

資料と公共性 : 2021年度研究成果年次報告書

岡崎, 敦

九州大学大学院人文科学研究院 : 教授

池上, 大祐

琉球大学国際地域創造学部 : 准教授

今井, 宏昌

九州大学大学院人文科学研究院 : 専任講師

多川, 孝央

九州大学情報基盤研究開発センター : 准教授

他

<https://doi.org/10.15017/4772780>

出版情報 : 2020-03-07. 九州大学大学院人文科学研究院

バージョン :

権利関係 :

デジタル環境下における情報専門職の専門性の検討 —デジタルアーカイブを軸として—

大沼 太兵衛

0. はじめに

ウェブに代表される近年の情報技術の進展及び各種オンライン情報資源の充実は、図書館・アーカイブズ・博物館・美術館等の記憶機関（memory institutions）の機能を大きく変容させている。本稿は、このような状況の下、現代の情報専門職（司書・アーキビスト・学芸員等）に求められる専門性について考察する。具体的には、デジタルコンテンツを扱う新たなサービスの典型例であるデジタルアーカイブの構築及び運用にかかる業務、並びにその各工程で必要となる技術・知識を俎上に載せる。もとより、情報専門職の専門性についての議論は長い蓄積を持つが¹⁾、本稿では基本的にそれらの議論には立ち入らない。本試論は、デジタルアーカイブという近年伸長著しい実践領域をめぐる課題に議論を限定することで、この問題に一つの側面から貢献を行うものである²⁾。

1. デジタルアーカイブとは

1.1. 定義

デジタルアーカイブは未だ歴史の浅い概念であることもあり³⁾、その定義・範囲に関して統一的な見解は未だ存在しない。定義の例としては、「有形・無形の文化財をデジタル情報として記録し、劣化なく永久保存するとともに、ネットワークなどを用いて提供すること。最初からデジタル情報として生産された文化財も対象となる」⁴⁾、「様々なデジタル情

¹⁾ 特に司書の専門性をめぐる近年の議論動向については、次の文献に詳しい。大城善盛「1980年代以降の司書職（司書職制度）論の歴史」『司書職制度の再構築：日本の図書館職に求められる専門性』日本評論社、2019年、15-87頁。

²⁾ 本稿の意見に係る部分は筆者の個人的見解であり、所属組織を代表するものではない。

³⁾ 一般的には、1990年代半ばに月尾嘉男・東京大学教授（当時）が初めて用いた造語であると説明されることが多い（総務省「デジタルアーカイブの構築・連携のためのガイドライン」2012年、6頁（https://www.soumu.go.jp/main_content/000153595.pdf）、等）。（以下、ウェブ情報の最終アクセス日は全て2022年2月12日）

⁴⁾ 日本図書館情報学会用語辞典編集委員会『図書館情報学用語辞典 第5版』丸善出版、2020年、162頁。

報資源を収集・保存・提供する仕組みの総体」⁵⁾、「社会が残すことを選択した／すべき知識情報基盤としてのデジタルデータとそれにまつわる仕組みの総体。真正性や永続性の確保、万人へのアクセス保障がその要件となる」⁶⁾等が挙げられる。最大公約数的な理解では、デジタルコンテンツの収集、蓄積、保存、組織化、オンライン提供、データ共有といった一連の機能を備えたシステムの総称とみなすことができ、記憶機関がその業務の一環として構築しウェブ上で公開することが多い。なお、これは本来、和製英語であり、欧米においては、このような総称としての術語は基本的に存在しない。英語を例に取れば、コンテンツの由来・性格や説明の力点の置き方により、“digital library”、“digital heritage”、“digital archives”、“digital preservation”等の呼称が使い分けられるのが通例である。

1.2. 現状

Project Gutenberg (1971 年開始)⁷⁾等の電子図書館を最初期の例として、資料のデジタル化は欧米先行により長い歴史を持つが、インターネット及びウェブの普及に合わせ、デジタルコンテンツをオンラインで提供するデジタルアーカイブは、特に 2000 年代以降に急増・拡大している。デジタル化資料の増加とともに、その性格もかつての優品主義から脱却して社会の知識インフラの一翼を担う存在となりつつあり、記憶機関における業務上の重要性も顕著に高まっている。また、特色ある地域資料を対象とした、地域の記憶の担い手としての役割も近年、デジタルアーカイブの機能として注目され始めているところである。市民の情報環境の「主戦場」がウェブ上に移りつつある現状において、各記憶機関にとってはデジタルアーカイブを整備することが、自機関のサービス及びプレゼンスの向上に直結するという状況が醸成されていると言える。

