

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（案）

— 「人と機械が読む時代」の知識基盤の確立に向けて—

令和 3 年 1 月 13 日

科学技術情報整備審議会

目次

I	本提言の位置付け	1
II	基本的な視点	1
1	研究・社会のデジタルシフト	1
(1)	「データ駆動型研究」の進展	1
(2)	新型コロナウイルス感染症拡大の影響	2
(3)	海外の動向	2
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	3
(1)	恒久的保存のための取組	3
(2)	利活用促進のための取組	3
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	4
1	全体の方向性	4
2	個別の取組の方向性	5
(1)	データのオープン化と教育等における利活用促進	5
(2)	資料のデジタル化・全文テキスト化等の推進	7
(3)	多様な文化資源の収集・保存	7
IV	おわりに	9
	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（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	10
	用語集	11

I 本提言の位置付け

本提言は、国立国会図書館における今後5年間を目途とした科学技術情報の整備の在り方についての基本方針を提言するものである。しかしながら、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。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の利活用を人工知能（AI）等を用いて促進する、など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を搭載した「機械」も「読者」とするときを迎えた。デジタルを前提とした社会に向けて、国立国会図書館は、我が国唯一の国立図書館として、あるいは国会に附属する立法補佐機関として、いかに科学技術情報を整備していくべきか、換言すれば、どのようにこれからの時代の知識基盤を確立していくべきかという観点から、本提言をまとめた。

なお、令和2年に改正された科学技術基本法¹及び国による次期科学技術基本計画（科学技術・イノベーション基本計画）策定に向けた議論においては、「Society 5.0」の実現に向けて、自然科学のみならず、人文学・社会科学を含むおよそあらゆる学問の領域を対象とし、これらの知見を横断的、総合的に活用して、国連による持続可能な開発目標（SDGs）やAIの倫理等の現代的諸課題の解決やイノベーションの創出に当たろうとしている。本提言においても、同法の対象及び国の基本計画の方向性を共有している。

II 基本的な視点

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

近年、国際的にオープンサイエンスの概念が普及し、学術ジャーナルのデジタル化・オープンアクセス化という従来の流れを推し進めただけでなく、研究サイクル・研究プロセスにおけるデジタル化の一層の浸透や、データを出発点として仮説・モデル・知識を生成する研究、すなわち「データ駆動型研究」の進展をもたらしつつある。

この「データ駆動型研究」は、自然科学はもとより、人文学・社会科学においても研究スタイルを変化させ、新たな研究の可能性を開きつつある。情報学や統計学の知見により研究手法を革新するとともに、人文学分野で生み出される大規模データを他分野に導入して新たな研究課題や知識を得ようとする近年の「デジタル人文学」はその好例である。その発展のためには、文献等の文化資源を蓄積した図書館のデジタル化が不可欠である。

こうした研究のデジタルトランスフォーメーション（DX）が進む中、査読を伴わず迅速に公開されるプレプリントの拡大や査読のオープン化など、従来の学術情報流通・コミュニケーションの限界を乗り越えようとする動きも見られるが、同時に、信頼性への

¹ 科学技術基本法等の一部を改正する法律（令和2年法律第63号）により令和3年4月1日から科学技術・イノベーション基本法に改められる。

懸念など、新たな課題を生み出している。また、オープンサイエンスの普及とともに、在野の研究者や市民を巻き込んだシチズンサイエンスも広がりつつあり、社会における科学や研究の在り方は大きな転換点を迎えている。その中で図書館に期待され、果たすべき役割も変容しつつある。

(2) 新型コロナウイルス感染症拡大の影響

折からの新型コロナウイルス感染症（COVID-19）の世界的拡大は、現状の学術情報流通・コミュニケーションの基盤のぜい弱性もあぶり出した。国立国会図書館や大学図書館による資料の閲覧・貸出し及び紙の複写物の提供といった、物理的な場所と資料等（フィジカル）に依拠したサービスの休止・縮小により、国内外の専門家や学生、一般市民への直接的な図書館サービスが途絶しただけでなく、我が国全体の教育・研究活動等の停滞を招いたことは、その顕著な事例である。

