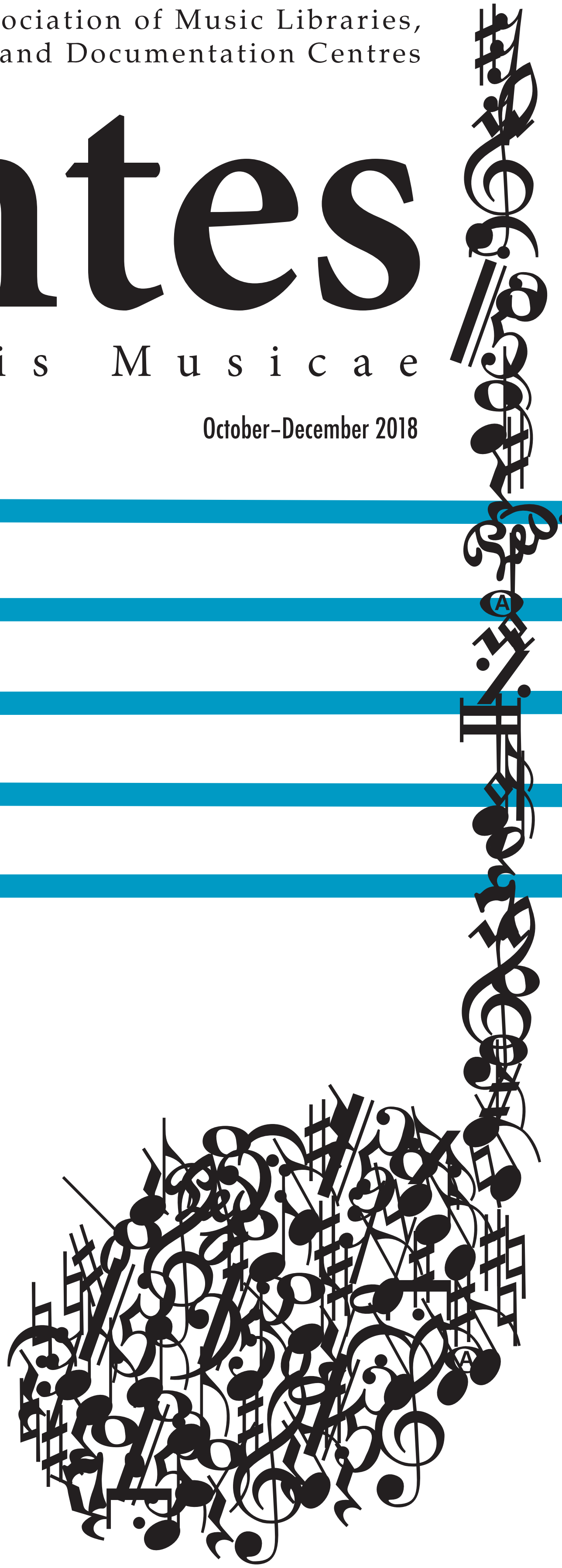
Journal of the International Association of Music Libraries, Archives and Documentation Centres

Fontes Artis Musicae

October–December 2018



ISSN 0015-6191 (print) 65/4
ISSN 2471-156X (online)

FONTES ARTIS MUSICAE

VOLUME 65/4, OCTOBER–DECEMBER 2018

CONTENTS

ANALYTICAL DESCRIPTION OF A DATA SHEET FOR AUDIO PRESERVATION: FROM MEANING TO DATA AND BACK

Federica Bressan

Introduction

This article makes a case for the importance of supporting the choice of audio preservation metadata with open documentation on the conceptual framework applied to the preservation workflow. A complete example is provided in the form of a data sheet, the design of which reflects the choices at conceptual level. A data sheet is a structured report normally enclosed in an audio preservation copy. A preservation copy, or preservation master (IASA-TC 04, 2004), is the audio equivalent of a diplomatic copy or facsimile. The main concepts discussed in the article are, the difference between content transfer and content description, the conceptual model of the audio document, the status of the audio document (cultural heritage or not?), and high- vs. low-profile resources.

This article does not add to the discussion on the nature of metadata, its importance in preservation, interoperability, and multilingualism. It focuses on the importance of the definition of a clear conceptual model (step 1) that is made explicit through open documentation (step 2). Most documentation on metadata is operative: there should be more emphasis on the concepts behind the data. This would reveal the motivation, the intention, and the limits of the framework. Eventually, the conceptual model will result in the choice of a metadata set: it is important to get there in the right order, ideas first, implementation second.