デジタルアーカイブは機関の大小を問わず多数存在するが、目下の代表的な大規模デジタルアーカイブの例を挙げれば、国立国会図書館デジタルコレクション⁸⁾、フランス国立図書館の Gallica⁹⁾、米国の民間機関による Internet Archive¹⁰⁾などがある。また、自らはコンテンツを持たず、参加機関の提供コンテンツへのアクセスを集約して提供するポータル

⁵⁾ デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会「我が国におけるデジタルアーカイブ推進の方向性」2017 年、2 頁。

(https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/houkokusho.pdf)

⁶⁾ 福島幸宏「図書館機能の再定置」『Library Resource Guide』31 号、2020 年、13 頁。

⁷⁾ <https://www.gutenberg.org/>

⁸⁾ <https://dl.ndl.go.jp/>

⁹⁾ <https://gallica.bnf.fr/>

¹⁰⁾ <https://archive.org/>

態をとるデジタルアーカイブとしては、ジャパンサーチ¹¹⁾、Europeana¹²⁾、Digital Public Library of America (DPLA)¹³⁾等がある。

2. デジタルアーカイブの構成要素

デジタルアーカイブは多様な要素技術の集合体であり、そのため構築・運用にあたって必要となる知識・技術も多岐にわたる。以下、柳による基本的枠組の提示¹⁴⁾の他、知的財産戦略本部のガイドライン 2 種¹⁵⁾、国立国会図書館による「資料デジタル化の手引き」¹⁶⁾、総務省のガイドライン¹⁷⁾等の先行する各種ドキュメントを基にデジタルアーカイブの一般的な構成要素を抽出・整理し、同時にその各局面において求められる知識・技術を検討する。

2.1. 企画・計画

デジタルアーカイブ構築にあたっては、予めその目的、コンセプト、構築のメリット、提供コンテンツ、想定されるユーザー層、予算（長期的なランニングコストも含む）、実施体制等について、組織のミッションとの整合性も踏まえた上で検討することが最低限必要となる。また、ここで策定した各種方針は体系的にドキュメント化しておくことが望ましい。いずれにせよ、組織における予算獲得のためには予算権者の理解と同意は不可欠であり、その意味でも必ず通過しなければならないプロセスである。

¹¹⁾ <https://jpsearch.go.jp/>

¹²⁾ <https://www.europeana.eu/>

¹³⁾ <https://dp.la/>

¹⁴⁾ 柳与志夫『デジタルアーカイブの理論と政策：デジタル文化資源の活用に向けて』勁草書房、2020年、174-181頁。

¹⁵⁾ デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会「デジタルアーカイブの構築・共有・活用ガイドライン」 (https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/guideline.pdf)、

及びデジタルアーカイブジャパン推進委員会及び実務者検討委員会「デジタルアーカイブのための長期保存ガイドライン（2020年版）」 (https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/pdf/guideline2020.pdf)

¹⁶⁾ 国立国会図書館「国立国会図書館資料デジタル化の手引 2017年版」 (http://dl.ndl.go.jp/view/download/digidepo_10341525_po_digitalguide170428.pdf?contentNo=1&alternativeNo=)

¹⁷⁾ 総務省、前掲書。

これを適切に行うためには、構築すべきデジタルアーカイブを現代の情報環境の中に正しく位置づけられる広い視野と情報収集能力に加え、企画立案能力、文書作成能力といった、一般的なプロジェクト管理の最上流工程で必要とされる素養が求められる。

2.2. デジタルコンテンツ

デジタルアーカイブには当然、その内容たるデジタルコンテンツが必要である。換言すれば、その工程にはデジタルコレクションの構築が必然的に含まれる。デジタルコレクションの構築にあたっては、当該デジタルアーカイブが扱う資料の範囲や収集方法について予め定めたコレクション構築方針を策定しておくことが望ましい。

デジタルデータの収集方法には大きく、アナログ資料のデジタル化とボーンデジタル資料の入手の2通りがあるが、ここでは現在までのデジタルアーカイブの主流となってきた前者（特に画像データの作成）について述べる。デジタル化にあたっては、外注・自機関実施のいずれの場合であっても、作業フロー全体を設計した上で、詳細な技術仕様を定める必要がある。そのためには、画像データ（画像フォーマット、解像度、色空間等）、機器（各種スキャナ、デジタルカメラ等）、保存媒体（光ディスク、HDD等）等についての一連の技術的な知識が要求される。

また、デジタル化にあたっては、対象資料それ自体についての知識、及び資料の扱いに習熟していることが求められる。資料の劣化状況、特殊な判型の資料の有無、権利面の制約（特に著作権保護期間中の資料や寄託資料の場合）等の確認も踏まえ、必要な事前作業（修復、解体、権利処理等）を施した上で、資料の扱いに習熟した要員によってデジタル化作業が行われる必要がある。

