

実務者検討委員会 第一次中間取りまとめ(案)

目次

はじめに	2
1. これまでの経緯	3
2. 議論の背景	4
3. 実務者検討委員会で示されたデジタルアーカイブの構築・連携の現状と課題	6
4. デジタルアーカイブ社会の実現に向けた施策の検討のために	8
5. 国の分野横断統合ポータル構築	16
6. 今後の主要検討課題	21
おわりに	21

(補足資料)

○共通メタデータフォーマット

- ・連携フォーマットについて

- ・ジャパンサーチ(仮称)の連携フォーマット詳細版

- ・ジャパンサーチ(仮称)利活用フォーマット

- ・分野横断統合プラットフォームのメタデータフォーマット仕様

○デジタルアーカイブアセスメントツール

(関連資料)

- 工程表(平成 29 年 9 月 5 日 デジタルアーカイブジャパン推進委員会決定)

- 我が国におけるデジタルアーカイブ推進の方向性(報告書) エグゼクティブ・サマリー

- デジタルアーカイブの構築・共有・活用ガイドライン(ガイドライン) - 概要 -

はじめに

現在、我々を取り巻く「サイバー」空間の環境は刻一刻と変化しており、利便性の高いサービスが次々と提供される反面、その担い手の多くが国際プラットフォームであることもあり、「サイバー」空間において日本の長い歴史や伝統に根差した文化の重層性や多様性が埋没してしまうのではないとの危惧が呈されることも多い。

一方、「リアル」社会においては、ライフデザインや働き方における価値観が多様化し、シェアリングエコノミーに代表されるように必ずしも「所有」に拘らず必要に応じた様々な価値を「シェア」といった行動が広く行われるようになってきている。

経済活動や社会生活などどの局面においても「サイバー」の存在感は増大する一方であるが、他方で、「リアル」と「サイバー」の結びつきを強化する技術でもある、IoTや人工知能(AI)、ビッグデータといった以前にはなかった新技術の開発・応用も進んできている。こういった新技術の進展は、データを媒介にした異業種同士や供給者と顧客の直接の結びつきを加速させ、コンテンツ¹やデータを距離、時間、費用、言語の制約なく共有することを可能とする。すなわち、モノづくりやコンテンツの作成・発信、マネタイズや評価・価値付けに至るまで誰もが容易に行うことができるようになる。

こういった流れを踏まえ、我が国の企業・大学・行政機関など様々な主体が保有する知的資産²を可能な限りオープンな形でシェアし、その知的資産を利活用することで「サイバー」から「リアル」へ誘導または補完する役割を担い、新たな価値を生む足がかりとなりうるデジタルアーカイブ³の意義について今一度共有した上で、我が国でも、デジタルアーカイブジャパンとして充実させていく必要がある⁴。

従前は、権利による保護も含め、データを独占ないし秘匿することによりその価値の最大化が図られることが多かったが、近年はソフトウェアの開発領域では、GitHubなどに代表されるような開発プロジェクトのための共有ウェブサービスの取組や学術領域におけるオープンサイエンスの取組などオープンな形でシェアを図ることで価値の最大化や効率化を図る流れが形成されてきている。

こういった流れは、今後さらに拡大されることが予想されるところ、合わせて、データの提供機関が真正性や正確性を保証した上でオープン化⁵することで、二次創作物の価

¹ 個々の社会・文化・学術情報資源。「デジタルコンテンツ」の他、アナログ媒体の資料・作品等を含む。

² メタデータや写真、動画などのコンテンツに限らず、3Dデータや研究データなどありとあらゆるデータをいう。

³ ここでは、様々な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。デジタルアーカイブで扱うデジタル情報資源は、「デジタルコンテンツ」だけでなく、アナログ媒体の資料・作品も含む「コンテンツ」の内容や所在に関する情報を記述した「メタデータ」や、コンテンツの縮小版や部分表示である「サムネイル/プレビュー」も対象とする。

⁴ 現在国連で提唱されているSDGs(Sustainable Development Goals:持続可能な開発目標)の観点から、いずれは各国で構築されているナショナル・デジタルアーカイブと連携することにより、国際的なデータ提供インフラとしても機能・貢献することも視野にいれておくべきとの指摘もあった。

⁵ 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状態をい

値の向上も期待できることばかりでなく、場所や時間を超えて正しい情報を見つけることも可能となる。

こうしたデータやコンテンツの共有基盤となるデジタルアーカイブジャパンの構築には、個々のアーカイブ機関⁶におけるデジタル化やシステム整備、権利処理などへの継続的な取組を基本としつつ、制度的課題の解決や関係者間での調整が必要なこともあり、そのためには相応のプロセスと時間が必要となる。平成 29 年度に開催された実務者検討委員会においては、平成 29 年 9 月のデジタルアーカイブジャパン推進委員会において合意した範囲・事項⁷について、議論や取組を進めてきており、その成果を本中間取りまとめにおいて取りまとめることとした。

1.これまでの経緯

デジタルアーカイブに関する取組としては、平成 29 年 4 月、「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⁸」において、報告書⁹及びガイドライン¹⁰が取りまとめられている。

また、これを踏まえ、「知的財産推進計画 2017」(平成 29 年 5 月知的財産戦略本部決定)において、知的財産戦略の一環として、2020 年に向け、デジタルアーカイブ構築に係る各種の施策に取り組むことが決定されたほか、「経済財政運営と改革の基本方針 2017」(骨太方針)及び「未来投資戦略 2017」(いずれも平成 29 年 6 月閣議決定)においても、デジタルアーカイブ施策の必要性や方向性が示されている。

このような施策の必要性及び方向性を受け、平成 29 年 9 月に第 1 回デジタルアーカイブジャパン推進委員会が開催され、様々な分野におけるデジタルアーカイブ構築の取組について工程表が決定されるとともに、デジタルアーカイブ推進に係る実務的課題に対応するため、実務者検討委員会を設置することが決定された。実務者検討委員会においては、関係省庁等連絡会・実務者協議会の報告書で示された施策等の取組状況について整理するとともに、デジタルアーカイブ構築・利活用に係る課題への対応策につ

う。なお、本ガイドラインは、PDM、CC0、CC BY、特に CC4.0 の利用条件での公開を推奨している。

⁶ アーカイブ/アーカイブズは、公文書館等を指す言葉として一般的に用いられてきたが、ここでは広い意味での記録機関全般を指す。社会・文化・学術情報資源である資料・作品等のコンテンツを収集し、その資源を整理(組織化)し、保存し、提供する機能を持つ機関・団体等をいう。博物館・美術館、図書館、文書館といった文化的施設のほか、大学・研究機関、企業、市民団体、官公庁・地方公共団体等を含む。なお、提供機能が限定的であり、一般への公開を想定していない機関等も含む。

⁷ http://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/suisin/dai1/gijisidai.html
資料 1 の 8 ページ「デジタルアーカイブジャパンの当面の対象範囲」及び資料 4 の 1 ページ「デジタルアーカイブジャパンの実現に向けた工程表(全体)」

