

# 映画データの長期保存を目的としたデジタルアーカイブシステムの試作

川島 健一<sup>1)</sup> 中西 智範<sup>2)</sup> 岡本 直佐<sup>2)</sup> 大関 勝久<sup>2)</sup> 竹村 裕樹<sup>1)</sup> 杉野 博史<sup>1)</sup>

<sup>1)</sup>株式会社エヌ・ティ・ティ・データ <sup>2)</sup>東京国立近代美術館フィルムセンター

東京国立近代美術館フィルムセンターは、映画データの長期保存という課題に直面している。この課題に対応すべく、情報システムの要件と、保存対象のデータに付与すべきメタデータ、およびその構成について、情報の長期保存のためのアーカイブを定義する OAIS 参照モデルを参考に、検討を行った。本稿ではその検討内容と、デジタルアーカイブシステムの試作について報告する。

## Prototyping of digital archive system for the long term preservation of film data

Kenichi KAWASHIMA<sup>1)</sup>, Tomonori NAKANISHI<sup>2)</sup>, Naosuke OKAMOTO<sup>2)</sup>, Katsuhisa OHZEKI<sup>2)</sup>, Yūki TAKEMURA<sup>1)</sup>, Hiroshi SUGINO<sup>1)</sup>

<sup>1)</sup>NTT DATA Corporation, <sup>2)</sup>National Film Center, The National Museum of Modern Art, Tokyo

National Film Center, The National Museum of Modern Art, Tokyo is facing the challenge of long term preservation of film data. To cope with the challenge, we research the requirement for information system, metadata which should be added, and their structure, considering OAIS reference model which defines the archive for long term preservation of information. In this paper, the result of that research and prototyping of digital archive system are reported.

### 1. はじめに

東京国立近代美術館フィルムセンター（以下、NFC）は、わが国で唯一の国立映画機関として、国内外の映画フィルムの収集、保存、復元と、これらについての調査・研究を行っている。その一環として、映画コンテンツの恒久的な保存と利用を目的として、そのデジタル化、およびデジタル修復事業を実施しており、その過程では様々なデジタルデータが生成される<sup>1)</sup>。また、現在多くの映画作品が、撮影から上映まで一貫してデジタルデータによって運用されている状況を踏まえると、近い将来にはボーンデジタルコンテンツの収集、管理も必要になることが想定される。

一方でこのようなデジタル化の流れに対し、デジタルデータの長期的な保存方法は確立されておらず、危機的な状況にあるという、所謂「デジタルジレンマ」の課題が提起されている<sup>2)</sup>。以上を踏まえると、映画コンテンツの長期保存に資する映像データの管理方法の確立が急務である。なおアナログ記録媒体としてのフィルムについては、温度・湿度を最適に保つことのできる専用の施設を備えることにより媒体の長期的保存に対応しているが、デジタル化とデジタル修復も保存

されたフィルムに対して実施されるため、これらを統合管理できることが望ましい。

これらの課題に対し NFC の BDC プロジェクト<sup>3)</sup>では、フィルムアーカイブにおける映像データの長期保存システムの調査研究事業を進めており、本稿ではその第一段階としての、映像データの長期保存のためのデジタルアーカイブシステムの試作について報告する。その業務要求としては、「映画コンテンツの100年以上先の理解と利用を可能とするための、メタデータと映像データの管理」と定義した。対象とする映画コンテンツの記録方式は、フィルムとデジタルデータの双方を含む。先ずもって長期保存の期間を100年としたのは、アナログフィルムを適切に管理すれば、カラーであればその期待寿命は100年とされるためである。※ ISO18901(2010)参照

試作の検討にあたっては、OAIS 参照モデル<sup>4)</sup>をインプットとした。OAIS 参照モデルは、情報の長期的な保存を目的として策定された標準であり、ISO14721(2003)として認証されている。情報システムと組織を含む、システムの枠組みを対象としており、アーカイブの責任や備えるべき機能、保存する情報について定義している。

