

第五期国立国会図書館科学技術情報整備基本計画

「第五期国立国会図書館科学技術情報整備基本計画」(以下「本計画」という。)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—」(以下「提言」という。)を受けて、国立国会図書館が今後5年間(令和3年度から令和7年度まで)に取り組むべき事項を定めるものである。なお、本計画は、平成28年3月に策定した「第四期国立国会図書館科学技術情報整備基本計画」(平成28年国図利1603252号。以下「第四期計画」という。)の後継に当たる。

目次

I 基本的な考え方	1
1 目的	1
2 「人と機械が読む時代」の知識基盤の確立に向けて	2
II 「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	3
(3) データの組織化・オープン化	3
(4) データの利活用促進のための環境整備	4
2 恒久的保存のための取組	4
(1) 資料収集の強化	4
(2) 未収資料・データのデジタル収集	4
(3) デジタル資料の長期保存	5
III 実施に当たって	5

I 基本的な考え方

1 目的

本計画は、提言を踏まえて、国立国会図書館が、デジタルトランスフォーメーション(DX)による社会変革を後押しし、少子高齢化や地方創生といった我が国の課題の解決や国連の持続可能な開発目標(SDGs)の達成に貢献していくために、オープンで信頼され、大規模災害や感染症流行といった緊急事態に対するレジリエンスを備えた知識基盤を構築し、人々の利活用に資する観点で、科学技術情報整備を進めることを目的とする。

国立国会図書館は、この知識基盤を常に、かつ、広く参照可能とすることで、人々の知的な創造活動を支援していく。

2 「人と機械が読む時代」の知識基盤の確立に向けて

提言では、現在の情報環境の変化として、自然科学のみならず人文学・社会科学においても、「データ駆動型研究」の進展に伴い研究スタイルが変化していることや、新型コロナウイルス感染症(COVID-19)の世界的流行が明らかにしたように、図書館のデジタル化及び情報通信技術(ICT)を活用した遠隔サービスがDXの観点からは不十分であったことを挙げている。また、海外の動向として、先行する欧米のデジタル化・オープン化等に係る法整備を踏まえた各種の取組や、近年進境著しいアジア諸国のアーカイブ整備・デジタル化等の状況に触れ、これらを参照しつつ我が国のあるべき姿を迫及していく必要があるとしている。

その上で、提言は、国立国会図書館の今後の取組について、「人」による利活用に加えて、人工知能(AI)等の「機械」による利活用という二つの方向性を示し、国立国会図書館はこれらの実現に向けて「利活用促進」及び「恒久的保存」のための基盤を整備していくべきである、としている。

第四期計画では、多種多様な資料・情報への長期的かつ広範なアクセスと利活用の基盤となる「深化型知識インフラ」の実現を目指し、資料のデジタル化及びデジタル化資料図書館送信サービスの拡大、テキストデータの活用、メタデータのオープンライセンス化等の取組が行われた。提言では、第四期計画の着実な進捗を評価する一方、全文テキスト化を含めたデジタル化の推進、オンライン資料の収集範囲拡大、電子情報の長期保存の本格的な実施といった今後国立国会図書館が取り組むべき課題も指摘している。

本計画は、提言で示された長期的な展望に立ち、「国立国会図書館ビジョン2021-2025」(令和3年国図企2101133号)の下、適切な財政措置を講じつつ、国立国会図書館が今後5年間にわたって「人と機械が読む時代」の知識基盤の確立に向けて取り組むべき内容を具体的に示すものである。

II 「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。
- 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。

(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。
- 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。
- 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。

(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。
- 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。
- 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。
- 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。
- 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。
- 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。

(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。
- 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。

- 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。

(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。
- 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。
- 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。
- 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する。

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。
- 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。
- ③ 我が国の学術情報でも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。

(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。

- 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。
- 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。

(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。
- 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。
- 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。

(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。
- 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。
- 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。

Ⅲ 実施に当たって

各事項の実施に当たっては、関係機関との連携・協力を密にし、国立国会図書館全体としてこれに取り組む。また、本計画に掲げた各事項について、年度ごとに進捗状況の確認と評価を行う。

なお、本計画に記載された情報環境をめぐる諸状況の変化が激しいことから、これに機動的に対応するために、本計画は必要に応じて見直すものとする。