⁸ https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/index.html

⁹ 我が国におけるデジタルアーカイブ推進の方向性(平成 29 年 4 月)「関連資料」中にサマリーを添付 https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/houkokusho.pdf

¹⁰ デジタルアーカイブ構築・共有・活用ガイドライン(平成 29 年 4 月)「関連資料」中に概要を添付 https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/guideline.pdf

いて議論を行った。

2. 議論の背景

様々なコンテンツのメタデータ¹¹整備やデジタルコンテンツ¹²の拡充をはじめとするデジタルアーカイブに関する取組を進めることは、文化の保存・継承・発展の基盤になるだけでなく、保存されたコンテンツの二次的な利用や国内外に発信する情報の基盤となる重要な取組である。観光、教育、学術、防災など様々な用途での活用を通じて、デジタルアーカイブの構築・共有と活用の循環を持続的なものとし、その便益を国民のものとしていくことで、我が国の社会的、文化的、経済的発展につなげることが重要である。効率的・効果的なデジタルアーカイブの活用を推進するためには、我が国が保有する様々なコンテンツへのナビゲーションを行い、様々な分野や地域の情報と活用者をつなぐための連携基盤としての役割を果たす「国の分野横断統合ポータル」¹³の構築も重要である。

現在、我が国の「アーカイブ機関」は、国、独立行政法人、地方自治体、民間等の多岐にわたっている。デジタルアーカイブの構築は一部で進みつつあるが、世界の動向を踏まえると、メタデータやデジタルコンテンツの提供は十分とは言えず、分野横断的なアーカイブの連携に関する取組についてはほとんど進められていない状況にある。

¹¹ データに関するデータを意味し、「サムネイル/プレビュー」や用語を統制するための語彙等も含むものとして広く定義されることもあるが、ここでは、コンテンツの内容、外形、所在等に関する記述等のデータをいう。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や URI 参照のデータを指す。

¹² アナログ媒体の資料・作品等をデジタル化した高品質なデジタルコンテンツ、あるいはボーンデジタルの作品(デジタルカメラの写真、電子書籍等)であればそのデジタルコンテンツ自体。

¹³ 我が国が保有する様々なコンテンツのメタデータに関し、分野横断的な検索を可能とするポータルサイト。「知的財産推進計画 2015」において、その構築の必要性が掲げられ、「知的財産推進計画 2017」では、国立国会図書館が、2020 年までに、国の分野横断型の統合ポータル「ジャパンサーチ(仮称)」の構築を目指すものとされている。

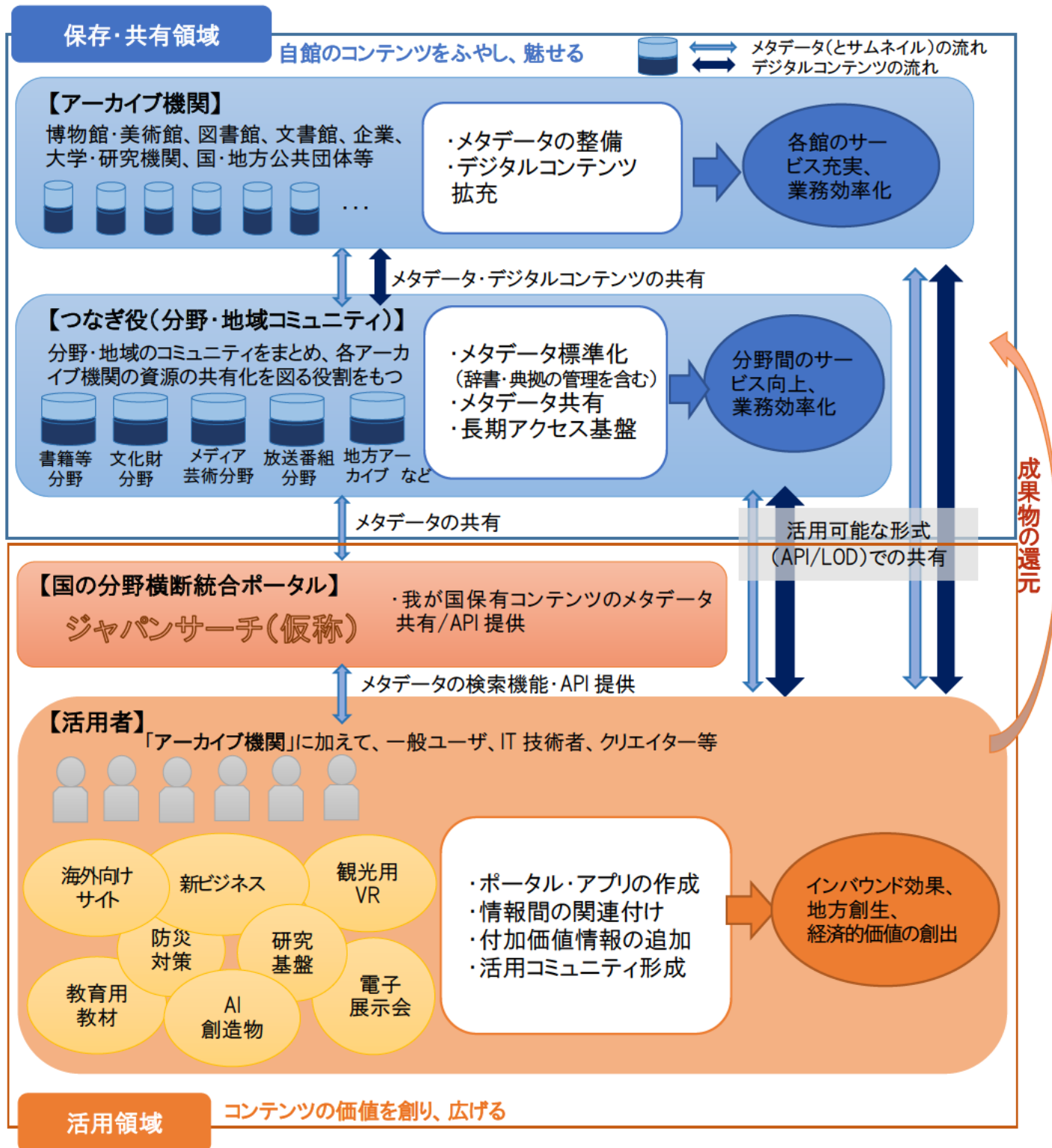


図1 デジタルアーカイブの共有と活用のために

こうした状況の下、関係省庁等連絡会・実務者協議会の報告書では、我が国におけるデジタルアーカイブの取組を推進するために、図1に示すように、分野・地域ごとの「つなぎ役」¹⁴を介して、「アーカイブ機関」と国の分野横断統合ポータル「ジャパンサーチ（仮称）」の間でメタデータの共有を進め、「活用者」¹⁵がジャパンサーチ（仮称）等からメタデータを共有し、さまざまな用途に活用するというサイクルの構築を目指す方向性が示された。また、アーカイブ機関は、関係省庁等連絡会・実務者協議会のガイドラインに沿って、デジタルアーカイブに関する取組（メタデータの整備、サムネイル/プレビュー¹⁶の作成、デジタルコンテンツの拡充、整備したメタデータやサムネイル/プレビューのオープン化、デジタルコンテンツの利用条件表示など）を行うことが望ましいこと、デジタルアーカイブ推進のためには、分野・地域のコミュニティの主体となるつなぎ役の果たす機能・役割が重要であることなどの方向性も打ち出された。

