Universiteit van Amsterdam
Faculty of Humanities, Department of Media Studies
Professional MA Preservation & Presentation of the Moving Image

The Archive as Database. Theoretical and practical analysis of an indexical relation.
A case study: the Zanni Collection

Supervisor:
Dr Giovanna Fossati

Author:
Francesca Morselli
Student number: 6101089

Amsterdam, 31 March 2011
1. INTRODUCTION ........................................................................................................... 4

   1.1. RESEARCH QUESTION AND HYPOTHESIS ......................................................... 5

   1.2. CONCEPT OF INDEXICALITY .............................................................................. 7

   1.3. CASE STUDY: DECONSTRUCTION AND RECONSTRUCTION ......................... 8

   1.4. SECOND LEVEL METADATA ............................................................................ 9

2. ARCHIVE, NARRATIVE, DATABASE, INTERACTION ............................................. 10

   2.1. CORRESPONDENCE BETWEEN THE ARCHIVE AND THE DATABASE ............. 10

   2.3. RELATION BETWEEN ITEMS ......................................................................... 11

   2.4. A FOOTSTEP FOR THE ACCESSIBILITY ............................................................ 12

   2.5. TOWARD METADATA AND INTERACTION ..................................................... 13

   2.6. CONCLUSION .................................................................................................... 15

3. THE ZANNI COLLECTION ......................................................................................... 17

   3.1. INTRODUCTION ................................................................................................. 17

   3.2. ARTEMIO ZANNI ............................................................................................... 18

   3.3. DESCRIPTION OF THE COLLECTION ............................................................. 19

   3.4. THE SKELETON OF THE COLLECTION ............................................................. 21

       3.4.1. AN OVERVIEW ............................................................................................ 21

       3.4.2. AIM OF THE TREE-TRUNK GRAPHICAL REPRESENTATION ....................... 22

       3.4.3. METHODOLOGY ......................................................................................... 23

       3.4.4. CATALOGUING ........................................................................................... 24

       3.4.5. DECONSTRUCTION ..................................................................................... 27

       3.4.6. RECONSTRUCTION ..................................................................................... 28

       3.4.7. DIFFERENTIAL ANALYSIS ....................................................................... 33

            3.4.7.1. HYPOTHESIS ...................................................................................... 33

            3.4.7.2. METHODOLOGY ............................................................................... 33
3.4.7.3. CASE ONE: ZANNI IN THE SMALL VILLAGE OF FELINA.................................................. 35
3.4.7.4. OBSERVATIONS TO CASE ONE .................................................................................. 37
3.4.7.5. CASE TWO: ZANNI IN AFRICA AND INDIA ............................................................... 38
3.4.7.6. OBSERVATIONS TO CASE TWO ................................................................................. 40
3.4.8. CONCLUSION ..................................................................................................................... 40

4. CONCLUSION .......................................................................................................................... 41

5. BIBLIOGRAFY ......................................................................................................................... 45
1. INTRODUCTION

The digitization process is today subverting the classical notion of the archive. Diana Taylor in her lecture *Archiving Performance: The Digital as Anti-Archive* discusses how the digital revolution is profoundly questioning the established practices of the archive, particularly referring to the notion of spatiality. (Taylor 2009) According to Whitney Trettien’s account of Taylor’s lecture:

“Whereas the archive of yesteryear was a physical, authorized place with a known (or at least knowable) set of institutional practices (i.e., "we will deposit this cultural object because it has some historical value"), the digital archive explodes its content across various platforms, placing it in variable contexts that subvert the known "thingness". (Trettien 2009)

In a similar context the film archivist perceives these new practices of data handling and organization as complex and extraneous tasks, preferring to directly handle the artefacts in the belief that these deserve more attention and care. In my opinion the archivist – and a hard-to-change idea of the archive – may not realize that the data (the digital copy or representation of the artefacts) deserve the same attention of the physical records.

This statement is justified by the notion of *originality*, referred to the analogue or digital nature of film. As Giovanna Fossati states,

“The authority of the object is one of the main motors in the discourse on the archival life of film and it lies at the foundation of the archival mandate of preserving material artefacts.” (Fossati 2009)

The nature of the archival object – the film – and its transformation into different matters is the basis of a discourse that considers the database as a repository of digital items.

Moreover, commenting on Benjamin’s theory of reproducibility, Fossati recognizes that besides being a serial product, when it enters the archive, the film originates a “newly
recognized authenticity” because it becomes heritage and museum artefact. (Fossati 2009)

Fossati concludes her analysis by stating that

“Whether it is a photographic, analogue duplication or a digitization (…) there is still continuity between analogue and digital with the accent lying, not on the photographic or digital process, but on the human mediation that executes it.” (Fossati 2009)

In the same way the hypothesis of this dissertation considers the archive and the repository of digital records as effects of the same phenomenon: in fact, the latter can be perceived as an encoded and intangible transformation of a physical archive.

1.1. RESEARCH QUESTION AND HYPOTHESIS

In this thesis I will treat the database as a cultural form and a product of the information society. I will try to demonstrate that the database has developed a double identity: on the one hand it is a tool for organizing information, on the other it has earned a certain cultural autonomy. If we consider the database as a part of the archival process, than it can be perceived as a metaphor for the individual and collective memory and the way we access the database reflects the way we access the archive.

In this essay I put forward the hypothesis of an indexical relation between the archive and the database: conceived as a network of entities (the digital records) linked together, the database can be designed as a projection and reference to the real archive.