Most existing metadata sets (standard or not) show an overlapping core. Speaking of audio metadata, for example, most will include the file duration and format—obvious choices. In this sense, choosing one over the other may not seem to make a big difference in practical terms. Different metadata sets can be construed as diverging “extensions”. However, examining the non-overlapping parts may reveal conceptual divergences that the adopter should take into consideration to see if the set meets the requirements of the archive.

Federica Bressan (federica.bressan@ugent.be) is a post-doctoral researcher at Ghent University, where she leads a research project on interactive installation art under the Marie Curie funding programme H2020-MSCA-IF-2015. She holds an MD in musicology and a Ph.D. in computer science. Previously, she held a post-doctoral research position at the Department of Information Engineering, University of Padova, Italy, where she coordinated the laboratory for sound preservation and restoration. The core of her research lies in the study of the relationship between historical audio recordings and their digital (digitised) representation—a relationship which is not only very complex, it also develops over time (context, digital philology, hermeneutic implications of encoding cultural objects).

This study was partially supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 703937. The author would like to thank the Centro per it dialetto romagnolo – Fondazione Casa di Oriani in Ravenna, Italy, for the permission to share the sample data sheet. Many thanks to Richard L. Hess for the interesting dialogue refining some technical points.

For (small to medium) institutions adopting existing metadata sets, one of the problems is that many sets are either specific to one aspect of the archive (administrative, copyright, etc.) and need to be combined or integrated to cover the archive requirements; or abstract, in that they provide general guidelines but not the specifics of the implementation—leaving space to misinterpretations. An asset of the example provided in this article is that it is complete: it makes the main principles of the conceptual framework explicit and it is ready for adoption. Readers can take the data sheet enclosed in Appendix A and use it as a template for their own, or they can to send a request to the author to receive the full documentation of the database schema where the data is stored. The database schema is also a reflection of the conceptual framework and would have made an equally-good example for discussion in this article; the data sheet was preferred only for reasons of space limit.

The subtitle of this article, "from meaning to data and back", summarises the essential message of this article: metadata are the concrete expression of a conceptual interpretation of the world; the choice of metadata is never neutral, it reflects our idea of document, archive, and cultural heritage (in increasing order of abstraction); it is important to ensure that the metadata translate the framework accurately and in a way that enables users "at the other end of history" to recompose the message we have "encoded" there. Discussions around metadata should never be about fields, but about the question "what idea of the world does your model express?"

Conceptual Framework

This section explains the building blocks of the conceptual framework within which the data sheet was designed. The "idea of the world" expressed in it, can be summarised in the following statement:

An audio recording is a historical document, part of the [audiovisual, multimedia] cultural heritage. As such, it deserves appropriate protection and valorisation. Whenever restoration is necessary, philological principles should be applied.

Preservation is a complex process that can be broken down into many steps, two of which should be very well distinguished: content transfer and content description. Content transfer is only concerned with the preparation of the audio preservation copy and is "blind" to content.

The data sheet presented in this article contains information about the source document (one of the inputs of content transfer), the preservation copy (one of the outputs), and the process itself.

The Preservation Workflow

Preservation is "the total sum of the steps necessary to ensure the permanent accessibility—forever—of documentary heritage" (Edmonson, 2002), where forever means decades or centuries, or "long enough to be concerned about the obsolescence of technology" (Hedstrom, 2002). It is important to distinguish between the steps involved in the process of preservation. The dichotomy between content and container characterising audiovisual materials requires that we leave behind the paradigm of "preserving the original" (applicable to most traditional cultural goods) in favour of a dynamic form of preservation ("Nothing has ever been preserved—it is only being preserved" [Edmonson, 2016]) predicated upon the "disembodiment" of the content from its [original] container. The preservation of the content results in the creation of the preservation

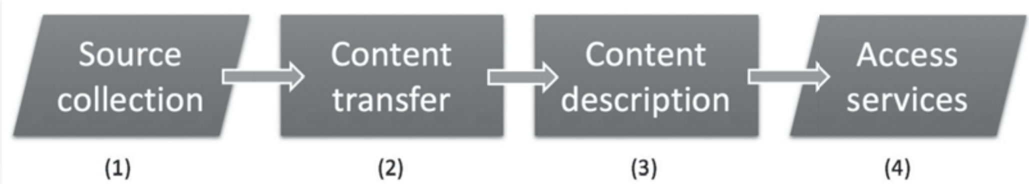


Fig. 1. Preservation workflow using the standard building blocks: (1) and (4) are input/output, (2) and (3) are processes.

copy or preservation master, an "artifact designated to be stored and maintained as the preservation master" (IASA [International Association of Sound and Audiovisual Archives], 1999), an authoritative surrogate of the original document (for the author's implementation of the preservation copy see Bressan and Canazza 2013, sec. 3.5). The preservation copy described here is an evolution of that described in Bressan and Canazza 2013, Figure 5. The preservation copy is an archival document in its own right, and it is not ready for access and use.

