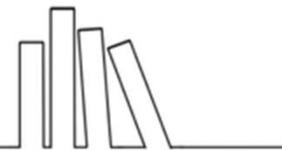


日本図書館研究会 2022年度図書館学セミナー
2022年11月19日



国立国会図書館の 個人向けデジタル化資料送信サービス

国立国会図書館 利用者サービス部サービス企画課 福林靖博



自己紹介

- 2002年～、国立国会図書館に勤務。
 - 閲覧/ILL、レファレンス（リサーチ・ナビ）、情報システム、予算、国会向け調査部門…
- 2018年～2022年、電子情報企画課
 - デジタル関連・情報システム事業の全体総括
 - 資料デジタル化及びその利活用に係る関係者協議
 - 第五期科学技術情報基本計画の策定
 - 2021年著作権法改正対応
- 2022年～、サービス企画課
 - 一般公衆向けサービス部門の全体総括
 - 個人向けデジタル化資料送信サービス、公衆送信に係る関係者協議

本日の内容

□ 国立国会図書館のビジョン

□ 所蔵資料デジタル化の推進

□ デジタル化資料の個人への送信

サービスの概要と利用状況、そして新「デジコレ」

□ デジタル化資料の利活用

全文テキストとイメージ、そしてジャパンサーチ

国立国会図書館のビジョン2021-2025

ビジョン 2021-2025

国立国会図書館のデジタルシフト

— 情報資源と知的活動をつなぐ7つの重点事業 —

ユニバーサルアクセスの実現



情報技術によって私たちの社会活動やコミュニケーションのあり方が変容し、さらに新型コロナウイルスが大きな影響を及ぼす中で、社会のデジタルシフトが加速しようとしています。この状況を踏まえ、国立国会図書館は、情報資源と様々な知的活動を的確につなげていくために、今後5年間を「国立国会図書館のデジタルシフト」推進期間と位置付け、7つの事業に特に重点的に取り組みます。これらの重点事業は、将来にわたる全ての利用者に多様な情報資源を提供するユニバーサルアクセスを実現する事業と、そのための恒久的なインフラとなる国のデジタル情報基盤の拡充を図る事業から構成されます。

国のデジタル情報基盤の拡充



国会サービスの充実

量的・質的に拡充したデジタル情報基盤と利便性を向上させた検索手段を用いて、さらに充実した国会サービスの提供を図ります。



インターネット提供資料の拡充

インターネットや身近な図書館で閲覧できるデジタル資料の拡充を図ります。そのための著作権処理や関係者との合意形成を進めます。



資料デジタル化の加速

デジタルで全ての国内出版物が読める未来を目指し、この5年間で100万冊以上の所蔵資料をデジタル化します。テキスト化も行い、検索や機械学習に活かせる基盤データとします。



デジタル資料の収集と長期保存

有償の電子書籍・電子雑誌の制度収集を開始し、著作権者や出版者の協力を得て、安定的収集を実現します。また、他機関のデジタル資料の収集・移管、再生困難なデジタル資料の形式変換等、多面的な取組によってデジタル資料の長期保存を目指します。



読書バリアフリーの推進

視覚障害等の理由で経書に困難がある利用者向けに、バリアフリー対応の資料の収集・検索・提供サービスと、利用しやすいテキストデータの製作支援を推進します。



「知りたい」を支援する情報発信

専門知識を活かして膨大な資料・情報をキュレーションし、効率的な調べ方のガイドや、知識を深めるための資料の紹介等、社会に役立つ情報を発信します。



デジタルアーカイブの推進と利活用

図書館の領域を超えて幅広い分野のデジタルアーカイブを連携させる「ジャパンサーチ」を通じて、多様な情報・データがオープン化され、活用が促進される環境づくりを支えます。

所蔵資料デジタル化事業の経緯

2000年	<ul style="list-style-type: none"> 資料デジタル化を開始。著作権処理を行いインターネットで公開（2～4万冊／年）
2008年	<ul style="list-style-type: none"> 資料デジタル化及び利用に係る関係者協議会発足
2009年	<ul style="list-style-type: none"> 著作権法改正（NDLで保存目的のデジタル化が明確化）
2009～ 2011年	<ul style="list-style-type: none"> 大規模デジタル化事業実施（2009年度、2010年度補正予算） 図書66万点、雑誌22万点、古典籍7万点、博士論文14万点、官報等のデジタル化実施。
2012年	<ul style="list-style-type: none"> 著作権法改正（図書館等への絶版等入手困難な資料の送信が可能に）
2014年	<ul style="list-style-type: none"> 図書館向けデジタル化資料送信サービス（図書館送信）開始
2015年	<ul style="list-style-type: none"> 災害関係資料のデジタル化（2014年度補正予算）図書約6万点、雑誌約2万点のデジタル化
2018年	<ul style="list-style-type: none"> 著作権法改正（外国の図書館等へも絶版等入手困難な資料の送信が可能に）
2019年	<ul style="list-style-type: none"> 外国の図書館等にも図書館向けデジタル化資料送信サービスを拡大 デジタル化内製の実験プロジェクト開始
2021年	<ul style="list-style-type: none"> 国内刊行図書のデジタル化（2020年度補正予算）、資料デジタル化推進室の設置 「資料デジタル化基本計画2021-2025」の策定 著作権法改正（特に第31条第3項：絶版等資料の個人（家庭）への送信） 国内刊行図書のデジタル化（2021年度補正予算）
2022年	<ul style="list-style-type: none"> 個人向けデジタル化資料送信サービス（個人送信）開始 国立国会図書館デジタルコレクションリニューアル（12月予定）

