など）
主題・区分	主題・分類／各分野のキーワード的共通認識のある区分（国宝、ドキュメンタリーなど）
識別子	コンテンツを特定するための識別子（ISBNなど）
言語	コンテンツの記述言語
サムネイル画像	コンテンツの特徴を確認するための画像（提供元とは別にサムネイルを保持する場合）
記述	コンテンツの物理的特徴・素材等の記述、個別項目に収録できない情報
上位コンテンツ	当該コンテンツがその一部である上位コンテンツ（公文書などの資料階層）

基本項目	内容
提供情報	コンテンツにアクセスするための情報。下記サブ項目からなる
提供者	コンテンツ（に関する情報）の提供者を識別するURI。保管者が別であればURIも。
リンク	コンテンツの紹介ページやアクセス情報が記載されたページのURL
オブジェクト	コンテンツのデジタル画像や音声・動画のURI
権利情報	コンテンツの利用に関するライセンス及び権利に関する情報（ライセンスのURIなど）
個別識別子	提供者・所有者が管理するアイテムとしての識別子（請求記号など）
ソース情報	ソースデータ（ジャパンサーチが連携フォーマットで受け取ったデータ）とその提供者に関する情報。
提供者	ソースデータの提供者（つなぎ役）
データ	プラットフォームが保持・提供するソースデータ
リンク	つなぎ役におけるソースデータの掲載ページのURL
更新日	収集元データの更新日又はつなぎ役による収集日

APIの提供

- 共通項目のデータの他に、連携機関から収集したそのままの形のメタデータもAPIで提供
- メタデータアナライザの結果（項目特性に関する情報）もAPIで提供
- 検索API以外に、活用者がメタデータの大量取得を可能とするAPIも用意

The image shows two side-by-side screenshots. The left screenshot is a web page from ColBase for the item '花鳥' (Kachou). It includes a 'デジタルコンテンツの二次利用条件' (Secondary use conditions) section with 'CC BY' license and '教育利用' (Educational use) selected. A red dashed arrow points from the 'メタデータの取得' (Metadata acquisition) button in the bottom left to the right screenshot. The right screenshot is the 'Snorql for Japan Search' API response for the item. It displays a table of metadata for the URL <https://jpsearch.go.jp/data/cobas-37911>. The table includes fields like 'rdf:type', 'rdfs:label', 'schema:name', 'schema:creator', 'schema:spatial', 'schema:temporal', 'schema:dateCreated', 'schema:description', and 'schema:image'. A small image of the artwork is visible in the top right corner of the API response.

Snorql for Japan Search	
Description of <https://jpsearch.go.jp/data/cobas-37911>	
rdf:type	type:絵画 (絵画)
rdfs:label	"花鳥"
schema:name	"Kachou"@en-jp "かちょう"@ja-kana "花鳥"@ja
schema:creator	chname:服部雪斎 (服部雪斎)
schema:spatial	place:東京 (東京都)
schema:temporal	time:1871 (1871年)
schema:dateCreated	"1871"
schema:description	"ID: 37911" "分類: 絵画 [Classification]" "品質形状: 絹本着色 [Item shape]" "員数: 1幅 [Number of parts]" "法量: 33.3×68.4 [Size]" "解説: 雪斎は現在展示中の『目八譜』に貝を描いた絵師。江戸時代、幕府医学館に関わる仕事をし、明治以降は田中芳男などと関わり博物局で描き続けた。追真的な画風を売り物にしていたが、本作品のような掛け軸の花鳥図はあまり確認されておらず、めずらしい。(いがくかん たなかよしお) [Commentary]"
schema:image	<https://colbase.nich.go.jp/uploads/colle ... 8277e89f2d969c22.jpg>

