User Requirements and Relational Modelling for a Non-Theatrical Cinema and Video-Art Cataloguing System

Petra Marlazzi[1], Lisa Parolo[1], Cosetta Saba[1], and Nicola Vitacolonna[2]

[1] University of Udine
La Camera Ottica Laboratory
Department of Humanistic Studies and Cultural Heritage
piazza Vittoria 41, 35170 Gorizia (GO), Italy
{petra.marlazzi,lisa.parolo,cosetta.saba}@uniud.it
[2] University of Udine
Department of Mathematics, Computer Science and Physics
via delle Scienze 206, 33100 Udine (UD), Italy
nicola.vitacolonna@uniud.it

Abstract. We describe an ongoing effort to design and implement a computerized cataloguing system for a laboratory dedicated to the restoration and archiving of non-theatrical cinema and video art. The goal is to evolve the current information system taking into account three different aspects: (i) national and international standards and workflows concerning conservation, cataloguing and archiving of film and contemporary art; (ii) specific needs emerging by daily experimentation in film and video restoration practice; (iii) the interoperability with film archives and contemporary art museums. A flexible conceptual Relational model based on Codd's RM/T is proposed as a first step towards the development of a system meeting the unique requirements of non-feature films and videos.

1 Introduction

The preservation, digitization and restoration of non-theatrical cinema and video art means, from a technical perspective, having to do with two moving-image carriers which differ, primarily, for the material they are made of, i.e., photochemical vs. electromagnetic. Both carriers are time-based, but while the former is 'isomorphic', human-readable and can be analysed without the help of a machine, the latter is 'polymorphic', machine-readable, and its content can only be checked using a compatible video-player. Moreover, even the same kind of carrier presents different physical formats (film: 8mm, super8; video: 1/2" open reel, U-Matic, etc…) and varies in ontological essence (e.g., theatrical or non-theatrical cinema, artwork or documentation)—which must be taken into account when experimenting and practising preservation and digitization workflows [2, 17, 22, 10, 21, 11, 19, 3, 9, 8, 13].

Once digitized, the specificity of the carrier (film or video) is partially lost. Digital and long-term preservation workflows are similar in a digital environment, and so is the cataloguing description. The ontological essence, though, must be preserved with minimal information loss, which is typically much more difficult for A/D conversion than for 'digital natives', hence requiring different protocols. The methodologies applied to obtain a digital version from an analogue source, too, must be carefully documented.

While digitization protocols have been widely discussed in the literature [4, 8–11, 13, 14, 17, 19, 21, 22], the cataloguing of non-theatrical cinema and videoart has so far received less attention. At the end of its activities our laboratory *La Camera Ottica* should be able to provide the institutions an interoperable database, keeping track of the salient phases that led to the creation of digital files. It is, then, always more necessary to build a database able to collect and organize all the conservative, administrative and historical/para-textual documentation [14, 4, 15, 6, 16].

There is currently a large number of standardized structures that can be used to start cataloguing cinema and contemporary art. Nevertheless, none of them is focused on the complexity and unique features of 'objects' such as non-theatrical cinema or video/time-based artworks, which can be considered at the boundary between cinema and art, thus requiring a multidisciplinary approach.

Most of the public Italian archives and museum institutions follow the ministerial OAC[3] (*Opere d'Arte Contemporanea*) regulation for cataloguing contemporary (and complex) artworks, which is a strictly hierarchical and very detailed collection of fields and subfields. Beside the lack of a high-level conceptual view of the data, the OAC specification presents many inadequacies in relation to the specific needs of 'moving images'. For what concerns theatrical and non-theatrical cinema, a national regulation is missing and Italian film archives usually refer to the FIAF Cataloguing Manual[4], which shares the same conceptual starting model as the FRBR[5] (*Functional Requirements for Bibliographic Records*) report and adopts many of the definitions in the EN 15744 and EN 15907 European Standards for cinematographic works.[6] As such, the FIAF reference does not contain recommendations focused on analogue videoart or, more generally, on time-based art. For what concerns art, most public European museums and archive institutions currently follow the guidelines of the DCA (*Digital Contemporary Art*) European project,[7] which embraces the same conceptual model as FIAF and FRBR, but also includes suggestions on digitization workflows.

Most current proposals recognise, more or less explicitly, the need to consider the cataloguing process at different levels of detail. Some of them (FRBR, EN 15907, FIAF) distinguish, in slightly different ways, among four main 'entities' (see Table 1), whose meaning can be briefly summarised as follows:

[3] `http://www.iccd.beniculturali.it`.
[4] `https://www.fiafnet.org/`.
[5] `https://www.ifla.org/`.
[6] `https://www.en-standard.eu`.
[7] `http://www.dca-project.eu`. See, in particular, DCA dossiers D3.1, D4.2, and D6.1.