This dissertation introduces and elaborates on four main concepts: archive, database, narrative and interaction. The first two – Archive and Database – are the subjects of the analysis, while Narrative and Interaction draw the relations between the subjects. In this model the four concepts equally contribute to the construction of a new structure where the physical archive and the intangible database are equally important. This is possible as they
both present a narrative structure, defined by the relations between their records, that facilitates the accessibility and the interaction by the end users. 1

In particular, the database presents a narrative structure that is not immediately visible as it is contained in the relations between the digital records. By definition, a database contains the information about data representations and the relations that link them together. 2 The incorporation of the relationships between items into the database is the reason of the meaningfulness of the data system: thanks to the network of relationships the database ceases to be a simple list, becoming a meaningful narrative path throughout data. The same mechanism happens in the human brain where the connections between neurons – the synapses – are as fundamental as the cells/items to be connected: a defect in their interconnections is the cause of many diseases. The synapses of the database function as creators of meaning by stating affiliation, dominance and other sets of relationship.

In the next chapters these theoretical premises will be applied to a case study in order to discover the relations between the archival records.

---

1 In this essay the term "end user" identifies any possible participant that accesses the archive and database. This includes both passive consumers and active researchers such as scholars and archivists.

2 The database speaks a universal language that is fundamental for writing and programming activities: while usually specific terms are reserved to experts, in this case they will be important in the last part of this essay when delineating a practical case among the presentation of an audiovisual collection.

An Entity is the elementary item of a database; it is the object whose data must be captured. A Subtype is a subdivision of an entity that has characteristics such as attribute or relationship in common with the more general entity. Subtypes are presented in the modeling diagrams as entities within other entities. An Attribute holds a particular piece of information about an entity; in other words, attributes are the characteristics of entities.

Finally, the Relationship defines how two or more entities are associated with each other; several types of relationships can be defined in a database: one to one, one to many and many to many. (Len Silverstone 2009)
1.2. CONCEPT OF INDEXICALITY

The following questions will act as guidelines for this research: What are the elements that describe the relation between archive and database? How can this intangible relation be analyzed? Can it be applied to real case studies?

*Index* comes from the Latin word *indicare*, which means to show, to let know. Indexicality is a term that belongs to language theory where the indexical signs such as demonstratives and pronouns can evoke a referential projection or realize a referential act (Pierluigi Basso Fossari 2006).

Semiologist Charles Pierce discussed the matter of photographic indexicality: he recognized photographs to be “exactly like the objects they represent”. This exactness between the two objects is possible thanks to what he defines as second-class signs, the *indices*: “The index is physically connected with its object; they make an organic pair”. (Peirce 1894)

The concept of indexicality was born within the domain of the language and was later embraced by Semiotics to express the relation between two contiguous objects connected through their representation.

Within the context of this research I am interested in investigating whether the indexical relation is applicable to physical and digital objects and therefore to archive and database.

Braxton Soderman, argued that:

“Indexicality can still be applied to digital images by shifting the focus from the connection between the object of reality and its photographic representation to the process that makes a photograph reproduce reality. (…) Similarly to photographic images that are forced to correspond to reality, digital images are produced under such circumstances that they are physically compelled to correspond point by point to a symbolic algorithm.” (Soderman 2007)

This statement provides a theoretical background to the hypothesis of the existence of an indexical relation between archive and database. As Soderman foresees, the correspondence
between physical and digital items can be concretized through codification: it is through a code that the two objects create a common language that permits their translation from one domain - the analogue - to another – the digital, and vice versa.

Consequently to Soderman’s theory, I will elaborate on the existence of an indexical process from the physical archive to an intangible database. In the following chapters I will describe the possible outcomes that such a process produces in an audio-visual environment.

### 1.3. CASE STUDY: DECONSTRUCTION AND RECONSTRUCTION

In this essay I will analyse the Zanni collection, acquired by Osservatorio Reggio Emilia in collaboration with Home Movies (Home Movies. Archivio Nazionale del Film di Famiglia) in 2008. At the time I was working at the same laboratory and had the possibility to handle the films and work with the collection’s inventory and catalogue; together with my colleagues we soon recognized that the Zanni collection was worthy to be preserved and presented to a wider audience.

Within this dissertation the Zanni collection will be re-analysed from the point of view of its catalogographic description: this means that instead of considering the films from their contextual perspective (sociological, historical or thematic), I will start the analysis from their constituting items, the metadata.

This analysis can be organized in two main phases: *deconstruction* and *reconstruction*

1. First, I will extract metadata form the film catalogue: this process can be considered as a *deconstruction* of the collection, as it condenses the films to an ensemble of metadata

---

3 www.osservatorioreggioemilia.org; www.homemovies.it
2. Following this, metadata will be recombined in order to outline a narrative order: I consider this phase a *reconstruction* of the collection as it reconnects the separated elements in a meaningful and consistent order. Moreover a tree-trunk design – instead of an anomic structure – will permit to organize the metadata.

1.4. SECOND LEVEL METADATA

Finally, the analysis of the case study will suggest the existence of a *meta-metadatum* – or *super-metadatum* – that would permit the connection of metadata according to thematic clustering. The identification of meta-metadatum within the database would permit the enhancement of thematic connections among the collections of different archives.
2. ARCHIVE, NARRATIVE, DATABASE, INTERACTION

2.1. CORRESPONDENCE BETWEEN THE ARCHIVE AND THE DATABASE

“The central artefact of this system is not the camera but the filing cabinet.” (Sekula 1989)

To which degree can we think of the archive and the database as contiguous concepts? This question raises many interesting issues as the two are elements of different nature: on one hand the archive is the custodian of physical records while the database is a set of items that represent the content and structure of the real archive. In other words, the database is a reflection that references the archive to which it is permanently connected.

The ability to store endless quantities of documents in an infinitely smaller space was presented to institutions as the digital miracle that would finally obsolete piles of paper; on the other hand, the intangibility of the digital file revived in scholars and the public opinion the Derridian fear of forgetfulness.

“What is kept in the archive of course can be erased, can be lost, and the very gesture which consists in keeping safe is always, and from the beginning, threatened by the possibility of destruction.” (Derrida 2002)

Often, electronic data is seen as a fragile medium that is not fully representative of all the complexity of the archival records: this is probably due to the intangibility of pure data and this can easily induce fear in manipulating such intangible items. As an example digital formats change standards rapidly and what is readable today may not be accessible in a few years. Derrida developed his theory on the forgetfulness and destruction impulses intrinsic to
the archive: I think that the apprehension created by dealing with data can be linked to this unconscious impulse.

