1. **標題（和文および英文）**

電子図書館サービスからナショナルアーカイブの構築へ

－LOD化によるデジタル文化財の利活用を目指して－

From digital library services to the construction of National Archives

－Aiming at utilization of digital cultural heritage by making LOD－

1. **著者名（日本語およびローマ字）**

中山　正樹

NAKAYAMA Masaki

1. **著者の所属機関名および部署名（和文および英文）**

元国立国会図書館　専門調査員・電子情報部長

（同志社大学大学院総合政策科学研究科　嘱託講師）

Former Director General of Digital Information Department of the National Diet Library

1. **著者の所属機関の住所（和文および英文），E-mail アドレス**

nakayama@slis.doshisha.ac.jp

1. **著者抄録（和文および英文）**

筆者は，国立国会図書館に在職中，電子図書館サービスシステムの設計・構築と運用に携わり，今後の方向性としてのナショナルアーカイブ構想の立案に関わってきた。

国立国会図書館は，電子図書館構想を策定し，「電子図書館サービス」を進展させてきたが，今後は，「電子図書館サービス」から，図書館の枠を越えてあらゆる分野の情報を関連付けて利活用する基盤として「知識インフラ」の構築へ発展させ，その一翼を担うこととした。その実現形が，「ナショナルアーカイブ」の構築である。

「電子図書館サービス」の設計・構築において，筆者が意識的に適用してきたセマンティックウェブに関連する技術を振り返る。また今後，ナショナルアーカイブの構築において，関係機関が連携して何を目指し，LOD化技術を含めたセマンティックウェブの技術を，どのように体系的に適用していくべきかを展望する。

The author, during his tenure at the National Diet Library, had been involved in the design, construction and management of digital library service system, and in making plans of National Archives initiative.

The National Diet Library, formulating “Electronic Library Concept”in 1998, has developed digital library services, and is now seeking to build "knowledge infrastructure" as the basis for associating the information in all fields beyond the framework of library and utilizing them.

This paper reflects on the technology related to the Semantic Web, which the author had consciously applied in the design and construction of digital library service system, and also puts in perspective what goals the NDL should aim at,in cooperation with related institutions for the construction of the National Archives, and how to systematically apply the technology related to the Semantic Web including LOD.

1. **著者付与キーワード（和文および英文）**

ナショナルアーカイブ, デジタルアーカイブ, 知識インフラ, セマンティックウェブ, Linked Open Data, クラウドコンピューティング, WebAPI電子図書館, 国立国会図書館, 国立国会図書館サーチ,

National archive, Digital archive, knowledgeinfrastructure, Semantic web, Linked Open Data, Cloud computing, Web API, Digital library, Natonal Diet Library(NDL), NDL Search

2015年8月29日

電子図書館サービスからナショナルアーカイブの構築へ

―LOD化によるデジタル文化財の利活用を目指してー

元国立国会図書館　専門調査員・電子情報部長

中山正樹

# はじめに

筆者は，民間企業，政府関係の独立行政法人において，情報システム化により，業務サービス及びソフトウェア開発の効率化に取り組んできた。2002年に国立国会図書館（NDL）に入館して以降は，電子図書館事業の企画立案，図書館サービスシステムの構築・運用に携わり，2011年より電子情報部長および情報化統括責任者（CIO）として，電子図書館事業，情報システム全般を統括し，次世代図書館サービスの構築に向けて検討を進めてきた。筆者の信条は，一貫して，労働集約的な業務を情報システムにより効率化して，人がより創造的な業務へ移行できるようにすることである。

図書館に限らず，様々な業種・業態の組織において，大量の情報の中から効率的に適切な情報を取り出す情報提供システムが構築・運用されている。情報システムの観点では，図書館システムは，そのような情報システムの中での１アプリケーション分野であり，情報提供サービスとしても，利用者にとってOneOfThemのサービスであると考えている。

筆者は，そのような基本的な考え方で，NDL以外の情報資源，技術ノウハウを活用して，利用者の情報取得，知識創造を効率化するサービスを，如何にして効率的に構築・運用できるようにするかを念頭に進めてきた。次世代図書館サービスシステムは，従来型の図書館の延長線で考えるのではなく，世の中の多くの情報提供サービスの中で信頼性の高い情報を提供する機関としての責任と義務を果たせるサービスの提供を目指すべきとしてきた構想を論述する。

昨今，クラウドコンピューティングの普及に伴って，ウェブ上の情報を意味的に理解できる形で公開・共有を目指すセマンティックウェブ技術の中で，構造化された情報同士をリンクさせるための技術や方法論の総称としてのLinked Open Data(LOD)が注目されてきた。この概念は，NDLが1990年代に構想した電子図書館構想の「地球規模の知的財産を誰でも容易に利用できるようにする」という理念の実現のために，電子図書館サービスの構築の仕組みの考え方や，NDLが個別に適用してきた要素技術の集合としてみることができる。

筆者が，電子図書館サービスの構築において，セマンティックウェブ，LODの考え方を，どのように適用してきたか，今までの歩みを顧みるとともに，今後，電子図書館サービスの発展形としての国の文化財全体の利活用を目指したNDLを中核とした情報保有機関によるデジタル文化財のアーカイブ構築において，LODの考え方がどのように適用されていくべきかの方向性について，技術的な観点にフォーカスして考察する。

# 電子図書館20年の歩み

1980年代から，海外の多くの先進的な図書館同様，日本でも電子図書館事業に取り組み，1994年に我が国で最初の大規模な電子図書館の実証実験プロジェクトがNDLにより実施された。その後，NDLは，2002年から本格的なサービスとして離陸し発展させて，現在に至っている。



図1電子図書館に関連する構想とサービスの変遷

## 電子図書館構想と実証実験

1994年1月，通商産業省（現：経済産業省）の高度情報化プロジェクト事業の一環で，情報処理振興事業協会（現情報処理推進機構（IPA）は，慶應義塾大学の湘南藤沢キャンパス(SFC)内に，情報基盤センター（CII）を設置した。1）当時としては高速のネットワーク（6Mbps），大容量サーバ（4TB）を配備して，各省庁の白書のSGML化や，各種統計情報を串刺しにして連関を分析する「新産業創造データベース事業」，学校教育の高度化を図るための教育ツールとしてのオーサリングシステム，エグゼキュータシステムを提供する「100校プロジェクト」とともに，NDLと連携協力して，パイロット電子図書館プロジェクトを開始した。筆者は，本事業には，IPA側の担当として参画した。

パイロット電子図書館プロジェクトの目的は，21世紀の高度情報社会において，地球規模の知的財産を誰でも容易に利用できるように，地球上に広く分散して個々に収集・蓄積されている知的資源を，空間的・時間的制約を越えてアクセス可能とする環境を提供するための実証実験であり，NDLは，全国公共図書館総合目録ネットワークプロジェクトとパイロット電子図書館実証実験プロジェクトを実施した。

総合目録ネットワーク実験は，出版物の総合目録と図書館間の相互貸借のためのシステムとして，NDLを含む58館の政令指定都市および都道府県立図書館が参加して実施された。共通のメタデータ記述規則としてMARCベースの総合目録共通フォーマット，通信プロトコルとしてファイル転送規約を設定し，各図書館からの大量の初期データ，週次の差分データを受け取ることにより，1,400万件の書誌データがRDBMSの１つであるOracleデータベースに蓄積された。ISBN等の識別子がないメタデータも多く，タイトル，著者等で同一の書籍のメタデータであることを認識する機械的な同定作業を行った。所定の時間内に処理するためには，レスポンスを低下させないデータベースのテーブル設計，SQL言語によるプログラミングでは様々な工夫が実証実験された。

パイロット電子図書館実証実験では，NDL所蔵の貴重書，明治期刊行図書・雑誌，出版社から提供を受けた資料等を，マイクロフィルムからのスキャニング，デジタルカメラでの撮影により，当時としては高精細画像（4,000×5,000ピクセル）で，tiffおよびJpegフォーマットで約1,000万ページ（CD-ROM約2000枚分）のコンテンツをデジタル化し，画像データベースに登録。個々の画像データにメタデータを付与してメタデータデータベースを構築して，効率的な検索・利用方法，ユーザインタフェース（UI），電子化データの高度利用，効果的な電子図書館の構築支援の手法等について実証実験を行った。その成果であるシステム及びデジタル化コンテンツは，その以降の電子図書館サービスのベースとして引き継がれている。

