

6

電子書籍等のデジタルコンテンツの 長期保存と、将来にわたっての利用の保証



— 文化的資産の保存に向けた関係機関との連携協力 —

中山正樹 (国立国会図書館電子情報部)



目指すところ

国立国会図書館 (NDL) は、納本図書館として、印刷出版物だけでなく、電子情報も含めて、資料・情報として収集保存し、将来にわたって利用を保証する責務を持っている。しかしながらそれらの資料・情報のすべてを収集することは不可能である。

NDL は、NDL および他機関が分散して収集・保存している資料・情報を、意味的に関連付け、一元的に検索・ナビゲートできるようにする。また、情報を文化的資産として、将来にわたって利活用できるようにすることを目指している。

使命・経緯

■ 経緯

NDL は、1994 年頃から電子図書館構築に向けた活動を行ってきた。1995 年にはパイロット電子図書館プロジェクトでの実証実験を実施し、2002 年、電子図書館サービスを中心的な機能とする関西館が開館し、電子図書館事業を開始した。

2004 年には、「電子図書館中期計画 2004」を策定し、デジタルアーカイブの構築、レファレンス情報等に到達するための仕組みの充実、さらに、所蔵場所によらず、一元的にアクセスできるポータル構築に取り組んできた。

この 8 年間、電子情報に対する取り組みは、国の高度情報通信ネットワーク社会推進本部 (IT 戦略本部) が掲げた e-Japan 戦略等で NDL の役割が示され、それも実現する形でデジタルアーカイブ構

築事業を進めてきた。その実現形が、2012 年 1 月にリニューアルオープンした国立国会図書館サービスシステムの基本的な形であり、印刷刊行物等の物理的な資料の収集整理と閲覧提供を管理する業務基盤システム、電子情報を統合的に収集保存するデジタルアーカイブシステム (DA システム)、NDL および他の機関が保有する印刷刊行物、電子情報等を一元的に検索・ナビゲートする国立国会図書館サーチ (NDL サーチ) 等で構成されている¹⁾。

■ NDL の使命と目標

NDL では、2012 年 7 月に、今後 5 年程度の使命として「私たちの使命・目標 2012-2016」を策定し、文化的資産としての収集・保存、迅速かつ確かなアクセス環境・手段の整備、国内外の関係機関と連携して、知識・文化の基盤を一層豊かにすること、さらに、東日本大震災に関しては、出版物に限らずすべての記録を後世に残すことを、柱になる目標として掲げた。

関係機関との連携の観点

さまざまな資料・情報を文化的資産として保存し、利活用を促進するためには、収集、組織化、保存、提供のシステム構築・運用、業務実施のあらゆる局面での関係機関との連携協力が不可欠である。連携協力の実施に当たっては、姿勢として合意形成にとどまらず、具体的なアクションとして、実施していくことが重要である。

収集の観点

■「インターネット資料」の許諾に基づく収集

Web サイトでは日々刻々と新しい情報が発信され、同時に消されていく。時に、機関・組織の改廃・合併により、サイトそのものが消失する。消されていく情報には、後世に残すべき文化的資産も多く含まれ、印刷出版物の形態をとっていないいわゆるポーン・デジタルのものもある。NDL では、2002 年から、各機関の協力のもと、個別の許諾に基づいて、Web サイトを収集し、時間軸で再現できるように保存している。

■公的機関の「インターネット資料」の制度的収集

2009 年 7 月に国立国会図書館法と著作権法が改正され、国、地方公共団体、国立大学等の公的機関が発信するインターネット資料について、個別著作権者の許諾なく収集できるようになった。この法改正に基づき、NDL は、2010 年 4 月から、公的機関の Web サイトの網羅的な収集を開始した。

公的機関の協力により、順調に収集が行われている。なお、国公立大学の機関リポジトリ等は、早期に消失されることなく保存・提供が保証されているとみなされるので、制度的収集は行わず、NDL サーチで、資料の所在場所にナビゲートして、閲覧利用を保証している。

■民間の「オンライン資料」の制度的収集

電子書籍、電子雑誌など従来の図書、雑誌に相当しネットワーク上を流通する電子情報を「オンライ

ン資料」と定義している。2012 年 6 月に、国立国会図書館法と著作権法が改正され、オンライン資料の収集が許諾なしにできるようになった。この法律は、2013 年 1 月に施行される。しかしながら、有償で提供されているオンライン資料に関しては、条件の整備等、時間を要するため、当分の間、無償で提供され、閲覧制限機能が実装されていないものを収集することとし、2013 年 7 月からの収集に向けて、運用の検討を行っている。これにより、無償の電子書籍、電子雑誌は、ネットワーク上から消えていっても、将来にわたって利用が可能になる。

