When Digital Remembers Analogue

Conservation Metadata for Analogue Film as Preservation Description Information in a Digital Archive

Michelle Lindlar
TIB German National Library of Science and Technology
Germany
michelle.lindlar@tib.eu
0000-0003-3709-5608

Merle Friedrich
TIB German National Library of Science and Technology
Germany
merle.friedrich@tib.eu
0000-0001-7158-8583

Miriam Reiche
TIB German National Library of Science and Technology
Germany
miriam.reiche@tib.eu

Abstract – While "Digitization is not Digital Preservation", the two processes are often closely connected. This is especially true for audiovisual (AV), where the analogue materials, much like their digital counterparts, come in a variety of formats and are only accessible via rendering equipment such as projectors. This indirect accessibility of analogue audiovisual makes autopsy a resource-intensive process. Nevertheless, it's a necessary process prior to digitization, as the digital objects' quality depends on the quality of its analogue source material. But is analogue conservation data helpful in the preservation process of the digital object as well? Can it be used as Preservation Description Information in the OAIS sense? TIB, the German National Library of Science and Technology, has authored a Cconservation Metadata Schema for Analogue Film as part of the large digitization and digital preservation project DELFT. The schema includes metadata which TIB considers important knowledge for the contextualization and interpretation of the digital AV object. This paper outlines motivation and background for this development, discusses the connection of analogue and digital in an overarching analogue conservation and digital preservation workflow and gives a detailed insight into the TIB Conservation Metadata Schema.

Keywords – audiovisual preservation, conservation metadata, preservation description information

Conference Topics – Scanning the New Development; Exploring New Horizons

I. Introduction

One of the most frequently used mantras of our domain is: "Digitization is not Digital Preservation". Even as large national heritage institutions are into their second decade of digital preservation activities, they still find themselves needing to repeat this in conversations with users, consumers, stakeholders and funding bodies [1], [2]. On the other hand, it is also widely accepted that digitization in combination with adequate digital preservation can be a vital strategy for keeping materials accessible for the long term. Especially in the domain of AV preservation, where the original carrier material - such as nitrate - can pose a risk to the archive in general, digitization is a frequent strategy for preservation alongside or in addition to reformatting to a different carrier material.

Knowledge of the creation process of a born-digital object, in particular aspects such as creating software and scanner model used, is widely accepted as important information of a digital object [3]. The Open Archival Information System (OAIS) refers to this information as Preservation Description Information (PDI), which is *"The information which is necessary for adequate preservation of the Content Information and which can be categorized as Provenance, Reference, Fixity, Context, and Access Rights Information"* [4]. Within the standard, scanned books are used as an example for provenance information within Preservation Description Information, stating that information such as metadata about the digitization process and a pointer to the (analogue) master version are essential knowledge to be captured [4].

But is that all we need to know about a digitized object? As opposed to born-digital, a digitized object is another representation of an analogue counterpart. As such, the two representations are uniquely connected to each other - not only by inheriting the same intellectual entity descriptors, but also when considering quality factors. After all, any digitization result is directly dependent on the quality of its analogue source material. Even though the practice of capturing and documenting conservation information about analogue preservation objects is not new, a structured method in form of a lightweight metadata schema has been missing.

The following paper foremost addresses the research question of how conservation metadata of an analogue source object can serve as valuable Preservation Description Information for the digitized target in a digital archive. It furthermore introduces TIB's lightweight schema for conservation metadata for analogue film, thus closing the gap around a structured approach

17th International Conference on Digital Preservation
iPRES 2021, Beijing, China.

DOI: 10.1145/nnnnnnn.nnnnnnn

to capture and document this information and pass it to a digital archive.

In Section II we briefly describe the DELFT digitization project which provides the backdrop against which the connection between analogue conservation metadata and Preservation Description Information was explored. Section III looks at the relationship between the analogue and the digital object in the context of AV digitization and preservation. We describe the data models used for both analogue and digital and how the two forms of representation relate to each other. This is the necessary framework against which analogue conservation information can be identified as relevant Preservation Description Information for the digital object. The section concludes with a list of three use cases for conservation-based Preservation Description Information. Section IV introduces the TIB Conservation Metadata Schema which we authored to support the need for analogue conservation-based Preservation Description Information within the context of DELFT. We give an overview of our data model, data dictionary and XML Schema for a light-weight metadata schema. We conclude this paper with an outlook into further work on Conservation Metadata for Film in a digital archive in Section V.

