

**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（素案）**  
— 「人と機械が読む時代」の知識基盤の確立に向けて— の概要

**I 本提言の位置付け**

- ・基本計画は今後 5 年間（令和 3 年度から 7 年度まで）を対象とするが、提言の射程は急速な技術の進歩などを想定し、意図的にやや長期的な展望を意識したものとした。
- ・デジタル技術が生活のあらゆる局面に浸透し、蓄積・流通するデジタルデータを AI 等により利活用する社会の変革が進みつつある今日、国立国会図書館は、「人」だけでなく、AI 等の「機械」も「読者」とするときを迎えたと捉え、同館の任務にかんがみ、これからの時代の知識基盤をどのように確立していくべきかという観点から取りまとめた。
- ・今年改正された科学技術基本法における定義や現在検討中の国による次期（第六期）科学技術基本計画の方向性も共有することとした。

**II 基本的な視点**

提言の背景となる国立国会図書館を取り巻く状況として、「研究・社会のデジタル・シフト」及び「第四期国立国会図書館科学技術情報整備基本計画の主な成果」を取り上げた。

**1 研究・社会のデジタル・シフト**

**(1) データ駆動型研究の進展**

- ・「デジタルトランスフォーメーション（DX）」は図書館を含めた学術情報全般の流通・コミュニケーションの構造変化を引き起こし、「データ駆動型研究」の進展をもたらした。その好例である「デジタル人文学」は、文献等の文化資源の蓄積をいかすことから図書館においても着目される。

**(2) 新型コロナウイルス感染症拡大がもたらした「ニュー・ノーマル」**

- ・新型コロナウイルス感染症の拡大は、我が国全体の教育・研究活動の停滞を招き、図書館の物理的な場所と資料（フィジカル）に依拠したサービスの限界を示したが、従来の資料のデジタル化や遠隔サービスも、デジタルトランスフォーメーションという面では結果的に不十分であった。

**(3) 海外の動向**

- ・オープンデータ及びオープンアクセスの促進のための法整備や学術リポジトリ等のオープンサイエンスに係る取組などが注目される。

**2 第四期国立国会図書館科学技術情報整備基本計画の主な成果**

第四期計画期間中に「深化型知識インフラ」の構築に向けた取組は着実に前進していると評価できる。一方で、今後の課題としては、全文テキスト化を含めたデジタル化の推進や、ジャパンサーチの利活用促進、オンライン資料の収集範囲拡大、データのオープン化の一層の推進、電子情報の長期保存の本格的な実施等が挙げられる。これらを踏まえつつ、

国立国会図書館は、次期計画において 1 で述べたような状況に対応する知識基盤の整備に取り組むべきである。

### Ⅲ 「人と機械が読む時代」の知識基盤の確立に向けた取組の方向性

総論としての「全体の方向性」と個別テーマである「個別の取組の方向性」で構成している。「個別の取組の方向性」については、利活用促進と基盤整備は表裏一体の関係にあると考えられるので、これらで切り分けることとはせず、利活用を前面に押し立てることとした。

#### 1 全体の方向性

- ・デジタルトランスフォーメーション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 SDGs にもつながる我が国の課題の解決に貢献していくために、オープンで広く信頼され、大規模災害や感染症流行といった非常事態に対するレジリエンス（しなやかさ）を備えた国の知識基盤の整備に取り組まねばならないとした。
- ・あらゆる人々が読める環境を整備し、調査、研究、教育等の場面で利活用してもらう方向性と、国立国会図書館が潜在的・顕在的に保有するデータを AI 等の機械が読める形式で提供することでデータ駆動型研究に貢献する方向性を提示した。

#### 2 個別の取組の方向性

##### (1) データのオープン化と教育等における利活用促進

- ・国立国会図書館が収集・整備・保存するデータ（メタデータや典拠情報を含む。）のオープンな利用条件での提供
- ・データ提供に際し「人」に対するインクルーシブなインターフェイスの整備や「機械」に対するインターフェイスのための取組の一層の充実
- ・コレクション間の相互連関のための基盤として、識別子も含む典拠情報等を戦略的に拡張・整備していく必要性
- ・物理的・地理的な制約を克服するため、著作権処理の推進によるインターネット公開資料の拡大への一層の取組
- ・図書館資料の複製物のデジタル送信や絶版等資料の利活用
- ・データ駆動型研究等の学術研究を目的とするデータの利活用を可能とする枠組みの整備
- ・フェロシップ制度や共同研究等による外部の知見や資源を取り込んだ利活用モデルの提示や、データを用いた研究等の推進
- ・国内の多様なデジタルアーカイブ資源の利活用という視点からの「ジャパンサーチ」の更なる連携拡充の重要性