## 国立国会図書館中期計画2004

NDLは，パイロット電子図書館プロジェクトの成果を踏まえて，2002年10月に開館された関西館を拠点として，近代デジタルライブラリー，インターネット資源の選択的収集事業（WARP），各種の電子展示会を公開・提供した。

2003年には，e-Japan重点計画2003，e-JAPAN戦略Ⅱ加速化パッケージ，（内閣官房IT戦略本部）において，「国のデジタルアーカイブ構想」，「ジャパン・ウェブ・アーカイブ構想」の実現を，また2004年には，e-Japan重点計画2004において「国立デジタル・アーカイブ・ポータル構想」を一層推進することが明記された。

　このような動きを踏まえて，NDLは，2004年2月に「電子図書館中期計画2004」（中期計画）を策定した。この中期計画において，デジタルコンテンツを広汎な利用者に提供するために，NDLが国のデジタルアーカイブの重要な拠点となるということ，また国内外の多様な利用者層の需要に応じ，日本のデジタル情報全体へのナビゲーションを行う総合サイトを構築し，利用者がワンストップで利用できるようにすることを目指すこととした。2）

### アクションプランとしての3つの柱

１つ目は， 「デジタル･アーカイブの構築」 である。これには，書籍のデジタル化の推進と，インターネット情報の収集と恒久的保存の２つがある。２つ目は，「情報資源に関する情報の充実」で，レファレンス情報などナレッジの蓄積を蓄積し提供提供すること，3つ目は，「デジタル･アーカイブのポータル機能」である。

### 電子図書館中期計画2004の実現イメージ



図2 NDL電子図書館サービスの全体像

図2は，2004年にNDLが作成した電子図書館サービスの実現のイメージである。

右下の「デジタルアーカイブ」は，デジタル化したリソースやインターネット情報などを蓄積し，長期保存する。

その左は，既存システムのNDL-OPACなどの「蔵書目録」は，これらが書誌情報やデータベース内のメタデータを提供する機能を備えることにより，デジタルアーカイブポータルから統合的に利用できるようにする。

上の「ポータル」は，これらNDLの資源だけでなく，左下にある，他機関のデジタルアーカイブや目録などの有用情報も合わせて案内できるようにするもの。

左上の「ポータルのインタフェース」は，GUIにより利用者へのニーズに合わせた検索手段の提供と，他システムがポータルの機能を利用するためのAPI機能を持つ。

また，右の，「主題情報，目次情報，解題情報，レファレンス情報など検索を支援するための情報，知識としての「ナレッジデータベース」である。

2004年に描いたイメージの実現形として，ポータルを発展させたものが，PORTAであり，また，現在の「国立国会図書館サーチ（NDLサーチ）」である。

## デジタルアーカイブの構築

資料のデジタル化に関しては，NDLは，2009年5月から大規模なデジタル化を開始し，2011年までに，冊子体としては230万冊，約2億枚の画像をデジタル化した。また，インターネット資料の収集に関しては，NDLは，2010年4月に国等のウェブサイトの制度的な収集を開始した。さらに，いわゆる電子書籍・電子雑誌に類するオンライン資料は，2013年1月から，無償でDRMによるアクセス制限が掛かっていない資料についてのみの収集を開始した。デジタル化および収集した資料の提供に関しては，著作権保護期間の満了もしくは著作権者から許諾を得た資料は，インターネット上に公開している。デジタル化した資料のうち，市場で入手困難な資料に関しては，公共図書館等で閲覧が可能になっている。ここでは詳細は割愛するので，他の文献を参照していただきたい。本稿では，保存システムの機能として，メタデータに関連した施策として，「長期保存システム」のみ紹介する。

### デジタル情報の長期保存（OAIS, 情報パッケージ(SIP, AIP, DIP)）

電子情報を長期的に保存するためのシステムの構築に当たっては，NDLは，この長期保存システムとして必要な機能やプロセスを抽象的に定めたモデルを参照しつつ構築している。

その一つとして，OAIS（Open Archival Information System）レファレンスモデル(ISO 14721:2003)モデルがある。

このレファレンスモデルに沿ったパッケージとして，米Hewlett-Packard 社のDspace，IBM社のDIASがあった。受入機能として適切なメタデータの付与，アーカイブ（保存）機能として定期的なエラーチェックや保存媒体の更新，データ管理機能としてメタデータの管理，保存管理・運営機能として長期保存計画の立案，提供者との交渉，マイグレーションの実施，さらに保存計画機能として旧式化の検知，対応策の提示，アクセス機能として再生環境の保存，エミュレーションが想定されている。

このうち，保存計画を実現する要素技術であるフォーマットレジストリやフォーマット識別ツールの開発が海外で行われている。

情報パッケージとしては，MODSをベースにしたメタデータをMetsで括ったSIP（受入れメタデータ），AIP（保存メタデータ），DIP（提供メタデータ）を想定した。

## デジタルアーカイブポータルのプロトタイプ

NDLは，中期計画での3つ目の柱であるポータル機能について，2004年10月から開発を開始した。様々なデジタルアーカイブ内の情報を統合検索する仕組みの実用性を検証するために，デジタルアーカイブポータルのプロトタイプシステムとして，2005年7月に試験公開し，機能の検証を開始した。これが，現在の分野を越えたナショナルアーカイブ構想の原点であったと考える。

### 開発当初のサービス要件とシステム化

#### サービス要件定義において

* 利用イメージとして，利用者の情報探索行動の全般をサポートすることを目指すこととした。
情報探索行動として，問題の定義，情報ニーズの識別，情報探索戦略の策定，情報の獲得（情報源の咀嚼，情報の抽出），情報の統合（情報の組織化，提示），評価（成果の評価，プロセスの評価）等のステップを想定した。このようにして目的の情報の存在を知るだけでなく，情報そのものの閲覧および入手先までナビゲートを目指すこととした。
* それを実現するために，各デジタルアーカイブの構築の在り方にも言及した。交換用の標準的なメタデータ記述要素，記述規則の適用，交換用APIの実装を求めることとした。また，巨大な知識ベースとして「意味ある情報資源」として利用できるように，本文内容を組織化，検索できるように研究開発，技術開発が進むことも求めた。

#### システム化において

* 先進的かつ将来標準的な仕様となることが見込まれる技術を適用すること。
* オープンソースソフトウェア（OSS）・パッケージソフトを活用することとして，特に，適用事例が多いOSSを活用し，可能な限り新規開発はせず適用したソフトウェアのカスタマイズは必要最低限とすることとした。
* Webサービスとして他サービスから利用しやすいものとすることとして，各々の機能は独立したWebサービス機能として提供し，図書館界のみならず，デジタルアーカイブの世界で標準となり得る仕様の採用を目指した。

## PORTAの構築

NDLは，プロトタイプでの検証を踏まえて，実用システムとして「PORTA」を構築し，2007年10月に正式公開した。3)

### システムアーキテクチャ

NDLは，システム基盤として，Web2.0時代のサービス指向アーキテクチャ（SOA）を適用し，仮想OS上でOSSを組み合わせ，Dublin Core（DC）に準拠した国際標準のメタデータ記述要素，記述規則，メタデータ交換の共通API（OAI-PMH，SRW（SRU/SOAP），RSS，Z39.50等）を利用して，商用を含めた外部サービスとマッシュアップによるサービス連携と，コンテンツマネジメントシステム（CMS）による機能追加の容易性の確保を目指した。先進的な検索機能として連想検索検索エンジン（GETA）等を利用したあいまい検索も実装した。

### 統合検索対象の拡大

統合検索対象は，画像イメージ形式で提供している近代デジタルライブラリーとテキスト形式で提供している青空文庫の統合検索から始め，同時代の資料が分散保存されている国立公文書館アジア歴史資料センターとNDL-OPAC，地域情報をまとめているアグリゲータサイトとしてデジタル岡山大百科等を，ワンストップで検索しコンテンツに辿り着けるようにした。