II. Background and Motivation - TIB's DELFT Digitization Project

Between 2017-2019, TIB conducted DELFT (**D**igitalisierung **E**thno**L**ogischer **F**ilmbes**t**and), a digitization project of the ethnological film collection of the former Institution for Scienctific Film (IWF). IWF wound down in 2010 and all film material was transferred to TIB in 2012. The Ethnological Collection is a small subset of the IWF collection, nevertheless it is an important part of the *Encyclopaedia Cinematographica (EC)*, a collection of films started in 1952 by former IWF director Gotthard Wolf as a systematic filmic documentation of rituals and behaviors. EC films stem from various domains, ethnology being one of them [5]. In total 1.953 films were digitized as part of the DELFT project. The majority of the analogue source material was 16mm film in both black and white as well as in color, many with optical sound and additional separate magnetic carriers for audio. A very small portion of the DELFT material was in DigiBeta, which is currently not covered in the conservation metadata schema. [1] The project was of particular interest to TIB's conservation and digital preservation department, as it was the first larger scale AV-digitization project at TIB. While the largest part of the work - the selection, autopsy and documentation of the material as well as the quality assurance (QA) post-digitization - was conducted by TIB's conservator and digitization itself was outsourced, the digital preservation team was involved in the project from the start. In addition to determining the target digitization format and codecs[2], tailoring workflows for semi-automatic quality control, defining the Intellectual Entity (IE) levels for this collection[3], defining Significant Properties for Audio and Video[4] and deciding on the structure for the digitized objects, the digital preservation team worked hand-in-hand with the conservator to understand the impact that quality and structure of the analogue source material can have on the digital master. The result of this work is the Conservation Metadata Schema for Analogue Film presented in Section IV of this paper.

In addition to the brief project description above, the following things need to be noted to better understand the scope and context the Conservation Metadata Schema was developed in. As a national subject library, TIB is not a film archive and analogue AV only makes up a very small portion of TIB's analogue collections. This small collection of analogue AV material does not contain cinematographic works, but only scientific or educational materials. Due to these factors TIB's processes cannot be compared to those of large national film archives, where analogue material is analyzed and documented frame by frame and elaborate post-processing takes place. Instead, TIB's main focus is securing the intellectual content in the best possible state under strict resource limitations.

III. Relationship between Analogue and Digital Representation

To understand how conservation metadata of an analogue source material can be useful Preservation Description Information for a digital object, we first need to take a look at the relationship between the analogue and the digital copy and how the perception of that relationship has changed over the course of the past 20 years.
The 2002 IFLA/ICA "Guidelines for Digitization Projects" clearly states that a digital image is not a preservation master. However, the authors also recognized that a digital reproduction can aid in conserving the analogue master by reducing access to the analogue original [6]. Only two years later, in 2004, the Association of Research Libraries (ARL) endorsed digitization as a valid preservation reformatting strategy for paper-based materials, stating that *"The time is right to adopt digitization as a reformatting strategy for preservation"* [7].
Nevertheless, the debate around digitization as a suitable preservation form continued (and does so into the present day). The 2014 report "Film Heritage in the EU" states that: *"The film heritage of the 20th century [...] cannot be conserved for the future by using digital media. Digitisation is not a preservation measure, and photochemical originals shall always be preserved"* [8]. However, not all institutions dismissed digitization as an unsuitable strat-

[1] Please refer to the DELFT project page (in German language) for further information about the project: `https://projects.tib.eu/delft`

[2] See (in English language) `https://wiki.tib.eu/confluence/display/lza/Digitization+of+AV-Material%3A+Extent+and+parameters`

[3] See (in English language) `https://wiki.tib.eu/confluence/display/lza/Data+Management` for further information on TIB's Archival Data Management

[4] See (in English language) `https://wiki.tib.eu/confluence/display/lza/Significant+Properties`

egy for preservation. Instead, the focus of the debate shifted from an emphasis on reformatting to issues of usability, usefulness and quality of the digitized material [9]. This is echoed in the 2019 Memoriav Recommendations "Digital Archiving of Film and Video: Principles and Guidance" which point out that the focus of digitization seems to be more often on possible applications and usability of the digital version instead of on digitization for preservation. The authors argue that, in fact, digitization of analogue AV is an increasingly necessary action for cultural heritage institutions to take, as analogue AV technology is nearly obsolescent for many carrier formats. While the guideline stresses that analogue materials shall continue to be preserved under best possible conditions after digitization, Memoriav also recognizes that in cases such as lost or destroyed originals, physical degeneration or loss of technical means / knowledge to handle material, the digital object can become the only surviving copy [10].

