

テーマ：国立国会図書館のデジタル化について

講師：中島 寛 氏（国立国会図書館電子情報部電子情報企画課

資料デジタル化推進室資料デジタル化推進係長）

令和5年8月29日（火）東京都立中央図書館において開催されたTLA講演会を再構成してまとめたものです。
挿入スライドは、配布資料として別途HPで公開しています（令和5年11月1日現在）。詳しくはそちらを御覧ください。

国立国会図書館の中島と申します。どうぞよろしくお願ひいたします。

肩書が長くて申し訳ありません。私の所属する資料デジタル化推進室は、2021年4月に新しく電子情報部の下に設置された部署です。以前は、各資料の所管課が個別に資料デジタル化を実施したり、補正予算が措置されて大規模にデジタル化を実施する際には臨時に体制を立ち上げて実施したりしてきましたが、近年、資料のデジタル化を当館の重点事業の一つとして取組を加速させていくこととなり、一つ基幹となる部署が必要ということで、資料デジタル化推進室が設置された次第です。



私は2004年に国立国会図書館に入館し、様々な業務を担当してまいりました。過去にレファレンス係長を2、3年務めておりましたので、もしかしたら皆さんの中に、何かやり取りをさせていただいた方がいらっしゃるかもしれませんが、本日

は当館のデジタル化についてというテーマでお話をさせていただきたいと思います。

国立国会図書館の概要

- 1948年設立
- 国会に属する唯一の国立の図書館
- 納本制度に基づく資料・情報の収集を核として、国会、行政・司法各部門、国民に対するサービスを実施

➤ 国立国会図書館法（昭和23年法律第5号）

前文「国立国会図書館は、真理がわれらを自由にするという確信に立って、憲法の誓約する日本の民主化と世界平和とに寄与することを使命として、ここに設立される。」

第2条「国立国会図書館は、図書及びその他の図書館資料を蒐集し、国会議員の職務の遂行に資するとともに、行政及び司法の各部門に対し、更に日本国民に対し、この法律に規定する図書館奉仕を提供することを目的とする。」

（参考）国立国会図書館について <https://www.ndl.go.jp/jp/aboutus/index.html>

1

まず、当館の概要について簡単に説明させていただきます。当館は1948年に設立されました。国会に属する唯一の国立の図書館であり、司書資格の有無を問わず、様々な専門分野を持つ職員が勤務しています。

当館は、納本制度に基づく資料・情報の収集を核として、国会、行政・司法各部門、国民に対するサービスを実施することを使命としています。国立国会図書館法の前文には、「真理がわれらを自由にするという確信に立って、憲法の誓約する日本の民主化と世界平和とに寄与することを使命として、ここに設立される」とあります。それから第2条には、「国立国会図書館は、図書及びその他の図書館資料を蒐集し、国会議員の職務の遂行に資するとともに、行政及び司法の各部門に対し、更に日本国民に対し、この法律に規定する図書館奉仕を提供することを目的とする」とあります。

私個人も、日本国民に広くサービスを提供するという視点で働いております。

施設		
※座席数は令和3年度末時点の数値。		
 東京本館 国会向けのサービスや来館サービスを提供、3施設の統括 収蔵能力 本館書庫：450万冊 新館書庫：750万冊 閲覧スペース 18,983㎡／1,018席	 関西館 遠隔サービスの拠点、来館サービスも提供 収蔵能力 本館書庫：600万冊 書庫棟：500万冊 閲覧スペース 4,265㎡／359席	 国際子ども図書館 児童書の専門図書館 収蔵能力 レンガ棟：40万冊 アーチ棟：65万冊 閲覧スペース 2,054㎡／124席
(参考) 建物と設備 https://www.ndl.go.jp/ky/aboutus/outline/design.html 国立国会図書館年報 https://www.ndl.go.jp/ky/publication/annual/index.html		

当館は、東京本館、関西館、国際子ども図書館の3施設からなります。東京本館は、国会議事堂のすぐ横にあり、国会向けのサービスと一般利用者向けの来館サービスを提供しながら、3施設の統括をしています。スライドの「収蔵能力」というのは、本1冊当たりの厚みが3センチだと仮定して、その施設の書架に何冊並べることができるかという数字です。関西館は、新しく建設された施設ですが、最近も新しく書庫棟を建設して収蔵能力を増強しており、今では東京本館と同じぐらいの収蔵能力を有しています。

なぜ関西館の収蔵能力を増強しているかというと、元々関西館が、図書館資料の増加に対応した収蔵書庫の確保を目的の一つとして建設されたことによります。デジタル化をした資料に関しては、デジタル化画像を利用いただくこととし、原則として資料原本は利用に供しません。東京本館の資料をデジタル化し、関西館に移送して保存するという整理になっています。ただ、関西館に移送したデジタル化済み資料原本に絶対に触れないというわけではありません。特に、昔デジタル化した資料の中には、画像では何が書いてあるか読めず、結局、資料原本を確認しなければならないものもあります。このような資料への対応も課題であり、現在の技術で再度デジタル化することも進めているところです。

それから、国際子ども図書館は、児童書の専門図書館という位置付けで、当館の資料のうち児童書に関してはこちらの館で所蔵しています。児童書についても、他の資料と同じようにデジタル化が行われています。

所蔵資料

※いずれも令和3年度末時点の数値。

➤ 有体物

図書	雑誌・新聞	その他非図書資料等
約1,193万点	約1,994万点	約1,435万点

➤ 無体物

インターネット資料 (ウェブサイト)	オンライン資料 (電子書籍・電子雑誌等)
約1.4万タイトル 約22万件 (データ量約2PB)	【民間発行】 85.7万点 【公的機関発行】 54.6万点

(参考) 数字で見る国立国会図書館 <https://www.ndl.go.jp/ky/aboutus/outline/numerically.html>
 国立国会図書館年報 <https://www.ndl.go.jp/ky/publication/annual/index.html>

3

当館の所蔵資料について説明します。図書約1,200万点、雑誌・新聞約2,000万点、その他約1,500万点を所蔵しています。先程ご紹介した収蔵能力の数字より多いですが、これは収蔵能力が1冊3センチで計算した数字であるためで、雑誌・新聞等は3センチもないのが基本ですし、図書でも官庁向けの小冊子等薄いものも全部1冊、1点という扱いになっていますので、実際には、有体物としては、これぐらいの数を所蔵しています。

それから、無体物として、インターネット資料やオンライン資料も所蔵しています。まずインターネット資料としては、WARPという事業で、インターネット上のウェブサイトそのものを収集しているものが、約1.4万タイトル、22万件あります。タイトルに対して件数が多いのは、WARPでは、1つのウェブサイトを、例えば3か月おき、1年おきといった形で収集しているため、1つのサイトを複数回収集しているためです。データ量も非常に大きく、2ペタバイトぐらいあります。このためのストレージの確保も進めているところです。