##### (2) 資料のデジタル化・全文テキスト化等の推進

- ・全ての資料の長期的なデジタル化の実現に向け、国立国会図書館による、国会情報を始め、国内出版物の刊行年代や分野を拡大してのデジタル化の一層の推進

- ・今後のデジタル化の全文テキスト化や画像抽出等の技術を基にした本文検索、画像検索といった所在検索に加えて XML 等による構造化等の推進、データセットとしての利活用
- ・技術的課題解消のため、外部の知見・技術を取り入れつつ、技術的開発に必要な体制・環境の整備
- ・デジタル化及び全文テキスト化の対象拡大に伴う関係諸機関との丁寧な協議、個人情報やプライバシーへの配慮、手続の透明性確保等の必要性

### (3) 多様な文化資源の収集・保存

- ・恒久的な収集・保存の仕組みが存在しない有償等オンライン資料収集への早急な対応
- ・地域資料や海外機関所蔵の日本関係資料等の未収資料のデジタル等による収集
- ・科学技術情報の重要な要素である研究データについて、国立情報学研究所や科学技術振興機構等の関係機関との分担による国全体としてのアクセスの向上
- ・WARP（国立国会図書館インターネット資料収集保存事業）の枠組みを活用したデータベースやデジタルアーカイブの構築のほか、公的機関の研究データや政策データ、地域資料や分野横断的なデータの承継
- ・「ジャパンサーチ」等を通じたメタデータの収集、デジタル化やメタデータの整備が進んでいない機関への支援
- ・収集したデータへの長期アクセス保証実現のための持続可能なアーカイブ基盤の整備

# 第五期科学技術情報整備基本計画策定に向けての提言（素案） — 「人と機械が読む時代」の知識基盤の確立に向けて —

海外のデジタルオープンデータ・オープンアクセスの取組  
「**新型コロナウイルスがもたらした「ニュー・ノーマル」**」

教育、地方創生、少子高齢化…  
SDGsにつながる我が国の課題解決への貢献

デジタルトランスフォーメーション  
による社会変革を後押し

## 利活用促進

**人**

**オンラインアクセス強化**

- ・図書館送信サービスの制度改善
- ・統合的なオンラインサービス構築
- ・アクセシビリティ対応 等

物理的・地理的な制約の克服

**国立国会図書館の取組**

- ✓資料デジタル化・全文テキスト化の推進
- ✓収集データの識別子整備、オープン化
- ✓著作権処理、図書館送信の拡充
- ✓ジャパンサーチ等の拡充・改善
- ✓教育シーン等での利活用モデル提示

**機械**

**AI等の機械による利活用**

- ・データセット・APIによる提供
- ・他領域データとの相互連携 等

データ駆動型研究への貢献

### 外部の知見や資源を取り込みつつ取り組む

**NII・JST等** 連携・分担

**国全体としてのアクセス向上**

- ・他の基盤でカバーできない分野のデータ等の承継・保存
- ・民間ウェブサイトの収集検討 等

- ✓従来の資料収集・購読の強化継続
- ✓未収資料デジタルデータ、メタデータ収集
- ✓閉鎖アーカイブ、データベースや横断的な研究データ等の承継・保存（WARP）
- ✓長期アクセスを保証する持続可能なアーカイブ基盤の整備

連携・支援 **類縁機関**

**持続可能な情報基盤の整備**

- ・メタデータ含めたデータ収集
- ・デジタル化、メタデータ整備、長期アクセス保証の支援 等

国会情報      有償・無償のオンライン資料      国内外の未収資料      研究データ閉鎖DB等      デジタルアーカイブ

## 恒久的保存

日本/日本語に係る多様な知的文化資源