

デジタルアーカイブの構築・連携のための ガイドライン (案)

平成24年2月15日

目次

はじめに	1
本ガイドラインの構成	2
対象とする読者	3
関係するガイドライン等	3
本ガイドラインで用いる用語	4
第1章 デジタルアーカイブの構築	6
1. デジタルアーカイブとは	6
2. デジタルアーカイブ構築の意義	6
第2章 デジタルアーカイブの連携	10
1. デジタルアーカイブ連携の意義	10
2. デジタルアーカイブ連携の推進：地域内連携支援モデル	12
第3章 デジタルアーカイブの実例	17
1. Museum（博物館・美術館）	19
2. Library（図書館）	25
3. Archives（文書館）	29
4. 実例まとめ	31
第4章 デジタルアーカイブ構築・連携の課題	32
1. 目標設定の課題	32
2. メタデータの課題	32
3. 技術的な課題	33
4. 人材と体制の課題	34

第5章 デジタルアーカイブの構築・連携の手引き	35
1. デジタル化とシステム構築の前に：自館の現状把握と準備	36
2. デジタルアーカイブを作ろう	48
3. デジタルアーカイブを未来に伝えよう：継続・人材・保存	63
参考資料 1 管理データの項目一覧	68
参考資料 2 資料情報の属性一覧	69
参考資料 3 エンコーディングスキーム一覧	71
参考資料 4 文化遺産オンライン 作品情報(文化遺産情報)項目一覧	75
参考資料 5 規格番号	77
参考資料 6 画像コンテンツのパラメータ決定フロー	78
参考資料 7 チェックシート	81
参考資料 8 発注ケース：館内利用者への所蔵資料公開	83

はじめに

近年、我が国においては、さまざまな形態のデジタルアーカイブが存在します。

ここでは、「デジタルアーカイブ」を「図書・出版物、公文書、美術品・博物館・歴史資料等公共的な知的資産をデジタル化し、インターネット上で電子情報として共有・利用できる仕組み」を指すものとします¹。パソコンやスマートフォンなどの情報機器とブロードバンド通信を使用し、博物館・美術館、図書館、文書館などの「知の記録組織」へ実際に行かなくても、インターネットを経由してさまざまな文書や絵・写真などを閲覧できるウェブサイトがその具体例です。

デジタルアーカイブを活用することで、いつでもどこでも、調べごとや学習・研究が行えるようになります。今までは広く公開されていなかった資料をデジタル化して公開することで、貴重な知的資産を誰もが見られるようになります。

また、各地の郷土資料など、特定の場所でしか知られていなかった資料をデジタルアーカイブで公開することにより、広く利用される可能性が高まります。資料の利用が増加するとその価値が広く認識され、ひいてはその資料が根ざす地域の活性化や観光の促進にもつながると考えられます。

しかし、デジタルアーカイブを構築している組織は、我が国に存在する知の記録組織のうち、ごく一部にとどまっています。またせっかく構築しても、その存在が利用者に知られていなかったり、デジタルアーカイブの内容が更新されなかったり、システムが旧式化して技術の進歩に追いついていない事例も見られます。

我が国ではブロードバンド基盤の構築により情報の流通環境は世界最先端の状況にあります。その一方で、上記の通り、デジタルコンテンツの蓄積・二次利用を支えるデジタルアーカイブの構築が遅れています。産業・経済、学術・研究、芸術・スポーツ、趣味・学習、行政等の生産性を向上させていくには、情報の生産・流通・利用・蓄積・二次利用の円滑な二重サイクルを形成していくことが重要です。

我が国においては1990年代から始まったインターネットの発展とともに、多くの組織においてデジタルコンテンツを蓄積し、デジタルアーカイブとして提供する努力が続けられてきました。国立国会図書館や国立公文書館等、我が国を代表する組織では大規模なデジタルアーカイブが構築されています。その一方、公共図書館や博物館、美術館等では90年代末ごろにはデジタルコンテンツ開発が盛んに行われましたが、財政環境の影響もあり、最近では目立った活動を見つげにくくなっています。

しかし、公共的な知的資産を収集保存する組織——「知の記録組織」における保有資源のデジタル化は、貴重な文化遺産に接する機会を国民に広く提供するものであり、そうした組織による直接的なサービスのみならず、デジタル資源の教育利用や観光の促進、地域産業振興への利用等が期待されます。また近年、ネットワークにつながる携帯型の読書端末、スマートフォンなどの登場により、大多数の国民が、デジタルコンテンツに容易にアクセスできる環境を手に入れつつあり、出版流通環境にも大きな変化の予兆が見られます。一般の利用者にとっては、

¹ 総務省「知のデジタルアーカイブに関する研究会開催要綱」2011年2月

新しい電子書籍から、ウェブページ、そして知の記録組織によるデジタルアーカイブのコンテンツまで多様な情報資源など、広い範囲の中から効率よく探し、目的に応じて適切なものを選び、必要に応じて適切な対価を支払って利用できることが望めます。したがって、知の記録組織のデジタルアーカイブ構築者にとっては、自組織のデジタルアーカイブを作り上げる上で、そうした新しい情報環境に適した機能を持たせることが求められています。

そうした組織においては、デジタルアーカイブの構築・連携の方策を具体的に検討するうえで、自組織の現状に即した運用マニュアルを持ち、デジタルアーカイブ構築・連携の方針・戦略を定める必要があります。加えて、各組織の状況に応じたデジタルデータの作成方法や、デジタルアーカイブの設計等に関する技術の詳細にも言及する具体的な運用マニュアルを策定しなければなりません。

しかし、適切な運用マニュアルを独自に策定することは個々の組織にとって容易ではありません。デジタルアーカイブ構築・連携を促進する上では、各組織が運用マニュアルを効率的に作成するために参考となるようなガイドラインが必要です。

これまでも、デジタルアーカイブの構築や運営について、多数の研究や実証実験が行われ、その成果が発表されてきました。しかし、内容が専門的であったり、高度かつ大規模な計画が前提となっているものが多く見られ、地域の知の記録組織にとっては活用が難しいものでした。

このような状況を踏まえ、図書・出版物、公文書、美術品・博物品、歴史資料等公共的な知的資産の総デジタル化を進め、インターネット上で電子情報として共有・利用できる仕組みを構築し、知の地域づくりを推進するため、地域の知の記録組織で活用していただくことを目標に本ガイドラインを作成しました。

本ガイドラインの構成

本ガイドラインは、以下の構成になっています。

「第1章 デジタルアーカイブの構築」では、デジタルアーカイブを構築するにあたっての前提を説明し、続く「第2章 デジタルアーカイブの連携」では、構築したデジタルアーカイブの効果を高める「連携」の考え方について説明します。ここまでで、デジタルアーカイブを構築・連携するための基礎的な知識を知ることができます。

「第3章 デジタルアーカイブの実例」では、第1章・第2章で紹介してきたデジタルアーカイブを実際に構築した例を取り上げます。第1章・第2章での説明の理解を深めるため、また実際にデジタルアーカイブを構築するときの参考とすることができます。

「第4章 デジタルアーカイブ構築・連携の課題」では、デジタルアーカイブを構築する各機関において個別に検討する必要がある点や、現時点で未解決であり、今後の継続した検討が必要な、デジタルアーカイブに関連する課題を示します。

最後の「第5章 デジタルアーカイブ構築・連携の手引き」では、連携可能なデジタルアーカイブを構築する際の手順を示します。

最初から通してお読みいただくと、デジタルアーカイブの基本的な概念から実践のヒントまでを知ることができます。

対象とする読者

本ガイドラインは、博物館・美術館、図書館、文書館などの、知の記録組織で業務に従事する方々に向けて作成されました。各組織の実情にあったデジタルアーカイブ構築の方法を考える上での参考となることを目的としています。また、資料管理システムの構築や知的資産のデジタル化・メタデータスキーマ作成等の標準的な機能要件を示すことで、実際にシステム構築・データ作成を行う上での参考となることも目的としています。特に、現在デジタルアーカイブを構築していない組織や、過去に構築したものの、その効果が実感できていなかったり、運営の方法を決めかねている組織に読んでいただきたい内容になっています。

また、後述する「地域内連携支援モデル」において重要な役割を担う、都道府県立図書館のデジタルアーカイブの担当者の方々にもお読みいただきたいと考えています。

そして、これらの機関からシステム構築等を請け負う企業の方々にも参考にしていただきたいと考えています。

ただし、デジタルアーカイブを構築する知の記録組織毎の事情により、効果的な構築の方法が様々に異なることに留意する必要があります。本ガイドラインを参考にしながら、一部の作業を簡略化あるいは省略したり、逆に量や品質をより高めることには何の支障もありません。

上記に限らず、デジタルアーカイブに登録される知的資産を作成する方々、さらにはデジタルアーカイブに興味がある一般の方々にも広くお読みいただけます。

関係するガイドライン等

本ガイドラインを読む上での前提となる、あるいは参考になるガイドライン等を以下に示します。必要に応じて参照してください。

No	タイトル	発行者	入手先
1	国立国会図書館資料デジタル化の手引 2011年版	国立国会図書館	http://www.ndl.go.jp/jp/aboutus/digitalguide.html
2	電子情報の長期利用保証に関する調査研究(平成14年度～平成16年度、平成18年度～平成22年度)	国立国会図書館	http://www.ndl.go.jp/jp/aboutus/preservation_02.html
3	全国の公文書館等におけるデジタルアーカイブ・システムの標準仕様書(平成21年3月)	国立公文書館	http://www.archives.go.jp/law/pdf/da_100118.pdf
4	メタデータ情報共有のためのガイドライン	メタデータ情報基盤構築事業	http://www.meta-proj.jp/A03.pdf

5	デジタルアーカイブ 権利と契約の手引き 契約文例 + Q&A 集	デジタルアーカイブ推進協議会	http://www.dcaj.org/jdaa/public/index.html
6	文化資源のデジタル化に関するハンドブック	東京大学大学院情報学環ほか	http://www.center.iii.u-tokyo.ac.jp/handbook

本ガイドラインで用いる用語

本ガイドラインで使用する主な用語について、以下に示します。

用語	説明
アーカイブ	古文書・公文書・記録文書の集合、もしくはその保管所、文書館を指す言葉です。 「アーカイブズ」や「アーカイブス」と表記されることもあります。
デジタルアーカイブ	本ガイドラインでは、図書・出版物、公文書、美術品・博物品・歴史資料等公共的な知的資産をデジタル化し、インターネット上で電子情報として共有・利用できる仕組みを指します。
メタデータ	「データに関する（構造化された）データ」。情報資源を組織化し、同定・識別や検索、管理・保存を行うために必要なデータの総称で、あるデータそのものではなく、そのデータに関連する情報です。データの作成日時・作成者・データ形式・タイトル・注釈、またそれらをまとめた目録・索引・抄録等の二次情報などが考えられます。データを効率的に管理したり検索したりするために重要な情報です。 ²
コンテンツ	内容、中身という意味の英単語です。媒体により記録・伝送され、人間が観賞するひとまとまりの情報、すなわち、映像や画像、音楽、文章、あるいはそれらの組み合わせを意味することが多くあります。具体的には、ニュース、小説、映画、テレビ番組、歌、ビデオゲーム、マンガ、アニメなどを指して使われます。 ³ 本ガイドラインの中では、知の記録組織が持つ資料を指します。

² 国立国会図書館『NDL デジタルアーカイブシステム・メタデータスキーマガイドライン』2007年、p.141。
IT用語辞典 e-Words : <http://e-words.jp/w/E383A1E382BFE38387E383BCE382BF.html> (平成 23 年 12 月 29 日確認) を参考にした。

³ IT用語辞典 e-Words : <http://e-words.jp/w/E382B3E383B3E38386E383B3E38384.html> (平成 23 年 12 月 29 日確認) を参考にした。

知の記録組織	本ガイドラインの中では、図書館・美術館・博物館・文書館等・大学・研究機関・国・地方自治体等、公共的な知的資産を収集保存する組織を総称して「知の記録組織 (Memory Organization)」と呼びます。
スキーマ	データベースの用語として古くから使われており、「情報の構造を定義するための記述」を指します。本ガイドラインの中では、項目の名前 (記述項目) や、項目が文字なのか数字なのか (記述形式)、必須、推奨、省略可能な構造的な制約などを定義した記述のことを指します。

第1章 デジタルアーカイブの構築

1. デジタルアーカイブとは

デジタルアーカイブという言葉は、平成8年に設立された「デジタルアーカイブ推進協議会（JDAA）」の準備会議の中で月尾嘉男氏（東京大学教授(当時)。平成14年～15年 総務省総務審議官、現在、東京大学名誉教授）から提案され、広報誌「デジタルアーカイブ」で初めて公表されました。デジタルアーカイブ推進協議会において、その概念は、「有形・無形の文化資産をデジタル情報の形で記録し、その情報をデータベース化して保管し、随時閲覧・鑑賞、情報ネットワークを利用して情報発信」というデジタルアーカイブ構想としてまとめられました。ここでは、図書・出版物、公文書、美術品・博物館・歴史資料等公共的な知的資産をデジタル化し、インターネット上で電子情報として共有・利用できる仕組みをデジタルアーカイブと呼びます。

2. デジタルアーカイブ構築の意義

(1) なぜデジタルアーカイブをつくるのか

デジタルアーカイブの基本的なメリットは、「誰でも、いつでも、どこからでも、有用な知的資産にアクセスできること」です。

知的資産は、「収集、利用・創造、公開、共有」が循環していくことが重要です。情報の価値は利用されて初めて生じます。利用されるためには適切に公開される必要があります。そして公開されたものを共有できる仕組みが、社会・文化の基盤となります。

知的資産には、フローとして流通していく側面と、図書館、博物館・美術館、公文書館（MLA機関⁴）を中心に蓄積・保存されるストックの両面があります。新たな知的資産の生産・流通を活発化するには、フローのための制度等の整備とあわせて、生産の基盤となるストックの部分の整備が欠かせません。

情報環境の変化、特に携帯型情報端末の普及により、どこにいてもコンテンツにアクセスできる環境が整いつつあります。こうしたなかで、デジタルアーカイブは、デジタルコンテンツとなった知的資産を効率よく利用者に届けることができます。

これまでの、モノやアナログ媒体の形で蓄積されてきた知的資産をデジタルに移行させることと、新たにデジタルで生まれてくる知的資産を最終的にストックして、次世代まで活用できるように蓄積・保存することの両面が、デジタルアーカイブの使命であり、デジタルアーカイブを構築する意義であるといえます。

しかし、デジタルアーカイブのメリットは、単に原資源をデジタル化してインターネット上に載せさえすれば得られるというものではありません。使いやすいインタフェースを介して、有用な知的資産を探し、閲覧し、そしてその知的資産を利活用することが利用者にとって容易に行えるものでなければなりません。

⁴ 博物館・美術館（Museum）、図書館（Library）、文書館（Archives）のそれぞれの頭文字をとった呼び方。

(2) デジタルアーカイブの構築によって何が得られるか

デジタルアーカイブを構築するMLA機関にとって考えられるメリットを示します。以下に館種ごとに代表的な例を挙げますが、これらとは別の目的や利点・利益を独自に検討し、デジタルアーカイブを構築することを制限するものではありません。

(a) 全ての館種にとって

どの組織にとっても、自組織の活動成果の普及・公開、資料の継続的保存・管理、資料の検索性の向上、広報活動の効果がります。地域の誇りである文化財は全ての館種に存在します。それらを連携して横断的に見せる「場」としてデジタルアーカイブが有効であり、地域起こし、地域復興、地域の絆に貢献します。

また、文化財とデジタル文化財が補完関係を保ち、素材としての文化財情報から新たな価値を創造することがデジタル化の役割と考えられます。

(b) 博物館・美術館にとって

博物館・美術館においては、展示室・展示環境の制約を取り払うことができる効果が重要であると考えられます。

例えば対象が細かく描写してある大型の資料のように、ガラス越しでは資料を詳しく見ることができないような資料については、超高精細デジタル資料を作成することで、より詳細に見ることができます。さらに、スペースの制約、展示期間の制限などで、展示室に置けない資料を見てもらうことができます。

美術館の場合、博物館に比べて著作権の保護対象となる資産が多く、すべてをインターネットに公開することが難しいことから、デジタルアーカイブ構築の効果が理解されにくいことが考えられます。しかし、所蔵目録の整理と情報公開が可能になることや、館内展示を目的とするデジタル高精細画像とPRを目的とした低画質の画像を使い分けるなど、デジタルのメリットを生かした活用法が考えられ、画像等の公開が限定的であっても、デジタルアーカイブ構築のメリットを享受することは十分に可能です。

デジタルアーカイブ構築・連携のメリットを尋ねたヒアリング調査⁵では沖縄県立博物館、ひめゆり平和祈念資料館、那覇市立壺屋焼物博物館等の博物館からのコメントとして、以下のメリットがあると報告されています。

- ・ 研究利用・メディア等の二次利用で、「この資料」と特定されるので、確認や連絡の負担が軽くなった。
- ・ 展示物の有無だけを聞くようなレファレンスは、ウェブで見てもらうことによって減少した。
- ・ 展示されていない資料を見せることができる。

⁵ 平成22年度総務省「我が国におけるデジタルアーカイブの構築・統合に関する調査研究報告書」知の記録組織へのヒアリング調査より

(c) 図書館にとって

公共図書館においては、資料のデジタル化とデジタルアーカイブの構築により、

- ・ 歴史的資料（そのままでは提供できない資料）の提供サービス
- ・ 失われつつある地域の知的資産の確保
- ・ アナログ情報へのアクセス（索引的データベース）の援助

などの効果が得られます。

大学図書館ではこれに加えて、活動成果の普及・公開を行うことができることも重要な点です。

デジタルアーカイブ構築・連携のメリットを尋ねたヒアリング調査⁶では沖縄県立図書館、沖縄県立図書館、青森県立図書館等の図書館からのコメントとして、以下のメリットがあると報告されています。

- ・ 自館にない資料を他館に紹介するときに便利。横断検索の利用者は図書館組織が多い。
- ・ レファレンス機能の強化という点で評価。資料の有無だけを聞くような、レファレンスを軽減している

(d) 文書館にとって

文書館においては、国内外を問わずいつでも所蔵資料を利用できる環境の整備・充実を図ることができます。また、電子的に作成・保存される公文書等の「保存」と「利用」への早急な対応は、文書館全体にとっての使命に関わる新たな要請であり、デジタルアーカイブの構築がその一助になることが考えられます。

デジタルアーカイブ構築・連携のメリットを尋ねたヒアリング調査⁷では北谷町公文書館、沖縄県公文書館等の文書館からのコメントとして、以下のメリットがあると報告されています。

- ・ 館の広報としての意味がある。来なければどんな資料がわからない、というのと、来なくてもどんな資料があるか見てもらえる、というのと大きく違う。

(3) デジタルアーカイブのコンテンツ

知の記録組織が提供するデジタルコンテンツの多くは、歴史的、文化的資料です。もともと知の記録組織が持つ資料がそうした性質を持つものであるということも理由のひとつではありますが、デジタル化の際に著作権上の問題が少ないものが選ばれているようです。また、大学等の研究機関が提供するデジタルコンテンツは、論文等の学術的資料が中心です。

一方、こうした歴史的・文化的・学術的資料は、一部の専門家にとってのみ利用価値があるのではないか、さらには、専門家であればデジタルアーカイブでなくても原情報への到達手段を持っているのではないか、といった批判があることも理解した上でデジタルア

ーカイブの構築を進めていかなければなりません。

そのため、コンテンツの従来の利用者にとってはより利用性を高め、そして従来は利用してこなかった潜在的利用者にとってもより魅力的なものにしなければなりません。

1990年代からの爆発的なネットワーク環境の進化にしたがって、もともと電子的に作られ発信・出版される、電子書籍等のコンテンツが爆発的に増加しています。利用者の視点からは、こうした電子書籍等もデジタルアーカイブのコンテンツも、どちらも重要な知的資産であり、ネットワーク上で有機的に組み合わせて利用できるようにすることが望まれます。

⁶ 平成 22 年度総務省「我が国におけるデジタルアーカイブの構築・統合に関する調査研究報告書」知の記録組織へのヒアリング調査より

⁷ 平成 22 年度総務省「我が国におけるデジタルアーカイブの構築・統合に関する調査研究報告書」知の記録組織へのヒアリング調査より

第2章 デジタルアーカイブの連携

1. デジタルアーカイブ連携の意義

(1) デジタルアーカイブの連携を考えるにあたって

デジタルアーカイブは、図書館・美術館・博物館・文書館等・大学・研究機関・国・地方自治体等の様々な知の記録組織で構築されています。したがって、デジタルアーカイブの内容はコンテンツを提供する組織に応じて様々であり、管理方法、提供方法ともに多様です。そのため、デジタルアーカイブの連携とは言っても、既存のデータベースを丸ごと統合するといったことを考えることは現実的ではなく、また、個別のアーカイブの持つ特色を失うような連携方法を考えることも適切ではありません。デジタルアーカイブの意義や重要性、基本的な考え方については前章で説明したとおりですので、以下では、現実的かつ将来指向の「デジタルアーカイブの連携」とは何かについて、読者の理解を助けるための基礎的視点を示します。

(2) デジタルアーカイブの連携の必要性

1990年代より先進各国の図書館、博物館・美術館等の組織を中心として、ネットワーク時代における知的資産のより高度で広範囲の利用を目的として、各組織が所蔵する知的資産をデジタル化し提供する、デジタルアーカイブの構築が進められてきました。多くの市民が知的資産に接する機会を増やし、教育機関等での教育・学習のための知的資産活用を進め、そして知的資産を利用した産業を育成する一方、貴重な実物の知的資産へのアクセスを限定することで、知的資産の劣化を防ぎ、適切な保存環境を維持することを目的としています。我が国では、国立国会図書館における大規模な図書館のデジタル化、国立博物館を中心とする高品位な資料のデジタル化、国立公文書館による歴史公文書のデジタル化等が進められてきています。

このように、単独でのデジタルアーカイブの構築は進展していますが、目を世界に転じると、ヨーロッパにおける Europeana⁸、アメリカ議会図書館を中心とする World Digital Library⁹、台湾の National Digital Archive Program¹⁰ など、図書館、博物館・美術館、文書館等が持つデジタル資料を連携して提供しているプロジェクトがあります。これに対して、我が国においては国立国会図書館の国立国会図書館サーチ(NDL Search)¹¹、国立情報学研究所による学術機関リポジトリ連携支援事業¹²のように、他組織のアーカイブからメタデータを収集することでデジタル資料を提供するサービスはまだ限定的です。

一般的に、インターネット上のコンテンツは、多様なサービスによって独自に提供されていて、すべてが整理されているわけではありません。そのため、必要な情報を効率よく入手するためには、「ポータルサイト¹³」を入口として利用します。同様に、デジタルアーカイブ

ブのコンテンツ利用に関しても、利用者にとって使いやすい入り口から入って、いろいろなサービスを統合的に利用できるようにする必要があります。すなわち、利用者の目から見ていくつものデジタルアーカイブがひとつのサービスとして統合されているように見えることが望まれます。それにはデジタルアーカイブ間の連携が必要です。

(3) デジタルアーカイブ連携の効果

デジタルアーカイブを構築し、それらを相互連携させることにより、以下の効果が得られると考えられます。

- ①形態や領域に関わらず日本の知的資産をデジタル化し、統合的な検索を実現し、それらの情報の流通性を高めることで、知的資産の保存が図れるのみならず、従来では想像できなかったような新しい利用方法を生み出すことができます。
- ②標準化された仕様に基づいてデジタルアーカイブを構築し、デジタルアーカイブへのアクセス方式(API¹⁴)を公開することで、構築されたデジタルアーカイブ間の連携が可能になり、国民にとって知的資産へのアクセス性が高まるのみならず、知的資産の活用を行うための新たなビジネスが創出されます。
- ③日本の知的資産を世界からアクセスしやすくすることで、文化教養面における知的資産利用を促進し、日本文化や研究の国際的発信に寄与します。また、国際環境下において日本の高等教育や観光産業などの振興を図ることができます。

(4) デジタルアーカイブ連携の実現に向けて

デジタルアーカイブの連携を実現するには、デジタルアーカイブの管理運営方針から、コンテンツの組織化方法や連携方法とそれを実現するためのメタデータ、そしてデジタルアーカイブの核となるデータベースの構成やデジタルデータの標準フォーマットといった技術的問題まで、いくつかの異なる視点から、デジタルアーカイブの連携、そして統合的利用環境の実現のための情報共有や標準化の方式について検討しなければなりません。そのため、連携可能性をベースにしたデジタルアーカイブ構築により、デジタルアーカイブの連携と統合的利用環境を可能にします。

これまでに、博物館・美術館、図書館、文書館等の連携（以下 MLA 連携）の議論と実践には多くの蓄積があり、国立国会図書館サーチ・文化遺産オンライン等、国による知的資産の共有基盤が既に整備されています。しかし、国立国会図書館サーチ・文化遺産オンライン等では、全国の組織を十分に網羅できていない状況にあります。

公共的な知的資産は、あまねく存在するものであり、多くの組織を網羅する必要があります。特に、災害による知的資産の消失へのバックアップとしての役割も考えるのであれば、できるだけ多くの組織の参加を促し、欠落が最小限になるような仕組みを考える必要があります。

⁸ Europeana <http://www.europeana.eu/portal/>

⁹ World Digital Library <http://www.wdl.org/en/>

¹⁰ National Digital Archives Program, Taiwan http://www.ndap.org.tw/index_en.php

¹¹ 国立国会図書館サーチ(NDL Search) <http://iss.ndl.go.jp/>

¹² 学術機関リポジトリ連携支援事業 <http://www.nii.ac.jp/irp/>

¹³ 利用者がインターネット上で必要とする機能やサービス、コンテンツ、ウェブサイトへのリンクなどを総合的

に案内する、玄関口の役割を持った Web サイトのこと。(IT用語辞典バイナリ: <http://www.sophia-it.com/content/%E3%83%9D%E3%83%BC%E3%82%BF%E3%83%AB%E3%82%B5%E3%82%A4%E3%83%88> (平成 24 年 2 月 12 日確認)を参考に記述した。)

¹⁴ Application Program Interface の略称 あるいはプラットフォーム向けのソフトウェアを開発する際に使用できる命令や関数の集合のこと (IT用語辞典 e-Words: <http://e-words.jp/w/API.html> (平成 24 年 1 月 20 日確認)を参考に記述した。)

こうした状況を踏まえ、本ガイドラインでは、「デジタルアーカイブの連携へのハードルを下げながら、全体的なデジタルアーカイブ構築のレベルの底上げを図る仕組みを作ること」を主眼としています。

メモ：MLA 連携とは？

博物館・美術館 (Museum)、図書館 (Library)、文書館 (Archives) の間の連携は、それぞれの頭文字をとって MLA 連携と呼ばれます。
これら3つの機関は、知的資産を扱っているという点では共通しているものの、これまでには保有する資産の性質や管理手法が異なるため、個別の機関として発展してきました。しかし、デジタルアーカイブの世界では原資料がデジタル化されるため、その区分があいまいになります。これを機に各機関が連携し、利用者に対して一体となってサービスを提供しようという機運が高まっています。