2.3. RELATION BETWEEN ITEMS

How important is it to outline relations between the entities in the database?

I consider the relation between the records in the archive as the starting point for the development of more complex and narrative relations, which will be the next step in this dissertation.

Among the authors that most treated this topic are Alan Sekula and Ernst Val Alphen. In The Body and The Archive (Sekula 1989) Allan Sekula recognizes this relationship and affiliation between the records within the archive. He describes the archive as being both a paradigmatic entity and a concrete institution at the same time; referring to a photographic archive, he defines it as a vast substitution set providing a relation of equivalent value between images. In his speculation, this technological medium introduces a kind of democracy within the archive: as an example, the digital camera that scans and records a digital representation of each image, indefinitely preserves every one in their mathematical form. I consider Sekula’s perspective on the relation between real and the indexical objects within as a step further in my research: electronic data within the archival realm has to be considered fundamental, not only as a tool to enrich and establish complex relations between records, but also because it redefines and reshapes the archive.

In Free Association and Anomic Archiving (Van Alphen 2006), Ernst Val Alphen largely debates on the inter-relation between objects in the archive. His position is interesting as he designs a signification model based on a psychological and subjective connection between the elements: if archiving is conceived as a process that creates a certain order – which means to create boundaries and coherence – it is then presumed that the unexpected remains out of this process. But, he states that it is impossible for every item to remain isolated and far from
the others, because its uniqueness is created from the relation with the other items. (Van Alphen 2006)

Van Alphen defines this mutual signification as a montage: an irrational interconnection of elements that cause a paradoxical effect:

“At a certain point, the individual components are deemed to be only another expression of those object which surrounds it. Uniqueness, specificity and individuality are destroyed within the process of archiving.” (Van Alphen 2006)

In this last paragraph van Alphen states that in this inter-relation between items, a collective significance emerges in a form that he describes as anomic (from greek, a-nomos: without law or rule), where this term refers to the potentially endless interrelation between objects and virtual entities. Van Alphen perceives the archive as an anomic ensemble of records, where a certain order and relation between the items exists, but only at an initial phase: this order would be then disrupted by the infinite and always increasing number of relations.

2.4. A FOOTSTEP FOR THE ACCESSIBILITY

The main objection to the use of databases is that they are cold and displaced since they consist of intangible and electronic files: data is perceived as a set of values in a chart. In this essay I stress the essential role of the database as an intermediate step between the archive and user accessibility. In particular I wish to demonstrate its narrative nature and the way its structure can influence the access and retrieval of information.

In the process of identifying the relations between the archive and the database, it is necessary to define what is the role of narrative in this context, and why it could facilitate the accessibility of an audiovisual collection.

Mieke Bal in Narratology: Introduction to the Theory of Narrative (Bal 1997) writes:
“Events, actors, time and location, together constitute the material of a fabula. In order to differentiate the components of this layer from other aspects, I shall refer at them as elements. These elements are organized in a certain way into a story. Their arrangement in relation to one another is such that they can produce the effect desired, be this convincing, moving, disgusting or aesthetic.” (Bal 1997)

As Bal describes the elements of a fabula as elements, she considers them in their primary function: as entities that interact with other entities in order to produce a certain effect, decided by the author. In her analysis Bal removes every attribute and ornament from the characters.

I associate Bal’s urge to define the components of the fabula as elements with the attempt of this essay to circumscribe units within the archive and the database, in order to organize them in a meaningful construction. From the database it is possible to extrapolate names, events, locations and dates, recomposing them according to the needs of curators, historians, journalists and scholars: in the same way Bal describes the process of extraction and re-ordering of narrative elements.

I share Mieke Bal’s idea that the fabula is the result of an interpretation of the text by the reader. Would it be possible to extend this interpretative process to an audiovisual text? How can the narrative nature and the structure of the information in the database influence the access and interpretation of end-users?

2.5. TOWARD METADATA AND INTERACTION

Metadata is data about data, examples may be keywords assigned to an image or the type of codec used to compress a file:

“Metadata is what allows the computer to see and retrieve data, move it from place to place, connect data with other data.” (Brouwer, Charlton and Mulder 2001).
Annemieke de Jong highlights the role of metadata in the access of a collection by the end user itself. In particular, she is interested in how metadata can be used in the field of audiovisual media.

While elaborating on the purpose and categorization of metadata, she states that metadata are always additional to the document content. They form the characteristics of that content, describe its processing and are therefore closely related, without being part of it. (De Jong 2003)

Secondly, she individuates four ways through which the end user can employ metadata in order to access audiovisual heritage:

1. To find material that corresponds to the user searched criteria
2. To identify an information object
3. To select an information object
4. To obtain access to an information object (De Jong 2003)

I consider metadata the basis for the creation of narrative paths through audiovisual content in the database; *Metadating* is what permits the user to access, interpret and re-shape the database.

Lev Manovich has dedicated an extensive research to the issues and challenges brought by the introduction of metadata. In particular, he focused on the creative potential of metadating and the consequent possibility to transform the archive from gatekeeper to knowledge generator.

“In terms of its creative and generative potential, metadating the image paradigm means following four related directions: inventing new systems of image description and categorization; inventing new interfaces to image collections; inventing new kinds of images which go beyond such familiar types as a digital photograph or a digital video; approaching the new super-human scale of visual data available, not as a problem but as an opportunity.” (Lev Manovich 2003)
With this statement Manovich puts forward the hypothesis of considering the Metadating process as a new paradigm to interface with reality: in his view this could open to a more collaborative culture where metadata are not created by a unique generator but are the results of cross connections and mutual feedback between content providers and end users. (Lev Manovich 2003)

I share the approach of Manovich to think of metadata from a creative and collaborative point of view. This new perspective allows me to extend the definition of metadata as an interaction and narrative tool in order to enhance the presentation and the access of digital collections.