資料デジタル化基本計画2021-2025

評価要素	<ul style="list-style-type: none">• 唯一性・希少性• <u>資料の利用機会の拡大</u>（インターネット公開や図書館・個人送信が見込まれるか）• 資料の劣化状況、保存の緊急性• デジタル化への社会的・学術的ニーズ• 国や世界の体系的なデジタルコレクション構築への貢献
対象資料	<p>日本で刊行された資料（外国刊行の日本語資料・日本関係資料も含む）</p> <ul style="list-style-type: none">• 図書（<u>2000年までに刊行されたもの</u>）※官庁出版物はそれ以降も含む• 雑誌（刊行後5年以上経過したもの）• 古典籍資料• 録音・映像資料• 博士論文• 他（憲政資料、日本占領関係資料、<u>日系移民関係資料</u>、地図、<u>新聞<試行></u>）
利用提供	<ul style="list-style-type: none">• 「国立国会図書館デジタルコレクション」で提供• <u>本文テキストデータの作成を推進し、全文検索を可能に</u>• デジタル化済み原資料は原則として利用停止• 公開範囲：館内限定・図書館送信・インターネット公開

https://www.ndl.go.jp/jp/preservation/digitization/digitization_plan2021.pdf

補正予算による所蔵資料デジタル化の加速

2020年度補正予算（第3号）

項目	概要
図書資料のデジタル化	1987年までに整理した国内刊行図書のデジタル化 約45億円 ※社会科学分野、人文科学分野の一部（大半が入手困難資料） ⇒館内公開&入手可能性調査中。入手困難なものは来年から送信予定
デジタル化設備の整備	館内で所蔵資料のデジタル化を行うための ブックスキャナ等導入
全文テキスト化の推進	デジタル化済み資料のOCRによる 全文検索用のテキスト化 OCR精度向上に向けた研究開発 ⇒来月から全文検索サービスを開始。OCRはOSSとして公開済
電子書庫機能の拡張等	ストレージ増強・国立国会図書館デジタルコレクションの改修

視覚障害者の方の読み上げ用にも

2021年度補正予算（第1号）

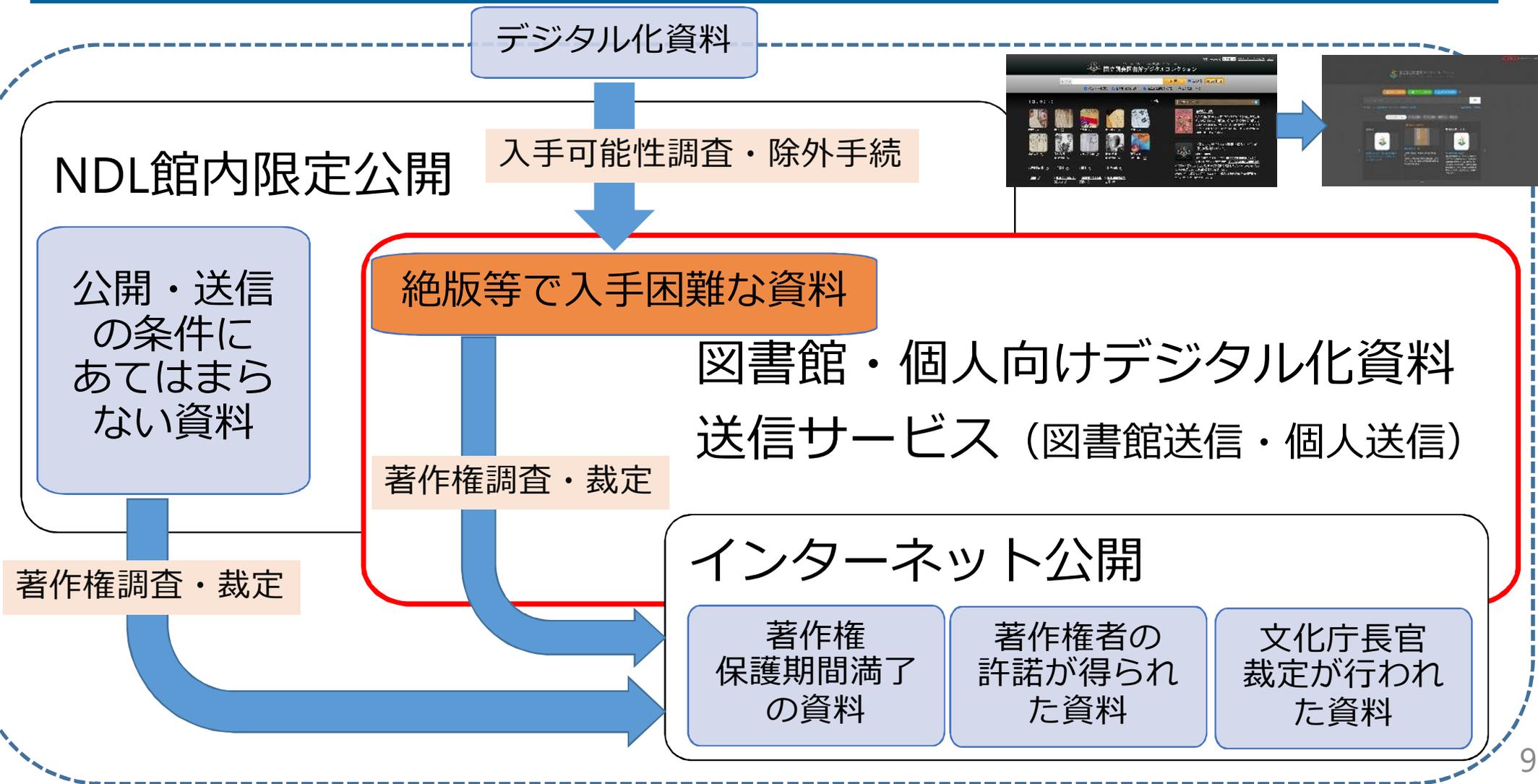
項目	概要
図書資料のデジタル化	1987年までに整理した国内刊行図書のデジタル化 約37億円 ※人文科学分野、自然科学分野の一部（大半が入手困難資料） ⇒来年、入手可能性調査実施予定。入手困難なものは再来年送信予定
全文テキスト化の推進	視覚障害者向け（読み上げ用）OCR処理プログラムの研究開発
電子書庫機能の拡張等	ストレージ増強・視覚障害者向けの全文テキストデータ提供機能の拡張