### 統合検索のために苦労した点

* データ提供機関に対して，統合検索への賛同を得ること。大きな意義は，各機関のコンテンツの利活用が促進されることであり，また，利用者の利便性が向上することであったが，サービスが横取りされると誤解されることが多々あった，また，サイトへのアクセスが増えることを不安視され，なかなか賛同が得られなかった。
* また，連携には，提供機関のサイトで標準プロトコルの実装が必要であったが，ほとんどのサイトで，外部提供インタフェースを持っておらず，各サイト側に実装が必要であったが，開発ベンダーが法外な開発費を提示し，その費用が捻出できない状況であった。
* さらに，メタデータのマッピングの調整においては，それぞれの機関において， Dublin Core（DC）をベースにした記述要素を適用できても，記述規則が異なり，統合しても内容として認識できない状況であった。特に，横断検索においては，メタデータの表記のゆれをカバーした検索はできず，調整には，膨大な時間を費やした。汎用的なメタデータのマッピング，記述規則等のルールの普及が必要であった。

### インキュベータとしてのPORTAの役割

筆者は，PORTAは唯一の統合検索サービスではなく，PORTAがインキュベータとなって，各利用者ニーズにあった様々な統合検索サービスが出現することを期待した。

* まず，継続して，デジタルコンテンツの可視化を目指すこととし，インキュベータの役割として，各データプロバイダに検索やサービス連携のためのAPIが実装されることの普及啓発活動を継続するとともに，外部機関においてPORTAの外部APIを利用した様々なサービスが出現することを期待した。
* また，利用者の情報探索行動において，統括可能情報の網羅性を高めるために，様々な機関から提供されるデータベースを分散デジタルアーカイブと位置付け，コンテンツの体系的な構築と，網羅性の高い統合検索サービスの実現を目指すこととした。
* さらに，単なるキーワードでの検索サービスではなく，意味的に関連付けるセマンティックウェブ・サービスでの連携を目指すこととした。

### Web3.0の時代を見据えて適用すべき技術と進め方

筆者は，今後の機能強化のため，セマンティックウェブ技術を中核として，適用すべき技術として3つに分類した。

#### 収集を容易にする技術

収集効率を高める技術（差分収集と再現技術を含む），収集品質を高める技術，収集したサイトイメージから著作物を切り出す技術の適用を想定した。

#### 管理を容易にする技術

* メタデータの付与を自動化・省力化する機能

書誌的事項は，情報の作成元が作成することを基本としつつ，管理，検索に十分なメタデータが付与されていない場合は，多少精度が低くてもメタデータは可能な限り自動付与する機能，自動付与されたメタデータの確認作業を省力化する技術の適用を想定した。

* セマンティックウェブ技術の活用

当時，集合知を利用するWeb2.0から，意味的情報を相互利用するWeb3.0が叫ばれた時期であり，Web2.0の要素として，Blog，Forksonomy，Ajax，Mashup等の技術，更に

Web3.0の要素として，RDF，SKOS，オントロジー，GRDDL，SPARQLの技術の適用を想定した。

* 蔵書管理指向からサービス指向へ

将来的には，目録を利用者側の視点から見直すモデル（FRBR）に基づくメタデータの再構築も想定した。

* 検索を容易にする技術

従来型のキーワードマッチングによる検索だけでなく，キーワードサジェストやクラスタリング検索等により，検索候補を推定する技術を活用した検索機能も想定した。

#### PORTA構築の進め方

筆者は，進め方として，図書館の枠を超えた国際標準，業界標準を積極的に適用するとともに，政府機関，民間を問わず，国内外の研究機関，研究者等と連携して調査研究を進め，関係機関との共同構築，オープンソースの利用，単独でのトータルシステムを目指すのではなく，各機関の研究成果を組み合わせてマッシュアップによるサービス構築を想定した。

まずは，海外ではIIPC，国内では様々な技術の研究開発性を保有しているJST，NII，NICT，AIST，IPA等の政府機関，大学の研究室との連携，有用な検索サービスを実施している商用ポータル機関（Google，Yahoo等）との連携によるウェブ協調型のサービスの提供を目指した。

## トータルな次期図書館システムを目指して

2009年当時の冊子体資料を扱う従来型の図書館システムの最大の問題は，運用・保守作業，機能拡張コストが肥大化して，機能拡張により新たな利用者ニーズに応えるための機能拡張ができなくなったことである。そのために，NDLは，下記のような方針により，全面リニューアルすることとした。

* 紙・デジタルを問わずトータルな次期図書館システムの構築と提供の実現を目指す。
* NDLサーチを，図書館ポータル，次世代OPAC，PORTA，ゆにかねっとを統合した入口機能とする。
* 新規開発を極力なくす。
* 実現すべき機能とレベルを明確にしたシステム化要件定義書により，見積もり内容を的確に判断する。
* 特定ベンダーしか提案できないRFPにならないようにして，競争原理が働く調達を行う。
* 従来型業務・サービスは，パッケージで実現できない業務は，業務の内容や流れを見直し，可能な限りパッケージの仕様に合わせる。

　　この方針に基づきシステム開発を進めた。

冊子体刊行物の収集・組織化業務及び，蔵書検索・申込システム「NDL-OPAC」を，パッケージに切り替えることによる開発・運用コストを大幅に削減し，デジタル化資料の来館者サービスシステム，国立国会図書館サーチ等，デジタル時代の次期図書館利用者サービスの基本機能を充実させた。

## 国立国会図書館サーチ（NDLサーチ）

NDLは，PORTAの後継の情報探索サービスシステムとして，2012年から正式に提供を開始した。のちに提起される「知識インフラ」のフレームワーク作りのベースとしたサービスであり，筆者が振り返るとLinked Data，Linked Open Dataと同一の方向性を指向したものとなっていた。

NDLは，システムの構築に当たっては，様々な視点を持つ専門家の参画のもとで，サービス要件を確定させ，サービス要件を満たす機能を実装するためのシステム化要件を定義した。段階的な実装を想定し，OSSを含めて，その時点での最適な適用技術，パッケージを示す技術標準適用ガイドラインを踏まえて，開発を進めた。

### 機能の概要

NDLは，FRBRのモデルに沿って同一著作，原作から様々な媒体や形式に派生して発行されている著作物を書誌事項で体系的に把握できるように著作単位でグルーピング表示する機能，検索キーワードが明確でなくても想定されるキーワードサジェスト機能，障害者向け機能，検索ワード，外部サービスを活用して内容表示を日中韓英に翻訳して表示する機能，個人ごとの検索要件，検索範囲をあらかじめ特定できるパーソナライズ機能，スマートフォン対応機能等を実装した。

NDLは，交換用メタデータとして，冊子体出版物以外の情報への適用を想定して従来の仕様をよりシンプルにしたDC-NDL(Simple)と，セマンティックウェブに適合した記述モデルであるRDFをデータの記述に使用し，FRBRのWork，Manifestation，Itemを体系的に定義したDC-NDL(RDF)を交換用メタデータ記述要素・規則として定義した。

NDLは，メタデータハーベスト，横断検索のためのWebAPIを設定し，そのAPIの実装を働きかけることにより，連携先の一層の拡大を図るとともに，外部システムでメタデータを利活用のためのWebAPIとして，検索結果や新着書誌情報をFeedly等のニュースリーダで確認できるように，RSS配信する機能，DC-NDL(RDF版)，JSON形式でダウンロードする機能等の実装を強化し，「情報ハブ」の役割を提供した。4)

また，著者名典拠，件名典拠情報に関しては，国立国会図書館典拠データ検索・提供サービス（WebNDLA）として，キーワード検索，NDC分類記号検索により検索できるほか，NDLサーチと連携し，関連キーワードの提示，典拠コントロールされた各標目による書誌データの再検索など多様な検索機能を提供した。また，全典拠データの個別ダウンロード，国立国会図書館件名標目表（NDLSH）の収録範囲となる典拠データを，RDF/XML形式, RDF/Turtle形式, JSON形式で，ダウンロードすることを可能とした。