オンライン資料というのは、電子書籍、電子雑誌等、電子的に出版された出版物を収集しているものです。最近本格的に、商業出版物の収集を始

めたところで、まだまだ数が少ないですが、民間発行のものが 85.7 万点、公的機関発行のものが 54.6 万点あります。

資料デジタル化の位置付け

・国立国会図書館ビジョン2021-2025

・5. 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。」

・資料デジタル化基本計画2021-2025

「デジタル化した資料を原資料の代替として提供することで原資料を保存し、検索の利便性や障害者を含むあらゆる人々の利用可能性を高め、また、関係機関等との有機的な連携により知識・文化の基盤を構築する」

国立国会図書館ビジョン2021-2025



当館における資料デジタル化の位置づけについて説明します。当館では、「国立国会図書館ビジョン 2021-2025」の 5 番目に、資料デジタル化の加速を掲げ、重点事業として取り組んでいます。その中で「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」としています。100 万冊の具体的な内訳等については後で説明させていただきます。

ビジョンを受けて策定した「資料デジタル化基本計画」では、「デジタル化した資料を原資料の代替として提供することで原資料を保存し、検索の利便性や障害者を含むあらゆる人々の利用可能性を高め、また、関係機関等との有機的な連携により、知識・文化の基盤を構築する」という目標を掲げています。

資料デジタル化：経緯

2000年	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（2008年までは2～4万点/年）
2009年	著作権法改正（第31条第2項新設）→予防的保存を目的としたデジタル化が可能に
2009～	大規模デジタル化事業実施（平成21年度、22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14万点等のデジタル化、著作権調査を実施
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	図書館向けデジタル化資料送信サービス（図書館送信）開始
2015年	災害関係資料のデジタル化（平成26年度補正予算） 震災・災害関係の図書約6万点、雑誌約2万点のデジタル化実施
2018年	著作権法改正（第31条第3項改正）→外国の図書館等への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 デジタル化内装の実験プロジェクト開始
2021年	国内刊行図書デジタル化（令和2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（国内刊行図書の範囲拡大や新聞の追加等） 著作権法改正（特に第31条第4項：絶版等資料の個人への送信）
2022年	国内刊行図書のデジタル化（令和3年度補正予算） 個人向けデジタル化資料送信サービス（個人送信）開始（印刷は2023年1月～）
2023年	国内刊行図書のデジタル化（令和4年度補正予算）

当館におけるこれまでの資料デジタル化の経緯についてお話させていただきます。以前、「近代デジタルライブラリー（近デジ）」とか、「貴重書画像データベース」という名前を聞いたことがある方がおられるかもしれません。実は近代という名前がついていたことには理由があります。

スライドに「2000 年に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（2008 年までは年に 2～4 万点）」とありますが、2008 年までは、明治期以降に刊行された資料のうち、著作権が切れていると確認できたものと、著作権者の許諾や文化庁長官の裁定を得られたもののみをデジタル化していました。明治期とそれに続く時代をスコープとしていたので「近代」なんですね。

それが、2009 年の著作権法改正により、当館は、資料原本の保存目的であれば、著作権者の許諾なくデジタル化を行えることとなりました。折しも成立した経済危機対策の平成 21 年度補正予算にデジタル化経費として 127 億円が計上されたことから、2009 年から 2011 年にかけて、大規模デジタル化事業を実施することとなりました。それまでは大体年に 1 億円程度の予算でデジタル化を行っていましたが、急に規模が大きくなったわけです。そのため、「経済危機対策大規模デジタル化実施本部」という体制を組みました。冒頭でもお話ししたとおり、臨時的体制です。

また、2012 年の著作権法改正により、当館が絶

版等の入手困難な資料を他の図書館等に送信し、送信先の図書館の利用者が閲覧したり、一部分の複写を受けることが可能になりました。それを受けて、当館では、2014年に図書館向けデジタル化資料送信サービス(図書館送信)を開始しました。送信対象となる絶版等資料を選定するプロセスは多段階にわたっていますが、本日は細かい説明は割愛させていただきます。

この後、平成26年度補正予算で災害関係資料のデジタル化経費が計上されたため、2015年に、震災・災害関係の図書6万点、雑誌2万点のデジタル化を実施しました。

2018年には、著作権法改正により、図書館送信サービスの送信先に、外国の図書館等が追加されました。それを受けて、2019年から外国の図書館等にも、図書館送信サービスを拡大しました。

それから、2019年に「デジタル化内製の実験プロジェクト開始」とあります。突然「内製」と言われても、何のことかなと思われるかもしれませんが。実はここまでのデジタル化は、基本的に外部の事業者へ委託して外注という形で実施していたのですが、この頃から、館内部でデジタル化を開始しました。その理由は後ほどご説明します。

2021年には、館全体のデジタル化を推進する戦略部署として私が所属する資料デジタル化推進室が設置されました。折しも、コロナ禍の状況下でデジタルコンテンツへのニーズが高まっていることを受け、当館所蔵資料のデジタルアーカイブ整備の経費として令和2年度補正予算に60億円が計上されたことで、大規模なデジタル化取り組むことになりました。その後、令和3年度補正予算、令和4年度補正予算でもデジタルアーカイブ整備の経費としてまとまった金額が認められ、国内刊行図書を中心に資料のデジタル化に取り組んでいるところです。

あわせて2021年には、「資料デジタル化基本計画2021-2025」を策定して、今後5年間でデジタル化の対象とする所蔵資料の範囲と優先順位、デジタル化の方法等についての考え方をまとめました。

また、2021年著作権法改正により、従来は当館から図書館等にしか送信できないこととなっていた絶版等資料について、一定の要件の下で、当館が利用者個人にも直接送信ができるようになりました。これを受けて、2022年5月から個人向けデジタル化資料送信サービス(個人送信)を開始し、2023年1月からは印刷(プリントアウト)もできるようになりました。

令和2～4年度補正予算の実施事項	
・ 令和2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 ※社会科学分野、人文科学分野の一部(大半が入手困難資料) 約30万点
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
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
・ 令和3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 ※人文科学分野、自然科学分野の一部(大半が入手困難資料) 約30万点
全文テキスト化の推進	視覚障害者等向け(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・国立国会図書館デジタルコレクションのリニューアル

2023年現在は、令和4年度補正予算により、国内刊行図書のデジタル化に引き続き取り組んでいます。規模としては約5,000万コマで、ページ換算すると大体1億ページ分ぐらいです。ここ3年間、同一の規模で継続して実施しています。

資料デジタル化基本計画2021-2025	
評価要素	- 唯一性・希少性 - 資料の利用機会の拡大(インターネット公開や図書館・個人送信が見込めるか) - 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	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) ・ 図書(2000年までに刊行されたもの) ※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、日系移民関係資料、地図、新聞<試行>)
利用提供	・ 「国立国会図書館デジタルコレクション」で提供 ・ 本文テキストデータの作成を推進し、全文検索を可能に ・ デジタル化済み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

(参考) 資料デジタル化基本計画2021-2025 https://www.ndl.go.jp/jp/preservation/digitization/digitization_plan2021.pdf