Work: an entity comprising the intellectual or artistic content and the process of its realisation.

Expression/Variant: an entity that may be used to indicate any change to content-related characteristics that do not significantly change the overall content of a Work as a whole.

Manifestation: a physical embodiment of a moving image Work/Variant. Manifestations include all analogue, digital and online media.

Item: a single physical copy of a Manifestation of a Work or Variant.

A restoration laboratory deals with such entities in a bottom-up fashion, starting from physical artifacts typically taking the form of collections of items. In addition to the above entities, for a video and film preservation laboratory, the need to identify and catalogue the *collection in process* must be recognised, which warrants the creation of a further entity/level (Table 1).

FRBR	DCA	FIAF	OAC	La Camera Ottica
-	-	-	-	Collection
Work	Work	Work	Main record	Work
Expression	[Expression]†	[Variant]‡	Level 1 record	Expression††
Manifestation	Manifestation	Manifestation	Level 2 record	Manifestation
Item	Item	Item	Level 3 record	Item

† Mentioned but not used. ‡ Optional. †† An associative entity (see Section 4).

Table 1. Comparison of conceptual entities for modelling moving image works.

2 Specific Non-Theatrical Cinema Requirements

What distinguishes a non-theatrical work from a feature movie is the purpose for which it is produced. The term "theatrical" is connected with the place where the official cinematographic production is shown and connotes a particular productive process—generally speaking, a kind of gauge (from 35mm upward), specific production machines with professional figures, and complex copyright laws. Therefore, the term "non-theatrical" means every audiovisual work where screenings often take place within a private institutional context (e.g., film club screenings, educational screenings). Non-theatrical cinema peaked until the arrival of electronic and, later, digital technology which changed the carrier, though not the purpose, as is revealed by the categories of such films: industrial, training, scientific, amateur, ethno-anthropological, advertising and experimental.

The transition to digital, for some scholars, meant the "death of cinematography", but, paradoxically, in the field of restoration and preservation it gave the possibility to show and watch vintage films again, especially those belonging to the often difficult to access non-theatrical heritage. This has brought to the attention of cultural institutions the necessity to conserve this inherently marginal cinematographic production. Moreover, the need for establishing specific guidelines to do that becomes even more compelling.

While the bibliographic models are suited for the needs of library catalogues, they are inadequate[8] for moving image (non-feature film) archives, as the latter typically have in their collections rare or unique copies of films, or intermediate, and thus incomplete, productions and unreleased material. Even if the cataloguing standards provided by the FIAF take into consideration the "technological advances revolutionising cataloguing, preservation and access practices",[9] they do so from a perspective that values mainstream productions and thus only released materials. Those that are not are largely lacking of guidelines—an exception being the broadcast industry (for which broadcast-specific metadata schemas exist like EBUCore and PBCore). There is a need, then, for cataloguing strategies that do not rely on existing publishing requisites.

That is because an 'archetype' is a vacuous or difficult to define concept in the creative process of cinematographic production, whose existence (to paraphrase Walter Benjamin) is inseparable from its mechanical reproduction and its material properties, which are intrinsic to its essence [11]. Thus, in addition to a reconceptualization of the hierarchical model usually followed by textual criticism (see Table 1), the other particular characteristics that accompany the cataloguing of these materials are:

- the concurrence between *work* and *item*, since the filmic object that identifies the work is unique, in the sense that it does not have any copies to begin with;
- untitled work, resulting from the absence of a publishing intent or other artistic considerations;
- significance not as a work unto itself, but in relation to the collection of origin and to other *items* contained in it. To that end, there is a need for a cataloguing expansion supporting extended relationships and, in particular, para-textual documentation.

The publishing aspect being lacking, if not outright non-existing, it is no longer necessary to adopt a philological approach to recover the text and its classification. As a result, the concept of 'expression/variant' (Section 1) becomes less relevant for non-theatrical work, since the (compensative, substitutive, dismissal, alternative) variants of a work [21] appear only in a second edition—being variations of the original text—but non-theatrical work does not have an official publication and, accordingly, it does not have an 'original' text. Such characteristics point our attention to the 'materiality' of cinema rather than its 'textuality', highlighting the evolution of the techniques and functions of the cinematographic medium.

[8] "While this shared bibliographic model works well for libraries, since many will have exact copies of the same publication, it does not provide all the functions that moving image archives need", *The FIAF moving image cataloguing manual*, 2016, p. 2.

[9] *The FIAF moving image cataloguing manual*, 2016, p. 2.

