

## 第五期国立国会図書館科学技術情報整備基本計画の実施状況

「第五期国立国会図書館科学技術情報整備基本計画」は、第 13 回科学技術情報整備審議会(令和 3 年 1 月 13 日開催)において国立国会図書館長に提出された「第五期国立国会図書館科学技術情報整備基本計画策定に向けての提言—『人と機械が読む時代』の知識基盤の確立に向けて—」を受けて、国立国会図書館が令和 3 年度から令和 7 年度までの 5 年間に取り組むべき事項を定めたものである。

ここでは、同計画において取り組むべき事項として掲げられたものについて、令和 3 年度から令和 6 年度末までの実施状況を取りまとめた。なお、特筆すべき実施状況として、次が挙げられる。

- ①令和 4 年 5 月に個人向けデジタル化資料送信サービスを開始した。
- ②補正予算による国内刊行図書のデジタル化等を進め、新たに、約 148 万点をデジタル化し、うち約 58 万点を送信サービス対象とした。
- ③令和 4 年 12 月に国立国会図書館デジタルコレクションをリニューアルし、機能の拡充等の実現に加え、デジタル化資料の全文テキストデータを検索可能とした。
- ④令和 6 年 1 月の新しい国立国会図書館サーチの公開により、国立国会図書館の所蔵資料や国立国会図書館デジタルコレクションのデジタル資料、国立国会図書館インターネット資料収集保存事業(WARP)で収集したウェブサイト、全国の公共・大学・専門図書館や学術研究機関等が提供する資料やデジタル資料の統合的な検索が実現し、調べ方のヒントや事例(リサーチ・ナビやレファレンス協同データベース)も一つのサービスから提供できるようになった。また、著作権法第 31 条第 2 項に基づく図書館等公衆送信サービスとして、PDF ファイルで複写物を提供する「遠隔複写(PDF ダウンロード)」サービスを令和 7 年 2 月に開始した。
- ⑤令和 6 年 1 月のみなサーチ正式版の公開により、様々な障害のある方が、いろいろな種類のアクセシブルな資料について、検索から利用まで簡単に行えるように、ユーザーインターフェースが改善された。また、デジタル化資料約 319 万点の全文テキストデータを視覚障害者等向けに送信し、音声読み上げや点字表示等に活用できるようになった。
- ⑥令和 5 年 1 月に有償等オンライン資料(電子書籍・電子雑誌)の制度収集を開始した。

### 目次

「人と機械が読む時代」に向けて取り組むべき事項.....	2
1 利活用促進のための取組 .....	2
(1) 利活用の基盤となる所蔵資料のデジタル化・全文テキスト化の推進.....	2
(2) 情報へのアクセス環境の改善 .....	4

(3) データの組織化・オープン化 .....	6
(4) データの利活用促進のための環境整備 .....	6
2 恒久的保存のための取組 .....	7
(1) 資料収集の強化 .....	7
(2) 未収資料・データのデジタル収集 .....	8
(3) デジタル資料の長期保存 .....	9
その他の事項 .....	10

## 「人と機械が読む時代」に向けて取り組むべき事項

### 1 利活用促進のための取組

#### (1) 利活用の基盤となる所蔵資料のデジタル化・全文テキスト化の推進

##### (a) 所蔵資料デジタル化の戦略的な推進

① 「資料デジタル化基本計画 2021-2025」(令和 3 年国図電 2103012 号)に基づき、所蔵資料のデジタル化を加速する。特に、和図書については対象年代を拡大し、2000 年刊行までのものを対象として、集中的にデジタル化を行う。また、和雑誌については、刊行後 5 年以上経過したものを対象範囲とし、社会的ニーズ及び学術的ニーズを踏まえ、雑誌記事索引採録対象誌及び学協会等からのデジタル化要望があるもののデジタル化を優先的に行う。

#### <実施状況>

- 2000 年までに刊行された国内刊行図書のうち約 148.6 万分冊、刊行後 5 年を経過した雑誌約 3.9 万分冊を外部委託により、図書・雑誌等約 4.5 万分冊を内製により、それぞれデジタル化した。
- 図書約 128.1 万点、雑誌約 5.2 万点の提供を開始し、うち約 58 万点を送信サービス対象に追加した。