Figure 1 shows part of the preservation workflow relevant to this article. An important distinction between block 2 (content transfer) and block 3 (content description) is that the approach in block 2 is determined by the model of the audio document vs. the approach in block 3 is determined by the model of the content in function of its final destination. Content transfer is blind to content in that the procedures applied to the carrier are the same regardless of the content (speech, music, repertoire, etc.). The data sheet described in this article is the same for all types of content (same metadata); content description varies in function of the type of content (metadata may or may not intersect). The type of expertise required to conduct steps 2 and 3 is different (mostly technical for step 2, spanning across any discipline for step 3, e.g., history, linguistics, musicology). The main input of step 2 is the original carrier (or source), and the main output is the preservation copy (including the data sheet presented here). The main input of step 3 is the preservation copy, and the main output is a collection of self-contained audio resources accompanied by their description—ready for content-specific semantic queries.

Content transfer (step 2) can be further broken down into preparation of the carrier, signal transfer, data processing, and archival (Bressan and Canazza 2013, sec. 2.5). Preparation of the carrier involves the optimisation of the physical state of the carrier (cleaning, restoration). Every aspect of this process is documented in a database (Bressan and Canazza 2013, sec. 4.1). The data sheet presented in this article contains a subset of these data.

The "disembodiment" of the content (audio information) from its original container (carrier) takes place in step 2; step 3 operates at a completely disembodied level. The physical limitations of the carrier do not matter anymore, and self-contained abstract units such as a song, an interview, a sport match commentary (i.e., the semantics that end users will use in their queries, or "how humans think") can be catalogued independently. For this reason, between step 2 and 3 a logical reorganisation of the audio materials might be desirable (Bressan and Canazza 2013, sec. 4). It is like breaking free from borders that have become unnecessary. Example: a series of eight music performances was segmented on fourteen tapes due to the limited duration of each tape. After digitisation, the segmentation can be eliminated by recomposing the audio material in eight digital files.

The mindset before digitisation is “we have 14 tapes” vs. after digitisation “we have 8 performances”. Vice-versa, the number of self-contained abstract units can exceed the number of carriers, going from “we have 14 tapes” to “we have 64 folk songs”. In this example, the reorganisation of the audio material will not entail cross-mixing/stitching together, but only trimming/cutting.

It goes without saying, the relationship between the original carriers and the self-contained units is maintained (in the database), so it is possible at any time to know how many self-contained units were recorded on one carrier, or on what/how many carrier(s) a given unit was originally recorded on. This clear separation, without loss of information, is the reason why the data sheet presented here is strictly limited to audio preservation metadata and does not contain any information or reference to the content. With reference to Figure 1, the ontological center of step 2 is the physical audio carrier; step 3 operates at a higher level of abstraction, free from the very notion of physical carrier.

The Importance of the Model

A model is a “simplified abstract view of a complex reality”. The human brain creates models of the world all the time, in order to reduce the complexity of the world to a manageable level. In the field of ecological psychology, it is an established fact that our ability to create a mental picture (model) of the environment implies that we discriminate between different elements in the environment, select the relevant ones, and order them in a hierarchy of value (Gibson, 1979). Analysis, selection and ranking are intrinsic to the functioning of the human brain. It follows that the way in which we engage with the world is not neutral, because it involves an act of interpretation (always based on limited knowledge).

Encoding is an interpretative act (Orlandi, 1997). Analog-to-digital content transfer (trans-coding) is both a practical and an interpretative act. It is necessary to decide what is relevant about the original document, and make sure it gets included in the preservation copy: some elements may be represented directly (pictures of the attachments), others indirectly (verbal description of the smell). The original document and the preservation copy are not the same entity, and loss of information is expected in the process of trans-coding. The “information density” (Brock-Nannestad, 2001) of the original is always higher. The purpose of the preservation copy is to minimise the loss of information, and to avoid the introduction of distortions in the information that is retained.