(例) ColBaseより「花鳥」（東京国立博物館所蔵） <https://jpsearch.go.jp/item/cobas-37911>

APIの活用例

- Europeanaとジャパンサーチをキーワード「葛飾北斎」で横断検索

SPARQL query (例)

```

PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX edm: <http://www.europeana.eu/schemas/edm/>
PREFIX ore: <http://www.openarchives.org/ore/terms/>
SELECT ?uri ?label ?image WHERE {
  BIND(chname:葛飾北斎 as ?cname)
  {
    ?cname schema:name ?ename.
  }
  FILTER(lang(?ename)="en")
  BIND(replace(?ename, " ", " ")
  as ?name)
  ?loduri owl:sameAs ?cname ;
  rdfs:isDefinedBy <http://dbpedia.org/>.
  <http://sparql.europeana.eu/> {
    {?uri dc:creator ?name .}
    UNION
    {?proxy
    dc:creator ?loduri ; ore:proxyFor ?cho .
    ?uri ore:proxyFor ?cho ; edm:europeanaProxy "false". }
    ?uri dc:title ?label
    OPTIONAL{?uri
    ore:proxyIn [edm:isShownBy ?image ]}
    } UNION {
    ?uri rdfs:label ?label;
    schema:creator/owl:sameAs? ?cname .
    OPTIONAL{?uri
    schema:image ?image}
    }
}

```

<http://data.europeana.eu/proxy/provi ... ce_3000059083755>	"Ehon Suikoden Suikoden yūshi no ezukushi"	-
<http://data.europeana.eu/proxy/provi ... ce_3000059090745>	"Hokusai gafu"	-
<http://data.europeana.eu/proxy/provi ... ce_3000045463718>	"Denshin gaky"	-
<http://data.europeana.eu/proxy/provi ... ce_3000045464019>	"Ehon teikin ōrai. Zen"	-
<http://data.europeana.eu/proxy/provi ... ce_3000135526584>	"Hokusai gashiki"	-
<http://data.europeana.eu/proxy/provi ... ce_3000135526673>	"Hokusai soga"	-
<http://data.europeana.eu/proxy/provi ... RC_KONB11_1_4656>	"Nagasakiya"	-
<https://jpsearch.go.jp/data/cobas-38421>	"獅子図屏風"	
<https://jpsearch.go.jp/data/cobas-49254>	"千繪の海・絹川はちふせ"	
<https://jpsearch.go.jp/data/cobas-49262>	"詩歌寫真鏡・在原業平"	
<https://jpsearch.go.jp/data/cobas-49264>	"詩歌寫真鏡・安倍の仲磨"	
<https://jpsearch.go.jp/data/cobas-49280>	"鶺鴒、白粉花"	
<https://jpsearch.go.jp/data/cobas-51308>	"百人一首うはか恵とき・源宗于朝臣"	
<https://jpsearch.go.jp/data/cobas-51327>	"諸國瀧廻り・木曾路ノ奥阿彌陀ヶ瀧"	
<https://jpsearch.go.jp/data/cobas-51332>	"百合"	

神崎正英「外部エンドポイントとの統合クエリ」 <https://www.kanzaki.com/works/ld/jpsearch/fedquery>

JPS利活用スキーマの詳細については...

- ジャパンサーチ「開発者向け情報」 <https://jpsearch.go.jp/api>

- 神崎正英氏ウェブページ

The Web KANZAKI「ジャパンサーチ非公式サポートページ for 利活用スキーマ」

<https://www.kanzaki.com/works/ld/jpsearch/>

The Web KANZAKI music & knowledge sharing Help updated 2020-01-09

ジャパンサーチ 非公式 サポートページ for 利活用スキーマ

ja en

ジャパンサーチの (主として利活用スキーマの) 応用や活用に関する情報を順次掲載していきます。

ジャパンサーチ (のRDF) を使う

- はじめてのジャパンサーチ利活用スキーマ

いろいろな利用のための材料やヒント

- ジャパンサーチ名鑑 (正規化名索引) - ジャパンサーチ名前正規化辞書から生成した索引
- Japan Search normalized name index - 正規化辞書から生成した英語名索引
- 都道府県別に見た特徴的タイプ - ジャパンサーチ収録アイテムのタイプを地域別に分析
- 日本関係外国語図書の書誌情報: JPSモデル版 - 利活用スキーマをオープンデータに応用する例として。マッピング説明あり
- メトロポリタン美術館RDF: JPSモデル版 - オープンアクセスCSVから日本関連約2万3千点を選び、利活用スキーマでRDF化
- クリーブランド美術館RDF: JPSモデル版 - オープンアクセスAPIから日本美術部門の作品データを取得してRDF化
- CSVからジャパンサーチ型RDFを生成する - 利活用スキーマが各種データに適用できることを試してみましょう
- 相模絵の場合: ジャパンサーチとWeb Annotation - アイテム記述に独自のメモを加えて活用。JPSノートからの変換も
- ギャラリーとWeb Annotation - JPSノートと同じデータ構造のギャラリーを注釈記述に変換。独自コレクションの出発点に