### サービスの状況

NDLサーチでは，人間文化研究機構の6構成機関の研究資源共有化システム（nihuINT），国立美術館の4つの美術館が所蔵している作品の総合目録，商用の電子書籍ポータル（hon.jp）との連携等のPORTAが担ってきたデジタル情報のポータルとしての役割を継続するとともに，ゆにかねっとが担ってきた各地の公共図書館蔵書の総合目録としての役割を引き継いだうえ，NDL-OPAC，インターネット資料収集保存事業（WARP），国立国会図書館デジタルコレクション，国会会議録検索システム，リサーチ・ナビ（調べ方案内），レファレンス協同データベース等のNDLの主要データベースを統合検索し，図書館，美術館，博物館，文書館，政府機関，新刊書籍の出版予定情報としてJPOの近刊図書情報，Japan Knowledge等の商用データベースなど，200DB，約1億件の書誌・メタデータを横断的に検索できるようにした。検索対象が必ずしもデジタル化されてインターネットで閲覧可能なコンテンツでなく，所在情報により入手手段へのアプローチに留まるものも多いが，今後各機関によるデジタル化が進めば，インターネットで閲覧できるコンテンツへのナビゲーションも充実すると思われる。

また，統合検索のほかに，検索キーワードを使った外部サービス検索として，WorldCat，WebCat Plus，Wikipediaでの再検索，更に入手手段として，各種オンライン書籍・電子書籍販売サイト，古本屋ポータルサイトへのナビゲートも行った。

最近の進展としては，公共図書館からのデータ提供方式のOAI-PMH / DC-NDL(RDF) への切替，ジャパンリンクセンター（JaLC）との連携によりDOIでの固定URLでの検索，WebNDLAからVIAFへのナビゲートも可能になっている。

### 連携方針・連携モデル

統合検索先の拡大においては，連携可能なところから進め，網羅性の確保のため，分野ごとの中核的なデータベースに標準的な連携機能の実装を働きかけてきたが，標準仕様が浸透してきた状況において，NDLは，改めて連携方針・連携モデルを提示し明確にした。

#### 連携実施計画策定の課題・背景

* 日本におけるメタデータ提供のプラットフォームとしての認知度が高まりつつある今，連携可能な範囲での拡張ではなく，目標と計画を設定して内外に示す必要がある。
* ナショナルアーカイブ（構想）におけるコンテンツ検索・提供機能を担うことを見据える必要がある。
* 図書館等の情報機関が保持するメタデータの集約及びAPIを通じた一般への提供により，政府が推進する「公共データの民間開放（オープンデータ）」の一翼を担う必要がある。

#### 連携方針・連携モデル

次の5項目を連携方針の柱とする。

* 日本の刊行物及び刊行物と同等の内容を有するコンテンツの網羅を志向。
* 公的機関，学術研究機関，図書館・文書館・博物館・美術館等の文化機関が作成し，インターネット上で提供している一次情報（コンテンツ），二次情報（メタデータ）及び参考情報等を対象とする。
* 一般利用者にとって有用性が高いコンテンツを持つシステムを優先。その際，一次情報の入手までの障壁が低いシステムを重視。
* APIを実装したシステムを優先。メタデータ授受に用いるフォーマットとして，DC-NDL(RDF)を推奨。
* 効率的に検索対象を拡大するために，個別のシステム（=data provider）との連携よりも，それらを集約した統合検索サービス（=aggregator）との連携を優先。

つまり，APIは普及期にありインキュベータとしての役割を終了し，既にAPIが実装されているシステムとの連携を拡大し，人間文化研究機構，JST，NII等のように，各所管において，統合データベースを提供しているサービスとの連携により，効率的に網羅性を確保するということである。

## 公共的書誌情報基盤

情報内容として整理・要約されている出版物に関する情報の共有化と利活用は最も重要視される。

公共的書誌情報基盤は，出版界が作成する出版情報（販売促進情報）を活用した図書館での書誌作成の省力化を目指すもので，出版界と図書館界が持つ著者，著作の書誌・書評等の情報の共通識別子によるLinked Data化により，出版情報の網羅的なデータベースの構築と共有が期待できる。

### 公共的書誌情報基盤の整備（2010年）

NDLが，出版関係機関と協力し，我が国を代表する標準的な書誌情報を作成・提供する公共的基盤を整備するものである。出版文化の基礎となる質の高い出版・書誌情報が，無償もしくは廉価にて，迅速またタイムリーに読者，利用者に届けられることを目的とするものであった。5)

また，この事業によりNDLの納本事務がより網羅的かつ円滑に行われることが期待された。

当時の想定は，次のようなものであった。

* 近刊情報は，出版物のISBNとNDL書誌IDが関連付けられるように，出版社からISBNが付与された書誌情報が，JPOを経由してNDLに送られ，JPNO, 書誌IDを付与して，近刊情報として公開する。その近刊情報をJPOを経由して取次に送ることにより，共通の識別子で近刊情報が管理される。
* 新刊情報は，取次において，JPO近刊情報と現物を突合し，出版情報追記して，新刊情報とする。その情報をNDLにおいて，近刊情報を新刊情報として置換え，最低限の書誌事項を追記して新刊情報として公開すること。

しかしながら，NDLでの実装においては，納本された新着情報の提供は，出版情報をインプロセス情報として提供することにより早まったものの，出版界と図書館界での情報の共有には繋がっていない。

### 全国書誌情報の利活用に向けた超党派勉強会での検討（活字文化議員連盟　2015年9月）

図書館などで広く活用されているMARCのもととなる書誌データを国民が無料で利用できるように政策や予算なども含め総合的な施策を検討するもの。

背景としては，NDLが作成する書誌データは，民間の書誌データを活用することで迅速な提供に向けて取り組んでいるが，民間の提供スピードには及ばず，公共図書館などでの利活用が進まない現状があり，2016年春をめどに課題のたたき台をとりまとめるとされている。

筆者が期待する結論としては，近刊情報，新刊情報が図書館で活用され，また，図書館での書誌情報，典拠情報が出版界での出版物に関するデータベースで利活用されることにより，出版情報を活用した書誌作成が効率化されるとともに，出版物の検索の網羅性が確保されることにより利用者が出版物を見つけやすくサービスの提供が容易になることである。

## 知識インフラの構築を目指して

新たな知識の創造のためには，分野を越えた知識の関連付けが必要であり，日本中に散在するコンテンツの所在を集中管理し，そこに検索をかければ，関連する全ての必要なコンテンツが得られるようにするものである。そこでは，単に情報を集めたものではなく，関連するものが有機的に結合され，ネットワーク的に統合化されたものであり，日本中にある芸術を含んだあらゆる学問・研究のコンテンツ，研究ツール，社会状況データ等が知識の形に組織化され，これらの知識・情報が公開され，全ての人が共有できる。

### 知識インフラとは

「知識インフラ」とは，情報資源を統合して検索，抽出することが可能な基盤で，国内の各機関が保有する情報を知識として集約し，新たな知識の創造を促進し，知識の集積・流通・活用と創造するサイクルの構築を目指すものである。

### 知識インフラの必要性

第４期科学技術基本計画「科学技術に関する基本政策について」（内閣府　総合科学技術会議2010年12月24日）が答申され，文献等研究情報のデジタル化，オープンアクセスの推進等とともに，「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」とされた。6)これを踏まえて，NDLは，2011年に，第三期科学技術情報整備基本計画－知識インフラ構築に向けて」を策定した。7)

### 知識インフラの構築の目的

知識インフラは，科学技術研究活動の実践を根本で支え，科学，技術，学術，文化活動によって生み出される多様なデータ，情報，文献，関連する情報が組織化され，それらへの迅速で適切なアクセスを可能にすることで，次の研究，開発，教育，その他の社会的・文化的実践へとつなげる動的サイクルを形成することを目的としている。つまり，情報の生産→流通→アクセス→再生産という知識の循環を促進するネットワーク，プラットフォームとなることを目指すものであり，また，組織や個別学術分野を越えた知識の融合を可能とし，学際的な新しい知識やイノベーションの創造を容易にするものである。