A good example of how our mental picture or model of an audio document may determine a different implementation of the preservation workflow is the concept of noise. Normally noise does not fit in our idea of desired signal (Brock-Nannestad, 2001). The ‘undesirability’ of noise is implicit in expressions like high quality, speech enhancement and sound restoration. In this hierarchy of values, a signal with less noise is ‘better’ than a signal with more noise. This reflects a will to ‘enjoy’ the sound, so it leans in favour of the aesthetic qualities of the listening experience. This should not be the approach for audio preservation. In audio preservation, there is no qualitative judgement of noise vs. desired signal. Technically, noise is signal. Discarding any portion of the signal constitutes a loss of information, which we want to minimise. Noise may carry useful information about the context of the recording, and it may enable a less invasive restoration—at a time when restoration becomes appropriate. The conceptualisation of the signal is not loaded with the value judgement of undesired vs. desired. The recording is not seen through the lens of its destination (e.g. aesthetic listening) but as historical document, to be transmitted ‘as is'. Every part carries information: the desired signal as 'primary' information, noise as 'secondary or ancillary' information (Brock-Nannestad, 1997).

A workflow where audio enhancement/restoration is allowed is legitimate, but it does not qualify as preservation in that it does not consider the audio document as a historical document. Once the model of the audio document is set, the workflow follows. 'What is the audio recording?' is an ontological question. If the answer is "a historical document", the principles of philology should apply. If not, creative choices can be made. It is important to understand that the question whether a de-noise could be applied before long-term storage is not a matter of aesthetic choices within the territory of preservation, but between preservation and other creative domains.

The Status of Sound Recordings: Cultural Heritage or Not

Audiovisual documents have two components, both of which are important regardless of the medium or format employed: the information content and the carrier on which it resides (Edmonson 2002, p. 8 and 60). The relationship between content and carrier may range from incidental to integral (Edmonson 2016, p. 21). The carrier is a product of industrial manufacture, the content represents something which "at some stage is or has been important to humans" (Brock-Nannestad, 2000). The carrier may or may not have aesthetic qualities, but it has a function: being played on a recorder. It is important that the carrier is in good condition because if it cannot be played it does not fulfil its purpose, preventing access to the content.

In his milestone text on the theory of restoration, Brandi (2005) defines restoration as the methodological moment of recognising a work of art. Restoration is an act of discernment: it does not depend on the qualities of the object, but on our (collective) perception of it. Brandi distinguishes works of art from other products of industrial manufacture, for which restoration is limited to re-establishing the object's functionality. A watch shows the time; when it breaks, we fix it so it shows the time again. The only purpose of the object lies in its function. Not so for audiovisual carriers: they have both a function and a cultural significance. They are multidimensional entities. An act of restoration that fixes the carrier without considering the implications on the document as a whole fails to "recognize" the document as part of the cultural heritage and violates the "documentary unity". (In this context it is useful to extend Brandi's concept of work of art to that of cultural heritage, as to include audiovisual materials, which were not considered in the 1960s.)

Considering restoration an act of discernment that does not depend on the object but on our (collective) perception of it (which is susceptible to change over time!), shows that whether we should consider an audio recording as cultural heritage or not cannot be determined by examining the recording's properties but rather our stance on it. This relates to the previous paragraph: the way in which we see the document is the 'model' we choose for it. Brandi's work shows us the consequences that derive from 'recognising' an object as work of art or simple product of industrial manufacture. It should be noted that audiovisual carriers show a hybrid nature in this sense, because technically carriers are industrial products. For a discussion on the critical encounter of art with industrial objects in the key of preservation (see Mimoso, 2009; Valentini, 2007).

High- vs. Low-Profile Resources

Preservation copies and the materials that derive from them (during content description) are the resources that scholars will use to build new theories and produce new knowledge. It is important that these resources be authentic, accurate and reliable (Duranti, 2012). Ideally, the end user should not feel the need to ask direct access to the original document, to obtain missing information or to verify its accuracy.

Factor et al. (2009) introduces the idea that the authenticity of digital resources cannot be evaluated by means of a boolean flag, but it is rather the result of a process. The process is never limited to the resource itself but extended to the information/document/record system (CASPAR Consortium, 2008). Hence the importance of recording (documenting) "what is done and the choices that are made" at any point in the preservation workflow: failing to do so "[infringes] the integrity of the work in the long term" (Edmonson 2016, p. 69).

A high-profile resource (Gigliozzi, 2003) is one that makes the conceptual framework explicit, and that declares its aims, scope and limitations. The documentation about content transfer must cover every step of the process and include information about the context where the transfer was carried out. In a field where quality is often associated with expensive equipment or big numbers, it is important to keep in mind that audio quality does not guarantee the resource's trustworthiness. Quality is not defined by perceptive criteria. A resource that does not inform the user about the equipment, the alignment levels, the expertise of the technical team, etc. is automatically declassed to low-profile—regardless of the audio resolution.