資料デジタル化基本計画の内容について説明します。デジタル化の対象とする資料の選定の評価

の要素としては、唯一性・希少性、資料の利用機会の拡大（インターネット公開や図書館・個人送信が見込まれるかどうか）ということを重視しています。また、資料の劣化状況、保存の緊急性、デジタル化への社会的・学術的なニーズ、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ということを評価の要素にして検討しています。

現在デジタル化の対象となっている資料群は、スライドに記載のとおりです。図書は 2000 年までに国内で刊行されたものを対象とし、その中で、内容等による取捨選択は行わず、形態等の問題でデジタル化が不可能なもの以外は基本的に全てデジタル化するという考え方です。時々、利用者や出版者の方から、特定の資料についてデジタル化する予定はありますかというようなお問合せをいただくことがあります。2000 年までに国内で刊行されたものかどうかという点で判断することになると回答しています。

雑誌は刊行後 5 年以上経過したものを対象にしていますが、特に学協会誌、国内の学協会が刊行・配布している雑誌について発行元から要望があった場合には、優先的にデジタル化を実施しています。

それから古典籍資料、録音・映像資料、博士論文、憲政資料等もデジタル化する方針です。これらは、それぞれの資料を所管している課が個別にデジタル化を実施しています。

デジタル化した資料は国立国会図書館デジタルコレクション（デジタルコレクション）で提供することとしています。著作権の状況等により、館内限定・図書館／個人送信・インターネット公開の 3 つの公開範囲で提供しています。本文のテキストデータの作成を推進し、全文検索を可能にしています。

デジタル化済みの資料原本は原則として利用停

止にしています。図書と雑誌については、先にご説明したとおり、関西館に移送し保存しています。

令和2～4年度の補正予算の背景

- ・コロナ禍を契機としたリモートアクセス要望の高まり
・[図書館休館対策プロジェクト](#)、[日本歴史学協会](#)、[日本出版者協議会](#) など
- ・内閣府知的財産戦略本部「[知的財産推進計画2020](#)」（2020年5月）
・「絶版等により入手困難な資料をはじめ、図書館等が保有する資料へのアクセスを容易化するため、図書館等に関する権利制限規定をデジタル化・ネットワーク化に対応したものとするについて、研究目的の権利制限規定の創設と併せて、権利者の利益保護に十分に配慮しつつ、検討を進め、結論を得て、必要な措置を講ずる。」（p68）
- ・自由民主党政務調査会知的財産戦略調査会「[国立国会図書館の図書等のデジタル化についての提言](#)」（2020年9月）
・2000年までに刊行された図書165万点について、5年間でデジタル化を。

先ほど触れた令和2年度からの補正予算の背景として、コロナ禍を契機とした各所からのリモートアクセス要望の高まりがありました。図書館が休館していても資料が見られるようにする必要性が提起され、これに応えるということで、個人送信とセットで実現したと考えていただくと分かりやすいかと思います。

内閣府知的財産戦略本部が取りまとめている「知的財産推進計画 2020」の中でも、「絶版等により入手困難な資料をはじめ、図書館等が保有する資料へのアクセスを容易化するため、図書館等に関する権利制限規定をデジタル化・ネットワーク化に対応したものとするについて、研究目的の権利制限規定の創設と併せて、権利者の利益保護に十分に配慮しつつ、検討を進め、結論を得て、必要な措置を講ずる」と記載されました。

更に、自由民主党の政務調査会知的財産戦略調査会「国立国会図書館の図書等のデジタル化についての提言」でも、2000 年までに刊行された図書 165 万点について、5 年間でデジタル化を実施すべきと記載されました。

令和2～4年度補正予算の実施事項

・ 令和2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 ※社会科学分野、人文科学分野の一部（大半が入手困難資料）約30万点
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
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

・ 令和3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 ※人文科学分野、自然科学分野の一部（大半が入手困難資料）約30万点
全文テキスト化の推進	視覚障害者等向け（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・国立国会図書館デジタルコレクションのリニューアル

令和2年度から4年度までの補正予算によるデジタル化での実施事項について説明します。

まず、令和2年度補正予算では、1987年までに整理された国内刊行図書約30万点のデジタル化を実施しました。東京本館書庫の特定の区画の資料をまとめてデジタル化した結果、分野としては、社会科学、人文科学のものが中心となりました。

2点目、デジタル化設備の整備として、ブックスキャナ等を購入し、内製デジタル化を実施するための作業室を館内に開設しました。

3点目として、全文テキスト化の取組を2件、行いました。1つは、デジタル化済みの資料についてOCRによる全文テキスト化を行うもので、外注により実施し、高い精度で文字認識されたテキストデータが納入されました。もう1つは、これからデジタル化する資料を全文テキスト化できるようOCRの研究開発を行うもので、同じく外注により実施し、高い精度で文字認識できるOCRプログラムが納入されました。

それから4点目として、ストレージの拡張や、デジタルコレクションのシステム改修を行いました。これは、大量のデジタル化を実施するためにはストレージの増強やシステム面での対応が必須となったためです。

令和3年度補正予算では、前年に引き続き、国内刊行図書のデジタル化を実施しました。また、全文テキスト化の推進としては、前年度に開発したプログラムについて、視覚障害者等向けに、よ

り正しく読み上げられるよう改善を行いました。この成果であるNDLOCRという処理プログラムを使って、デジタル化資料の全文テキスト化を進めようとしているところです。

さらに、電子書庫機能の拡張等として、再度必要分のストレージの増強を実施しました。

2023年1月には、デジタルコレクションのリニューアルも行いました。これにより、令和2年度補正予算により作成した全文テキストデータが検索できるようになりました。

令和2～4年度補正予算の実施事項

・ 令和4年度補正予算（第2号）

項目	概要
図書資料のデジタル化	国内刊行図書のデジタル化 ※本館書庫1階採集資料（1988～1995）等のデジタル化、約40万点 ※2000年までに刊行された図書付録等のデジタルデータ抽出、約1.5万点
全文テキスト化の推進	逐次刊行物に対応した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 図書付録等のデジタルデータの利用環境の整備 ※デジタルデータ抽出のための装置の開発に向けた技術調査、エミュレーション（過去の動作環境の構築）に係る技術検証等



写真1 国立国会図書館蔵書での資料のデジタル化作業
写真2 国立国会図書館蔵書での資料のデジタル化作業
写真3 国立国会図書館蔵書での資料のデジタル化作業
写真4 国立国会図書館蔵書での資料のデジタル化作業
写真5 国立国会図書館蔵書での資料のデジタル化作業
写真6 国立国会図書館蔵書での資料のデジタル化作業
写真7 国立国会図書館蔵書での資料のデジタル化作業
写真8 国立国会図書館蔵書での資料のデジタル化作業
写真9 国立国会図書館蔵書での資料のデジタル化作業
写真10 国立国会図書館蔵書での資料のデジタル化作業

今年度も、令和4年度補正予算を基に、引き続き国内刊行図書のデジタル化を実施しています。3年目となり作業の形が固まってきたということもあり、点数が40万点とちょっと増えています。