3. 実務者検討委員会で示されたデジタルアーカイブの構築・連携の現状と課題

実務者検討委員会においては、関係省庁及び主要アーカイブ機関から、デジタルアーカイブを構築・利活用していく上での現在の取組状況や今後の課題などについて、次のとおり報告が行われた。

○書籍等分野（国立国会図書館）

書籍等分野のつなぎ役として、「国立国会図書館サーチ」を運用中。資料のデジタル化、メタデータのオープン化¹⁷、デジタルコンテンツの二次利用条件表示の促進等の課題があるが、これらについては国全体としての取組が必要である。

○公文書分野（独立行政法人国立公文書館）

国立公文書館では、「国立公文書館デジタルアーカイブ」と「アジア歴史資料センターデジタルアーカイブ」の2つを運用中。前者はデジタル化の更なる推進、後者はデータベースの拡充が、それぞれ課題となっている。

○文化財分野（文化庁伝統文化課）

文化遺産オンラインを国立情報学研究所と運営している。文化遺産オンラインは、国立国会図書館サーチと連携するとともに、独立行政法人国立文化財機構の国

¹⁴ Europeana の「アグリゲーター」、DPLA の「ハブ」に相当する役割・機能を果たす機関のことをいい、分野・地域コミュニティにおけるメタデータを集約し、API 等による提供を行う機関。メタデータの標準化、用語の統制（辞書・典拠・シソーラスの管理）を行い、コミュニティにおけるメタデータ整備やデジタルコンテンツ作成への支援、長期アクセス保証のための基盤提供、活用コミュニティの形成などの役割を担う。単独の組織が担う場合もあれば、行政、NPO や大学等と分担して行う場合もありうる。

¹⁵ デジタルアーカイブの様々なデータを活用する者。「アーカイブ機関」が自らのために活用することに加え、一般ユーザ、IT 技術者、クリエイターなど、様々な機関・団体・個人がデータの活用者になりうる。

¹⁶ コンテンツの要約又は一部分の表示。縮小した画像（サムネイル）、本文テキストの一部表示や数秒程度の音声・動画（プレビュー）等をいう。

¹⁷ 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状態をいう。なお、ガイドラインでは、PDM、CC0、CC BY、特に CC4.0 の利用条件での公開を推奨している。

立博物館所蔵品統合検索システム (ColBase) とも連携しており、文化財分野の他のデータベースとの連携にも努めていく。文化遺産オンラインでは、国指定文化財に係る文字情報は全てデジタル化の上で公開しているが、画像等の資料デジタル化や二次利用条件表示の促進等が課題である。

○文化財分野 (独立行政法人国立美術館)

今年度から来年度にかけて、国立美術館の所蔵品の来歴等メタデータの精緻化を進めている。また、公開されている情報資源のゲートウェイを構築する計画で、将来的には、「ジャパンサーチ (仮称)」や文化遺産オンラインとの連携を進める予定である。

○文化財分野 (独立行政法人国立文化財機構)

国立文化財機構の4つの国立博物館の所蔵品を検索できる「国立博物館所蔵品統合検索システム (ColBase)」を運営している。メタデータやサムネイルは CC BY¹⁸ で公開している。現在、国立国会図書館サーチ、文化遺産オンラインとの連携に向けた調整を進め、ColBase のメタデータ及び画像の充実に努めている。

○メディア芸術 / メディアコンテンツ分野 (文化庁芸術文化課)

メディア芸術データベースの 2020 年正式公開を目指し、現在、国立国会図書館サーチとの API¹⁹ 連携を検討中。補助金によるメディア芸術作品のアーカイブ化支援、一部作品のサムネイル画像の表示、利用条件表示にも取り組んでいる。専門的な能力を持つ人材の育成が課題である。

○自然史・理工学分野 (独立行政法人国立科学博物館)

生物多様性の分野では、地球規模生物多様性情報機構 (GBIF) が世界規模のポータルを運営している。日本では、「S-Net (サイエンス・ミュージアムネット)」というポータルを国立科学博物館が運営している。また、科学史資料や産業技術に関する歴史的資料の所在については、国立科学博物館が運営する「産業技術史資料データベース」で公開している。これら活動の継続・推進・教育、データ等のデジタル化、非生物 (岩石・鉱物等の地球科学系、産業技術史系) データへの対応が課題である。

○人文学分野 (大学共同利用機関法人人間文化研究機構)

人間文化研究機構に属する6機関等のデータベースを検索できる「統合検索システム nihuiNT」を運営し、国立国会図書館サーチとも連携している。nihuiNT のライセンスは、現状は各機関に委ねられているが (例えば、国文学研究資料館の歴史的典籍のデータは CC BY-SA²⁰)、ガイドラインに沿ってできるだけオープンにするよう

¹⁸ 原作者のクレジット (氏名、作品タイトルなど) を表示することを主な条件とし、改変はもちろん、営利目的での二次利用も許可される最も自由度の高い CC ライセンス。

¹⁹ Application Programming Interface。オペレーティングシステムやアプリケーションソフトが、他のアプリケーションソフトに対し、機能の一部を利用できるように提供するインターフェースのことをいう。

²⁰ 原作者のクレジット (氏名、作品タイトルなど) を表示し、改変した場合には元の作品と同じライセンス (こ

努めている。データ化の促進とデータの統一性の確保(データクレンジング)が課題である。

○放送番組分野(公益財団法人放送番組センター)

放送ライブラリーで公開する番組のメタデータを検索・閲覧できるシステムを運用しているほか、ドラマのメタデータをジャパンサーチ(仮称)プロトタイプと連携(サンプルデータの提供)する作業を進めている。また、公開番組の視聴ニーズに対応するため、図書館などの公共施設や大学の講義の中で利活用する事業を進めている。

○放送番組分野(日本放送協会)

NHK アーカイブスが保存している映像のポータルサイトを公開しており、「みのがしなつかし」「みちしる」など6分野から選択して視聴が可能。放送番組センターによるジャパンサーチ(仮称)プロトタイプへのサンプルデータ提供への協力に加え、NHK アーカイブスの一部データとジャパンサーチ(仮称)プロトタイプとの連携についても検討を進めている。