#### デジタル化資料の提供点数

	R3.3 月末	R4.3 月末	R5.3 月末	R6.3 月末	R7.3 月末
図書	約 97 万点	約 99 万点	約 158 万点	約 186 万点	約 233 万点
雑誌	約 134 万点	約 135 万点	約 136 万点	約 137 万点	約 140 万点
総計(※)	約 276 万点	約 281 万点	約 343 万点	約 390 万点	約 451 万点

※図書・雑誌以外の資料も含む。

② 博士論文、録音資料・映像資料、古典籍資料、日本近現代政治史料、国会情報等の各資料群のデジタル化に引き続き取り組むほか、新たに新聞のデジタル化にも着手する。

#### <実施状況>

- 博士論文、古典籍資料、日本近現代政治史料、国会情報、新聞、児童書等の資

料群について、外部委託で約 22.4 万分冊(約 1,344.7 万コマ)、内製で約 3 万分冊(約 115.7 万コマ)のデジタル化を実施した。なお、このうち新聞については、日本新聞協会との間でデジタル化試行について合意のうえ、進めている。

- 録音・映像資料について、カセットテープ及びソノシート約 0.25 万点を外部委託により、LD 及び VHS ビデオテープ約 1.0 万点を内製により、それぞれデジタル化した。

(b)デジタル化資料の全文テキスト化等の実施

- ① 全てのデジタル化資料(古典籍資料、録音資料・映像資料等を除く。)の全文テキスト化に取り組む。

<実施状況>

- 古典籍等を除くデジタル化資料約 319 万点のテキストデータを作成した。
- NDLOCR(当館デジタル化資料に特化した独自開発の OCR(光学式文字認識))を利用してデジタル化資料からテキストデータを作成するシステムを構築し、令和 6 年 1 月から稼働開始している。

全文テキストデータを作成したデジタル化資料の点数

	R3.3 月末	R4.3 月末	R5.3 月末	R6.3 月末	R7.3 月末
累積点数	約 5 万点	約 247 万点	約 247 万点	約 269 万点	約 319 万点

- ② 生成されるテキストデータは、著作権、個人情報、プライバシー等に配慮した上で検索利用(単なる文字列検索にとどまらない高度な検索も想定する。)に供するほか、AI の学習用データセットとしての提供及び視覚障害者等への提供にも取り組む。

<実施状況>

- デジタル化資料約 319 万点の全文テキストデータを、リニューアルした国立国会図書館デジタルコレクションで検索を可能とともに、みなサーチによって視覚障害者等が利用できるようになった。
- 次世代システム開発研究室による実験サービスである次世代デジタルライブラリーで、著作権の保護期間が満了した図書約 28 万点、古典籍資料約 8 万点を全文検索可能とともに、OCR テキストデータを誰でも自由にダウンロードできる API で提供した。

- ③ 国立国会図書館内の開発研究体制を整備し、外部のオープンソースコミュニティや有識者等の知見も取り込み、レイアウト認識処理を含む光学式文字認識(OCR)の精度向上や、全文テキストデータの構造化に向けた調査研究に取り組み、その成果をできる限りオープンな利用条件で提供する。

<実施状況>

- NDLOCR 及び NDL 古典籍 OCR(デジタル化した古典籍に特化したもので、令和 4 年度に内製で開発)の開発・改善(視覚障害者等の読み上げ用テキストの作成を含

む。)を行い、プログラムをオープンソース(CC-BY)として公開した。

- ノートパソコン等の一般的な家庭用コンピュータや OS 環境で利用可能な、古典籍資料用の軽量 OCR「NDL 古典籍 OCR-Lite」を研究開発・公開した。

## (2) 情報へのアクセス環境の改善

- ① 内閣府を始めとする関係府省・機関と協働し、我が国の幅広い分野・地域のデジタルアーカイブと連携する「ジャパンサーチ」の一層の拡充及び海外も含めた活用促進に取り組む。