クエリの応用

- MLAエンドポイントとの統合クエリ - Europeana, DPLA, 英国博物館, BnF, Wikidataなどの横断検索
- 絵入源氏物語を電書リーダーで読む - 各行メタデータから動的に電子書籍ファイルを生成し、青空文庫と合わせて表示
- 二十一代集を電書リーダーで読む - 歌集、歌人単位で検索した勅撰和歌メタデータから動的に電子書籍ファイルを生成
- 万葉集を電書リーダーで読む - Wikisourceの万葉集データから生成したジャパンサーチモデルRDF。電書リーダーでも表示
- ジャパンサーチに見るWikidataの作品 - Wikidataに登録のある主要作品で、ジャパンサーチにも含まれるものを選んで作者別に紹介
- IIFマニフェスト検索 - ジャパンサーチに登録されているIIFマニフェストを検索し、IIFコレクションを生成

クエリ例はSPARQLエンドポイント解説の他、Snorql for Japan Searchの入力エリア下にもポップアップリストがあります。さらにその下に用意した「どこ」「だれ」などの簡単入力でも基本検索クエリを実行できるEasySPARQLを試し、それを元に加工するのも良いと思います。



ジャパンサーチ試験版 デモ

JAPAN SEARCH BETA 検索キーワードを入力

久隔清音馳恋無傳
安和且慰下情
大阿闍梨所示五八詳卷
一百廿神仏并方圓篇并
茅名今奉和得
仁岳志伏乞令聞
其所撰岳義並其
若視其和詩志
著筆之文難段後代惟示
妾曲必造初時奉上
謹附貞聰仁子奉状和兩
私仁四年十月廿五日弟
雄範閑系 法前

最澄筆 尺牘(久隔帖)

最澄(伝教大師、767~822)が高雄山寺(今の神護寺)の空海のもとにいた愛弟子の泰範に宛てた書状。現存する唯一の最澄自筆書状で、「久隔漏音(久しく御無沙汰を)」と書き出しているところから、「久隔帖」と呼ばれて名高い。