最後に、出来上がった成果物については、求める仕様を満たしていることを確認することにより、適切な品質管理を実現しなければならない。ここでは、検品のためのツールや体制が整っていることが条件となる。

2.3. メタデータ

デジタルアーカイブのみならず、電子情報の流通全般においてメタデータは極めて重要な機能を持つ。作成・収集されたデジタルデータは、然るべきメタデータの付与・組織化の工程を経ることによって、初めて発見・識別・入手が可能となる。

一般に、メタデータには記述メタデータと管理メタデータの2種類がある。アナログ資料をデジタル化した場合、原資料のタイトルや責任表示といったメタデータ（書誌情報等）は記述メタデータとして引き継がれ、デジタル化作業日や権利情報等の管理上の内部情報は管理メタデータとして新たに作成されることが多い。

幅広いデータの共有・連携（2.7.参照）を視野に入れるのであれば、メタデータは何らかの標準的なスキーマ（語彙・記述規則）に沿って作成されることが望ましい。原資料のメタデータが準拠するスキーマに倣い、デジタルデータのメタデータも標準的な概念モデル

(FRBR、ISAD(G)、CIDOC CRM 等)、語彙・記述規則 (Dublin Core、EAD、METS 等) から選択されることになるため、採用するメタデータスキーマに関する概念的な理解は最低限必要となる。また、実際のメタデータは XML や JSON といった形式で記述 (エンコーディング) されることが多いため、これらデータ記述言語についても把握しておく必要がある。

さらに、近年はデジタルデータへの長期的なアクセス保証の観点から、永続的識別子 (Persistent Identifier) をメタデータとして付与することが増えている。永続的識別子の基本的な仕組みを理解した上で、DOI (Digital Object Identifier)¹⁸⁾、ARK (Archival Resource Key)¹⁹⁾等の選択肢から、付与対象や目的に応じて適切なものを採用する必要がある。

2.4. データの保存

一般にアーカイブは資料の恒久的な保存及び利用を保証する仕組みであり、デジタルアーカイブにおいてもデータの長期保存は本質的な課題の一つである。

長期保存にあたっては、まず保存媒体そのものの寿命の問題がある。あらゆる媒体には固有の耐用年数がある。光ディスク、磁気テープ、HDD、クラウド等、それぞれの特性やメリット・デメリットを理解した上で適切な媒体を採用することが必要である。また、データの保全のためには、適切な保存環境 (温湿度等) の整備、複数箇所での保存 (バックアップの作成)、適切な検査の実施 (ハッシュ値の取得、定期検査等)、必要に応じた媒体変換といった作業も行われる必要がある。

次に、ソフトウェア・ハードウェア両面における再生環境の旧式化・陳腐化の問題がある。旧式 OS や古いアプリケーション、又は既に廃れた機器を動作条件とするコンテンツの場合、長期の利用保証を行うためには、データ自体の保全に加え、マイグレーション (媒体変換やファイル形式の変換) やエミュレーション (旧式環境のソフトウェアを現在の環境で利用可能とする技術) といった措置を講ずる必要がある。

デジタルデータの長期保存については、各種の国際標準がある。代表的なものとしては、概念モデルである OAIS 参照モデル²⁰⁾、メタデータ標準である PREMIS²¹⁾等が挙げられる。また、英国の Digital Preservation Coalition が作成・維持している Digital Preservation Handbook²²⁾に代表される優れた手引きも複数存在する。こうした国際的な動向に目を配り、知見を蓄えておくことも重要である。

¹⁸⁾ <https://www.doi.org/>

¹⁹⁾ <https://arks.org/>

²⁰⁾ ISO14721 として国際標準化されている (<http://www.oais.info/>)。

²¹⁾ <https://www.loc.gov/standards/premis/>

²²⁾ <https://www.dpconline.org/handbook>

このように、長期保存はそれ自体で一つの専門領域を成す射程の広さを持っており、高度に技術的な内容も含む。デジタルデータ管理の実務を担う情報専門職には、一通りの知識を持つことが求められる。

2.5. システム構築・運用

デジタルアーカイブは情報システムであるため、開発工程は必須である。開発方法には、ゼロから開発を行うスクラッチ開発、オープンソースのコンテンツ管理システム（CMS）の利用、既存の商用パッケージの導入、クラウドサービスの利用、といった選択肢がある。最終的なシステム仕様は、実現すべき要件・予算等を総合的に勘案して決定される。ユーザーインターフェース（UI）や検索システムの検討もこの段階で行うことになる。