2. デジタルアーカイブ連携の推進：地域内連携支援モデル

(1) 「地域内連携支援モデル」の背景

多くの知の記録組織では、何らかのメタデータを整備していながら、様々な理由から、公開していない、また公開していても国立国会図書館サーチ等の共有基盤に連携していません。デジタルアーカイブの全体的な底上げを図る上では、まず、これら多くの公開・連携されていないメタデータを、可能な限り共有基盤上に連携してもらうための方策を講じる必要があります。

知的資産のメタデータを連携していく上で、以下のような課題があります。

- ①各組織のデジタルアーカイブ構築に関する現状に大きな差異がある
- ②全国レベルの共有基盤につながるメリットが十分認識されていない
- ③各組織独自のメタデータから共有可能なメタデータへの変換が難しい

以上の課題を解決しながら、現時点で、より多く、またより網羅的にメタデータを公開・連携し得るモデルの一つとして「地域内連携支援モデル」を提案します。

(2) 概要

都道府県レベルの機関が、個々の機関と全国レベルの共有基盤との連携について支援を行うモデルを、「地域内連携支援モデル」とします。

現在、メタデータを公開・連携していない組織では、直接国立国会図書館サーチ等の全国レベルの共有基盤と連携することに困難があると考えられます。

全国レベルでの連携は、「誰がどのように利用するのか」という具体像が見えにくく、メリットが感じられない一方、「どのように使われるかわからない」という不安がある、との声が聞かれました¹⁵。

しかし、仮にデジタルアーカイブの連携に、都道府県レベルの機関が協力してくれるならば、以下のようなメリットが見えてきます。

(a) 網羅性がある

- ・ 境界が明確で、遺漏がない
- ・ 都道府県立組織として図書館/博物館/文書館等が概ね設置されている

(b) 知的資産共有のメリットがわかりやすい

- ・ 図書館/博物館では、資料の相互貸借が頻繁に行われている単位である
- ・ 「地域」という同じ主題を扱う組織が多い
- ・ すでに都道府県内図書館の総合目録という形で運用されている
→他組織資料の発見や相互貸借を容易にするツールとしてわかりやすいメリットがある

(c) 組織間の連絡がとりやすい

- ・ 図書館は寄贈依頼や市町村立図書館への貸出、博物館は都道府県博物館協会、文書館等・文書管理部署は公文書管理法への対応等の研修などを通じて、すでに連絡体制が築かれている
- ・ 個々の組織の担当者が互いに顔の見える関係にある
→データの更新・メンテナンスが継続されやすい
→相談や事例の紹介等を通じて、「どう利用されるかわからない」といった不安をある程度解消できる

例えば都道府県立図書館では、都道府県内の連携をとりまとめる以下のような役割が期待できます。

(a) 網羅性が相対的に高い

- ・ 全ての都道府県に存在する
- ・ 地域資料収集や市町村立図書館への貸出を通じて県内各市町村と網羅的・定期的・継続的な連絡がある

(b) 既存のシステムを利用できる

- ・ ほぼ全ての都道府県で既に資料検索システム・横断検索システムを持っている

(c) 組織のミッションと合致する

- ・ 主題にとらわれず地域内のあらゆる知的資産を扱う組織のミッションがある
- ・ 他組織の支援、他組織との連携が組織のミッションになり得る

(d) 継続性が相対的に期待できる

- ・ 相対的に人材（特に、資料の管理・整理にあたる専門的人材）が豊かである

都道府県立レベルの機関が「県内連携の窓口」かつ「全国レベルの統合基盤への窓口」として、ノウハウの中間的集約及び各館の支援を行うことによって、現状よりも網羅的に、より多くの組織が連携基盤に参加できると考えられます。

¹⁵ 平成 22 年度総務省「我が国におけるデジタルアーカイブの構築・統合に関する調査研究報告書」p.109「知の

記録組織へのアンケート調査」設問 3-5 より

なお、ここまでの説明では「都道府県立レベルの機関」と記述しましたが、都道府県立の機関だけがノウハウの中間的集約及び各館の支援の役割を担えるというわけではありません。各地域の事情によっては、例えば県立の機関よりも市立の機関のほうが適している場合があります。ノウハウの蓄積状況や人材配置を考慮し、その都度最適な役割分担を検討することが重要といえます。

(3) 方針

地域内連携支援モデルの基本方針を以下に記します。

- ・ 規模や目的、扱う資料、地域性が大きく異なる各組織の実情に合わせた柔軟性をもたせる
- ・ 地域の網羅性と主題の多様性を高めるために、単に知的資産の数だけでなく、参加する組織の数がなるべく増えるよう配慮する
- ・ 可能な限り既存の資源を利用し、低コストかつ持続可能な仕組みとする
- ・ 質の向上や新たなサービスの展開を一挙に進めることは難しいため、まずは、既存のメタデータを、より多く、公開・連携の基盤にのせていくことを優先する

(4) 対象

以下の機関を対象とします。

- (a) 資料のメタデータを整備してはいるがウェブ上で公開していない機関
一部図書館の特殊コレクション・多くの博物館/文書館等/行政情報センターなど
- (b) 資料のメタデータを体系的に整備していない機関
一部の博物館/文書館等/行政情報センター/研究機関

(5) 実施手順

「地域内連携支援モデル」は、具体的な方式やシステムを示しているわけではありませんので、さまざまな実践の形が考えられます。

ここではその取組の一つとして行われた、秋田県立図書館と秋田県の4つの「知の記録組織」による実証実験¹⁶の概要を紹介します。

(a) 実証実験の目的

- ① 各館で異なるメタデータ（目録・書誌情報）の項目自体は統一せず、意味の対応付けを行うことで連携が実施できることを実証する。
- ② 図書館、文書館、美術館、博物館等複数のデジタルアーカイブ間におけるメタデータの相互利用に関する技術について、平成22年度新ICT活用サービス創出支援事業の成果である「メタデータ情報基盤」システムを活用した実証環境を構築して検証する。
- ③ 中小規模館におけるデジタルアーカイブ構築・連携に係る課題を整理する。

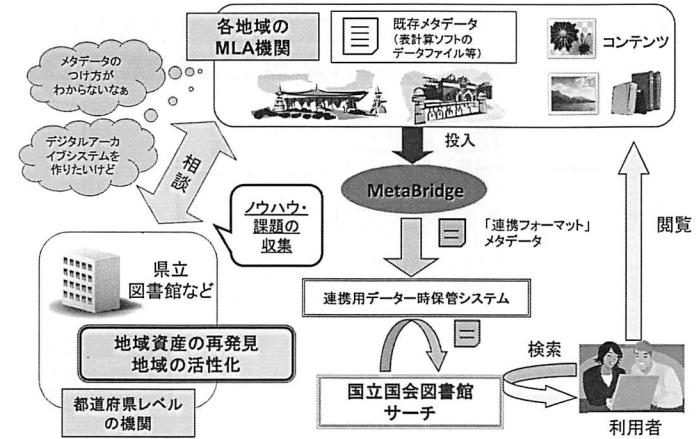


図 2-1 地域内連携支援モデル 実証実験概念図

(b) 実証実験の流れ

【参加館】

- ・ 目録データ（メタデータ）の整理・提供
- ・ 目録データの項目定義をメタデータ情報基盤(MetaBridge)¹⁷へ登録

【都道府県立図書館】

- ・ 参加館への業務的支援（業務的な質問への回答、ノウハウの収集・伝承）

【事務局】

- ・ メタデータ変換サポート（メタデータ情報基盤への登録作業に関して技術的な支援）
- ・ メタデータ情報基盤からのデータ変換システム構築・提供

【中央機関】

- ・ ポータルサイト(国立国会図書館サーチ¹⁸)への連携データ登録

図 2-2 に、実証実験の実施の流れを示します。

¹⁷ メタデータ情報基盤構築事業で構築されたメタデータレジストリ"Meta Bridge"
<http://www.metabridge.jp/infolib/metabridge/menu/>（平成24年1月20日確認）

¹⁸ 国立国会図書館サーチ(NDL Search) <http://iss.ndl.go.jp/>（平成23年12月29日確認）

¹⁶ 平成23年度総務省「我が国におけるデジタルアーカイブの構築・連携に関する調査研究」より

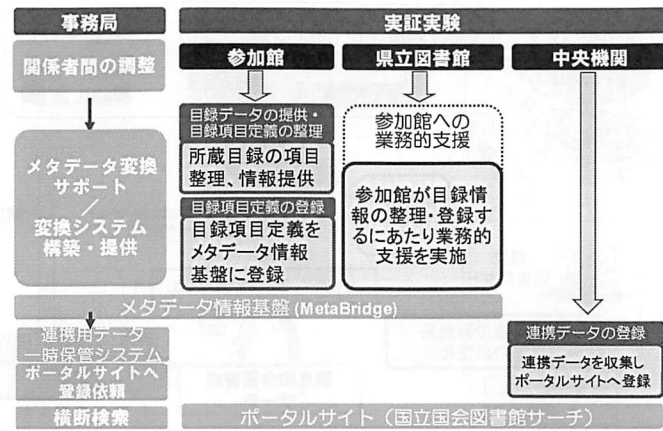


図 2-2 実証実験 実施の流れ

(c) 実証実験の成果

- ① 参加館である、博物館・美術館・文学資料館・文書館の4つの館種でそれぞれの目録項目定義を作成できた。項目の統一をせずに、Dublin Core¹⁹への連携項目に対して意味の対応付けを行うことで、国立国会図書館サーチへの連携を可能とする状況となった。
- ② 各館で作成した目録項目定義をMetaBridgeに登録することにより、それぞれの館で作った収蔵目録データから連携用のデータへと変換が可能となることが実証できた。
- ③ 「地域内連携支援モデル」を用いた取組を行う上での課題を整理した。
 - ・コミュニケーション・人材の課題
メタデータという言葉一つ取っても「言葉の意味がわからない」等の意見があった。館種ごとの専門用語などもあるため、言葉一つ取っても注意・配慮が必要である。また、推進役となる組織が人材や体制を整えておくことも必要である。
 - ・現場職員の作業負荷の課題
システムの操作を館職員の方に行ってもらう場合には、メニュー構成の複雑さや専門用語の難しさが問題となる。今後の継続のためには、メニューの簡略化や言い換えなど、できるだけ作業負荷を軽減する工夫が必要である。
 - ・システム機能面での課題
今回はMetaBridgeと国立国会図書館サーチの間に、連携用データの一時保管システムを事務局にて準備した。より広い範囲の利活用のため、MetaBridgeの周辺サービスを増やす推進活動も必要である。

¹⁹ ダブリンコアと読む。1995年にウェブ上のリソースの発見とメタデータの相互運用性を目的として提案され、その後標準化が進められたメタデータスキーマ。15の記述要素集合と55の記述要素からなる、シンプルなメタデータスキーマである。

第3章 デジタルアーカイブの実例

本章では、すでに構築されているデジタルアーカイブの実例をいくつか取り上げながら、これからデジタルアーカイブを構築しようとされている各地域の Museum(博物館・美術館)、Library(図書館)、Archives(文書館)にとって参考となる取組を紹介します。

デジタルアーカイブの構築・運営が最も進んでいるのは、国レベルの大規模な機関です。

・Museum(博物館・美術館)

東京国立博物館では、日常の収蔵品管理業務の中にシステムを組み込むことで、写真撮影やデジタル化及びシステムへの入力が自然に行えるような業務フローを構築し、デジタルアーカイブの充実に役立っています。また、独立行政法人国立文化財機構では、東京国立博物館の他全4館の国立博物館の収蔵品を鑑賞できる「e 国宝」²⁰を公開しています。

独立行政法人国立美術館では、東京国立近代美術館をはじめとする4つの国立美術館の収蔵品を、リーフレットをめくるようなユーザ・インターフェースで気軽に鑑賞できるデジタルアーカイブ「遊歩館」²¹を公開しています。また、「所蔵作品総合目録検索システム」²²では、4館の所蔵目録を詳細に検索することができます。

・Library(図書館)

国立国会図書館では、(1)デジタルアーカイブの構築 (2)情報資源に関する情報の充実 (3)デジタルアーカイブのポータル機能の充実に目標を掲げ、電子図書館サービスを行っています。貴重資料を電子化して公開している「近代デジタルライブラリー」²³、「国立国会図書館のデジタル化資料」²⁴や、国内に多数存在するデジタルアーカイブ等を一括検索するポータルサイト「国立国会図書館サーチ」²⁵などを構築・運営しています。

・Archives(文書館)

国立公文書館では、平成17年4月からインターネットを通じて「いつでも、どこでも、だれでも、自由に、無料で」、館所蔵の歴史公文書等の目録情報の検索、デジタル画像の閲覧が可能で「国立公文書館デジタルアーカイブ」²⁶を公開しており、平成22年3月1日からは、より分かりやすく、使いやすいデジタルアーカイブとしてリニューアルして運用しています。

上記のような大規模デジタルアーカイブの取り組みに対し、地域のデジタルアーカイブ構築では、その構築・運営のために使える人材や予算などの資源が非常に限られている場合がほとんどです。

そのためこの章では、市区町村あるいは都道府県レベルでのデジタルアーカイブの実例を詳

²⁰ e 国宝・国立博物館所蔵 国宝・重要文化財 <http://www.emuseum.jp/> (平成23年12月29日確認)

²¹ 国立美術館 IAINMoA 遊歩館 <http://search.artmuseums.go.jp/yuuuhokan/> (平成23年12月29日確認)

²² 独立行政法人国立美術館 所蔵作品検索 <http://search.artmuseums.go.jp/> (平成23年12月29日確認)

²³ 近代デジタルライブラリー | 国立国会図書館 <http://kindai.ndl.go.jp/> (平成23年12月29日確認)

²⁴ 国立国会図書館のデジタル化資料 <http://dl.ndl.go.jp/> (平成23年12月29日確認)

²⁵ 国立国会図書館サーチ (NDL Search) <http://iiss.ndl.go.jp/> (平成23年12月29日確認)

²⁶ 国立公文書館「デジタルアーカイブ」<http://www.digital.archives.go.jp/> (平成23年12月29日確認)

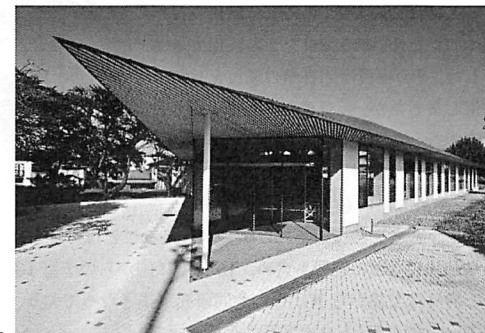
しく取り上げその特長を説明するとともに、限られた資源の中での構築・運営の工夫などを紹介します。

1. Museum（博物館・美術館）

(1) まちとしょテラソ（小布施正倉・小布施人）

- はじめに

平成 21 年 7 月に町立図書館が「まちとしょテラソ」としてリニューアルオープンされました。これまで親しまれた町の図書館であること、待合せの場という意味を込めた「まちとしょ」そして「世の中を照らしだす場」「小布施から世界を照らそう」などの「照らそう」から「テラソ」を加えて「まちとしょテラソ」という愛称がつけられました。



この「まちとしょテラソ」は「学びの場」「子育ての場」「交流の場」「情報発信の場」という 4 つの柱による「交流と創造を楽しむ、文化の拠点」という理念のもとで建築されました。設計段階から町民と一緒に作られた図書館であり、人と人、地域と地域をつなげ、「面」で文化拠点となる図書館を目指しています。

- デジタルアーカイブの取組

システム名	デジタルアーカイブ事業
URL	http://obusesshoso.info/
公開した日時	平成 19 年 12 月
格納資料の種類	地域美術館・博物館の収藏品
格納資料の件数	151 件（おぶせミュージアム・中島千波館） 50 件（高井鴻山記念館）の収蔵作品
アクセス数	現在特に取得していない

同館の館長は、就任当時から、デジタルアーカイブの構築を構想していました。デジタルアーカイブの重要性を説明し、町の人達と一緒に小布施町のデジタルアーカイブを作ってきました。

小布施町のデジタルアーカイブのコンセプトは「100 年後のおもてなし」「100 年後への贈り物」という理念で取り組んでいます。



取組当初の体制は図書館職員、参加機関の担当者、町民などの10人程度の委員会を作成しました。その当時からNIIの中村佳史特任研究員にもアドバイザーとして参加してもらい進めていました。

人材については、県の緊急雇用支援が ついたため、臨時職員を2名採用できました。その職員は図書館の職員として採用しましたが、デジタルアーカイブも兼務で行うことを前提で採用しています。すべてを委託すると自館になにも残らず、結果的に何もできない状況を作ってしまうと考え、できるだけ職員が行えることは行うようにしています。費用面ではデジタル化や勉強会などの経費も含め、19年度に国からの予算1000万円を確保し準備を行いました。

構築は当初からNIIとの共同研究事業となっており、開発はNPO法人の連想出版に委託しています。ソフトウェア等は連想出版の開発したものをベースで利用したため、開発費の負担は、Webページを小布施町用にいくつか作成した費用にあたる約50万円程度に抑えることができました。運用費は年間約400万円程度で、これには追加コンテンツ作成費用、新機能の開発費用、各種物品購入費用が含まれています。

小布施正倉は文化遺産オンラインとシステム連携を行っています。デジタルアーカイブを構築したことにより、「受け身」ではなく「提案」型として情報発信ができるようになりましたが、システム連携することで、より広く多くの方々への情報発信になることをメリットと感じています。

現在は、巻物のデジタル化に取り組んでおり、今年度は巻物全体のデジタル化を行い、次年度はその巻物の解説を作っていく計画です。Web画面上のジョグダイヤルを操作することで本物の巻物を巻くような体験ができるインタフェースを準備し、ページの拡大もできるなど、子供達にも興味を持てるデジタルアーカイブを今後も作って行きたいと語る館長の夢は尽きることがありません。常に新しいこと、楽しいことへ取り組む姿勢により生み出され受け継がれていくデジタルアーカイブは、他の機関への参考になると言えるでしょう。



高井鴻山記念館



おぶせミュージアム・中島千波館

(2) 『信州デジくら』（長野県デジタルアーカイブ推進事業）

● はじめに

長野県では、地域の祭り等の継承が担い手不足により困難になっている状況や、県立歴史館、県立長野図書館等の県所蔵物が劣化して保存が必要になっている状況から、長野県に存在する歴史・文化・自然等の貴重な社会的資産をデジタル化して蓄積、保存し次世代に継承するとともに、インターネットで広く公開することにより、地域アイデンティティの再認識や地域づくりにつなげていくことを目的として、デジタルアーカイブに取り組んでいます。

平成18年度から事業の検討を始め、関係部局によるワーキンググループ、県民の意見聴取を行い、平成20年度に関係団体等14名からなる長野県デジタルアーカイブ推進協議会を設置しその協議を経て、平成20年6月に「長野県デジタルアーカイブ基本構想・推進計画」を策定しています。



● デジタルアーカイブの取組

システム名	信州デジくら（長野県デジタルアーカイブ総合情報システム）
URL	http://digikura.pref.nagano.lg.jp/
公開年月日	平成22年4月1日
格納資料の種類	県立長野図書館・県立歴史館・長野県信濃美術館所蔵物、 県民投稿・市町村提供の映像 など
格納資料の件数	2,415件（平成23年12月末時点）
アクセス数	PV：約10,000件/月平均

『信州デジくら』は、公募による「長野県デジタルアーカイブ推進事業」の愛称です。長野県の歴史・文化・自然等の貴重な社会的資産をデジタル化して保存する「蔵」であり、大切に守り伝えられてきたものや暮らしそのものを、生きた形で次世代へ伝えていくという意味が込められています。

推進計画に掲げる「貯めるだけでなく、使えるデジタルアーカイブの構築」を目指す姿として、研究会やワーキンググループで検討し、平成21年度地域ICT活用交付金（2,300万円）

を補正予算で活用し、県立長野図書館・県立歴史館・信濃美術館が登録機関として参加するシステムを構築しています。

ASP²⁷サービス利用（年間約150万円）によりシステムを運用し、平成22年4月からサイト公開しています。

『信州デジくら』システムは、参加する各館が所蔵品管理システムとして利用できるようなっているため、別の所蔵品管理システムを有しない機関にとってはシステム投資を抑えられるメリットがあります。

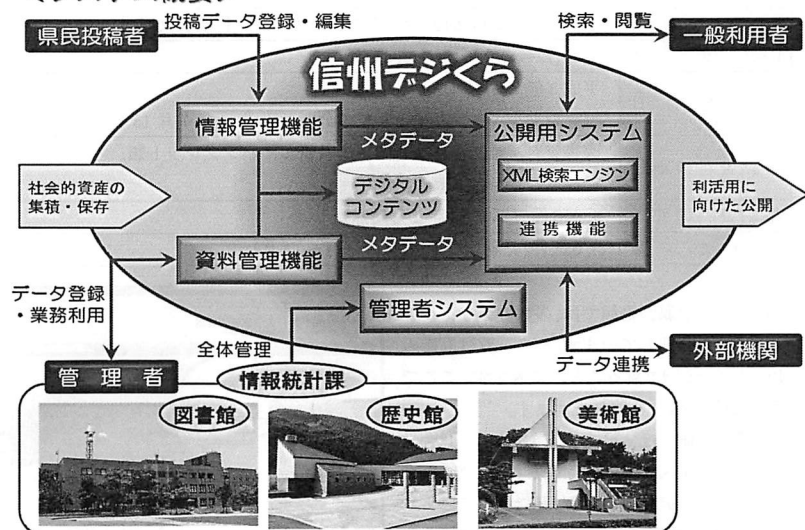
また、公開システムとしての機能を有し、同一のシステム上で、公開・非公開の制御ができるため、デジタルアーカイブのために作業が大きく増えることがないシステム構成になっています。

このようなシステム整備をした点は、デジタルアーカイブの継続的運営の観点からも、今後構築を行う他の機関でも参考になるものと言えるでしょう。

また、他のデジタルアーカイブとの連携機能も有していますが、平成24年1月から国立国会図書館サーチとの連携を始めたところであり、連携の効果が判明するのはまだこれからの状況です。

「地域アイデンティティの再認識と地域づくり」を基本方針として、県の事業推進の下、地域活性化を図ろうとするこのモデルは、地域内連携支援の一つとして注目したい実例です。

<システム概要>



(提供:長野県企画部情報統計課)

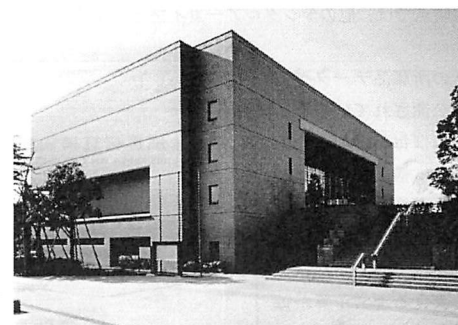
²⁷ インターネットを通じて顧客にビジネス用アプリケーションをレンタルするサービス。主に Web ブラウザーから ASP 事業者のサーバにインストールされたアプリケーションを利用する形態が多い。(IT用語辞典 e-Words: <http://e-words.jp/w/ASPE382B5E383BCE38393E382B9.html> (平成24年1月20日確認)を参考にした。)

(3) 練馬区立美術館

- はじめに

練馬区立美術館は昭和60年10月に開館し、日本の近現代美術を中心に企画展を開催するとともに、2,200件を超えるコレクションを持っています。

同館は、近・現代という『いま』を生きている美術作品を中心に、美を「みる・みせる・つくる」、つまり「鑑賞・発表・制作」という美術館の機能をフル活用し、活動をつづけてきました。ほかの美術館とはちょっと異なる斬新な視点・大胆な切り口で新たな美の知



見を広げ、独自性を追求しています。また、開かれた美術館として設立された同館では、施設利用申込を行うことでギャラリー・創作室を利用することができます。

同館の『ときめきの美 いま 練馬から』というキャッチフレーズには、より多くの区民や美術ファンに愛される美術館へ、そのことが練馬区民の誇りとなり自慢となる美術館へ、という思いが込められています。

- デジタルアーカイブの取組

システム名	所蔵品データベース
URL	http://www.nerima-art-museum.jp/
公開年月日	平成17年3月ごろ
格納資料の種類	近現代の地元の作家を中心とした日本画、版画、彫刻・工芸など
格納資料の件数	約1,700件(平成24年1月現在)
アクセス数	約37,000~71,000件/月(平成23年)

同館は施設がそれほど大きくないため、常設展示をするスペースがありません。このような環境の中でも自館所蔵のコレクションを公開し、利活用を図ることを目的として、所蔵品データベースを構築・公開しました。

デジタルアーカイブという名称ではありませんが、所蔵品をインターネットに公開する取組としては比較的古くから(平成17年)行われており、先進的な実例と言えます。公開当初には新聞にも取り上げられました。²⁸

同館の所蔵品データベースシステムは、内部事務用の所蔵管理システムと同時に整備されているものの、当初から公開用のシステムを構築することを目的としていたことが特筆されます。

システムとしてはASPサービスを利用しており、館内にはサーバーを置いていません。データ更新も含めて運用を委託することで、負担の低減を図っています。

²⁸ 朝日新聞(東京版)平成17年3月25日付

構築費用は、内部管理・公開用のシステム（ポジフィルムからのデジタル化を含む）の初年度構築を約 240 万円で行い、毎年追加する収蔵品のデジタル化も含めた年間の運用費用に約 140 万円を計上して運用しています。

現時点では、他のデジタルアーカイブとの連携はされていません。

この所蔵品データベースで公開されている作品は、ほぼ毎年追加されています。構築初年度だけでなく継続した取組が行われている点は参考にしたい実例です。また、これを支える運営予算が毎年継続していることも重要な点と言えます。

公開されている作品は、地元練馬区や近隣の区にゆかりの作家からの寄贈作品が多いそうです。

同館では、デジタルアーカイブでの情報配信を含め、地元作家の情報発信を積極的に行う取組が評価され、担当の学芸員と作家との信頼関係が築かれることで作品が集まっています。所蔵作品の情報を公開し、広く情報発信することで、美術館の発展が促進されるばかりでなく、地域の活性化にも貢献している実例です。



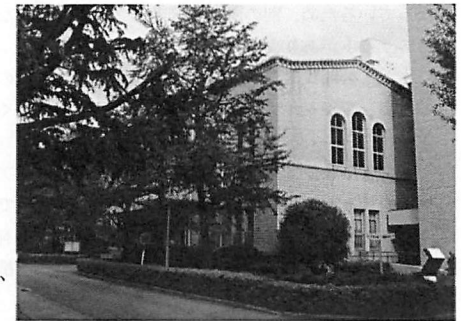
2. Library (図書館)

(1) 神戸大学 震災文庫

- はじめに

神戸大学は、「開放的で国際性に富む固有の文化の下、『真摯・自由・協同』の精神を発揮し、人類社会に貢献するため、普遍的価値を有する「知」を創造するとともに、人間性豊かな指導的人材を育成する」ことを使命とする、神戸市所在の国立の総合大学です。

附属図書館は 4 キャンパスの 9 図書館及び大学文書史料室を運営するとともに、所蔵資料などの知的資産や教育・研究成果のデジタルアーカイブを広く世界に発信し、社会に対する同大学の知的資源のポータル（窓口）機能を果たすことを目指しています。