A low-profile resource is useless at best, and potentially dangerous when it does undetected. If one low-profile resource is found in a collection, the trustworthiness of the entire collection is compromised.

A sample data sheet is included in the appendix to this article. In the next sections, a detailed description of each section of the data sheet is presented. It may be useful to keep the data sheet at hand while reading.

Logical Structure

A preservation copy, and the descriptive sheet accompanying it, includes essential information about:

(1) the source document
(2) the preservation copy
(3) the process [of content transfer]

(1–2) are outputs of block 2 in Figure 1; (3) is what happens inside the block.

The data sheet is meant to be self-explanatory and self-contained. It is structured in four sections, some of which are further divided in subsections (e.g., section 3). Information about point (1–3) is mostly found in the respective sections 2–4, but since it is interconnected, sometimes elements of it will appear in different sections (e.g., track configuration, and information about (1) is found in section 3).

Section 1. List of documents enclosed in the preservation copy. This is a complete list of the documents enclosed in the preservation copy. By document here, we understand "each element present in the preservation copy", practically speaking files. The preservation copy is self-contained and complete, and constitutes a "documentary unity" in itself.

Section 2. Description of the source document. This section includes information about the source document that was partly known in advance (e.g., provenance), partly obtained during visual inspection (e.g., model brand), and during replay (e.g., tape transfer speed). It includes a detailed report on the state of preservation of the audio carrier.

Section 3. Description of the preservation copy. This section includes information about the ownership and the institutional context in which the copy was produced, and a detailed description of each track contained in the preservation copy. This description is very important because the ratio between carrier sides and number of files is not necessarily always 1:1 (see subset. 4.3.2).

Section 4. Technical scheme of the transfer system. This scheme is a visual summary of the content transfer setup, including a description of the equipment, the connections and the data flow.

Fields Description

The data that populate the fields are retrieved from the preservation database. They were ingested at different stages of the content transfer process. The sheet is automatically generated with a function of PSKit (Preservation Software Kit), a software developed by the author (Bressan and Canazza 2013, sec.4.1), which queries the database and formats the output in LATEX (a programming language for high-quality typesetting)[1].

The data sheet begins with a cover page reporting: (1) the signature of the preservation copy, (2) a note about the nature of the data sheet itself, and (3) a table of contents. Here is the table of contents:

1. List of documents enclosed in this preservation copy
2. Description of the source document
    2.1 General information
    2.2 State of preservation of the source document
3. Description of the preservation copy
    3.1 General information
    3.2 Description of the audio files
        3.2.1 Audio file 1, 2, N

(The number of audio files stored in a preservation copy can vary, so in this table of contents subsection 3.2.1 contains a generic list. In a real example, subsections 3.2.2, 3.2.3, etc. would appear accordingly.)

    3.3 Photographic documentation
    3.4 Video documentation
    3.5 Tape configuration
4. Technical scheme of the transfer system
    4.1 List of documents enclosed in this preservation copy

The purpose of this list is to make sure that no element is deleted or misplaced in the preservation copy. All elements are relevant to the preservation copy, in that it is a self-contained unit. Should some of these elements be missing, the integrity of the preservation copy would be compromised.

For each type of document, the file extension and relative path are specified. The path is relative to the preservation copy itself, which in the implementation described in this article is stored on disk as a “folder”.

The list includes the data sheet itself (in pdf format), the audio files (in wave format), the checksums of the audio files (in xml format), the pictures and scanned images of the

1. Information about LATEX can be found at https://www.latex-project.org/about/, accessed 15 October 2018.

source document (in jpeg format), and the video recording of the process of signal extraction (in mov format).

4.2 Description of the source document

This section contains general information about the source document, followed by a detailed report on its state of preservation.

4.2.1 General information

General information about the source document include the following fields:

Archive. This is the document's provenance. The database stores detailed information about the document's legal owner, the contact person, the address and so on. In this field, only the official name of the institution is reported. Archives may or may not be subdivided into fonds and series. The database supports this option when a new archive entry is created. If the archive has subdivisions, more fields are added after "Archive". In the sample data sheet in Appendix A, the document belongs to the fond "Tullia Magrini" of the archive "Centro per Dialetto Romagnolo".