これは、国立国会図書館を含めた我が国の学術情報機関がこれまで行ってきた資料のデジタル化や情報通信技術（ICT）を活用した遠隔サービスが、デジタルトランスフォーメーションという観点からは結果的には不十分であったことを示すものと言える。今後新型コロナウイルス感染症が終息し、これまでのようなフィジカルに依拠したサービスが復活するとしても、それだけを従来と変わりなく継続するということはもはやあり得ない。デジタル化や ICT を活用した遠隔サービスをたゆみなく推進していくことが、全ての人々にとって使いやすい図書館であるために不可欠であることを我々は改めて認識したのである。

(3) 海外の動向

米国著作権法フェアユース条項や、欧州におけるデジタルアーカイブ促進のための権利制限規定、オープンデータ及びオープンアクセスの促進のための指令など、欧米諸国のデジタル化やオープン化等に係る法整備等は総じて我が国の先を行くものである。各国の国立図書館においては、これらを踏まえつつ、学術リポジトリやオンライン出版プラットフォームの整備といったオープンサイエンスに係る取組、クラウドソーシングによるデジタル化資料のテキスト化及び全文テキストデータを活用した外部研究者との共同研究等の取組が進められている。

また、近年は、中国、台湾、韓国やシンガポールにおいても大規模な学術リポジトリやデジタルアーカイブが整備されている。例えば、中国では国家図書館が地方公共図書館の所蔵資料のデジタル化・共有を進めるなど、アジア諸国の進境には著しいものがある。

これらの取組も参照しつつ、我が国のあるべき姿を追求していく必要がある。

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

国立国会図書館は「第四期国立国会図書館科学技術情報整備基本計画」（以下「第四期計画」という。）において、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を可能とする基盤となる「深化型知識インフラ」の実現を目指し、様々なコンテンツを生み出し蓄積する「恒久的保存のための領域」と、コンテンツをより利活用しやすく整備する「利活用促進のための領域」の二つの領域の充実とこれらをつなげる役割を果たす、としていた。領域別にまとめた主な成果は次のとおりである。

(1) 恒久的保存のための取組

○デジタル化の推進

平成 29 年度に科学技術関係資料を対象としたデジタル化に係る予算を措置し、前年度より当該予算を倍増した（約 2.3 億円）。図書、雑誌（和洋の国内学協会誌を含む。）及び博士論文を中心にデジタル化を着実に進めたほか、録音資料（カセットテープ、SP レコード）・映像資料（レーザーディスク）等のデジタル化にも着手した。ただし、デジタル化は所蔵資料のうち和図書・和雑誌等の 5 分の 1 程度にとどまっている。

○電子情報資源の長期アクセス保証

電子情報資源の長期保存・長期アクセス保証に係る取組として、NII-ELS（国立情報学研究所電子図書館事業）で維持困難となった学術情報等の保存や、WARP（国立国会図書館インターネット資料収集保存事業）の収集対象の拡大、パッケージ系電子出版物の長期保存に係る調査、USB メモリー・M0（光磁気ディスク）資料等のマイグレーション作業を行った。また、関連して、「国立国会図書館東日本大震災アーカイブ」（以下「ひなぎく」という。）において、存続が困難となった各地のアーカイブの承継に着手した。

○文献相当の国内情報資源の網羅的な収集

有償等オンライン資料収集については、関係団体との協議や実証実験を行い一定の進展は見られたものの、制度に基づく収集は実現しておらず、オンライン資料の流通増に対して収集できない資料が広がっている。

(2) 利活用促進のための取組

○デジタル化資料の利活用促進

「図書館向けデジタル化資料送信サービス」の国内参加館を大幅に増加（約 400 館→約 1,200 館）させた。また、令和元年度から海外の図書館等への送信を行うこととなった。ただし、著作権保護期間の延長もあり、インターネット公開点数の伸びは頭打ちとなっている。