保存の観点

■所蔵資料のデジタル化

NDL が収集保存している印刷出版物は、経年劣化が進むとともに、閲覧・複写提供により劣化が加速される。原本保存のために、2009 年に著作権法が改正され、著作権者の許諾を得ないで所蔵資料をデジタル化することが認められた。2009 年度補正予算で、NDL 所蔵資料のデジタル化経費として約 127 億円が計上され、大規模にデジタル化を実施した。このデジタル化で、NDL 所蔵資料の 1/4 程度はデジタル化できたが、残りも引き続きデジタル化を進めていく必要がある。しかしながら、国の予算が厳しい状況において、今後も継続的に大量のデジタル化を行う用途は立っていない。

■文化的資産の保存

NDL が収集したパッケージ系電子出版物、インターネット資料、もしくはデジタル化した電子情報も、国の知識・文化の基盤となる資料・情報であり、データを失うことはあってはならない。

電子書庫としてのストレージは、東日本大震災アーカイブのためのものも含めると、現在においても 2PB（ペタバイト）の容量となる。今後、さらに増加する電子情報の利用を保証するためには、大きく 2 つの観点がある。1 つは、物理的に読めなくなないように保存（物理保存）すること、もう 1 つは、



ファイルの内容が読めなくならないように保存（論理保存）することである。

物理保存について、現時点において半永久的に保存できる記録媒体は実用化されていないので、膨大なデータを物理的に読めるようにしていく仕組みの確立が課題である。現在、GlusterFS というシステムの適用を試行している。これは、寿命が5～10年で数TB（テラバイト）程度の容量の磁気ディスクを備えたPCを並列に配置したスケーラブルな大容量分散ファイルシステムで、順次容量の大きな磁気ディスクに置き換えることで、少しずつ媒体変換を進め、かつ、必要な容量を確保できる仕組みである。また、このシステムにより、大規模災害に備えたディザスタ・リカバリ対策として、複数の分散したセンタで同期する仕組みの実装も想定している。近い将来には、クラウドサービスを活用し、複数の民間クラウドサービスを組み合わせ、相互に同期させることで災害時も含めたデータの保存を図ることも想定する。

論理保存については、ITの発展とともに、さまざまな国際標準、業界標準のフォーマット仕様を適用した電子情報が存在するが、将来にわたって読めるようにすることは大きな課題である。

この問題は、保存の使命を持つNDLだけでは解決が不可能である。新しい媒体、フォーマット仕様を開発してきた技術者・研究者の方々、国際標準・業界標準を策定してきた機関、その仕様を適用したアプリケーションやコンテンツを開発、販売してきた企業など、さまざまな関係者が、現在の利用者のみならず、文化的資産として後世においても利用されることを想定して、仕様の共通化、マイグレーション等に協力して取り組んでいただきたい。



組織化の観点

■ 組織化の意義

組織化とは、利用者が迅速、的確かつ容易に検索できるように、メタデータ（書誌データを含む）を付与して整理することである。

1つの著作物が、単行本として出版され、のちに文庫本となり、またさまざまな形態の電子書籍として、派生して流通しているが、体系的に整理されたメタデータが付与されていると、利用者属性（知識レベル、嗜好等）、利用環境（PC、モバイル、アクセス場所等）を考慮して、コンテンツを的確に選択できるようにすることが容易になる。

外形的な情報によるメタデータのみならず、セマンティックWeb技術等を駆使して本文テキストからの組織化も行えると、より利便性の高い検索サービスが実現できる。

■ 組織化の連携協力

図書・雑誌の出版者、博物館、文書館、図書館等のいわゆるMLA機関、その他著作物を提供するすべての機関が、語彙の違いを吸収できる共通のメ

タデータ記述規則を適用し、意味的に関連付けられることが重要である。関係機関で協力してメタデータの相互交換の仕組みを構築する必要がある。

また、爆発的に増加する電子情報には、従来の印刷刊行物のように人海戦術的な精緻なメタデータの付与は困難である。自動的にメタデータを付与する技術、本文情報も含めて組織化する技術等、大量のデータを構造化・マイニングするための研究開発とその成果の実用化が期待されている。

電子書籍出版社等との連携

電子書籍は、印刷出版物の延長にあるものであり、文化的資産の1つの形態である。

現在、電子書籍出版は、ビジネスとして立ち上がろうとしている。NDLは、電子書籍によって読者人口が増えて、出版全体の市場が拡大し、出版ビジネスが加速されるように支援するとともに、電子書籍の利用を将来にわたって保証することが役割と考える。そのためにも、民間の市場経済活動を阻害することなく、市場拡大のために、出版界と下記のようなさまざまな連携協力を検討している。