- デジタルアーカイブの取組

システム名	神戸大学附属図書館 デジタルアーカイブ 【震災文庫】
URL	http://www.lib.kobe-u.ac.jp/eqb/
公開年月日	平成 8 年 7 月 19 日 (デジタルギャラリーとしての公開は平成 11 年 7 月 15 日)
格納資料の種類	図書資料、雑誌資料、新聞・広報紙類、パンフレット類、一枚もの資料、写真資料等
格納資料の件数	4,906 件 (平成 24 年 1 月時点)
アクセス数	3,035 件/月平均 (平成 23 年 4 月～12 月)

神戸大学附属図書館では、多数のデジタルアーカイブを公開しています。その一覧は同館のホームページ²⁹で見ることができます。ここでは、「震災文庫」を取り上げます。

「震災文庫」は、平成 7 年 10 月に、同大学の社会科学系図書館内に開設されました。災害復興や地震研究・防災対策などに役立てるよう、阪神・淡路大震災にかかわるあらゆる資料を可能なかぎり収集し、被災地をはじめ全国の利用



²⁹ 神戸大学附属図書館 デジタルアーカイブ <http://www.lib.kobe-u.ac.jp/da/> (平成 24 年 1 月 20 日確認)

者に広く提供することを目的としています。また、資料は後世に確実に伝えるべく、記録としてすべて保存されています。また、オープンとはほぼ同時にホームページが公開され、資料の目録が公開されました。

その後、来館できない利用者への資料提供のため、資料内容そのもののデジタル化がすすめられました。現在、チラシなどの一枚もの資料、写真、図書、広報紙、音声、ビデオ映像など多種多様な資料が公開されています。

目録（メタデータ）は全件が公開されています。また、所蔵資料の約1割にあたる4,900件強が、一次情報（全文・画像）までデジタル化され公開されています。

予算は、科研費や文部科学省の電子図書館予算を活用しました。

同文庫は大学の附属図書館に設置されていますが、一般の利用者に開放されています。資料もその性質上、市民からの寄贈によるものが多いため、市民の方へのサービスと考えて構築しています。付与するメタデータの記述の作成は、寄贈者の方の納得感があるように配慮しています。

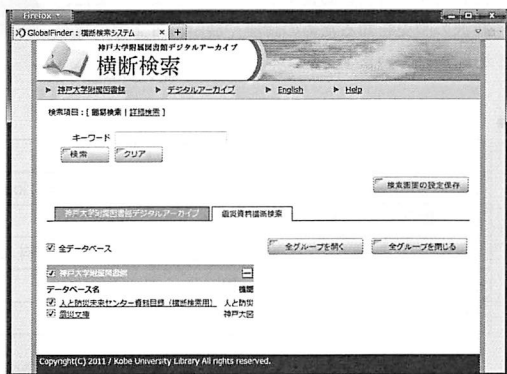
寄贈者が資料を見に来館することもあり、「一緒につくりあげてきた」という意識が利用者と図書館の双方にあり、真の市民参加がなされている実例といえます。

同図書館のデジタルアーカイブでは、「新聞記事文庫」が国立国会図書館サーチと連携しています。震災文庫は、「阪神・淡路大震災記念 人と防災未来センター」と連携し、関連資料の横断検索を実現しています。平成24年からは、兵庫県立図書館も参加します。

このようにデジタルアーカイブの連携は、中央機関のポータルサイトと連携する方法だけでなく、類縁機関と連携することで利用の便を図る方法も有効です。

震災文庫担当者のお話によると、「デジタルアーカイブは試行錯誤の積み重ね」であり、「その時の最善を判断し、過ちを正す」ことが重要であり、また「外部の専門家を信じる——最も正しそうな方を選び出し、信頼して依頼する」こともまた重要です。

取組の中で改善を図っていくこと、自分たちでやるべきことと外部の専門家に任せることを切り分けたうえできちんと依頼することは、今後デジタルアーカイブを構築する機関にとって、参考になる考え方といえます。



(2) 萩市立萩図書館

● はじめに

昭和49年に開館した旧萩市立萩図書館は老朽化のために新築することになり、平成21年秋から新図書館の建設工事が行われ、平成23年3月21日に新図書館が開館しました。新図書館は2階建てで児童館を併設しており、「萩あいふらり」という愛称がつけられています。延べ床面積は約3800平方メートルで、図書館部分の広さは旧図書館の約2倍になり、座席数は3倍の160席程度となりました。そして図書に関しては、自由に読むことができる本が今までの約8万冊から14万冊へと格段に増えました。郷土資料や明治維新史関連書籍などの貴重な図書も調査・研究コーナーに並んでいます。



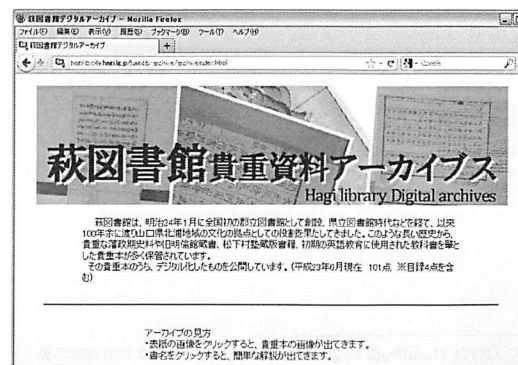
また開館に併せて、「萩市電子図書館」が始まりました。公立図書館で電子書籍を貸し出す電子図書館サービスは全国で3番目の正式導入となるということで、注目を集めています。

● デジタルアーカイブの取組

システム名	萩図書館貴重資料アーカイブス
URL	http://hagilib.city.hagi.lg.jp/hagilib-archives/archiveindex.html
公開年月日	平成23年3月21日
格納資料の種類	萩図書館所蔵の江戸・明治期を中心とした貴重書と資料目録
格納資料の件数	101件（平成23年6月時点）
アクセス数	現在は取得していない。

新しい萩図書館の開館を機に、所蔵している維新史関連を中心とした貴重資料をデジタル化、データ化することにより、情報発信することを目的として取組をスタートした。同年8月からはASPサービスの公開とは別に、萩図書館のホームページにアーカイブスを公開し利用登録者以外からも閲覧できるようにしました。

構築予算には、財団法人図書館振興財団の助成事業を活用しま



した。構築は萩図書館とNPO³⁰法人（館の運営受託団体）との協働体制で行いました。

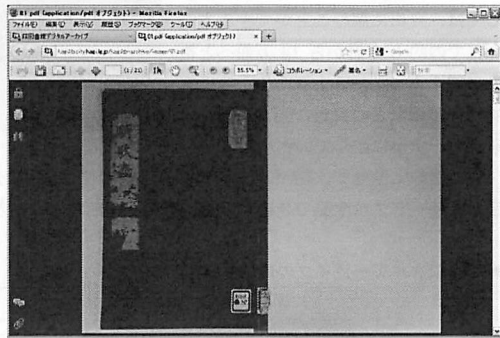
構築後の効果として、今回の取組でアーカイブ化した資料等を所蔵していることを今まで知らなかった利用者に関心を持ってもらうことができた点が挙げられます。また、当初のアーカイブのシステムに電子図書館のプラットフォームを利用したため、アーカイブも含めた電子図書館に対する視察や問い合わせが相次いで寄せられたそうです。

費用面で見ると、初年度のシステム構築、データ変換とその年の運用費も含めて約250万円、構築費用とは別でデジタル化（撮影及びデータ化など一式）の費用に約450万円を要しています。また、今後の継続の際にデジタル化の費用を抑えるため、初年度の費用にてデジタルカメラ、撮影台、記録用パソコン、スキャナ類を整備し、自館でのデジタル化作業に備えています。

年間の保守費用は、ASPサービス運用費用およびデジタル資料の掲載費用を合わせ、約90万円となっています。

現在の所、他の団体等のシステムとの連携はされていません。連携を行いたいと考えているものの、技術的な問題（プラットフォームやデータ共有化）があって進んでいない状況です。

開始して間もない同館のデジタルアーカイブですが、今後環境が整えば、関係機関との間で相互閲覧や検索といった連携ができるように取り組む意向があるとのことで、今後も引き続き注目すべき事例です。



³⁰ 「NPO (NonProfit Organization)」とは、様々な社会貢献活動を行い、団体の構成員に対し収益を分配することを目的としない団体の総称（内閣府NPOホームページ: <https://www.npo-homepage.go.jp/about/npo.html> (2012年1月31日確認) を参考にした。)

3. Archives (文書館)

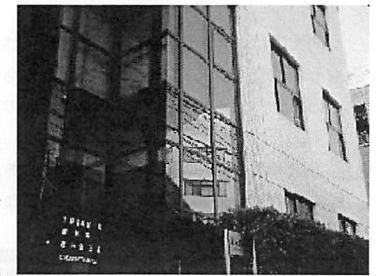
(1) 藤沢市文書館

- はじめに

藤沢市文書館は、郷土の文化を将来へ継承し、先人の歴史に学び現在を知ることで生活を豊かにするため、貴重な文書や記録を散逸させるような事態を阻止し、市民共有の財産を後世に伝えるとともに、地域への理解を深め、現在の行政に反映させるための機関として、次の役割を掲げて活動しています。

- (a) 歴史・行政情報の提供の充実に努める
- (b) 古文書等、地域記録史料の保存継承をはかる
- (c) 藤沢市の歴史を知るための事業を行う
- (d) 行政の記録を残す

同館は、「江の島縁起」「北条早雲文書」といった地域の古文書や郷土資料、行政資料、公文書などを保存しており、地域の歴史が分かる資料館です。市民資料室で資料を閲覧することができます。また、藤沢の歴史を知るための講座や展示、古文書の講座も開かれています。³¹



- デジタルアーカイブの取組

システム名	藤沢市文書館
URL	http://digital.city.fujisawa.kanagawa.jp/
公開年月日	平成22年4月
格納資料の種類	写真、絵葉書、古地図、展示資料など藤沢ゆかりの歴史資料
格納資料の件数	1000点程
アクセス数	月約2500ページ

藤沢市文書館のデジタルアーカイブは平成21年度に試験的な公開を開始し、平成22年4月に正式公開しました。

デジタルアーカイブを公開した目的は、文書館と公文書管理などその役割を知ってもらうための啓発活動です。また、文書館が持つ資料の活用も目的の一つと考えており、学校教育や生涯学習としての利用も期



³¹ 藤沢市観光課・社団法人藤沢市観光協会 <http://www.fujisawa-kanko.jp/spot/spot03.html>

待しています。

したがって、同館のデジタルアーカイブは一般市民向けとして構築されています。デジタルアーカイブの資料を見てそのページをプリントアウトして来館し、閲覧の請求を行う人が最近多くなっています。資料の所在を広く知らせることができる点と、あらかじめ特定されたデータの閲覧請求を受けることで館の職員が効率よく対応できることの2点がメリットといえます。

デジタルアーカイブの構築事業は、藤沢市相互提案型協働モデル事業により採択された NPO 法人湘南市民メディアネットワークと、平成 20 年度より準備を進めてきました。

予算が非常に厳しい中で、委託を受けた担当者の方々の熱意とボランティア的な支援でここまでの機能を実装してもらったと感謝の言葉を館長も述べていました。

見てわかりやすい、興味を持てることが大事であるという言葉通り、非常に美しい映像で、見て楽しいデジタルアーカイブとなっています。(写真1、写真2：「デジタルコンテンツ>デジタル展示>藤沢市の70年」の操作画面)

また、何度もアクセスしてもらうために、コンテンツをできるだけこまめに段階的に更新、拡充し、陳腐化させないことが重要で、そのためには事業の継続性が欠かせないと館長は話しています。

構築の基本方針や取組の姿勢をはじめとして、参考になるポイントが多い事例です。



写真 1



写真 2

4. 実例まとめ

これまでに紹介した事例の特徴を以下にまとめます。参考となる事例を探すための手掛かりとしてください。

機関名	デジタルアーカイブ名	概要
まちとしょテラソ(小布施町立図書館)	小布施正倉	小布施町立図書館が中心に支援を行い、町の人達と一緒に小布施町のデジタルアーカイブを作っている。小さい規模でも共同研究や実証実験を活用し、様々な補助を有効に使うことでデジタルアーカイブを構築、維持している事例
長野県デジタルアーカイブ推進事業	信州デジくら	県及び関係団体等からなる協議会での協議・検討を通じて、地域アイデンティティの再認識と地域づくりにつなげていくことを目的として、県(図書館、歴史館、美術館)所蔵物のデジタル化等のデジタルアーカイブを推進している事例
練馬区立美術館	所蔵品データベース	所蔵品をインターネットに公開する取組としては比較的古くから(平成 17 年)行われており、先進的な事例と言える。長期間に渡り継続している点や、広く情報発信することで、美術館の発展が促進されるばかりでなく、地域の活性化にも貢献している事例
神戸大学	震災文庫	災害復興や地震研究・防災対策などに役立つよう、阪神・淡路大震災にかかわるあらゆる資料を可能なかぎり収集し、被災地をはじめ全国の利用者に広く提供することを目的として明確に定め、関係機関とシステム連携を行っている。コンテンツの収集から市民とともにアーカイブを作り上げた事例
萩市立萩図書館	萩図書館貴重資料アーカイブス	新しい萩図書館の開館を機に、所蔵している維新史関連を中心とした貴重資料をデジタル化、データ化することにより、情報発信することを目的としてスタートした。電子図書館と共存した今後のデジタルアーカイブの形の一つとして注目したい事例
藤沢市文書館	藤沢市文書館(FUJISAWA DIGITAL ARCHIVES)	文書館と公文書管理などその役割を知ってもらうための啓発活動を目的として、見てわかりやすい、興味を持てることをコンセプトにサイトを構成している。継続的なコンテンツの更新を行い何度もアクセスしてもらうことを念頭に置いた取組姿勢などが参考となる事例

第4章 デジタルアーカイブ構築・連携の課題

デジタルアーカイブを構築・連携する上でのさまざまな課題について、平成 21 年度に実施した「我が国におけるデジタルアーカイブの構築に関する調査研究」ではデジタルアーカイブを構築する上での技術的な課題を抽出し、平成 22 年度に実施した「我が国におけるデジタルアーカイブの構築・統合に関する調査研究」では、デジタルアーカイブを連携する上での課題を抽出しました。

本年度は昨年までの課題を踏まえ、デジタルアーカイブを構築し運用して行くために各機関があらかじめ検討しておくべき課題を整理し記載します。また、第 5 章の手引きには課題に対して他館の事例などを記載していますので、参考情報として利用して下さい。

1. 目標設定の課題

(1) デジタル化・デジタルアーカイブの目標設定

各機関によりデジタル化の目的やデジタルアーカイブの構築の目的は異なるものです。それぞれの機関内でその目的や到達点は違いますが、長期利用・保存の観点から考慮した目標設定をしておく必要があると考えられます。

(2) デジタル化対象の判断基準

未デジタル化資料のデジタル化については、現物の保存状態の悪化等、デジタル化対象物の状態を第一に考慮すべきですが、現在の利用者のニーズも含めデジタル化対象の判断基準を決めておく必要があると考えられます。

2. メタデータの課題

(1) 利用のためのメタデータの検討

デジタルアーカイブに格納されたコンテンツは、利用されやすい存在でなければなりません。コンテンツの利活用および連携を容易にするためには最低限、連携する項目定義や語彙については、標準的なものを採用するように館内で検討しておく必要があると考えられます。

(2) 保存のためのメタデータの検討

構築後に長い時間が経過すると、構築当初の目的やデータ作成時に使用された技術などの情報が失われ、オリジナルの状態を再現することに支障が生じる懸念があります。各館でデジタルアーカイブへ格納する対象物をデジタル化した時に、デジタル化した時の情報を残しておく必要があると考えられます。

3. 技術的な課題

(1) ファイルフォーマットの採用基準

デジタルアーカイブを構築する上で最適なファイルフォーマットは各デジタルアーカイブの目的や予算により異なります。国際標準化された規格（デジュール・スタンダード）、標準化されていないが事実上の標準になっている規格（デファクト・スタンダード）など幾つか挙げられますが、どちらが優れているというのではなく、使用目的に応じた適切なファイルフォーマットを検討する必要があります。どの規格を採用するにしても将来的にデータ変換を考慮したファイルフォーマットを検討する必要があります。

(2) ファイルフォーマットの変更基準

技術の変遷に対応するため、一度採用したファイルフォーマットを変更することが考えられますが、ファイルフォーマットの変更実施を判断する基準を検討する必要があります。

(3) 保存媒体の選択

ファイルを保存するための媒体に何を使用するか検討する必要があります。DVD³²、Blu-ray Disc³³等の光ディスクが一般的ですが、後でデータが読めない、保存場所を取られる等の課題があります。ハードディスク³⁴は価格が年々下がっていることもありますが、障害時には全てが読み出せなくなってしまうことなど、バックアップや冗長化については慎重な検討が必要です。

(4) 長期利用・保存のための方法検討

デジタル化の技術については、技術革新が現在も進んでおり、今は再生可能なデジタルデータでも将来に渡り問題なく再生できる保障はありません。つまり、昔作成されたファイルフォーマットを再生するソフトウェアが利用できなくなり、表示も印刷もできない現象が起こることがあります。また、デジタルデータを保存した媒体の寿命やハードウェアの寿命で利用できなくなるなど、長期的な観点での利用や保存のためには何らかの対応を行う必要があります。

日本、海外でその手法について調査研究が進められており、それらの動向の把握と各機関でもどのようなタイミングでどのような手法を使って長期利用・保存の対応を実施するかが課題と考えられます。

³² Digital Versatile Disk の略称 データ記録メディアとして利用される光学ディスクの一種（IT用語辞典 e-Words: <http://e-words.jp/w/DVD.html>（平成 24 年 1 月 23 日確認））

³³ 現行の DVD を超える容量を実現する「次世代 DVD」規格の光学ディスクのこと（IT用語辞典 e-Words: <http://e-words.jp/w/Blu-ray20Disc.html>（平成 24 年 1 月 23 日確認）を参考に記述した。）

³⁴ パソコンを初めとするほとんどのコンピュータに搭載されている、代表的な外部記憶装置のこと（IT用語辞典 e-Words: <http://e-words.jp/w/HDD-2.html>（平成 24 年 1 月 23 日確認））

4. 人材と体制の課題

(1) 人材に求められる能力

デジタルアーカイブを構築・推進する人材に求められる知識は多岐に渡ります。しかし、都道府県レベルでの機関においても、そのような人材を確保することは難しく、地方の組織になると更に深刻な状況です。これまでのデジタルアーカイブの構築事例で見ても、推進役となる人の存在が大きく、人依存となっていることが課題と考えられます。

(2) 組織の予算と維持

デジタルアーカイブを長期に渡り維持するため、担当者の異動や予算配分の停止によってデジタルアーカイブの運用が停止しないように、中長期的な予算措置が求められます。また、デジタルアーカイブの性質上、増え続けて行くデータを保持するための予算確保などが検討課題と考えられます。

第5章 デジタルアーカイブの構築・連携の手引き

この章の注意点

この「手引き」の章では、できる限り実践的なガイドライン例とするため、仮の読者像を設定しました。その読者がデジタルアーカイブを立上げる業務を与えられたという想定で、デジタルアーカイブを作っていくための全体像をつかんで頂くことを目的としています。

■ 仮想読者の環境

地方の博物館・美術館や図書館、文書館でデジタルアーカイブを推進する職員

■ デジタルアーカイブの対象

自館の所蔵品、地域の貴重資料

1. デジタル化とシステム構築の前に：自館の現状把握と準備

デジタルアーカイブを構築すると言っても、どのようなことを準備・用意すればよいかかわからないといった声が聞かれます。一概に「こうやればできます」と言う方法はありませんが、以下にデジタルアーカイブを構築、運用する際に押さえておくポイントを、3つの段階に分けて記載します。

段階	プロセス		知の記録組織	組織・機関 地域支援	支援(委託)企業
準備段階	自館の現状把握と準備	(a) 計画を策定する	◎	○	○
		(b) 対象物を選定する	◎	○	
		(c) 著作権等の処理を行う	◎		
構築段階	デジタル化 (a・b) システムの構築 (c・d)	(a) 対象物をデジタル化する	◎	○	◎
		(b) メタデータスキーマを整理する	◎	○	○
		(c) 基本機能・連携機能を確定する	◎	○	
		(d) デジタルアーカイブを構築する	◎	○	◎
運用段階	デジタルアーカイブの継続運用	(a) 計画を再検討する	◎	○	○
		(b) 人材と教育	◎	◎	○
		(c) 長期利用・保存	◎	○	○
		(d) アウトリーチ	◎	○	

- ◎ 主体的に行う実施者（実施主体）
○ 実施主体を補佐する役割の関係者

(1) 計画を策定する

(a) 全体方針の検討

デジタルアーカイブを構築する目的、提供方法、利用対象者などを考え、実際に使える予算と構築にかかる想定予算を比較検討して、計画を整理しておきましょう。

その中でも最初に目的や方針を明確にしておくことが重要です。いくつかの事例を紹介いたします。

小布施町立図書館 まちとしょテラソ Web サイト³⁵より

『100年前を伝え、100年後へサービス』

小布施町の文化（文化財的な文化、生活＝小布施ならではの風土的文化）の収集を行い、そしてまとめ公開（発信）することにより、人づくりや町づくりに役立たせることができると考えます。その収集された情報を蓄積するシステム、デジタルアーカイブ（デジタル化された「保存記録」や「記録保存館」）は、未来へのタイムカプセルとなることと同時に、私たちが、先人から学び、今をどのように生き抜いていくのかという羅針盤の役目を果たす事となると考えます。このことから、町民の皆さんと小布施にとってのデジタルアーカイブを考え、協働でデジタルアーカイブ事業を行っています。

「長野県デジタルアーカイブ基本構想・推進計画」（2008年6月長野県策定）³⁶より

『地域アイデンティティの再認識と地域づくり』

- 伝統文化と情報通信関連産業の連携による文化・産業基盤の整備地域の伝統文化を再発見し共有するとともに、コンテンツ産業をはじめとした情報通信産関連産業の活性化を促すことで、様々な分野の連携・協働による文化創造・産業創出のための基盤整備に資する。
- 世界に向けた長野県の魅力の発信による観光振興資料等から読み取れる確かな地域特性を共有し、長野県の魅力を世界に発信していくことで、県内外からの実物への興味・関心を高め、長野県の観光振興に資する。
- 発信能力を持ち地域を支える人材育成地域を見直し、地域の魅力を発見し理解して、新たな価値を創造しながら発信することで、地域アイデンティティを次世代に継承・発展させ、地域を支えることができる人材育成を行う。

³⁵ 小布施町立図書館 まちとしょテラソ <http://machitoshoterrasow.com/index.html>（平成24年2月12日確認）

³⁶ 長野県デジタルアーカイブ基本構想・推進計画(本文)

http://www.pref.nagano.lg.jp/kikaku/josei/da/koso_1/koso.pdf（平成24年2月12日確認）

(b) コストの検討

前項で検討したデジタルアーカイブ構築の全体方針から、利用対象者などを考え、実際に使える予算と構築にかかる想定予算を比較検討して、自館で構築できるデジタルアーカイブの仕様を整理しておきましょう。

① デジタル化のコストの考え方

デジタル化のコストは、デジタル化の目的、資料の種類、状況、委託する量や期間によって大まかに決まります。立体物や、ふすまのような大きなものは異なりますが、書籍などの平面で一般的な資料のデジタル化の考え方について説明します。

(ア) 目的の分類

デジタル化の目的を大きく分類すると、以下のように分けることができます。

目的により、画像の精細度が変わり、連動してコストも変わります。

以下の分類は、上から高精細画像が必要となる順に並べてあります。

1. 特殊な研究目的のもの
2. 印刷で精密な複製をつくるためのもの（特に色の管理が重要になります）
3. 印刷して読め、OCR³⁷でテキスト抽出できるもの
4. Web で読め、OCR でテキスト抽出できるもの
5. Web で読むためのもの

それぞれ、カラーかグレースケール³⁸かによってコストが変わります。ほかの条件が同じならば、グレースケールのほうが安価です。

メモ：OCRのコストについて

画像の精細度を検討するための使用目的の例として、OCRの実施有無をあげましたが、OCR関連の作業自体のコストも検討する必要があります。

OCRの精度は完璧ではなく、似た形の文字を間違えたり、文章のレイアウトを正しく認識しないことがあります。「OCRした文章は検索用である」と割り切って誤りを許容する、という考えもありますが、そうでなければ処理後には校正が必要になります。校正をどこまで行うか（ルビの対応など）がコストに影響を与えるため、OCRの処理を行う場合には注意が必要です。

³⁷ OCR:(Optical Character Recognition, 光学文字認識)とは、画像データ上にある文字と思われる部分を解析し、コンピュータ上で扱える文字(テキスト)データに変換することを示します(朝日新聞「知恵蔵2009」)

³⁸ コンピュータ上で色の表現方法の一つ。画像を白から黒までの明暗だけで表現し、色の情報は含まない「モノクロ」のこと。(IT用語辞典 e-Words <http://e-words.jp/w/E382B0E383ACE383BCE382B9E382B1E383BCE383AB.html> (平成24年1月30日確認))

(イ) 資料の種類・状態

資料の種類や状態により、デジタル化作業に要する時間や機材が異なりますので、コストが変わります。以下の分類では、下にいくほど金額を低く抑えられます。

1. 貴重書のようなものや、資料の状態が悪く扱いにくいもの
2. 製本雑誌のように厚いものや、大きさが不揃いなもの
3. それほど貴重とは言えないが、裁断できないもの
(非接触/接触可でデジタル化するかどうか異なる)
4. 裁断できるもの(フィーダ³⁹付きスキャナで取り扱ってよいもの)

(ウ) 量と期間

デジタル化のコストはその多くが人件費であるため、取扱に慣れることで、生産性が大きく向上します。量が少なく定常ラインがなければ、低い生産性しか出ないため、コスト高になってしまいます。生産性が決まると量と期間で作業者の人数が算出できます。ただし、作業者が多いほど熟練者の割合が少なくなり、品質が悪化するため、生産性が低下する傾向があります。

他にも、納品画像データの種別・媒体、メタデータ・目次データの作成の有無やその形式、作業場所の指定(あるいは作業場所の環境の指定)の有無や、一度に持ち出せる数量などに関する条件など、様々な仕様がコストに大きく影響します。また、大規模なデジタル化を行う場合には、物流に関する時間とコストを計算に入れる必要があることなども注意が必要です。

このようにデジタル化と言ってもコストの幅が大きいため、目的の設定、資料の調査、仕様の確定、サンプル調査によって最適な選択をすることが重要です。

以下に、デジタル化のコストについて検討する助けとなる事例を紹介します。

³⁹ 印刷用紙を自動的に供給する装置。シートフィーダのこと。(IT用語辞典 <http://www.sophia-it.com/content/%E3%82%B7%E3%83%BC%E3%83%88%E3%83%95%E3%82%A3%E3%83%BC%E3%83%80> (平成24年1月20日確認))

デジタル化のコストについて (1) ～慶應義塾大学メディアセンターの経験から～

デジタル化の対象物や品質によるコストの違いの例を、慶應義塾大学メディアセンターの経験からのコメントで紹介します。

- ① 取り扱いの難しい貴重な資料を、カラーマネージメントをきちんと行ない、デジタル複製可能な品質レベルのデジタル化を行った場合
1 ページ 3000 円前後 の見積もりが出ることが多い。
- ② 裁断可能で読めればよくて、OCR をかけて校正しない場合、安価な業者では、300 ページの本が 一冊 200 ～600 円程度 = 1 ページ 2 円以下で可能である。
- ③ 同じように、裁断可能で、OCR をかけて校正はしない場合でも、オンデマンドプリント[※]用の画像などの場合は、デジタルデータへの検査・加工・信頼度を担保し、1 ページ 100 円以上になる場合もある。