○テキストデータの活用

国立国会図書館の次世代システムの開発研究部門において、機械学習を活用したテキスト化精度向上の研究に取り組み、その途中成果として、デジタル化資料の全文検索機

能を搭載した「次世代デジタルライブラリー」の公開、テキスト化したデータの学習用データセットとしての公開等を行った。また、日本点字図書館と協力して共同校正システムを用いたテキストデータ化実証実験や、「ひなぎく」での全文検索機能の一部提供を行った。

○多様なコンテンツのメタデータの統合的検索機能の提供

図書館だけでなく、公文書館や美術館、博物館等の多様な分野のデジタルアーカイブからメタデータを収集し、統合検索機能を提供する「ジャパンサーチ」を、内閣府を始めとする関係府省等と協力の上、国立国会図書館の開発研究部門が中心となって開発し、令和2年8月に正式版を公開した。

○メタデータのオープンライセンス化・標準化の推進

国立国会図書館作成書誌データの無償化（CC BY 互換）を行い、オープンデータセットとして提供を開始した。また、「ジャパンサーチ」の「共通メタデータフォーマット」を策定するとともに、メタデータのオープンライセンス化（原則 CC0）及びメタデータ API（外部提供インターフェイス）機能による利活用を促進した。このほか、国立国会図書館が所蔵資料をデジタル化したものへのデジタルオブジェクト識別子（DOI）の付与や、DC-NDL（国立国会図書館ダブリンコアメタデータ記述）の維持・普及を行った。

以上をまとめると、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。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、電子情報資源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、次期計画において前節で述べたような状況に対応する知識基盤の整備に、適切な財政措置を講じた上で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

1 全体の方向性

国立国会図書館は、その使命を果たすため、デジタル技術を活用し、また、ポスト・コロナのデジタルを前提とした新しい社会に適合した方法により、科学技術情報の整備を促進していくべきである。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我が国の課題の解決や SDGs の達成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。この取組を推進することにおいてこそ、これからの社会において、我が国がこれまで蓄積してきた知的活動の成果をあまねく人々によって常に、かつ広く参照すること、そして判断のよりどころとすることが可能となる。すなわち国の知識基盤としての国立国会

図書館の本来的な役割をよりよく果たすことができるのである。

その際、冒頭に述べた「人」と「機械」という二種類の「読者」（利活用のチャンネル）から逆算して、求められる国立国会図書館の取組を整理していくことを提案したい。具体的には、国内外の専門家や学生、一般市民といったあらゆる人々が読める（オンラインによるアクセスが可能な）環境を整備して調査、研究、教育等の多様な場面で利活用してもらうという方向性と、国立国会図書館が潜在的・顕在的に保有するデータを「機械」が読める（利活用可能な）形式で提供することでAI等を活用した「データ駆動型研究」に貢献するという方向性が考えられる。両者は相互に独立して存在するものではなく、「機械」が読むことも最終的には「人」に資することとなると言える。この二つの方向性を実現するための、利活用促進と恒久的保存のための基盤の整備が、引き続き国立国会図書館の取組の核心となる。その際、これらの取組が、国立国会図書館の中心的任務の一つである立法補佐機能の深化・高度化にも資するものとするべきである。

「利活用促進のための領域」においては、全文テキスト化等を射程に入れた、国会情報を含む資料のデジタル化を戦略的に推進するとともに、メタデータを含むそれらのデータの組織化・オープン化に取り組むべきである。また、レジリエンス確保の観点からも、著作権処理の加速化によるインターネット公開資料の増加や、著作権法改正の動き等も踏まえ、「図書館向けデジタル化資料送信サービス」の利用者への直接送信を図ること等により、図書館資料へのアクセスの容易化に取り組むべきである。更に、教育シーンでの利活用モデルの構築や実践支援、「ジャパンサーチ」等の拡充による情報アクセス環境の改善が求められる。これらの実現のためには、特に制度・技術面において、外部の知見を取り込みつつ取り組むことも必要である。