2.6. CONCLUSION

In this chapter I portrayed the relation between the archive and the database. My aim is not to describe the technical characteristics of the database and the way metadata function: these topics are more properly discussed among technicians and informatics.

On the other hand, database and metadata are relevant to a wider range of professionals, among which the archivist: why should he be involved in a debate over the database? What is the connection between his role and the way information is organized in the database?

First, I outlined a fundamental connection between the archive and the database, which I identified as an indexical relation, where the real objects in the vaults and the entities in the database communicate and interact.

Secondly, I focused on metadata as an instrument to open to the accessibility of the database and therefore of the archive. Through this perspective, the objects establish a narrative relation, so that - while accessing the collection – the end user is able to activate an interactive process and navigate the whole collection through multiple paths.

The archivist is deeply involved in this process: the border between the archive and the database is vanishing and potentially, the database becomes the archive. The entities in the
database represent the records in the archive to such an extent that, accessing the database equates to accessing the archive itself.
3. THE ZANNI COLLECTION

3.1. INTRODUCTION

In the second part of this dissertation I will demonstrate how the theoretical issues raised in the previous chapters can be applied to a case study. In particular I will refer to three key concepts developed previously in this essay: the indexical relation between archive and database; metadata as the key element for the accessibility and interaction with the archive; finally, I will introduce the new concept of meta-metadata as a tool to connect thematically more archives.

In this chapter I will analyse the Zanni collection by decomposing and recomposing it, in order to reveal its metadata structure. The collection will be first decomposed identifying film formats, anthroponyms, toponyms, dates and events; then it will be recomposed following the order of the elements – the metadata – as it emerged in the films.

The metadata structure will be represented graphically with the shape of a tree trunk: the idea is that of creating a network where the film elements are connected in order to show their relationships. This graphical representation is useful for this analysis as the nodes represent the metadata and the cobweb their relations.

The second part of the case study will be dedicated to reconstructing the film collection, previously fragmented in its minimal component. I will operate a differential analysis, designing two narrative paths from the tree-trunk in order to demonstrate different possible interpretations of the collection.
3.2. ARTEMIO ZANNI

The Zanni Collection was rediscovered in 2008 when the nephew of Artemio Zanni approached the Italian Home Movies association and the university of Modena and Reggio Emilia. For decades, he had been keeping hundreds of films and pictures belonging to his uncle, the priest Artemio Zanni.

These documents portray a man that was a point of aggregation for the life of a small mountain community but also an eccentric religious guide that made his motto out of simplicity and a genuine religious belief. Zanni arrived in the small village of Felina, on the mountains of the north of Italy, just after the end of the Second World War in 1945. Before this, he was a war priest and during this period he promised to many soldiers that he would have taken care of their children, in case they shouldn’t survive. So he did.

As soon as he arrived in Felina he inaugurated an orphanage for those children who lost their parents during the war. The orphanage that was named Casa Nostra – Our Home – turned into an active part of the community: all the children were involved in activities connected with the locals, such as the summer harvest, the restoration of the church or the celebration of religious festivities. Among all these activities, the entire village was enthusiastic to participate in Zanni’s activities as a filmmaker.

Although living in a small isolated community, Zanni appears to have been a curious and inquisitive soul when adopting a 16mm camera soon after the Second World War. His first approach with film was both of surprise and excitement: this is evident in the way he portrayed the faces of the orphans and of his own family, as he develops his filming style. These portraits are exceptionally powerful as they depict the intense gaze of these people; nonetheless these images represent a significant historical trace.

The Zanni filmic production evolves through the time to different film format and moves to different places: as soon as Zanni enlarges his borders by visiting Africa and India, he brings
with him an 8 and super8 camera. He films different landscapes and traditions and sends the film back to the orphanage in Italy. The films were conceived as a communication tool to make the orphans aware of his missions and also to permit them to experience different countries and people through the films.

3.3. DESCRIPTION OF THE COLLECTION

The Zanni Collection is valuable for its richness, heterogeneity and historical relevance. It represents an ideal case study for this research as it portrays a wide variety of characters, places, situations and events: this will permit the end-user to access and actively navigate the collection through a rich set of metadata.

The collection is composed of 97 films. Among them are 69 Super8, 22 Regular 8mm, 5 16mm and 1 magnetic audio. (Home Movies. Archivio Nazionale del Film di Famiglia; Relabtv. University of Modena and Reggio Emilia. 2009/2010)

Artemio Zanni started filmmaking as soon as he arrived to the village of Felina in the late Forties, just after the end of the Second World War. According to local historian Giuseppe Giovannelli (who has also been a guest of the orphanage), Zanni’s first shots date back to the late Forties, when he borrowed from another priest a 16mm camera to film his family and a trip on the mountains with the first guests of Casa Nostra. His last shots are dated around the early Eighties and depict his missions in Africa and India, where he contributed to the construction of a leper colony.

Zanni filmed for more than thirty years, giving back to us a rich image of himself - an eccentric priest, dynamic and passionate - but also of the people and landscapes he was surrounded by. Moreover, his filming technique developed through time, to the point that it became part of the social activities of the orphanage: the camera was passed from hand-to-hand among his numerous collaborators and friends.
The Zanni collection develops the following major themes: the local community (Felina, Casa Nostra, the locals) and its activities; Zanni’s family and finally the missions in India and Africa.
3.4. THE SKELETON OF THE COLLECTION

3.4.1. AN OVERVIEW

Figure one. The original metadata structure of the Zanni collection
3.4.2. AIM OF THE TREE-TRUNK GRAPHICAL REPRESENTATION

This graph – a tree trunk – represents the skeleton, the metadata structure of the Zanni Collection. Metadata are grouped according to their thematic proximity and connection. As an example, Felina, Travels, Missions and Family represent the macro-themes that then progressively specify until a singular event, place or persons are identified.