3 Specific Video Artworks Requirements

As mentioned in Section 1, video artworks are partially or completely recorded using magnetic tapes. Such support has widely varying characteristics, which depend, among the rest, on the storage format, which can be open-reel or cassette (U-Matic and VCR), on the variety of different brands (mostly Philips, Memorex and Sony) and standards used for recording (PAL, NTSC, SECAM), and on the age of the tape. The life of magnetic tapes is variable, but it can be estimated in about thirty years [19, 3, 15, 16], which is a relatively fast decay time.

Today's best and most internationally agreed way for accessing and storing this particular type of artwork is to digitize the contents of the tapes. In general, though, digitization can only be performed after a technical restoration process of the analogue media and a historical contextualization of its contents through the analysis of para-textual documentation. Digitization then enables the production of copies using different coding formats, according to the foreseen purpose of each copy, and also the non-trivial task of long-term preservation storage (Hard Disk, LTO, Cloud) [20]. As it happens with film, this entails the current and future proliferation of analogue and digital copies, versions and variants, whose 'authenticity' and inter-relationships we should always be able to verify and re-establish. Each single aspect of the physical (diagnostic and historical) analysis within which to proceed to the first hypotheses regarding the state of conservation and contents of the tape must be identified and recorded in a catalogue system. While in the latest OAC regulation there is no indication concerning, in particular, the 'material' description of audiovisual tapes, some essential fields that describe time-based analog and digital carriers can be found in DCA dossiers [13, 16].

Moreover, both the DCA guidelines and the OAC regulation prescribe that technical interventions should be documented through both a final and more generic report and so-called *Preservation Metadata*. The OAC document, in particular, provides a paragraph (Conservation and Intervention) with data directed to diagnostics and restoration activities. However, although some parts of an intervention may be generalizable for all the elements of a single collection, other aspects, such as the empirical evaluation of each tape or the cleaning and baking processes, are defined *ad hoc*. A crucial feature of a cataloguing system, then, is to provide support for defining and recording *workflows*, which allows users to check the condition and validity of each phase of a restoration work, even in the long term [13, 16, 20].

Before moving from analog to digital signal, a fundamental aspect is the definition of the digitization quality parameters, which depend on the purpose for which each digital copy is created (conservation, access, preservation), on the physical characteristics of the analogue material, and on the available instrumentation. As a rule, the product of the migration is an *archival master file*; once the content is verified through technical and historical analysis (which may bring to the need of digital restoration), the next intervention consists in the creation of a *production master file* and (possibly several) *derivative files* (also called *access copies*). Production and derivative files may be the result of

post-production activities, such as editing, color correction, digital restoration, addition of titles, A/V encoding, and so on, each of which must be tracked in order to be verifiable and reversible [3, 16]. The compression parameters, in particular, must be carefully chosen according to the foreseen purpose of each copy, pondering the trade-off between quality and storage space (which ranges from a few to several hundreds GBs per copy).

The completion of the video preservation tasks is not the end of the cataloguing process of an artwork. As the DCA guidelines explain, the ultimate digital object to be catalogued will be composed of four fundamental parts: the work of art (abstract), its digital and analogue manifestations, documentation (layout plants, certificates and contracts) and contextual information (actors, locations, event, dates, etc…). Such parts can be linked to each other through administrative metadata; besides, each part requires its own specific descriptive metadata.

The structure proposed by OAC is deficient especially if the need is to describe genealogies between analogue and digital versions, variants and copies, the original and digital carrier and format in which audiovisual components are stored [12, 15, 18, 13]. OAC proposes a stratification distinguishing just between 'main records' (*scheda madre*) and 'secondary records' (*schede figlie*), that is, the components and under components of a complex artwork (see Table 1). Relationships between components are defined in a group of fields named 'Relation'. The DCA research project, on the other hand, proposes three 'closed' hierarchical levels—*Work*, *Manifestation* and *Item*—sharing in this way the conceptual model of the FIAF cataloguing manuals (see Table 1). DCA guidelines, however, recognizing the uniqueness of each artwork, do not retain the *Expression/Variant* level; rather, they emphasise the description of the relations between two artworks that might share the same items and manifestations.

4 Summary of RM/T

We believe that building a comprehensive cataloguing system with future interoperability in mind requires first of all a rigorous conceptual and extendible view of the data. To that aim, we propose a formal specification and classification of the core entities described in the previous sections based on RM/T, which was developed as a more sophisticated version of the basic Relational Model for advanced users in need to model complex domains [5]. In what follows, we assume familiarity with the basic Relational Model (RM) [1].

RM/T, similarly to RM, includes not only the definition of the data structures, but also a rich variety of operators on those structures, and a number of integrity rules, without committing to a particular implementation. Besides, it supports incomplete descriptions (by having only a subset of properties defined for an entity) and offers extended support for several semantic concepts. The explicit support at the conceptual level and a set of new powerful algebraic operators is what sets RM/T apart from RM and makes it suitable for modelling complex data management requirements.