JAPAN SEARCH BETA

ぜひ、一緒に
お手元のスマホで
ご覧ください



<https://jpsearch.go.jp/>

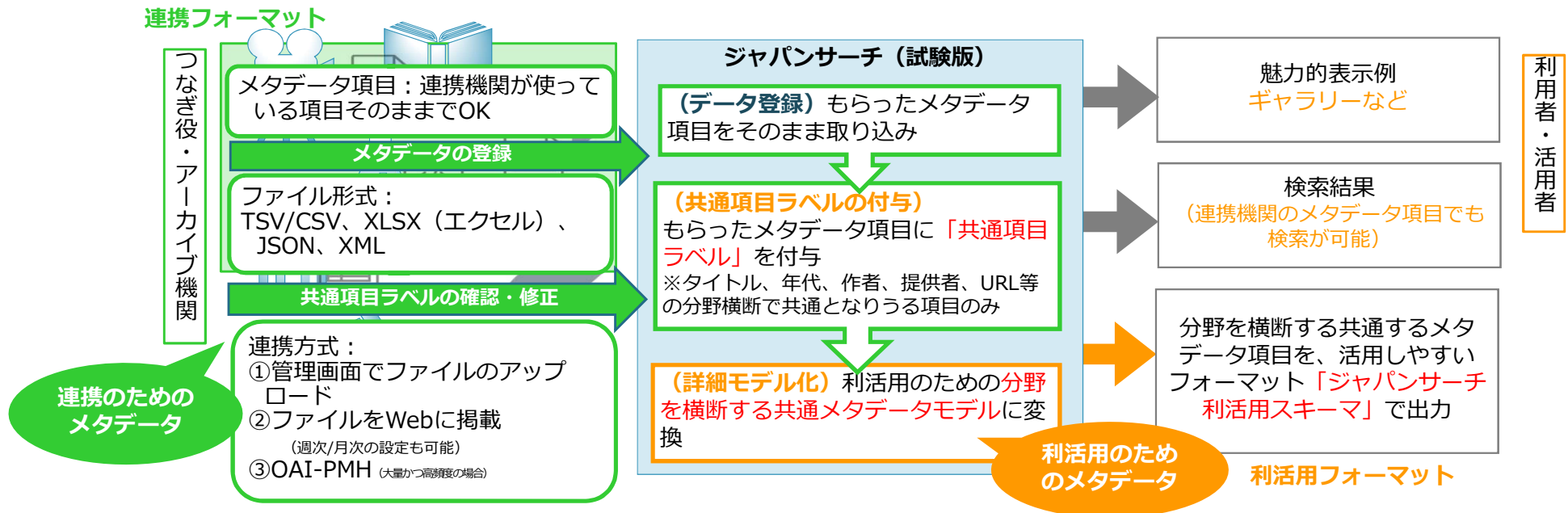
ジャパンサーチのメタデータ連携の仕組み

実際の連携手順

1. 連携を希望される場合は、「[お問合せフォーム](#)」からご連絡ください。
2. 上述の連携方針に照らして、実務者検討委員会において連携の可否を決定します。
3. 実務者検討委員会で承認後、国立国会図書館とデータの取扱いに係る文書の取り交わしを行います。
4. 国立国会図書館にサンプルデータを送付いただきます（メタデータの登録作業前に、システム上、データ登録に支障がないかを確認します）。
5. 国立国会図書館から、組織登録・メタデータ登録の具体的な手順をご案内しますので、それに沿って登録作業を進めていただきます。（詳細はデータの登録方法をご参照ください。）
6. メタデータの登録作業完了後、ジャパンサーチ上で貴機関のデータが公開され、検索可能となります。また、「ギャラリー」「テーマ別検索」の作成も可能となります。