What are the implications of these different positions on analogue vs. digital on digital preservation? If we accept the position that digitization is a valid reformatting strategy, then this strategy automatically comes with the responsibility for digital preservation. Following that thought, the digital might be considered merely another representation of the analogue in which case the digital representation needs to be as faithful to the original as possible. But how does that fit in with feature-rich possibilities of AV digital formats? While different language tracks were stored on separate data carriers in analogue, digital AV containers provide a framework which can hold many audio tracks between which the user can seamlessly switch during playback. Is the digital object indeed a representation of the analogue original? Or is it something different? In order to find answers to these questions we need to explorer modelling options for the analogue source and the digital target in an archival context.

A. Where in the file is reel 1? Modelling analogue and digital AV objects in an archival context

When first exploring the different analogue objects for DELFT from a digital preservation point of view, we were surprised by the different descriptive layers of each object. There was an overarching Signature that was valid for all different language versions of the film. Each language version had an own Media Asset Management Identifier (MAM-ID) And each film consisted not only of 1-n film reels plus 0-n sepmag audio carriers, but each of these reels and sepmags had multiple copies as well! To better understand the complexity of "a film" and its different levels, we turned to the Functional Requirements for Bibliographic Records (FRBR) model. FRBR is a conceptual descriptive model for bibliographic records introduced by IFLA in 1998 [11]. Though not without criticism, FRBR is widely adopted within the library world and also embedded in the cataloging rules RDA (Resource Description and Access). At the core of FRBR are the "Entities - Group 1", also referred to as the WEMI model, which define and describe hierarchical relationships between *Work*, *Expression*, *Manifestation* and *Item* (see Figure 1). This relationship has also been transferred to EN 15907 - "Film Identification - Enhancing Interoperatiblity of Metadata - Element Sets and Structures", to better suit cinematographic works. The EN-standard removes the *Expression* level and introduces a *Variant* entity in its place [12]. As our collection does not cover cinematographic works, but scientific films we focus on WEMI in this context of our analogue AV collection and within the scope of this paper.
WEMI can be summarized as different hierarchical levels of products of intellectual or artistic creation. The IFLA report lists a number of attributes that may be captured for each level.

1. Attributes of the WEMI levels

Figure 2 shows a mapping of the DELFT objects to the WEMI entities. *Work* is the overarching intellectual creation - in our case ä film about Criollo (Venezuela, Carabobo) - Dance of the Herdsmen at Mariara". FRBR attributes for the *Work* entity are *title and form* (e.g., "Criollo (Venezuela, Carabobo) - Dance of the Herdsmen at Mariara", "Film"), *numeric designation* (e.g., Signature "E 1234") and *intended audience* (e.g., "IWF - Encyclopaedia Cinematographica")[5]. These attributes can be considered the most basic information about the intellectual object and are naturally of high value to the digital form as well. Especially the relation between "intended audience" and the well-known digital preservation concept of the "designated community" are notable attributes. On the analogue side, the *Work* is broken down into *Expressions*. As briefly presented in Section II several language versions may be available for each film. The FRBR attribute *language* is simultaneously the differentiating factor between separate *Expressions* on the analogue side as well as key metadata for the DELFT collection on the digital side. [6] While attributes at the *Work*

[5] In the case of the DELFT object the intended audience is inherited from the collection "IWF - Encyclopaedia Cinematographica", i.e. scientists and students from the domain of ethnology interested in rituals and behavior. See also section II

[6] Further examples for FRBR attributes at the *Expression* level are *recording technique* and *extent*. However, these play no role in the TIB use case.

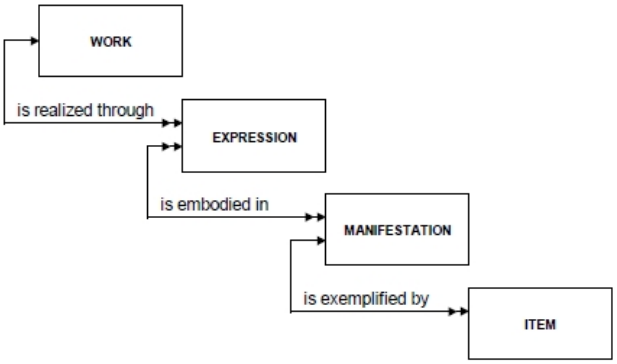

Figure 1: FRBR Entities Group 1 - WEMI Model [11]