## 2. 情報システムへの要求

### (1) 保存情報の理解容易性の担保

試作にあたっては、OAIS 参照モデルが規定する要求の一つとして、「保存する情報はそれ単独で、情報生産者の助けを借りることなく、指定コミュニティにとって理解できること」を重視する。モデルではその実現のためとして、保存する対象の情報に対して付与すべきメタデータが定義されており、それらをパッケージ化したものを「情報パッケージ」として定義する。(図 1 に示す。「コンテンツデータ」が保存対象であり、それ以外の要素はそのメタデータであると言える。)これは、単に保存対象の情報が維持されるのみでは十分でなく、それが何を意味するのかについて記述する情報とセットで保管することで初めて、長期にわたって対象を「理解」および「利用」できるとの考えに基づくものであり、本試作システムにおいても基本的な要件となる。

なお試作にあたって前提とする「指定コミュニティ」としては、「映画コンテンツの恒久的な保存と利用」という目的を共有する他のフィルムアーカイブとの相互運用も視野に、将来的には FIAF 加盟の各フィルムアーカイブを想定するが、まずは NFC 所属のアーキビストを、長期保存する情報の「理解」と「利用」の主体と想定する。

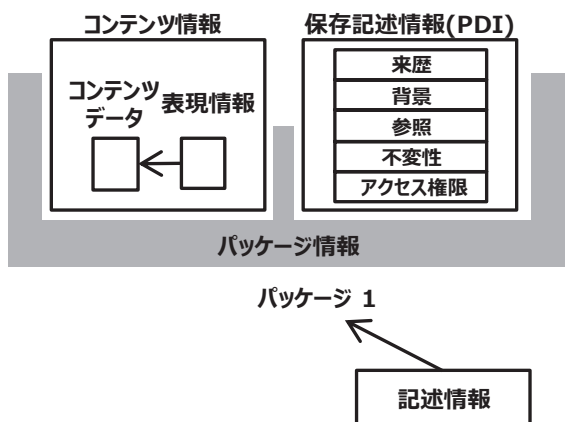


図 1：情報パッケージ概念図

### (2) 保存情報の発見可能性の担保

国際フィルムアーカイブ連盟 (以下、FIAF) 5) は世界の映画アーカイブ機関が加盟する国際組織であり、映画の文化遺産としてのアーカイブ、次世代への継承等を活動目的としている。FIAF は 1991 年に、フィルムアーカイブのためのカタログリングルールを策定した 6)。その目的は、全世界の映画コンテンツに関する目録データの相互コミュニケーションを支援することであった。図書館目録は主に出版と版数の発見をサポートするために構築されているが、映画の目録はプレプリント、マスタープリント、未発表素材のようなユニークもしくは希少性の表現を含むことがあ

り、その点で差別化されている。2016 年に改訂された最新のルール 7) では、近年の映画アーカイブ分野における、技術的な進歩による環境の変化を取り込んで、Work-Variant-Manifestation-Item のエンティティで構成される拡張 4 階層モデルを策定した (図 2)。これは FRBR 8) の概念モデルを参考にしており、従来のフィルムだけでなく、ビデオテープやデジタルファイル (LTO, Blu-ray 等) についても表現が可能となった。

OAIS 参照モデルでも規定される、保存対象の情報の発見手段としては、アーカイブにおける目録データの構築と実装が不可欠である。指定コミュニティの前提に照らすと、試作システムではこの FIAF のカタログリングルールを採用することが妥当である。

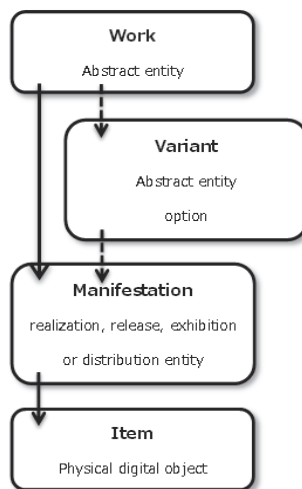


図 2：FIAF による拡張 4 階層モデル

### (3) 大容量データの保存

映画コンテンツのデジタル映像データは、ドキュメントや画像に比べて 1 つのコンテンツが占めるデータ容量が格段に大きくなる。上映用圧縮データの生成元となる非圧縮のデータであれば、映画 1 作品分が数テラバイトに至ることが通常である。これらの保存を想定する以上、大容量のデータ保存手段が必要である。保存媒体についてはその単価が下がってきているとは言え、容量に比例してリソースが必要であり、そのコスト抑制が求められる。