### <実施状況>

- 令和 3 年度以降、30 連携(つなぎ役)機関、155 データベースと連携した。令和 6 年度末時点で、55 連携(つなぎ役)機関、269 データベース、連携メタデータ数は約 3,100 万件となった。

ジャパンサーチの連携状況(累積)

	R3.3 月末	R4.3 月末	R5.3 月末	R6.3 月末	R7.3 月末
連携(つなぎ役)機関	25	33	39	49	55
連携データベース	118	170	202	230	269
連携データ	約 2,200 万件	約 2,530 万件	約 2,830 万件	約 2,950 万件	約 3,100 万件

※連携データベースは、データベースの統廃合による削除等があったため、実施状況に記載の新規連携データベース数とは一致しない。

- ② 書籍等分野のデジタルアーカイブのつなぎ役である「国立国会図書館サーチ」、蔵書検索・申込システムである「国立国会図書館オンライン」、国立国会図書館のデジタル資料の保存・提供の基盤である「国立国会図書館デジタルコレクション」等について、データ面及び機能面を拡充する。

### <実施状況>

- 国立国会図書館サーチについて、令和 3 年度以降、新規連携(34 機関)及び API による連携方式への切替え(11 機関)を行うとともに、連携維持のため、データ入替え、データ項目(サムネイル URL 等)追加等の取組を実施した(156 機関)。
- 令和 4 年 12 月に国立国会図書館デジタルコレクションをリニューアルし、機能の拡充等を実現した。

- ③ 各情報資源への適切なナビゲートを実現するため、次世代の統合的なオンラインサービスを整備する。その際、視覚障害者等へのインクルーシブな対応も強化する。

### <実施状況>

- 令和 6 年 1 月の新しい国立国会図書館サーチの公開により、国立国会図書館の所蔵資料や国立国会図書館デジタルコレクションのデジタル資料、国立国会図書館インターネット資料収集保存事業(WARP)で収集したウェブサイト、全国の公共・大学・専門図書館や学術研究機関等が提供する資料やデジタル資料の統合的な検索が実現し、調べ方のヒントや事例(リサーチ・ナビやレファレンス協同データベース)も一つのサービスから提供できるようになった。
- 令和 6 年 1 月のみなサーチ正式版の公開により、様々な障害のある方が、いろいろな種類のアクセシブルな資料について、検索から利用まで簡単に行えるように、ユーザーインターフェースが改善されたほか、OCR で作成した全文テキストデータの提供により、音声読み上げや点字表示等に活用できるようになった。

- ④ 著作権法改正の動向を踏まえた上で関係団体・府省等と協議を行い、デジタル化資料のうち絶版等資料に係るデジタル化資料送信サービスの提供範囲を図書館等から個人に拡大する。

＜実施状況＞

- 令和 4 年 5 月に個人向けデジタル化資料送信サービスを開始し、令和 5 年 1 月にはプリントアウト機能の提供を開始した。サービスの利用規約に同意した利用者の累計は、令和 6 年度末時点での約 30 万人である。
- 令和 6 年度末時点の図書館・個人送信資料の提供点数は約 204 万点である。

図書館・個人送信資料の提供点数

	R3.3 月末	R4.3 月末	R5.3 月末	R6.3 月末	R7.3 月末
提供点数	約 152 万点	約 153 万点	約 184 万点	約 179 万点	約 204 万点

※R4.5 の個人送信資料サービス開始以前は、図書館送信資料としての提供点数。

- ⑤ 図書館資料の複写物の送信サービス(複写物の電子ファイルでの送信)については、著作権法改正の動向を踏まえた上で関係団体・府省等と協議を行うとともに、制度設計及びその実施に取り組む。

＜実施状況＞

- 令和 4 年度から関係者協議を行い、サービス開始に向けた当館の各種準備を進めていたところ、図書館等公衆送信補償金管理協会から、図書館等公衆送信サービスへの対応が当面行えない旨の告知があったことから、当初予定していたサービス開始時期を見直すこととなったが、令和 7 年 2 月に PDF ファイルで複写物を提供する「遠隔複写(PDF ダウンロード)」サービスを開始した。