The fundamental conceptual construct in RM/T is the *entity*. Informally speaking, an entity is any object, relationship or concept in the real world that has a relevant role for the information system to be built. Entities can be grouped into *entity types* (or simply *types*) via a form of abstraction from instances to classes (e.g., all persons may belong to a Person type).

Entities (and their types) can be partitioned, according to their role, into: *characteristic entities*, whose purpose is to describe multi-valued properties of entities of other types; *associative entities*, which denote relationships among other entities, and *kernel entities*, which are none of the above. Along an orthogonal dimension, entities (and their types) may be organized into taxonomies via set-inclusion based generalization (*subtyping*), e.g., every Person *is an* Agent. Types at the root of a taxonomy are called *inner types*. Finally, entities may be perceived at different levels of granularity: an entity may be viewed as a whole or as an aggregation of other (simple or aggregate) entities, e.g., a collection of movies. When an entity (type) represents an aggregation of other entities, it is called a *cover (type)*. Note that an entity in general may be a member of more than one cover (even of the same type).

Entities (of any type) are modelled using a countable domain $\mathcal{ED}$, whose elements are called *surrogates*, each one being a representative of one and only one distinct entity of the modelled reality. Two surrogates anywhere in the database are equal if and only if they denote the same real-world entity.[10] For each type T (irrespective of its role), a unary relation on $\mathcal{ED}$, called an *E-relation*, is defined, which asserts the existence of an entity of type T. Besides, for each characteristic type C, an additional binary relation on $\mathcal{ED} \times \mathcal{ED}$ is defined, binding each entity of type C to the entity it describes; and for each associative type establishing a relationship among n other (not necessarily distinct) types, an $n+1$ relation on $\mathcal{ED} \times \cdots \times \mathcal{ED}$ is defined, in which one attribute identifies the associative entity and the remaining n attributes refer to the other interrelated entities.

Entities typically also have simple (immediate) properties that describe them; in particular, in most cases they (should) have simple properties *identifying* them, i.e., subsets of properties assuming a unique value for each distinct entity. Simple properties are modelled using *property relations* (*P-relations*), i.e., n-ary relations $P(S : \mathcal{ED}, A_1 : D_1, \ldots, A_n : D_n)$, where S identifies (to the system) the entity being described, and each attribute A_i is defined on a suitable domain D_i (which constrains the admissible values for property A_i). If $D_i = \mathcal{ED}$ for some i then P is called a *designative* relation. We adopt the view that property relations should be decomposed into *minimal meaningful units* [5]: normalization theory can be used to determine such groupings.

The characteristic types providing a description of a given kernel or associative type T form a strict hierarchy (the *characteristic tree* of T). So, an RM/T model is essentially a collection of characteristic trees, whose nodes are further connected by many-to-many relationships (via associative relations) or one-to-many relationships (through designative relations). We will further assume that,

[10] Of course, such bijection can be enforced by the system only as long as natural keys can be defined for the entities involved—RM/T does not require them, though.

within a single characteristic tree, all non-surrogate attributes and all surrogate attributes not referring to the same type have distinct names.

RM/T also maintains explicit meta-information about the database schema in a collection of *catalog relations* (see Fig. 3), which includes the following: relation **PG** ties each P-relation to its E-relation; relation **CG** relates each characteristic type to the type it describes; relation **AG** stores the fact that a type *sub* is part of the definition of an associative type *sup* via attribute *att*; relation **SG** describes the immediate subtypes *sub* of each generic type *sup*; and relation **KG** specifies which types *sub* may be members of cover types *sup*.

RM/T enforces a number of integrity constraints in addition to those that are part of RM. Some of them should be obvious from the foregoing description (e.g., a tuple t may appear in a P-relation only if the corresponding E-relation asserts the existence of the entity described by t; referential integrity on designative attributes; and so forth). One constraint that will be useful to keep in mind is that *every occurrence of a surrogate anywhere in the database must appear in at least one E-relation.* Another important constraint is that each characteristic entity is existent-dependent on the entity it refers to (which is not required for one-to-many relationships in general). See [5] for the full list of RM/T constraints.