それから、2000年までに刊行された図書付録等1.5万点のデジタルデータ抽出を実施しています。昔よく、本にCD-ROMが付いているものがあつたかと思いますが、その類の資料です。2000年ぐらいのCD-ROMの中には、現在の環境ではもう読めなくなってしまっているものもありますので、デジタルデータを抽出して、デジタルコレクションに載せていくことを進めています。

また、全文テキスト化の研究開発と書いてありますが、逐次刊行物に対応したOCR処理プログラムの開発を行っています。

更に、電子書庫機能の拡張等として、ストレージの増強や、図書付録等のデジタルデータの利用環境の整備を行っています。CD-ROMのデータを抽

資料デジタル化に係る業務

【デジタル化作業】

①資料選定 → ②スキャン → ③品質検査 → ④データ登録 → デキスト化 → 公開

デジタル化とは切り離して業務

【画像とメタデータ】

(1) デジタル
画像、(目次)、管理データ
HDD LTO
(2) データ登録
保存
国立国会図書館サーチ
(3) メタデータ
書誌データ
閲覧室
(4) データ利用
インターネット公開
図書館・個人送付
国立国会図書館デジタルコレクション

納品
外注
自動登録
内蔵

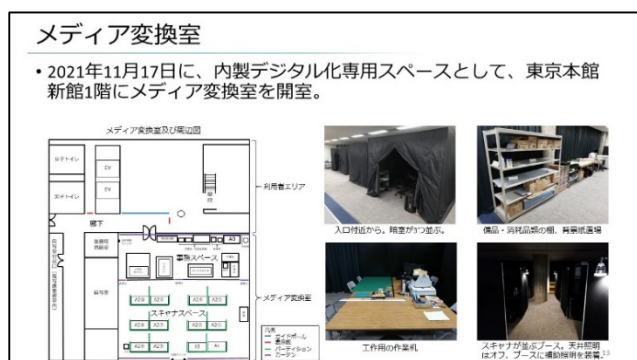
デジタル化済みデータの流れは、スライド下部に「画像とメタデータ」として図示しているとおりで、外注と内製とで流れが若干異なります。

ら補完する形で作成しています。

デジタル化データの登録・公開

デジタル化データ（図の左にある画像、メタデータ、目次データ等）は、外注・内製の場合とも、まず、当館内部で開発した検品用のシステムに登録されます。そこで検品を行い、品質に問題がなければ、デジタルコレクションのシステムに転送します。デジタルコレクションのシステムは業務系と提供系の2つに分かれており、まず業務系の方で確認した後、皆さんがご覧になっている提供系、つまりデジタルコレクションの画面の方で公開するという流れになっています。

– 7 –



内製デジタル化の専用スペースとして使用している、東京本館のメディア変換室の様子をご紹介します。

室内には、入口付近に暗室がいくつか並んでいて、奥は全体が暗幕で覆われ、スキャンニングのためのブースが設置されています。スキャナの種類としては、オーバーヘッド型のブックスキャナがメインで、A3サイズ対応が1台、A2サイズ対応が10台、A1サイズ対応が1台あります。この他に、A0サイズ対応のフラットベッドスキャナ、A2サイズ対応のV字型スキャナ、マイクロフィルム用スキャナ、ADFスキャナ等もあります。ADFというのは、自動紙送り式のもので、背を切ることができる資料や、バインダー綴じの資料等に使用しています。

テキスト化の推進

- 2019年1月施行の改正著作権法（第30条の4、第47条の5）により、画像データからのテキストデータ作成と所在検索サービスが実施可能に
- デジタル化資料約247万点（2020年12月時点）の画像から、OCR処理により本文のテキストデータを作成（文字認識率（図書・雑誌）は平均96.86%）
- サービスへの活用
 - 2022年12月 国立国会図書館デジタルコレクションで全文検索用データとして利用開始
 - 2023年3月 視覚障害者等用データ送信サービス（みなサーチβ版）で提供開始

約247万点の内訳

種別	内容	点数（冊数）
図書	1968年までに受け入れた図書 - 震災・災害関係資料の一部（1969年以降受入分も含む）	97万点
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）	132万点
博士論文	1988年（一部）～2000年に送付を受けた論文	15万点
その他	官報等	2万点
計		247万点

※ 2021年以降にデジタル化した資料も、今後、順次テキスト化予定

テキスト化の状況について説明します。

2019年1月施行の改正著作権法第30条の4により、画像データからのテキストデータ作成と所在検索サービスが実施可能になったことを受け、テキストデータを作成することになりました。

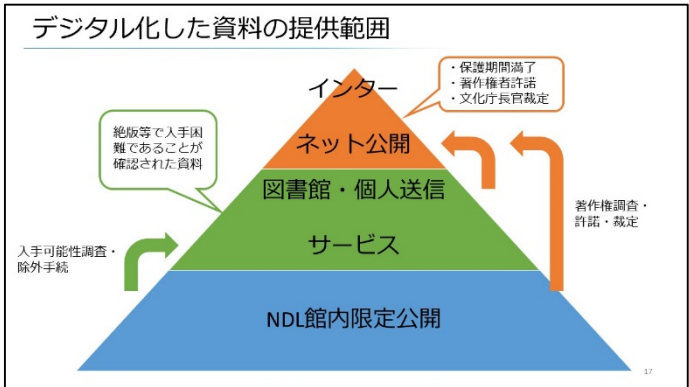
令和2年度補正予算により、デジタルコレクションで2020年12月時点に公開されていたデジタル化資料約247万点の画像から、OCR処理によって、本文のテキストデータを作成しました。図書・雑誌の文字認識率は平均96.86%です。

これらのテキストデータのサービスへの活用としては、先にもお話したとおり、2022年12月にデジコレで全文検索用のデータとして、スニペット表示されるような形で提供しています。

さらに、2023年3月に、視覚障害者等用データ送信サービス「みなサーチβ版」でも提供を開始しています。こちらでは、視覚障害者等に対し、テキストそのものを提供するというような形になっております。

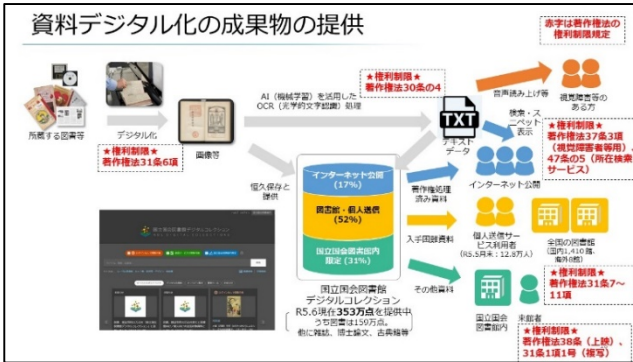
247万点の内訳は、2020年までにデジタル化した資料であり、1968年までに受け入れた図書等が中心です。その後にデジタル化したものはまだOCR処理が終わっていないため全文検索ができないので、御注意いただきたいと思います。