### 知識インフラの機能

知識インフラにおいては，文字データだけでなく多様な形式で表現されるデータや情報を対象とし，収集，保存，識別，組織化，検索，表示，公開といった機能を実現させる必要がある。利用者は，大量のデータに対して特定条件に適合したデータだけを抽出したり，多様な分野の情報を一括して検索したり，自分の関心に合わせて実体のリンクやネットワークを形成したりといったことが自由にできることが求められる。

また，単語等による検索だけでなく，自動分類や収集された全体を見通した上での体系化や秩序化がなされることが期待される。8)

### 知識情報基盤の構築モデル

NDLは，NDLサーチの前身のPORTAと，NDLデジタルデポジットシステムが進めてきた概念をベースとして，知識情報基盤の構築モデルを描き，後の「東日本大震災アーカイブ（ひなぎく）」により，大震災の分野に絞った「知識インフラ」の実現を目指した。

「ひなぎく」の実現に当たっては，情報の可視化技術，情報の収集の効率化技術，情報の組織化技術，情報の集合知化技術，情報検索技術，閲覧・表示技術等の次世代技術の研究開発成果の活用なしには実現できないと考える。9)



図3 知識情報基盤の構築モデル

## 東日本大震災アーカイブ

2011年3月には，東日本大震災が発災し，甚大な被害がもたらされた。国は，この大震災に関連する災害現象そのもの，災害前・災害直後・復興の過程，災害時の対応，他地域・次世代への教訓等のあらゆる記録を後世に残すこととした。10)NDLは，それを実現するために，大震災関連にフォーカスした知識インフラの構築の実現形の先行事例として，2013年3月には，大震災に関するあらゆる記録，記憶を保存し，一元的に検索できるようにする「東日本大震災アーカイブ（ひなぎく）」を構築した。11)

大震災アーカイブの構築は，従来の図書館の枠を越え「あらゆる情報，記録を収集，永久保存して，提供する」というNDLの使命を達成するために，今後数年間で取り組むべき事項の多くが含まれていると考える。

### 収集の基本的な考え方

大震災に関連する情報のすべてを，NDLだけで収集し保存することは不可能であり，国全体で分担して収集し，分担して保存する形を目指すべきである。

#### 収集範囲

国全体で．様々な機関が保有している大震災前の記録，大震災後の事象・被害，状況の記録，今後の復旧・復興の記録等，過去から，現在，未来に亘って，可能な限り収集・保存する。

#### 収集方法

* NDLによる直接収集

NDLは，制度的に収集可能な記録のほか，他の機関が保存の対象としていない記録等については，積極的に受け入れる。

* 他機関による記録等の保存の推進

記録の保有機関，アーカイブ機関，NDLのいずれかで記録等を保存し，NDLは，その所在情報を把握できるようにする。

* メタデータの収集又は検索の機械的連携

NDLは，現在のNDLサーチと同様に，可能な限り，メタデータの収集，もしくは横断検索による機械的連携を行う。

* 働き掛け・協議

NDLは，国等の各機関で，国全体としての記録等の保存の必要性の認識を共有し，関係機関に協力を要請する。

### 東日本大震災アーカイブのシステムイメージ

NDLは，これまでの成果であるNDLサーチ，ウェブアーカイブシステム，デジタルデポジットシステム等の既存のシステムをベースに，機能拡張する形でシステムを構築した。

構築に当たっては，効率的，効果的に進められるように，また，利便性の高いサービスにするために，次世代技術の実用化実証実験に取り組み，成果の積極的な活用が図れるように調査研究を進めた。



図4 東日本大震災アーカイブのシステムイメージ

* 映像・観測データ等を受け取ることを想定して，処理能力，ストレージ容量を必要に応じて増強できる仕組みとして，分散処理サーバ(Hadoop)，分散ファイルシステム(GlusterFS)を導入した。
* 大震災アーカイブ自身が，災害で消失してしまわないように，ディザスタリカバリも考慮した。
* 画像・映像なども的確に検索ができるように，明確なメタデータが付与されていない情報にも可能な限りメタデータを自動付与する仕組みの導入を検討した。
* 本文も含めたテキストの全文インデキシング等を試行した。
* 次世代の図書館システムのモデルとなることを指向した。

### 「ひなぎく」での課題

構築・運用においては様々な課題が顕在化した。

#### 技術面での課題として，

* 大規模分散ファイルシステム，分散処理システムを適用したが，実際には，想定ほどコンテンツが集まっておらず，将来的な実用化レベルでの検証ができていない。
* 計測機等から得られた生データ，分析の中間データなどのファクトデータの永久保存も目指しているが，情報に関するメタデータまでしか収集できていないため，ビッグデータへの対応の検証ができていない。
* ビッグデータとしてのウェブアーカイブから大震災関連の情報を自動的に切出し，コンテンツ単位で精緻な検索をする仕組みが構築できていない。
* 各アーカイブにおいて，検索およびハーベスト用のWebAPIが未実装なものが依然として多い。
* 収集前データに永続的識別子がない。
* 各アーカイブのメタデータとの記述規則の差異，付与されたメタデータに記述要素の不足が多い。
* 不完全なメタデータへの自動付与機能はない。
* 利活用のための検索・閲覧機能として，本文テキスト，イメージ認識技術等，内容を関連付けた検索技術が未成熟である。

#### 人材面での課題としては，

* 専門分野に関する知見，文化資産の収集・保存・修復・公開の技能の不足。
* 文化資産を取り扱うための知識・技能の不足（プリザベーションエンジニア，コーディネータ，エンベデッドライブラリアン）。
* デジタル技術を活用したアーカイブ化のための知見の不足（アーキビスト，法規担当）。
* システム開発・運用管理の一般的な知識・技能の不足（システムライブラリアン，ITエンジニア）。

#### 協力体制の課題として

* 縦割り行政のため，立法府であるNDLが進めることに警戒感。
* 「放送アーカイブ」とする構想に対する報道圧力・事後検閲の可能性に対する警戒感。

#### 制度面・運用面での課題として

* 情報公開法では，各文書に対して最低保存期間が設定されるが，運用では行政文書管理簿に掲載されていない軽微な資料の保存期限は1年未満とされ，1年を経過すると廃棄される。
* 公文書管理法では，軽微な資料を含めて，歴史的公文書と指定されなければ公文書館に移管されない。
* 肖像権，プライバシー権，人格権等の権利処理がされなければ提供できないが，NDLは，提供を前提としない，いわゆるダークアーカイブとしての収集は行わないため，継続して維持できなくなったアーカイブのコンテンツを預かることもできない。
* 国有財産法，財政法の解釈では，アーカイブシステム機器等の運用維持が困難になったアーカイブ機関のために，Webサーバやストレージを無償で貸出すことができない。
* 現在のインターネット情報の制度的収集は国等のサイトに限定されているため，民間および個人のサイトは個別許諾により収集しなければならないが，個別に許諾手続きが必要なため，悉皆的な収集が困難。民間サイト，個人サイトは，ハーバード大学と連携し，インターネットアーカイブ社に収集を依頼している。

このように課題は多いが，協力体制，制度面・運用面の課題は，大震災に限らない共通の課題であり，今後，国全体のナショナルアーカイブ構想の議論の中で，具体的な解決策の検討が進められることを期待している。

# 今後10年で目指すところ（2015年～2024年）

NDLは，このような歩みで電子図書館事業を進展させてきたが，筆者は，今後10年を踏まえて，社会の要請，情報技術の実用化動向を想定した図書館サービスの姿と，その実現に向けて実施すべき事項を明確にし，速やかに対応していく必要があると考える。

東日本大震災アーカイブを踏まえて，分野を越えた知識インフラの実現形として，あらゆる記録を情報として集約し，相互に関連付けて知識化し，将来にわたって利用を保証するとともに，「社会・経済的な価値を創出」できる「新たな知識の創造と還流」の仕組みを現実的に構築していく時期と考える。