Label for preservation. When a document enters the workflow for content transfer, it is assigned a label—regardless of the signature by which it is known in the source archive (see next point). The label "for preservation" is instrumental to the process of content transfer. A consistent naming convention allows for a better processing and management of the files. These labels can be compared to hospital bracelets: codes, in the place of names, make it easier for the hospital system to deal with the patients. The naming convention applies to all the documents in the workflow described in this article, regardless of their provenance. The convention consists in five characters that identify the archive name, and three digits for a progressive number relative to the same archive. In the sample data sheet in Appendix A, the document's label for preservation is MAGRN037 (for "Magrini").

Signature in the original archive. Another reason to assign a temporary label for preservation is that not all documents were given a signature in their archive, or sometimes they were given with two or more, as a result of successive cataloguing choices. Archives are complex entities, often with a long history, and they bear the signs of changes in political, financial or social positions of the management. It is not realistic to expect orderly and organised sets of audio documents. This should not discourage their care, on the contrary, effective solutions must be devised to compensate for this state of affairs—and digitisation is an excellent time to do it.

Carrier type. Eight types of carriers are considered: compact cassette, compact disc, digital audio tape, microcassette, minidisc, open-reel tape, phonographic disc, and digital non-audio media (this includes flash drives, CD-R, etc.).

Archiving procedure. Cassette (for compact cassette, digital audio tape, microcassette), disc (for compact disc, minidisc, and phonographic disc), open reel (for open-reel tapes), and hard disk drive.

Accompanying material. There are many reasons why the information on the boxes might not help understanding the carrier's content (see two points below). At the same time, transcoding the text from the boxes to the computer would be an interpretative act (see Figure 2), which is minimised during content transfer (see subsection 2.2), and preferably postponed to content description. The default text in this field is "See enclosed images of the source document".

Text and notes. Same as previous (see Figure 2). Default text, "See enclosed images of the source document".

Container type. Material of the container/box. Options include plastic, paper/cardboard, metal or "missing".

Container brand. In case of missing container type, this field does not apply.

Carrier brand. Brand and model are separated (see next field) to increase future possibilities to manipulate the data (queries, etc.).

Carrier model. See previous.

Tape width. Normally 1/4 inch.

Reel flange diameter. Example: 5 in (12 cm).

Type of recording. Analog or digital.

Recording technique. Magnetic, optical or electro-mechanical.

Equalisation curve. This is the equalisation curve applied at recording time. This is normally not known, therefore "unknown" is a common value. Even when the curve is specified on the box, it should be possible to verify this information. This field requires certainty. In case of doubts, it is better to write a note in the "Notes on the signal" below.

Noise Reduction. Same as previous.

Number of sides. Determining the format (specifically track configuration) of an audio tape can be a challenging task (see subsection 4.3.2), and so can be expressing it in a concise and non-ambiguous way. The strategy adopted by the author requires three fields, this one and the next two. This strategy is able to distinguish between standard cases and also critical cases. An exhaustive explanation is not possible here but an example is provided in Figure 3.

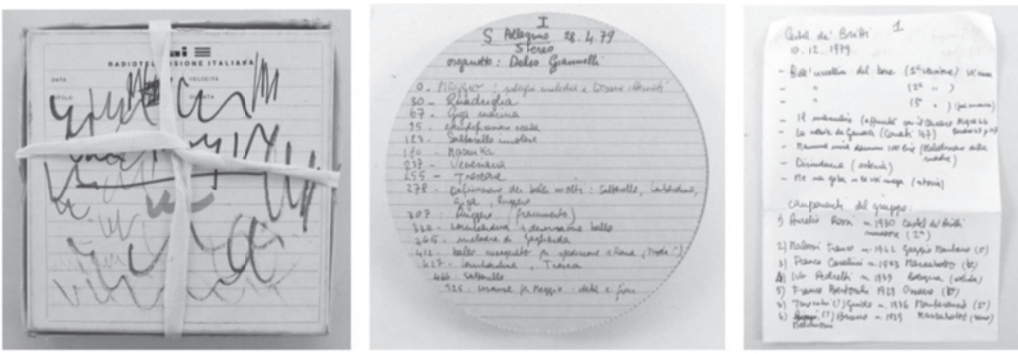
Number of tracks (channels) for each side. Typical values are 1, 2, 4. See previous field.

Signal type. Typical values are mono, stereo, or quadriphonic. See previous field.

Tape transfer rate. Two observations: this information is certain, i.e., it is learned or confirmed during signal extraction. If the speed is indicated on the box and it clashes with the actual speed, the latter should be noted here. Second observation: by design, there is only option for the transfer speed, however there are tapes showing multiple transfer rate speeds. Considering the low number of tapes with multiple speeds, the author decided to keep the model simple and provide only one field for this information. This means that there is a limitation to the manipulation of this information (questions like "how many tapes show different speeds?" or "what is the carrier with the most speed changes?" cannot be automatically processed). However, no information is lost because eventual speed changes are thoroughly annotated in the "Notes on the signal". For example, "The tape was recorded at two different speeds (7 1/2 ips and 3 3/4 ips), therefore the preservation copy contains two separate files. At the beginning of the recording the right channel is silent". Otherwise typical value: 7 1/2 ips (19 cm/s)—in both metric systems.