所蔵資料デジタル化事業の現況

資料	これまでの取組	インターネット公開	図書館/個人送信	NDL館内限定	合計	昨年度からの取組
図書	明治期以降、1968年までに受け入れた図書 震災・災害関係資料の一部（1968年以降に受け入れたものを含む。）	36万点	54万点	38万点	128万点	2000年までに刊行・受入したもの（対象：約170万冊） ※官庁出版物は2000年以降も含む ※5年間で100万冊以上のデジタル化を目指す
雑誌	明治期以降に刊行された雑誌（刊行後5年以上経過したもの）	2万点	82万点	51万点	135万点	刊行後5年以上経過した雑誌 →学協会等からデジタル化要望があるものを優先する。
博士論文	1990～2000年度に送付を受けた論文	2万点	13万点	2万点	16万点	1989年度以前に送付を受けたもの。
新聞	未提供	-	-	-	-	試行（日本新聞協会と合意したもの等。ただし、商用データベース等で提供されているものは対象外。）
その他	古典籍、地図、官報、録音・映像資料、憲政資料、日本占領関係資料等	19万点	4万点	10万点	31万点	（継続） 新たに、日系移民関係資料を追加
※2022年11月時点の提供点数		57万点	152万点	102万点	311万点	

<https://www.ndl.go.jp/jp/preservation/digitization/index.html#situation>

個人送信対象資料の概念図



デジタル化資料の図書館への送信

2012年	著作権法改正（第31条第3項新設）→図書館等への絶版等入手困難な資料の送信が可能に	（2014年～）図書館送信サービス開始
2018年	同法改正（第31条第3項）→外国の図書館等へ絶版等入手困難な資料の送信が可能に	（2019年～）外国の図書館等に拡大
2021年	同法改正（第31条第3項ほか）→個人（家庭）への絶版等入手困難資料の送信が可能に	（2022年5月～）個人への送信サービス開始

対象資料	<ul style="list-style-type: none"> 絶版等の理由で入手困難な資料 NDLが入手可能性調査を行い市場で流通していないことを確認 ※ただし、漫画・絵本・商業出版による雑誌・管理委託著作物等は除外 出版者・著作(権)者等の申し出により一定の除外基準に該当するものを送信対象から除外
対象施設	<ul style="list-style-type: none"> 著作権法第31条に規定する「図書館等」 ＝公共図書館、大学図書館、国公立博物館・美術館、国公立の研究機関の図書館 公益法人設置の図書館（個別指定）、公益法人立の博物館・博物館相当施設 ※司書または司書に相当する職員の配置が必要 2019年度から外国の図書館にも送信開始（2018年著作権法改正） 関係者協議での合意に基づき、NDLによる要件確認・承認が必要
参加館数	国内1,390館＋海外6館（2022年11月1日現在）

「[国立国会図書館のデジタル化資料の図書館等への限定送信に関する合意事項](#)」に基づき運用

個人向けデジタル化資料送信サービス

- 「国立国会図書館による入手困難資料の個人送信に関する関係者協議会」で協議
- 2021年12月に「国立国会図書館のデジタル化資料の個人送信に関する合意文書」を公表