Finally, one of the more interesting aspects of RM/T is its extended Relational Algebra, which allow users to formulate queries that are somewhat independent of the schema of a database. We consider the following operators:

1. $\mathsf{NOTE}(R)$ is the name of the relation R (i.e., a string). The inverse operator is $\mathsf{DENOTE}()$ (i.e., $\mathsf{DENOTE}(\mathsf{NOTE}(R)) = R$).
2. $\mathsf{TAG}(R) \doteq R \times \{\mathsf{NOTE}(R)\}$.
3. $\mathsf{COMPRESS}\,(\cdot, \mathcal{R})$ is the relation obtained by repeated pairwise application of associative and commutative operator $\cdot$ to the relations in the set $\mathcal{R}$;
4. $\mathsf{APPLY}(f, \mathcal{R}) \doteq \{\, f(r) \mid r \in \mathcal{R} \,\}$, where f maps relations into relations;
5. $\mathsf{CLOSE}(R)$ is the transitive closure of (binary) relation R.
6. $\mathsf{PROPERTY}(R)$ groups into a single relation E-relation R and all its immediate properties.[11]

We find it convenient to define additional operators $\mathsf{GRAPH}()$ and $\mathsf{LGRAPH}()$. Given the name of any associative E-relation A, $\mathsf{GRAPH}(A)$ is a binary relation on schema $(s\colon \mathcal{ED}, t\colon \mathcal{ED})$ representing the symmetric closure of the graph of the association denoted by A. Then, $\mathsf{LGRAPH}(A) \doteq \mathsf{TAG}(\mathsf{GRAPH}(A))$. For instance, if a, b and c are related via a ternary association A then $\mathsf{LGRAPH}(A)$ contains the tuples (a, b, A), (b, a, A), (a, c, A), (c, a, A), (b, c, A), and (c, b, A).

5 A Case Study: *Do You Remember This Movie?*

To give an example of the complexity we face when dealing with particular kind of non-theatrical cinema and time-based artworks, we will consider *Do You Remember this Movie?* by Luigi Viola as a case study. In 1979, the video was

[11] This is a derived operator. See [5] for a formal definition.

recorded on an analogue U-Matic cassette. It depicted the artist while watching a home movie he had made with his family a couple of years before. The institution owning the video had many copies of the same artwork in several different (VHS and DVD) cartridges. Some time after the acquisition by our lab, we found out that, in 1982, the artist had produced a 'remake' of the video with the same content and title, but with different production technologies. Besides, further documental research revealed that the first version of *Do You Remember this Movie?* had been featured as part of a multi-media installation entitled *I looked for… (da Alice 1977)*, which was presented during a collective exhibition in 1980 [7] (see Fig. 1). Eventually, we also found the film of the home movie which had been projected in both versions of *Do You Remember this Movie?*.

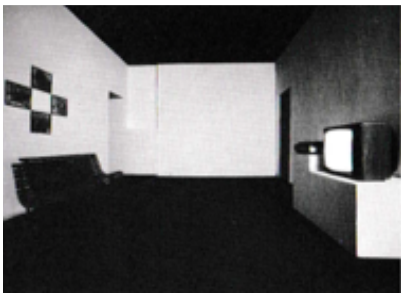

Fig. 1. The only known installation of *I Looked for… (da Alice 1977)*.

Fig. 2. A frame from *Do You Remember This Movie?* (1979).

We face, then, a set of complex relationships among: one 'original' artwork created in 1979; a different version with the same title dated 1982; a home-movie film (with no title) made between 1975 and 1976 by the same artist; and a multimedia installation with a temporal component constituted by four different videos, among which *Do You Remember this Movie?*, and a spatial component, including the room, a green bench and four photographs (see Fig. 1). Note also that the multimedia installation exists only through the para-textual documentation and the testimony of the artist.

"A good catalog is designed to demonstrate to catalog users a number of different kinds of relationships among its records." [23]. Yet, the whole reconstruction of the conservation and exhibition history of the artwork is the most difficult part of cataloguing. Establishing complex relationships also makes retrieving data more difficult. We claim that RM/T can help users designing systems in which such problems are elegantly solved.

The core of our model is based on four inner kernel relations corresponding to the types Collection, Item, Manifestation and Work. The inestricable relationship between abstract works and their concrete manifestations is captured by treating WORK as a cover type having MANIFESTATION instances as members. According to such point of view, a work is essentially just a less granular view of a set of

manifestations. Similarly, MANIFESTATION and COLLECTION are also cover (in fact, partition) types with ITEM instances as members.

The Variant/Expression type is modelled as an associative relation between works (possibly with additional constraints, such as the requirement that it is a strict partial order). Using an associative type circumvents the inherent semantic overlap between 'expressions/variants' and 'works', which would occur if Variant were treated as kernel. It also removes the transitive dependency (works have many variants or manifestations, and variants have many manifestations) implied by the FIAF proposal.