The graph contains the objects (or entities, according to the database language) and their relations. On one hand every entity can be perceived in its autonomy, as the signifier of an external-real object that is its reference in the film. On the other hand the graph can be perceived as an ensemble of dynamic relations established between the entities: every node of the graph can be seen as connected with other nodes creating different relations. The way metadata are structured has an important impact on the presentation and access of an audiovisual collection.

Annemieke De Jong (De Jong 2003) offers the technical background to this hypothesis:

“The more highly structured metadata are, the more that structure can be exploited for searching, manipulating and interrelating with other information objects. Effective organizing of metadata requires an underlying structure: a metadata model or schema. A metadata model is a graphical representation of the arranged order of the metadata-elements that are compiled in the metadata dictionary. The metadata model defines the hierarchical and associative relationships between different components of an information object and their attached media, according to the internal structure of the objects themselves and the way end-users wish to approach the information.” (De Jong 2003)

The graphical visualization of the Zanni collection has several advantages.

1. First, it is possible to access to the information hidden in the collection at a first glance: year, format, places, persons, events etc….
2. Secondly the graph presents the metadata of the collection in a consistent way, showing relations and affiliations between items.

3. Finally, I hypothesise that this kind of visualization helps curators to research and manage the collections: as the films are defined as data, it becomes easier to search for the information within.

This graph also represents the connections with the theoretical background previously introduced, because it can be identified as the key element between the film and the database: it is the graphical representation of their *indexical relation*.

Looking to an audiovisual collection from this different perspective and not only from a filmic, textual, historical or social point of view, can open new interesting possibilities for the organization and retrieval of the collection. Instead of approaching the archive as an unorganized ensemble of films, I would rather suggest to outline the relations starting from a metadata structure, its skeleton; it will then be possible to extend and translate the relations that have emerged, in narrative and curatorial selections.

### 3.4.3. METHODOLOGY

Among archives and collecting institutions, metadata retrieval is today a highly debated issue. On one hand web portals and the increased accessibility of digitized collections requires a standard for metadata, trough which collections can be retrieved trough the same set of data. On the other hand the increasing amount of objects that require digitization also require automatic systems for the extraction of metadata.

According to Annemieke de Jong, metadata can be generated in three ways:

- By automatic indexing of elements of the content (e. g. by extracting image elements in video, keywords in the spoken word and automated detection of programme genres).
• By implicit generation of elements during the creation and encoding of the shots (hour
and date as registered by the camera, time codes and frame numbers when material is
being digitised).
• By manual or semi-automatic addition of external information, such as catalogue
description, keywords etc.

The future of metadata leads to automatically harvested metadata (Metadata Harvesting) 4,
and systems capable of extracting huge quantities of data at the same time. This would help
to relieve from the painstaking and time consuming effort of a manual process as it
recognizes and collects automatically metadata from audiovisual content.

In the rest of this chapter I will rather focus on the way I extract metadata from one film of
the Zanni collection, practicing what was introduced in the previous pages: deconstructing in
order to see micro and macro relations between the elements, and reconstructing to have an
overall perspective of the archive..

3.4.4. CATALOGUING

My exercise starts with the film cataloguing.

“Castellazzo, Bassa Reggiana. Film di Famiglia di Don Zanni. Ritratti di parenti. A partire da
00:02:08 Felina, ritratto di Savio e Pietra Crotti, con la madre all’uscita di “Casa Nostra”. A partire da
00:02:59 Castellazzo, scuola Elementare. Ritratti di scolari. A partire da 00:04:21 Cusna, Gita al
Cusna 1951.” (Home Movies. Archivio Nazionale del Film di Famiglia; Relabtv. University of
Modena and Reggio Emilia. 2009/2010)

4 For more insight on this topic see: Carl Lagoze, Herbert Van de Sompel, Michael Nelson and Simeon Warner,
«The Open Archives Initiative Protocol for Metadata Harvesting,» Open Archives, 2008,
http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm.
Below are a few frames extrapolated from the same film: Savio and Pietra Crotti with their mother (Figure two), the Primary school of Castellazzo and a few students (Figures three and four), the first guests of Casa Nostra to the Cusna Mount (Figure five).

Figure two. Savio and Petra Crotti (and their mother) in front of their house in Castellazzo

Figure three. Portrait of a Primary School Student in Castellazzo
Figure four. A few students in front of the Primary School with Leone Crotti

Figure five. A young boy refreshing with the water of a small torrent on the Cusna Mount
3.4.5. DECONSTRUCTION

The Metadata resulted from the cataloguing of the Zanni Collection were collected and with them a preliminary classification was created: *anthroponyms* (name of person), *toponyms* (the name of a place) and *generic*.

Below is the list of metadata extracted from the films of the Zanni Collection.