## 「知的財産政策ビジョン」での今後10年の目標

我が国では，「知的財産政策ビジョン」（2013年6月7日知的財産戦略本部）等において，文化資産のデジタルアーカイブ化を推進する政策も含めて，今後10年を見据えた知的財産関連の政府の取組としての目標を掲げた。12)政策としては，従来の事業モデルの延長線上での「改善」だけでなく，事業モデルそのものを創造・転換する「イノベーション」を創出すること，あらゆる分野の知的財産を対象としたビッグデータも含めたクラウド化，オープン化，更にユーザが共同でコンテンツを制作するクラウドソーシングの普及が見込まれている。そのための知財人財の育成・確保，コンテンツの権利処理の円滑化，電子書籍の普及促進，データの収集・蓄積・分析による多様な付加価値の創造の研究開発，日本の伝統や文化に根ざした魅力あるコンテンツ・製品などの発掘・創造等の取組も示された。

## 電子書籍ナショナルアーカイブ

有形の出版物を中心とした物としてのアーカイブ機能の延長に，無形の情報としての電子書籍のナショナルアーカイブがある。



図5 電子書籍分野のアーカイブの機能モデル

### 電子書籍分野のアーカイブの役割

筆者は，電子書籍分野のナショナルアーカイブでの必要な機能を5つに分解し，それぞれの出版界と図書館界の役割分担を例示した。

長尾前館長により，2008年に提案され，改めて2015年3月に示された電子出版物流通に関する「長尾モデル」13）として提案されたモデルと類似したプラットフォームである。

提供用の商用コンテンツは，電子書籍サイトが保持し提供する。NDLは，DRMによるアクセス制限が掛かっていないコンテンツを収集し，ダークアーカイブとして恒久的保存の役割を担い，電子書籍サイトがコンテンツを保持できない，もしくはサービスを継続できない場合，NDLの恒久保存用アーカイブを利用して電子書籍サイトから提供する。また，NDLが原資料保存のために画像イメージでデジタル化したコンテンツを，出版界が二次利用して電子書籍化・プリントオンデマンド（POD）化し，電子書籍サイトから提供することにより，いわゆる絶版本の電子書籍化を促進する，さらに出版物・著作権者の情報は，出版界・図書館界の共通のデータベースとして維持・更新し，相互に利活用することも想定したモデルである。

### 電子書籍分野のアーカイブの機能モデル

電子書籍に関するナショナルアーカイブは，コンテンツの生成機能，収集・一時保管機能，保存機能，権利情報・管理情報の収集・管理機能，配信・流通機能の5つの機能を想定する

　コンテンツの創出から収集・保存，配信・流通に至る全体の流れを一元的に集約するため，恒久保存のアーカイブと共に，ナショナルアーカイブ全体のメタデータを集約するデータベースにおいては，多様な主体が多様なデータを扱えるよう，柔軟かつ多層的なデータ構造を実現する。具体的には，著作物・著作者・出版者等の書誌情報，販売データ・販売者・所蔵機関等の所蔵情報，目次・索引・シソーラス等の情報探索情報等の多様なデータを一元的に管理可能な仕組みとすることを想定した。

## 国の文化資産のアーカイブ

以下は，筆者がNDL在職中に，知識インフラの実現形として国のデジタル文化財のアーカイブシステムの基盤として想定したものであり,現時点において，NDLを含めて関係機関のコンセンサスが得られたものではない。

### アーカイブに関連した国の活動の方向性

国は，「知的財産政策ビジョン」（2013年6月7日知的財産戦略本部決定）に基づく国の施策の中で，図書館に直接関連する計画として，電子書籍化と利活用の促進に関する構想（電子書籍と出版文化の振興に関する議員連盟），デジタル文化資産の保存・活用の基盤の整備に関する構想（デジタル文化資産推進議員連盟）14)，2020年までの6年間の文化芸術分野での振興に関する基本的な方針（文化庁）15)，学術情報の公開と共有の拡充に関する計画（文部科学省）16)，大規模災害の記録と記憶の保存などのアーカイブの構築等を検討している。個々の施策の目的は異なっていても，対象とする文化的資産は相互に関連するものであり，知識インフラを目指した仕組みとして，国全体で文化関係資料の価値を高め，新たな文化や情報を生み出す社会基盤として，「恒久保存・継承・公開・活用」が可能なナショナルアーカイブとなることが望まれる。「資料・情報を文化資産として収集・保存する」ということは，従来からの出版物に相当する情報の範囲ではなく，美術館，博物館，文書館等が保有する無形・有形の文化財をデジタル化した情報を含め，インターネット上で流通している著作物全てを文化資産としてアーカイブすることであると筆者は考える。

文化財のナショナルアーカイブの構築と運用に関しては，様々な機関において，必要性の認識の共有やアーカイブ立国宣言の提案17)，制度的な課題解決のために，アーカイブ基本法の法制化，推進体制作りなどが議論されている。また，具体的なサービスシステムの仕組みとして，技術的課題解決のための研究開発も進んでいる。18)

図6 文化財を含めたナショナルアーカイブの機能イメージ

恒久的保存基盤の部分は，文献も含めて，すべての文化財，災害情報が相互に関連付けられて蓄積されている。恒久的保存基盤を活用して新たな知識創造活動が行われ，知識創造活動の成果を，様々な目的に応じて利活用されていくことを想定する。

### 恒久的保存基盤

恒久的保存基盤は，恒久保存と利活用のための共通プラットフォームとして，１つの機関にすべてを集約するのではなく，各分野のアーカイブを集約する拠点が中核となって分散アーカイブを構築し，各機関の情報を相互に持ち合って，障害，災害に備えるとともに，情報のマイグレーションを行うことにより，将来にわたって利用を保証する仕組みである。その分散アーカイブを集合して，あたかも１つのアーカイブとして見えるようにして網羅性，完全性を確保し，個々の情報同士を意味的に関連付けて，情報間のネットワークを構築することを想定する。

このようにネットワーク化された情報に対して，分野を越えて網羅性を保証した検索GUIとして，本文の全文検索，あいまい検索，シソーラス検索などを組み合わせた検索で情報を取り出すだけでなく，取り出された情報から芋づる式に関連する情報を取り出せるようにすることである。

### 知識創造基盤

知識創造基盤は，それぞれの分野の専門家のみならず，広く国民も含めて，様々な分野の網羅的な知識を活用して，新たな著作物を創造する場である。創造活動を支援する基盤として，情報全体の基本情報としてのメタデータを付与する活動，記事，章節項，文節等の単位で組織化・構造化する活動，情報間を意味的に関連付けるための基本情報として，用語辞書，典拠，シソーラス辞書等を作成する活動を想定する。

新たな知識を創造する活動は，まず，恒久的保存基盤に格納された網羅的な情報を活用して新たな知識を創作する活動がある。関連付けて利用できる情報の幅が広がるため，より高度な創造性が期待できる。

また，歴史的な文化財や現代文化を映像化，画像化，テキスト化する活動，構造化された情報に解題情報等を付与する活動，情報間を意味的に関連付ける活動，テーマを設定してデジタルギャラリを構築する活動等が含まれる。

ここで生成された情報は，新たな知識として恒久的保存基盤に蓄積されていくことを想定する。

### 知識利活用基盤

「見るだけのアーカイブ」から「使い，創り，繋がり，伝えるアーカイブ」として，広く国民による新たな知識の創造，新産業の創出，地域活性化，防災・減災，教育活用，教養・娯楽，観光，国際文化交流等，様々な利用者毎の目的に応じて，恒久的保存基盤に格納された一次情報，知識創造基盤で創出された二次的情報を有機的に組み合わせて，利用できるようにする基盤である。網羅的な情報から，利用目的に応じてあらかじめ適切に絞り込み，利用者の属性，スキル，利用場所に応じて，様々なGUIを用意して，利用者が必要とする情報，参考となる関連する情報を容易に得られるようにするもので，レファレンスサービスによる情報探索支援，オンラインレファレンスなども含まれる。

### 運用基盤

データ保存課題，人材不足問題，資金不足問題，権利処理コスト問題など，デジタルアーカイブの活性化を阻害する課題を制度的に克服する方策(人材育成策・予算措置・権利処理に関する法改正など)を導き出し，解決を図っていく中核的な役割を担う推進母体(司令塔)としての組織・体制も必要である。この基盤では，ナショナルアーカイブ全体の戦略企画，デジタル情報の保存や利活用のための調査研究，研究開発，デジタル化支援，アーカイブに所蔵された資料に関する知識，読解力とIT技術の知識等も備えた高度な専門的人材の育成，孤児著作物の権利処理や，絶版作品のデジタルアーカイブ化における所有権，肖像権問題も含めた権利情報DBの構築を促進する等の役割を持つことを想定する。