上記の各分野における取組のほか、分野横断的な取組として、内閣官房明治 150 年関連施策推進室が「明治 150 年」ポータルサイトを立ち上げ、明治時代に関連するデジタルアーカイブ情報の集約を進めている。

4. デジタルアーカイブ社会の実現に向けた施策の検討のために

実務者検討委員会においては、デジタルアーカイブに関する関係省庁等連絡会・実務者協議会の報告書で示された課題への対応策についても検討が行われた。

(1)「デジタルアーカイブ社会」とは

報告書において提言された「デジタルアーカイブ社会」とは、「デジタルアーカイブが日常的に活用され、様々な創作活動を支える社会・学術・文化の基盤となる社会」と位置付けられる。すなわち、デジタルアーカイブによって、日々生み出される様々なデータが共有され誰でも簡単にアクセスできること、また、それが日常的に利活用されやすい条件で提示されており、誰でも自由に様々なシーンで新しいコンテンツが生み出せる社会である。

平成 29 年 11 月、知的財産戦略推進事務局が産学官のデジタルアーカイブ関係者を一同に集めたフォーラム(産学官フォーラム)を開催したが、そこでは、玉石混交の情報の中から自ら必要とする正確な情報だけが取れる社会、アーカイブとしてどこかに置かなくとも、自然と蓄積される仕組みが実現される社会を目指すこと、そのためには利活用する人がデータを提供するアーカイブ機関へのリスペクトをもって一定のルールやモラ

のライセンス)で公開することを主な条件に、営利目的での二次利用も許可される CC ライセンス。

ルに従いながら利活用を行う必要があるといった意見があった。今後も引き続き産学官フォーラム等の機会を通じ、広く社会一般に施策の方向性について周知するとともに、関係者との情報共有、意見交換を行いながら、目指すべき姿と、その実現のために必要とされる課題解決の方向性を探っていくことが必要である。



図2 デジタルアーカイブ社会のイメージ(例)

(2)デジタルアーカイブ構築の推進

(メタデータの整備)

デジタルアーカイブを活用していくためには、コンテンツにメタデータを付与し、サムネイル／プレビューの整備を進めていくことが必要となる。しかし、規模の小さいアーカイブ機関においては、人手やノウハウの不足により、これらの整備が進んでいないため、行政やつなぎ役を通じた支援が必要とされている。また、つなぎ役に求められる役割の一つに分野内のメタデータ標準や用語の統制といったものがあるが、分野の知識とメタデータの知識双方を有する専門家でなければ分野のメタデータ標準を作ることが難しいといった意見もあった。

これらの課題を解決するためには、当該分野において主導的立場にある大規模なアーカイブ機関が、自らのメタデータ付与の基準や方法を公開していくといった解決策²¹も考えられる。また、AIによる顔認識、物体認識、文字列認識などの機能を活用してメタデータを自動付与するといった新たな技術開発が進んでいることから、学習用データの提供をしやすくする等これらの技術開発を支援することや、技術の共有化を図るといった取組も考えられる。

²¹ (今後のヒアリングの結果、各分野で推奨すべき取組があれば、参考資料の一覧に付す)

(デジタルコンテンツの拡充)

書籍や公文書などの一部分野ではコンテンツのデジタル化が進んでいるものの、全体としてインターネット上で利用できるデジタルコンテンツは欧米と比較すると少ないといわざるを得ない。

デジタルコンテンツを拡充していく際には、アーカイブ機関が活用者のニーズを把握すること、限られたリソースの中で何をデジタル化するかという優先順位をつけることが重要である。また、デジタルリマスターなどの技術も活用することでより品質のよいデジタルコンテンツの作成も可能となってきたことから、そのような最新デジタル技術の活用も検討すべきである。

そして、小規模機関においてもデジタルコンテンツの拡充を図っていくためには、小規模機関がコンテンツのデータ整備に責任を持ち、つなぎ役が大きなプラットフォームを用意することで、役割・負担を分担することが望ましいとの意見もあった。

(地域アーカイブの構築)

地域アーカイブの取組という観点から、現在デジタル化について議論されているが、これは「氷山の一角」にすぎず、地方には多くの資料が眠っているが予算も人もないため、そもそも貴重な資料の保全が難しい状況であるとの指摘や、地方自治体が設置した視聴覚ライブラリーに保管された資料や、市史編さんのために集積された資料などのデジタル化が進んでいないとの指摘があった。地方自治体に対する調査²²においても、デジタルアーカイブ構築に関する方針や計画を有するところはほとんどなく、一部アーカイブ機関において所蔵資料のデジタル化やデータ公開に取り組んでいるものの、来館者の増加など顕著な効果がみられるまでには至っておらず、予算や人員、スキルが不足していることを課題として挙げる回答が多かった。

国や地方自治体、大学、企業、NPO などすべての主体において、地域アーカイブを推進するための基盤となる人材の育成にそれぞれが取り組んでいかなければならない（「(5)人的基盤の整備」参照）。

その上で、地域のデジタルコンテンツを増やしていくためには、地域住民が当事者意識をもち、各アーカイブ機関と地域住民が協働してデジタルアーカイブを構築するプロセスが重要であり、それがその後の利活用の創出にも繋がる。

具体的には、県域・圏域の数か所に、誰でも利用可能なスキャナ等の機器を備えたスタジオやアプリケーションの提供等の支援策を講じる、地域住民自らが知りたい、利用したいと思えるアプリを提供することによって、自動的に地域のデータがアーカイブされる仕組みを行政側で用意するといった取組を推進していくことが必要である。

²² 「地方自治体における知的財産戦略等の策定・活用に関するアンケート」（平成 29 年 9 月知的財産戦略推進事務局実施）

(持続可能なデジタルアーカイブの構築)

これまでに構築されたデジタルアーカイブの中には、保存されることなく消えていったものも多く、これら消えゆくデジタルアーカイブの長期的なアクセス保証も大きな課題の一つである。

予算面では、構築当初からサーバシステムなどのリプレイス・運用保守に係る経費を確保する方法を考える必要がある。その際、基金化・クラウドファンディング利用などを検討し、ビジネスとして成り立つような形を模索することも考えられる。

システム面では、固定 URL は必要ではあるものの無効になれば意味がなくなるため、何十年か後の人たちが使うという意識を持って、情報が長期的に有効であるよう工夫すべきとの意見があった。この点、ガイドラインにおいても言及されているが、長期アクセスを保証する DOI (Digital Object Identifier)²³等の永続的識別子の付与・普及といった取組も考えられる。

また、あるデジタルアーカイブを他機関で引き受けて、従前通り公開していく場合に、著作権者や肖像権者から改めて許諾を得ることが難しく、公開できなくなることがよくあるとの指摘があった。この点については、DOI 等の永続的識別子の付与に加えて、デジタルアーカイブを構築、公開する際に、包括的に権利処理することが望ましいが、いったん公開されたものについて、追加の権利処理を簡便に行えるような仕組みの構築の検討も望まれる。