## 3. 長期保存システムのサブシステム構成

情報の発見手段としての目録データモデル、および「情報パッケージ」に必要な各種のメタデータは、本事業における最適な設計の試行に加えて、今後の長期的な運用も考慮すると、長期保存システムのポリシーに従って、かつ特定の製品に依存しない形で、自由度高く実装、変更されなければならない。指定コミュニティのナレッジの継承や、

NFC 外のフィルムアーカイブへの拡大も考慮すると、ポリシーでメタデータスキーマや語彙セットを規定し、必要に応じて保守しなければならないためである。それらの標準についても、年々改訂されると想定される。また、特に目録データについては、頻繁なデータアクセスとデータ更新の要求に対して、高速に応答する必要がある。

一方で前述のとおり、映像データの保存には大容量の保存媒体が不可欠であるが、技術動向とコストを考慮して、その時々において最適な装置を選択できることが望ましい。従って、少なくとも保存のための装置は、サブシステムとして分離することが必要である。

以上より長期保存システムの構成としては、各種メタデータを柔軟に定義、実装できるカタログシステム製品をサブシステム C として、映像データを保管するための製品と組み合わせることによって実現することとした。試作においては、コスト面で優位性のある LTFS ストレージシステム（サブシステム A）と、映像のデジタルデータを保管する CMS（サブシステム B）をこれに充てることとした。（図 3）前者はファイルシステム領域を提供するものであり、後者は SOAP 等の API を備える事例である。これらと連携するサブシステム C としては、メタデータスキーマと語彙、およびその変換プログラムを柔軟に定義することのできる、多様な入出力 API を備える AMLAD9)を採用した。AMLAD は OAIS 参照モデルに準拠した、コンテンツ管理および提供の機能を備えるデジタルアーカイブ製品であるが、本事業ではそのカタログ機能部分を提供する。

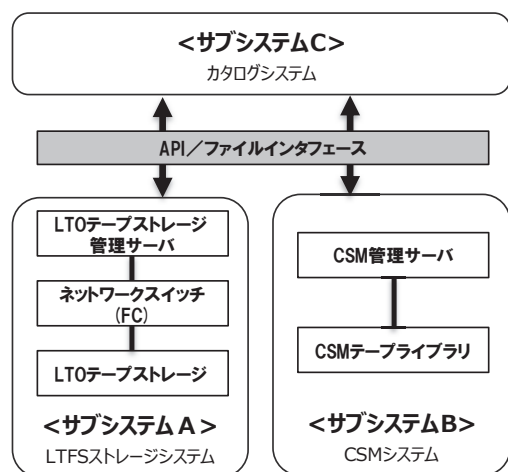


図 3：長期保存システムの構成

#### 4. 映画目録データモデルの実装

発見可能性の担保の要求に従い、拡張 4 階層モデル（図 2）を基本として目録データモデルを、サブシステム C において実装した。Work/

Variant, Manifestation, Item の 3 つのクラスを定義し、それぞれ XML ベースのメタデータスキーマを備えた。メタデータ要素としては拡張 4 階層モデルが定義する ID, 各種 Type, Title 等の基本セットに加え、「注記」、「読み」等の NFC で必要な管理項目についても追加を行った。加えて各階層間は ID の持ち合いによって、関連を表現した。

アーカイブ情報としてのメタデータは、クラス定義ごとに用意された関係データベース上のテーブルに格納される。製品の機能により、メタデータスキーマの各フィールドから必要な検索インデックスを別途生成することにより、メタデータを直接走査せずに検索応答を行う。