※ コピー機の技術を応用し、少数数でも安価で高品質な出力を実現するデジタル印刷方式。

上記のように、デジタル化のコストを出す場合は要件を検討した上で計算する必要があります。表 5-1 に、日本の MLA 機関での事例を記載します。

以下の事例の価格は、いずれも上記 (ウ) にて説明したとおり、デジタル化の量や環境など様々な要件によって変わることにご注意ください。

表 5-1 デジタル化コストの参考数値

秋田県立図書館					
※量や環境など様々な要件により単価は大きく変動します。 表中の数値は何らの基準を表すものではありません。					
対象物	役務内容	数量	単価	価格	日付
貴重資料	・フィルム撮影 ・フィルムからデジタル化 (白黒)	マイクロ 10万コマ	約 150 円	約 1,780 万円	平成 9 年
		6×7 (中判) 100 コマ	約 6,000 円		
		4×5 (大判) 約 200 コマ	約 11,000 円		
秋市立萩図書館					
※量や環境など様々な要件により単価は大きく変動します。 表中の数値は何らの基準を表すものではありません。					
対象物	役務内容	数量	単価	価格	日付
明治維新史関連書籍	メタデータ作成 デジタル化、(JPEG、PDF、独自形式 .iek への変換) システム登録 ※電子図書館システム構築時	100 冊 約 5,400 コマ	—	約 450 万円	平成 23 年

② システム構築のコストの考え方

システム構築のコストを検討する場合は、そのシステムの要件を決めておく必要があります。機能についての詳細は次節の「基本機能・連携機能を確定する」に記載していますが、各館で必要とする要件から、想定する機能を抽出するための資料を巻末「参考資料7 チェックシート」として付けました。

またそのチェックシートを使って仮想の博物館で要件を検討していく流れを「参考資料8 発注ケース：館内利用者への所蔵資料公開」に示していますので併せて参考にしてください。

なお、各自治体の MLA 機関におけるシステムの調達にあたり、政府調達の基本指針と自治体が定める情報システムの指針やガイドライン等を参照しておきましょう。基本指針の資料と、参考となる自治体から公表されている指針例等を併せて以下に記載します。

表 5-2 政府調達の基本指針及び自治体ガイドライン等

資料名	URL
「情報システムに係る政府調達の基本指針」(2007年(平成19年)3月1日 各府省情報化統括責任者(CIO)連絡会議決定)	http://www.meti.go.jp/policy/it_policy/tyoutatu/
「情報システムに係る政府調達の基本指針」実務手引書(第2版)(2007年(平成19年)7月1日 総務省行政管理局)	http://www.soumu.go.jp/main_content/000141665.pdf
自治体例) 岡山県情報システム調達ガイドライン	http://www.pref.okayama.jp/page/detail-89152.html
自治体例) 浦安市情報システム調達指針	http://www.city.urayasu.chiba.jp/item18900.html

表 5-3 システム構築コストの参考数値

■システム構築費			
※規模や機能など様々な要件により価格は大きく変動します。表中の数値は何らの基準を表すものではありません。			
組織名称	役務	価格	日付
秋田県立図書館	図書館システム構築 1式 ※電子ライブラリーの費用は一部分	約2,000万円	平成16年
萩市立萩図書館	システム構築(初年度の運用費含む) データ変換 システム構築期間 H22.11~H23.3上旬	約250万円	平成23年
練馬区立美術館	内部管理・公開用のシステム(ポジフィルムからのデジタル化を含む)	約240万円	平成17年

③ 保守費の考え方

保守費は利用しているシステムの利用形態により変わります。ASP サービスなどを利用している場合と、自館にサーバなどの設備を持つオンプレミスとでは保守にかかる費用は変わってきます。システムの規模や、保守に含める作業範囲によっても変わります。

また、サービスレベルの設定の仕方によっても、コストに影響を与えます。特に土日祝日等通常の企業が休日に当たる日の対応や、24時間365日の対応など、人件費へ直接影響するサービスレベルを設定すると、コストに大きく影響が出ます。

以下にいくつかの事例を紹介します。

表 5-4 システム保守コストの参考数値

■保守費			
※規模や作業内容など様々な要件により価格は大きく変動します。表中の数値は何らの基準を表すものではありません。			
組織名称	役務	価格	年度
小布施町立図書館	小布施正倉, 小布施人百選 システム運用、ソフトウェア開発、データ作成を含む	約338万円/年	平成23年
秋田県立図書館	図書館システム保守 1式 ※デジタルライブラリー保守費含む デジタル化追加費用など ソフト費+ハード費	約240万円/年	平成23年
長野県デジタルアーカイブ推進協議会	ASP サービス利用料	約150万円/年	平成22年
練馬区立美術館	ASP サービス利用料(館内・公開システムの保守、画像・メタデータの更新も含む)	約139万円/年	平成23年
萩市立萩図書館	ASP サービス利用料 (50点のデジタル資料掲載経費含む)	約95万円/年	平成24年(案)

(2) 対象物を選定する

対象物の選定にあたっては、対象となるモノそのものについての知識を持ち、その価値についてある程度評価できる人材が、その選定過程に関与することが望ましいと考えられます。

さらに、デジタル化したあとの活用まで考えて選定することを考えるのであれば

- ・学校教育等で活用することを想定する場合には教育関係者
- ・地域ブランドの立ち上げ等も想定する場合には地元企業

など対象物を選定する段階から巻き込むことも検討してみると良いでしょう。事例でも挙げています「信州デジくら」の場合は、長野県デジタルアーカイブ推進協議会を設置しその協議により対象物を選定しています。

しかし、一方でデジタルアーカイブの導入を検討している機関には、知的資産が手をつけられずに残っているケースが多いことが様々な機関へのヒアリング⁴⁰から課題として挙がっています。但し、自館内にある全ての知的資産をデジタル化するには予算や人の制約があり

⁴⁰ 平成22年度総務省「我が国におけるデジタルアーカイブの連携に関する調査研究」より

困難であるため、デジタルアーカイブに収録する対象について優先順位を決めて選定する必要があります。

対象物の優先順位づけとして下記の2つが挙げられます。

- ・ 一次情報（現物）そのものが経年劣化により今後利用できなくなる知的資産
- ・ その地方独特の知的資産（工芸品、その地方で生まれた有名な作家の資料など）

さらにその中でも優先順位をつけたい場合は、人気のある資産（対象物）を選ぶことをお勧めします。人気があるということは、すなわち様々な人達がそれを見たいと思う物です。そのような知的資産がデジタル化され公開されることにより、その資産と所蔵館を宣伝することができ、来館のきっかけにもつながるなどの効果が見込まれます。

現に、デジタルデータ（写真）をインターネット上で販売している会社では、デジタル化方針の一つとして「人気がある物」を優先順位の一つにしています。⁴¹

また、次項で説明するように、公開の許諾を得やすいものを優先して選ぶのも一つの方法です。

以下に、一般的に挙げられるデジタルアーカイブの対象物の例を示します。

- ・ 建築物や美術工芸品(絵画、彫刻、工芸品、書跡・典籍、古文書、考古資料、歴史資料)などの有形文化財
- ・ 演劇、音楽、工芸技術などの無形文化財
- ・ 生活の中で用いられた有形民俗文化財
- ・ 写真や映画、ビデオで撮影したテープ
- ・ 特殊コレクション・貴重書（主に図書館）
- ・ 写真・会誌（主に博物館）



神戸大学附属図書館「震災文庫」のような阪神・淡路大震災に関する文献・資料など、震災にかかわるあらゆる資料を収集し公開している所もあります。このようにテーマを絞って資料を公開することも効果的です。

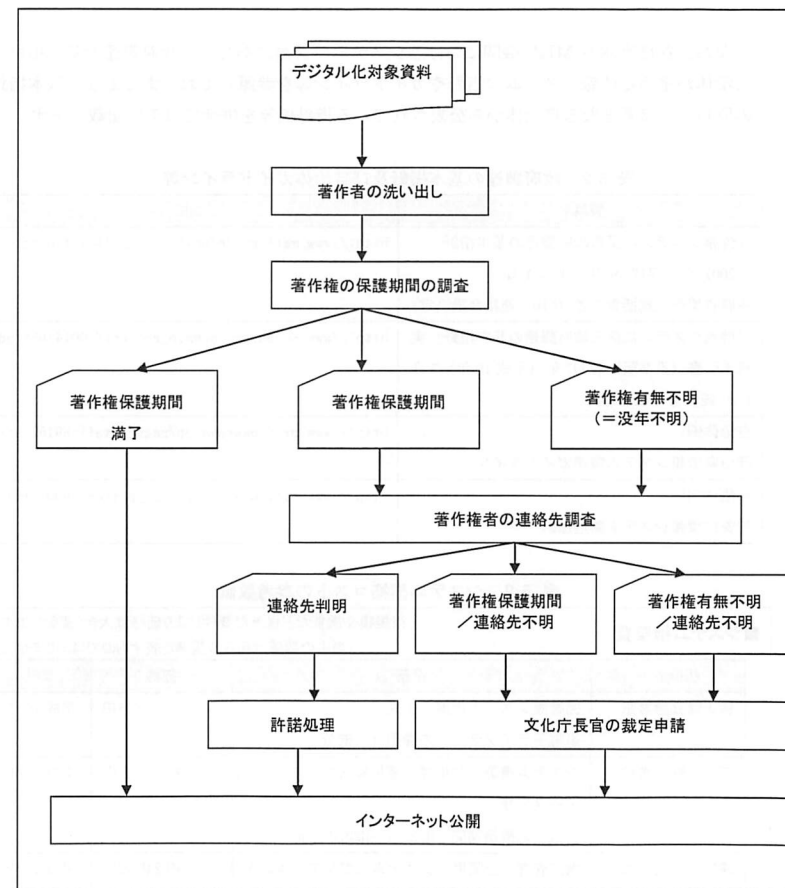
また、対象物の選定対象に対しても目録をはじめからつくるのではなく、既存のものを利用するという視点は重要です。対象物の点検と資料の状態、目録の有無などの整備とできるだけ一連の流れで実施することが良いと言われています。

⁴¹ 平成 23 年度総務省「我が国におけるデジタルアーカイブの構築・連携に関する調査研究」より

(3) 著作権等の処理を行う

一般的に対象物をデジタル化しデジタルアーカイブとして公開する場合、知的財産権やプライバシーについて考えておくことが重要です。

著作権処理については、国立国会図書館で平成 23 年 3 月に改訂版として報告されている、「国立国会図書館資料デジタル化の手引 2011 年版」⁴²に具体的な処理の流れが記載されています。ここではその概念図を例として記載します。各項目の詳細は本文 88 ページを参照して下さい。



⁴² 国立国会図書館 資料デジタル化の手引 2011 年版 P88 参考資料 4 著作権処理
<http://www.ndl.go.jp/jp/aboutus/pdf/digitalguide2011.pdf>

著作権契約の基本的な考え方や具体的な条項の趣旨を知りたい場合は、文化庁の「誰でもできる著作権契約マニュアル」⁴³が参考になります。また、著作権の許諾の例として、デジタルアーカイブ推進協議会(JDAA)の活動成果物、デジタルアーカイブ<権利問題と契約文例>⁴⁴にある、契約文例(1)「素材に著作権が存在する場合の、素材の著作権者とデジタルアーカイブ構築・運営者との間での著作物の利用許諾契約書」の例を抜粋して記載します。

利用許諾契約書(案)

<作品(静止画)の著作権者>(以下「甲」という)と<デジタルアーカイブ構築・運営者>(以下「乙」という)とは、甲が著作権を有する作品を利用許諾することについて、次の通り契約を締結する。

第1条(定義)

本契約において、次の用語は、以下に定める意味を有する。

- (1) 本件作品とは、甲が著作権を有する別紙1欄記載の作品をいう。
- (2) デジタルアーカイブとは、デジタルデータを収録素材として構築されたデータベースであって、乙が保有するものをいう。
- (3) 本件データとは、デジタルアーカイブに収録された本件作品をデジタル化したデータをいう。
- (4) 紹介とは、第三者に対し、本件データの利用を促すために、本件データを提示することをいう。
- (5) 提供とは、本件作品の利用許諾を得た第三者に対し、本件データの複製物を引渡すことをいう。

第2条(利用許諾)

乙は、以下の各号に定める範囲で、本件作品を非独占的に利用することができる。

- (1) デジタルアーカイブの構築
本件作品をデジタル化し、データベースに含ませることによりデジタルアーカイブを構築すること。
- (2) 本件データの紹介又は提供
デジタルアーカイブに収録された本件作品を印刷物又はCD-ROM、MO、DVD等の電子記録媒体に複製して譲渡し、又はインターネット等のネットワークで公衆送信することにより、紹介又は提供すること。

第3条(再利用許諾)

乙は、本契約の期間中以下の各号に定める範囲で、本件データを第三者からの申し込みに応じて随時再利用許諾することができる。

- (1) 印刷物を制作し頒布、販売すること
 - (2) CD-ROM、MO、DVD等の電子記録媒体による作品を制作し頒布、販売すること
 - (3) 放送番組を制作し放送すること
 - (4) インターネットのウェブサイトを作成し公衆送信すること
 - (5) 前各号に伴う宣伝広告活動に利用すること
2. 前項による再利用許諾の場合、乙の第三者に対する当該再利用許諾の有効期間は最長〇年間を超えないものとし、期間を延長する場合は、本契約による有効期間の条件等を考慮し、甲乙協議してその取り扱いを決定するものとする。
3. 再利用許諾を希望する第三者が、本件データを改変して利用する場合については、その都度甲乙協議して取り扱いを決定するものとする。

上記以外に

- ・ 契約文例(2)「素材の所有者とデジタルアーカイブ構築・運営者との間での素材の使用承諾についての契約書」
- ・ 契約文例(3)「デジタルアーカイブ構築・運営者と利用者との間でのデジタルデータ等

⁴³ http://www.bunka.go.jp/chosakuken/keiyaku_manual.html

⁴⁴ <http://www.dcaj.org/jdaa/public/kenri/kenri.html>

の利用許諾契約書」

などの文例があります。また、本文中にはコメントや変形例の記載もありますので、是非一読することをお勧めします。

MLAの所有物ではない、寄託資料や地元の個人・団体等が所蔵する資料をデジタル化する場合には、著作権上の権利はない場合でも、何らかの取り決めをしておいた方が、その後の関係を良好に保つためには良い場合もあります。

また、一般市民から写真などを新たに収集する場合の許諾の例として、知のデジタルアーカイブに関する研究会(第3回)で紹介された、Yahoo! Japan 東日本大震災 写真保存プロジェクトの投稿ガイドライン⁴⁵を示します。

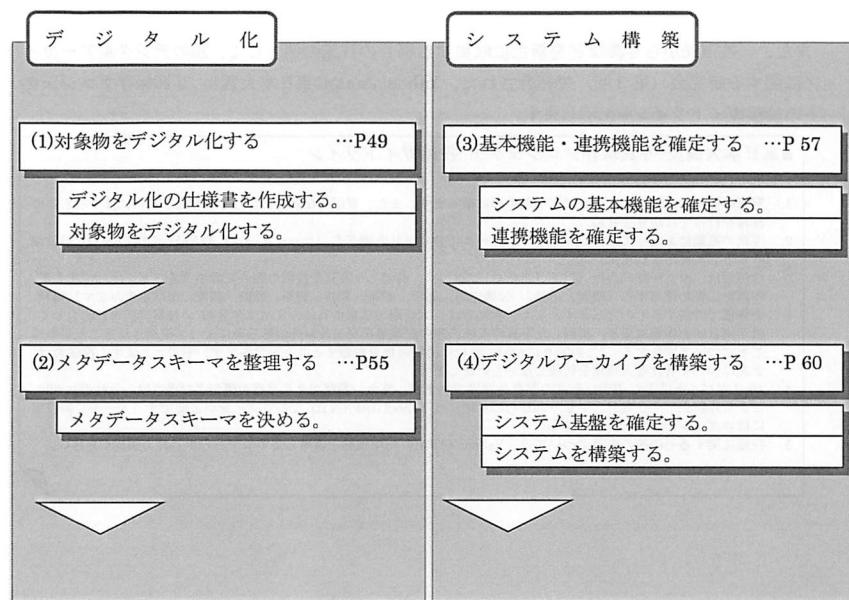
■東日本大震災 写真保存プロジェクト 投稿ガイドライン

1. 写真の投稿は、本企画の趣旨に沿ったものに限り、また、宣伝的な要素を含むものや公序良俗に反するものの投稿は行いません。
2. 写真の掲載にあたっては、Yahoo!プロフィールで設定された表示名(ニックネーム)やYahoo! JAPAN IDが投稿者名として表示されます。
3. 投稿者は、ヤフー株式会社(以下「当社」)に対して、投稿した写真や投稿の際の記載事項を日本国内外で本企画の趣旨に従って使用する(複製、上映、公衆送信、展示、頒布、譲渡、貸与、翻訳、翻案、出版を含みます)権利を無償で許諾するものとします。また、投稿者は、本企画の広報の目的、または非営利の復興支援の目的もしくは学術目的の場合に限り、投稿した写真や投稿の際の記載事項が当社以外の第三者によって使用されることがあることについて同意する(この限りにおいて当社に再許諾権を許諾することを含みます)ものとします。投稿者は、上記について著作人格権を行使しないものとします。
4. 10メガバイト以下のJPEG形式の写真を投稿できます。また、投稿できる写真の幅および高さは、それぞれ200ピクセル以上です。なお、フォームからの投稿にはYahoo! JAPAN IDでのログインが必要です(メールでの投稿にはログインの必要はありません)。
5. 投稿に関するその他の事項については、Yahoo! JAPAN 利用規約を適用しますのであらかじめご確認ください。

⁴⁵ 東日本大震災 写真保存プロジェクト ガイドライン <http://archive.shinsai.yahoo.co.jp/upload/> (平成23年12月29日確認)

2. デジタルアーカイブを作ろう

ここでは、デジタルアーカイブを実際に作るために必要となる実務的な作業を記載します。各機関の目的や規模により個々の仕様も様々となります。システムの仕様にデジタル化も密接に関係しますので、ここでは並行した形で検討を行うことをお勧めします。一般的に必要と思われる作業をそれぞれ順に記載しますので、適宜必要な箇所を参照して、自館のデジタルアーカイブの構築に利用して下さい。



(1) 対象物をデジタル化する

(a) デジタル化の仕様書を作成する

デジタルアーカイブの構築を担当する方が最初に考えることは、デジタルアーカイブに格納する対象物をどのようにデジタル化するかという点です。しかし、各機関で独自に対象物をデジタル化することは技術的や人材的な面で困難であることが多く、企業への委託を行うことが一般的な選択です。従って、企業へ対象物のデジタル化を委託するための仕様書を作成することが必要となります。「国立国会図書館 資料デジタル化の手引 2011年版」などの資料を中心に、仕様書を作成するために押さえておきたい技術や知識について記載します。

① デジタル化方法の検討

対象物をデジタル化する方法にはいくつかの方法があり、目的と対象物に応じて適切な方法を選択することが重要です。

デジタルアーカイブで対象物となる物はテキスト、静止画像、音声、動画などさまざまな形態が考えられますが、ここでは特に多いと考えられる静止画像や資料、書籍について記載します。「最新の技術と図書館サービス 第6章 資料のデジタル化と図書館」⁴⁶には、静止画像のデジタル化の方法が幾つか挙げられています。

- ・フィルム撮影し、フィルムからデジタル化する方法
メリット：中間生成物としてフィルムが残り、資料の長期保存が可能となる
デメリット：撮影とデジタル化の二回の作業が必要であり、費用面など負担増になる
 - ・資料を直接スキャンしてデジタル化する方法
 - ・デジタルカメラ撮影によってデジタル化する方法
メリット：作業が一回ですみ费用的に安価、色校も一回になり原資料の再現性が高い
デメリット：中間生成物がないため、資料の長期保存に注意が必要となる
- 資料を直接スキャンしてデジタル化する方法は、フラットベッドスキャナ⁴⁷などを使って手軽にデジタル化が可能で少量のデジタル化資料を比較的安価に提供・公開したい場合に適した方法と言われています。

上記以外の方法もありますが、それぞれの方法に対して、メリット・デメリットがありますので各機関の目的等に併せてデジタル化する方法を選んで下さい。

② 画像ファイルフォーマットの検討

保存用のファイルフォーマットとしては非圧縮の TIF(ティフ)を採用している機関が多くあります。しかし、一方では国立国会図書館・国立公文書館などの機関で保存用として JPEG 2000(ジェイペグ 2000)⁴⁸が採用されており、JPEG(ジェイペグ)には無い利点がある

⁴⁶ 大串夏身 山崎博樹「最新の技術と図書館サービス 第6章 資料のデジタル化と図書館」青弓社 2007年 P232

⁴⁷ 原稿をガラス台に固定し、下から光を当てて読取装置を動かして画像を読み取るスキャナのこと

(IT用語辞典 e-Words)

<http://e-words.jp/w/E38395E383A9E38383E38388E38399E38383E38389E382B9E382ADE383A3E3838A.html> (平成24年1月23日確認)

⁴⁸ 世界標準規格は参考資料5に規格番号記載

多いことも事実です。JPEG に比べ、表示する際にプラグイン等のソフトウェアが別途必要となるようなシステムの制約があるため、判断は必要ですが選択肢として注目するフォーマットの一つと言えます。他のフォーマットとしては、慶應義塾大学メディアセンターでは、デジタルカメラで撮影したデータは RAW(ロウ)⁴⁹データを保存用としている事例もあります。RAW データを保存するメリットは撮ったデータをそのまま「生」データとして保存するため、画質が劣化せず、修整を柔軟に行えることが挙げられます。デメリットはカメラメーカーによって形式が異なるため、メーカーのサポート終了等で表示ができなくなるなど、メーカー依存が高まる可能性がある所です。知らない間にソフトウェアがサポートされなくなることがないように、定期的を確認するなどの注意が必要となります。

「国立国会図書館 資料デジタル化の手引 2011 年版」では、「保存用画像」「提供用画像(中間サイズ)」「提供用画像(拡大サイズ)」「サムネイル画像」の 4 種類が説明されています。各館の目的や方針に応じて利用する画像の種類を選択しますが、保存用・提供用の 2 種類のフォーマットは最低限決めておくことを推奨します。

機関名	ファイルフォーマット(画像)
東京国立博物館	保存用：TIFF(非圧縮)、提供用：JPEG
国立国会図書館	保存用：JPEG 2000、提供用：JPEG/JPEG 2000 併用
国立公文書館	保存用：JPEG 2000、提供用：JPEG

長期保存を考えた場合には、保存用として、**後で加工できるようなフォーマットで残しておく**ことがポイントと考えられます。(TIFF の非圧縮も有効な保存フォーマットと言えますが、ファイルサイズが JPEG 2000 などと比べると約 2 倍近く増えますので、保存媒体のコスト等も含め考慮が必要です)

画像ファイルフォーマット(提供用)では JPEG が最も多く利用されています。文書関連のフォーマットでは ISO の規格にもなった PDF/A が多く利用されています。

後で加工できるという点では、文字情報のみの場合、TEXT も有効です。

③文字コードの検討

ファイルフォーマットが TEXT であるなしに限らず、システムで採用する文字コードについても選択しておく必要があります。平成 21 年 3 月に国立公文書館が公開している『デジタルアーカイブ・システム標準仕様書』⁵⁰の 13 ページに文字コードについての「標準仕様の考え方」が記載されていますので、以下に参考のため記載します。

- ・ UCS-4 を表現できる文字コードとする。
- ・ 符号化形式は UTF-8 または UTF-16 とする。

文字コードを単に UCS-4 などと指定しただけでは、中国・台湾の漢字も入力対象に含まれるようになります。特に戦前の資料など旧字体が多い資料を入力した場合に、見かけがより近いということで、入力者によっては日本の文字コードでは入力できない漢字を使ってしまうことが考えられます。

これらの漢字は、検索の際に支障が起きる場合があります。システムの機能で字形の差を吸収できればよいですが、必ずしも希望通りに吸収できるとは限りませんし、追加機能としてコストがかかることも考えられます。文字コードは UCS-4 にしておいたとしても、実際に使う範囲は日本の漢字の範囲内にしておいた方が無難だといえます。ただし、扱う資料に中国・台湾のものが多いなどの事情があれば、使用を検討する必要があります。

このようなテキストの入力仕様については、メタデータを実際に作成する時に必要となりますので、個々のデジタルアーカイブの事情を考慮しながら、あらかじめ決めておくことを推奨します。「国立国会図書館 資料デジタル化の手引 2011 年版」61 ページ 参考資料 2「目次のテキスト化」には目次データのテキスト化に当たり、文字コードや入力する文字に係るルールの例が記載されています、入力仕様として利用できる部分が多いため、以下に参考のため記載します。

文字コード・字体・旧字

目次データの作製の前に、入力する文字に係るルールを定める必要がある。次にルールの例を示す。

- (1) 文字コードは、Unicode (UTF-8 符号化) 及びデータ中の区切りとして使用する LF (U+000A) を使用する。CJK 互換漢字及び CJK 互換漢字補助は、CJK 統合漢字に置き換えて入力する。ただし、対応する文字が無い U+FA0E、U+FA0F、U+FA11、U+FA13、U+FA14、U+FA1F、U+FA21、U+FA23、U+FA24、U+FA27、U+FA28、U+FA29 は除く。
- (2) 原則として、旧仮名遣いやカナなどを含め、表記のとおりに入力すること。ただし、漢字の異体字、記号付きアルファベット等で表記のとおり入力できない場合は、置き換えても意味を損ねない同等の文字と判断できる文字に置き換える。
- (3) ローマ数字は、半角アルファベットの組合せに置き換えること。例えば、「Ⅱ」は「II」というように「I」を 2 文字使って表記する。小文字の場合は、小文字で「ii」と表記する。
- (4) 判読不能な文字については、該当する本文ページの標題を参照し、字形を同定する。また、元の文字が意味の上から明白な場合、類推作業を行う。その結果、字形を同定又は類推できた文字については、当該文字を入力し、[] で囲んだ形に置き換える。判断がつかない文字については、白四角 (□) を用い、該当文字数の分だけ並べる。文字数不明の場合は、「□・□・□」と入力する。文字コードにないものについては、「=」で置き換えて入力する。
- (5) 上記のほか、○や□で囲む文字、記号 (数式等を含む)、合成文字等についても入力の基準を

⁴⁹ デジタル写真における「生」(未加工)の状態のこと (IT用語辞典 BINARY: <http://www.sophia-it.com/content/RAW> (2011 年 12 月 29 日確認))

⁵⁰ http://www.archives.go.jp/law/pdf/da_100118.pdf

④解像度の検討

作成するデジタルデータの用途に応じて、スキャン時の適切な解像度を決めておきましょう。「国立国会図書館 資料デジタル化の手引 2011年版」での例と、慶應義塾大学メディアセンターでの「デジタル複製」の例を記載します。

① 保存用画像

次のいずれかの値に従って画像を作製する。

- ・原資料に対して 300~400dpi
- ・画像の縦又は横が 2,000~6,000 ピクセル程度

② 提供用画像(中間サイズ)