(3) データの公開・オープン化の促進

データの公開・オープン化、アーカイブ機関側にインセンティブが働かないとの意見がある一方、その存在が、利用者だけではなく、広く認知されることで、コンテンツそのものの価値の向上に寄与するとともに、アーカイブ機関のブランド力の向上に寄与するとの意見もある。実際、アメリカのメトロポリタン美術館やオランダのアムステルダム国立美術館などが、自館や自館所蔵資料の価値向上のために、相当数の所蔵資料の高精細画像を CC0²⁴で公開している。

アーカイブ機関によっては、著作権や肖像権等が問題となって、所蔵コンテンツを全て無料で公開することが難しい場合もあるため、各アーカイブ機関の判断で適切な利用条件を設定できるように全体のシステムを構築していく必要がある。ただし、所在情報だ

²³ コンテンツの電子データに付与される国際的な識別子。ISO 26324:2012 として国際標準化されている。DOI の登録には国際 DOI 財団の認める DOI 登録機関への参加が必要であり、我が国の DOI 登録機関としてはジャパンリンクセンター (JaLC) がある。なお、DOI はデータの管理主体が変わっても識別子自体は変わらないため、この観点からも持続可能なデータの提供を行うことに資するものである。
<https://japanlinkcenter.org>

²⁴ 全ての著作権等の権利を放棄することを意味する。これは、著作権に基づいて訴訟を起こす権利、逸失利益等が出て不法行為に基づく損害賠償を求める訴訟を起こす権利も含めて放棄し、著作者人格権など放棄できない権利については行使しないことを約束するといったことを意味する。

けであっても公開することに価値はあるとの意見もあり、メタデータの整備・公開・オープン化はできる限り対応することは求められる。

地域アーカイブのコンテンツの公開については、プライバシーにも配慮し、公開の範囲を当該地域に限定することについて許諾を得るといった方法も選択肢の一つとして検討すべきである。また、教育研究目的に限定した利用や、公開後一定の年数を経過したものについては、権利処理を要することなく一定の範囲内での利用(特定の施設内での限定公開等)を可能にするなどのゆるやかな合意形成を社会全体で図れないかといった意見もあった。

また、サムネイル/プレビューの作成・公開については、美術や写真の著作物を展示するアーカイブ機関による利用に係る制度改正に向けた準備が進められているところだが、利活用促進のためには、それ以外の著作物を所蔵するアーカイブ機関についても、自館の所蔵資料/収蔵品のサムネイル/プレビューを公開できるような制度が期待されるところの意見があった。

オープン化に関しては、多くのコンテンツホルダーからは、CC0 や CC BY といった条件で公開することにより、どのように利活用されたかが分からなくなるおそれがあるとの指摘があった。しかし、CC0 や CC BY といったクリエイティブ・コモンズ・ライセンスは、著作権の権利情報による利用条件を簡潔に示すもの²⁵であり、典拠表示などを利用条件とすることを否定しているものではないことから、過度な恐れをもって、オープン化を回避することは望ましくないものと考えられる。他方で、こういった恐れに対して、活用者は、適用されているライセンスや利用条件をよく確認し、順守しなければならないのはもちろんのこと、コンテンツ自体の価値をさらに高め、データ提供者にとってもメリットにつながる形で活用することが求められる。その上で、データを集約・提供する機関においては、サイトポリシーなどで、典拠等の情報を明らかにするよう活用者に求めることで、データ提供者の懸念を一層払しょくすることも可能となると考えられるところ、ジャパンサーチ(仮称)においてもそのような対応を進めていくことで、一層のオープン化が図れるのではないかといった意見もあった。

(4) 利活用の促進

(市民参加型のアーカイブ)

デジタルアーカイブの利活用を進めるに当たっては、アーカイブ機関だけでなく、地域住民等が協働で構築するプロセスが重要である。市民参加型のアーカイブ構築を通じて、一人一人の市民がデジタルアーカイブの使い手であると同時に作り手でもあるという

²⁵ なお、クリエイティブ・コモンズ・ライセンス等の著作権ライセンスは、著作権者自ら保有する権利の利用許諾を行うものであるため、著作権保護期間が満了しているコンテンツや、著作物としての要件を満たさない創作性のないデータなどに適用したとしても、その二次利用に著作権法による保護を超えた追加的な制約を課すものではないことに留意が必要。(ガイドライン p.19 より)

意識を持ち、アーカイブの作成と利活用の循環を生み出すことが可能となる。市民に見える形で地域がボトムアップで関与することが地域の活性化にも繋がるほか、アーカイブされたコンテンツについても、様々な情報が付加されて新たな価値が生まれてくる。特に様々な活用事例が周知されることにより、地域コミュニティにおいて無価値とされていた資料の価値について再考され、それが他のコンテンツのメタデータの整備やデジタル化に繋がるといった循環が期待される。

また、例えば、自然史・理工学系のデータの多くは、一点ではあまり価値は無いが、大量に集めることで価値が出るという性質がある。それを一人で集めていくことは不可能であるが、皆で同じ粒度のデータを集めることで価値あるデータを生み出すことが可能となる。実際に、欧米やオーストラリアなどでは、市民科学(シチズンサイエンス)の活動によって、生物・自然観測のデータ収集が活発に進められている。

現在、市民参加型デジタルアーカイブの作成環境を提供するソフトウェアもオープンソースで開発・公開されつつあり、これらをうまく活用することでデジタルアーカイブのエコシステム(循環)を生み出していくことが考えられる。過度なコストをかけずに、適時適切なデジタルアーカイブが構築される、いわばエコシステムといえるような仕組みを作ることができれば、既存アーカイブの長期運用のみならず、新たなデジタルコンテンツの作成や発信、利活用も含めた取組を持続的なものとすることが可能と考えられる。国や地方自治体には、市民参加による地域アーカイブの構築や活用の取組、そのための個々の継続的な活動を後押しすることが求められている。

(具体的な利活用のシナリオ)

デジタルアーカイブの利活用を促すためには、単にデータを集約するだけでなく、その使い方(シナリオ)を具体的に示すことや、利活用しやすいプラットフォームでデータを公開することを検討することが必要である。その際、単にデータを公開するだけでなく、よりユーザに身近でかつ意外性のある見せ方を工夫することが重要である。例えば、江戸料理レシピを料理レシピ検索サイト「クックパッド」で写真付きで公開することで、予想もしなかった大きな反響があったという事例がある。また、民間事業者や団体、大学等においても利活用モデルの模索を行っていることが多く、産学官での協力・連携も検討していく必要がある。

その他、個別分野における利活用のポイントは以下のとおりである。

- ・教育分野での利用を促進するためには、メタデータに学習指導要領に記載されている文言や教科書の見出しを付与するなど、単元との対応関係を明確にし、教員に使ってもらいやすくする工夫をすると、効果的である。
- ・観光分野においては、地域のコンテンツをデジタル化し、観光に生かす取組は既に各地で始まっている。提供が容易なデータから漸次オープン化することによって、観光客の誘客から地方創生に繋げていくことが可能となる。