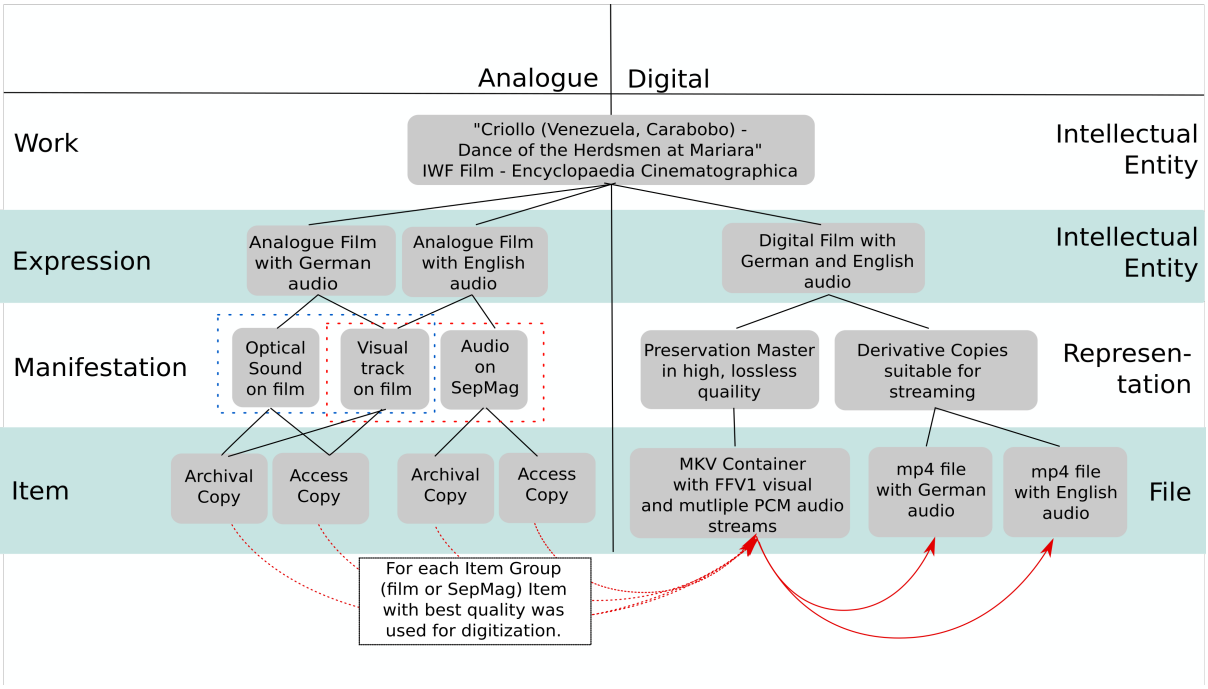

Figure 2: TIB Mapping of analogue to WEMI and digital to PREMIS

and *Expression* levels are mainly of descriptive nature, thus making them logical candidates for good-practice descriptive metadata on the digital level, attributes on the *Manifestation* and *Item* levels primarily pertain to the extent, form and quality of the carrier material. In other words, metadata about *Manifestations* and *Items* are the analogue equivalent to what digital preservation considers technical metadata. As exemplified in Figure 2, the *Manifestation* is the realization of the *Expression* in a concrete carrier form. This is not to be confused with the carrier itself, which is captured and described at the *Item* level. As such, attributes at the manifestation level are *form of carrier* (e.g., "16 mm") and *playing speed* (e.g., "24 fps"). Multiple copies exist for each carrier and while they all share the same attributes of the *Manifestation*, the condition of each copy is described at the *Item* level (e.g., "splice count 0").[7] The relevance of these attributes to the digital object will be further discussed in Section III B.

2. *WEMI to PREMIS or two separate worlds?*

As described in the previous section, the hierarchy of the analogue AV objects handled during the DELFT project is complex. However, what appears to be overwhelming, especially when coming from the digital world, can be made sense of using the FRBR WEMI model. But how do we transpose this to the digital archive? And do we have to keep this structure? TIB uses the PREMIS data model within its digital archive. Each AIP contains one *Intellectual Entity (IE)*. Each *IE* can contain 1-n *Representations* and each *Representation* can contain 1-n *Files* [8]. Each *IE* can be embedded in structural *IEs* or collections to form a nested IE model. Figure 2 shows how the different digitization results were mapped to the data model. The top level IE is synonymous to the WEMI *Work*. From there on, however, analogue and digital differ due to different functionalities and structure of the digital objects. Where analogue needs separate *Expressions* and *Manifestations* to model the language versions, the digital model needs only one: the digital film in both languages as the core Intellectual Entity within the digital archive. On the *Representation* level, the digital archive also differs significantly from the analogue - while Archival and Access Copies exist on the *Item* level, they are both of the same quality based on technical capabilities - they do differ based on conservation status, making one more suitable for digitization than the others. In the digital archive *Preservation Master* and *Derivative Copies* have different quality criteria per design. While the *Preservation Master* is of the highest quality, its large file size makes it often an unsuitable candidate for access. *Derivative Copies* are smaller files at the price of compression and codec change. This is a concept that - at least in the DELFT use case - did not exist in the analogue model. The assumption that the PREMIS *Representation* does not align with the WEMI *Manifestation* is also acknowledged in the PREMIS Data Dictionary [13]:

> *In FRBR a manifestation entity is 'all the physical objects that bear the same characteristics in respect to both intellectual content and physical form' In the PREMIS model a Representation is*

[7] In practice, the conservator chose the *Items* with the highest quality as the digitization source. Structured condition information was only captured for that *Item*.

[8] For further information on TIB's digital archive data model, see (in English language): `https://wiki.tib.eu/confluence/display/lza/Data+Management`

a single instance of an Intellectual Entity held in a preservation repository; note the difference in multiplicity '(all' versus 'a single instance').

If not only digital and analogue data models differ, but furthermore even the structure and features of individual items, as in the case with multiple audio tracks, is there still something that digital needs to know about analogue? Or is digital an entirely different object that does need to be aware of its provenance in the analogue world?

B. What does digital need to know about analogue?

As described in Section II TIB's conservator and the digital preservation team worked hand-in-hand in deciding which conservation information shall be documented in a structured way to be archived as Preservation Description Information alongside the digital object. This section briefly describes the key information pieces and the rationale behind them when thinking about the connection between analogue and digital. A number of key technical information pieces that the Conservator considered vital information turned out to also be key information pieces about a digital AV object. Criteria such as colour space, colour range, duration, framerate and audio or silent are part of the list of Significant Properties TIB has defined for Video and Audio [9] and share the same values for analogue and digital. This information is therefore not include in the listing below.

The following information was defined as relevant knowledge between analogue and digital:

1. Identifier of analogue *Item(s)* chosen as a digitization source [10]
   Rationale: As shown in Figure 2 there are several copies for each *Item*. While they are designated as Archival Copies or Access Copies[11], the viewing results showed that in several cases the Access Copies were of better quality than their Archival counterparts. Capturing identifier information about the copy used allows to match analogue source against a digital target. Since TIB is preserving the analogue as well as the digital object, this is important context information. In addition, all information found on the film container, such as the year of the copy process or the processing laboratory, are captured as vital technical provenance information.
2. Material of the analogue carrier
   Rationale: Each carrier material (for film typically cellulose acetate, cellulose nitrate or polyester) has specific requirements for storage but also for scanning processes. Cellulose acetate, which is the carrier material for all DELFT films, can curl and warp as a result of deterioration. Curling and warping can have significant impacts on the digitization result of the scanner.
3. Visual autopsy results indicating potential problems in the scanning process resulting in blur
   Rationale: Since cellulose acetate can warp during playback, it needs to be ensured that the scanner used for digitization can handle these anomalies. The conservator conducted a visual autopsy of the material and also visually checked the playback result for every item in the collection. During the visual autopsy the conservator noted the degree of deformation (none, small, medium, strong). This information can help contextualize perceivable blurs in the digital object. Additionally, warping is an indicator for the degree of deterioration of the analogue material. With use of viewing equipment the conservator also captured information about splice count and perforation damage, factors which can influence the scanning result in similar ways to deformation.
4. Automatic analysis results
   Rationale: Depending on the utilized equipment, information such as shrinkage values can be automatically generated during scanning. In the case of shrinkage, some scanners can capture information such as min, max and average shrinkage values. Like the results of the visual autopsy conducted by the conservator, these automatically generated values are helpful information to contextualize the degree of film deformation and potential blurs in the digital result.
5. ph-Value
   Rationale: The conservator also ran ph-tests on a subset of the analogue objects. ph-test results and the date they were conducted help understand the rate at which the analogue material in the collection degrades. It is also a quality indicator for the analogue material as the ph-value can be an indicator of the vinegar syndrome which ultimately destroys the carrier material beyond repair.

The documentation of mechanical damage such as scratches was also discussed and, in general, deemed helpful but could not be realized as part of DELFT due to resource constraints. In particular, options of structured descriptions such as "across 3 frames" and "from bottom left to upper right" will take some thorough exploration first.