#### 5. 情報パッケージの実装

##### (1) パッケージ化手法

OAIS 参照モデルでは情報パッケージのパッケージ化の具体的な手法については規定していないが、大きく分類すると、コンテンツデータを内包する「物理的パッケージ」と、コンテンツデータへの参照を内容とする「論理的パッケージ」が想定される。ただし保存対象がフィルム等の物理媒体である場合は、必然的に後者となる。例えば米国会議図書館の研究プロジェクト事例 10)では、物理的パッケージ手法として検討を行い、情報の保存と送受信において、自己展開可能な Zip フォーマットを推奨している。また、あるオブジェクトに関する記述、管理、構造の各メタデータを 1 つの XML に内包するメタデータスキーマ標準であり、OAIS が規定する情報パッケージとして利用できるとされる METS 11)は、対象のオブジェクトがデジタルコンテンツである場合、そのファイルのバイナリデータを XML に内包するか、参照のみを記録するかを選択することができる。バイナリデータを記録する場合は METS 内で情報パッケージが完結し、これは物理的パッケージに分類できる。

本試作においては、3 章で示すシステム構成下では、各種メタデータと映像データの管理は、異なるサブシステムで実行されるため、映像データについての情報パッケージのまとまりとしては、図 4 の通りとなる。しかしこのように一つの情報パッケージを構成するデータがサブシステムを跨いで分散していると、システム障害によってその一部が失われる可能性は高くなる。また媒体によるデータの流通を考慮しても、映像データとメタデータを同一の実体として備えておくことが望ましい。よって、情報パッケージに望まれる本来の機能である、それ単体での理解容易性と利用可能性の担保を念頭に置き、サブシステム C より映像データ側に必要なメタデータをコピーおよび同期し、物理的情報パッケージを形成する。

（図 5）具体的には、サブシステム A においては、

同一ディレクトリ内に映像データとメタデータを格納する。サブシステム B に対しては、API を通じてメタデータを送付し、これを単独オブジェクトとして格納の上、映像データとまとめて AXF (Archive eXchange Format) ファイルとして成型する。

なお、国内では国立国会図書館デジタルコレクション 12)において、画像コンテンツをメタデータと同梱した物理的情報パッケージを、要求に応じて出力する機能の事例がある 13)。

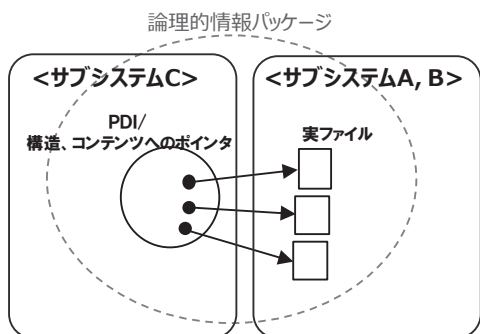


図 4：映像データの論理的情報パッケージ概念図

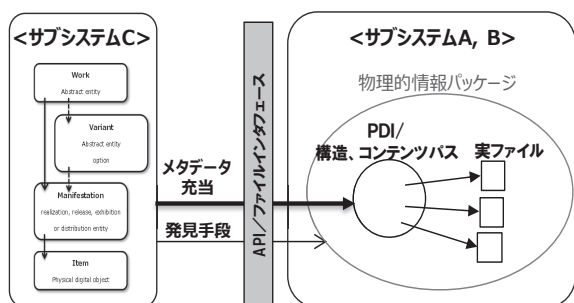


図 5：映像データの物理的情報パッケージ概念図

(2) 来歴

情報パッケージが備えるべきデータのうち、保存記述情報の「来歴」は、保存対象の情報の真正性と信頼性を担保する。例えばフィルムをデジタル化してデータを生成した場合、デジタルデータの来歴として生成元のフィルムの識別子を持つことで、情報パッケージ単独でその由来が理解できる。試作システムにおいて来歴として記録すべき内容について、OAIS 参照モデルと、NFC のフィルムアーカイブ活動に係るワークフローをインプットとして、検討を行った。結果を表 1 に示す。(なお、「アセット」とは、フィルム以外の媒体に格納された映像コンテンツを指す。これには電子的な動画データファイルを含む。)

なお、フィルムアーカイブ活動のワークフローの特徴として、1つのコンテンツについてオリジナル素材を基点として、撮影、各種補正、保存、上映等の目的ごとに連続的にデータの変換を実施する点がある。これには、サブシステムCのカ

タログデータモデルにおいて Relationship, Event, Agent の各クラスを定義し、映像コンテンツの変換の前後の関係性を、イベントとセットで記録することで対応する。新しいデータの生成を伴わないイベントも含め、これらを情報パッケージに自動的にコピーすることで、来歴の実装とする。(図 6)