# 文化財でのLOD化の目指すところ

従来は，NDLは，同分野の情報の網羅性を意識して，各機関の情報に付けられたメタデータによる網羅的な情報探索，情報検索を効率的に実践できることに注力してきた。今後の国の文化財全体のナショナルアーカイブでは，異分野の様々な情報も含めて情報を関連付けることにより，思わぬところで自分の興味を喚起する情報に出会え，新たな発想の知識の創出が可能になる環境の実現が重要であると筆者は考える。

Linked Data化は，様々な情報を関連付けて知識として再利用するための具体的な実現方法を示したものであり，LOD化が進むことにより，あらゆる情報が関連付けられて利活用が可能になる。筆者は，LOD化の動きが，分野を越えて進展することにより，文化財全体のアーカイブがナショナルアーカイブとして実現されていくものと考える。

## クラウドコンピューティングの世界でのサービスの連携（2007年想定）

図7は，筆者が2007年に想定した，クラウドコンピューティングの普及を見越した将来像である。

　様々な機関から提供されているWebサービスが，連携することにより，より効果的なサービスとなっていく。サービスが連携できる情報提供環境として，当時の大きなパラダイムシフトの一つとしてのクラウドコンピューティングは，今や個人の情報管理のインフラとして普及している。

クラウドコンピューティングは，ネットワークコンピューティング，グリッドコンピューティング，SaaS（Software as a Service），Webサービス等の技術を融合させた基盤である。

個別の組織が，単独でコンピュータ資源（ハードウェア，ソフトウェア）や情報を抱えてサービスを提供し，利用者が，個別のサービスを利用するような形ではない。クラウドコンピューティングの普及により，様々な機関の情報の共有だけでなく，個別のサービスを組み合わせて利用することが容易に実現できる環境が整った。

各情報保有機関のサービスも，このパラダイムシフトに対応していかなければ，孤立したサービスになっていく。クラウドコンピューティングが普及期に入り，また，デジタルトランスフォーメーションの考え方が一般化してきた社会において，各機関が保有するデジタル情報の管理の仕方，顧客サービスのあり方を再考する必要がある。今後ますます，情報と情報が情報群，サービスとサービスがサービス群，人と人がユーザ群として有機的に関連付けられていく。

いわゆるWeb2.0の時代では，パーソナライズ，人と人の関係を活用して情報を選択してきた。また，サービスも先進的な機関が提供するWebサービスを個別に利用してきた。Web3.0の時代以降は，それに加えて，情報と情報，サービスとサービスの関係を組み合わせて，ユーザ群と情報群が多対多の関係で，より的確な情報の活用が可能になってくる。

従来のように個別のデータベースを検索し，コンテンツを閲覧するような形から，クラウドコンピューティングの世界での複数のリソースから得た情報に，分析情報などの付加価値を付けて，利用者へ助言するような内容を知識として提供するサービスが実現されつつある。例えば，近年のカーナビの情報は，単に目的に行くルートを情報として提供しているのではなく，様々な手段により動的に変化する状況を集約し，その結果を踏まえてのリルート情報や，ルート上の様々な位置でのスポット情報などを合わせて提供している。ナショナルアーカイブの枠組みの中で，Linked Open Data化が進むと，更に利便性の高いサービスとして発展する。

このようなインターネットの世界での情報提供の状況の変革が，NDLで進めているウェブアーカイブ，情報探索サービスのあり方にも大きく影響してくる。従来のウェブアーカイブは各機関のウェブ情報を収集保存してきているが，このようなサービスの形をそのまま保存することは困難であり，また，関連付けられたメタデータ等，どのようにしてサービスの中にある情報を収集していくのかを考える必要がある。

クラウドコンピューティングの普及により，想定したサービスが現実になってきたが，今後，NDLは，情報を収集し将来にわたって保存する使命をどのように果たしていくか，考えていく時代が来ているように思える。

図7 クラウドの世界でのサービスの連携

## 文化財のアーカイブ構築のために何に留意するべきか

最後に，ナショナルアーカイブの構築を目指す中で，司令塔的な役割を果たす組織および文化財を保有する各機関は何に留意すべきか，筆者が想定する事項を列挙する。

* 今後10年を見据えて，個別の組織の論理ではなく，利用者視点でサービスを構築する。
* アーカイブの網羅性の確保のために，業種業態毎のアグリゲータ機関を設置してメタデータを集約する。更に，Linked Data化するために，情報を一意にする識別子を付与する。
* メタデータはもとより，コンテンツをOpen Data化する。
* 分野を越えたアーカイブ構築のために，メタデータの記述要素，記述規則および情報交換の通信規約（API）は，共通仕様を実装する。
* メタデータおよび情報内容の用語表現の違いを吸収する辞書を構築する。
* アーカイブの構築・運用に必要な人材を育成・確保する。

これを実現するするために，国の役割は次の点にある。

* 知的財産政策ビジョン，知財計画に掲げた施策を確実に実施する。
* 文化資源のアーカイブの利活用に必要な制度を整備する。
* アーカイブの構築・運用に必要なスキル・知識を習得するための体系的なカリキュラムと教材を作成するとともに，教育体制を確立する。

# まとめ

NDLは，今までは「電子図書館構想」（1998年）の実現を目指し，具体的な実施の柱立てとして「電子図書館中期計画」（2004年）を策定し，デジタルアーカイブ構築，情報に関する情報の充実，デジタルアーカイブポータルの構築を行ってきた。

筆者は，電子図書館としての知識・情報・文化の新しい基盤を含めて，図書館の枠を越えた基盤作りが，「知識インフラ」であり，知識インフラの実現形の１つが「ひなぎく」，知識インフラ構築の全体の実現形が，「文化資源のナショナルアーカイブ」と位置付けている。1990年代から目指す方向性は明確であり，構築に当たっては，セマンティックウェブ技術の適用を常に意識して進めてきており，具体的な実現手段として浸透しつつあるLinked Data化は，ナショナルアーカイブの構築においても，実現手段として中核的な技術の一つとなると考える。

# おわりに

個々の文化財の保有機関でのデジタル化が進む状況において，文化的資産をあらゆる人々が将来にわたり享受，活用できるようにし，人々の創造的な活用に貢献するためには，個々に情報資源の「見える化」をするだけでなく，全体のそれぞれの施策が同一の方向性を持って，相互に資源を補完し合って，社会全体でより効率的な利活用の保証に取り組む必要があり，ナショナルアーカイブは重要な基盤の役割を果たすことになる。ナショナルアーカイブの構築の先には，世界レベルでの「インターナショナルアーカイブ」があり，その構築へと発展することを目指していくべきと考えている。

今後10年のデジタル情報化の進展を見据えると，現行の図書館業務は根幹から変革が求められると思われる。知識創造のための情報の多くがインターネットにより，入手できる状況において，知識創造を支援する図書館の役割と必要な機能を再検討し，速やかに対応していかなければ，情報提供機関としての図書館は存立意義を失う可能性がある。