雑誌等は比較的新しい年代までデジタル化を行っており、OCR処理が終わっているものの比率が高いです。全文検索で、新しい語は雑誌ばかり引っかかってくる、というのはこういう理由により



ます。

デジタル化した資料の提供範囲は、インターネットからログインなしで閲覧可能なもの、図書館・個人送信サービスで閲覧可能なもの、館内限定公開の3通りです。館内限定公開のものから、絶版等で入手困難であることが確認されれば図書館送信・個人送信になり、著作権保護期間満了のものはインターネット公開ということになります。



こちらの図は資料デジタル化の成果物の提供フローと根拠となる法令についてまとめたものです。デジタルコレクションでの提供状況は、インターネット公開が17%、図書館・個人送信が52%、館内限定が31%となっています。

視覚障害者等用のデータ送信に関しては著作権

法第37条第3項において認められています。ただし、スクリーンリーダーによる読み上げに対応している電子書籍が市販されているもの等は除きます。

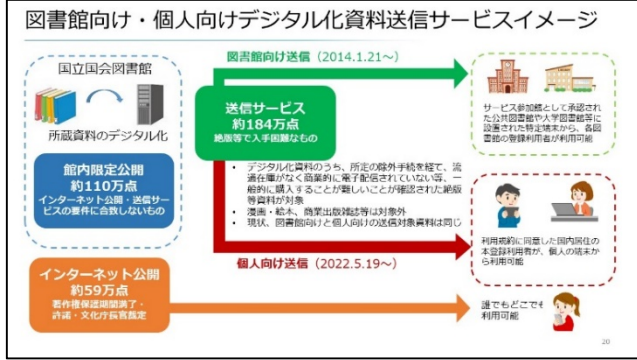
それから、検索結果のスニペット表示や所在検索サービスについては、著作権法第47条の5によって可能になりました。

資料デジタル化：提供状況（2023年6月現在）

資料種別	インターネット公開資料	図書館・個人送信対象資料	NDL館内提供資料	合計	年代・概要
図書	36万点	85万点	38万点	159万点	明治期以降、1987年までに整理された図書等
雑誌	2万点	82万点	52万点	136万点	国内発行雑誌のうち、変化した雑誌や学術雑誌等（発行後5年以上経過したもの）
古書籍	8万点	2万点	-	9万点	寄書集・準寄書集、江戸期以前の和漢書等
博士論文	1万点	14万点	2万点	17万点	1989～2000年度に国内の大学から送付を受けた学位論文
新聞	-	-	8万点	8万点	明治期以降に発行された新聞のうち、資料保存のためにデジタル化したもの等
録音・映像関係資料	-	-	1万点	1万点	カセットテープ、ソノシート、レーザーディスク、録音、手稿等
その他	11万点	2万点	10万点	23万点	官報、憲政資料、地図、プランニング、日本占領関係資料、歴史的音源等
合計	59万点	184万点	110万点	353万点	

(参考) デジタル化資料提供状況 <https://www.ndl.go.jp/jp/preservation/digitization/index.html#situation>

こちらの表は、2023年6月時点の資料種別ごとの提供状況をまとめたものです。図書はインターネット公開、図書館・個人送信の対象資料が結構あります。雑誌はインターネット公開の資料は少ないのですが、図書館・個人送信の資料は結構多く、図書と同じくらいの点数になっています。このように、図書館・個人送信サービスを利用いただければ、古い資料に関しては多くのものをご覧いただけるようになっています。



図書館・個人送信サービスのイメージですが、デジタル化資料のうち、所定の除外手続を経て、流通在庫がなく商業的な電子配信もなされていないなど、一般的に購入することが難しいと確認された絶版等資料が対象となります。

所定の除外手続としては、当館が入手可能性調査を行った後、送信候補とした資料のタイトル等の一覧を一定期間公表し、出版関連団体、出版者、著作（権）者等からの除外の申し出を受け付けます。除外基準に該当したものを除き、公開するという流れになっております。

なお、漫画、絵本、商業出版雑誌等は対象外となっています。また現状、図書館向けと個人向けの送信対象資料は全く同じになっています。

図書館向けデジタル化資料送信サービス	
(経緯)	2009年 著作権法改正（第31条第2項新設）→国立国会図書館の保存目的のデジタル化が可能に →衆・参の附帯決議や文化審議会中間総括を受け、関係府省でデジタル化資料の利活用について議論 →第一段階として、市場における入手が困難な出版物に限定して公共図書館等への送信を行う方針に 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 →参議院附帯決議も踏まえ、市場への影響に留意した運用を関係府省協議会で協議、合意事項取りまとめ 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
(概要)	対象資料 デジタル化資料のうち、絶版等の理由で入手困難なもの（★後述） 対象施設 ・ 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 ※ 寄託または司書に相当する職員が必要 ・ 参加館となるには、国立国会図書館による要件確認・承認が必要 利用方法 ・ 参加館の登録利用者が、参加館内の特定端末から国立国会図書館デジタルコレクションにアクセス ・ プリントアウト可能（一部図書館は閲覧のみ） 運用 資料デジタル化及び利用に係る関係者協議会において取りまとめた「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」に基づく
(参考)	図書館向けデジタル化資料送信サービス https://www.ndl.go.jp/jp/use/digital_transmission/index.html 資料デジタル化及び利用に係る関係者協議会 https://www.ndl.go.jp/jp/preservation/digitization/consent.html#anchor01 23

図書館送信サービスについて、これまでの経緯を改めて御説明します。2009年の著作権法改正により当館の保存目的のデジタル化が可能になりました。その成果物の有効活用を図るようという衆議院・参議院の附帯決議を受け、関係府省で利活用についての議論が行われ、第一段階として、入手困難な出版物に限定して、公共図書館に送信を行うことにしようということになりました。

これを受けて再度著作権法が改正され、図書館送信が可能になりました。運用方法について関係者と協議を行った後、2014年にサービスを開始し、2018年、19年には外国の図書館等にも拡大するということになりました。

個人向けデジタル化資料送信サービス	
(経緯)	2020年 コロナ禍における研究者・学生等からの米館等利用できる図書館サービスへのニーズの高まり →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」 2021年 著作権法改正（第31条第4項はか）→個人への絶版等入手困難資料の送信が可能に →関係者協議会で協議、合意事項を取りまとめ 2022年 個人向けデジタル化資料送信サービス開始（5月19日～、閲覧のみ） 2023年 プリントアウト機能提供開始（1月18日～）
(概要)	対象資料 デジタル化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）（★後述） 対象者 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 ※ 氏名・居住先・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、家族・郵送・オンラインにより受付。海外居住者への送信も検討予定。 利用方法 ・ 個人の端末から利用可能ID・PWでログインして国立国会図書館デジタルコレクションにアクセス ・ 閲覧（ストリーミング）とプリントアウトが可能 ・ 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 運用 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において取りまとめた「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事項」に基づく
(参考)	個人向けデジタル化資料送信サービス https://www.ndl.go.jp/jp/use/digital_transmission/individuals_index.html 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 https://www.ndl.go.jp/jp/preservation/digitization/consent.html#anchor02