<table>
<thead>
<tr>
<th>Antroponimi</th>
<th>Toponimi</th>
<th>Generiche</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alba Castellari</td>
<td>Africa</td>
<td>Animali</td>
</tr>
<tr>
<td>Amos Zaneli</td>
<td>Alpi</td>
<td>Anziane</td>
</tr>
<tr>
<td>Anna Maria</td>
<td>Appennini</td>
<td>Anziani</td>
</tr>
<tr>
<td>Battistino</td>
<td>Austria</td>
<td>Asilo</td>
</tr>
<tr>
<td>Beniamino</td>
<td>Basso</td>
<td>Automobili</td>
</tr>
<tr>
<td>Berta</td>
<td>Barceliona</td>
<td>Ballo</td>
</tr>
<tr>
<td>Bianca Morelli</td>
<td>Basilica di San Paolo Fuori le Mura</td>
<td>Bambine</td>
</tr>
<tr>
<td>Bigio</td>
<td>Basilica di San Pietro</td>
<td>Bambini</td>
</tr>
<tr>
<td>Braglia</td>
<td>Bolzano</td>
<td>Bande</td>
</tr>
<tr>
<td>Bresciani</td>
<td>Braiola</td>
<td>Barca</td>
</tr>
<tr>
<td>Bussi Enrico</td>
<td>Campania</td>
<td>Bassa</td>
</tr>
<tr>
<td>Cati</td>
<td>Carpineti</td>
<td>Befana Parmigiana</td>
</tr>
<tr>
<td>Cardinale Baron (?)</td>
<td>Casa Nostra</td>
<td>Calcio</td>
</tr>
<tr>
<td>Cardinale Sergio Pignedoli</td>
<td>Castelazzaro</td>
<td>Calcincolo</td>
</tr>
<tr>
<td>Casolari</td>
<td>Castelnuovo ne' Monti</td>
<td>Campagna</td>
</tr>
<tr>
<td>Ciro Cassinadri</td>
<td>Castelsardo</td>
<td>Capeo</td>
</tr>
<tr>
<td>Chiara (Cotti?)</td>
<td>Catalogna</td>
<td>Carnevale</td>
</tr>
<tr>
<td>Clementina Santi</td>
<td>Cattedrale di Orvieto</td>
<td>Cerimonie</td>
</tr>
<tr>
<td>Davoli</td>
<td>Cattedrale di Santa Maria del Fiore</td>
<td>Chiesa</td>
</tr>
<tr>
<td>Domenico Cocconi</td>
<td>Centrofrica</td>
<td>Chiese</td>
</tr>
<tr>
<td>Don Armando Zanni</td>
<td>Cerreto</td>
<td>Cineprese</td>
</tr>
<tr>
<td>Don Breno</td>
<td>Cesenatico</td>
<td>Cocomeri</td>
</tr>
<tr>
<td>Don Carlo</td>
<td>Cinque Terre</td>
<td>Colonia</td>
</tr>
<tr>
<td>Don Giglioli</td>
<td>Como</td>
<td>Corteo</td>
</tr>
<tr>
<td>Don Gilberto Baroni</td>
<td>Congo</td>
<td>Dipinti Religiosi</td>
</tr>
<tr>
<td>Don Gino</td>
<td>Cortina</td>
<td>Donne</td>
</tr>
<tr>
<td>Don Giorgio</td>
<td>Cusana</td>
<td>Festa</td>
</tr>
<tr>
<td>Don Giulio Righi</td>
<td>Emilia-Romagna</td>
<td>Fiumi</td>
</tr>
<tr>
<td>Don Giuseppe Monelli</td>
<td>Escorial</td>
<td>Giovani</td>
</tr>
<tr>
<td>Don Leone Zanni</td>
<td>Farno</td>
<td>Giro d'Italia</td>
</tr>
<tr>
<td>Don Orlandi</td>
<td>Febbio</td>
<td>Gite</td>
</tr>
<tr>
<td>Don Orlando Giora</td>
<td>Felina</td>
<td>Laghi</td>
</tr>
<tr>
<td>Don Romano</td>
<td>Firenze</td>
<td>Lago</td>
</tr>
<tr>
<td>Don Umbro</td>
<td>Francia</td>
<td>Macchine Agricole</td>
</tr>
<tr>
<td>Don Vallolini</td>
<td>Gange</td>
<td>Macchine Fotografiche</td>
</tr>
<tr>
<td>Edda</td>
<td>Genova</td>
<td>Macellazioni</td>
</tr>
<tr>
<td>Enzo Filippi</td>
<td>Germania</td>
<td>Mare</td>
</tr>
<tr>
<td>Enzo Piccinini</td>
<td>India</td>
<td>Matrimonio</td>
</tr>
<tr>
<td>Euridice</td>
<td>Indonesia</td>
<td>Messa</td>
</tr>
<tr>
<td>Famiglia Bussi</td>
<td>Italia</td>
<td>Militari</td>
</tr>
<tr>
<td>Fanti</td>
<td>Kenya</td>
<td>Missioni</td>
</tr>
<tr>
<td>Filippi</td>
<td>Lago di Garda</td>
<td>Monumenti Indiani</td>
</tr>
<tr>
<td>Francesca</td>
<td>Lago d'Iseo</td>
<td>Nave</td>
</tr>
<tr>
<td>Franceschino</td>
<td>Lago del Ventrass</td>
<td>Neve</td>
</tr>
<tr>
<td>Francesco Del Rio</td>
<td>La Mancha</td>
<td>Oceano</td>
</tr>
<tr>
<td>Fulvio Pignedoli</td>
<td>Lazio</td>
<td>Palloncini</td>
</tr>
<tr>
<td>Gabriele</td>
<td>Lendia</td>
<td>Parenti</td>
</tr>
<tr>
<td>Gabriella Coriani</td>
<td>Levizzano</td>
<td>Pasti</td>
</tr>
<tr>
<td>Gino Persia</td>
<td>Liguria</td>
<td>Processione</td>
</tr>
<tr>
<td>Giuliana Zanni</td>
<td>Lombardia</td>
<td>Pullman</td>
</tr>
<tr>
<td>Giuseppe Battistessa</td>
<td>Lourdes</td>
<td>Religiosi</td>
</tr>
<tr>
<td>Giuseppe Giovanelli</td>
<td>Madagascar</td>
<td>Religiosi</td>
</tr>
<tr>
<td>Lana</td>
<td>Madrid</td>
<td>Ritratti</td>
</tr>
<tr>
<td>Leo</td>
<td>Marola</td>
<td>Scuola</td>
</tr>
<tr>
<td>Leonie</td>
<td>Matahwa</td>
<td>Spagna</td>
</tr>
<tr>
<td>Leonino</td>
<td>Milano</td>
<td>Stampe Religiosi</td>
</tr>
<tr>
<td>Lilly</td>
<td>Monaco di Baviera</td>
<td>Statue Religiosi</td>
</tr>
</tbody>
</table>
Figure six. Original Metadata set of the Zanni Collection

### 3.4.6. RECONSTRUCTION

The list of metadata created from the cataloguing of the films accomplished the first part of the process that I described as a *deconstruction* of the collection.