なぜ	<p>コロナ禍における研究者・学生等からの来館せず利用できる図書館サービスへのニーズの高まり ⇒ 出版者協議会、図書館休館対策プロジェクト、日本歴史学協会等からのデジタル化資料公開範囲拡大の要望 ⇒ 「知的財産推進計画2020」「図書館関係の権利制限規定の見直し（デジタル・ネットワーク対応）に関する報告書」</p>
いつ	2022年5月19日開始 （改正著作権法施行は2022年5月1日）
だれが	国立国会図書館
なにを	<p>国立国会図書館がデジタル化した資料のうち絶版等により入手困難なもの ⇒ 図書館送信と同じ範囲（国立国会図書館による入手可能性調査及び出版者・権利者等による除外確認を経たもの。また、3月以内に復刊予定のものも除く） ⇒ 他の図書館から提供された、国立国会図書館未収かつ入手困難なデジタル化資料の提供も可能（*）</p>
だれに	<p>国立国会図書館登録利用者のうち国内在住かつ当該サービスの利用規約に同意した者 ⇒ 既に6.8万人超 ⇒ 氏名・現住所・生年月日が確認できる身分証明書による本人確認が必要。本登録のみ。 ⇒ 利用者登録は、来館・郵送・オンラインにより受付。 ※海外への送信は来年度以降に検討予定。</p>
どのように	<p>利用者がID・PWにより「国立国会図書館デジタルコレクション」にアクセス ⇒ 国立国会図書館オンラインへのログイン時に利用規約への同意が必要。 ⇒ 学校図書室等の公の施設で100インチ以下のディスプレイを用いてデジタル化資料を見せることも可能。 ⇒ 当面はストリーミング（閲覧）のみ。プリントアウト（制限なし）は2023年1月下旬から開始予定。 ⇒ 2022年12月にデジタルコレクションを全面リニューアル予定。デジタル化資料（約247万点）の本文の検索も可能となる予定。</p>
いくらで	無料

(*）国立国会図書館未収かつ入手困難資料のデータ収集事業へのご協力をお願い <https://www.ndl.go.jp/jp/preservation/digitization/data-acceptance.html>

NDL未収入手困難資料のデータ収集

- NDL未収、かつ絶版等により入手困難資料となっている資料について公共図書館・大学図書館等がデジタル化したデータを当館にご提供いただき、国立国会図書館デジタルコレクションのデジタル化資料として長期にわたって保存するとともに、オンラインでの幅広い利用に供することを旨とするもの。
- 収集したデジタル化データは国立国会図書館デジタルコレクションを通じて、各資料に応じた範囲で利用提供。
⇒図書館・個人への送信が可能。
- 著作権法第31条1項3号に基づく。
- 「国民の情報アクセスを確保する観点から…望ましい」
 - 「図書館関係の権利制限規定の見直し（デジタル・ネットワーク対応）に関する報告書」（令和3年2月）
 - 「国立国会図書館のデジタル化資料の個人送信に関する合意文書」（令和3年12月3日）



資料の保存

- ▶ ニュース
- ▶ 所蔵資料の保存
- ▶ 資料のデジタル化について
- ▶ 電子情報の機能的保存と利用
- ▶ 保存技術
- ▶ IFLA/PAC/アジア地域センター
- ▶ マニュアル・パンフレット・冊子資料

国立国会図書館未収かつ入手困難資料のデータ収集事業へのご協力のお願い

国立国会図書館では、当館が所蔵していない入手困難資料について、公共図書館・大学図書館等からデジタル化データの形で収集する事業を推進しています。このページでは、日本国内の図書館の方を対象に、当該事業の概要をご説明するとともに、デジタル化データ提供へのご協力をお願いするものです。

事業の概要

本事業は、当館未収かつ絶版等により入手困難となっている資料について公共図書館・大学図書館等がデジタル化したデータを当館にご提供いただき、[国立国会図書館デジタルコレクション](#)のデジタル化資料として長期にわたって保存するとともに、オンラインでの幅広い利用に供することを旨とするものです。