To account for mixed media entities such as *I Looked For…* and *Do You Remember This Movie?*, we must recognise that the former is *not* a manifestation of the latter, nor a variant; instead, the latter is an essential *constituent* element of the former. Mixed media items are *composed of* A/V items and extra-filmic material (e.g., photographs, objects, spaces). Making such composition relationships explicit is essential for cataloguing, because reuse is common in video art and non-theatrical cinema—in fact, we argue that Composition is at least as important as Variant. Compositions may be specified at any level of granularity (items, manifestations, works), hence they must be modelled using three (not independent) associative types: ITEMCOMPOSE, MANIFCOMPOSE, WORKCOMPOSE. For instance, *Do You Remember This Movie?* is a component of *I Looked For…* at the work level; the 1978's version is a component of the (only) manifestation of *I Looked For…*; but, as it happens, we do not know any relationship at the item level.

Finally, for each defined entity type, an arbitrary number of properties and characteristics can be defined.[12] Many RM/T queries make no assumption on the number and structure of P-relations or on the size and depth of characteristic trees. We conclude this section by showing a few such queries.

1. Retrieve *all* the immediate properties of the original version of *Do You Remember This Movie?*:

$$\begin{aligned} W &\leftarrow \sigma_{title='Do\ You\ldots'}(\mathsf{PROPERTY}(\textsc{Work}) \bowtie \mathsf{PROPERTY}(\textsc{Title})) \\ V &\leftarrow \delta_{vwork\#\rightarrow work\#}(\pi_{vwork\#}(\textsc{VariantWork})) \\ \mathit{Answer} &\leftarrow W \bowtie (\pi_{work\#}(\textsc{VariantWork}) \setminus V) \end{aligned}$$

   Here, δ is the *renaming* operator [1].
2. Retrieve the characteristic tree of 1982's *Do You Remember This Movie?*:

$$\begin{aligned} R &\leftarrow \pi_{sub}(\sigma_{sup='\textsc{Work}'}(\mathsf{CLOSE}(\mathbf{CG}))) \cup \{\mathsf{NOTE}(\textsc{Work})\} \\ S &\leftarrow \mathsf{COMPRESS}\,(\mathbb{X}, \mathsf{APPLY}(\mathsf{PROPERTY}(), \mathsf{APPLY}(\mathsf{DENOTE}(), R))) \\ \mathit{Answer} &\leftarrow \sigma_{title='Do\ You\ Remember\ldots'\wedge year=1982}(S) \end{aligned}$$

   The result is a flattened version of the characteristic tree of the given work.
3. Find the persons who are the subject of works they have authored:

$$\begin{aligned} C &\rightarrow \sigma_{role='author'}(\mathsf{PROPERTY}(\textsc{Credits}) \bowtie \mathsf{PROPERTY}(\textsc{Role})) \\ \mathit{Answer} &\rightarrow \delta_{subject\#\rightarrow agent\#}(\mathsf{PROPERTY}(\textsc{HasAsSubject})) \bowtie C \end{aligned}$$

[12] For a (somewhat contrived) example, see App. A.

CATR

relname	*reltype*
Collection	*E*-relation, inner kernel
Item	*E*-relation, inner kernel
AnalogItem	*E*-relation, kernel
DigitalItem	*E*-relation, kernel
Manifestation	*E*-relation, inner kernel
Work	*E*-relation, inner kernel
Variant	*E*-relation, associative
WorkCompose	*E*-relation, associative
ManifCompose	*E*-relation, associative
ItemCompose	*E*-relation, associative
WorkType	*P*-relation
WorkTitle	characteristic relation
…	…

AG

sub	*sup*	*att*
Work	Variant	*work#*
Work	Variant	*vwork#*
Work	WorkCompose	*work#*
Work	WorkCompose	*complex_work#*
…	…	…

PG

sub	*sup*
WorkType	Work
WorkYear	Work
VariantWork	Variant
…	…

CG

sub	*sup*
WorkTitle	Work
Unit	Item
Intervention	Unit
…	…

KG

sub	*sup*
Item	Collection
Item	Manifestation
Manifestation	Work
…	…

SG

sub	*sup*
DigitalItem	Item
AnalogItem	Item
Person	Agent
Organisation	Agent
…	…

AG (cont.)

sub	*sup*	*att*
Subject	HasAsSubject	*subject#*
Work	HasAsSubject	*work#*
Work	Credits	*work#*
Agent	Credits	*agent#*
…	…	…

Fig. 3. (A small part of) the RM/T catalog.

Since an entity may belong to several types, to assert that a person is the subject of a work it is sufficient to insert its surrogate into a kernel type Subject. An associative type HasAsSubject then may relate any subject with any work.

4. Find the works related *in any way* to 1982's *Do You Remember This Movie?*:

$$A \leftarrow \pi_{sup}(\sigma_{sub='\text{Work}'}(\mathbf{AG} \bowtie_{sup=sup' \wedge sub=sub' \wedge att \neq att'} \delta_{*\rightarrow *'}(\mathbf{AG})))$$
$$B \leftarrow \pi_{work\#}(\sigma_{title='Do\ \ You\ldots' \wedge year=1982}(\text{WorkYear} \bowtie \mathsf{PROPERTY}(\text{Title}))$$
$$Answer \leftarrow \delta_{work\#\rightarrow s}(B) \bowtie \mathsf{COMPRESS}\,(\cup, \mathsf{APPLY}(\mathsf{LGRAPH}(), A))$$

This query returns tuples of the form (w_1, w_2, R), where w_1 is the surrogate of the specified video, and w_1 is directly related to w_2 via relation R.