次のいずれかの値に従って画像を作製する。

- ・原資料に対して 300~400dpi でスキャンした保存用画像を圧縮して作製する。
- ・通常、保存用画像の縦・横 4 分の 1 程度のサイズで、画像の縦又は横が 500~1,000 ピクセル程度の画像を作製する。

③ 提供用画像(拡大サイズ)

次のいずれかの値に従って画像を作製する。

- ・原資料に対して 300~400dpi でスキャンした保存用画像を圧縮して作製する。
- ・通常、保存用画像の縦・横半分程度のサイズで、画像の縦又は横が 1,000~5,000 ピクセル程度の画像を作製する。

④ サムネイル画像

通常、画像の縦又は横が 100~300 ピクセル程度の画像を作製する。

デジタル複製・印刷について：慶應義塾大学メディアセンターの場合

一次資料と同じサイズで、高い精度での印刷をすることを前提とした場合、原寸に対し 400dpi が確保できることを目安とする。

高精細な複製の手順については、カラーマネジメントの方法や出力装置からの要求、入力装置の性能、撮影時の状況に応じて変化させる必要があり、一概には規定できない。

ICC プロファイル等のコントロール手順を理解し手順どおり行うことを前提として以下のような対応が必要となる。

- (1) デジタル化作業者は、ディスプレイの調整 (calibration) を行い、高演色蛍光灯の使用、外光遮断等によって室内環境を整える。
- (2) 入力機器のプロファイルを作成する。
- (3) 画像の撮影を行う。
- (4) 撮影した画像を印刷する場合は、印刷業者へ上記環境を伝え、AdobeRGB への変換を行い、データを渡す (印刷の場合は CMY へ変換することもある)。

Web 用途の場合は sRGB へ変換することが多い。

資料が公開されてから技術も変化しています。技術の変化に合う適切な解像度を利用して、各種の画像を作製することを検討しましょう。

巻末(「参考資料 6 画像コンテンツのパラメータ決定フロー」)に、選択肢ごとの判断の目安を記載していますので、解像度などを検討する時の参考にしてください。

⑤管理データの検討

デジタルデータを長期的に利用・保存して行くためにも、デジタル化した時の情報は重要な情報となります。保存した時の情報が詳細にわかると、再度その保存データを使ってデジタル化する際にオリジナルに近い状態で復元できる可能性が高まります。また、現在のフォーマットから新しいフォーマットへの変換を行う上でも重要な情報として利用が期待されています。

項目例としては、画像データ自体に関するもの (フォーマットタイプや圧縮方法、カラースペース⁵¹など)、もう一つが画像データの作製に関するもの (所蔵者、スキャナ ハードウェア、スキャナ ソフトウェアバージョン、作成日) が例として挙げられています。

上記の他に「国立国会図書館 資料デジタル化の手引 2011年版」では、画像フォーマット別の圧縮率、カラースペース、階調、スキャニング単位の定義、媒体などの記載があります。

管理データの項目を考える際には、そのデータを公開するかどうかを併せて決めておくようにしましょう。管理データは、公開のために作るメタデータとは性質が異なるものです。たとえば美術館や博物館で、所蔵品を新たに購入した場合にその購入元や価格を記録しておくことが考えられますが、そのような情報を無制限に公開する必要はないと考えられます。同様に、寄託品などで寄託者のご住所や連絡先を記録することも考えられますが、そのような個人情報は一般的に公開すべきでないものです。

ここまでがデジタル化を行うために必要な知識と仕様書を記載する際の作業の概要となります。参考としてあげた資料を適宜参照しながら、自館の要求に沿った仕様書を準備しましょう。

なお、自館で作成した仕様書に基づいてデジタル化作業をする企業が、その館の要求に合ったデータを本当に作製可能か事前に検証しておくことが望ましいと言えます。

評価項目の例として、東京大学大学院情報学環と凸版印刷株式会社との共同研究プロジェクトにおいて作成された、「文化資源のデジタル化に関するハンドブック」⁵²には、評価項目の例が挙げられていますので、参考までに記載します。

⁵¹ 各色を数値の組合せで表現する方法または表現可能な色域のこと。色空間とも呼ばれる。

(IT用語辞典 BINARY)

<http://www.sophia-it.com/content/%E3%82%AB%E3%83%A9%E3%83%BC%E3%82%B9%E3%83%9A%E3%83%BC%E3%82%B9> (平成 24 年 1 月 23 日確認) を参考に記述した。)

⁵² 東京大学大学院情報学環 凸版印刷株式会社 文化資源のデジタル化に関するハンドブック 詳細版 P53 http://www.center.iii.u-tokyo.ac.jp/wp-content/themes/mssa-default/images/111101_dch_shousai.pdf (2012 年 1 月 31 日確認)

主観評価の実施 評価項目

画像の評価は例えば以下のような観点から評価を行います。

- ・ 色調
- ・ 諧調
- ・ 解像度
- ・ 形状
- ・ ノイズ
- ・ 鮮鋭度

デジタル化では特に色の管理について注意が必要となるため、「文化資源のデジタル化に関するハンドブック」の P45 カラーマネジメントも参照するとよいでしょう。

(b) デジタル化作業

デジタル化作業は委託業者側の作業が主となることが多いですが、発注者側でも作業の概要や流れを押さえておくことは重要です。ここでも、「国立国会図書館 資料デジタル化の手引 2011 年版」を参照して作業項目を中心に例を挙げます。

- ①原資料の授受・運搬
- ②事前調査
- ③スキャニング
- ④デジタルデータの品質検査
- ⑤デジタルデータの編集
- ⑥管理データの作製
- ⑦データの格納媒体の選択とチェック方法

本項での作業で注意する点として、以下の 2 つの項目を挙げることができます。

④デジタルデータの品質検査

発注者側での対応が必要となる作業です。前項でも記載した通り評価項目を事前に決めておきましょう。

⑥管理データの作製

デジタルデータを長期利用・保存するためにも、重要な情報となる管理データ（デジタル化した時のメタデータ）がきちんと作製されていることが重要です。各機関で必要な管理データ項目を決めて、作製しましょう。巻末の「参考資料 1 管理データの項目一覧」に管理データ項目の例を記載しましたので、参照してください。

参考：OCR の仕様の考え方

OCR (Optical Character Recognition, 光学文字認識) とは、画像データ上にある文字と思われる部分を解析し、コンピュータ上で扱える文字(テキスト)データに変換することを示します。※1

国立国会図書館のデジタル化ではこの処理を行っていませんので、紹介した手順には出てきませんでしたが、特に書籍をデジタル化するには有益な手法です。

慶應義塾大学メディアセンターでの実例を紹介します。

- ・原則として、OCR では 1 文字あたりに要求される DPI 数からスキャニング仕様が決まる。
- ・また、ルビ、欄外、記号をどのように扱うかによっても仕様を決める必要がある。
- ・利用する OCR ソフトウェアによって特性があるため、日本語 OCR の精度の評価は難しい。
サンプル処理による評価を行い、スキャニング仕様へ反映させる必要がある。

※1 出典：朝日新聞「知恵蔵 2009」

(2) メタデータスキーマを整理する

(a) メタデータとは

メタデータは、一般に「データに関するデータ」と定義されます。

博物館、美術館では所蔵品の目録を作成しています。所蔵品そのものではなく、所蔵品に関するデータである目録の情報は、メタデータと同じ意味と考えられます。

メタデータが必要な理由の例として、目の前にラベルの付いていない裸のボトルがあることを想像してください。ボトルの中身を知るためには、常にボトルの中の液体を飲む必要がありますが、何も書いていないボトルの中の液体を飲めるでしょうか。しかし、適切なラベルが付いていれば、飲まずとも中身を知ることが可能であり、また安心して飲むことができます。同様のことが、中身が全く分からないファイルや、ウェブページにも当てはまります。ネット上では、自分が使うものに関する情報、すなわちメタデータを頼りにしているいろいろなことを行っているのです。絵画と本では目録の項目が違うように、対象物によってメタデータの項目が違う場合もあれば、人によってもつけ方が違います。従って、ある一定の規則を決めてメタデータを作成する必要があります。そのメタデータの記述項目や記述形式のことをメタデータスキーマと呼びます。

(b) メタデータスキーマを決める

一つの組織（機関）内でメタデータスキーマを決めることはさほど難しくはありません。業務の中で目録を作成する等と同様のことを行っているケースが多くあります。

しかし、デジタルアーカイブの連携や統合的利用を可能とするためには、連携可能性をベースとしたメタデータスキーマを決めることが求められます。すべてを 1 つのメタデータスキーマに統一することは難しいですが、様々な機関においてメタデータの標準化が図られており、表 5-5 のようなメタデータスキーマを参考にしながら、自館のメタデータ

キーマを決めておきましょう。

メタデータ自体が整備されている場合には、既存のデータを、新たに決めるスキーマに当てはめる作業が必要となります。できればコンテンツごとに一意で重複のない ID を設定することが望ましいです。

以下の表 5-5 に機関別メタデータスキーマ一覧を記載します。

表 5-5 機関別メタデータスキーマ一覧

機関名	システム名	メタデータスキーマ	参照先
東京国立博物館		ミュージアム資料情報構造化モデル (資料情報の属性一覧)	参考資料 2
国立国会図書館	国立国会図書館サーチ (NDL Search)	国立国会図書館ダブリンコア	参考資料 3
	国立国会図書館の デジタル化資料	メタデータ記述 (DC-NDL)	
国立公文書館	国立公文書館デジタル アーカイブ	国立公文書館 EAD 定義 第 1.07 版	国立公文書館 ホームページ ⁵³
文化庁・N I I	文化遺産オンライン	文化遺産オンライン 作品情報(文化遺産情報)項目一覧	参考資料 4

また、平成 22 年度 総務省「新 ICT 利活用サービス創出支援事業」にて構築されたメタデータ情報基盤システムである「MetaBridge」では、様々な機関のメタデータスキーマの例が登録されています。自館のメタデータスキーマを作る上で参考となる情報も多くありますので一度参照して見ると良いでしょう。

メタデータスキーマは各館によって異なる項目があってもよいのですが、連携(主に検索用)に必要な項目は必ず含めておくことを推奨します。以下に、平成 21 年 3 月に国立公文書館が公開している『デジタルアーカイブ・システム標準仕様書』⁵⁴の横断検索のために必要な最小限の目録情報として挙げられている 5 項目を示します。この 5 項目は他システムへ連携するための項目としても十分参考となる項目です。

Dublin Core の要素を参考に、横断検索のために必要な最小限の目録データの項目として、「ID」「年代(作成年度)」「資料名」「作成者(部署名)」「備考等」の 5 項目を設定する。

上記以外に、最近では「時間」「場所(緯度・経度)」なども有効な検索項目と言われており、文化遺産オンラインでは「場所(緯度・経度)」をメタデータの項目として設けています。

⁵³ デジタルアーカイブ 関連資料 <http://www.digital.archives.go.jp/support/reference.html> (2012 年 1 月 20 日確認)

⁵⁴ http://www.archives.go.jp/law/pdf/da_100118.pdf

(3) 基本機能・連携機能を確定する

自館の所蔵品のデジタルデータ作成及びメタデータの準備ができました。最後にインターネットで公開するための仕組みが必要となります。

構築すべきシステムの仕様は、システムの目的や収録するデータの量、予想されるアクセス数などにより大きく異なるため、統一的なガイドラインを示すことは困難です。

ここでは、デジタルアーカイブの機能仕様について平成 21 年 3 月に国立公文書館が公開している『デジタルアーカイブ・システム標準仕様書』から機能に関連する部分を参照し記載します。

(a) システムの基本機能を確定する

① 目録データの管理・登録機能

・ データベース

目録データの登録や管理については、データの編集や抽出を効率的に行うことが求められます。この点からリレーショナル・データベース⁵⁵のようなデータベースを導入することが有効であると考えられます。商用ソフトは費用的な面での負担も大きいと想定されるため、オープンソースのデータベースを採用することが推奨されています。

・ 検索エンジン

登録された目録の項目ごとに検索を行えることと、アーカイブされた対象物に文字情報があれば、検索文字からその対象物の全文検索が可能な検索エンジンを採用することが推奨されています。

・ データの登録、編集、削除機能

目録データの登録、編集、削除機能は必要不可欠な機能です。また多くの機関で表計算ソフトなどを用いて目録データを作成しているため、その表計算ソフトで作成した目録データあるいは CSV (Comma Separated Values、カンマ区切り形式) などで一括登録が可能な機能を設けることが推奨されています。

・ 公開・非公開機能

著作権などの様々な要因で所蔵資料の公開については、注意する必要があります。これを踏まえ、目録データに公開・非公開の制御ができる仕組みが必要と考えられます。

② 業務機能

・ 利用状況の把握機能

利用状況を把握することは利用者側の潜在的な要求を捉える上で重要な機能と考えられます。対象物に対してのアクセス回数やアクセス元の情報(国別など)が確認できる仕組みが必要と考えられます。

活用例としては、美術館では次の展示会への出展候補を決めたり、博物館ではコーナーの新設などの情報としての利用も考えられます。

⁵⁵ リレーショナルデータモデルに基づいて設計されたデータベース。データの結合や抽出を容易に行うことができる。(IT用語辞典 e-Words: <http://e-words.jp/w/HD:2.html> (平成 24 年 1 月 23 日確認) を参考に記述した。)

③画像データの登録・管理機能

- ・ 画像データの公開・非公開機能

画像データも目録情報と同様に著作権などの様々な要因により注意が必要です。画像データに公開・非公開の制御ができる仕組みが必要と考えられます。

- ・ 画像データ・目録データの連携機能

目録情報を閲覧した際に関連する画像データが閲覧できるなどの目録データと画像データを連携できる仕組みが必要と考えられます。

- ・ 画像データの配信・閲覧機能

画像データの配信・閲覧機能は、特別なソフトウェアを必要とせず表示できる仕組みが必要と考えられます。

④情報配信に必要な機能

Web ブラウザから資料の検索や参照を可能とする機能が重要と考えられます。

以下の点も参考にして構築しておく、よりシステム連携が容易になると考えられます。

個別の資料のページはシステムのトップページや検索からだけではなく、外部から直接参照できるようにしておくことが望ましいです。このためには外部から直接指せるような URL を使い、システムの更新などでも変更しないような安定した URL が望ましいです。

Web ページの記述においては、内容（メタデータ）の構造とページの見目の体裁のための構造を分けて記述すると、後述の多様な閲覧デバイス対応が容易になります。

また、Web ページの他に、メタデータを RDF(Resource Description Framework)で記述したものを公開すると、他のサービスからのメタデータの利用やシステム連携が容易になります。

⑤ユーザ・インターフェース

ユーザ・インターフェースの基本は誰もが使いやすいことです。キーワードによる検索は登録されているデータの前提知識があればよいですが、必ずしもそうとは限りません。階層構造を表現したユーザ・インターフェースを採用することも必要と考えられます。また、今後はスマートフォンやタブレット端末など PC 以外での様々な閲覧デバイスからのアクセスが増えてくると考えられます。多様化する閲覧デバイスにも対応した、使いやすいユーザ・インターフェースを採用することも検討しておきましょう。

(b) 連携機能を確定する

デジタルアーカイブを構築・公開することは重要ですが、一つの機関だけでは多くの人からのアクセスには限界があります。それを様々な人からアクセスしてもらうためにデジタルアーカイブ同士を連携させることが今後重要と考えられます。

但し、どのデジタルアーカイブと連携するかに応じて、必要となる環境や準備する内容

が変わってきますので、ここでは連携するにあたり、検討すべき項目や機能について記載します。

連携機能を構築後に別途追加することも可能ですが、検討は最初に行っておくことを推奨します。また、費用や開発効率を考えると、システムの基本機能の構築と同じタイミングで構築することが望ましいと言えます。

①連携機能の検討

連携する機関に応じたシステムの情報を収集しておきましょう。連携データの形式は連携先のシステムに依存するため、その機関がどのようなメタデータ形式を使っているのか、一括のデータ登録ツールがあるのかなど、少し細かな点についても確認しておくことで連携をスムーズに進めることができます。

以下に連携機能の種類とその説明を記載します。

- ・ 横断検索機能

横断検索機能とは 1 度の検索キーワードの入力で複数のシステムに対して検索命令を送り、その結果を表示する機能です。横断検索機能を利用するためには横断検索に対応したプロトコルを実装する必要があります。

横断検索プロトコルとしては、国際標準の Z39.50、SRU/SRW、OpenSearch といった種類のプロトコルがあります。

- ・ 一括提供機能

主にメタデータを他のシステムへと連携させる機能を指します。

メタデータ交換のためのプロトコルに、OAI-PMH(Open Archives Initiative Protocol for Metadata Harvesting)があります。国立国会図書館の国立国会図書館サーチや、国立情報学研究所の学術機関リポジトリデータベースにも導入されています。

表 5-6 に、利用されるプロトコルの一覧を記載します。

表 5-6 収集、検索に利用される主なプロトコル⁵⁶

種別	プロトコル	説明
収集	OAI-PMH	メタデータを自動収集（メタデータ・ハーベスティング）し、データを交換するためのプロトコル Open Archives Initiative・Protocol for Metadata Harvesting の略称
検索	Z39.50	情報検索を行うための通信プロトコル ANSI/NISO 標準及び、ISO 国際標準（ISO 23950） 米国議会図書館が標準を管理している
	SRU/SRW	情報検索を行うための通信プロトコル Search/Retrieve via URL/Search/Retrieve Web Service の略称
	OpenSearch	情報検索を行うための通信プロトコル 米アマゾン子会社の A9.com 社が提唱する検索エンジン提供者のための技術で XML（Extensible Markup Language）フォーマットを利用

(4) デジタルアーカイブを構築する

(a) システム基盤を確定する ―クラウドサービスの検討

これまでではサーバやソフトウェアを購入しそれを自館内やデータセンター等に設置する「オンプレミス」と呼ばれる形態が一般的でした。様々な外部環境の変化もあり、現在は「クラウドコンピューティング」の利用が急速に進んでいます。

クラウドという用語はさまざまな意味で使用されていますが、ここでは、情報処理推進機構（IPA）の「中小企業等におけるクラウドの利用に関する実態調査報告書」⁵⁷の定義に従うことにします。以下に引用します。

⁵⁶ 平成 21 年度 委託調査「我が国におけるデジタルアーカイブの構築に関する調査研究」調査報告から

⁵⁷ 独立行政法人情報処理推進機構「中小企業等におけるクラウドの利用に関する実態調査報告書」2011 年 5 月 30 日、http://www.ipa.go.jp/security/fv23/reports/sme_guide/documents/sme-cloud_report.pdf（平成 24 年 1 月 31 日 確認）

表 1 「クラウド」の定義

用語	定義
クラウド	クラウドコンピューティング、クラウドサービスを区別することなく示す場合に用いる総称
クラウドコンピューティング	大規模データセンターにおいて仮想化等の技術を用いてコンピュータの機能を用意し、それをインターネット経由で自由に柔軟に利用する仕組みの総称
クラウドサービス	クラウドコンピューティングに基づいて、サービスの形で提供される IT の機能

つまり、SaaS・PaaS・IaaS 等、ユーザに提供される機能・サービスをクラウドサービス、クラウドサービスを構成する仕組み、技術、サービスの総称をクラウドコンピューティングと定義する。

それぞれに利点・欠点がありますが、オンプレミスの場合は、その館のニーズに合ったハードウェアやソフトウェアを導入できる利点がある反面、コストが高額となる傾向が強い点が欠点として挙げられます。

クラウドサービスでは、提供される機能の範囲内であれば、自館で構築・運用するよりもコストを低く抑えられる利点があります。「中小企業等におけるクラウドの利用に関する実態調査報告書」の、クラウドサービス導入後の利用者に対する調査結果によると、「導入費用（設備費用）が小さい」「運用コストが小さい」の評価で「満足度高」が過半数を占めており、コストメリットでの評価が高かったことが報告されています。

しかし、サービスの提供がいつまで継続するか分からない等、サービスを提供する企業側の体制面などの課題が考えられます。自館の要望に合った機能を備え、かつ信頼性の高いシステム基盤を選択しましょう。

メモ：クラウドコンピューティングの種類

SaaS (Software as a Service) ユーザが必要とする機能を必要な分だけサービスとして利用できるようにしたソフトウェア、またはその提供形態
PaaS (Platform as a Service) アプリケーションソフトウェアが稼働するためのハードウェアや OS などのプラットフォームをサービスとして利用できるようにした提供形態
IaaS (Infrastructure as a Service) コンピュータシステムを構築および稼働するためのハードウェア(サーバやストレージ)やネットワークをサービスとして利用できるようにした提供形態 IST(米国立標準技術研究所)の定義より

第 5 回知のデジタルアーカイブ研究会「クラウドの視点から見たデジタルアーカイブ」発表資料より

(b) システム開発手法を確定する

システム構築に関するもう一つのポイントは、システム開発の手法です。

自館の要件に合わせて一から開発する「スクラッチ開発」と、既にできあがったソフトウェアをそのまま使用したり、若干の修正を加えて利用する「パッケージ開発」の、大きく分けて2つの選択肢があります。

クラウドサービスを利用する場合、IaaSはスクラッチ開発に、SaaSはパッケージ開発に適用できます。ASPサービスもこのパッケージ開発のケースになります。PaaSは両者の中間に位置します。

パッケージ開発の利点は、スクラッチ開発に比べて安価に、また導入期間をかけずに使用を開始できる点です。ただし、ファイルフォーマットやデータ形式があらかじめ限定されていることが多いため、利用するパッケージがデジタル化で検討したファイルフォーマットやメタデータ形式を採用できるのかなどの注意が必要です。

以下に、デジタルアーカイブにおけるクラウドのメリットと今後の課題について「知のデジタルアーカイブに関する研究会」で報告された内容を参考情報として記載します。

<デジタルアーカイブにおけるクラウドのメリット>

- ・ 基盤サービスとしては選択肢となる。
(ストレージサービス、バックアップサービス、インフラサービス)
- ・ 初期投資の削減
- ・ 余剰資産が不要
- ・ 運用コストの削減

<クラウドの課題>

- ・ 事業者やサービスによって標準化されていないため、システムの移行が困難
- ・ サービスやセキュリティの指標が統一されていない

第5回知のデジタルアーカイブ研究会「クラウドの視点から見たデジタルアーカイブ」発表資料より

(c) システムを構築する

各機関の状況により、調達範囲は流動的になると考えられるため、何をどの範囲で調達・委託するのか、また自館で行う作業は何かを明確にして依頼を行いましょう。

システムを構築する際に注意する必要があるポイントとして、以下が挙げられます。

- ・ デジタル化したデータの導入方法
- ・ 旧システムからのデータの移行作業方法と役割
- ・ システム利用方法の教育

3. デジタルアーカイブを未来に伝えよう：継続・人材・保存

(1) 計画の再検討

デジタルアーカイブは構築して終わりではなく、継続していくことが重要です。そのため、数年単位の計画を立てることが望まれます。最初に立てた計画も構築後に定期的に見直しを行い、出て来た課題等を含めて再度計画を立て直すサイクルを確立することが重要です。

デジタルアーカイブを作ることがそれぞれの館での目的ではありません。したがって、デジタルアーカイブを無理なく継続して続けていくために、日常の業務の中で対象物のデジタル化や所蔵目録作成の際にデジタルアーカイブの元になる資源を作成するような仕組みを作ること(業務の改善)も検討してみてください。

(2) 人材と教育

人材については、どの機関でも共通の課題として挙げられている項目です。また根本的な解決策がなかなかないこともこの問題の特徴と言えます。

デジタルアーカイブという長期的な事業を継続するためには、「人的資源」をどのように確保し育てていくかがデジタルアーカイブを支える一つの要素であると考えられます。現時点で具体的な解決方法を示すのは難しいですが、いくつかのポイントを記載します。

(a) 人材の育成

さまざまな支援を外部から得たとしても、それらを管理・指導する人材が組織内にいなければ、継続的な活動につながらないことは容易に想像がつきます。現状では、専任の担当者を確認することも難しく、デジタルアーカイブの専門的な知識を持っていない職員にもその役を兼務してもらわざるを得ないことが予想されます。

資料のデジタル化やデジタルアーカイブに関する研修は、国立国会図書館、国立公文書館で行われることがあります。その他、各都道府県などの単位で開催される研修や勉強会の機会を活用し、知識の習得に努めることが望ましいといえます。

以下に、平成23年度に、国立国会図書館で実際に行われた「資料デジタル化研修（基礎編）」の実施要項⁵⁸を示します。

平成23年度資料デジタル化研修（基礎編）のご案内

国立国会図書館では、公共図書館の職員等を対象に資料デジタル化研修（基礎編）を開催します。

この研修は、これからデジタルアーカイブを構築する機関の職員や、新たにデジタルアーカイブの担当となった職員が、資料のデジタル化についての基礎的な知識を習得し、所属機関のデジタル化事業に役立てることを目的とします。

具体的には、次の3点の習得を目指します。

- (1) デジタルアーカイブの全体像をつかむ。
- (2) デジタルアーカイブを構築するうえでの事業の進め方を理解する。
- (3) 資料をデジタル化する際に必要となる、技術や権利についての基本的な知識を得る。

日程： 平成23年9月29日（木）、30日（金）

会場： 国立国会図書館関西館 第1研修室

対象： 公共図書館等の職員または地方公共団体職員で、「資料のデジタル化」についての基礎知識の習得を目指す者。1機関からの参加は原則1名。定員50名。

(b) 人材の確保

震災復旧のためのボランティアによる支援がメディアでも取り上げられ、注目を集めていました。デジタルアーカイブの運営に関しても、ボランティアの方々による支援を受けながら進めていくことが考えられます。

その地域の歴史にも詳しい高齢者の方もいれば、カメラを趣味に持つ方もいます。必要な知識を持っている地域住民の協力を念頭に置いた、住民参加型のデジタルアーカイブを目指すことも1つの方法です。

この場合職員にとって重要な能力は、企画を行い、ボランティアの支援を募り、取りまとめる力であるといえます。

(c) 組織と予算の維持

海外での事例により、複数の機関が連携する共同体制の構築が有用であると示されています。一部の特定の機関に依存するのではなく、組織の違いを超えた共同・協力のネットワークを組み、共同事業における予算確保、や複数の組織として全体で情報を維持する体制を作るなど、いわゆる地域のコミュニティ作りも重要な取組と言えます。

⁵⁸ 平成23年度資料デジタル化研修（基礎編）のご案内
http://www.ndl.go.jp/jp/library/training/guide/1191881_1485.html（2012年1月20日確認）

(3) 長期利用・保存

以下に、『電子情報の長期的保存とアクセス手段の確保のための調査報告書』⁵⁹で、実際に情報を読み取ることができなくなった事例などを参考に、長期保存にあたっての一般的な留意点を挙げます。

(a) 媒体の寿命に対する保存サイクル・バックアップの検討（マイグレーション⁶⁰）

慶應義塾大学メディアセンターでの取組例
デジタル化作業の中に保存サイクルを考えて実施している。

- ・撮影したデータはPCの中で保存・加工を実施している。
- ・作業が一区切りした段階で、①共有ディスク装置(RAID5)に移す。
- ・作業が一通り終わったら、②DVDと③バルク（「ばら積み品」という意味で、パソコンの部品や周辺機器のうち簡素な包装で販売されている形態を指す）のハードディスクにバックアップをとる。