収集対象となるデジタル化データ

収集の対象となるのは、以下の2つの条件を満たす資料について、各所属機関でデジタル化したデータです。主に国内発行の図書・雑誌を想定しています。

1. 国立国会図書館が所蔵していないこと
2. 絶版等で入手困難であること

※「よくあるご質問」もご確認ください。

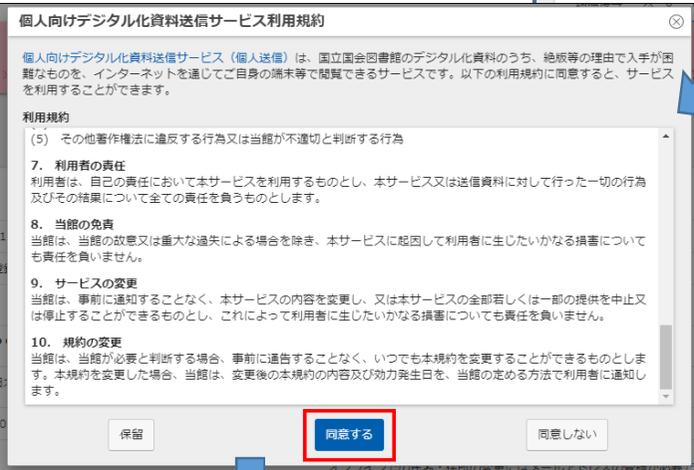
▶ ページの先頭へ

<https://www.ndl.go.jp/jp/preservation/digitization/data-acceptance.html>

個人送信サービスの流れ（12月中旬まで）



① 国立国会図書館オンラインでログイン
※初回のみ。IDない場合、新規登録。



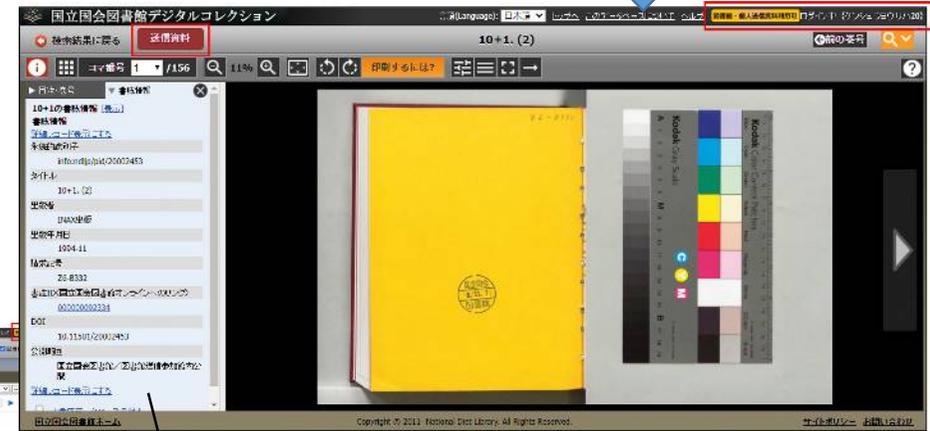
② 利用規約に「同意」
※初回のみ。



③-A 国立国会図書館オンラインの検索結果からアクセス



④-A 再度ログイン



⑤ デジタルコレクションで閲覧



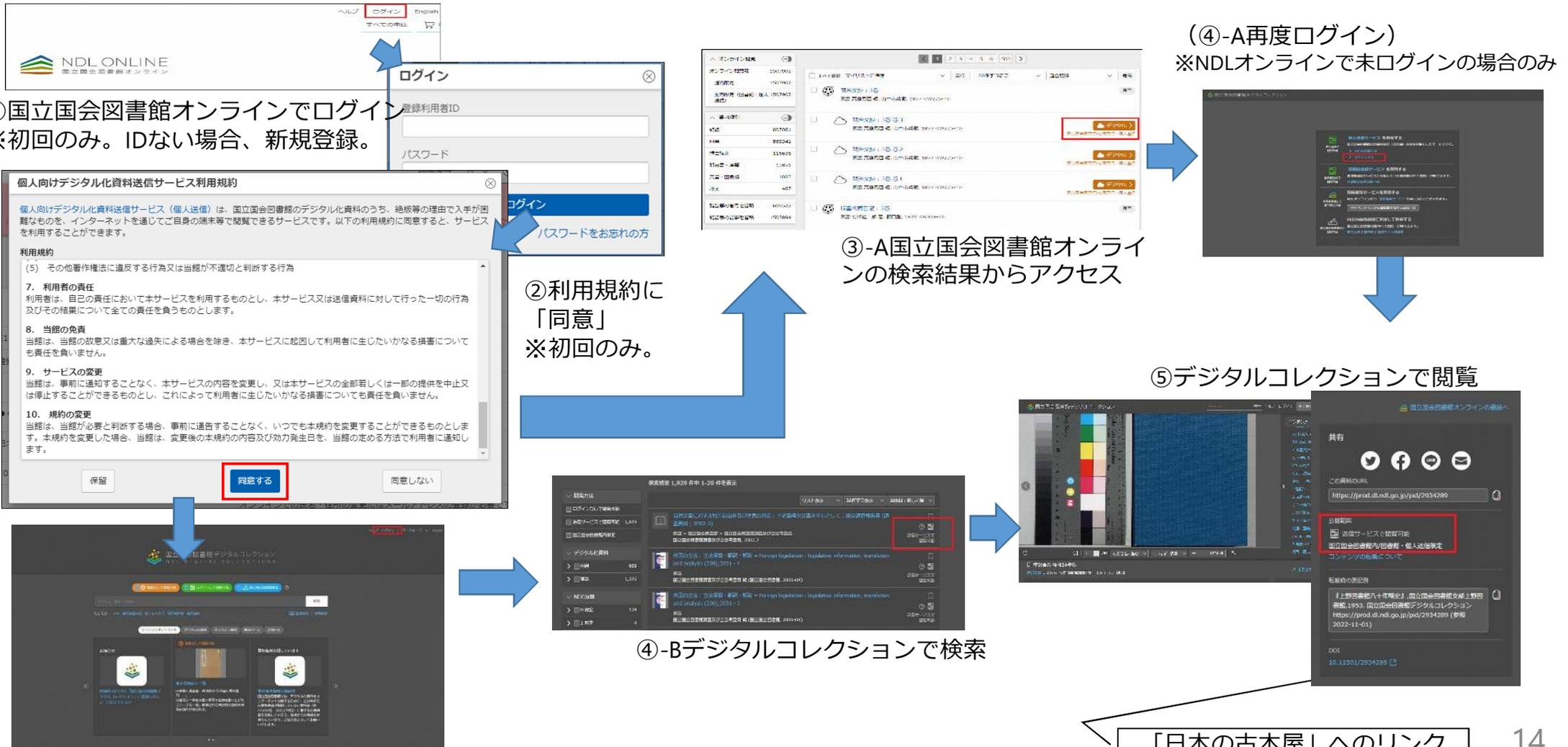
③-B 国立国会図書館デジタルコレクションでログイン



④-B デジタルコレクションで検索

「日本の古本屋」へのリンク

個人送信サービスの流れ（12月中旬から）



個人送信サービスの利用状況

利用規約に同意した利用者数	<ul style="list-style-type: none"> 10月末時点で約6.8万人 ⇒同時にオンラインでの利用者登録を開始した影響も大きい。
利用（閲覧件数）	<ul style="list-style-type: none"> 月間約26.5万件（6-9月の平均） ⇒図書館送信の8倍程度で推移
分野別の利用状況	<ul style="list-style-type: none"> 9月は、図書が約51%、雑誌が約48%。図書の利用うち約68%が人文系（提供は約59%）。 ⇒日本十進分類法の類目別に集計すると、<u>歴史・地理分野と芸術分野の利用が、提供割合に比して多い。</u>毎月ほぼ同じ傾向。 ※図書館送信サービスの閲覧や、来館者の漫画を除いた閲覧利用と、概ね同じ傾向。
アクセスランキング	<ul style="list-style-type: none"> 一位：5月は『怪獣ウルトラ図鑑カラー版』、7月と9月は『故吉田茂国葬儀記録』。 ⇒時事ニュースやSNSのトレンドと連動する傾向も見られる。 ⇒『人事興信録』『大日本職業別明細図』は安定的にアクセス...人名や会社要録、企業情報、国勢調査のような、実用性の高いデータが閲覧される傾向。<u>調査・研究目的か。</u>