- ⑥ デジタル化資料の著作権処理を加速し、インターネット公開資料の一層の拡大を図る。

＜実施状況＞

- 令和 3 年度以降の累計で、保護期間満了を確認した和図書等の約 1.3 万点、及び

権利者から公開要望を受けた和雑誌等の約 0.8 万点をインターネット公開した。

### (3) データの組織化・オープン化

- ① 「識別子に係る実施方針」(令和 2 年国図電 2001067 号)に基づき、関係機関と協力して、国立国会図書館が作成する書誌データの相互運用性の向上や、「ジャパンサーチ」による文化情報資源のリンクトデータ基盤形成への貢献等に取り組むとともに、これらの取組の一層の拡充に向けた検討を行う。

#### ＜実施状況＞

- 令和 3 年度以降の累計で、VIAF(バーチャル国際典拠ファイル。各国の国立図書館等からの典拠データ同定・名寄せした国際典拠ファイル)に約 21.9 万件の典拠データを提供するとともに、Wikidata にジャパンサーチの正規化 URI 約 0.8 万件を登録した。

- ② 国立国会図書館がデジタル化した資料の画像データや全文テキストデータの二次利用の枠組みを整備するほか、書誌データのオープンデータセットの継続的な提供にも取り組む。

#### ＜実施状況＞

- オープンデータセットの更新(国立国会図書館デジタルコレクション書誌情報等)・追加(参考図書紹介)を行った。
- 国立国会図書館内・図書館送信限定公開デジタル化資料の画像データ提供の試行を行った(令和 3 年 7 月から令和 6 年 3 月まで)。
- 「NDL イメージバンク」、「近代日本人の肖像」等の電子展示会を通じて、著作権保護期間を満了した画像データの利活用可能性を拡げる取組を進めた。
- 著作権の保護期間が満了した図書約 28 万点、古典籍資料約 8 万点については、次世代デジタルライブラリーから OCR テキストデータを誰でも自由にダウンロードできる API を提供している(再掲)。

- ③ 国立国会図書館が保存するデジタル化資料、オンライン資料等のデジタル資料へのデジタルオブジェクト識別子(DOI)付与に、引き続き取り組む。

#### ＜実施状況＞

- 令和 3 年度以降の累計で、デジタル化資料等約 132 万件に DOI を付与した。

### (4) データの利活用促進のための環境整備

- ① 若手研究者等を対象とする共同研究プログラム等により、デジタル化資料から抽出した大量の画像データや全文テキストデータ、「国立国会図書館インターネット資料収集保存事業」(WARP)から抽出したデータ、「ジャパンサーチ」で収集したメタデータ等、国立国会図書館が作成又は収集したデータを外部の機関や有識者等が利活用し、協働できる環境の整備に取り組む。

#### ＜実施状況＞

- デジタルライブラリーカフェ等のイベントの開催や外部イベントへの協力により、当館デ

一タ利活用のための情報発信を行った。

- ジャパンサーチで収集したメタデータについては、RDF データへの変換を継続し約 2,860 万件を提供している。
- NII における大規模言語モデル(LLM)の構築に協力するため、WARP で収集保存したファイルの URL リストを、NII に提供した。

② 関係機関や外部の有識者等とも協力の上、「ジャパンサーチ」の利活用機能等も用いて、上記のデータを活用した教育シーン（オンライン教育を含む。）での利活用モデルの構築や実践支援、一般市民の主体的な参加を促し、その知的な活動を後押しするようなプログラムの開発に取り組む。

#### ＜実施状況＞

- デジタルアーカイブの教育活用に関する講演やワークショップに積極的に参加し、ジャパンサーチのデータを使った利活用の促進に努めた。

③ 前 2 項の取組をコーディネートし、橋渡し役となる人材の育成に取り組む。

#### ＜実施状況＞

- 利活用事例の共有や各種イベントの開催、「「デジタルアーカイブ活動」のためのガイドライン」の掲載等を通じて、橋渡し役となる人材の育成に必要なコンテンツの充実や事例の共有に努めた。

④ メタデータの外部提供インターフェイス(API)による提供については、新しいプロトコルへの対応、項目の追加、新規フォーマットへの対応等を行い、より利活用しやすくなる。