C. Use cases for Analogue Conservation Metadata based Preservation Description Information

For the key information identified, three main use cases for analogue source material conservation metadata in a digital archive can be identified:

[9] See (in English anguage): `https://wiki.tib.eu/confluence/pages/viewpage.action?pageId=93608961`

[10] Please refer to Section III A for a definition of *Item* within the scope of this paper

[11] There are also cases of several archival copies such as positive and negative and several Access Copies.

1. **Contextualisation of anomalies during playback of digital material**
   As described above, issues such as deforming and shrinkage of the analogue can have a direct impact on the digitization result. Information about the conservation state can help a consumer of the digital better understand such anomalies during playback. If there is blur in large parts of the film, additional information about, e.g., shrinkage can contextualize that this is due to damage of the carrier and not a part of the original film as intended.

2. **Informing an overarching analogue and digital preservation strategy**
   The exact identification of the analogue digitization source (on the *Item*-level) as well as markers such as ph-value, that can be used to check for current state of deterioration, allow for a targeted overarching analogue and digital preservation strategy. If the analogue copy becomes no longer available, e.g., due to substantial deterioration, the digital archive can easily be notified that the respective digital object is now the only remaining copy. Depending on a digital archive's policy, this can result in a higher preservation level being assigned.

3. **Decision criteria for digitization and digital preservation in collaborative environments**
   Legislative and funding bodies such as the EU explicitly ask for cooperation in restoration and conservation of AV material [14]. As the quality information of the analogue source material also informs the quality of the digital target, this information can be used in cooperative efforts. If an institution would like to know whether a digitized copy of equal or better quality than their own material already exists, the answer can be given based on the conservation metadata.

The three use cases illustrate in which context conservation metadata plays a crucial role not only in digitization but also the digital preservation process. We therefore decided to accompany the digital object with conservation metadata captured on the analogue *Item*-level.

IV. TIB Conservation Metadata for Analogue Film

Desktop research and conversations with different national and international AV-collecting institutions showed that structured criteria for analogue material exist, but target large AV-Archives, which have, as described in section II, different resources and demands than TIB[12]. Furthermore, no metadata schema for structured conservation film metadata could be found. It was therefore decided to write the Conservation Metadata Schema for Analogue Film with the following design goals:

1. Cover all conservation criteria for analogue film which were identified as relevant and feasible for the Conservator to collect
2. Provide a framework in which these criteria can be described in a structured machine- and human-readable form
3. Apply controlled values / vocabularies where possible
4. Structure the information in a way that make them understandable in a digital object's context
5. Keep the framework light-weight
6. Keep the framework extendable
7. Provide information in German and English language to allow for community re-use and feedback

XML was chosen as the structure for the Conservation Metadata as it best fulfills the requirements listed above and is the preferred metadata serialization form in TIB's archival workflows. The resulting data dictionary and XSD schema files have been made publicly available in a GitHub Repository [13]. The data model and schema shall be briefly introduced here.

A. Data Model Overview

As one of the key requirements is to structure the information in a way that makes it understandable in a digital object's context, it seems sensible to use the hierarchical IE structure shown in Figure 2 as the basis for the data model. Within the Conservation Metadata Data Model, the highest hierarchical level is that of the digital Archive IE, e.g., in Fig 2 "Digital Film with German and English audio". Attributes assigned to that level are the Media Asset Management System Identifiers for the analogue *Expressions* and the Signature for the *Work*. All conservation metadata about the data carriers, i.e. the reels and audio, are allocated in the representation entity. The rationale behind this is that all information captured here, i.e., Conservation Information as described in Section III B, pertains to the Preservation Master Representation. As all Derivative Copies are generated from the Preservation Master, they automatically inherit this information.
A brief graphical overview of the structure can be seen in Figure 3. The Representation level of the Preservation Master is also the section where the Preservation Description Information needs to be allocated to contextualize the archival object. OAIS describes Preservation Description Information as *"information which is necessary for adequate preservation of the Content Information"* [4]. Content Information, in return, is *"A set of information that is the original target of preservation or that includes part or all of that information. It is an Information Object composed of its Content Data Object and its Representation*

[12] A good example of this is the exhaustive FIAF Moving Image Cataloguing Manual [15]