- ・地域における利用に当たっては、市民を巻き込んでデジタルアーカイブを構築していくプロセスを模索することが必要である。
- ・防災利用の観点からは、災害時だけでなく、平時においても、デジタル情報を共有・利活用し、情報認識の統一を図っておくことが必要である。
- ・その他、デジタル化された歴史資料等を回想法²⁶に利活用することも考えられる。

(今後の活用の可能性)

現在検討が進められている著作権法の見直しにおいて、柔軟な権利制限規定の下で、所在検索サービスや情報分析サービスなどの利用が可能になる方向で議論が進められている。これらのサービスが権利制限で認められることで、デジタルアーカイブの利活用の幅が一層広がることになり、その動きを注視する必要があるとの意見があった。

例えば、文字情報を含む画像のテキストデータ化によって、その検索利便性は向上する。また、ジャパンサーチ(仮称)で集約されるメタデータに含まれるコンテンツの所在情報から大量のデジタルコンテンツにアクセスし、AI がそれらのコンテンツを集約・解析することで、新たなデータを生み出したり、時代や地域を超えた関連性の高いコンテンツの発見をしたりすることも可能となる。こうしたことにより、ストック化(死蔵)していたデジタルコンテンツをフローに戻すこと(活用)がより容易になると考えられる。

(5) 人的基盤の整備

デジタルアーカイブの利活用を進めるためには、自館の所蔵コンテンツに付加価値を見出し、発信するキュレーターや、各地域の事情を分かった上でデジタルアーカイブを構築し、国際標準とすり合わせていくことのできる人材の育成が必要不可欠である。こうした人材には、単なる知識だけでなく、コーディネートや企画、マネジメントに関する能力だけでなく著作権法や個人情報保護法さらには、人権やプライバシーといった権利に対する理解も求められており、一朝一夕で養成できるものではない。

現在、アーカイブ機関の外側にある第三者の専門人材を活用する仕組みや、経験を積んだスタッフを一定程度広域的に活用する仕組みなども含め、多様な方策を検討する必要がある。また、地域のアーカイブにおいては、アーカイブ機関だけでなく、大学や企業、NPO などが連携してその地域固有の課題と一緒に取り組んでいくことも考えられる。こうした取組と AI などの新技術を活用することによって、限られた人的リソースを適切な業務²⁷に振り向けることが可能となり、現場の負担軽減にも役立つと考えられる。

²⁶ 認知症やうつ病のリハビリテーションの一つとして活用する方法で、過去の出来事や思い出を想起したり、語ったりすることで病状の進行を遅らせることを目的とするもの。

²⁷ 従来からの業務だけではなく、デジタルアーカイブの構築・発信などの業務についてもより積極的に行うことが可能となると考えられる。

(6)意識啓発

多くのアーカイブ機関では、来館者数や貸出し数を業績評価における重要な成果指標としており、所蔵品の高精細な写真や動画といったデジタルコンテンツをインターネット上に公開すると、来館者数が減るのではないかという意識があると指摘されている。しかし、デジタル時代においては、むしろ、優れたコンテンツを発信することにより、多くの情報の中に埋没することなく文化的資産を継承・保存・活用していくことが可能となるのであり、所蔵品やアナログ情報を適切に保存するだけでなく、デジタル情報を保存、発信していく重要性について認識を共有することが必要である。実際、海外では、「デジタルアーカイブ」よりも、「デジタルプリザベーション」が用いられることが多く、どのような規模のアーカイブ機関であっても、デジタル化、情報の保存に取り組むことと、可能な範囲でそれらの情報を提供していくことについて意識することが必要となる。

このような認識を周知していくためにも、ガイドラインを効果的な形で普及させ、アーカイブ機関の現場で活用してもらう必要がある。

(7)評価の枠組みの整備

来館者数やアクセス数等の従来型の指標だけでは、デジタルアーカイブの取組が適正に評価されないため、関係省庁等連絡会・実務者協議会では、評価指標に関しても議論が行われた。

本委員会は、デジタルアーカイブの構築・共有・活用のための活動に関して、組織的な取組からシステム面も含めてバランスよく自己点検・評価するための指標・ツールについて検討し、「デジタルアーカイブアセスメントツール」(補足資料参照)を作成した²⁸。このツールは、各組織において求められるレベルが異なることを踏まえ、項目ごとに三段階のレベルを用意した。アーカイブ機関やつなぎ役が、自らのミッションや役割を認識したうえで、必要な項目を選べるようにしている。このツールを使って、自らの立ち位置を確認することが可能となることが考えられる。

標準モデル: 小規模な機関であっても行う必要のある水準

先進モデル: 各アーカイブ機関がそれぞれのミッションや事業に照らして目指す水準

つなぎ役モデル: つなぎ役の役割を持つ機関が目指す水準

今後、このツールの改善の必要性を確認しつつ、本委員会において、引き続き普及に向けた検討を行うとともに、内容の見直しが必要になった場合には適宜検討を行っていく必要がある²⁹。

²⁸ 補足資料「デジタルアーカイブアセスメントツール」参照。

²⁹ 本論点に関する議論の中では、デジタルアーカイブに係る基本法を整備することで、デジタルアーカイブ構築・利活用の推進を図るとともに、アーカイブ機関内での取組促進を促すべきとの意見もあった。

5. 国の分野横断統合ポータル構築

2020 年を目標に、国の分野横断統合ポータルとして「ジャパンサーチ(仮称)」を国立国会図書館が中心となって構築するために、本委員会での検討によって打ち出された方向性と、今後一層の連携促進を図る上での課題をここに示す。

(1) 連携の方針

国の分野横断統合ポータルであるジャパンサーチ(仮称)とアーカイブ機関との連携は、原則として、つなぎ役を通じて連携することが望ましい。ただし、現時点においてメタデータ集約を行うつなぎ役が明確ではない分野³⁰・地域が多い。そこで、それらの分野・地域については、次の条件に当てはまるアーカイブ機関との直接連携を検討することとする。

- ・国の機関であり、当該分野におけるコンテンツを幅広くカバーしているアーカイブ機関
- ・公益に資する目的のため、当該分野におけるコンテンツを幅広くカバーしているアーカイブ機関
- ・唯一性・独自性の高いコンテンツ群を塊として扱う分野・地域を代表するアーカイブ機関
- ・その他(本実務者検討委員会において適当と認められるアーカイブ機関)

(2) つなぎ役の明確化と支援

誰がつなぎ役であるか明確化することは、ジャパンサーチ(仮称)の構築に向けて大変重要な課題であると同時に、各分野・地域におけるデジタルアーカイブの構築・共有・活用の促進のためにも、必要な取組である。

つなぎ役には、ジャパンサーチ(仮称)と各アーカイブ機関を結ぶこと(メタデータの集約)以外にも、次のとおり、メタデータの標準化やオープン化の推進、分野内における人的基盤の構築、法務的な課題への対応など様々な役割が求められている。