ジャパンサーチ（試験版）の連携方法

連携機関の作業負担を軽減しつつ、多種多様なメタデータの利活用を可能にするための共通メタデータフォーマット（連携フォーマット/利活用フォーマット）を策定。

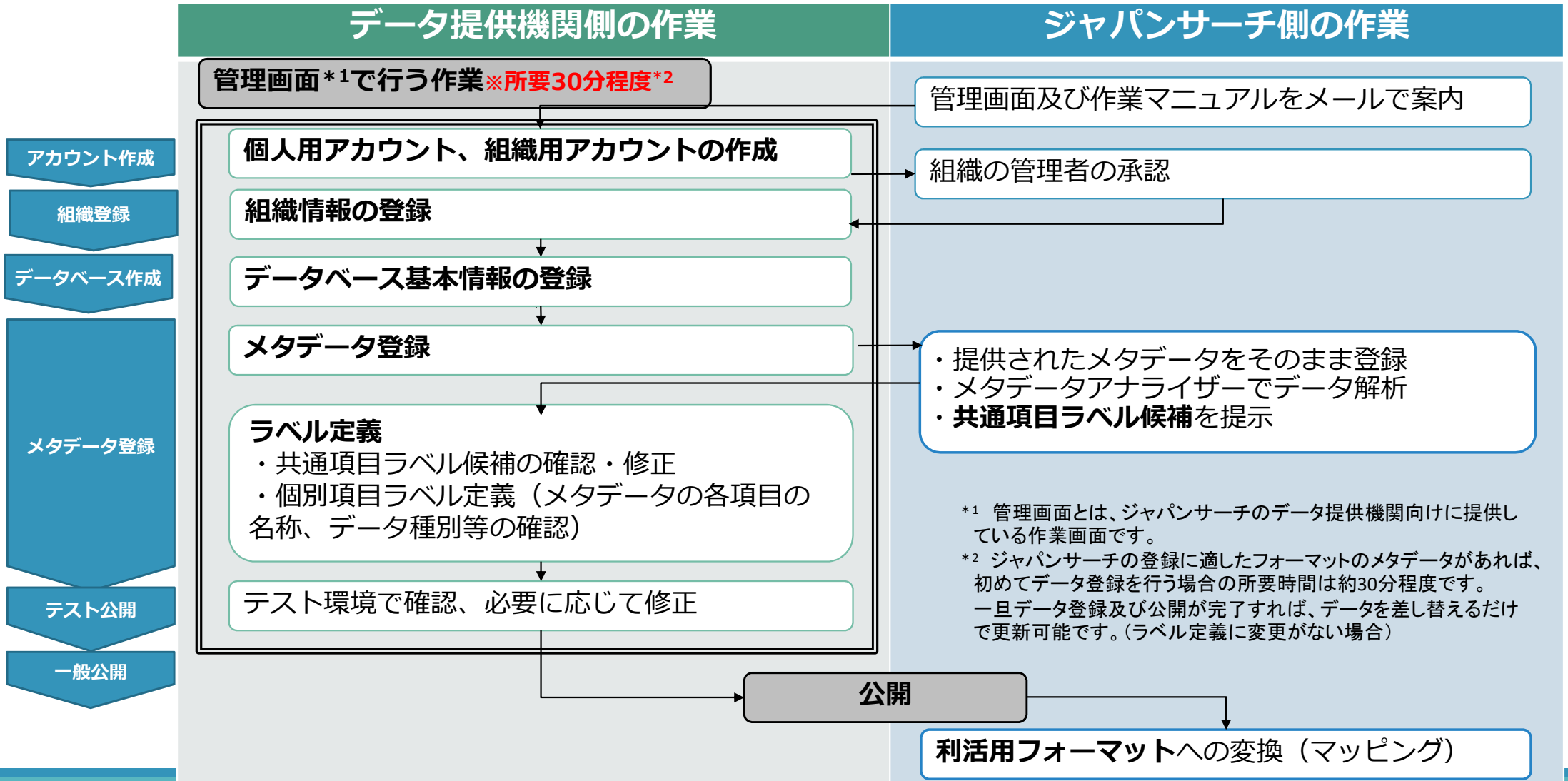


➤ メタデータは原則CC0 ※国際的流通を意識（Europeanaと同レベル）

※著作物性のあるもののみCC BYでも可。民間機関のメタデータなど原則に対応できない場合はデータベースごとに利用条件が分かるよう設定可能
※活用者には、CC0であっても出典等の情報を明記するよう、サイトポリシーでお願い

詳細は「連携をご希望の機関の方へ」を参照 <https://jpsearch.go.jp/cooperation>

メタデータ連携の流れ



データベースの基本情報の入力

必須	ID	データベースのID	各機関で付与。ジャパンサーチ上のデータベース紹介ページのURLに使用
必須	データベース名 (日/英/カナ)	データベースの名称 (日本語/英語/全角カナ)	
	データベースの説明 (日/英)	データベースの説明 (日本語/英語)	100字以上は折り畳み表示
必須	カテゴリ	データベースが扱うコンテンツの種別。選択式	データベース当たり一つが推奨だが、複数選択可
	サブカテゴリ	データベースが扱うコンテンツの種別。選択式のカテゴリを補うための情報。自由記入	複数選択可
	メタデータの権利表示	メタデータの権利情報/二次利用条件について、例外の場合のみ記載。原則はCC0 (著作権のある項目のみCC BYでも可だが、その場合も項目名とCC BYの記載が必要。)	自由記入。クリエイティブ・コモンズライセンス、政府標準利用規約等。権利情報について記述されている外部資源へのリンク (URL) でも可
	サムネイル画像の権利表示	対象のサムネイル画像がある場合のその権利情報/二次利用条件	
	コンテンツの権利表示	対象のデジタルコンテンツがある場合の権利情報/二次利用条件。	
	コンテンツの権利区分	対象のデジタルコンテンツの二次利用条件で、検索・絞り込み用に15種類からの選択式となっている。アイテムにより権利区分が異なる場合は (メタデータの) 共通項目で定義することが可能。	
	コンテンツ公開状況	デジタルコンテンツのアクセス範囲を選択式 (「ウェブ公開」「限定公開」「デジタルコンテンツなし」) で設定。アイテムにより公開状況が異なる場合は (メタデータの) 共通項目で定義することが可能。	
	コンテンツ種別	コンテンツの内容にあてはまるものを1つ選択	画像/動画/文書/表形式/音声/3D/コンテンツによって異なる から選択
	データベースのURL	データベースのURL	
	提供機関 (データベースの管理者)	データベースの管理者 (オーナー) が自動で入る。	
	代表画像	データベースの代表画像	