表 1：来歴として記録するイベント

No.	イベント	No.	イベント
1	フィルムの受入	11	フィルムのデジタル化
2	アセットの受入	12	マイグレーション
3	フィルムの上映	13	アセットの冗長化対策
4	アセットの上映	14	データ完全性チェック
5	フィルムの移動	15	媒体、フォーマットの陳腐化チェック
6	アセットの移動	16	アセットの見読性チェック
7	フィルムの外部貸出	17	デジタルコンテンツのフィルム化
8	アセットの外部貸出	18	メタデータの項目追加・変更
9	フィルム及びアセットの情報参照	19	参照用プロキシ作成
10	フィルムのマイグレーション (複製・不燃化)	-	-

映画目録データモデル

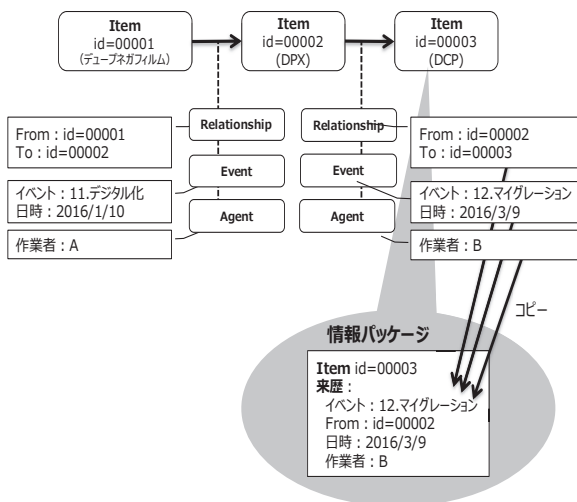


図 6：イベント情報の管理と来歴としての記録

(3) 表現情報

情報パッケージにおいて「表現情報」は、コンテンツデータと共にコンテンツ情報を形成する。

(図 1) その役割は、「指定コミュニティにとって、コンテンツデータを理解できるようにすること」とされ、具体的にはコンテンツデータがビット列であれば、それをデコードするために必要な情報である。これは Alex Ball 14)も指摘するように元来再帰的であり、情報パッケージ単体でその内容を理解するためには、表現情報自身にも、またその他各種のメタデータにも、その表現情報が必要である。例えばテキストデータで記述されたメタデータも、これを読み取るための文字コード情報が必要であり、その結果得られた日本語の内容理解のためには、日本語語彙の知識が必要と

なる。しかしこの点については、OCLC and CRL 15)が事例を示すように、外部の信頼できるフォーマットレジストリ(例えば 16))の仮定と、その利用に係る指定コミュニティのナレッジに依拠するのが現実的である。すなわち、映像データの表現情報としては、そのファイルフォーマットを一意に特定できる情報を格納する。

(4) フィルムに対する保存記述情報

試作システムが対象とするフィルムについては、特に古いものであれば基本的に所謂「一点物」であり、その修復の経緯はそれぞれに異なる。例えば『忠次旅日記』17)のように、過去に複雑な編集を経ているものもある。本事業ではシステムの試作にあたり、このようなフィルム修復の記録について、フレーム単位で管理可能な標準化を検討した。しかし有識者参加の研究会において、そのような記録のデータ量が膨大になること、また修復等作業の外部委託を考慮したときの、標準化の困難さについて指摘を受けた。

その代替として、外部委託時の事業者が作成するレポート類や、NFC 内部で作成するレポート等のドキュメントを、対象のフィルムに関連付く形で確認することができれば、想定する指定コミュニティにとって有効であると考えられる。よってドキュメントファイルのクラス定義と、フィルムに対応する情報パッケージから、ドキュメントクラスを参照するメタデータスキーマを実装した。

(図 7)