- (ア) 分野/地域の独自性を反映したポータルの整備・提供
- (イ) (分野/地域における、以下同)メタデータの集約、API 提供
- (ウ) メタデータの整備推進
- (エ) メタデータの標準化、用語の統制(辞書・典拠・シソーラス)
- (オ) デジタルコンテンツ等の二次利用条件³¹の整備・オープン化の推進
- (カ) 所蔵資料/収蔵品等のデジタル化のための技術や法務上の業務支援

³⁰ 「知的財産推進計画 2017」に記載があるのは、書籍等、文化財、メディア芸術、放送番組のみ。

³¹ 平成 29 年 4 月「デジタルアーカイブの構築・共有・活用ガイドライン」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 15～16 ページにおいて、既に利用条件の基本的な考え方は示されているため、これを前提として議論を進める必要がある。

(キ) コンテンツの長期保存・永続的アクセス保証(データホスト)への協力

(ク) 意識啓発・人材育成

(ケ) 活用促進のための取組

『我が国におけるデジタルアーカイブ推進の方向性』(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、平成 29 年 4 月、p.21)より抜粋

つなぎ役は一つの機関が全てを担う必要はない。行政が中心となって産学官が連携して役割を整備し、つなぎ役を支援していくことが求められる。特に地域のつなぎ役については、地方自治体が主体的に推進することが重要であり、社会教育施設や大学の役割と併せて考えていくという観点も必要との意見があった。

具体的には、つなぎ役は、実務者検討委員会での議論等を踏まえて、毎年作成される「知的財産推進計画」の中で明確化することが考えられる。

(3) 共通メタデータフォーマット³²の策定

ジャパンサーチ(仮称)にデータを提供する機関の作業負担をできる限り軽減しつつ、さまざまな分野の多様なデータの利活用を促進することも可能にするために考えたものが、ジャパンサーチ(仮称)の共通メタデータフォーマットである。

共通メタデータフォーマットには、ジャパンサーチ(仮称)との連携のためのフォーマット(連携フォーマット)と、集約されたメタデータを利活用するためのフォーマット(利活用フォーマット)の二つがある(補足資料「共通メタデータフォーマット」参照。)。全体のデータの流れのイメージは図 3 のとおりである。

³² 「共通メタデータフォーマット」とは、これまで、知的財産推進計画 2017 において「標準メタデータフォーマット」と呼称していたジャパンサーチ(仮称)上で利用するメタデータのフォーマットと同義であり、各分野・地域共通で使うメタデータフォーマットであることを明確にするために、今後は、この呼称を使用することにする。

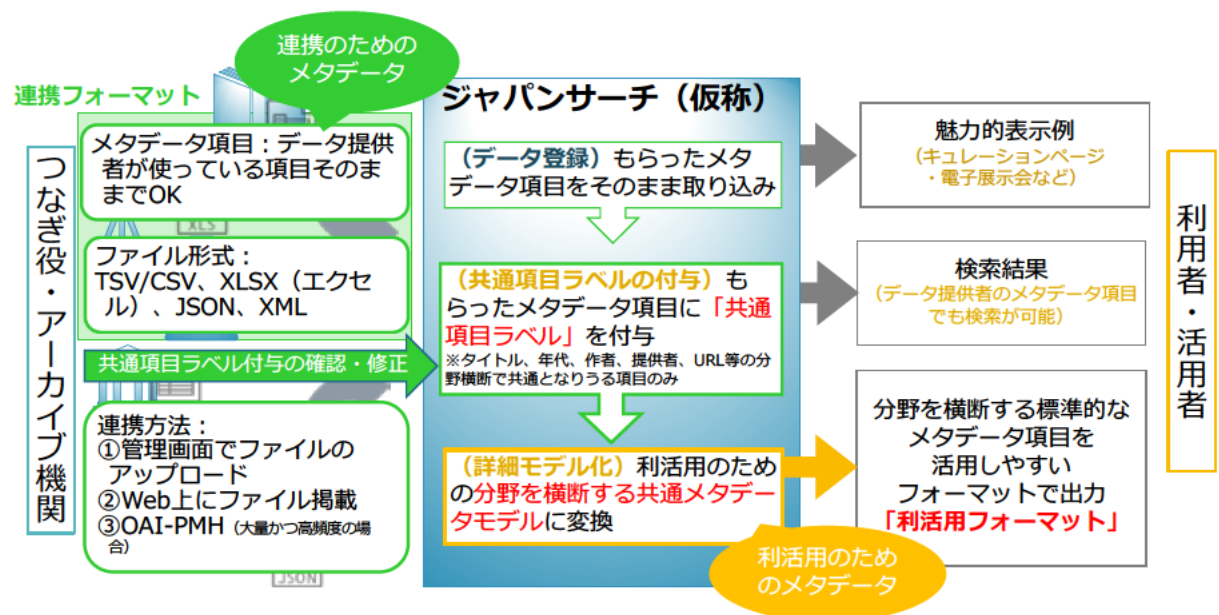


図3 ジャパンサーチ(仮称)におけるメタデータの連携と利活用まで

（連携フォーマット）

多種多様なメタデータ連携フォーマットのメタデータ項目については、各データ提供機関が持っているメタデータ項目をそのままの形でジャパンサーチ(仮称)に提供可能とする。必須項目は管理用のIDと名称/タイトルだけとし、その他は任意とする。ただし、名称/タイトルのヨミや英語名、最終更新日、URL等のデータをもっている場合は、その項目に共通項目ラベル(表1参照)を付与し、それらを同じラベル名で検索することを可能にする。表1に挙げた項目以外の項目は、各アーカイブ機関が示した項目名をそのまま用いることとなる。連携フォーマットについては補足資料を参照のこと。

なお、ファイル形式は、様々な形式に対応する。連携方法は、ジャパンサーチ(仮称)で用意する管理画面からファイルをアップロードする方法や、ウェブ上にファイルを掲載する方法を利用することを主として想定する。

表1: 連携フォーマットの共通項目ラベル一覧

	項目名	内容
必須項目	ID	提供元データ内での一意のID(管理番号)
	名称/タイトル	レコードの名称。検索結果表示に使われる
あれば必須で付与する項目	名称/タイトルヨミ	レコードの名称の読み(カタカナ・平仮名共)
	名称/タイトル英語	レコードの英語名称又はローマ字
	最終更新日	データの最終更新日(日付型)
	URL	レコードのリンク先のURL(つなぎ役と提供元共)

	サムネイル画像 URL	サムネイル画像の URL
	コンテンツ URL	デジタルコンテンツの URL
	提供者	オリジナルのコンテンツの提供者
	コンテンツの権利表示	データベース定義の情報と異なる場合のみ
	寄与者(人物)※	作者、発行者、出演者など
	時間(日付)※	制作年、対象時期など
	場所※	発行地、制作地など