「恒久的保存のための領域」においては、従来の資料収集の強化継続に加え、未収資料について、メタデータのみならず原資料をデジタルデータで収集して恒久的に保存するとともに、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、分野横断的な研究データ等の承継にも引き続き取り組む必要がある。また、図書館関係のみならず様々な分野のデジタルアーカイブのメタデータの収集に、「ジャパンサーチ」等を通じて取り組むべきである。その際には、関係諸機関と分担しつつも、国立国会図書館が日本及び日本語に係る全体の知識基盤整備において主導的な立場を果たしていかなばならないだろう。

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。）は、我が国のみならず世界に開かれたものとして、セキュリティにも配慮しつつ、「人」も「機械」もいつでも利活用（再利用）できる形式かつ可能な限りオープンな利用条件で提供されるべきである。仮に利用に何らかの制約を課さざるを得ない場合は、二次利用

条件の整備が必須である。提供に際しては、全文検索の実現も含めて各情報資源に適切にユーザをナビゲートするような統合的なオンラインサービスの再構築や検索エンジン等民間によるサービスを通じてのアクセスを視野に入れた、あらゆる「人」に対するインクルーシブなインターフェイスの整備（障害者や高齢者等への対応を含む。）が必要である。同時に、API や機械学習等のためのデータセットとしての提供といった「機械」に対するインターフェイス整備のための取組の一層の充実も視野に入れるべきである。あわせて、例えば所蔵機関をまたいだ関連資料のひも付けや地理情報等他分野のデータとの連携等、コレクション間の相互連関のための基盤として、関係機関とも調整しながら識別子も含む典拠情報等を戦略的に拡張・整備していく必要がある。特に識別子等を活用した多様なコンテンツの利用促進という面では、利用者により近いサイドでコンテンツへのアクセスを効率化するアプリケーション（ミドルウェア）の試験的開発等の具体的取組も想定されよう。

また、物理的・地理的・時間的な制約を克服するためにも、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に一層取り組むべきである。あわせて、図書館資料の複製物のデジタル送信や絶版等資料の「図書館向けデジタル化資料送信サービス」による利用者への直接送信の開始に当たっては、権利者の利益保護にも配慮し、文化庁を始め関係諸機関・諸団体とも協議しつつ取り組むべきである。「データ駆動型研究」等の学術研究を目的とする場合には、権利上制約のあるものも含めたデータの利活用を可能とする枠組みの整備も求められる。

一方で、ただ環境を整備するだけではデータの利活用が進まないのも事実である。これを進めるためには、国立国会図書館においても若手の研究者や大学院生等を主たる対象としたフェロウシップ制度や民間も巻き込んだ共同研究等により、「データ駆動型研究」の最新動向を踏まえた外部の知見や資源を広く取り込むことを通じて、従来の学術界のみならずシチズンサイエンスの進展にも寄与していくべきである。

また、急激に ICT の導入が進んでいる教育の現況を踏まえ、デジタル化資料や WARP、「ジャパンサーチ」等を活用したアクティブラーニング等教育シーン（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に取り組むことは極めて重要である。特に、長期化する新型コロナウイルス感染症への対応により、学校や図書館への立入りが引き続き制限される中、著作権法改正により可能となる絶版等資料の「図書館向けデジタル化資料送信サービス」による利用者への直接送信及び図書館資料の複製物のデジタル送信は、生徒・学生の家庭等での主体的な学びに大きく資するものであり、充実したサービス提供体制を構築すべきである。

さらに、国立国会図書館ならではのデータを用いた教育・研究、あるいは、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ジェクト等を推進することが考えられる。これらの推進のためには、国立国会図書館において適切な組織の整備・強化を行うとともに相応の人材を育成、確保する必要がある。