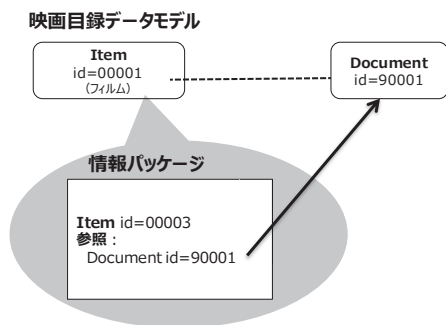


図 7：試作長期保存システム処理シーケンス

6. システム動作検証

(1) 業務シナリオ動通

以上の実装によって、試作システムの動作検証を行った。検証ではサブシステム C と B の間での SOAP API を用いた、XML ファイルの送受信によるシステム連携を前提とした。その問合わせは、サブシステム C から B に対して、アーカイブ対象データの情報の定期的な問合せを行い、差分をカタログに反映させる方式をとる。

実証ではカタログ情報の登録を経て、フィルムの修復、デジタル化や、ボーンデジタルデータの登録、フィルム化、上映等の基本ワークフローを

シナリオとしてシステムを動作させた。結果としてサブシステム間の協調処理によって、長期保存システムとして運用が可能であることを確認した。図 8 に、「アセットの受入」、および「メタデータの登録・修正」シナリオにおける、協調処理シーケンスを示す。

(2) 情報パッケージの生成

上記(1)の業務シナリオ動通の結果として、対象の情報パッケージ内のデータが、設計通り生成・同期されることを確認した。

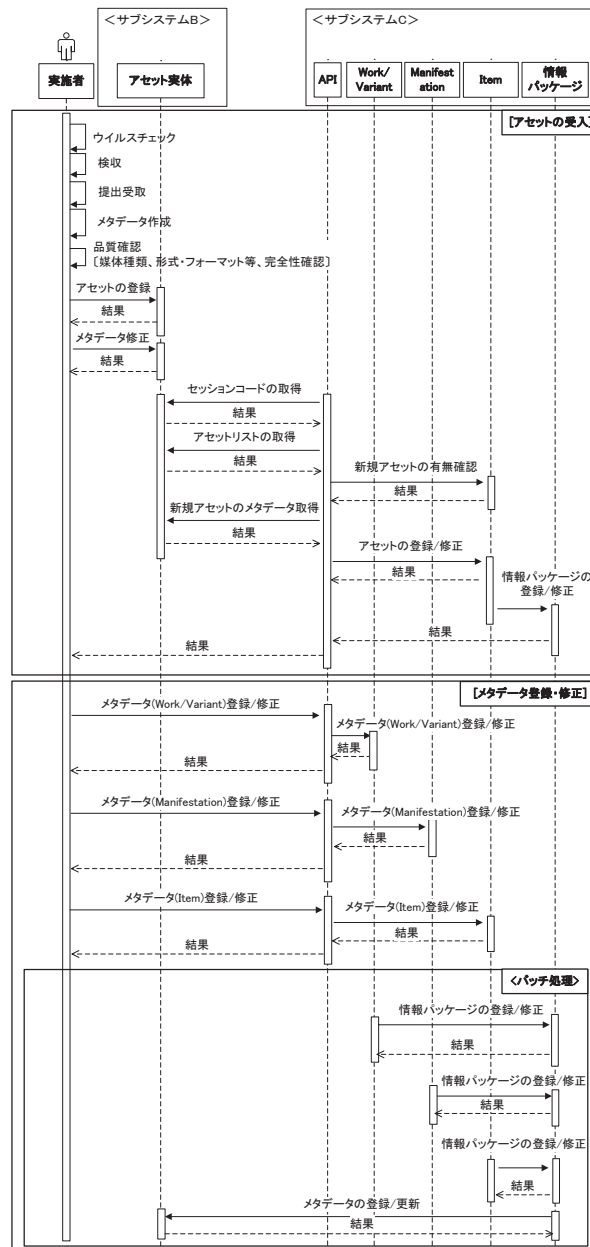


図 8：試作長期保存システム処理シーケンス (アセットの受入、およびメタデータの登録・修正)

## 7. おわりに

「映画コンテンツの100年以上先の理解と利用を可能とするための、メタデータと映像データの管理」という業務要求に対し、保存情報の理解容易性と発見可能性の担保、大容量データの保存というシステム要件を定義し、実装を行った長期保存システムの試作について報告した。サブシステム間の協調処理によって業務シナリオを検証し、情報パッケージの生成・同期を確認した。今後、以下の点についてさらに調査・研究を進める。