個人送信サービスは、コロナ禍における図書館の休館等を契機として、研究者・学生等から来館せずに利用できる図書館サービスへのニーズが高まったということを受け、政府の「知的財産推進計画2020」や「図書館関係の権利制限規定の見直しに関する報告書」等において、検討が始まったものです。著作権法が改正され、運用方法について関係者と協議を行った後、2022年に閲覧サービスを開始し、2023年にはプリントアウトの機能も提供を開始しました。

送信対象資料	
・ (特定) 絶版等資料に係る著作物(著作権法第31条第7項及び8項)	
・ 具体的な送信対象資料は、資料デジタル化及び利用に係る関係者協議会（デジタル化した資料の利用提供方法等について著作権者・出版者団体、大学、図書館等、関係の団体や機関と協議する場）で取りまとめた「国立国会図書館のデジタル化資料の図書館等への限定配信に関する合意事項」に基づき、以下のとおり運用。個人送信にも同様に適用される。	
・ デジタル化資料のうち、流通在庫がなく商業的に電子配信されていない等、一般的に図書館等において購入が困難である資料を送信対象とする。	
・ デジタル化した図書、雑誌、博士論文を送信候補資料とし、送信対象を入手困難な資料に限定するため、3段階の除外手続（入手可能性調査、事前除外、事後除外）を行う。	
・ 著作物が市場（オンデマンド出版及び電子書籍を含む。）において流通している場合（おおむね3か月を目安として流通予定であることを確認した場合を含む。）は送信対象から除外。このほか、著作権等管理団体事業者の管理著作物、著者本人から要請があった著作物も除外。	
・ 漫画・絵本、商業出版雑誌（関係者と合意が得られたものを除く）は送信留保。	
(参考)	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 https://www.ndl.go.jp/jp/preservation/digitization/consent.html#anchor02.pdf 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事項 https://www.ndl.go.jp/jp/preservation/digitization/kojinshoshin_agreement.pdf 図書館向けデジタル化資料送信サービス（図書館送信）に係る除外手続 https://www.ndl.go.jp/jp/preservation/digitization/distribution.html 23

送信対象資料については、著作権法では「(特定) 絶版等資料に係る著作物」とされています。具体的には、資料デジタル化及び利用に係る関係者協議会、これはデジタル化資料の利用提供方法等について、著作権者・出版団体、大学、図書館等、関係の団体や機関で協議する場となっているものですが、そこで取りまとめた「国立国会図書館のデジタル化資料の図書館等への限定配信に関する合意事項」に基づいて、スライドに記載のとおり運用しています。個人送信にも同様に適用されています。

一つずつ見ますと、まずは、デジタル化資料のうち、流通在庫がなく、商業的に電子配信されていない等、一般的に図書館等において購入が困難である資料が送信候補となります。デジタル化した図書、雑誌、博士論文から、送信対象を入手困難な資料に限定するため、3段階の除外手続（入手可能性調査、事前除外、事後除外）を行うということになっています。また、著作物が、市場（オンデマンド出版及び電子書籍を含む。）において流通している場合（おおむね3か月を目安として流通予定であることを確認した場合を含む。）は送信対象から除外すること、このほか、著作権等管理団体事業者の管理著作物、著者本人から要請があった著作物も除外することとなっています。

それから、漫画・絵本、商業出版雑誌等に関しては、送信が留保されています。

公共・大学図書館からのデジタル化データの寄贈

公共図書館等における所蔵資料のデジタル化

- ・ 絶版等で入手困難で貴重な資料について、資料保存のためのデジタル化が可能
⇒ 郷土資料等のデジタル化の著作権処理は不要

国立国会図書館による送信サービスの拡充

- ・ NDL未所蔵資料について、所蔵機関がデジタル化したデータを寄贈
- ・ 図書館・個人送信サービスで送信することが可能
- ・ 寄贈館のメリット：郷土情報等の発信力強化・テキスト化

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

寄贈：第1回文化遺産国際会議（平成27年11月1日）
第13・年度文化遺産国際会議・第13回文化遺産国際会議の経緯について
<http://www.ndl.go.jp/cultural/heritage/13th/13th.html>
2015年11月1日（平成27年7月7日）
<https://www.ndl.go.jp/cultural/heritage/13th/13th.html>

また、当館では、公共・大学図書館等からのデジタルデータの寄贈も受け付けております。この「所蔵資料のデジタル化」というのは、各図書館の所蔵資料のうち、絶版等により入手困難で貴重な資料については、資料保存のためのデジタル化を許諾なく行うことが認められていることを指します。これは著作権法31条の1項2号に基づくものです。

そのようにデジタル化されたもののうち、当館に所蔵がない資料について、データを寄贈いただければ、デジタルコレクションで公開し、図書館・個人送信サービスで送信することが可能になりま

す。

寄贈館側のメリットとしては、郷土情報等の発信力の強化や、全文テキスト化による検索可能化などがあります。その代わり、デジタル化は各機関でやっていただく形となっています。詳細はスライド下部のリンク「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御協力のお願い」をご参照ください。

読書バリアフリーの推進

- ・ 視覚障害者等の方を対象として、デジタル化資料からテキスト化したデータの提供を開始（2023年3月28日～）
- ・ 視覚障害者等用データ送信サービスに登録した視覚障害者等個人の方や同サービスの送信承認館は、国立国会図書館障害者用資料検索（愛称：みなサーチ）β版（2023年3月28日～試験公開中）から、全文テキストデータ約247万点をダウンロードして利用可能。

みなサーチβ版
<https://mina.ndl.go.jp/>

※正式版は、2024年1月公開予定。

（参考）国立国会図書館障害者用資料検索「みなサーチ」β版にて、視覚障害者等の方のデジタル化資料の全文テキストデータをダウンロードできるようになりました（ニュース）
https://www.ndl.go.jp/news/ty2022/230328_01.html

それから、読書バリアフリーの推進に向けた取組も実施しています。2023年3月から、視覚障害者等の方を対象として、デジタル化資料からテキスト化したデータの提供を開始しました。視覚障害者等用データ送信サービスに登録した視覚障害者等個人の方や、同サービスの送信承認館は、国立国会図書館障害者用資料検索、愛称「みなサーチ」のβ版から、全文テキストデータ約247万点をダウンロードして利用可能となっています。

「みなサーチ」は実のところ、ちょうど2、3日前に広報を始めた、当館の統合的オンラインサービスというものの一環をなすものです。現在は、国立国会図書館オンライン（NDL オンライン）と国立国会図書館サーチ（NDL サーチ）というシステムが別個に存在しており、NDL オンラインは当館の所蔵資料を検索するもの、NDL サーチは全国の図書館等の所蔵資料を検索するもの、というすみ分けになっています。NDL サーチの検索結果からでもNDL オンラインの方に最終的に飛べるので、NDL サーチを使っている方もいらっしゃると思い