参考 メタデータの登録

◆ メタデータ項目

メタデータ項目は自由だが、最低限の必須項目（ID、タイトル）がある
 (⇒詳細は「共通項目ラベル」参照)

◆ 登録方法

※OAI-PMHは、大量かつ高頻度の場合に想定

ファイルのアップロード		管理画面から、手動でファイルをアップロードする方法
ファイルをWebに掲載	ファイル取得	Web上にファイルを掲載し、管理画面でそのURLを指定すると、ジャパンサーチ側がファイルを取得しに行く方法
	ファイル定期取得	上記のファイル取得を、指定した頻度（日次・週次・月次）で自動実行し、データを自動更新する方法

◆ 対応ファイル形式

TSV・CSV	文字コードはUTF-8。同じデータ項目であれば複数ファイルも可。
XLSX	複数シート・結合セルには対応できない。XLS非対応。
JSON	1行1レコードのJSON Lines形式を推奨。他の構造にも対応予定。
XML	1行1レコード形式を推奨。他の構造にも対応予定。

ラベルの定義

共通項目ラベル

ID	必須	日本全国書誌番号
名称/タイトル	必須	タイトル
名称/タイトル英語	あれば必須	
名称/タイトルヨミ	あれば必須	タイトルよみ
最終更新日	あれば必須	
URL	あれば必須	NDLサーチへのリンク (URL)
サムネイル画像URL	あれば必須	
IIIFマニフェストURL	あれば必須	
所蔵機関	あれば必須	
所蔵機関URL	任意	
コンテンツの権利区分	あれば必須	
コンテンツ公開状況	あれば必須	
人物/団体	任意	著者 × 出版者 × +

◆ 個別項目ラベルの定義

(オリジナル) ソースデータの項目の名称、データ形式等の定義の確認を行う。必要に応じて項目の説明の追加が可能。

◆ 共通項目ラベルを付与する項目

必須	ID	オリジナルデータ内での一意のID (管理番号)
	名称/タイトル	レコードの名称。検索結果表示に使用される
あれば必須	名称/タイトル英語	レコードの英語名称又はローマ字
	名称/タイトルヨミ	レコードの名称の読み (カタカナ・平仮名問わない)
	最終更新日	データの最終更新日 (日付型)
	URL	レコードのリンク先のURL (つなぎ役のメタデータ表示先)
	サムネイル画像URL	サムネイル画像のURL (幅640~1280pixel)
	IIIFマニフェストURL	IIIFマニフェストのURL
	所蔵機関	オリジナルのコンテンツの所蔵機関
	コンテンツの権利区分	データベース定義の情報と異なる場合のみ。選択式
	コンテンツ公開状況	データベース定義の情報と異なる場合のみ。選択式
任意	人物/団体	作者、発行者、出演者など (複数可)
	時間/時代	制作年、対象時期など (複数可)
	場所	発行地、制作地など (複数可)
	座標 (緯度・経度)	座標
	所蔵機関URL	オリジナルのコンテンツの所蔵機関のURL
	解説	オリジナルのコンテンツに関する解説
	デジタルコンテンツURL	デジタルコンテンツの公開URL
	デジタルコンテンツ種別	データベース定義の情報と異なる場合のみ。選択式

Relational Databaseとelasticsearch

Relational Database (RDB)