Take-up reel wind type before treatment. Tails-out or Heads-out for most tapes. If tail and head cannot be determined (for example, for two-sided tapes), simply "2-sided tape" is specified. For compact cassettes, options are "beginning of side A or 13", or "other" if the tape was not fully rewound on either side (see Figure 3).

Take-up reel wind type after treatment. This should be "tails-out", or the opposite in which the tape was found. "Print-through is a problem of analog audio recording tape. It occurs when a strongly recorded (magnetized) section of tape is embedded in a tape wind pack next to laps with low recorded magnetization. With time, the strong magnetic signals will imprint copies of themselves on the weakly magnetized adjacent laps. [. . .] When recordings are stored heads out, the print-through information precedes the



(a) Container. (b) Container. (c) Attachment.

Fig. 2. Examples of containers with text (a-b) and attachment (c).

recorded information and is most disconcerting. When recordings are stored tails out, the print-through may become less obvious because an echo is less objectionable than a pre-echo. Storage of analog audio tape tails out has the additional advantage of requiring a rewind, which decreases the print-through level" (AES, 1997 [r2008]). A secondary legitimate option is to store tapes in the opposite direction in which they were found, for example, to avoid spoking (IASA [International Association of Sound and Audiovisual Archives], 1999; Casey, 2008).

Notes on the original carrier. Comments about the physical carrier or its attachments.

Notes on the signal. Comments about the audio signal. Some examples were given in the previous fields.

4.2.2. State of preservation of the source document

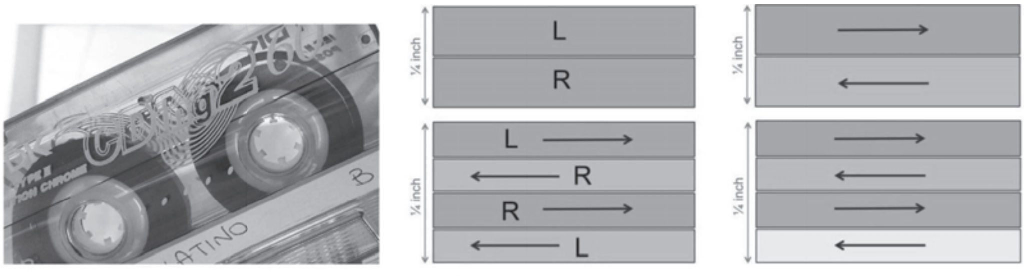
This section contains the report of the assessment of the state of preservation of the source carrier. It refers to the state of preservation of the carrier right before content transfer, before any physical restoration. Most signs and symptoms of degradation can be detected during visual inspection, but a few are integrated at a later time because they can only be detected during replay. The controlled vocabulary used for this assessment is the result of a comparative study carried out by the author (Bressan, 2018). Please see the enclosed sample data sheet for an example of report, and refer to Bressan (2018) for the explanations.

4.3. Description of the preservation copy

4.3.1. General information

Signature. A preservation copy is in its own right a new item in the digital archive, therefore it has a signature. In the naming convention adopted by the system described in this article, the signature consists of the root "cc", for "copia conservative" (preservation copy in Italian), an underscore and the signature of the source document the copy was derived from. For example, cc_MAGRN037.

Provenance (Archive). Normally this matches the field "Archive" in subsection 4.2.1, but it can occasionally be different if the organisation that stores the all-digital archive decides that it is convenient to reorganise the logical structure. Archives normally have an internal structure, but with digitisation it is also often the case that multiple archives are grouped together in the same digital platform, and even if not functionally associated to



(a) Compact Cassette. (b) Track configuration. (c) Track configuration.

Fig. 3. Compact cassette with a winding corresponding to the option “other” (a), and evidence of windowing, a “deformation of the layers of tape within the tape pack to the extent where light can be seen through it” (IASA [International Association of Sound and Audiovisual Archives], 1999). Examples of common track configurations for 6.35 mm (1/4-inch) reel-to-reel audio tape (b-c). Number of sides (i.e., playback directions) is 1 for (b)-top, 2 for (b)-bottom and (c)-top, 4 for (c)-bottom. Number of tracks (channels per side) is 2 for (b)-top and -bottom, 1 for (c)-top and -bottom. Signal type is stereo for (b)-top and -bottom, mono for (c)-top and –bottom; (c)-bottom demonstrates a typical recording practice, although not standardised. Drawings are schematic and actual track size is not shown.

one another, they are managed together (IFLA – Audiovisual and Multimedia Section, 2002, p. 7). To facilitate this management, archives might be renamed. This name is useful to retrieve resources in the digital repository; it is not necessarily the legal name.