～～～

■参考文献

1. 田屋裕之.「パイロット電子図書館システム」の現状と課題.出版ニュース.1996, 96 1/上・中,p.14-17
2. 中山正樹.国立国会図書館電子図書館中期計画2004の実施に向けて.情報の科学と技術. 2004,Vol.54,No.9,p.453-460.
3. 久古聡美・吉田曉・中山正樹. Web2.0世代のデジタルアーカイブポータルの提供を目指して.情報管理. 2006,Vol.49,No.6,p.313-323.
4. 国立国会図書館. “機能概要≪国立国会図書館サーチについて”. 国立国会図書館サーチ, 　<http://iss.ndl.go.jp/information/function/>,　(accessed 2016-1-15).
5. 国立国会図書館. “日本全国書誌の在り方に関する検討会議について”. 国立国会図書館インターネット資料収集事業(WARP). [http://warp.da.ndl.go.jp/info:ndljp/pid/8703998/www.ndl.go.jp/jp/library/data/jnbconf\_20100303.html](http://warp.da.ndl.go.jp/info%3Andljp/pid/8703998/www.ndl.go.jp/jp/library/data/jnbconf_20100303.html), (accessed 2016-1-15).
6. 内閣府科学技術政策・イノベーション担当. “第4期科学技術基本計画（平成23年8月19日　閣議決定）”. 内閣府. <http://www8.cao.go.jp/cstp/kihonkeikaku/4honbun.pdf>,(accessed 2016-1-15).
7. 国立国会図書館. “第三期科学技術情報整備基本計画”. 国立国会図書館インターネット資料収集事業(WARP). <http://dl.ndl.go.jp/view/download/digidepo_8262067_po_basic_plan03.pdf?contentNo=1>, (accessed 2016-1-11).
8. 長尾真.未来の図書館を作るとは.長尾真, 2012, 27p
9. 中山正樹. 国立国会図書館におけるデジタルアーカイブ構築－知の共有化を目指して－. 情報管理. 2012,Vol.54,No.11,p.715-724.
10. 内閣官房東日本大震災復興構想会議. 復興への提言～悲惨のなかの希望～. 2011-06-25, 第12回東日本大震災復興構想会議. 57p. <http://www.cas.go.jp/jp/fukkou/pdf/kousou12/teigen.pdf>, (accessed 2016-1-15).
11. 国立国会図書館. “東日本大震災アーカイブ構築プロジェクト”. 国立国会図書館.<http://kn.ndl.go.jp/static/about/project>,(accessed 2016-1-15).
12. 内閣官房知的財産戦略本部. “知的財産政策ビジョン”.首相官邸政策会議.<https://www.kantei.go.jp/jp/singi/titeki2/kettei/vision2013.pdf>,(accessed 2016-1-15)
13. 長尾真. “「電子図書館と電子出版の今後」JEPAセミナー”. 電子出版協会.<http://www.slideshare.net/JEPAslide/20150303-drnagao-45401397>, (accessed 2016-1-15)
14. デジタル文化財創出機構.“「国立デジタル文化資産振興センター(仮称)」設立構想提言”.文化庁文化関係資料のアーカイブに関する有識者会議.<http://www.bunka.go.jp/seisaku/bunkashingikai/kondankaito/bunka_archive/01/pdf/shiryo_5-2.pdf>,(accessed 2016-1-15)
15. 文化庁.「文化芸術の振興に関する基本的な方針（第4次基本方針）」.文化庁文化審議会.<http://www.bunka.go.jp/koho_hodo_oshirase/hodohappyo/pdf/2015052201.pdf>,(accessed 2016-1-15)
16. 日本学術会議情報学委員会. 我が国の学術情報基盤の在り方について －SINET の持続的整備に向けて－.日本学術会議.<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-t192-2.pdf>,(accessed 2016-1-15).
17. 「アーカイブ立国宣言」編集委員会 編. アーカイブ立国宣言. ポット出版, 2014,272p.
18. 中山正樹.国立国会図書館のサービスシステムの歩みと新たな方向性の模索－電子図書館事業20年の歩みー. 国立国会図書館月報. 2015,Vol.648,p.18-24.

**■図表，写真のキャプション**

図1 電子図書館に関連する構想とサービスの変遷

図2 NDL電子図書館サービスの全体像

図3 知識情報基盤の構築モデル

図4 東日本大震災アーカイブのシステムイメージ

図5 電子書籍分野のアーカイブの機能モデル

図6 文化財を含めたナショナルアーカイブの機能イメージ

図7 クラウドの世界でのサービスの連携

≪討議報告≫

LOD化によるデジタル文化財の利活用を目指して

**（発表：中山正樹）**

**和中　幹雄（大阪学院大学）**

フォーラム当日の中山氏による発表題目は，「LOD化によるデジタル文化財の利活用を目指して」であった。本誌掲載の論文ではそれは副題目となり，発表内容をより端的に示す題目を用いている。この発表の趣旨は，「国立国会図書館（NDL）の電子図書館事業の20年間を情報技術の観点から振り返り，今後のナショナルアーカイブに向けた展望と課題を概観したもの」ということができる。

フォーラム当日のプレゼン資料には，「NDL電子図書館サービスの全体像」「長期保存システムの概念モデル」「国立国会図書館サーチのコンセプト」「文化財を含めたナショナルアーカイブの機能イメージ」「ナショナルアーカイブにおけるLOD化」などなど，数多くのモデル図やイメージ図で溢れ，個々のスライド一枚を取ってみても，優に一つの発表内容となるものが多く見られた。そのため，予定された時間で終了できず，発表の趣旨の理解も不十分だったように感じられた。しかしこの点については，本誌掲載の論文を読めば氷解するであろう。

彼の発表が自分の職歴の紹介から始まるのも故なしとしない。NDLの電子図書館事業は，1994年に情報処理振興事業協会（IPA）との協力で実施したパイロット電子図書館プロジェクトによって始まる。彼はこの事業にIPA側の担当者として参画する。その後，関西館が開館し電子図書館サービスが開始される2002年にNDLに入館し，2004年に，デジタル・アーカイブ構築とデジタル・アーカイブポータルの考え方を明示した「電子図書館中期計画2004」策定の中心メンバーとなり，その後，PORTAからNDLサーチの開発に携わり，2015年3月末に，電子図書館事業や情報システム全般を統括する立場にある者として定年退職を迎える。このように，NDL電子図書館事業の20年と中山氏の図書館システムに関わるキャリアは全く重なるからである。

1994年から始まる現在までの20年間のシステム構築に当たっては，セマンティックウェブ技術の適用を常に意識して進めてきたと彼は述べている。つまり，LODも含めたセマンティックウェブの要素技術をどのような考え方で個別に適用してきたかを振り返ることにより，今後の方向性を示したのが本発表であったと言える。

情報技術の観点からの発表ではあるが，特に今後のナショナルアーカイブ構築に向けた課題整理において，内閣に設置された知的財産戦略本部による「知的財産政策ビジョン」などの国の政策を背景とした提言が多く見られる。彼は，発表を終わるにあたって，次のように述べている。「LOD化されたナショナルアーカイブ構築に当たっては，それらの意義・理念の確認，認識の共有，草の根的な試作など，やれるところからやるという段階から，大所高所の方針・計画に基づいた，具体的な実用サービス構築を行うという次の段階に移行する必要がある。そのためには，司令塔が必要である」と。

我が国にどのような司令塔が可能であろうか。

彼が示したモデル図やイメージ図は，これからの図書館を考える上で示唆に富む情報が数多く含まれている宝の山と言える。

司会者の仕切りが悪かったこともあり，討議時間はほとんど取れず，質疑もNDLサーチの個別問題に終始したことは残念であった。当日のプレゼン資料にあるさまざまなモデル図やイメージ図を基にまとめられた本誌掲載の論文をお読みいただき，今後，活発な議論がなされることを期待したい。

NDLサーチをめぐる質疑は，次のようなものであった。

Q.ディスカバリーサービスのNDLサーチで，NDL-OPACの代替ができるのか？

A．NDLサーチのGUIは，ナショナルアーカイブの情報発信基盤（知識利活用基盤）で全ての情報を検索対象とする。来館利用者向けのNDL-OPACは，現在はALEPHで提供しているが，今後は，NDLサーチで，あらかじめ来館者向けに当館所蔵の図書・雑誌，デジタル化コンテンツに絞り込んだGUIを用意すれば，現在のNDL-OPACと同等のサービスが提供できる。現に，児童書OPACは，NDLサーチで，GUIを変えて児童書に絞り込んで提供している。

Q．新聞総合目録がNDLサーチに取り込まれて使いづらくなった。

A．総合目録と言われていた，ゆにかねっと，児童書総合目録，新聞総合目録は，全て統合しNDLサーチでGUIを提供するとしてサービスを構築したが，新聞総合目録は，提供する情報のレベル，検索手法が他と大きく違い，相当無理があったのは確かである。

Q．NDLサーチの検索結果のデフォルトの「適合度順」はよくわからない。

A．書誌情報における検索キーワードの個数，出現頻度，データベースごとの優先度を数値化して，数値の高いものの順に表示している。ロジックの完成度はまだ高くない。