期間は特に決まっていないが、数年毎で②の媒体は廃棄して、①共有ディスク装置のデータから再度新しい媒体(その時代での最適なメディア)にバックアップする。

デジタル情報資源は、記録されている媒体の経年劣化により、データが読み出せなくなることがあります。一般に10年程度とされていますが、媒体の品質や、記録時の方法、保存環境により大きく異なるため、寿命を迎える前に、媒体を網羅的かつ定期的に交換するサイクルを業務の中に位置づけることが望ましいです。また、情報資源作成時に、複製物を作成し、バックアップをとることも重要な対策です。

(b) ハードウェア・ソフトウェア環境の保存・更新（マイグレーション/エミュレーション⁶¹）

デジタル情報は、常に人間が知覚できる状態にある紙媒体の情報とは異なり、特定のハードウェア（読み取り装置）・ソフトウェアの環境が揃わなければ、人間が知覚可能な情報にはなりません。しかし、ハードウェア・ソフトウェアとも、技術的進歩等に伴って非可逆的に更新され、デジタル情報資源が利用可能な環境を、長期的に維持することは困難です。

ハードウェアの更新に対しては、後継となる記録媒体へのマイグレーションを、記録媒体の寿命に対する措置と同時に検討することが望ましいです。また、ソフトウェアの更新

⁵⁹ 国立国会図書館『電子情報の長期的保存とアクセス手段の確保のための調査報告書』2004年。

⁶⁰ データ移行・データ変換のこと。（国立国会図書館 電子情報の長期的な保存と利用のFAQ “マイグレーション” “エミュレーション”とは何ですか？ http://www.ndl.go.jp/jp/aboutus/preservation_03.html#q09（2011年12月29日確認）を参考に記述した。）

⁶¹ 電子情報の利用に必要な技術的環境を擬似的に再現すること。（国立国会図書館 電子情報の長期的な保存と利用のFAQ “マイグレーション” “エミュレーション”とは何ですか？ http://www.ndl.go.jp/jp/aboutus/preservation_03.html#q09（2011年12月29日確認）を参考に記述した。）

に対しては、後継や代替となるソフトウェアに対応したファイルフォーマットへのマイグレーションと同時に、エミュレーションによる対応もあり得ます⁶²。

(c) 大規模な災害に備えたバックアップデータの分散保管

地震や水害の多発する日本において、大規模な災害や火災による記録媒体そのものの消失を無視することはできません。複製が容易なデジタル情報資源の特性を生かし、バックアップデータを遠隔地にも置き、消失のリスクを分散することは、長期保存の重要な対策として検討されるべきです。ただし、データを全て海外に置くことには、メリットとデメリットとを考慮し、慎重に考えることも必要です。

(d) 長期利用・保存のためのメタデータ

デジタルアーカイブ構築後に長い時間が経過すると、構築当初の目的やデータ作成の日時、使用技術等の情報が失われ、長期利用・保存に支障を生じることがあります。これらの情報の保存について検討しておくとい良いでしょう。本ガイドライン 54 ページで説明したように、デジタルデータを長期利用・保存するためにも、重要な情報となる管理データ（デジタル化した時のメタデータ）がきちんと作製されていることが重要になります。

(4) アウトリーチ

アウトリーチとは「手を伸ばして取る、手を差し伸べる」などという意味で、福祉などの分野における地域社会への奉仕活動、公共機関の現場出張サービスなどの意味で使用されています。

デジタルアーカイブにおいても、「構築して終わり」にするのではなく、広く活用してもらうための活動を行うことが望ましいと考えられます。

広く活用してもらうための活動の一つとしては、広報活動が挙げられるでしょう。秋田県立図書館では古くからデジタルアーカイブの構築を行った機関として様々な所で紹介されています。図書館の利用者への PR、ビジネス支援サービスの中での紹介に加えて、教育委員会や学校にデジタルアーカイブの活用を働きかけるなど、館外への広報活動による認知度向上の重要性を認識し、うまく活用している事例の一つと言えるでしょう。

秋田県立図書館の利活用の事例として、デジタルアーカイブのコンテンツが地域の製品のブランド化と結びついた例を記載します。

秋田の三関というところで取れるさくらんぼで「太陽の分け前」という商品があるが、このパッケージにデジタルアーカイブで紹介されている「菅江真澄遊覧記」の一節が書かれている。

この地区の土壤の良さがうたわれている記述を引用することで、江戸時代の頃から語り継がれてきた歴史を一つの付加価値として、さくらんぼのブランド化に役立てている。

デジタルアーカイブで公開したことにより、地域資料がビジネスに活用された事例である。

（前略）三関のサクランボの美味は、日本でも有数。またこの地区の土壤の良さを「旭さし、夕日輝く木のもとに、黄金千両漆徳おく」と菅江真澄が「雷の出羽路」に記している。これは800年のむかし、平泉の藤原秀衡の時代に、金鶏山のうたいし童謡にあったのを、雄勝地方で土の良さをうたったものである。（後略）



知のデジタルアーカイブに関する研究会（第1回）
「地域資料のデジタルアーカイブの現状 ー秋田県立図書館の事例からー」の発表より

⁶² 内閣府『平成20年度電子公文書等の管理・移管・保存・利用システムに関する調査』2009年に、一般的なファイルフォーマットごとに長期保存の対応策が提案されている。

参考資料 1 管理データの項目一覧

出典：『国立国会図書館 資料デジタル化の手引 2011 年版』

参考資料 3 デジタル化仕様書サンプル 表 9 管理データ(2)の項目一覧

項目名	説明	記入内容
MIMETYPE	画像データに関連付けられた MIME タイプの名称を入力する。	image/jp2
SourceType	画像データを作製するためにスキャンされたアナログの資料の媒体を指定する。	book
ImageProducer	画像データの組織レベルプロデューサーを特定する。	受託業者名を入力する。
HostComputer	画像データの作製時点で使用したコンピュータ名を入力する。	画像データの作製時点で使用するコンピュータ名を入力する。
OperatingSystem	画像データの作製時点で使用したオペレーティングシステムを入力する。	画像データの作製時点で使用するオペレーティングシステムを入力する。
OSVersion	画像データの作製時点で使用したオペレーティングシステムのバージョン番号を入力する。	画像データの作製時点で使用したオペレーティングシステムのバージョン番号を入力する。
ScannerManufacturer	画像データの作製に使用したスキャナのメーカー名を入力する。	画像データの作製に使用したスキャナのメーカー名を入力する。
ScannerModelName	画像データの作製に使用したスキャナの機種名を入力する。	画像データの作製に使用したスキャナの機種名を入力する。
ScannerModelNumber	画像データの作製に使用したスキャナの型番を入力する。	画像データの作製に使用したスキャナの型番を入力する。
ScannerSoftware	画像データの作製に使用したキャプチャソフトウェア名を入力する。	画像データの作製に使用したキャプチャソフトウェア名を入力する。
ScannerSoftwareVersionNo	画像データの作製に使用したキャプチャソフトウェアのバージョン番号を入力する。	画像データの作製に使用したキャプチャソフトウェアのバージョン番号を入力する。
DateTimeCreated	画像データを作製した年月日を入力する。「YYYY-MM-DD」と年月日を入力する。	入力する年月日は納入期限日とする。
DateTimeProcessed	画像処理した年月日を入力する。「YYYY-MM-DD」と年月日を入力する。	入力する年月日は納入期限日とする。
ProcessingAgency	画像処理した画像データの組織レベルプロデューサーを特定する。	受託業者名を入力する。
ProcessingSoftwareName	画像データを編集又は変換するのに使用した画像処理ソフトウェア名を入力する。	画像データを編集又は変換するのに使用した画像処理ソフトウェア名を入力する。
ProcessingSoftwareVersion	画像データを編集又は変換するのに使用した画像処理ソフトウェアのバージョン番号を入力する。	画像データを編集又は変換するのに使用した画像処理ソフトウェアのバージョン番号を入力する。

参考資料 2 資料情報の属性一覧

出典：『ミュージアム資料情報構造化モデル 資料情報の属性一覧』

性格	属性名	役割
識別・特定	1 識別子	記述単位を一意に識別する記号、番号。
	2 資料番号	組織によって資料に付された記号、番号。
	3 名称	資料の名前、呼称、タイトル。
	4 分類	資料の分野、種別。
	5 用途	民俗・考古資料などで資料が本来持っていた機能。
	6 様式	資料が作られているスタイル、流派。
物理的特性	7 品質形状	「材質」「技法」「形状」をまとめて記述する。 この3つをそれぞれ記述する場合は省略。
	8 材質	資料を構成する材料、材質。
	9 技法	制作に用いられている技法。
	10 形状	資料の形状の類型。
	11 員数	資料の数量、点数。
	12 計測値	数値で表現できる計測値。寸法や重量。
	13 部分	資料の部分、下位の記述単位への参照。
	14 保存状態	資料の保存状態。
	15 付属品	資料に付属する物品。付属文書や箱。
	16 印章・銘記	資料に直接書き込まれた文字や印。
履歴	17 制作	資料の制作、成立に関する情報。
	18 出土・発見	資料の出土、発見に関する情報。
	19 来歴	資料の伝来、所有、使用の歴史。
	20 取得	購入、寄贈などにより資料が管理下におかれることになった際の記録。
	21 整理・処分	移管、売却、破壊、盗難などにより資料が管理下におかれなくなった際の記録。
	22 受入	寄託、借入などにより資料を受入れた際の記録。
	23 調査	資料の調査履歴。
	24 修復	資料の修復履歴。
	25 展示	資料を公開した際の記録。
	26 所在	資料が保管されている場所。収蔵庫、貸出先などを含む。
	27 価格評価	資料に対する価格評価の履歴。
	28 受賞・指定	資料が受けた賞の履歴や文化財指定の履歴。
関連・参照	29 権利	所有権、著作権、複製権など権利についての記述。
	30 関連資料	他の資料への参照。関連する記述単位への参照も含む。
	31 文献	関連する文書、刊行された図書、論文等への参照。
	32 画像	写真などの視覚的二次資料。
	33 記述ノート	その他の情報についての文章による記述。
	34 記述作成	記述の作成者、変更履歴など。

参考資料3 エンコーディングスキーム一覧

出典：『国立国会図書館ダブリュンコアメタデータ記述（DC-NDL 2011年12月版）付録 語彙一覧表』

作品項目入力例 長谷川等伯「松林図屏風」（東京国立博物館所蔵）

性格	No	属性名	記述例
識別・特定	2	資料番号	A10471
	3	名称	松林図屏風
	4	分類	絵画
物理的 特性	9	技法	紙本墨画
	11	員数	6曲1双
	12	計測値	各縦 156.8 横 356.0
履歴	17	制作	長谷川等伯 16世紀
	28	受賞・指定	国宝
関連・参照	33	記述ノート	草稿ともいわれるが、霧に包まれて見え隠れする松林のなげない風情を、…

※実際には、より多数の項目を使い、構造化されたデータで表現されています。

語彙のカテゴリー				掲載場所	
語彙の表示名	プロパティ URI	QName	定義等	NDL Metadata Terms	Applica- tion Profile
1 標準的なプロパティ					
1.1 情報資源の記述に使用する語彙					
1.1.1 複数の記述に共通して使用する語彙					
Transcription	http://ndl.go.jp/dcnl/terms/transcription	dcndl:transcription	読みまたは題字形	10	11
1.1.2 タイトルに関する語彙					
DCTERMS Title	http://purl.org/dc/terms/title	dcterms:title	タイトル(リテラル(変数・関数ではない、文字列・数値等のデータ)で記述する場合)	-	11
DC Title	http://purl.org/dc/elements/1.1/title	dc:title	タイトル(値と誤みをセットで表現する場合)		11
Alternative	http://ndl.go.jp/dcnl/terms/alternative	dcndl:alternative	別タイトル(タイトルの別言語または別文字による表示形)	10	13
Series Title	http://ndl.go.jp/dcnl/terms/seriesTitle	dcndl:seriesTitle	シリーズタイトル	10	14
Series Alternative	http://ndl.go.jp/dcnl/terms/seriesAlternative	dcndl:seriesAlternative	別シリーズタイトル(シリーズタイトルの別言語または別文字の表示形)	11	15
Part Information	http://ndl.go.jp/dcnl/terms/partInformation	dcndl:partInformation	単行書の構成レベルに相当する情報	11	16
Part Title	http://ndl.go.jp/dcnl/terms/partTitle	dcndl:partTitle	単行書の構成レベルに相当する各著作のタイトル	11	17
Volume Title	http://ndl.go.jp/dcnl/terms/volumeTitle	dcndl:volumeTitle	多巻ものの各巻タイトルまたは刊行物の部編名	12	18
Volume	http://ndl.go.jp/dcnl/terms/volume	dcndl:volume	巻次または部編番号	12	19
Alternative Volume	http://ndl.go.jp/dcnl/terms/alternativeVolume	dcndl:alternativeVolume	別タイトル部編番号	12	20
Alternative Volume Title	http://ndl.go.jp/dcnl/terms/alternativeVolumeTitle	dcndl:alternativeVolumeTitle	別タイトル部編名	13	21
Uniform Title	http://ndl.go.jp/dcnl/terms/uniformTitle	dcndl:uniformTitle	統一タイトル	13	22
1.1.3 作成者に関する語彙					
DC TERMS Creator	http://purl.org/dc/terms/creator	dcterms:creator	当該情報資源の作成者(構造化又は URI により表現する場合)	e	23
DC Creator	http://purl.org/dc/elements/1.1/creator	dc:creator	当該情報資源の作成者(文字列で表現する場合)	-	25
Creator Alternative	http://ndl.go.jp/dcnl/terms/creatorAlternative	dcndl:creatorAlternative	作成者の別名または異なる形	14	25
Series Creator	http://ndl.go.jp/dcnl/terms/seriesCreator	dcndl:seriesCreator	当該情報資源が属するシリーズに対し著作権を持つ実体	14	26
Part Creator	http://ndl.go.jp/dcnl/terms/partCreator	dcndl:partCreator	単行書の構成レベルに相当する各著作に著作権を持つ実体	15	27
Volume Creator	http://ndl.go.jp/dcnl/terms/volumeCreator	dcndl:volumeCreator	当該情報資源の当該する巻または部編に対し著作権を持つ実体		28
1.1.4 寄与者に関する語彙					
DCTERMS Contributor	http://purl.org/dc/terms/contributor	dcterms:contributor	情報資源の整理に何らかの寄与、貢献をした実体(構造化又は URI により表現する場合)	-	29
DC Contributor	http://purl.org/dc/elements/1.1/contributor	dc:contributor	情報資源の整理に何らかの寄与、貢献をした実体(文字列で表現する場合)	-	30
1.1.5 版に関する語彙					
Edition	http://purl.org/dc/terms/edition	dcndl:edition	当該情報資源が属する版	15	31
Edition Creator	http://purl.org/dc/elements/1.1/editionCreator	dcndl:editionCreator	当該情報資源が属する版に対し著作権を持つ実体	15	31
1.1.6 出版に関する語彙					
DCTERMS Publisher	http://purl.org/dc/terms/publisher	dcterms:publisher	出版社・頒布者に関する情報(構造化又は URI により表現する場合)	-	32
DC Publisher	http://purl.org/dc/elements/1.1/publisher	dc:publisher	情報資源の成立に何らかの寄与、貢献をした実体(文字列で表現する場合)	-	33
Publication Place	http://ndl.go.jp/dcnl/terms/publicationPlace	dcndl:publicationPlace	出版地・頒布地	16	34
1.1.7 主題に関する語彙					
DCTERMS Subject	http://purl.org/dc/terms/subject	dcterms:subject	情報資源の内容を表す統制語彙、分類記号、フリーキーワード(構造化又は URI により表現する場合)	-	35
DC Subject	http://purl.org/dc/elements/1.1/subject	dc:subject	情報資源の内容を表す統制語彙、分類記号、フリーキーワード(文字列で表現する場合)	-	36
1.1.8 注記等に関する語彙					
DCTERMS Description	http://purl.org/dc/terms/description	dcterms:description	注記	-	38
DCTERMS Abstract	http://purl.org/dc/terms/abstract	dcterms:abstract	要約・抄録等	-	39
DCTERMS Table Of Contents	http://purl.org/dc/terms/tableOfContents	dcterms:tableOfContents	目次情報	-	39
1.1.9 日付に関する語彙					
DCTERMS Date	http://purl.org/dc/terms/date	dcterms:date	当該情報資源のライフサイクルにおける何らかの事象の日付	-	40
DCTERMS Available	http://purl.org/dc/terms/available	dcterms:available	利用可能日(利用可能期間)	-	41
DCTERMS Created	http://purl.org/dc/terms/created	dcterms:created	作成日	-	42
DCTERMS Date Accepted	http://purl.org/dc/terms/dateAccepted	dcterms:dateAccepted	論文や記事などの受理日	-	42
Date Captured	http://ndl.go.jp/dcnl/terms/dateCaptured	dcndl:dateCaptured	当該情報資源を採取・保存した日	16	43
DCTERMS Date Copyright	http://purl.org/dc/terms/dateCopyright	dcterms:dateCopyright	著作権が発効した日	-	44
DCTERMS Date Submitted	http://purl.org/dc/terms/dateSubmitted	dcterms:dateSubmitted	論文や記事などの提出日	-	44
DCTERMS Issued	http://purl.org/dc/terms/issued	dcterms:issued	出版年月日	-	45
DCTERMS Modified	http://purl.org/dc/terms/modified	dcterms:modified	更新日	-	46
DCTERMS Valid	http://purl.org/dc/terms/valid	dcterms:valid	有効期限(有効期間)	-	46
1.1.10 言語に関する語彙					
DCTERMS Language	http://purl.org/dc/terms/language	dcterms:language	情報資源の記述言語	-	47
Original Language	http://ndl.go.jp/dcnl/terms/originalLanguage	dcndl:originalLanguage	原文の言語	17	48

用語の表示名	プロパティURI	QName	定義等	複製場所 NLL Metadata Term Profile
1.1 著人・団体等の利用の記録に使用する語彙	http://ndl.go.jp/	creator	出版者の氏名に関する情報	29 95
1.2 制御的なコンテンツ	http://ndl.go.jp/	alternateVolume		
2 情報源の記録に使用する語彙	http://ndl.go.jp/	title	Title of the work or other identifying information	29 96
2.1 1.1.1 形式で表現できない場合の語彙	http://ndl.go.jp/	alternativeTitle	Alternative titles of the work	29 97
2.2 情報の複製に関する語彙	http://ndl.go.jp/	seriesTitle	Series title of the work	29 98
	http://ndl.go.jp/	seriesAlternative	Alternative series titles	30 98
	http://ndl.go.jp/	seriesVolume	Series volume or issue information	30 99
	http://ndl.go.jp/	seriesVolumeAlternative	Alternative series volumes	30 99
	http://ndl.go.jp/	seriesVolumePart	Series volume part information	30 100
	http://ndl.go.jp/	seriesVolumePartAlternative	Alternative series volume parts	31 101
	http://ndl.go.jp/	seriesVolumePartVolume	Series volume part volume information	31 102
	http://ndl.go.jp/	seriesVolumePartVolumeAlternative	Alternative series volume parts volumes	32 103
	http://ndl.go.jp/	seriesVolumePartVolumePart	Series volume part volume part information	32 104
	http://ndl.go.jp/	seriesVolumePartVolumePartAlternative	Alternative series volume part volumes parts	32 104
	http://ndl.go.jp/	seriesVolumePartVolumePartVolume	Series volume part volume part volume information	33 105
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumeAlternative	Alternative series volume part volumes parts volumes	33 106
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePart	Series volume part volume part volume part information	34 107
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartAlternative	Alternative series volume part volumes parts volumes parts	34 108
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolume	Series volume part volume part volume part volume information	35 109
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumeAlternative	Alternative series volume part volumes parts volumes parts volumes	36 110
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePart	Series volume part volume part volume part volume part information	36 111
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePartAlternative	Alternative series volume part volumes parts volumes parts volumes parts	37 112
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePartVolume	Series volume part volume part volume part volume part volume information	37 112
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePartVolumeAlternative	Alternative series volume part volumes parts volumes parts volumes parts volumes	37 113
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePartVolumePart	Series volume part volume part volume part volume part volume part information	39
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePartVolumeAlternative	Alternative series volume part volumes parts volumes parts volumes parts volumes	39
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePartVolumePart	Series volume part volume part volume part volume part volume part information	39
	http://ndl.go.jp/	seriesVolumePartVolumePartVolumePartVolumePartVolumeAlternative	Alternative series volume part volumes parts volumes parts volumes parts volumes	40

用語の表示名	プロパティURI	QName	定義等	複製場所 NLL Metadata Term Profile
1.1.1.1 形態学に関する語彙	http://ndl.go.jp/	format	形態学に関する形態学またはデジタル形態学の記述	49
DCTERMS Extent	http://ndl.go.jp/	extent	形態学に関する形態学に関する情報	49
DCTERMS Type	http://ndl.go.jp/	type	形態学に関する形態学の記述に関する記号	49
Metadata Type	http://ndl.go.jp/	type	形態学に関する形態学の記述に関する記号	17 52
1.1.1.2 識別に関する語彙	http://ndl.go.jp/	identifier	識別に関する識別子	52
SourceIdentifier	http://ndl.go.jp/	sourceIdentifier	識別に関する識別子の識別子	52
DCTERMS Bibliographic Citation	http://ndl.go.jp/	bibliographicCitation	識別に関する識別子の識別子	18 55
1.1.1.3 関係に関する語彙	http://ndl.go.jp/	relation	関係に関する関係	57
DCTERMS In Relation	http://ndl.go.jp/	inRelation	関係に関する関係	57
DCTERMS Has Version	http://ndl.go.jp/	hasVersion	関係に関する関係	58
DCTERMS Has Variant	http://ndl.go.jp/	hasVariant	関係に関する関係	59
DCTERMS Is Replaced By	http://ndl.go.jp/	isReplacedBy	関係に関する関係	60
DCTERMS Replaces	http://ndl.go.jp/	replaces	関係に関する関係	61
DCTERMS Requires	http://ndl.go.jp/	requires	関係に関する関係	62
DCTERMS Requires Part	http://ndl.go.jp/	requiresPart	関係に関する関係	62
DCTERMS Has Part	http://ndl.go.jp/	hasPart	関係に関する関係	63
DCTERMS Is Referenced By	http://ndl.go.jp/	isReferencedBy	関係に関する関係	64
DCTERMS References	http://ndl.go.jp/	references	関係に関する関係	65
DCTERMS Has Format	http://ndl.go.jp/	hasFormat	関係に関する関係	67
DCTERMS Has Format Part	http://ndl.go.jp/	hasFormatPart	関係に関する関係	68
DCTERMS Conform To	http://ndl.go.jp/	conformTo	関係に関する関係	70
DCTERMS Source	http://ndl.go.jp/	source	関係に関する関係	71
DCTERMS See Also	http://www.w3.org/2000/01/rdf-schema#seeAlso	seeAlso	関係に関する関係	72
OWL Same As	http://www.w3.org/2002/07/owl#sameAs	sameAs	関係に関する関係	73
1.1.1.4 範囲に関する語彙	http://ndl.go.jp/	range	範囲に関する範囲	74
DCTERMS Coverage	http://ndl.go.jp/	coverage	範囲に関する範囲	75
DCTERMS Spatial	http://ndl.go.jp/	spatial	範囲に関する範囲	75
DCTERMS Temporal	http://ndl.go.jp/	temporal	範囲に関する範囲	76
1.1.5 発行者/出版者の関係に関する語彙	http://ndl.go.jp/	publisher	発行者に関する発行者	76
Publication Periodicity	http://ndl.go.jp/	publicationPeriodicity	発行者に関する発行者	18 77
Volume Range	http://ndl.go.jp/	volumeRange	発行者に関する発行者	18 77
1.1.6 博士論文の著作に関する語彙	http://ndl.go.jp/	thesis	博士論文に関する博士論文	19 78
Degree Name	http://ndl.go.jp/	degreeName	博士論文に関する博士論文	19 78
Dissertation Number	http://ndl.go.jp/	dissertationNumber	博士論文に関する博士論文	19 79
Date Granted	http://ndl.go.jp/	dateGranted	博士論文に関する博士論文	20 80
Degree Grantor	http://ndl.go.jp/	degreeGrantor	博士論文に関する博士論文	30 80
1.1.7 雑誌記事の著作に関する語彙	http://ndl.go.jp/	issue	雑誌記事に関する雑誌記事	20 81
Issue	http://ndl.go.jp/	issue	雑誌記事に関する雑誌記事	20 81
Publication Name	http://ndl.go.jp/	publicationName	雑誌記事に関する雑誌記事	21 82
Publication Volume	http://ndl.go.jp/	publicationVolume	雑誌記事に関する雑誌記事	21 83
Page Range	http://ndl.go.jp/	pageRange	雑誌記事に関する雑誌記事	21 83
1.1.8 デジタル化した資料の特長に関する語彙	http://ndl.go.jp/	digitalized	デジタル化した資料に関するデジタル化	22 84
Digitized Publisher	http://ndl.go.jp/	digitizedPublisher	デジタル化した資料に関するデジタル化	22 84
Data Digitized	http://ndl.go.jp/	dataDigitized	デジタル化した資料に関するデジタル化	23 85
1.1.9 情報源の範囲・入手に関する語彙	http://ndl.go.jp/	availability	情報源に関する情報源	86
DCTERMS Availability	http://ndl.go.jp/	availability	情報源に関する情報源	87
DCTERMS Access Rights	http://ndl.go.jp/	accessRights	情報源に関する情報源	87
Availability	http://ndl.go.jp/	availability	情報源に関する情報源	87
DCTERMS Rights	http://ndl.go.jp/	rights	情報源に関する情報源	88
DCTERMS Rights Holder	http://ndl.go.jp/	rightsHolder	情報源に関する情報源	88
1.1.20 その他の語彙	http://ndl.go.jp/	thumbnail	サムネイルに関するサムネイル	89
FOAF Thumbnail	http://www.foaf.org/	thumbnail	サムネイルに関するサムネイル	89
1.2 情報源の創作者情報に関する語彙	http://ndl.go.jp/	creator	創作者に関する創作者	23 90
Holding Agent	http://ndl.go.jp/	holdingAgent	創作者に関する創作者	25 91
Call Number	http://ndl.go.jp/	callNumber	創作者に関する創作者	25 91
Local Call Number	http://ndl.go.jp/	localCallNumber	創作者に関する創作者	25 92
Holding Issues	http://ndl.go.jp/	holdingIssues	創作者に関する創作者	26 93
Absent Issues	http://ndl.go.jp/	absentIssues	創作者に関する創作者	26 93
1.3 情報源の管理情報に関する語彙	http://ndl.go.jp/	catalogingRule	情報源に関する情報源	27 94
Cataloging Rule	http://ndl.go.jp/	catalogingRule	情報源に関する情報源	27 94
Record Status	http://ndl.go.jp/	recordStatus	情報源に関する情報源	27 95

作品項目入力例：長尾真 著. 電子図書館. 新装版. 岩波書店, 2010.3.