#### ＜実施状況＞

- NII・JST・NDL 連絡会議に設置されたメタデータ WG の活動の一環として、メタデータ設計者向けに、メタデータの作成・流通経路の適切な選択ができるように作成した「メタデータ流通ガイドライン」の公開・整備を行った。
- 国立国会図書館サーチによる提供データについて、令和 3 年度以降の累計で、36 機関のライセンスを明示した。

## 2 恒久的保存のための取組

### (1) 資料収集の強化

① 我が国の有償等オンライン資料の収集・保存・利用のため、制度収集を開始し、これらの基盤の整備に取り組む。また、制度収集開始までの間、引き続き任意提供による有償等オンライン資料の収集に努める。

#### ＜進捗＞

- 国立国会図書館法等の改正により、令和 5 年 1 月から民間発行の有償又は DRM 付きの電子書籍や電子雑誌（有償等オンライン資料）の制度収集を開始し、令和 6 年度末までに約 0.3 万点を収集した。

●安定的な制度運用のため、制度の周知・広報や関係者との協議を行うとともに、収集した有償等オンライン資料の公開までにかかる期間を短縮するため、国立国会図書館内の横断的な体制で同資料を組織化する試行を開始した。

- ② 外国資料については、オープンアクセスなど刊行・流通形態の変化に留意しつつ、引き続きコアジャーナルを中心とした電子ジャーナル等の契約を着実に行う。また、電子ジャーナルのバックファイルを整備するとともに、セーフティネットの観点から永続的なアクセスの確保に努めた上で、大学等とのネットワークの構築に協力する。

＜進捗＞

●外国雑誌について、冊子から電子ジャーナルへの切替えを進めるとともに、ScienceDirect 等のバックファイルの整備を行って利便性を高め、科学技術関係資料費による外国雑誌の電子ジャーナルの購読タイトル数は令和 6 年度末までに 1,309 タイトル増加した(冊子は 193 タイトル減少)。あわせて、予算逼迫及び急激な円安、物価上昇等を踏まえ、購読中の冊子及び電子ジャーナルの定期的なタイトル見直しを行うことで、コアジャーナルの購読確保に努めた。

- ③ 我が国の学術情報である海外刊行の国内学協会誌の調査及び収集に引き続き努める。

＜進捗＞

●海外刊行の国内学協会誌(海外の学協会との共同編集・刊行を含む。)について、国内外の情報源を参考・調査し、未所蔵又は寄贈中止が判明した 15 タイトルを新規に冊子又は電子ジャーナルで購読した。

(2) 未収資料・データのデジタル収集

(a) 未収資料のデジタルデータ及びメタデータの収集

- ① 未収(国立国会図書館が所蔵していない)の地域資料や、海外の機関が所蔵する日本関係資料等について、関係機関と連携・調整の上、デジタルデータでの収集・保存・提供に積極的に取り組む。そのため、受入れの条件や手続等を整備する。

＜実施状況＞

●令和 4 年 7 月の本格実施以降、ウェブサイト等を通じた協力依頼や関係団体との協力を通じて、公共及び支部図書館から累計約 0.7 万点の収集を行った。なお、このうち入手困難であることが確認された約 0.2 万点は送信対象に、著作権に問題のない約 0.5 万点はインターネット公開となっている。

- ② 「ジャパンサーチ」、「国立国会図書館サーチ」等を通じて、未収資料や未連携のデータベース等のメタデータの収集、組織化及び提供に取り組む。

＜実施状況＞

●書籍等分野のつなぎ役として、国立国会図書館サーチの連携拡張によって、メタデータの収集・組織化を継続し、令和 6 年度末時点で、当館提供分以外の書籍等分野のメタデータ約 370 万点をジャパンサーチで提供するに至った。

- ③ 資料のデジタル化やメタデータの整備が進んでいない機関に対する支援策を検討する。

＜実施状況＞

- 資料デジタル化等に関する研修及び情報交換会を行い、各機関への技術・ノウハウ等の支援に努めた。また、情報交換会の実績を踏まえて開催したオンラインフォーラムで各地の取組を全国に共有した。