情報パッケージに定義された残りの要素のうち、「不変性」と「アクセス権限」として必要なデータを検討する。特に前者は、長期的にデジタルデータのビット列が破損していないかどうかのチェックを担い、具体的には保存対象の「コンテンツデータ」から算出したダイジェスト値が該当する。OAIS参照モデルは、その定期的なチェックを推奨するが、大容量である映画データについてのダイジェスト値算出が、現実的に実施できるかどうか、確認する必要がある。

次に、特に情報パッケージに記述するメタデータの、標準について検討する必要がある。本稿の報告時点では独自フォーマットによって実装しているが、指定コミュニティを、フィルムアーカイブ活動を担う他の機関に拡大していくためには、また、指定コミュニティの知識ベースの前提を、安全に軽減するためには、信頼できるメタデータスキーマ標準や語彙セットを定義もしくは採用する必要がある。

これらの課題に対応した上で、情報パッケージの、単独での理解容易性の検証を深める必要があると考えている。

**謝辞** 本事業にご協力いただき、貴重なご助言をいた「映画アーカイブにおける映像データの長期保存システム研究会」メンバーの方々に、感謝申し上げます。

## 参考文献

- 1) とちぎあきら：デジタルを通してフィルムが見えるーフィルム・アーカイビングの現場から、日本写真学会誌, Vol 72, No. 1, pp.16-22 (2009).
- 2) The Science and Technology Council of the Academy of Motion Picture Arts and Sciences : The Digital Dilemma (2007).
- 3) 岡島尚志：BDCプロジェクトの開始(2015) .  
([http://www.momat.go.jp/fc/wp-content/uploads/sites/5/2016/07/NFC121\\_BDCp02.pdf](http://www.momat.go.jp/fc/wp-content/uploads/sites/5/2016/07/NFC121_BDCp02.pdf)) (参照 2016-11-01).
- 4) The Consultative Committee for Space Data System : Reference Model for an Open Archival Information System (OAIS) Recommended Practice Magenta Book (2012).  
(<https://public.ccsds.org/Pubs/650x0m2.pdf>)

- 5) International Federation of Film Archives. (<http://www.fiafnet.org/>) (参照 2016-11-01).
- 6) International Federation of Film Archives : The FIAF Cataloguing Rules For Film Archive (1992).
- 7) International Federation of Film Archives : The FIAF Moving Image Cataloguing Manual (2016).
- 8) Munich K.G. Saur Verlag : Functional Requirements for Bibliographic Records, Final Report (1998) .
- 9) 株式会社エヌ・ティ・ティ・データ : AMLAD (Advanced Museum Library Archives Deposit). (<http://www.amlad.jp/>) (参照 2016-09-07).
- 10) Library of Congress : Archival Information Package (API) Design Study (2002).
- 11) Library of Congress : METS Metadata Encoding & Transmission Standard. (<http://www.loc.gov/standards/mets/>) (参照 2016-09-07).
- 12) 国立国会図書館デジタルコレクション. (<http://dl.ndl.go.jp/>) (参照 2016-11-01) .
- 13) 木目沢司：「国立国会図書館デジタルコレクション」のOAIS参照モデルへの準拠状況、情報管理, Vol 58, No. 9, pp.683-693 (2015) .
- 14) Alex Ball : Briefing Paper: the OAIS Reference Model(2006).  
(<http://www.ukoln.ac.uk/projects/grand-challenge/papers/oaisBriefing.pdf>) (参照 2016-11-01) .
- 15) RLG & OCLC : Trustworthy Repositories Audit & Certification: Criteria and Checklist(2007).  
([https://www.crl.edu/sites/default/files/d6/attachments/pages/trac\\_0.pdf](https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf)) (参照 2016-11-01).
- 16) HARVARD LIBRARY : Global Digital Format Registry.  
([http://library.harvard.edu/preservation/digital-preservation\\_gdfr.html](http://library.harvard.edu/preservation/digital-preservation_gdfr.html)) (参照 2016-11-01).
- 17) 板倉史明：『忠次旅日記』のデジタル復元、NFC ニューズレター, 2011年8月-9月号, pp.2 (2011)