なお、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では、プラットフォームとしての「ジャパンサーチ」の更なる連携拡充は極めて重要である。「ジャパンサーチ」による分野や産官学等の垣根を横断した連携により、「国立国会図書館サーチ（又はその後継サービス）」により集約する書籍等分野のデジタルアーカイブ利活用への相乗効果、海外へのコンテンツ発信、教育シーンでの活用、各地域・分野のコンテンツの掘り起こしという視点でも成果を期待できることから、積極的に取り組むべきである。

(2) 資料のデジタル化・全文テキスト化等の推進

(1)で述べたデータの利活用を推し進めるためには、紙等のアナログ媒体に記載された情報のデジタル化が必須であることから、全ての所蔵資料のデジタル化を長期的には実現すべきである。しかし、当面は、国内刊行出版物について刊行年代や分野、資料群に応じた対象の優先順位付けを行いつつ、年代・資料群共に対象をより拡大して着実に進めるべきである。一方で、国会の附属機関として、国会情報についてはより積極的にデジタル化を行い、国会と国民をつなぐ取組に努めるべきである。

しかし、「人」が読むにせよ、「機械」が読むにせよ、もはや従来の「画像化（PDF化）」だけでは不十分である。デジタル化資料については、全文テキスト化や写真などの画像抽出等の技術を基にした全文検索、画像検索といった所在検索に加えてXML等による構造化等を行うべきである。あわせて、これらのデータについてデータセットとしての利活用も視野に入れるべきである。その際、光学式文字認識（OCR）やレイアウト認識等の技術開発が課題となるが、開発したプログラムをオープンソースとして公開する等を通じて外部の知見・技術も取り込みながら技術的開発を進めていくための体制・環境の継続的・発展的な整備が必要である。また、関係諸機関・諸団体とも丁寧な協議を行いながら進める必要があるだろう。もとより、この取組は、視覚障害者等のアクセシビリティ改善にも資するものである。

ここで述べたデジタル化及び全文テキスト化は、本提言に盛り込まれた取組の出発点となるものであることから、重点的に取り組んでいくことを期待する。なお、デジタル化及び全文テキスト化の対象拡大に伴い、全文検索からのオプトアウト等による個人情報保護やプライバシーへの配慮等も必要となる。透明性が確保された手続と基本的な判断基準が求められる。

(3) 多様な文化資源の収集・保存

(1)や(2)で扱うデータをより広く利活用してもらうためには、多様かつ大量のデータの収集・保存（長期アクセス保証）が欠かせない。国立国会図書館が日本及び日本語に係る多様な文化資源の収集を行うに際しては、国立情報学研究所や科学技術振興機構といった国内関係機関と分担しつつ、日本及び日本語に係る知識基盤、すなわち「ナショナルコレクション」を構築するという視点を持つ必要がある。

まず、従来の有体物・オンラインの資料収集・購読については、継続・強化に努めるべきである。とりわけ、かねて課題となっている有償等オンライン資料の収集については、恒久的な収集・保存のための仕組みが存在しない現状のままでは我が国の知識基盤に大きな欠落を残すことになるため、早急な対応が必要である。ただし、当面の進め方として、学術情報について先行的に収集を開始する等、戦略的に推進する方向性も考えられる。外国刊行資料については、引き続きコアジャーナル中心の電子ジャーナル等の購読及び海外刊行の国内学協会誌の収集に努めるべきであるが、今後はオープンアクセスの進展に応じて、柔軟に見直す視点も求められる。