That list represents the whole film collection, but something is still missing: the relations between the elements. A database cannot exist without the relationships among its entities. Relations between objects are what permit to define this last process as a *re-construction*: this means that the elements within the database can be combined in different ways.

---

5 The metadata set of the Zanni collection has been realized by *Osservatorio Reggio Emilia* in 2010 according to the guidelines developed by *Home Movies. Archivio nazionale del Film di Famiglia*. 
This new possibility opens to an active fruition of the collection and to the idea of a participating user, playing and interacting with the films. Also scholars or researchers can benefit from this perspective, as the film is perceived as a dynamic ensemble, less restricted and immobile than through a linear reading.

For this reason, I think that visualizing the way metadata (entities) combine together (or recombine) helps to glimpse existing and possible relations.

Figure seven. Format and Year in the metadata visualization graph

Figure eight. Missions in the metadata visualization graph
Figure nine. Travels in the metadata visualization graph

Above are presented three sections (for an overview of the graph see p. 21) of the tree trunk chart that display technical (format, year, colour, stock, sound) and thematic (travels, family, missions, landscape, events etc.) data selections. Narrative paths can be outlined between these selections. As an example, Figure eight presents films which common theme is *Mission*. From here three more ramifications depart: *India, Africa* and *Indonesia* (the
countries we have proof of in the films). Choosing India as an interesting topic to research on, it is possible to see in the graph the events that took place there: Funeral, Inauguration of a leper colony, Construction of a leper colony, Visit to ancient monuments, Trip on the Ganges and Transportation of a Madonna statue on the river Ganges.

This example ends here but it is be possible to display more connections, such as the city or Village where the event took place (the connection with the toponyms), the people that were involved (the connection with the antroponyms), the year (the connection with the timeline branch). I could compare this event with others that happened in the same year or focus on the technical aspect of the film, such as the format or the stock, and deducing more information from this technical data. In Figure ten is an example of how multiple elements within the graph can be connected following different thematic tracks.

The same process applies to the rest of the collection and can create numerous connections between the elements: the same entity can be presented in many ways, depending on the context and the choice of the archivist/curator on how to present the collection.
Figure ten. Relations between metadata in the same film
3.4.7. DIFFERENTIAL ANALYSIS

In order to demonstrate how differently organized data can reflect and result in almost infinite curatorial selections, I will outline and compare two narrative paths designed within the same metadata structure. The aim of this exercise is that of showing the quantity of possible interpretations within the same audiovisual collection: the elements in the metadata structure link together composing a network that change shape every time an element is chosen instead of another.

3.4.7.1. HYPOTHESIS

In this exercise I will describe the existence of meta-metadata as it emerges from the differential analysis. The curatorial paths previously described as a possibility for multiple interpretations within the collection, can be redefined as metadata of metadata: this implies the existence of super-categories in the metadata set. The relation between metadata is therefore recursive; in other words, metadata are related to each other and these relations can themselves be categorized through metadata.

3.4.7.2. METHODOLOGY

For this task, I decide to focus on the figure of the priest Artemio Zanni. As it emerges from the description I gave of him in the chapter 3, he was a strong, ambitious, supportive and modest personality. I was impressed by his ability to feel comfortable in totally different environments: the small village on the Italian mountains – Felina – and the African and Indian communities where he stayed for long time, helping with the construction of a leper colony. Therefore this case study will focus on Artemio Zanni in the two communities, with the aim of comparing his personality and his gestures in such different environments.

First, I went through the cataloguing of the whole film collection and identified the metadata that suggested me the theme I was looking for. At the same time, I helped myself by visualizing and checking the actual content of the films.
Secondly I experimented with the metadata structure and combined the entities in classes to create a story. At this point, I already had identified the elementary entities.

With the metadata and the images I shaped two different narrations: the first, referring to the life and gestures of Zanni within his mountain community and the second referring to his travels and life during the missions in Africa and India. My hypothesis is that the two aspects of Zanni, the local and the international one, are related by his common actions: in other words, it is possible to identify the same gestures even in totally different environments.

Finally, I extracted key-frames from the films, in order to visualize the images of the narration I created: every frame represents a key moment and a key concept of my selection as they serve as a guide for further elaborations.
3.4.7.3. CASE ONE: ZANNI IN THE SMALL VILLAGE OF FELINA
Figure eleven. Metadata: Felina, Zanni, harvest

Figure twelve. Metadata: Felina, Zanni, restoration, old church

Figure thirteen. Metadata: Felina, Zanni, Sick People, Casa Nostra
3.4.7.4. OBSERVATIONS TO CASE ONE

1. From the analysis of the metadata structure it is already clear that a rich amount of elements is available: places, people and events. One noticeable element is that, among all the events not only religious festivities are mentioned, but also daily and seasonal activities such as the harvest, the carnival celebrations and a soccer match. Artemio Zanni also invented new festivities in order to strength the bonds between people in the village.

2. The figure that emerges from the analysis of the metadata structure is that of an active and helpful man. This emerges also from the still images extracted from the films: in Figure eleven Zanni helps the local farmers during the harvest; in Figure thirteen he takes part to the annual lunch organized to comfort sick people of the village; in Figure twelve Zanni plays an active part in the restoration of the old church of Felina.
3.4.7.5. CASE TWO: ZANNI IN AFRICA AND INDIA
Figure fifteen. Metadata: Africa, Zanni, African Village, Man

Figure sixteen. Metadata: Africa, Zanni, exploration, torrent

Figure seventeen. Metadata: India, Zanni, meal, children
3.4.7.6. OBSERVATIONS TO CASE TWO

1. First, I focused on the period during which these films were produced: during the Sixties and Seventies, the period during which Zanni started and later intensified his missions in Africa, India and Indonesia. The format he mostly uses is the Super8, which was largely popular during the Sixties.