その他の動き	<ul style="list-style-type: none"> テーマ別のまとめの登場（刀剣や鉄道まとめサイト、俳句雑誌での特集など） 古書店販売サイト（「日本の古本屋」）への導線リンクの利用の可視化 ⇒図書館送信開始時に、古書店団体との協議を踏まえて設けたもの 出版業界の絶版本への再注目？

※NDLの既存サービスへの影響の有無は、現時点では不明（例えば、コロナ禍で来館利用者は減）。ただし、これまでの資料利用の傾向から推測するに、現時点での影響は限定的か。

デジタル化資料の全文テキストデータ

(2021年度の取組)

- 「国立国会図書館デジタルコレクション」搭載のほぼ全てのデジタル化資料（約2億2300万画像コマ）の全文テキスト化（検索用）
- オープンソースとして公開可能なOCR処理プログラムの研究開発

(2022年度以降の取組)



1. 全文検索サービス（スニペット表示も）の提供

- NDLラボ上の「次世代デジタルライブラリー」で実験後、「国立国会図書館デジタルコレクション」で247万点分の全文テキストデータを提供（2022.12～）

2. 視覚障害等者向け（読み上げ用）のOCR処理プログラムの研究開発

- 市場で流通しているものを除外のうえ、視覚障害者等用データ送信サービスから提供

3. NDL-OCR（プログラム&データセット）の公開（NDLラボ/GitHub）

- CC BYでOCR処理プログラムをOSSとして公開（2022.4～） ⇒OCRの更なる精度向上へ
- 膨大なテキストデータの提供 ⇒様々な研究活動へ

4. テキストデータの効果的な利用提供方法の調査研究

- 資料中の図版キャプションの自動抽出による「キャプション検索」、1行単位の「類似文章検索」など
⇒NDLラボ上で実験公開。5月にNDL Ngram Viewer公開（デジタル化資料本文中に特定検索語が表れる頻度を列挙し、時系列で可視化） <https://lab.ndl.go.jp/ngramviewer/>



次世代デジタルライブラリー（実験システム）

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

● 主な機能

- 全文テキスト検索
- 資料中の図版（図・挿絵・写真等）の自動抽出及びその一覧表示
- 類似図版検索
- 見開き2頁画像の自動分割による1頁表示
- 紙の変色の自動除去（白色化機能） など

● 検索対象：

国立国会図書館デジタルコレクション
(<http://dl.ndl.go.jp/>)でウェブ公開している著作権
保護期間満了（PDM）の図書・古典籍約33万点
【本文テキスト】 PDMの図書28万点+古典籍6万点
【資料中の図版】 全て