[13] See: `https://github.com/TIB-Digital-Preservation/FilmConservationMetadata`

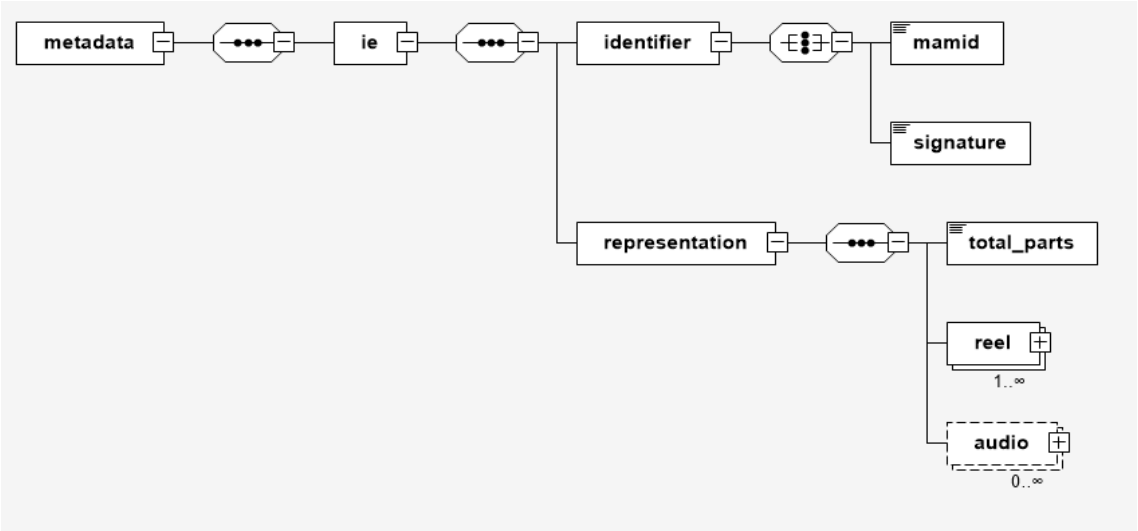

Figure 3: Main entity view of TIB Conservation Metadata

Information." [4]. As such Preservation Description Information clearly relates to more than just the Data Object in form of a file - it informs the preservation target in its current as well as all future generation versions within the archive. Due to this, conservation characteristics at the *Item* levels of a film reel (e.g., Archival Copy) are lifted from the analogue / WEMI lowest hierarchical level to a much more central information role within the archival digital object.

In-line with the key criteria chosen by the conservator and the digital preservation team and described in section III B, the following elements are captured for reel and audio:

- Reel - part number
- Reel - copy identifier
- Reel - carrier material
- Reel - deformation
- Reel - shrinkage (as a container with sub-elements Date, min-value, max-value and average)
- Reel - ph-test (as a container with sub-elements Date and value measured)
- Reel - perforation damage
- Reel - splice Count
- Audio - audio stream number
- Audio - signal base

The overall structure as well as all elements are described in a Data Dictionary which has been made available via a Github repository[14]. Table 1 shows the structure used to describe each Conservation Metadata Schema element within the Data Dictionary.

[14]Film Coservation Metadata Data Dictionary and xsd are available at `https://github.com/TIB-Digital-Preservation/FilmConservationMetadata`

Table 1: Template for description of Conservation Metadata elements

Name	Content Description
rationale	Why is there a need for the element?
diagram	A diagram of the element and sub-elements
properties	Cardinality (if defined), content (complex if element consists of child elements, simple if no child elements), necessity (mandatory, required,optional)
annotation	Annotations in German and English
children	Child elements
attributes	Attributes, their type and necessity (mandatory, required,optional)
source	Section of the xsd-schema
example in XML	Example

B. XML Schema

As described above it was decided to serialize the Conservation Metadata schema in XML. While the data dictionary includes XML examples, it can of course also be serialized in different forms. As TIB uses XML as the main format for metadata within the digital archive, an XSD Schema file was written and made available via github alongside the data dictionary.

V. Conclusion and Outlook

DELFT was the first large AV-digitization project for TIB. Conservator and digital preservation team worked hand-in-hand every step of the way, ensuring that good practice of both domains were fulfilled. A special focus was placed on the question of provenance between analogue and digital. In TIB's overarching conserva-

tion and digital preservation process, the connection between analogue and digital needs to be maintained and preserved. Information about the exact copy used for the digitization process as well as information about that copy's conservation state has been identified as relevant Preservation Description Information which informs about provenance and potential anomalies within the digital object. We have presented three use cases for conservation metadata as Preservation Description Information (see Section III C:

1. Contextualisation of anomalies during playback of digital material
2. Informing an overarching analogue and digital preservation strategy
3. Decision criteria for digitization and digital preservation in collaborative environments