語彙のカテゴリ	QName	記入例
語彙の表示名		
1 標準的なプロパティ		
1.1 情報資源の記述に使用する語彙		
1.1.1 複数の記述に共通して使用する語彙		
Transcription	dcndl:transcription	デンシ トショカン
1.1.2 タイトルに関する語彙		
DCTERMS Title	dcterms:title	電子図書館
DC Title	dc:title	電子図書館
1.1.3 作成者に関する語彙		
DC TERMS Creator	dcterms:creator	長尾, 真, 1936- ナガオ, マコト
DC Creator	dc:creator	長尾真 著
1.1.5 版に関する語彙		
Edition	dcndl:edition	新装版
1.1.6 出版に関する語彙		
DCTERMS Publisher	dcterms:publisher	岩波書店
Publication Place	dcndl:publicationPlace	JP
1.1.7 主題に関する語彙		
DCTERMS Subject	dcterms:subject	電子図書館
1.1.9 日付に関する語彙		
DCTERMS Date	dcterms:date	2010.3
DCTERMS Issued	dcterms:issued	2010
1.1.10 言語に関する語彙		
DCTERMS Language	dcterms:language	jpn
1.1.11 形態等に関する語彙		
DCTERMS Extent	dcterms:extent	127p ; 19cm
Material Type	dcndl:materialType	図書
Price	dcndl:price	1500 円
1.1.12 識別に関する語彙		
DCTERMS Identifier	dcterms:identifier	http://ndl.go.jp/dcndl/terms/ISBN 9784000057035
1.1.19 情報資源の利用・入手に関する語彙		
DCTERMS Audience	dcterms:audience	一般

※実際のデータの一部抜粋です。

参考資料4 文化遺産オンライン 作品情報(文化遺産情報)項目一覧

出典：『「文化遺産オンライン構想」成果報告フォーラム ～文化遺産とデジタル・アーカイブの最前線を知る～』文化庁文化財部伝統文化課普及指導係, 2011. p.71

2011/12/1

日本語項目名	データ型	説明
文化遺産コード	英数	文化遺産を一意に表すコード(文化遺産オンライン側で登録時に付与)
公開可否	英数(コード)	公開ステータス管理用(公開、非公開の別など)
所蔵館コード*	英数	所蔵館の館 ID(文化遺産オンライン側で付与したもの)
識別コード	テキスト	各所蔵館における作品識別コード(各館独自の作品 ID)
文化遺産名称*	テキスト	作品名称
文化遺産名称かな	テキスト	作品名称よみがな
員数	テキスト	員数に関する情報。数詞も含む
材質・構造・技法	テキスト	材質・構造・技法に関する情報。単位も含む
サイズ	テキスト	サイズに関する情報。単位も含む
解説文	テキスト	作品に関する解説など
作者名	テキスト	制作者名など
作者名かな	テキスト	制作者名のよみがななど
作者生没年	テキスト	作成者の生没年、例：1900-1981 など
西暦	テキスト	作品の制作された時期(西暦)
分野	テキスト	各所蔵館による分類名など
時代	テキスト	作品に制作された時期(時代、年号など)
世紀	テキスト	作品の制作された時期(世紀)
地域	テキスト	地域に関する情報。
指定区分	テキスト	「重要文化財」、「国宝」などの国指定区分名称
URL	テキスト	所蔵館 Web サイト内におけるこの文化遺産の個別ページ URL など
所在地	テキスト	所在地、考古遺物の場合は出土地など
所有者	テキスト	所有者や来歴に関する情報など
制作者生没年	テキスト	制作者の生没年、例：1900-1981 など
その他の情報	テキスト	初出、公開日などの情報、考古遺物の場合は遺跡名など
西暦 FROM(半角数字)	数字	作成時期の西暦年。複数年の場合は最初の年。紀元前の場合は先頭に'-'を追加
西暦 TO(半角数字)	数字	作成時期の終了西暦年。紀元前の場合は先頭に'-'を追加
分野コード	数字(コード)	選択項目コード一覧参照
時代コード	数字(コード)	選択項目コード一覧参照
世紀 FROM	数字	制作時期の世紀。複数年の場合は開始世紀(現在は未使用)
世紀 TO	数字	制作時期の世紀。複数年の場合は終了世紀(現在は未使用)

地域コード	数字(コード)	選択項目コード一覧参照
指定区分コード	数字(コード)	選択項目コード一覧参照
市区町村	数字(コード)	選択項目コード一覧参照(公開準備中)
緯度	数字	作品の所在地における緯度
経度	数字	作品の所在地における経度
保管者コード	英数	作品を保管している館の ID (主に文化庁提供の国指定文化財情報において利用)
関連リンク	テキスト	任意のアンカーテキストと URL のペア。2 件まで登録可
Gallery 公開可否	数字	公開：1 非公開：0

※必須項目

- ・文化遺産名称
- ・所蔵館コード

文化遺産オンライン情報登録 Q&A

入力編「全ての項目を埋める必要があるのでしょうか？」⁶³

「全ての項目を埋める必要はありません。文化遺産名称だけでも登録は可能です。ただし、様々な検索項目を用意しているため、より多くの項目を記入しておくことで、利用者が検索した際にヒットして閲覧される可能性が増えます。解説文も入力しなくても登録は可能です。しかし文化遺産データベースでは、全文検索を行っているので、解説文中に様々なキーワードがあると、ヒットする可能性が高くなり、利用者の目に触れる機会が増えます。」

作品項目入力例 「高井鴻山記念館 妖怪漢詩」

日本語項目名	入力内容	日本語項目名	入力内容
文化遺産コード	180715	分野	その他美術品 — 書 (36)
公開可否	公開する (1)	時代	江戸 (12)
所蔵館コード [※]	12715 : (高井鴻山記念館)	地域	長野県 (20)
文化遺産名称 [※]	妖怪図 妖怪漢詩		
文化遺産名称かな	ようかいず ようかいかんし		
材質・構造・技法	軸装		
サイズ	137.0×66.4cm		
作者名	高井鴻山		
作者名かな	たかいこうざん		

⁶³ 文化遺産オンライン情報登録 Q&A <http://bunkaedit.nii.ac.jp/faq.html>

参考資料 5 規格番号

本ガイドラインで記載した世界標準規格の規格番号を以下に記します。

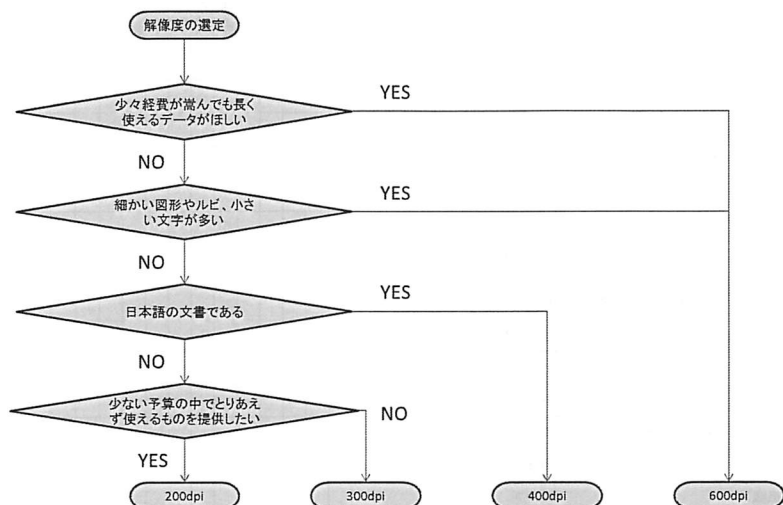
No	名称	規格番号
1	JPEG	ISO/IEC 10918、ITU-T T.81、JIS X 4301
2	JPEG 2000	ISO/IEC 15444-1、ITU-T T.800
3	PNG	ISO/IEC 15948
4	PDF	ISO 32000
5	Dublin Core	ISO 15836、NISO Z39.85
6	HTTP	IETF RFC2616
7	Z39.50	ISO 23950 - 1998、JIS X0806

参考資料6 画像コンテンツのパラメータ決定フロー

ここでは、各館の用途や条件から、デジタル化したコンテンツの画像ファイルの仕様を選ぶフローを記載します。

数値そのものは確定的な情報ではありませんので、考え方の目安としてお使い下さい。

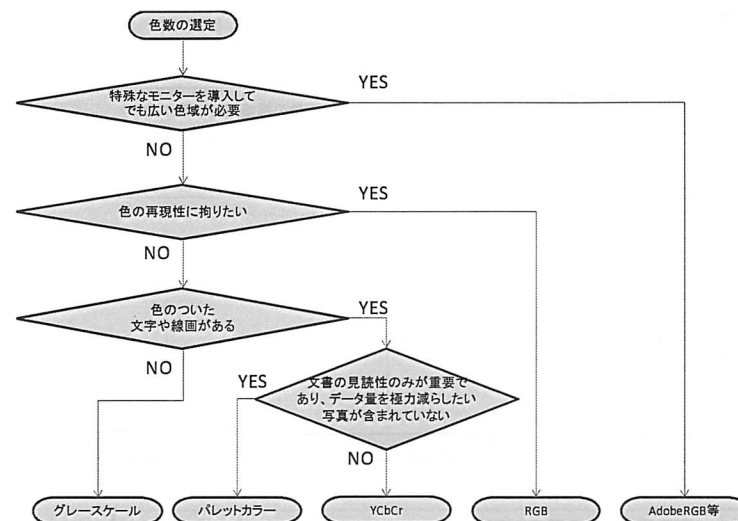
解像度の選定判断例



【注意点】

- ①データ量（ストレージコスト）は200dpiを基準にすると、300dpiで2.25倍、400dpiで4倍、600dpiで9倍と増加します。
- ②上記のフローは、保存用データの解像度を決定する場合のものです。
提供用データの場合は、データ量が増加すると、ストレージのコストに加えて配信サーバ側の負荷や表示にかかる時間、必要なネットワークの速度も大きく変わります。
提供用画像の解像度は上記で決定したものよりも低く、また非可逆圧縮の画像フォーマットを採用してデータ量を削減する必要があります。

色数の選定判断例



- ・選定結果の左から、ファイル容量が少なく済む順、かつ色の再現度が低い順に並んでいます。
- ・YCbCrはY成分のみを読みだすことでグレースケールとして扱えるため、使いまわしが良いといえます。

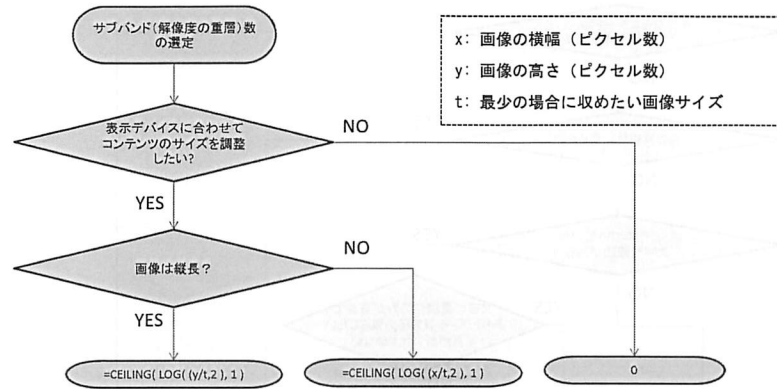
【用語説明】

No	用語	説明
1	カラー空間 (色空間)	表示・印刷できる色の範囲やその表現方式のことで、機器の特性によって表現できる色が異なることから、様々な方式がある。同じ画像でもカラー空間が異なると色の質等が異なるので注意が必要である。
2	グレースケール	コンピュータ上で色の表現方法の一つ。画像を白から黒までの明暗だけで表現し、色の情報は含まない「モノクロ」のこと。
3	パレットカラー	使用する色を決まった数だけあらかじめ選んで通し番号をつけておき、色の指定に番号を用いる色表現方式。色数には256色が使われることが多く、画像の容量を小さくするのに有効な方式である。 ⁶⁴
3	YCbCr	代表的なカラー空間の一つ。Yは輝度、Cbは青さ、Crは赤さの度合いを示す。JPEGやMPEG等DCT変換によって画像圧縮を行う場合は、RGB形式の画像データを圧縮するのではなく、YCbCrに変換した画像データを圧縮する。
4	RGB	代表的なカラー空間の一つ。コンピュータのスクリーンの多くは、RGB（赤・緑・青）の3色の組み合わせで色を作り出す。
5	AdobeRGB	Adobe Systemsが定義した色空間（カラー空間）のことである。1998年に発表された。基本的にCRTディスプレイで色を表現することを想定しているsRGBは、表現できる色の範囲に一定の限りがある。一方AdobeRGBは、sRGBに比べても遥かに広い範囲の再現領域を持ち、きめの細やかな色彩の表現が可能。印刷物に対する適合性や色構成の厳密性も高く、特にDTPの分野などでは長らくAdobeRGBが標準的に用いられている。ただしAdobeRGBを再現できる機器は比較的高価であり、専ら業務用の機材で採用されている。

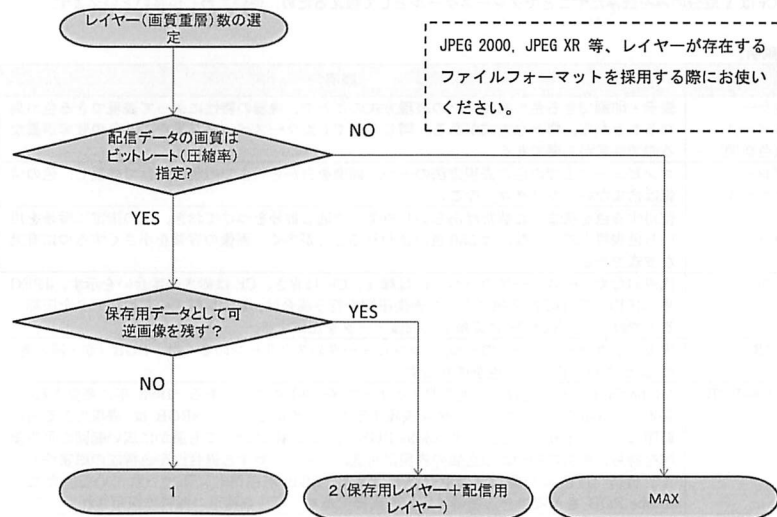
(特記無き物の出典：「国立国会図書館 資料デジタル化の手引 2011年版」)

⁶⁴ IT用語辞典 e-Words: <http://e-words.jp/w/E38391E383ACE38383E38388E382ABE383A9E383BC.html> (平成24年2月10日確認)

サブバンド(解像度の重層)数の選定判断例



レイヤー(画質重層)数の選定判断例 (JPEG 2000, JPEG XR 等の場合)



「MAX」に入る値は、画像フォーマットにより異なります。

参考資料7 チェックシート

項目	想定される選択肢	システムに依存する機能										コンテンツフォーマットに依存する機能							
		認証	著作権制御	DRM	デジタル標準	デジタル標準	目録データの取得・入力	目録とコンテンツの紐づけ	メタデータ連携	クラウド	オンプレミス	Unicode	動的配信	圧縮性能	可逆性	多層画質	多層解像度		
目的	サービス	館内限定で利用																	
	範囲	インターネット経由の外部利用																	
	閲覧対象	著作権	著作権切れ限定																
		著作権あり	○	○	◎	◎													
	リーチ	目録に限定				◎	◎												
		目録に紐づけられたコンテンツを公開				◎	◎	◎											
	利用形態	閲覧環境	PCのみ																
		携帯型情報端末を含む												◎	◎				○
		プリントアウト	許可する	○	○	○													○
	保存・継承	2次利用	考慮する															◎	○
考慮しない																	◎		
組織	政府系				◎														
	民間																		
リソース	初期投資	潤沢												△	△	△			
		厳しい				◎								◎	◎	◎			
	メンテナンス体制	充分												△	△	△			
厳しい													◎	◎					
管理	品質管理	自動化する																◎	○
		自動化しない																	
外部連携	考慮する				◎	○						◎							
	考慮しない																		

※1 (Digital Rights Management) デジタルデータとして表現されたコンテンツの著作権を保護し、その利用や複製を制御・制限する技術の総称
 ※2 デジタル標準: 公的な機関が認証することにより規格となっている標準のこと
 ※3 オンプレミス: 自社で用意した設備でソフトウェアなどを導入・利用すること
 ※4 クラウド: ネットワーク上にあるサーバのサービスを活用できるというコンピューティング形態を指す
 ※5 Unicode: 1993年に国際標準化機構(ISO)で ISO/IEC 10646の一部(UCS-2)として標準化された文字コード体系
 ※6 可逆性: 条件を変えるとその変化と逆の方向に変化が起こってもとの状態に戻る
 ※7 多層画質: 複数の層それぞれの画質
 ※8 多層解像度: いくつもの層それぞれの解像度

「チェックシート」の利用方法

◎	4点	選択したカラムの合計が4点以上の場合、要求仕様上の考慮が必要になる
○	2点	
△	1点	

「項目」の列には、最大3つのレベルで仕様項目が示されています。小項目ごとに、「想定される選択肢」が2つずつありますので、自館の要望に近い項目の行を選択します。

選択した行に◎（4点）、○（2点）、△（1点）があれば、それをチェックしておきます。

選択が終わったら、機能（列）毎に点数を合計し、仕様書に考慮すべき機能を洗い出します。

目安として、合計が4点以上になった機能は、特に要求仕様上の考慮が必要といえます。

点数が多い機能は、そのシステムにとって重要な機能といえます。4点に満たない箇所でも点数が加算されていれば、考慮は必要になります。

例えば、サービス範囲として「インターネット経由の外部利用」を選んだ場合は認証（2点）、閲覧権制御（2点）、セキュリティ（DRM）（4点）、データベース（4点）、Unicode（4点）と言った形で機能に点数が加算されていきます。

外部連携に「考慮しない」を選んだ場合は、すべて0点です。これを機能毎に集計します。

上記の条件で、システム仕様書を検討したいと言った場合は、セキュリティ（DRM）、データベース、Unicodeの最低3機能については、仕様書として特に記載すべき機能となります。

具体的な利用例は参考資料8に記載がありますので参照して下さい。

参考資料8 発注ケース：館内利用者への所蔵資料公開

ここで出てくる館や要望は架空のものであり、
例示するシステム構成を推奨するものではありません。
あくまでも、考え方の参考例としての記載です。
その旨ご留意の上、お読み下さい。

【背景・前提】

展示スペースの関係からこれまで来館者に閲覧してもらうことができなかった郷土資料の書籍（A5またはB5、平均80頁、約500冊）をスキャンし、館内に設置したタブレット端末で閲覧してもらおうと考える郷土資料館があります。

(1) チェックシートを使ってシステム要件のスタート地点を確認する

最初に、参考資料7の「チェックシート」を使用し、検討する必要があるシステム要件を確認します。

チェックシートに従って要件を挙げていくと下記のようにになりました。

項目	意向
サービスの範囲	来館料を徴収した利用者だけに館内のタブレット型の携帯端末でコンテンツを公開する
著作権	考慮する必要がない
公開する内容	コンテンツ
閲覧端末	館内に設置したタブレット型携帯端末
プリントアウト	考慮しない
2次利用	考慮しない
長期保存・継承	時間経過とともに陳腐化する仕様は避けた
組織	政府系ではない
初期投資	厳しい
メンテナンス体制	厳しい
品質検査	数が少ないので自動化の必要はない
外部連携	考慮しない

次ページに要件を検討した時のチェックシートを記載します。

要件となる選択肢に色をつけています。

項目	想定される選択肢	システムに依存する機能										コンテンツフォーマットに依存する機能					
		認証	閲覧権制御	著作権	セキュリティ(DRM)	データベース	デジタル標準	目録データの取得・入力	目録とコンテンツの紐づけ	メタデータ連携	オンプレミス	クラウド	Unicode	動的配信	圧縮機能	可逆性	多層解像度
目的	サービス範囲	館内限定で利用															
		インターネット経由の外部利用	○	○	◎	◎							◎				
	閲覧対象	著作権切れ限定															
		著作権あり	○	○	◎	◎											
	リーチ	目録に限定				◎	◎										
		目録に紐づけられたコンテンツを公開				◎	◎	◎									
	利用形態	閲覧環境															
		PCのみ 携帯型情報端末を含む											◎	◎			○
	プリントアウト	許可する	○	○	○												○
		許可しない															
保存・継承	2次利用														◎	○	
	考慮しない														◎		
組織							◎										
	政府系 民間又は政府系以外																
リソース	初期投資	潤沢										△	△	△			
		厳しい					◎					◎	◎	◎			
	メンテナンス体制	充分										△	△	△			
		厳しい											◎	◎			
管理	品質管理	自動化する													◎	○	
		自動化しない															
外部連携		考慮する				◎	○			◎							
		考慮しない															

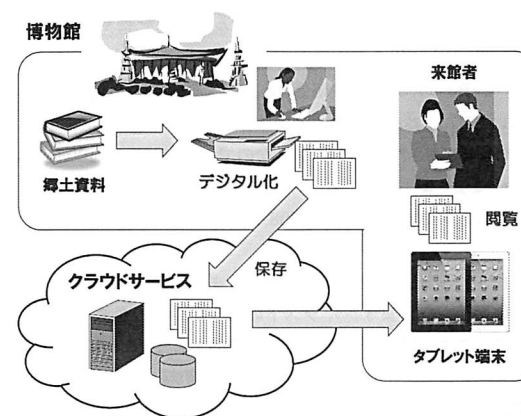
◎ 4点 選択したカラムの合計が4点以上の場合、要求仕様上の考慮が必要になる
○ 2点
△ 1点

項目	点数	判断
認証	0	不要
閲覧権制御	0	不要
著作権管理	0	不要
セキュリティ(DRM)	0	不要
データベース	4	必要
デジュール標準	4	必要
目録データの取得・入力	4	必要
目録とコンテンツとの紐づけ	4	必要
メタデータ連携	0	不要
オンプレミス	0	不要
クラウド	8	必要
Unicode 使用	12	必要
動的配信	4	必要
圧縮性能	4	必要
可逆性	4	必要
多層画質	0	不要
多層解像度	2	考慮

以上によりシステム要件定義の出発点は下記のようになります。

- データベース
- デジュール標準
- 目録データの取得・入力
- 目録とコンテンツとの紐づけ
- クラウド
- Unicode 使用
- 動的配信
- 圧縮性能
- 可逆性
- 多層解像度

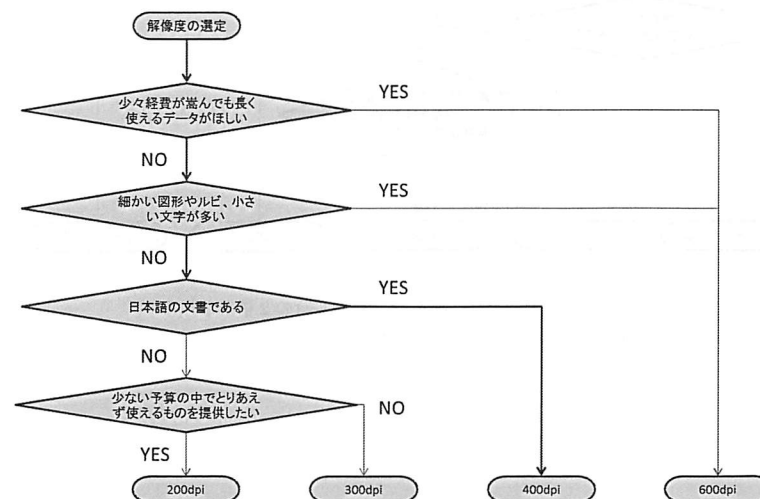
システム構成のイメージ図を、以下に示します。



(2) コンテンツの形式を確認してシステム要件を見直す

せっかく作るのですから、作製したデジタルデータは長く使いたいところです。その一方で「高くつく」というのは避けたいところです。また、コンテンツのほとんどは墨書でそれほど細かい図形や文字は見当たりません。

そこで解像度としては思い切って 400dpi を選択しました。



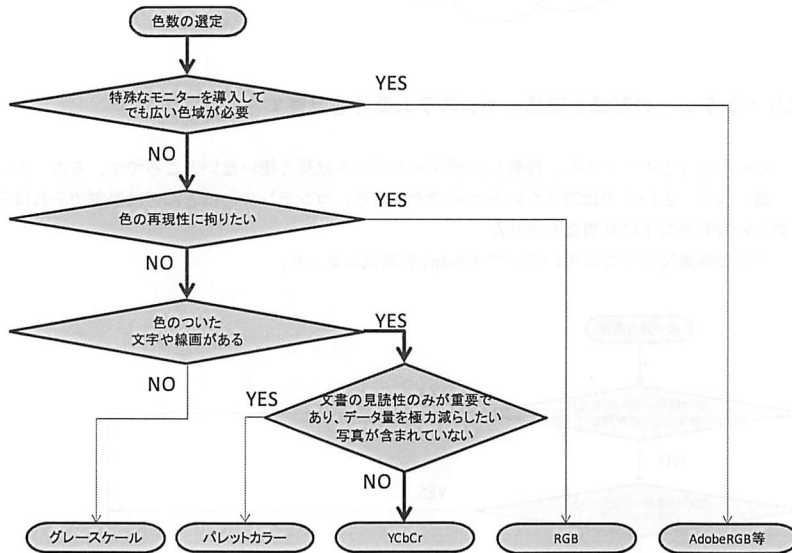
ここで、あらかじめ画素数を計算しておきましょう。コンテンツはA5またはB5なので大きい方のB5を例にとってコンテンツあたりの画素数を概算しておきます。

B5サイズ見開き(=B4サイズ)の横幅は14.33インチですから、400dpiの場合の横方向画素数は 400×14.33 で5,732ピクセルになります。

一方、B5サイズ画像の高さは10.12インチなので、同様にして400dpiの場合の縦方向画素数は4,048ピクセルになります。

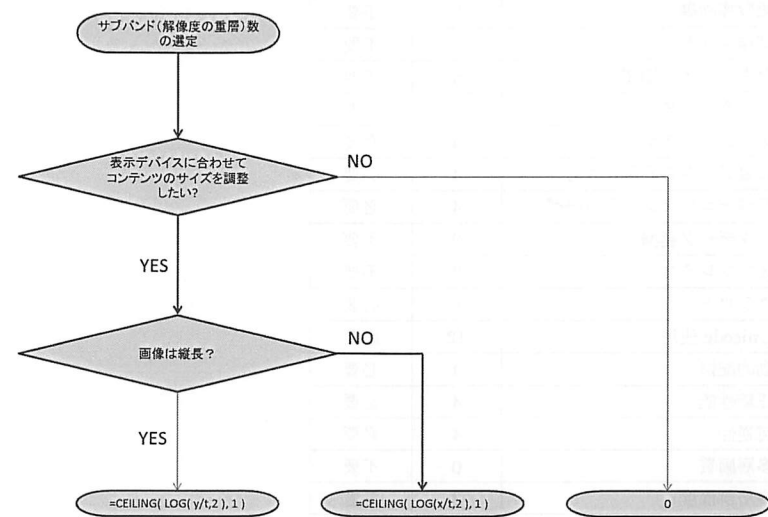
そうすると全体では $5,732 \times 4,048$ で23,203,136(約23M)ピクセルとなります。

次に色数を考えます。写真と言えるほどではありませんが、多色使いのページもあればグラデーションの美しい墨絵もあります。イラストのように少ない色数ではせっかくのコンテンツが泣きます。そこで、ここでは写真程度の色数を選びました。