※あれば原則付与するが、分野の特性に応じて付与しないこともありうる。

(利活用フォーマット)

多種多様なメタデータをまとめて利活用するに当たって、シンプルで使いやすいものであると同時に必要十分な情報を提供するためにはどのようなモデルがよいかといった視点から、Europeana 等の海外事例も踏まえて、利活用フォーマットを検討した。

ジャパンサーチ(仮称)では、連携フォーマットによる連携後に、データ提供機関から提供されたメタデータ項目の一部について、ジャパンサーチ(仮称)のシステム側で、分野共通に利用可能なメタデータモデルに変換し、「利活用フォーマット」で出力できる仕組みを用意する。利活用フォーマットの概念整理及び実際のフォーマット(案)については補足資料を参照のこと。

なお、利活用フォーマットの主要項目は表 2 のとおりである。未だ分野の標準的なメタデータフォーマットがない場合に、どのような情報をメタデータとして整備する必要があるかを検討するに当たって、参考になるものと考え、次に示すものである。分野の必要に応じて検討することが望ましい。

表2:利活用フォーマットの主要項目一覧

基本項目	内容
タイプ	コンテンツの基本区分(書籍、文化財など大きく情報を区分する枠)
名称	タイトル、別名、読みなど検索対象とする名前
寄与(者)関係	コンテンツに寄与した人／組織(作者、発行者、出演者など)
場所関係	場所に関する情報(発行地、制作地など)
時間関係	時間に関する情報(制作年、対象時期など)
主題・区分	主題・分類／各分野のキーワード的共通認識のある区分(国宝、ドキュメンタリーなど)
識別子	コンテンツを特定するための識別子(ISBN など)
言語	コンテンツの記述言語
サムネイル画像	コンテンツの特徴を確認するための画像(提供元とは別にサムネイルを保持する場合)

記述	コンテンツの物理的特徴・素材等の記述、個別項目に収録できない情報
上位コンテンツ	当該コンテンツがその一部である上位コンテンツ（公文書などの資料階層）
提供情報	コンテンツにアクセスするための情報。下記サブ項目からなる
提供者	コンテンツ（に関する情報）の提供者を識別する URI。保管者が別であれば URI も。
リンク	コンテンツの紹介ページやアクセス情報が記載されたページの URL
オブジェクト	コンテンツのデジタル画像や音声・動画の URI
権利情報	コンテンツの利用に関するライセンス及び権利に関する情報（ライセンスの URI など）
個別識別子	提供者・所有者が管理するアイテムとしての識別子（請求記号など）
ソース情報	ソースデータ（ジャパンサーチ（仮称）が連携フォーマットで受け取ったデータ）とその提供者に関する情報。
提供者	ソースデータの提供者（つなぎ役）
データ	プラットフォームが保持・提供するソースデータ
リンク	つなぎ役におけるソースデータの掲載ページの URL
更新日	収集元データの更新日又はつなぎ役による収集日

共通メタデータフォーマット全体を通じたメタデータの変遷のイメージは、図4のとおりである。

なお、今回示した資料は、技術的な側面からの記載が中心となっており、普及においては、分かりやすい資料が必要との指摘があった。そういった点も含め、要望に応じて適宜検討を行っていくことが必要である。

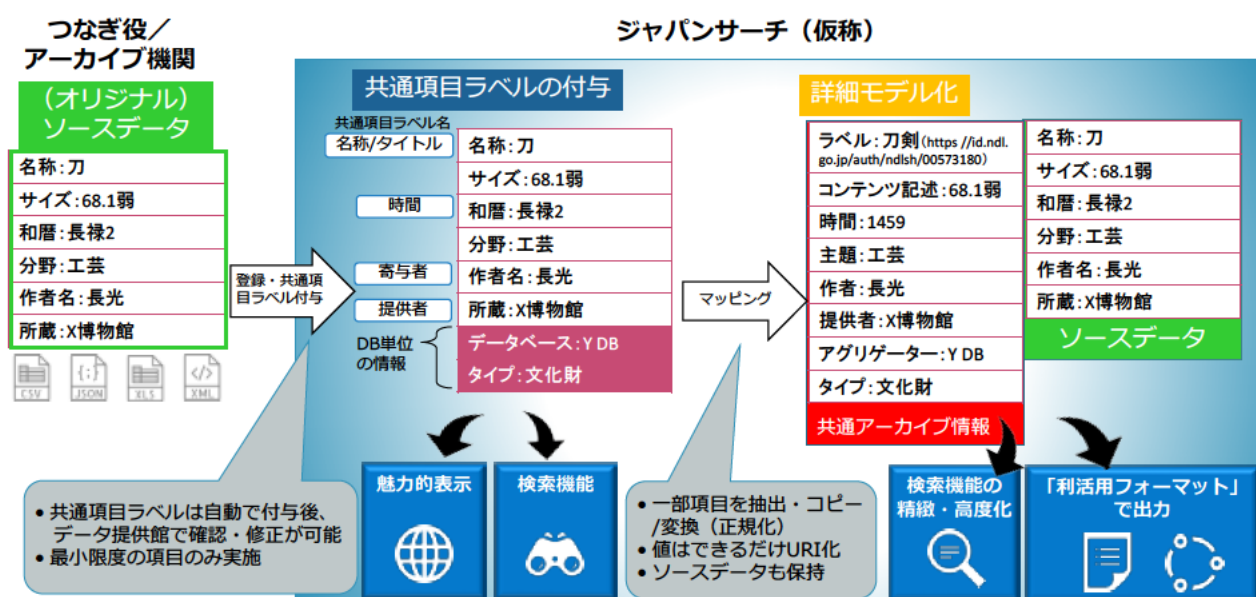


図4 ジャパンサーチ(仮称)における、データ変遷のイメージ

6. 今後の主要検討課題

- (1) 利活用モデルの模索やそのための制度的課題の整理
- (2) 新技術を活用したデジタルアーカイブの構築の在り方の検討
- (3) 長期利用保証の在り方の検討
- (4) ジャパンサーチ(仮称)上での共通メタデータフォーマットを踏まえた、各分野におけるメタデータの在り方の検討
- (5) 各分野・地域におけるつなぎ役の役割や分担の明確化、つなぎ役に対する国の支援策の検討
- (6) メタデータ等のオープン化の実施、コンテンツの二次利用条件表示の促進策の検討(望ましい権利表記の共有等)

おわりに

デジタルアーカイブ社会の実現により、文化の保存・継承のみならず、観光、地方創生、教育研究、ビジネスへの利用など、様々な効果が期待できるが、それに向けた本格的取組は始まったばかりである。

本実務者検討委員会においては、今後も、引き続き、6に掲げた課題を中心に、課題整理や、その解決に向けた対応策について議論を進め、各アーカイブ機関やつなぎ役の取組が円滑に進むよう必要な施策の方向性について検討をしていくこととする。