6 Concluding Remarks

None of the current national and international standards and regulations, while critical to ensure future interoperability with other databases and institutions, suit the specific needs for cataloguing complex 'objects' like non-theatrical cinema and video/time-based art. An accurate description of the complex relationships among the several entities involved in the restoration process and well-defined, system-supported, digitization and cataloguing workflows are two of the key elements that we have identified as crucial for a useful cataloguing system ready for interoperability. We have proposed an extendible conceptual model as a foundation for such a system. Our model can be easily implemented in relational DBMSs (although support for RM/T is currently lacking) or mapped into metadata schemas.

References

1. Abiteboul, S., Hull, R., Vianu, V.: Foundations of Databases. Addison-Wesley (1995)
2. Bolter, J.D., Grusin, R., Grusin, R.A.: Remediation: Understanding New Media. MIT Press (2000)
3. Bordina, A.: La conservazione dell'arte video: teorie, strategie e tecniche. Ph.D. thesis, University of Udine (2008)
4. Carlos, M.: A comparison of Scanning Technologies for Archival Motion Picture Film. Master's thesis, Staatliche Akademie der Bildenden Künste Stuttgart (2013)
5. Codd, E.F.: Extending the database relational model to capture more meaning. ACM Transactions on Database Systems 4(4), 397–434 (1979)
6. Ernst, W.: Digital Memory and the Archive. University of Minnesota Press (2013)
7. Fagone, V. (ed.): Camere incantate. Video, cinema, fotografia e arte negli anni '70. Milano, Palazzo Reale (1980)
8. Flueckiger, B.: Material properties of historical film in the digital age. NECSUS—European Journal of Media Studies 1(2), 135–153 (2012)
9. Fossati, G.: From Grain to Pixel: The Archival Life of Film in Transition. Amsterdam University Press, 2nd edn. (2011)
10. Gracy, K.F.: Documenting the process of film preservation. The Moving Image 3(1), 1–41 (2003)
11. Herr, L.: The digital dilemma: Strategic issues in archiving and accessing digital motion picture materials. Academy of Motion Picture Arts and Sciences (2007)
12. Laurenson, P.: Authenticity, change and loss in the conservation of time-based media installations. Tate Papers Autumn 2006 (2006)
13. Martone, P. (ed.): Tra memoria e oblio. Percorsi nella conservazione dell'arte contemporanea. I timoni (2014)
14. Mazzanti, N.: Challenges of the digital era for film heritage institutions. Publications Office of the European Union (2012)
15. Noordegraaf, J., Saba, C., Le Maître, B., Hediger, V.: Preserving and Exhibiting Media Art: Challenges and Perspectives. Amsterdam University Press (2013)
16. Parolo, L.: For a History of Italian Video-Art in the Seventies: the Cavallino Gallery Video Archive (1970–1984). Historical-Critical Examination of Sources and Research of New Digital Cataloguing and Archiving Methods. Ph.D. thesis, University of Udine and University Roma 3 (2017)
17. Read, P., Meyer, M.P.: Restoration of Motion Picture Film. Butterworth-Heinemann, 1st edn. (2000)
18. Saaze, V.v.: Installation Art and the Museum: Presentation and Conservation of Changing Artworks. Amsterdam University Press (2013)
19. Saba, C.G.: Arte in Videotape: Art/tapes/22, collezione ASAC - La Biennale di Venezia. Conservazione, restauro, valorizzazione. Silvana editoriale (2007)
20. Traczyk, T., Ogryczak, W., Pałka, P., Śliwiński, T.: Digital Preservation: Putting it to Work. Springer (2017)
21. Venturini, S.: Il Restauro Cinematografico: Principi, Teorie, Metodi. Campanotto Editore (2006)
22. Wheeler, J.: Videotape Preservation Handbook. In: AMIA (2002)
23. Yee, M.M.: Moving Image Cataloging: How to Create and How to Use a Moving Image Catalog. Libraries Unlimited (2007)

A Example RM/T Instance for a Video Art Catalogue

Mandatory properties such as internal or standard identifiers (i.e., natural keys) are omitted for simplicity. Attributes ending with # are defined on $\mathcal{ED}$. Surrogate keys have been underlined only when there is more than one surrogate attribute.