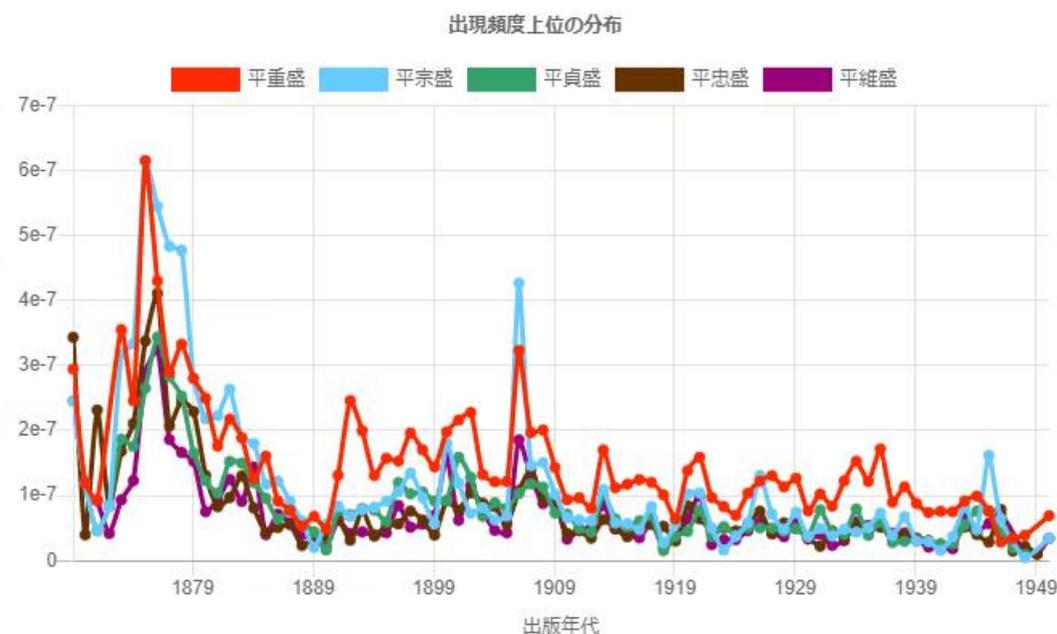
● Chrome、Firefox推奨

NDL Ngram Viewer (実験システム)

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

- 主な機能
 - 出版年代ごとのキーワードの可視化
 - 正規表現検索のサポート
- 検索対象：
国立国会図書館デジタルコレクション(<http://dl.ndl.go.jp/>)でウェブ公開している著作権保護期間満了(PDM)の図書約28万点
- Chrome、Firefox推奨

検索クエリ「平.盛」



年内に検索対象をデジタル化図書雑誌全件に拡大予定

NDLイメージバンク

<https://rnavi.ndl.go.jp/imagebank/>

- 国立国会図書館所蔵の浮世絵、雑誌、図書などから、選りすぐりのビジュアル資料を紹介するオンライン展示。
- 歌川広重や川瀬巴水の描いた美しい日本の風景画や、大正ロマンを象徴する画家竹久夢二の描く可憐な美人画の人気作のほか、雛祭り、お花見、花火など四季の風物を描いた江戸の浮世絵、文様・図案・デザイン等も多数収録。
- 著作権保護期間を満了した画像を利用しているので、年賀状、オリジナルグッズづくりなど、様々な用途に活用可能。



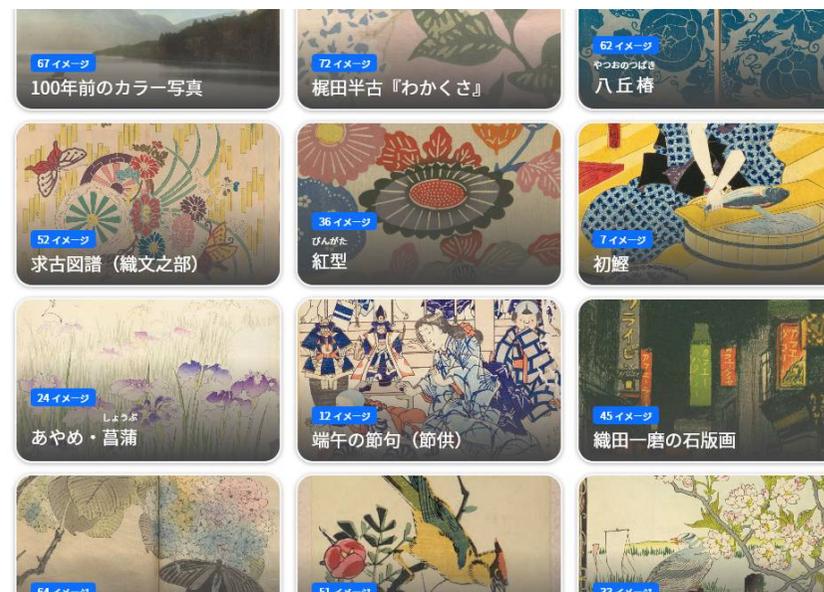
NDLイメージバンク テーマ一覧 コラム一覧 ご利用について

NDLイメージバンク

インターネットもテレビもない時代、書物は多くの人々にイメージを届ける媒体でした。
NDLイメージバンクでは、浮世絵、図書、雑誌などの様々なメディアに載ったイメージをお届けします。



