

長期保存・長期利用保証の検討について(進め方)

1. 目的

「第二次中間取りまとめ」(平成 31 年 4 月 実務者検討委員会)において、長期保存・長期利用保証の在り方については、本委員会で引き続き主要検討課題として取り組むこととしており、また、実際にアーカイブ機関が取り組むべき内容がわかるようガイドラインが必要との意見が出されている(参考資料 2)。そこで本年度は、長期保存・長期利用保証に関するガイドラインの作成に取り組むこととする。また、必要に応じて、「デジタルアーカイブアセスメントツール」の該当部分(参考 1)の修正も行う。

2. 検討の進め方

課題整理のため、長期保存・長期利用保証に係る取組の現状と課題について、構成員(機関)へのヒアリングを実施する。実務者検討委員会でヒアリング結果を共有した上で検討を行い、令和元年度の成果物として「長期保存・長期利用保証ガイドライン」を取りまとめる。取りまとめに当たっては、平成 29 年 4 月「デジタルアーカイブの構築・共有・活用ガイドライン」(参考 2)の内容も含めたものとする。

3. スケジュール

令和元年 8～10 月	第一次ヒアリングの実施
11 月	第 8 回本委員会で課題整理(※本日)
11～12 月	第二次ヒアリングの実施
12 月頃	第 9 回本委員会でガイドライン(素案)の検討
令和 2 年 3 月頃	第 10 回本委員会でガイドライン(案)の検討
4 月頃	ガイドラインの公開

4. 検討状況

現在までに、第一次ヒアリング対象として、次の 6 構成員(機関)からヒアリングを行った。

国立国会図書館
国立公文書館
国立美術館
国立映画アーカイブ
東京国立博物館
国立歴史民俗博物館

5. 第二次ヒアリングのお願い

その他の各構成員(機関)の取組についても、情報収集のご協力をお願いしたい。

(候補)

国立科学博物館
放送番組センター
日本放送協会

(参考1) デジタルアーカイブに関するアセスメントツール

	アーカイブ機関		つなぎ役モデル
	標準モデル (小規模な機関で行うことが推奨される水準)	先進モデル (各機関のミッション等の必要に応じて目指す水準)	(分野・地域コミュニティのつなぎ役の役割をもつ機関が目指す水準)
5 持続可能性の担保	※ガイドライン pp.11-13 2(4)長期アクセスの保証のために		
①メタデータ管理 (識別子付与)	長期にわたり安定的にデータを保存・管理するため、個別のコンテンツを判別し認識できる識別子(重複しない一意の管理番号)を付与している。	長期にわたり安定的にデータを保存・管理するため、個別のコンテンツを判別し認識できる識別子(重複しない一意の管理番号)を付与している。	当該コミュニティ内に属する組織・団体に対し、機関識別子の付与・管理を行っている。また、当該コミュニティに対し、コンテンツへの永続的識別子の付与・普及に努めている。
②アクセス保証 (メタデータ)	公開メタデータについて詳細表示ページを用意し、永続的な固定 URL でアクセスできる。(一部でも可。つなぎ役等のポータルを通じて提供されている場合を含む。)	公開メタデータ全てについて詳細表示ページを用意し、DOI(デジタルオブジェクト識別子)等の永続的識別子を付与するなどして長期アクセス保証を意識して公開している	公開メタデータ全てについて詳細表示ページを用意し、DOI(デジタルオブジェクト識別子)等の永続的識別子を付与するなどして長期アクセス保証を意識して公開している
③アクセス保証 (コンテンツ)	公開しているサムネイル/プレビュー又はデジタルコンテンツは、永続的な固定 URL でアクセスできる(つなぎ役等のポータルを通じて提供されている場合を含む。)	公開しているサムネイル/プレビュー又はデジタルコンテンツは、永続的な固定 URL でアクセスできる(つなぎ役等のポータルを通じて提供されている場合を含む。)	公開しているサムネイル/プレビュー又はデジタルコンテンツは、永続的な固定 URL でアクセスできる(つなぎ役等のポータルを通じて提供されている場合を含む。)
④データ移行性	メタデータも含め、サムネイル/プレビューやデジタルコンテンツの管理において、特定の機器(システム、メディア)等に依存しないデータ形式で保持している、又はデータの抽出を可能としている。	データ形式の確保に加えて、デジタルアーカイブで提供しているデータについて、組織統廃合時のデータ移行や他機関へのデータ譲渡に対応できるよう、包括的な権利処理を行っている。	データ形式の確保に加えて、デジタルアーカイブで提供しているデータについて、組織統廃合時のデータ移行や他機関へのデータ譲渡に対応できるよう、包括的な権利処理を行っている。
⑤データ管理	データの更新(修正・削除等)作業や抽出作業の取得等を簡便に行うことができるようにしている。	データの更新(修正・削除等)作業や抽出作業の取得等を簡便に行うことができるようにしている。	データの更新(修正・削除等)作業や抽出作業の取得等を簡便に行うことができるようにしている。
⑥保存フォーマット	デジタルコンテンツの保存用に作成するマスターデータについて、特定製品等に依存せず、仕様書等が公開され、かつ広く普及している(国際標準等で定められた)データ形式を採用している。	デジタルコンテンツの保存用に作成するマスターデータについて、特定製品等に依存せず、仕様書等が公開され、かつ広く普及している(国際標準等で定められた)データ形式を採用している。	
⑦データ保存	バックアップによる提供データの複製、データ保存場所の分散等により万が一に備えた保存の体制を整えている。	(システムリプレースも視野に入れた)保存計画を立てた上で、提供データに加え、保存用のマスターデータについても、バックアップや保存場所の分散等により万が一に備えた保存の体制を整えている。	(システムリプレースも視野に入れた)保存計画を立てた上で、提供データに加え、保存用のマスターデータについても、バックアップや保存場所の分散等により万が一に備えた保存の体制を整えている。
⑧システム安定性			当該コミュニティに属するアーカイブ機関自らのデータの整備・公開ができるよう、安定的な統合プラットフォームを構築・運用している。