A.1 Some Kernel Entities and their *P*-Relations

Collection

coll#
k_1

CollectionName

coll#	*coll_name*
k_1	Fondo Cavallino

Item

item#
i_1
i_2
i_3
i_4

ItemCollection (cover member)

item#	*coll#*
i_1	k_1
i_2	k_1
i_3	k_1
i_4	k_1

ItemManifestation (cover member)

item#	*manif#*
i_1	m_1
i_2	m_2
i_3	m_3
i_4	m_4

AnalogItem

analog#
i_1
i_2
i_3

AnalogData

analog#	*base extent*
i_1	triacetate 6474 ft

DigitalItem

digital#
i_4

DigitalData

digital#	*container*
i_4	MPEG

Work

work#
w_1
w_2
w_3
w_4

WorkType

work#	*work_type*
w_1	Home Movie
w_2	Video Art
w_3	Mixed Media
w_4	Video Art

WorkYear

work#	*year*
w_1	1976
w_2	1979
w_3	1980
w_4	1982

Format

format#
f_1
f_2
f_3

FormatInfo

format#	*carrier*	*format*
f_1	video	U-Matic
f_2	film	16 mm
f_3	video	H.264

Manifestation

manif#
m_1
m_2
m_3
m_4
m_5

ManifWork (cover member)

manif#	*work#*
m_1	w_1
m_2	w_2
m_3	w_3
m_4	w_4
m_5	w_2

ManifFormat (designative)

manif#	*format#*
m_1	f_2
m_2	f_1
m_4	f_1
m_5	f_3

Subject

subject#
s_1
s_2
w_1
a_1

SubjectDescription

subject#	*description*
s_1	Family
s_2	Carnival

Agent

agent#
a_1
a_2
a_3
a_4
a_5

AgentAddress

agent#	*address*
a_3	via della video arte 78
a_5	via delle Scienze 205

Person

person#
a_1
a_2
a_3

PersonName

person#	*first*	*last*
a_1	Luigi	Viola
a_2	Paolo	Cardazzo
a_3	Lisa	Parolo

Organisation

org#
a_4
a_5

OrganisationName

org#	*org_name*
a_5	Galleria del Cavallino
a_6	La Camera Ottica

A.2 Some Characteristic Entities and their *P*-Relations

Title
title#
t_1
t_2
t_3
t_4

TitleWork	
title#	*work#*
t_1	w_2
t_2	w_2
t_3	w_3
t_4	w_4

TitleDetails		
title#	*title*	*title_type*
t_1	Do You Remember This Movie?	preferred
t_2	Do You Remember This Film?	draft
t_3	I Looked for… (da Alice 1977)	preferred
t_4	Do You Remember This Movie?	preferred

Unit
unit#
u_1
u_2

UnitItem	
unit#	*item#*
u_1	i_1
u_2	i_1

UnitDetails	
unit#	*unit_details*
u_1	reel 1
u_2	reel 2

Intervention
int#
j_1
j_2

InterventionUnit	
int#	*unit#*
j_1	u_1
j_2	u_1

InterventionDescription	
int#	*intervention_description*
j_1	manual cleaning
j_2	scanning

Ownership
own#
o_1
o_2

OwnershipColl	
own#	*coll#*
o_1	k_1
o_2	k_1

OwnershipOwner	
own#	*agent#*
o_1	a_5
o_2	a_6

OwnershipAcquisition	
own#	*acq_date*
o_1	1976/1/1
o_2	2003/1/1

Paratext
para#
p_1
p_2

ParatextManifestation	
para#	*manif#*
p_1	m_1
p_2	m_1

ParatextDescription	
para#	*paratext_description*
p_1	see Fig. 1
p_2	catalogue

A.3 Some Associative Entities, with Properties and Characteristics

Variant
variant#
v_1

VariantWork		
variant#	*work#*	*vwork#*
v_1	w_2	w_4

VariantNotes	
variant#	*variant_notes*
v_1	Some parts remade

ManifCompose
mc#
e_1

ManifComposeManif		
mc#	*manif#*	*complex_manif#*
e_1	m_1	m_2

WorkCompose
wc#
e_1

WorkComposeWork		
wc#	*work#*	*complex_work#*
e_1	w_1	w_2

Credits
credits#
c_1
c_2

CreditsAgentWork		
credits#	*agent#*	*work#*
c_1	a_1	w_2
c_2	a_2	w_2

HasAsSubject
has#
h_1
h_2
h_3

WorkHasAsSubject		
has#	*subject#*	*work#*
h_1	w_1	w_2
h_2	w_1	w_4
h_3	a_1	w_2

Role
role#
r_1
r_2

RoleCredits	
role#	*credits#*
r_1	c_1
r_2	c_1

RoleDescription	
credits#	*role*
r_1	author
r_2	producer