YCbCrで表現すると、データ量(バイト数)はピクセル数の2倍となります。先述の約23Mピクセルの画像の場合には約46Mバイトとなります。

次に解像度の重層(多層解像度)について考えます。必要な層の数(分割数)を計算します。



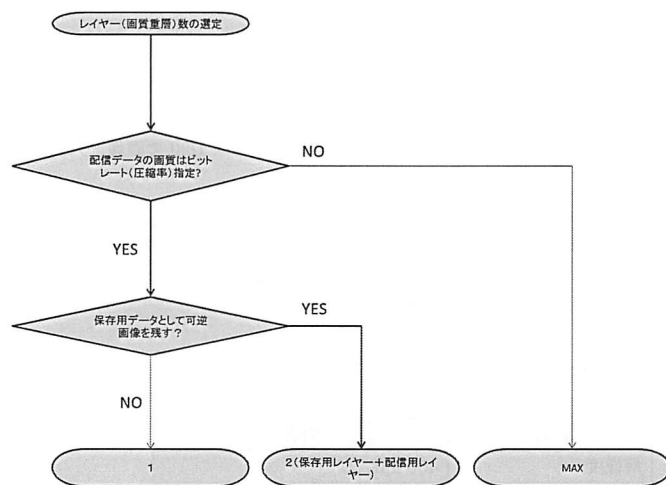
ここまでの作業で、今回対象とするコンテンツの横方向画素数が5,732ピクセルであることは判っています。

閲覧デバイスとして考えているiPad2は横方向1,024ピクセルですから、これを式に当てはめてエクセルで計算してみると、 $=\text{CEILING}(\text{LOG}(5732/1024, 2), 1)$ で3となります。

この場合には画像データから横幅がそれぞれ5,732、2,866、1,433、716ピクセルの画像を取り出せることとなります^{※1}。一番小さい横幅716ピクセルの画像は、iPad2の長辺ピクセル数である1,024ピクセルよりも小さいため、画像全体を画面に収めることができます。

※1 分割数が3(2の3乗分)から計算した場合(5732÷2=2876、5732÷4、5732÷8=716.5)

次にレイヤー数（画質重層）を考えましょう。「(1) チェックシートを使ってシステム要件のスタート地点を確認する」の結果を見ると画質の重層は必要ないことになっています。ところが、後に選択するフォーマットによっては費用をかけずにこの特性を得ることができます。追加の経費なしにより良いものが得られるのであればそれに越したことはありませんので、念のために必要なレイヤー数を検討しておきます。



今回、画質の重層は必須要件ではありませんので、画質の重層が可能なフォーマットを選択した場合にのみ2レイヤーを選択します。この場合はレイヤーが2層なので、保存用のデータファイルから一層目に対応する配信用の簡単に画像を取り出すことができます。

そうでないフォーマットを選択した場合には、いったん保存用のデータを作成して、それを配信用のデータに変換する処理が必要です。

(3) フォーマットを選択する

次に配信用画像の圧縮率を決めます。今回は平均80頁×500冊=40,000頁、見開きでスキャンするのであれば20,000画像となりますので、目視検査が可能な分量と言えます。(1分間に2枚の画像を確認できるとして2日程度の作業です)。対象とするデータ量が多くて目視検査が事実上不可能な場合には、品質検査を自動化して画像ごとに最適な圧縮率を決定しますが、今回はそこまでする必要がないため、固定で圧縮率を決めてしまうことにします。

さて、圧縮率を決めるにあたってはその前にフォーマットを決めておく必要があります。何故なら画像フォーマットはそれぞれに拠って立つ理論が異なり、それが時として顕著な圧縮性能の差となって現れてしまうからです。例えば1994年に規格化されたJPEGと2000年に規格化されたJPEG 2000とでは、圧縮性能に概ね2倍の差がある(つまりJPEG画像とその半分のサイズのJPEG 2000画像とを比べた時に画質が同じくらいになる)ことが知られています。

場合によっては非標準のフォーマットを選択することもあります。フォーマットの安全性(互換性、継続性、特許リスク)と競争による低コスト化の観点から、またはWTOの政府調達協定の観点から、特別な理由がない限りデジュール標準(3大国際標準化機関の規格)を採用することをお勧めします。

非可逆圧縮が可能なデジュール標準の画像フォーマットとしては、JPEG、JPEG 2000、JPEG XRが代表的です。なお、JPEGという文字列は3大国際標準化機関が合同で策定した画像規格であることを示しています。

これら選択肢の中からどれを採用するのかを決めるにあたって、各規格の特性を考えてみます。

JPEGを選択する場合の一番のメリットは、どのブラウザからでも直接見ることができる点にあります。拡大・縮小や移動などの操作が多少不便であっても構わないのであれば、特別な配信の仕組みを用意する必要がありません。一方で、サーバのディスクを大きく消費すること、通信に時間がかかること、タブレットなどの携帯デバイスに対応するのが難しい(iPad2を例にとると、扱える画像の上限サイズとして縦横それぞれ1024ピクセルと規定されているため、より大きな画像を扱う場合には配信サービス側に対応することになりますが、JPEGフォーマットはそのような処理に向いていません)こと、あらかじめ圧縮率を決められないこと、保存用に別フォーマットのデータを用意する必要があることなどがデメリットとして挙げられます。

JPEG 2000はデータ提供者が必要とする機能や性能を盛り込んで策定されたフォーマットで、大規模な画像アーカイブではこのフォーマットを採用した例が多く見られます。メリットとしては、圧縮性能が良いこと、合理的な動的配信が可能であること、それにより通信量を抑え携帯デバイスに対応して使いやすいユーザ・インターフェースを提供できること、保存用画像から配信用画像を取り出せること、画質制御を行いやすいことなどが挙げられています。

方で、このフォーマットの画像を直接表示できるブラウザが今のところ Safari のみであることから、ブラウザからの要求に応じて情報を取り出し動的に配信する仕組みが必要となること、JPEG 2000 からの配信用データ取り出し用のソフトウェア・ライブラリには有償のものが多く、ライセンス費用の考慮が必要になることがデメリットとして挙げられています。

JPEG XR の"XR"は eXtended Range の略であり、色域の向上と複雑度の低減を目指して策定されたフォーマットです。メリットとしては、精密な色表現が可能であること、デジカメなどの携帯デバイスにも搭載可能なこと、圧縮性能が良いこと、保存用画像から配信用画像を取り出せること、画質制御を行いやすいことなどが挙げられます。一方でこのフォーマットの画像を直接表示できるブラウザは今のところ Internet Explorer のみであることから、ブラウザからの要求に応じて情報を取り出し動的に配信する仕組みが必要となること、配信用データ取り出し用のソフトウェア・ライブラリに有償のものが多く、ライセンス費用の考慮が必要になることがデメリットとして挙げられています。また、開発・動作環境が Windows 系の OS に偏る傾向がある点も、考慮すべきポイントです。

画質の劣化が視覚的に目立たない圧縮率としては、JPEG で 1/8～1/20 程度、JPEG 2000 および JPEG XR では 1/16～1/40 程度を目安とすることが多いようです。

さて、フォーマットの選定に移りましょう。

まずコストについて考えてみます。今回作成する画像データは、前述のとおり 1 枚当たり 46M バイトの画像が 20,000 画像あります。1/20 に圧縮するとして、 $46M \times 20,000 \times 1/20 = 46G$ バイトが必要な計算です。市販のハードディスクなら 1 万円程度支払えば十分なサイズのものを購入することができますが、クラウドに配置するとなるとデータ量が毎月の経費に影響してくるので注意が必要になります。大手のオンラインストレージサービスでは現在のところ 1G バイトあたり年間¥2,500～¥10,000 の価格帯でサービスを行っているようです。これを参考にすると、1/20 圧縮を採用する場合には年間¥115,000～¥460,000 の経費がかかる計算となります。1/40 圧縮を採用するのであればこの経費は半分になります。1/20 圧縮した JPEG の画像と、1/40 圧縮した JPEG 2000 もしくは JPEG XR の画質はほぼ同じと見なせるので、圧縮性能の観点からは JPEG をやめて JPEG 2000 か JPEG XR を採用するのが得策と言えます。

一方で、配信時の計算量についても考える必要があります。JPEG ならば基本的なウェブサービスで十分である場合が多いのですが、JPEG 2000 や JPEG XR を使う場合には閲覧者の要求に合わせて動的にデータを作る必要がありクラウドではそこにも経費がかかってきます。価格はクラウドサービスごとにばらつきが大きい上に利用頻度の影響も受けますので一概には言えませんが、大手 PaaS の標準プランを参考にすると、年間¥76,000 程度となっているようです。

結局のところクラウドのコスト面だけを考えたのでは、圧縮性能の高い JPEG 2000 や JPEG XR を選べば良いのか、あるいは配信時の計算量が少ない JPEG を選べば良いのか、判断がつか

きません。どちらが有利かはクラウドサービスの料金体系に依存してくるのです。

なお、クラウドサービスの種類や特徴については、第 5 章 2(4)(a)「システム基盤を確定する - クラウドサービスの検討」にて解説していますので、参照しておいてください。

もっとも、今回はタブレット型携帯端末での閲覧を実現するために動的配信が必須条件になっていますから、いずれにせよ JPEG 2000 か JPEG XR かのいずれかのフォーマットを採用することになります。

ここでは、特に色域の向上や複雑度の低減は必要としていないため、大規模アーカイブの例に倣って JPEG 2000 を採用することとします。配信用画像の圧縮率はとりあえず目安の最小値である 1/40 として、目視検査により不合格となった画像については 1/20 の画像と入れ替えることにします。

この入れ替えを簡便に実施するために、レイヤー数を 3 に改めます。保存用画像の画質、配信用画像の画質(1/40)、目視検査の結果入れ替える画像の画質(1/20)、と、1 つの画像データから、3 種類の画質の画像を取り出すことを考えての選択です。

以上により、作成する画像データの詳細が決まりました。ここにそれらをまとめておきます。

項目	内容
解像度	400 dpi
画像フォーマット	JPEG 2000 (ISO/IEC 15444-1)
色表現	YCbCr
解像度の重層数	3
画質の重層数	3 (可逆・1/20・1/40)

詳細が決まったら、その内容に問題がないかを確認するため、サンプルデータの作成と、画質の確認を行います。画像の作成を業者へ外注する場合は、その見積と併せて依頼することも考えられます。

画像作成業務の提案を募る際には、その審査においても、サンプルを提出させることは大変重要なポイントです。そのため、仕様に品質に関する記述を明記しておく必要があります。

画像データの仕様だけでなく、仕様の目的や、検討の過程を含めて仕様に明記しておく、理解が容易になります。

上記の仕様は次項にて仕様案として使用します。

(4) 画像データの投入方法を検討する

大前提として、画像データの作成・入力に際しては、資料に極力ダメージを与えない方法を選択する必要があります。デジタル化したからといって原資料を捨て去るわけではありません。デジタルアーカイブ構築作業の効率を追い求めるあまり、資料を傷つけてしまつては本末転倒です。原資料の性質や保存状態を十分に考慮し、画像データの投入方法を検討しましょう。

画像データの入力作業は外注して実施するのが一般的ですが、予算が少ないのであれば館員が自分で入力することも選択肢の一つです。本例では平均 80 頁の書籍 500 冊が対象ですから、見開きで入力するなら、500 冊×40 画像で結局 20,000 画像の作成が必要であることが判っています。スキャナを使って 1 分間に平均 2 画像を入力するとすれば 167 時間、交代で 1 日 8 時間入力作業にあたるとすれば 2 1 日間の作業になる計算です。業務多忙の折厳しいものはありますが、今回は予算を重視して館員が自分で画像データの投入を行うこととします。

画像データの投入にはスキャナやパソコンといった機材と処理を行うためのソフトウェアが必要になります。これらを新規に調達することも選択肢のうちですが、近年ではスキャナ機能が付いたプリンタやパソコンがほとんどの施設に設置されていますので、可能であればそれらを流用することも選択肢になります。その場合にはそれらの機材の名称、スペック、オペレーティングシステムなどを要件に記載しておくことで、設置済みの機材が流用できる提案を求めることができます。将来機材を入れ替えても困らないように、機材固有の名称は例として書き、スペックを記載した上で「相当以上」と添えておくことで安心です。

動作環境

システムに対し、館員が操作を行う際に使用する PC・スキャナ等の環境を以下に示す。提案システムの各機能は、下記の環境において稼働可能とすること。

・ PC

CPU : Core i5 2.66GHz 相当以上 (例 : ソニー VPCZ139FJ/S)

メモリ : 4GB 以上

ハードディスク : 128GB 以上

OS : Windows 7 Home Premium Service Pack 1 64 ビット 相当以上

・ スキャナ

フラットベッド型 CIS 1200dpi 相当以上

(例 : キヤノン PIXUS MX420 (有線 LAN 接続))

クラウドサービスの機能を使用する場合、一般的にはブラウザ上で操作を行うことが多いですが、専用のアプリケーションを使用することもあります。あるいは、事前にパソコン上で動作するソフトウェアでファイルを加工してから、そのファイルをクラウドサービスに登録するという形態も考えられます。一般的に、ブラウザ上のみで操作する場合はシステム導入が簡単な点が長所であり、専用アプリケーションを使用する場合は、業務に特化した便利な機能や操作性が提供されるという長所があります。

ここでは、システム全体の機能として画像の登録ができれば、実現形態は問わないこととし、業者の提案にゆだねることにします。

画像データ登録機能

提案システムには館員が画像データを作成してクラウドに登録する機能が含まれているものとする。

本機能により生成される画像データは下記の仕様を満たすものとする。

- ・ 解像度 : 400 dpi
- ・ 画像フォーマット : JPEG 2000 (ISO/IEC 15444-1)
- ・ 色表現 : YCbCr
- ・ 解像度の重層数 : 3
- ・ 画質の重層数 : 3 (可逆・1/20・1/40)

(5) 書誌データの内容を検討する

閲覧利用者が 500 冊の書籍の中から 1 冊を選ぶのは骨の折れる仕事です。まして、今回はタブレット型情報端末を使って本を選んでもらうので、端末の画面に選択肢を表示して並べる領域にも限りがあります。そこで、検索の機能を提供して効率よく候補を絞ってもらう必要があります。そのような検索のサービスを提供するにあたっては、何を使って検索をするのが効率的で使い勝手が良いのか十分に検討しておく必要があります。

同様のサービスを見ると、概ね以下に示すような項目が検索対象に設定されているようです。

- タイトル
- 著者名
- 著者名の読み仮名
- 出版社
- 出版年
- 分類 (NDC⁶⁵など)

⁶⁵ 日本十進分類法 の略。図書分類法の一。米国のデューイ十進分類法を参考にして、日本で考案された十進分類法。(小学館『デジタル大辞泉』)

本例では、とりあえず上記項目を検索項目としておきます。

また、候補の表示にあたっては書籍の内容を分かりやすく伝える必要があります。そのためには検索項目に挙げられなかった情報を表示させる必要があるかもしれません。逆に、検索には使っても表示には必要ない項目があるかもしれません。更に、候補の一覧表示とは別にコンテンツ個別の詳細説明を表示させた方がよい場合もあります。

今回は一覧表示と詳細表示との2画面を設けて、それぞれ下記の項目を表示させることとします。

- 一覧表示
 - タイトル
 - 著者名
- 詳細表示
 - タイトル
 - 著者名
 - 出版年
 - 出版社
 - 内容紹介
 - 司書から一言

ここまでの情報をまとめて書誌情報にすることとします。これらの情報はデータベースに格納します。データベースの構成などは要求仕様を含める必要がないためここでは考えません。

検索機能

提案システムには閲覧利用者が検索を行って閲覧コンテンツの候補を絞り込む機能、およびその結果を一覧表示する機能が含まれるものとする。

本機能において検索対象とする書誌情報を以下に掲げる。

- ・タイトル
- ・著者名
- ・著者名の読み仮名
- ・出版社
- ・出版年
- ・分類

検索結果を一覧表示する機能において表示対象とする書誌情報を以下に掲げる。

- ・タイトル
- ・著者名

また、提案システムは書誌情報を詳細に提示する機能を含むものとする。

本機能において表示対象とする書誌情報を以下に掲げる。

- ・タイトル
- ・著者名
- ・出版年
- ・出版社
- ・内容紹介
- ・司書から一言

なお、提案システムにはこれら書誌情報を管理するデータベースが含まれるものとする。

(6) データベースにデータを投入する方法を決める

データベースに何を登録するかを決めたら、そこに対応するデータを投入しなければなりません。その方法もここで決めておきます。

まず書誌情報の入力について考えます。データの投入を外注する方法もありますが、今回は分量も少なく書誌の量も500冊と少ないので自分で入力することとします。そのためには書誌データを追加する仕組みが必要であり、データ投入用の機材とソフトウェアが必要になります。これは発注要件に含めなければなりません。ただし、多くの場合は既に館内にパソコンが設置されているはずですから、それを流用できるのであれば機材の購入をやめて、ソフトウェアの動作環境を発注要件に含めます。

次に書誌情報とコンテンツとの紐づけについて考えましょう。こちらはシステムの中身まで把握しないと入力のしようがありませんし、無理をしてそうした作業を行ったとしても誤りが発生しやすく効率が良くありません。画像データを投入後にこの処理を行うのであれば、書誌情報も画像データも揃っている状態なので処理を自動化できるはずですが、この処理も忘れずに発注要件に含めます。

書誌情報登録・更新・削除機能

提案システムにはデータベース上の書誌情報を登録・更新・削除する機能が含まれるものとする。

(7) 目録とコンテンツとの紐づけを検討する

閲覧利用者が表示中の書誌情報から画像情報に移動できるようにするため、書誌情報とコンテンツとを結びつける必要があります。また、本例は動的に画像を配信するため、コンテンツの各ページにあたる画像を指し示すためのデータベースが必要になります。前項と同様、データベースの構成などは要求仕様を含める必要がないためここでは考えません。

書誌情報とコンテンツとを紐づける機能

提案システムには閲覧利用者が書誌情報と紐づけられたコンテンツを閲覧する機能が含まれるものとする。また、コンテンツ内の画像情報はページ順に閲覧利用者に提示されるものとする。

提案システムにはこれらの管理情報を格納するデータベースが含まれるものとし、併せてクラウドに登録された画像データを書誌情報に紐づける機能も提供されるものとする。

(8) 目視検査の方法を検討する

「(3) フォーマットを選択する」では目視検査を行うことと、その結果画質が不十分であったと判断された場合に配信用画像データを差し替えることを決めました。ただし、まだその方法を決めていません。方法を考えて要求仕様に反映させる作業が残っています。

「(4) 画像データの投入方法を検討する」で既に検討した通り画像データの投入はクラウドに対して行われます。目視検査は、その過程で行うことも、投入が終わってから行うことも可能です。それぞれについてメリットとデメリットを考えてみます。

画像データを作成してクラウドに投入する過程で目視検査を行う場合のメリットとデメリットを考えてみるとざっと下記ようになります。

- メリット
 - 工程数が少ないため確認作業の漏れが起りにくい
 - 配信用データ差し替えの機能が不要となるのでシステムがシンプルになる
- デメリット
 - データ作成の工程が複雑になるため作業の精度が低下する
 - データ作成と目視確認の工程を分離できないため分担作業が困難になる
 - 目視確認の機会がデータ作成時に限られるため複数人による検査ができない

データ投入が終わってから目視検査を行う場合のメリット・デメリットはちょうど逆になります。置かれている状況や判断する人の価値観にもよりますが、どうもデータ投入後に目視検査を行うほうが、利便性が高いようです。従って本例ではデータ投入後に目視検査を行う方式

を採用することになります。

そうするとクラウド側に目視検査とデータ差し替えのサービスが必要になりますので、これを要求仕様に盛り込みます。

目視検査および画像データ差し替えの機能

提案システムにはクラウド上に配置された画像データを、端末からの指示に応じて目視で確認し、必要に応じて差し替えを行う機能が含まれるものとする。

(9) 閲覧の方式を検討する

さて、これでサービス開始にあたって用意する機能や必要な手順の定義が整いました。いよいよ目的とする閲覧サービスの定義です。

最初に述べたように、本例は館内にタブレット型端末を設置して、来館者に郷土資料の書籍の閲覧をサービスすることを目的としています。また、「(1) チェックシートを使ってシステム要件のスタート地点を確認する」において、クラウドを使ってサービスする方針を決めています。あとは閲覧用に準備するタブレット型携帯端末の選定です。

タブレット型端末の選定にあたっては、調達価格のほかセキュリティや操作性の考慮が必要になります。本例では普及が進み操作に慣れた利用者が比較的多い iPad2 を 4 台用意し、3 台を館内閲覧用に、1 台を予備とすることになります。

なお、このサービスは来館者にのみ行うことを目的としています。それ以外の iPad2 利用者に勝手に閲覧されることを防ぐ目的から、要求仕様では認証についても言及します。

閲覧機能

提案システムにはクラウド上に配置された書誌情報およびそれに紐付けられた画像データを館内に設置したタブレット型携帯端末に配信し閲覧させる機能を含むものとする。また、館外の端末からのアクセスに対してはこれを制限する機能を合わせて提供するものとする。また、本提案システムには閲覧用のタブレット型端末として Apple 社製 iPad2 が 4 台含まれるものとし、併せて館内に既設の無線 LAN 環境への接続設定作業および閲覧ソフトウェアが含まれるものとする。

(10) データ管理の方法を検討する

さて、ここまで登場したデータを整理してみましょう。大きく分けるとデータベース上に格納するものと画像情報を取り出すデータとが登場しました。

クラウド上に配置されたデータベース上には書誌データおよびそれと画像データとを紐付けるデータとがありました。それらはクラウド上のサービスで使われるのみであるため、他に配置する積極的な理由は見当たりませんが、少なくとも災害や事故に備えたバックアップを持つておく必要はあります。バックアップ用のデータは稼働場所と異なる場所（できるだけ遠く離れた場所）に配置する必要があります。今回は定期的に容量の大きな光ディスクにコピーを作成して館内に保管することにします。

一方の画像データについてはさらに複雑です。オフラインで管理すべき長期保存用のデータ、目視検査を反映した差し替えのためにクラウド上に配置するデータ、そして同じくクラウド上に配置される配信用のデータ、の3種類について考えなくてはなりません。

幸い今回は長期保存用の画像データから簡単に配信用のデータを取り出せる仕様のフォーマットを選択していますので画像データについてはバックアップを考慮する必要がありません。しかし、一方で差し替えのためにクラウド上に配置するデータについては経費節減のため、不要になり次第削除する必要が生じ、そのためにはこの画像を消去する機能が必要となります。長期保存用の画像データはデータベースと同様に大容量の光ディスクに格納して館内に保存するものとします。

なお、光ディスクとしては現時点で最大容量である 128GB の BDXL を使用するものとし、必要な機材も同時に調達するものとし、これを既出のパソコンに接続して使用することとします。

ただし、BDXLに限らず、採用した技術規格が早期に陳腐化してしまうことも考えられます。そのような場合に代替手段が取れるように、要求仕様は「任意の容量で分割してバックアップ用のファイルを生成する」といった機能にしておきます。

データ管理機能

提案システムにはクラウド上に配置されたデータベースを読み込み、任意の容量に分割して別媒体上にバックアップする機能が含まれるものとする。

また、提案システムには生成した画像データを長期保存用として 128GB 以上の容量をもつ光ディスク等に格納するための機能が含まれるものとする。

さらに、提案システムは、クラウド上に配置した差し替え用画像データを任意の時点で削除する機能を提供するものとする。

これらの機能を実現するため、本提案には 128GB の BDXL ディスクの読み書きが可能で「動作環境」で記述した環境に接続可能なディスクドライブの調達が含まれるものとする。

なお、バックアップは「動作環境」で記述した端末上で当面実施するが、端末を変更することもあり得る。調達するディスクドライブは、端末を変更しても極力使用可能であるような、汎用的な接続方式とすること。

(11) 利用状況の把握について検討する

以上で目的を遂行するための要求項目は揃いました。残るはこのサービスの評価を行うにあたって必要となる情報の取得についての考慮です。ここではその指標として全体のアクセス数およびコンテンツごとのアクセス数を用いることとします。

アクセス機能

提案システムにはコンテンツごとおよび全体のアクセス数を月ごと、曜日ごと、時間帯ごとに集計する機能が含まれるものとする。

集計結果は画面で閲覧できるほか、集計データをダウンロードできる機能を有すること。

(12) 全体を見直す

これで一通りの要件定義が終わりました。全体を見直して要求項目の過不足や不整合を検討します。必要に応じてセキュリティ要件などを追加すると良いでしょう。

(13) 予算感を得る

要求仕様書がまとまった段階で、可能であれば入札に関わらない、複数の関係者に相談するなどして予算感を得ます。予算感のほか、複数業者による入札が可能な仕様になっているかどうかを確認できると良いでしょう。ここからそれぞれの組織に適った調達のプロセスが開始されることになります。

以上

別添資料 「知のデジタルアーカイブに関する研究会」名簿

(敬称略、五十音順)

新 麗 (あたらしい)	株式会社 IJ ソリューションインスティテュート技術研究所主幹研究員
安達 文夫 (あだちふみお)	国立歴史民俗博物館 (大学共同利用機関法人人間文化研究機構) 教授
入江 伸 (いりえしん)	慶應義塾大学メディアセンター課長
植村 八潮 (うえむらやしお)	社団法人日本書籍出版協会理事
大内 英範 (おおうちひでのり)	東京大学史料編纂所特任助教
大場 利康 (おおばとしやす)	国立国会図書館関西館電子図書館課長
岡本 明 (おかもとあきら)	NPO 法人知的資源イニシアティブ理事、株式会社寿限無代表取締役
小川 恵司 (おがわけいじ)	凸版印刷株式会社事業開発・研究本部総合研究所情報技術研究室室長
加茂 竜一 (かもしゅういち)	一般財団法人デジタル文化財創出機構研究主幹
神門 典子 (かんだのりこ)	国立情報学研究所教授
杉本 重雄 (すぎもとしげお)	筑波大学大学院図書館情報メディア研究科教授
武田 英明 (たけだひであき)	国立情報学研究所学術コンテンツサービス研究開発センター長・教授
田中 久徳 (たなかひさのり)	国立国会図書館電子情報部電子情報企画課長
田良島 哲 (たらしまさとし)	東京国立博物館学芸研究部調査研究課書跡・歴史室長
常世田 良 (とこよだりょう)	社団法人日本図書館協会理事
鳥越 直寿 (とりごしなおひさ)	メタデータ情報基盤構築事業メタデータ情報基盤事業検討会委員
丸山 信人 (まるやまのぶひと)	社団法人日本雑誌協会デジタルコンテンツ推進委員会幹事
水谷 長志 (みずたにたけし)	独立行政法人国立美術館本部情報企画室長/東京国立近代美術館企画課情報資料室長
宮澤 彰 (みやざわあきら)	国立情報学研究所教授
盛田 宏久 (もりたひろひさ)	大日本印刷株式会社教育・出版流通ソリューション本部デジタル推進部部長
山崎 博樹 (やまざきひろき)	秋田県立図書館主任図書専門員兼企画広報班長
八日市谷 哲生 (ようかいちやてつお)	独立行政法人国立公文書館公文書専門官

(オブザーバー参加)

文部科学省生涯学習政策局社会教育課

文化庁文化財部伝統文化課

経済産業省商務情報政策局文化情報関連産業課

(庶務)

総務省情報流通行政局情報流通振興課