2. Zanni visited these far countries to help building leper colonies where this was a major disease. In any case, while visiting Africa, India and Indonesia, Zanni was also interested in meeting and living with the local people and to take part in their daily activities. Moments of exploration and relax with friends or the locals are also part of these travels.

3.4.8. CONCLUSION

A look to Artemio Zanni in his small community having fun with a few children in the years after the war and the same Zanni caressing a child, in Congo, in the late Sixties. Even if these two images belong to different historical, geographical and temporal contexts, Artemio Zanni is always present in the images, not only physically but also in his benevolent and charismatic presence. What is visible in the films is that Zanni behaves always in the same spontaneous and active way no matter where he is: he always seems to be completely integrated with the surroundings.

Through the analysis of the metadata it becomes clear that they represent key-points of a narration over Zanni’s life in a small mountain community or in the desert in Africa: every event, every place, every person represent new opportunities of investigation in the development of a curatorial selection of the collection.

By considering thematic selections connecting clusters of metadata it was possible to create curatorial paths trough the collection. In my opinion this was possible thanks to the existence
of meta-metadata that would group the metadata according to their thematic affiliation: in this way new interpretative layers can be created (meta-metadata are the result of an interpretation) stratifying onto each other.

4. CONCLUSION

In this thesis I put forward the hypothesis of the existence of an indexical relation between the archive and the database, where the last can be perceived as a reference and an intangible projection of the concrete archive. Potentially, this relation opens to new possibilities as both the archivist and the final user can work with the data while the physical items remain safely stored in the vaults of the archive.

The hypothesis of an indexical relation between archive and database facilitates the accessibility to audio-visual collections: the possibility to outline relations between the entities of the database means to design effective relations among the items in the archive.

Secondly, I considered metadata as the basic elements through which interaction can be established between the database and the end user: while the metadata set is the minimal unity that represents the archive, the metadata item is the minimal unit within the database that makes its accessibility possible. From this perspective, the database can be considered as an ensemble of metadata and only through their manipulation it can result as interactive.

Metadata can be linked together in order to compose all possible narrative sequences.

While analysing the Zanni collection, I stressed the relation between archive and database by working directly on the metadata structure in order to create narrative paths through the collection.

1. First, through this case study I demonstrate the indexical relation between database and archive: working on the metadata structure is equivalent to working directly on the
films of the collection. The selection of frames that have been proposed to exemplify the metadata selection demonstrate this equivalence.

2. On the other hand, the analysis of the case study permitted to elaborate on the concept of *meta-metadata*: these represent macro categories of metadata identifiable within the narrative paths proposed among the hypothesis of this research.

These points raise some important questions. How does the role of the audio-visual archivist change? Should the archivist shift its focus on database and metadata? In the context previously outlined the role of the archivist becomes part of the database: as the archive and the database are indexical signs - therefore each of them refers and exists because of the other – also the archivist would shift his attention from the physical to the intangible archive.

Moreover, if we consider the *meta-metadata* as part of the archive, also its interpretation becomes part of the archive: the re-use of metadata collections within the archive enriches and creates an abstracted representation of the archive itself.

Sarah Cook underlines the *political aspect* of the networked culture that has changed the role of the curator. She states that rather than playing the role of the exhibition caretaker, curators act as filters seeking opportunities for meaningful exchange between artists and the community. (Cook 2008) While in Cook’s analysis this produces a democratizing role in the exchange between the archive and the public, in this case also the archivist, through the use of database and its metadata, shifts his role to a more curatorial one.

The creation of *meta-metadata* gives also the possibility of creating connections between a network of archives: for example, *Europeana* enables people to explore the digital resources of Europe's museums, libraries, archives and audio-visual collections. This is possible because 1500 cultural institutions standardize their metadata and super-metadata models,

---

6 Europeana is a web portal that gives access to thousands of European cultural resources. Around 1500 European institutions provide content to Europeana: their collections, merged together, permit the end user to explore the European cultural and scientific heritage.
enabling the user to search for a certain theme across every database, through one simple research.

A recent report of the Comité des Sages\textsuperscript{7} over the Digitization of Europe’s Cultural Heritage, stressed the necessity to give accessibility to the European cultural heritage through the Europeana Web platform:

“The Europeana portal should become the central reference point for Europe's online cultural heritage. Member States must ensure that all material digitised with public funding is available on the site, and bring all their public domain masterpieces into Europeana by 2016. Cultural institutions, the European Commission and Member States should actively and widely promote Europeana.” (European Commission 2010)

A few years before (in 2003) in the Report of the Comité des Sages, Diane Whittaker wrote:

“Common portals seem to be the answer to the problems of searching across databases in the museum, archives and library sector” (Yarrow, Clubb and Draper n.d.)

The cited sources confirm how much the attention in terms of accessibility of cultural heritage will focus in the near future on the concepts of sharing and aggregation: Europeana will probably be the main European portal fulfilling these requirements. For these reasons, meta-metadata will become ever more important in order to aggregate metadata and to translate metadata sets from one institution to another.

\textsuperscript{7} The "Comité des sages" comprised Maurice Lévy (Chairman and Chief Executive Officer of advertising and communications company Publicis), Elisabeth Niggemann (Director-General of the German National Library and chair of the Europeana Foundation) and Jacques De Decker (author and Permanent Secretary of Belgium’s Royal Academy of French language and literature).
While archives safeguard their physical records, they can nevertheless share it with all the people that participate to their data-driven processes and through information sharing, promote innovation and collaboration.
5. BIBLIOGRAFY


Thanks:

I want to thank my colleagues and friends of *Home Movies. Archivio Nazionale del Film di Famiglia* and *Osservatorio Reggio Emilia* for the possibility to work on the Zanni collection and for their support during my research.

My thanks go to Giovanna Fossati, thesis supervisor, for her constructive feedback in structuring my research and to Julia Noordegraaf and Marie Baronian for their inspiring lectures and guidance during P&P Master Programme.

Finally I wish to express my deepest gratitude to my family and to Edoardo for their endless support and for sharing my interests and ambitions during my studies.