ますが、システムが2つあるというのが分かりにくいですし、NDL サーチでヒットしたものは全部国立国会図書館に所蔵されていると誤解される場合がすごく多いようです。これら2つのウェブサービスを統合する予定なのですが、その統合的オンラインの下にあるものの一つが、「みなサーチ」ということです。



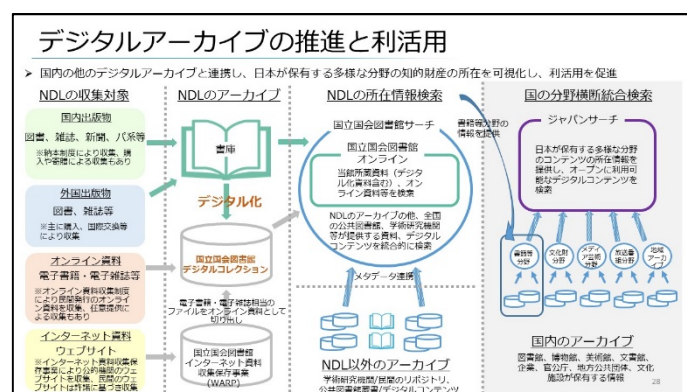
デジタル化資料の利活用の工夫としては、国立国会図書館所蔵の浮世絵、雑誌、図書などから、選りすぐりのビジュアル資料を紹介した「NDL イメージバンク」や、近代日本の形成に影響のあった、政治家、官僚、軍人、実業家、学者、文化人などの肖像写真を紹介する「近代日本人の肖像」、いろいろなテーマのミニ電子展示「本の万華鏡」等を作成、公開しています。

特に「NDL イメージバンク」や「近代日本人の肖像」には、著作権フリーのものしか入っていないので、これらを使って、各図書館でも、色々なことができますと思います。また、利活用を図る一つのサンプルと思っていただいてもいいかもしれません。



さらに別の、利活用を図る取組としては、ジャパンサーチもあります。ジャパンサーチでは、日本国内の様々な分野のデジタルアーカイブのメタデータを横断的に検索することができます。Google のように全てのウェブサイトを検索するというのではなくて、ジャパンサーチと連携している機関のデジタルアーカイブの中から、どういうものがあるか調べたり、調べたコンテンツをまとめて自分のページを作ったりということができるプラットフォームです。

ジャパンサーチには、多くの機関のデジタル化情報が蓄積されており、他機関の情報と一緒に提供できれば、さらに役に立つというものもあります。こうして、当館のデジタルコレクションや全国書誌のデータも、ジャパンサーチと連携するようにしています。



この図は、ちょっとごちゃごちゃしていますが、デジタルアーカイブの推進と利活用ということで、左から右に行くにつれて、国全体を俯瞰していくような構成になっているスライドです。

一番左は、当館の収集対象を記載しています。国内出版物のほか外国出版物、オンライン資料や、ウェブサイト等のインターネット資料の収集も行っています。ちなみに外国出版物については、選書して購入しているので、他のものと比べて量的には少ないです。

当館がデジタル化した資料や収集したオンライ

ン資料は、全て、デジタルコレクションに登録されます。次に、そのメタデータを「国立国会図書館サーチ」に登録します。ここには当館以外の機関、全国の図書館等のデータも連携して入ってきます。ジャパンサーチはこれに加えて、他の連携機関のアーカイブを検索できるという枠組みになっています。

ジャパンサーチの使い方は、なかなかイメージが難しいとは思いますが、レファレンス等でピンポイントに何かを探そうと思って使うというよりは、既にデジタル化されている資料の見せ方、可能性を広げるという使い方の方が合うように思います。どちらかというと、「何か面白い切り口で見せられないか」とか「こんな関連資料もあったんだ」というように、ちょっと遊び心を持って、紹介する使い方がいいのかもしれないなど、個人的には思っております。

次世代に向けて：NDLラボの取組

- 次世代の図書館システムの開発に資する要素技術の実証実験を行うためのウェブサイト「NDLラボ」を運営。
 - 実験の成果を公開することにより、我が国の次世代図書館サービスの利便性向上に貢献することを目指す。
- 提供中の実験サービス例
- 次世代デジタルライブラリー <https://lab.ndl.go.jp/dl/>
全文テキスト検索や画像検索が可能なデジタルライブラリーであり、AI等技術を活用した図書館サービスの実験場。令和4年度古典籍資料のOCRテキスト化実験の成果物である古典籍資料約8万点分全文テキスト検索可能。
 - NDL Ngram Viewer <https://lab.ndl.go.jp/ngramviewer/>
OCRによって作成されたテキストデータから、出版年代ごとの単語及びフレーズ出現頻度を可視化・列挙することができるサービス。

それから次世代に向けた取組として、NDL ラボというウェブサイトで、次世代の図書館システムの開発に資する要素技術の実証実験を行っています。実験の成果を公表することによって、我が国の次世代図書館サービスの利便性向上に貢献することを目指しています。

NDL ラボで提供中の実験サービスとしては、次世代デジタルライブラリーがあります。ここで全文検索や画像検索など、新しい技術の実験を行っています。2020年までは、ここでしかテキスト検索ができませんでしたが、現在はデジタルコレク

ションでも提供しています。また、画像検索についても、ここで実験したものをデジタルコレクションに実装しています。

それから「NDL Ngram Viewer」というものがあります。OCRによって作成されたテキストデータを基に、単語、フレーズの出現頻度を出版年代ごとに可視化・列挙することができるものです。これも、新しい技術を実験していこうということの一環で生まれた成果を公開しているものです。

デジタル資料の長期保存

- 国立国会図書館が収集した資料は、現在と未来の読者のために、国民共有の文化的資産として永く保存され、日本国民の知的活動の記録として後世に継承される。
 - 収集したCD・DVD等の「パッケージ系電子出版物」やインターネット上で発信される電子情報も、長期的に保存し利用を保証する必要があるが、媒体のぜい弱性、再生装置の入手困難化、再生ソフトウェア等技術の陳腐化等、課題は多い。
 - 所蔵資料をデジタル化したデータの保存も今後の課題。長期利用のためには維持管理が必要。
 - 「国立国会図書館デジタル資料長期保存基本計画2021～2025」（2021.3）を策定。最近の取組は、以下のとおり。
- パッケージ系電子出版物（CD/DVD、USBメモリ、フロッピーディスク等）のマイグレーションによる保存対策の開始
 - デジタル化資料の保存用データ（光ディスク約18万枚等）の状態調査と、LTOへの媒体移行の開始
 - 各種の調査を実施
 - ⇒「デジタル資料の長期保存に関する国内機関実証調査報告書」「光ディスクの状態調査の手法に係る基礎的調査報告書」公開（2022.6）
 - ⇒「フロッピーディスクの長期保存対策に関する調査報告書」公開（2023.4）
- （参考）電子情報の長期的な保存と利用 <https://www.ndl.go.jp/jp/preservation/dlib/index.html>