また、未収資料の収集にも取り組むべきである。例えば、地域資料や海外機関所蔵の日本関係資料等、国立国会図書館が未収の紙媒体資料については、デジタルデータでの収集を視野に入れて取り組むべきである。さらに、研究サイクルのデジタル化やオープンサイエンス、EBPM（証拠に基づく政策立案）の推進を背景に、科学技術情報は、論文だけではなく、その根拠となる研究データも重要な要素となり、大学、研究機関、学協会等のデータプラットフォームの整備が進みつつある。研究データ等の科学技術情報に関して、国立国会図書館は、国立情報学研究所や科学技術振興機構等の関係機関と分担し、国全体として科学技術情報へのアクセスを向上させることが望まれる。この場合、国立国会図書館においては、他ではカバーできない分野、例えば、公的機関の研究データや政策データ、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データの収集・保存、WARP 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承継等を担当することが考えられる。国立国会図書館が、可能なところから承継条件の整理等を行い、アーカイブの支援に取り組むことに期待したい。また、民間ウェブサイトの制度的収集についても、我が国全体の課題として検討を進めるべきである。

メタデータについても、第三者による他機関・分野のデータをつなげる利活用を見据えた上で、オンライン資料へのDOI付与等の識別子の整備と併せて、「ジャパンサーチ」等を通じた収集を引き続き行うべきである。その際、デジタル化やメタデータの整備が進んでいない機関については、国立国会図書館が整備を支援することも検討すべきである。

その上で、これらの収集したデータへの長期アクセス保証を実現するために、災害等に対するレジリエンスを備えた持続可能なアーカイブ基盤の整備に努めるべきである。国立国会図書館に対する信頼感は、この点の実行により担保されるものであろう。この分野においても、国立国会図書館がこれまでの取組で得た知見をいかして、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を醸成する等、各機関（各識別子の国内登録機関を含む。）への支援を検討すべきである。

IV おわりに

「真理がわれらを自由にする」という崇高な理念の下に設置されている国立国会図書館は、その理念を常に念頭に置きつつ、同時に社会及び技術環境の変化に即して、時には大胆な変化を恐れず、どのようにその使命を実現するかということを考えなければならない存在であるということを改めて確認したい。

「人と機械が読む時代」を迎えつつある今日、国立国会図書館が人々に信頼され得る知識基盤であり続けるために取り組まなければならないことは、本提言において述べたように多岐にわたっている。「人と機械が読む時代」の特徴は、「人」が通覧することが不可能なほど多様で膨大なデータをAI等の機械が扱うようになることである。「機械」を読者とすることで、国立国会図書館がこれまで蓄積してきた膨大な資料から検索により必要な知識を瞬時に見つけ出すことを可能にするにとどまらず、例えば日本及び日本語に係るあらゆる分野の知見が集約された大量のデータの中から「人」では気付くことが難しいパターンを見いだすことを可能とすることで、新たな研究や課題解決への道が開かれていこう。

データがあらゆる社会・経済活動の基盤となる現代において、国立国会図書館がこれまで蓄積してきた資料を「機械が読める」形でデジタル化し、信頼に値するデータとして利活用可能とすることの意義は計り知れず、これに最優先に取り組む必要がある。これを着実に進めるために、社会からも広く理解されやすい、マイルストーンとなるような具体的な目標を設定することを次期基本計画の策定に際して強く期待する。

また、オンライン資料なくして今日の学術情報流通は十全なものとはなり得ない。それゆえ、国内で生み出される有償等オンライン資料を含むオンライン資料の国立国会図書館における収集と恒久的保存の意義が広く国民に理解され、これまで紙の資料をベースに形成されてきた我が国の知識基盤が、オンライン資料についても同様に国立国会図書館に形成されることを期待する。同時に、この基盤に欠落が生じることがないように、あらゆる関係者がその社会的責任を自覚して責務を果たすことを願う。

本提言が、これからの国立国会図書館にとって意義ある指針として、有効に活用されることとなれば幸いである。

「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（案）
 ー「人と機械が読む時代」の知識基盤の確立に向けてー」の概要



用語集

API（外部提供インターフェイス）

Application Programming Interface の略。システムが提供するデータや機能を外部のソフトウェアから呼び出して利用するための手順やデータ形式等を定めた仕組みをいう。アプリケーションの開発者は、API に従って機能を呼び出す短いプログラムを記述すれば、自分でプログラミングすることなくその機能を利用したアプリケーションを作成することができる。