(参考2:「デジタルアーカイブの構築・共有・活用ガイドライン」より)

(4) 長期アクセスの保証のために

(識別子の付与方法)

長期にわたり安定的にデータを保存・管理するため、アーカイブ機関は、個別の資料・作品や情報を判別・認識できる識別子(2(1)で示した管理番号)を付与することが求められる。

識別子の付与には、2種類の方法がある。組織内で決めたルールに基づき、一つは手作業で一意的となる(重複しない)管理番号を付与する方法、もう一つは機械的に管理番号を付与する方法である。手作業で管理番号を付与する場合は、管理番号が重複しないように運用することが求められる。識別子(管理番号)は数字でもそれ以外でもよいが、国際的な流通を意識する場合、漢字等は使わない方がよりよい。

表形式により手作業でメタデータを整備している場合は、第一列目に、同一ファイル内で重複しない管理番号をつけるなどすればよい。ファイル内で一意であれば、「ファイル名」+「管理番号」を機関における識別子として利用できる。複数のファイルを複数のフォルダごとに管理している場合には、それらを「/(スラッシュ)」でつないで「フォルダ名/ファイル名/管理番号」などとすればよい。

機械的に管理番号を付与する場合は、一意に特定するための管理番号を付与するルールとなるよう、プログラムなどを調節する必要がある。データベースシステムを用いてメタデータ等を整備している場合は、一意の管理番号が自動で付与されるように設定することがよい。

手動にせよ機械的な付与にせよ、一度付与した識別子(管理番号)は、原則として変更しないようにしなければならない。同じ番号を別のデータに付与するような使い回しをしてはならないし、機械的な付与においては、削除した場合に自動で新たな番号が振り直されることのないよう気をつける必要がある。

なお、メタデータを新規に整備する場合、自館の所蔵品目録や展示会図録等から全作品の一括付与を仮に行う方法なども考えられる。その識別子を用いて自ら保有する作品情報や解説等のデータを整備・公開するなどの活用が可能になる。

(識別子の種類)

識別子の付与は、データを共有する際の相互運用性の担保においても重要な取組である。URI(Uniform Resource Identifier)の付与については「3(5)データ共有の方法」を参照のこと。

このほか、デジタルコンテンツの識別子として、国際標準規格であるDOI(Digital Object Identifier)がある。アーカイブ機関が個別のコンテンツごとに URL を指定してアクセス可能なウェブページを提供している場合は、国際的な流通促進を考慮して、DOI を付与することも考えられる。DOI は永続的な識別子であるため、DOI のメタデータのメンテナンスが必要であるが、URL が変更されても長期アクセスが保証される点にメリットがある。DOI の付与には、我が国で唯一の DOI 付与機関であるジャパンリンクセンター(JaLC)¹の会員になる必要がある。

機関を特定するための識別子としては、国際標準規格である ISIL(International

¹ <https://japanlinkcenter.org/>

standard identifier for libraries and related organizations)があり、我が国では、国立国会図書館が付与・管理を行っている²。つなぎ役が機関コードの管理を行う場合に ISIL の活用が考えられる。

(システム持続可能性・安定性の確保)

アーカイブ機関は、デジタルアーカイブの持続的なアクセスの維持に関して、一定の配慮を行う必要がある。プラットフォームやサイトについて、安定性を確保するため、システム更新や、組織の統廃合によってデータが失われないよう、データ作成時・システム構築時から、画像・映像等のデジタルコンテンツやサムネイル/プレビューは、メタデータも含め、特定の機器(システム、メディア)等に依存しないデータ形式とし、データ移行性を確保することが重要である。ストレージ装置や各種デバイス、システム自体は数年ごとのリプレースが必須であり、そのための経費、さらにメンテナンスに従事する人員の確保も含め、運用面のコストをデジタルアーカイブ構築時にある程度見込んでおくことが求められる。

(長期アクセスの保証)

アーカイブ機関は、保有するコンテンツの保存に加え、メタデータやサムネイル/プレビューを含むデジタル情報資源の保存にも取り組むことが求められる。

2(1)で既述したメタデータ項目については、それぞれの項目の意味が将来的にも把握できるようドキュメントを整備し、維持する必要がある。

デジタルコンテンツの保存のためには、長期的な利用可能性を考慮し、特定製品等に依存せず、仕様等が公開され、かつ広く普及している(国際標準等で定められた)フォーマットを採用する必要がある。また、2(3)で既述したとおり、デジタル化の際のドキュメント等の記録を残しておくことも求められる。

さらに、メタデータ、サムネイル/プレビュー、デジタルコンテンツなど関連する全てのデジタル情報資源にいえることであるが、3章で述べるようにデータの共有を促進することにより、複数箇所にデータが存在する状況を作り出すことも、保存の有効な方策である。災害や大規模なシステム障害等への対応可能性が高まることから、データの分散化・複数化を進めることが求められる。

なお、長期的な保存を考えたアーカイブの仕組みに関する標準については、欧米では OAI参照モデル(参考資料「確認すべき標準・ガイドライン等」参照)を基にした各種ガイドライン等が整備されており、我が国でも留意することが望ましい。

² <http://www.ndl.go.jp/jp/library/isil/>

なお、現在 ISIL は図書館中心の登録であるため、他の分野に属する機関の登録への広がりが望まれる。