The Conservation Metadata Schema for Analogue Film presented in this paper is naturally only the first iteration of what TIB deems as important analogue-based Preservation Description Information for the digital representation. The materials digitized in DELFT were fairly homogeneous in regards to analogue format - therefore the elements included in the first (1.0) as well as current (2.0) version of the data dictionary have a strong focus on conservation information pertaining to 16 mm cellulose acetate film. These elements will be expanded once we embark on the digitization of different formats. Furthermore, some of the information captured is dependent upon the capabilities of the scanning equipment used. While in the DELFT project the vendor used a Spinner S by MWA which can automatically measure shrinkage values. Since then, other digitization processes have taken place using scanning equipment which does not have this capability, resulting in these criteria being changed from "mandatory" to "optional" in the schema. While the Conservation Metadata currently focuses heavily on film, we see room for improvement regarding audio. Lastly, we would like to explore how the information contained in the conservation metadata can be used in automatic workflows for the digital object. One idea is to also include the carrier format, e.g. 16 mm, in the Conservation Metadata schema. This information would give context information about the digitized object which can be used to pipeline the object to the correct validation workflow such as a MediaConch policy[15], where institutional digitization parameters for 16 mm are checked against.

We are confident that this work will continue into the future and that there is much more of analogue that can be used in the preservation of digital.

[15]For TIB's MediaConch policy used in DELFT, see: `https://mediaarea.net/MediaConchOnline/publicPolicies`

Acknowledgment

The DELFT project was funded by BMBF, the German Federal Ministry of Education and Research.

References

[1] J. Knight, *Digitisatiton does not equal digital preservation!* Online: Future Proof – Protecting our digital future. A State Archives and Records initiative for the NSW Government, 2011. [Online]. Available: {`https://futureproof.records.nsw.gov.au/digitisation-does-not-equal-digital-preservation/`}.

[2] P. Zhou, *Digitization is not digital preservation*, dpc World Digital Preservation Day 2017 Blog, 2017. [Online]. Available: `https://www.dpconline.org/blog/wdpd/digitization-is-not-digital-preservation`.

[3] Library of Congress, "Sustainability of Digital Formats; Planning for Library of Congress Collections: Still Images » Tags for TIFF, DNG, and Related Specifications," [Online]. Available: `https://www.loc.gov/preservation/digital/formats/content/tiff_tags.shtml`.

[4] CCSDS, *Reference Model for an Open Archival Information System (OAIS) - Magenta Book*, CCSDS, 2012. [Online]. Available: `http://public.ccsds.org/publications/archive/650x0m2.pdf`.

[5] M. Taureg, "The development of standards for scientific films in german ethnography ethnography," *Studies in Visual Communication Studies in Visual Commu*, vol. 9, no. 1, pp. 19–29, 1983.

[6] J. McIlwaine, et al., "GUIDELINES FOR DIGITIZATION PROJECTS for collections and holdings in the public domain, particularly those held by libraries and archives," IFLA Preservation and Conservation Section, Tech. Rep., 2002.

[7] K. Arthur, et al., "Recognizing digitization as a preservation reformatting method," ARL Association of Research Libraries, Tech. Rep., 2004.

[8] E. Commission, "Film heritage in the eu. report on the implementation of the european parliament and council recommendation on film heritage 2012-2013," European Commission, Tech. Rep., 2014. [Online]. Available: `https://digital-strategy.ec.europa.eu/en/library/european-commissions-report-film-heritage`.

[9] K. K. Matusiak and T. K. Johnston, "Digitization for preservation and access: Restoring the usefulness of the nitrate negative collections at the american geographical society library," *The American Archivist*, vol. 77, no. 1, 241–269, 2014. [Online]. Available: `https://digitalcommons.du.edu/lis_facpub/1/`.

[10] Memoriav, *Digital Archiving of Film and Video: Principles and Guidance*, Version 1.2, ser. Memoriav Recommendations. 2019, 2019.

[11] IFLA Section on Cataloguing, *Functional Requirements for Bibliographic Records - Final Report*, ser. UBCIM Publications – New Series. K . G. Saur, 1998, vol. 19.

[12] CEN, *Film identification - Enhancing interoperability of metadata - Element sets and structures*, CEN, Ed. 2010.

[13] PREMIS Editorial Committee, *PREMIS Data Dictionary for Preservation Metadata*, Version 3.0. 2015.

[14] European Parliament and Council, "RECOMMENDATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 November 2005 on film heritage and the competitiveness of related industrial activities," 2005, [Online]. Available: `http://data.europa.eu/eli/reco/2005/865/oj`.

[15] T. R. Natasha Fairbairn Maria Assunta Pimpinelli, "The fiaf moving image cataloguing manual," 2016, [Online]. Available: `https://www.fiafnet.org/images/tinyUpload/E-Resources/Commission-And-PIP-Resources/CDC-resources/20160920%20Fiaf%20Manual-WEB.pdf`.