ID	タイトル	分類	その他
1	X	Y	Z
2	A	B,C	
3	P		Q

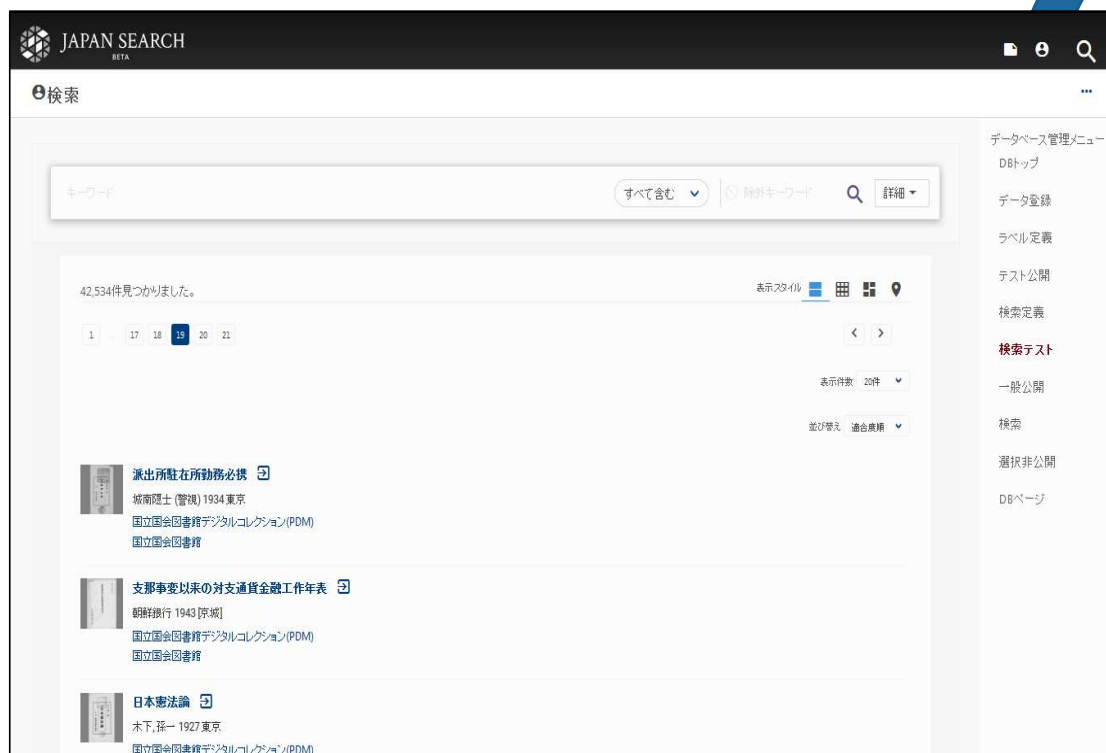
- 列（テーブル）を定義して、1行1データで値を入れていく。
- 1対多対応の構造がある場合には、テーブルを分ける必要がある場合も。
- データ構造を決めないと、システムが作れない。

elasticsearch

<pre>ID:1 { title:X, category:[Y], extent:Z }</pre>	<pre>ID:2 { title:A, category:[B,C] }</pre>	<pre>ID:3 { title:P, extent:Q }</pre>
---	---	---

- Elasticsearchはdocument store型のデータベースエンジンで、任意の構造のJSONを保存し、ID指定で取得することができる。
 - 事前に構造を決めなくても良い
- 検索のためには、検索用の転置インデックスが生成される。

検索テスト



検索テスト画面



検索テストの時点では関係者のみ確認できます。
⇒必要に応じて修正後、一般公開へ

連携に向けて、アーカイブ機関側で必要な整備

1. メタデータの整備

「名称/タイトル」「ID」のみ必須ですが、次の点にもご留意ください。

- ジャパンサーチからとぶリンク先ページの「固定URL」
- 「サムネイル画像URL」
- （可能なら）名称/タイトルのローマ字（英語）

2. 二次利用条件の設定

メタデータ、サムネイル、デジタルコンテンツの利用条件及び公開についてご検討ください。

- ジャパンサーチでは、メタデータは「**原則CC0**」

※国際的流通を意識（Europeanaと同レベル）

※著作物性のあるもののみCC BYでも可。民間機関のメタデータなど原則に対応できない場合はデータベースごとに利用条件が分かるよう設定可能

※活用者には、CC0であっても出典等の情報を明記するよう、サイトポリシーでお願い

- サムネイルはCC0/CC BY相当に
- デジタルコンテンツのウェブ公開を増やし、可能ならCC BY相当に

各機関のご判断により、ウェブ公開有無や利用条件の設定ができますが、可能な範囲で、オープンな利活用にご協力をお願いします。

（参考）自機関のデジタルアーカイブの状態を点検される場合は、「**デジタルアーカイブアセスメントツール**」（「第二次中間取りまとめ」（実務者検討委員会 平成31年4月））をご利用下さい。

ご清聴ありがとうございました



ジャパンサーチ試験版（公式）
[@jpsearch_go](https://twitter.com/jpsearch_go)

#ジャパンサーチ
で検索！



フィードバックをお持ちしています
<https://jpsearch.go.jp/contact>