- ・NDL デジタル化コンテンツの二次利用の促進
- ・電子書籍サイト等、商用サイトへの案内の強化
- ・電子書籍ビジネスのプラットフォーム整備への協力
- ・電子書籍フォーマットの共通化
- ・電子書籍に対する永続的識別子の付与
- ・公共図書館での利用環境の共通化
- ・著作権管理センターの構築・運用の協力

新たな取り組み

■ 東日本大震災アーカイブ

東日本大震災アーカイブは、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・次世代への教訓等を記録として網羅的に収集し、後世に残すものである。大震災の記録は、従来からの収集対象である印刷刊行物にとどまらず、ピラ類、写真、動画、音声はもとより観測記録等、多種多様である。また、記録を保有している機関もさまざまであり、早期に収集保存に着手しなければ、散逸の恐れがある。記録を保有もしくは集約している関係府省、博物館・美術館、図書館、文書館、企業および先行して震災アーカイブを構築・運営している組織と協力して、網羅的な収集と保存を進めたい。

大震災アーカイブポータルに関しては、既存のデジタルアーカイブシステムをベースに、分散アーカイブを構築し、また、統合的に利用できるポータ

ルを構築している。また、見せ方も記録の日時、場所も意識した閲覧機能の実装を目指している。

■ 知の共有化に向けた連携

国の第4期科学技術基本計画で示された「知識インフラ」は、知の共有化を目指す分野を問わないモデルであり、2012年1月にリニューアルしたサービス・システムの延長線上にあるものである。

また、東日本大震災アーカイブは、コンテンツ、システムとともに、分野を特定した「知識インフラ」の実現形であり、既存のサービスをベースに、必要な機能を実装する。このアプローチは、国の施策としての「ビッグデータの利活用」「知の共有化」に繋がる。

政府の施策との連携

知的財産推進計画2012(2012年5月知的財産戦略本部)²⁾の「戦略2：日本を元気にするコンテンツ総合戦略」では、「電子書籍の本格的な市場形成」および「コンテンツのアーカイブ化とその活用促進」において、関係府省とNDLが協力して取り組むべき事項の内容とスケジュールが示されている。

「電子行政オープンデータ戦略」(2012年7月IT戦略本部)³⁾では、オープンガバナンスの方向性として、国民共有財産である公共データを積極的に公開すること、機械判読可能な形式で公開し、営利目的、非営利目的を問わず活用を促進することが示されている。

また、内閣府総合科学技術会議の科学技術イノベーション政策推進専門調査会⁴⁾において、第4期科学技術基本計画の重点化課題「新たな産業基盤の創出」の重点的取り組みとして「大規模情報(ビッグデータ)の利活用の基盤技術の開発・標準化・普及促進」が明確化され、ビッグデータの収集・蓄積・分析等の研究開発および国際標準化を進めるとされている。NDLのデジタルアーカイブ構築、東日本大震災デジタルアーカイブポータル構築は、まさにビッグデータを扱うシステムであり、これらの

施策での技術開発，実証実験の成果を活用していきたい。

関係機関を繋ぐ役割を果たす

NDLは、「私たちの使命・目標 2012-2016」を策定し，具体的な実施計画の策定作業を進めている。

NDLは，関係機関との連携により，国としての資料・情報を，ビッグデータとして利活用できることを目指していく。

今後，さまざまな業種・業態で情報を発信者している機関同士，それらの情報を発信している機関と情報の利用者同士，また，膨大な情報を高度に処理・活用するための研究開発・技術開発を行っている組織同士を繋ぐ役割を果たしたいと考えている。

繋ぐに当たっては，関係機関間の利害調整ではなく，未来志向でより創造性を持って，資料・情報の

権利保持者の権利を尊重し，将来的な利活用の拡大を目指して，共存共栄で協力・分担して進めることが大切である。

参考文献

- 1) 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，情報管理，Vol.54, No.11, pp.715-724 (2012)。
- 2) 知的財産戦略本部：知的財産推進計画 2012，知的財産戦略本部会合議事次第，<http://www.kantei.go.jp/jp/singi/titeki/120529/gijisidai.html> (2012)。
- 3) 高度情報通信ネットワーク社会推進戦略本部：電子行政オープンデータ戦略，http://www.kantei.go.jp/jp/singi/it2/pdf/120704_siryoku2.pdf (2012)。
- 4) 総合科学技術会議：第 5 回科学技術イノベーション政策推進専門調査会議事次第，<http://www8.cao.go.jp/cstp/tyousakai/innovation/5kai/index.html> (2012)。

(2012 年 8 月 31 日受付)

● 中山正樹 m-nakaya@ndl.go.jp

2002 年国立国会図書館入館。デジタルアーカイブおよびポータル構築に従事。現在，館の電子情報関連事業，情報システム関連業務を統括。