システムの構築・稼働後は日常的な運用保守が必要であり、サービスレベル及びセキュリティレベルを維持しつつ、障害対応、ソフトウェアの定期的なアップデート、機器のリプレース、サポート切れへの対応、といった作業が継続的に行われる必要がある。ここでも恒常的な予算・実施体制が確保されていることがサービス継続の条件となる。

システム導入・運用保守を委託業者が行う場合であっても、発注者及び管理者として、ICTの基礎知識があることが望ましい。特に公的機関の場合、適正な価格で適切な物品・役務を調達するためには、技術的な事項について受託業者と正確なコミュニケーションを行い、適切な判断を下せるようにするための基礎的な素養は必須と言える。

2.6. 権利・ライセンス

デジタルデータの公開にあたっては、著作権法の遵守が必要であるとともに、適切なライセンスの付与・表示を行うことが望ましい。著作権保護期間内のコンテンツや権利関係の不明な資料の場合、公開に向けて権利処理（個別の権利者からの許諾取得、文化庁長官裁定による一括処理等）が必要になることもある。

また、一般にユーザーにとっては、デジタルアーカイブ上で発見したコンテンツをどこまで二次利用可能であるのか（商用利用が可能なのか、改変可能なのか、著作権者のクレジット表記が必要なのか等）は必ずしも自明ではなく、そのことが、利用における放恣や、逆に過度の萎縮を招く可能性がある。ユーザーにとっての利用可能性をライセンスの形で明示しておくことは、デジタルコンテンツの健全な利活用を促し、ひいては社会の知識基盤・創造基盤としてのデジタルアーカイブの所期の目的に適うことになる。現在、一般的にライセンスの表示にあたっては、クリエイティブ・コモンズ（Creative Commons）²³⁾の体系に基づく表示（CC BY、CC0等）が行われることが多い。

²³⁾ <https://creativecommons.org/>

2.7. データ交換・他機関との連携

デジタルアーカイブの持つ大きな可能性の一つは、機関を超えた連携やデータ交換にある。前述のとおり、共通の形式でメタデータを用意しておくことにより、異なるシステム間でのデータの互換性を実現する相互運用性(interoperability)を担保することが望ましい。前述のとおり、ジャパンサーチ、Europeana、DPLA といった大規模プラットフォームは、参加機関からメタデータを集約して統合的な検索とコンテンツへのアクセスを提供するサービスであり、参加機関にとっては、このようなプラットフォームにデータ提供を行うことで、広い利活用を期待することができる。

また、各種の API (Application Programming Interface) をシステムに実装することで、データを広く利活用してもらえらる仕組みを作ることも重要である。例えば、OAI-PMH²⁴⁾に対応することで第三者によるメタデータの機械的収集 (harvesting) を可能にすること、画像データを IIIF²⁵⁾ 準拠とすることで画像データの相互運用性を担保すること、RDF²⁶⁾ 形式でメタデータを整備することでリンクトデータとしての活用につなげること、等が良く行われる。

2.8. 利活用促進

デジタルアーカイブは、想定するユーザー層に正しくリーチし、利活用されることが重要であり、そのためにウェブ上における可視性向上の努力が求められる。可視性向上のためには、SNS 等での情報発信による広報活動や、前項で述べた各種プラットフォームとのデータ連携が有効である。また、利用者教育の視点から、利用法についての発信や研修²⁷⁾を行うことも効果的である。

デジタルアーカイブについては、その提供コンテンツや UI に目が向きがちであるが、正しく利活用されることは組織内での継続的な予算の確保にもつながり、それがシステムの持続可能性にも直結することになる。また、デジタルアーカイブにブランディング効果を期待する場合、利用促進は一層重要となる。そのため、この点における取組は軽視されてはならない。

²⁴⁾ <https://www.openarchives.org/pmh/>

²⁵⁾ <https://iiif.io/>

²⁶⁾ セマンティックウェブ実現のためのデータ記述形式。主語・述語・目的語からなるトリプル (triple) と呼ばれる形式で情報資源を記述する。現在は W3C 勧告で定められている (<https://www.w3.org/RDF/>)。

²⁷⁾ 柳、前掲書、222 頁。

3. 考察

3.1. 現状の課題

以上、デジタルアーカイブの構築・運用に際しての要素及び必要な専門知識について概観してきた。情報専門職の関わる業務の中でデジタルアーカイブに関するものだけを取り上げても、マネジメント的側面から技術的側面に至るまで、多様かつ高度な専門性が要求されることが明らかとなった。