(b) データベース、研究データ等の収集

- ① 存続が困難となったデータベースやデジタルアーカイブの継承に取り組む。「国立国会図書館東日本大震災アーカイブ」におけるアーカイブの継承を継続するほか、その経験を踏まえ、これらを継承する際の課題や条件等を検討する。

＜実施状況＞

- 東日本大震災に係るデータベースのうち閉鎖した 3 機関のデータを承継し、公開したほか、閉鎖予定の連携先と承継に向けて調整を行った。また、「東日本大震災アーカイブシンポジウム」等を通じて、継承の取組や課題の周知を図った。

- ② 研究データについては、国立情報学研究所、科学技術振興機構等との分担を前提に、地域資料や特定の研究分野に結び付かない分野横断的なものを対象とする等、国立国会図書館が担うべき役割の整理を進め、収集・保存・提供する際の課題や条件等を検討する。

＜実施状況＞

- NII 及び JST と協力して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を開催し、その成果として「メタデータ流通ガイドライン」の公開・整備を行った。(再掲)

- ③ 民間ウェブサイトの収集については、近年の動向を踏まえて課題を整理する。

＜実施状況＞

- 収集対象の選定マニュアルの整備に向けて、諸外国の動向やこれまでの収集実績を踏まえて課題を整理し、今後取り組むべき内容を検討した。

(3) デジタル資料の長期保存

- ① 「国立国会図書館デジタル資料長期保存基本計画 2021-2025」(令和 3 年国図電 2102181 号)に基づき、国立国会図書館が所蔵するパッケージ系電子出版物の長期利用が可能となるよう、国立国会図書館内の実施体制を整備するとともに、マイグレーション等の作業に取り組む。

＜実施状況＞

- 内製により FD 等のマイグレーションを進め、FD についてはほぼ完了した。令和 3 年度からの累計で、FD 約 0.8 万点のほか、USB メモリ、RW 型光ディスク、MO、MD 等のマイグレーションを行った。

- 光ディスクについては、技術調査等のうえ外部委託によりマイグレーションを行うとともに、マイグレーションシステムを開発し、劣化状況調査を行った。

- これらの当館が所蔵する主要な記録媒体について、マイグレーション手法を確立し、マイグレーションのための設備、体制を整えた。

② デジタル化資料、オンライン資料等のデジタル資料に関しては、合理的かつ安定的な保存環境を整備するとともに、これらの長期保存対策に必要な技術的調査研究を行う。

#### ＜実施状況＞

●デジタル化資料の保存用画像データの安定的な保存のために、これまで使用してきた破損や劣化のリスクがあり、数量的に継続的なマイグレーション（媒体移行）が困難な光ディスクや HDD から、大容量でマイグレーションが比較的容易な LTO（データ保存用磁気テープ）への移行作業を軌道に乗せた。令和 3 年度以降、累積で光ディスク約 9 万点、HDD 約 250 台を移行した。

●デジタルデポジットシステム及び WARP について、クラウド、オンプレミス機器及び LTO のそれぞれの特性や、災害復旧、長期保存性、経済性を考慮してデータの保存方法を検討・決定した。

③ 前項で得られた知見を踏まえ、関係機関と長期保存に関する課題等の共有を可能とするコミュニティの形成を図る。

#### ＜実施状況＞

●長期保存に係る国内機関の実態、光ディスク及び FD の長期保存対策技術に係る調査を行い、報告書を公開した。

#### その他の事項

- 関係団体等と協力し、商用の電子図書館サービスを視覚障害者等の方が使う際に必要なアクセシビリティ要件を整理したガイドライン「電子図書館のアクセシビリティ対応ガイドライン 2.0」※を公開し、公立図書館等への普及活動を行った。
- 生成 AI の利活用可能性を検討する内部体制を立ち上げ、検討を進めている。

※「電子図書館のアクセシビリティ対応ガイドライン 2.0」は、令和 5 年 7 月に策定・公開した「電子図書館のアクセシビリティ対応ガイドライン 1.0」を更新して、令和 7 年 5 月に公開した。