DC-NDL

National Diet Library Dublin Core Metadata Description（国立国会図書館ダブリンコアメタデータ記述）の略。メタデータ記述に用いる語彙の国際標準"Dublin Core"（ダブリン・コア）を基に、国立国会図書館が独自の拡張を加えたもの。

EBPM（証拠に基づく政策立案）

Evidence-Based Policy making の略。定まった定義はないが、日本では、政策目的の明確化、政策手段の選択、政策効果とコストの検証など、政策プロセス全般を対象として、統計等データをはじめとする関係する様々な証拠（エビデンス）を活用し、政策形成を改善することをいう。

Society 5.0

「サイバー空間（仮想空間）とフィジカル空間（現実空間）を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する、人間中心の社会（Society）」をいう。狩猟社会（Society 1.0）、農耕社会（Society 2.0）、工業社会（Society 3.0）、情報社会（Society 4.0）に続く、日本が目指すべき未来社会として、第5期科学技術基本計画で提唱された。

オープンアクセス

論文等の学術情報が、インターネットを通じて自由にかつ無償で提供されていること。閲覧、ダウンロード、コピー等適法な用途で利用する場合には問題なく利用できること。

オープン化

インターネットを通じて広く公開されており、商業利用も含めて、目的に応じた活用可能な条件が明示され、手続を要せずに提供されるデータが第三者に自由に利用できるようになっていること。

オープンサイエンス

定まった定義はないが、情報通信技術を活用して科学を変容させることであり、公的研究資金を用いたものをはじめ、研究成果（論文、生成された研究データ等）をインター

ネットの上で広く共有することをいう。オープンアクセスや研究データのオープン化、論文査読プロセスのオープン化、シチズンサイエンスなど、科学と社会の相互信頼を確保しながら、イノベーションの創出にもつなげることを目指したサイエンスの新しい方法論やスタイルを模索する様々な活動を包括的に捉える概念とされる。

オープンソース（ソフトウェア）

インターネット等を通じて、プログラミング言語で書かれたソースコードを無償で公開し、一定のルールの下で誰でもソフトウェアの改良や再配布を行えるようにしたソフトウェアのこと。

オープンデータ

国、地方公共団体及び（民間）事業者が保有するデータのうち、加工、編集、再配布等の利活用が容易にできるように、インターネット等で公開され、営利目的、非営利目的を問わず無償で二次利用可能なルールが適用された機械可読なデータをいう。

オープンライセンス

インターネットを通じて広く公開し、かつ利用条件を商用利用についても制限なく自由に無償で利用することを可能として、そのことを明示すること。

オプトアウト

事前に了解を得るのではなく、事後的に取扱い方法の修正を受け付ける仕組みをいう。

オンライン資料

インターネットで公開される電子情報で、冊子体の図書又は逐次刊行物に相当するもの（電子書籍、電子雑誌等）をいう（国立国会図書館法第 25 条の 4）。国立国会図書館は、国立国会図書館法に基づき、民間のオンライン資料を収集している。ただし、現在は、無償かつ DRM（技術的制限手段）が付いていないものを制度収集の対象としており、有償又は DRM が付いているもの（「有償等オンライン資料」という。）はその対象外となっている。

クラウドソーシング

不特定多数の人（crowd）に委託（outsourcing もしくは sourcing）することの造語であり、インターネットを介して人手やスキルを活用する仕組み。

クリエイティブ・コモンズ・ライセンス（CC ライセンス）

著作権者が著作物の利用条件を簡便に意思表示できるようにするためのツールであり、国際非営利団体クリエイティブ・コモンズが提供している。このうち、CC0 は、権利者が、その著作物に対して有するすべての権利を法令上認められる最大限の範囲で放棄し、パブリック・ドメインに提供することを意味する。また、CCBY は、クレジット（原作者の氏名、作品タイトル等）を表示することを主な条件として、改変のほか、営利目的