手彩色の風景写真

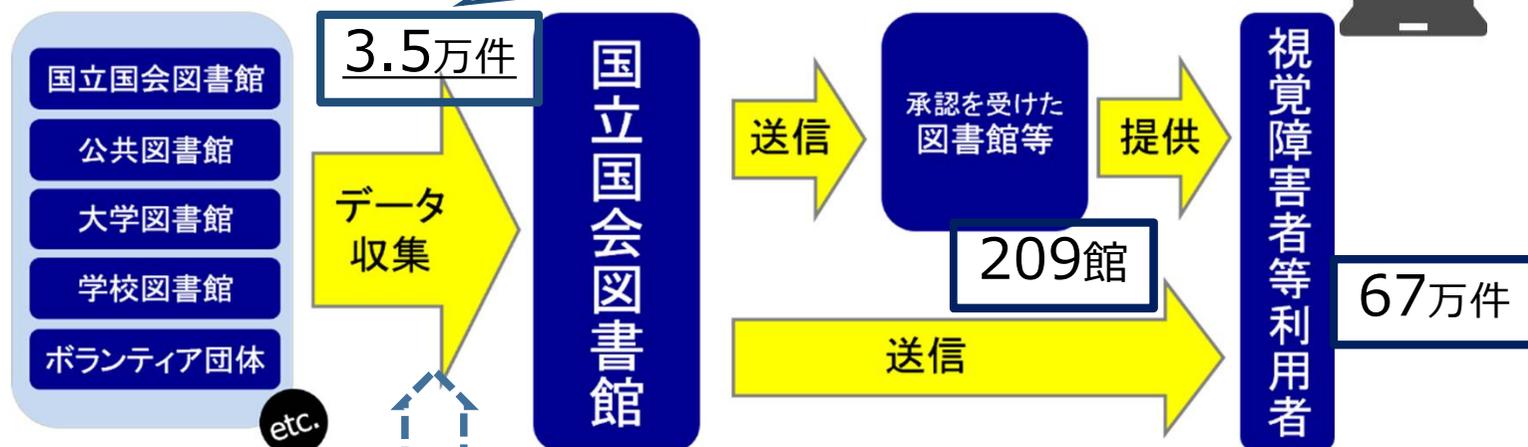


視覚障害者等へのテキストデータ提供

視覚障害者等の読書環境の整備の推進に関する法律（読書バリアフリー法）（R元）
⇒「視覚障害者等の読書環境の整備の推進に関する基本的な計画」（R2）
（対象期間：R2～R6）

視覚障害者等用データ送信サービス

依頼による学術文献テキストデータ製作
校正済データ51タイトル
未校正データ123タイトル /FY2021



NDL
「障害者サービス実施計画 2021-2024」

FY2022～（予定）未校正テキストデータ
247万件から 市販アクセシブル電子書籍等除く

ジャパンサーチ

● デジタルアーカイブの検索・閲覧・活用基盤

さまざまな分野のデジタルアーカイブと連携し、我が国の多様なコンテンツの**メタデータ**をまとめて検索・閲覧・活用

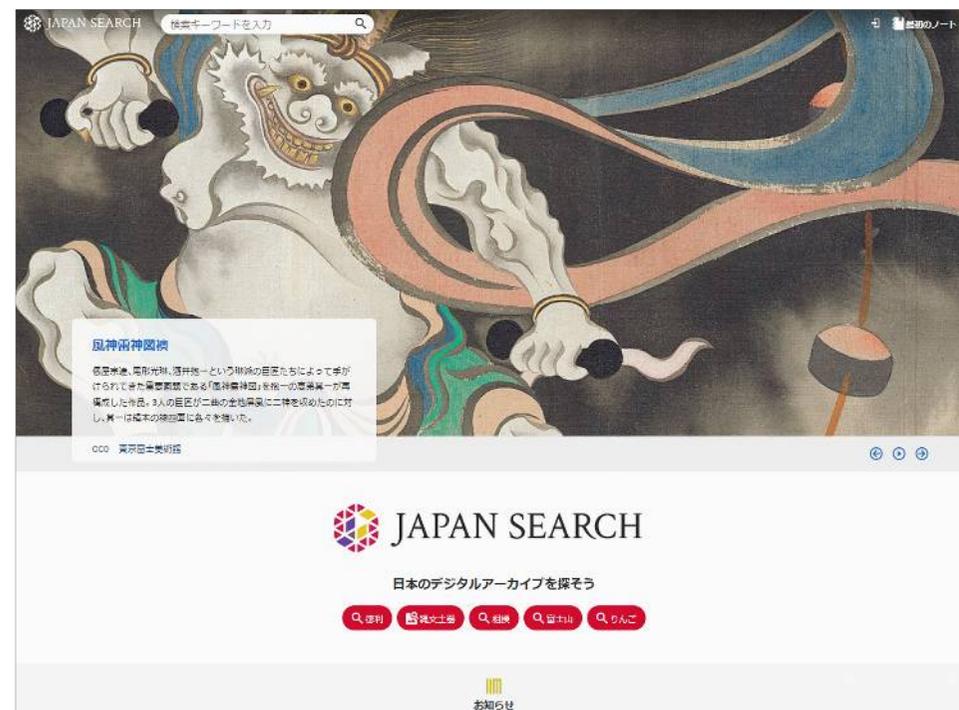
- **37**連携（つなぎ役）機関から、**192**データベース・メタデータ約**2,600**万件を検索可能（2022年11月現在）

● 国全体の取組

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）
 システムの開発・運用担当：国立国会図書館

● 実験的な技術の取込

類似のサムネイル画像の検索
 ローマ字読みの自動生成



ジャパンサーチトップ画面
<https://jpsearch.go.jp/>

参考

NDLのサービスの見取り図