加えて、デジタル環境下において情報専門職の関わる業務は、当然、デジタルアーカイブのみにとどまらない。一例を挙げるならば、2018年の著作権法改正によって、いわゆるビッグデータの機械学習等への利用が権利制限規定に盛り込まれた²⁸⁾。人だけでなく機械もユーザーとして想定すべき時代に既に入っており、データサイエンスや AI といった実践領域が記憶機関の先端的な業務に加わりつつある²⁹⁾。

このように情報専門職に求められる能力は拡大の一途をたどっており、目下、その専門性の再定義と涵養が問われている局面にあると言えるが、翻って日本の現状を顧みると、「中でも人材不足は深刻である。（中略）残念ながら、デジタルアーカイブ化が拡大している各分野で、これらの能力を有し開発・運用・活用の担い手になる人材は不足しており、人材育成は喫緊の課題」³⁰⁾という状況にある。最後に、この点について若干の考察を付言して本稿の締めくくりとしたい。

3.2. デジタル時代の情報専門職の専門性涵養に向けて

人材の育成にあたっては、第一に情報専門職の資格要件の現代化が長期的には有効と考えられる。現在、日本の代表的な情報専門職の国家資格である司書³¹⁾及び学芸員、並びに国立公文書館が審査を行う認証アーキビストのいずれにおいても、その取得・認証にあたって必要とされる履修科目要件・認証要件は、デジタル時代の要請という点から見た場合、

²⁸⁾ 文化庁「著作権法の一部を改正する法律（平成30年法律第30号）について」（https://www.bunka.go.jp/seisaku/chosakuken/hokaisei/h30_hokaisei/）

²⁹⁾ 例えば Europeana は戦略方針で技術的イノベーションの例として AI の活用を謳い、メタデータ品質の自動改善等を具体的な目標として掲げている（“Europeana Strategy 2020-2025”（https://pro.europeana.eu/files/Europeana_Professional/Publications/Europeana%20Strategy%202020-%20-%202025.pdf））。

³⁰⁾ 三宅茜巳・井上透・松家鮎美「デジタルアーカイブと人材育成：知識基盤社会を支える人材-デジタル・アーキビスト-育成教育」『デジタルアーカイブ学会誌』2(4)、2018年、376-377頁。

³¹⁾ 司書は、正確には図書館法上の図書館（公共図書館）の専門職員の資格であるが、大学図書館・各種専門図書館等においても、事実上の資格要件となっていることが多い。

必ずしも十分な内容となっていない³²⁾。これら公的資格の付与要件の再検討と現代化がなされれば、デジタル時代の専門職人材育成の政策的な裏付けとなると考えられる。

第二に、個々の人材養成という視点では、現職者向けの学習・研修機会の拡充を望みたい。情報専門職については、専門職団体、公的機関、有志等の主催によって各種の研修・講習・勉強会が活発に行われる文化が既に存在するが、例えばデジタルアーカイブの基礎知識を習得できる機会は未だ少ない。特に中小規模の機関にとっては、量・質ともに必ずしも組織内の OJT が十全に行える状況にあるとは限らないこともあり、開かれた継続的な研修機会の確保が望まれる³³⁾。大学におけるリカレント教育の一環として、関係科目を開講する形式も考えられる。

最後に、第二の点とも関わるが、組織を超えた情報共有の重要性の可能性について述べたい。多岐にわたるデジタルアーカイブの構成要素も、個別に取り出せば内容はある程度共通化されていると言うことができ、業務フローの標準的なモデル化や技術のコモディティ化も進みつつある。知識・ノウハウを持つ機関・研究者・情報専門職個人がそうした情報を積極的に発信・共有し、あわせて情報専門職同士の情報交換と協働を可能とする実務者コミュニティがさらに醸成されることによって、国内の情報専門職全体の能力の底上げにつながっていくことが可能になるだろう。

³²⁾ 司書の例を挙げれば、図書館法施行規則の定める司書科目のうち情報技術を直接的に主題とする科目は「図書館情報技術論」（2単位）のみである。

³³⁾ デジタルアーカイブに特化した研修の例としては、NPO 法人 IRI 知的資源イニシアティブによる各種セミナー (<https://www.iri-net.org/event/digital-archives-class/>) 等があり、国立国会図書館は資料デジタル化や権利処理といった各論に関する講師派遣研修及び遠隔研修を実施している (<https://www.ndl.go.jp/jp/library/training/index.html>)。また、特定非営利活動法人日本デジタル・アーキビスト資格認定機構の認定する「デジタル・アーキビスト」資格及び当該資格取得のための講習会があるが、現状では民間の一資格に留まり、認定養成機関も未だ少数の大学に限られる (<https://jdaa.jp/>)。