最後に、デジタル資料の長期保存についてお話をさせていただきます。

当館では、資料保存のためのデジタル化を行っているため、デジタル化が済んでも、紙の資料は捨てずに保存します。また、CD・DVD等のパッケージ系電子出版物やインターネット上で発信される電子情報も、長期的に保存し利用を保証する必要がありますが、媒体のぜい弱性、再生装置の入手困難化、再生ソフトウェア等技術の陳腐化等、課題は多い状況です。

例えば、WARPでアーカイブされているものを見ても、15年、10年前のインターネットの動くページは、どのサイトもだいたいFlashを使っていたので、今は閲覧できなくなっていますよね。昔のものをそのまま取っておけば見られるというわけにもいかない。

紙は比較的媒体の寿命が長いのですが、フロッピーディスクなどデジタル資料は早期に利用できなくなってしまうこともあるので、これらの長期

保存について喫緊の課題として取り組んでいます。マイクロフィルム、ビデオテープ、カセットテープ等のアナログ媒体も同様で、再生機器の維持、媒体の補修等が難しくなっています。緊急度の高い媒体から順次、デジタル化やマイグレーションに着手する必要があります。

所蔵資料をデジタル化したデータの保存も今後の課題です。長期利用のためには、維持管理が必要とありますが、当館では、具体的には LT0（磁気テープ）とクラウドとを比較検討しました。当館は保存するデータの量がとても多く、クラウドにデータを出し入れするコストが非常に高くなることから、LT0 を選択しました。

LT0 については、業界団体が、温度 60 度・湿度 80% の環境に 3 か月間置いた後、読み取る実験をしています。これは 50 年分に相当する負荷ということであり、これでも読み取れたという結果が出たそうですが、全ての LT0 が 50 年後も利用できるかと保証されているわけではありませんし、その前に再生装置の方が終わりを迎えると思われます。このため、定期的に媒体変換をしていかななくてはならないものと認識しています。

こうした状況を踏まえ、当館では「国会図書館デジタル資料長期保存基本計画 2021-2025」を策定し、パッケージ系の電子出版物の CD、DVD、USB、フロッピー等のマイグレーションによる保存対策を始めています。また、デジタル化資料の保存用データが格納されている光ディスク、主に Blu-ray の状態調査や、LT0 への媒体変換も進めています。

それから、デジタル資料の長期保存に関する調査等を行って、報告書などを公開しています。一番新しいところだと、「フロッピーディスクの長期保存対策に関する調査報告書」を公開しました。

最後に参考として、当館のホームページ等に

色々に掲載している資料がありますので、ご参照いただければと思います。

(参考)

- JEPASEMINAR 国立国会図書館のデジタルシフト「ビジョン2021-2025」
<https://www.jepa.or.jp/sem/20210421/>
- JEPASEMINAR ジャパンサーチ(JapanSearch) 概要紹介
<https://www.jepa.or.jp/sem/20210720/>
- JEPASEMINAR 国のデジタル情報基盤の充実に向けてー国立国会図書館「ビジョン2021-2025」の取組から
<https://www.jepa.or.jp/sem/20220202/>
- JEPASEMINAR 国立国会図書館デジタルコレクションのリニューアル
<https://www.jepa.or.jp/sem/20230228/>
- NDL遠隔研修 「資料デジタル化の基礎」
https://www.ndl.go.jp/jp/library/training/remotedigi_basic_2019.html
- NDL遠隔研修 「デジタル化資料の権利処理と利活用」
https://www.ndl.go.jp/jp/library/training/remotedigi_copyright_2019.html
- NDL遠隔研修 「デジタル資料の長期保存に関する基礎知識」
https://www.ndl.go.jp/jp/library/training/remotedigi_digitalpreservation.html
- 第24回図書館総会、国立国会図書館主催フォーラム「#NDL全文使ってみた〜「次世代デジタルライブラリー」&「NDL Ngram Viewer」」
<https://lab.ndl.go.jp/event/fff2022/>

私からのお話は以上になります。どうもありがとうございました。(拍手)

《質疑応答》

○参加者 A 氏 今日は、貴重なとても興味深いお話をありがとうございました。一つ教えていただきたいのですが、読書バリアフリー推進のための「みなサーチ」ではデジタル資料のテキスト化をしたものを提供していらっしゃると思いますが、現在テキスト化の精度が 96.8%とのことでした。近デジを見たときに、古い資料は、かなり「ゲタ」と呼ばれる変換できなかった文字の代替記号(≡)が入っていたのですけれども、この「みなサーチ」を提供する際には、このゲタ文字を補完したりしていらっしゃるのでしょうか。音声読み上げのとき、その部分が読めないのではないかと思います、お尋ねします。

○中島氏 これは、「みなサーチ」に限らず、OCR後のデータを人手で直しているかどうかということかと思いますが、現状では都度の修正は行っておりません(別途製作依頼は可能)。ご要望があるだろうというのは認識しておりますが、膨大な点数があり容易に対応できるわけではないこともあり、未校正であることをご理解いただいた上で利用いただく形としています。

○事務局 2000 年までに刊行された本を 2025 年

までにデジタル化していくという計画があるとのことでしたが、そうすると 25 年前までに刊行された資料は、ある程度デジタルで利用できるという状態かと思います。なかなか、将来的な見通しは難しいと思うのですけれども、今後、ある時点で、どのくらい前までの資料が利用できるかという期間というのは、長くなったり短くなったりする可能性があるでしょうか。

○中島氏 当館では、2000 年までの図書約 170 万点のデジタル化を目指すとして、補正予算を要求し、措置されてきました。これは図書に限ったものであり、例えば雑誌については、直近 5 年より前に刊行されたものを、少しずつ、別の予算でやっています。

その先どうするのが課題ではあるのですが、図書約 170 万点のデジタル化が実現できたとしたら、あと 2、3 年ぐらいで、2000 年までの国内刊行図書に関しては、ほとんどが、デジタルコレクションで中身まで検索できるようになります。そこは覚えて帰られると良いのかもしれません。

○参加者 B 氏 大学図書館に勤めています。本学では、明治期の、特に医学系の資料をいくつか持っていて、おそらく国会図書館では持ってないものもあるのかなと思っています。デジタルアーカイブということを考えるに当たって、2 つ方向性があると思うのですが、国会図書館としては、デジタル化した資料のデータを寄贈という形で提供してもらうのと、独自のデジタルアーカイブで公開した上で、ジャパンサーチで連携してもらうのと、どちらのほうが良いのだろうかというのを聞きたいです。

○中島 当館としては、どちらの方向性もありえる話です。一般的には、一箇所ですべて見られることが重要ではないか、という気がします。

ただ、著作権のある資料については、個別のデ

ジタルアーカイブだと、複写やダウンロードも含めた利用提供の部分が複雑になるかと思います。

国立国会図書館については、送信サービスでの提供について法的な根拠がありますので、おそらく、寄贈していただく方が使いやすくなるのではないかと思います。

本日は、ありがとうございました。(拍手)