の二次利用を許可することを意味する。

コアジャーナル

ある専門分野において重要度が高く、利用が多い雑誌の一群のこと。

光学式文字認識 (Optical Character Recognition: OCR)

画像データを機械・ソフトウェアに読み取らせ、テキストデータを生成すること。

持続可能な開発目標 (Sustainable Development Goals: SDGs)

2015年9月の国連サミットで採択された「持続可能な開発のための2030アジェンダ」に掲げられた2030年までに持続可能でよりよい世界を目指す国際目標をいう。地球上の「誰一人取り残さない (leave no one behind)」ことを掲げ、17の目標と169のターゲットから構成される。その前身にあたる2001年に策定されたミレニアム開発目標 (MDGs) と異なり、発展途上国だけでなく、先進国における取組みも求めている。

シチズンサイエンス

職業科学者ではない一般の市民によって行われる科学的活動、市民参加型のサイエンスをいう。日本では、社会課題の解決に重きを置いた「市民科学」と呼ばれる活動もあるが、シチズンサイエンスは、市民科学に加えて、学問体系における科学的規範に則った知識生産も包含する、より広範な科学的活動とされている。

デジタルアーカイブ

博物館・美術館、図書館や公文書館をはじめとする様々な主体が保有する多様な分野のデジタル情報資源を収集・保存・提供する仕組みの総体をいう。デジタルコンテンツに限らず、アナログ媒体の資料等も含むコンテンツのメタデータやサムネイル等も扱う。

デジタルオブジェクト識別子 (Digital Object Identifier: DOI)

コンテンツの電子データに付与される国際的な識別子。国際DOI財団が管理しており、日本では、科学技術振興機構 (JST) が事務局を務め、国立国会図書館、国立情報学研究所 (NII)、物質・材料研究機構 (NIMS) の4機関が共同運営するジャパンリンクセンター (JaLC) が、同財団の認定を受けたDOI登録機関となっている。DOIは、各機関固有のプレフィックスと個々のコンテンツを特定するサフィックスの文字列から構成される。DOIをウェブブラウザ等に入力するとペアで保管されているコンテンツの所在情報 (URI) に変換される。このため、URLに変更があっても永続的なアクセスが可能となる。文献のほか、研究データ等にも付与される。

デジタルトランスフォーメーション (DX)

広くデジタル化による社会経済の変革をいう。文部科学省の科学技術・学術審議会学術分科会・情報委員会の提言『コロナ新時代に向けた今後の学術研究及び情報科学技術の振興方策について』(令和2年9月30日) (https://www.mext.go.jp/b_menu/shingi/mext_00)

[537.html](#))では、研究のデジタルトランスフォーメーションを次のように定義している。

「ネットワーク・計算資源等の情報システム基盤の整備・高度化、データや学術情報のデジタル基盤の整備・高度化、研究施設・設備の遠隔化・スマート化等を推進することにより、多様な研究形態や、効果的・効率的な研究プロセスの実現、更には、データ駆動型研究や AI 駆動型研究をはじめとする情報科学技術を活用した新たな科学的手法の進展等を促進することをいう。」

電子情報資源

電子的方法や磁気的方法で記録された情報全般のこと。インターネット資料、オンライン資料、デジタル化資料、研究データ等を含む。

プレプリント

査読前論文のこと。

マイグレーション

長期保存のために行うファイルフォーマットの変換や媒体変換等をいう。

メタデータ

コンテンツの内容・特性、所在等を簡便に記述・表現したデータ。図書館における書誌データ、アーカイブ機関の収蔵品等の目録データ、文化財の基礎データ等のテキストデータのほか、URI 参照のデータ等の電子情報資源の管理のために必要なデータも含む。

リポジトリ

研究成果である論文など、大学や研究機関等において生産された電子的な知的生産物の保存や発信・提供を行うためのインターネット上のアーカイブシステムをいう。