Creation date of this preservation copy. The date on which the new record was created in the database. This field is filled in automatically. Completing a preservation copy, from creation to the final validation of the data, might take up to several days. This date does not give information on how long the working process for this document was. The date format is: YYYY-MM-DD.

Venue of content transfer. The name of the institution (research lab, private business, public service, etc.) where the content transfer was carried out.

Supervisor. The person in charge of the content transfer, the direct responsible for every intervention during the process. Not the head of the unit or the legal responsible of the institution. This is who personally carried out the digitisation. It normally should be an employee of the institution mentioned in the previous field.

4.3.2. Track configuration

Before signal extraction, the format of the recording should be determined: for open reel tapes, this mainly includes the track configuration and tape transfer speed—but also the equalisation curve and the noise reduction system applied at recording time. This section of the data sheet contains a description of the track configuration, specifically the number of tracks (or channels) in which the tape width is divided, and the playback directions (indicated by arrows, or assumed left-to-right without arrows; see Figure 3). Non-standard uses of the tape parameters to extend the duration of the tape are not uncommon: for example, recording two mono signals on a stereo device (by muting one channel at a time, see Figure 3c-top), or up to four mono channels on a double stereo or quadriphonic device. Changing the tape transfer speed helps to extend the duration of the tape too, but speed is not represented in Figure 3. The examples in Figure 3 are all rather common: the necessity to include an image to represent the track configuration and playback directions was raised during a research project where all these possibilities were mixed on the same tape (including multiple speed changes), including parts of the tape recorded in both directions on the same tracks. Because the images do not represent all the relevant parameters (e.g., speed changes), extra information is included in the "Notes on the signal" (subsection 4.2.1), and in the naming convention of the audio files. The note in the example in Appendix A reads: "The recording begins at speed 1 7/8 ips, and switches to 7 1/2 ips at minute 7:07 of the file MAGRN037_ve1178". The audio files in the preservation copy are MAGRN037_ve1712.wav, MAGRN037_ve1178_chL.wav, and MAGRN037_ve1178_chR.wav. This is a case where a stereo tape produced three digital audio files, instead of one as expected. Why? As the note says, there is a speed change during the recording (from 1 7/8 ips to 7 1/2 ips). A speed change does not have an easy parallel in the preservation of traditional archive materials like books. It is not self-evident how speed changes should be dealt with in a philological approach to preservation. Considering that no "stitching" should occur in the audio stored in the preservation copy, the solution adopted by the author provides that the entire length of the tape be read at every speed used on that tape side (in this case, there is only one side). This will proceed two audio files, both of which will contain part of the audio at the wrong speed. However, this will maintain the information of the speed change, without altering the frequency content of the audio file, and allow for "stitching" and processing at a later time (which should occur before the material is delivered to the end users). This preservation copy contains three files because the transfer at the lower speed produced a file the size of which exceeds the recommended limit for safe storage, therefore the right and left channels were saved separately. This is not considered illegal intervention because the files can be recombined without any loss of information.

4.3.3. Description of the audio files

This text opens the section: "This preservation copy contains [number] audio file(s). Each file begins with a print of the noise floor of the active system (timing: 00:00:00–00:00:05 sec)".

The print of the noise floor is a short recording (five seconds) of the noise produced by the digitisation setup when all the equipment is running but no signal is being recorded: it is expected to be negligible.

Except for the fields marked with a star (see below), every feature is automatically extracted from the signal. Metadata extraction is performed with PSKit implementing JHOVE, an extensible software framework for performing format identification, validation, and characterisation of digital objects[2].

For each audio file in the preservation copy, the following sets of information is provided:

File name. Name of the audio file. For the naming convention, see subsection 4.3.2.

Duration (hh:mm:ss). Total duration of the audio file, including the five seconds of the print of the noise floor of the active system.

Size. Human readable size of the file on disk (e.g., 678MB or 2.1GB).

File extension. The file extension as it appears after the last dot in the file name: "wav".

Format (MIME Type). This field contains a pointer to a controlled vocabulary in the database, where MIME codes are stored[3]. Typically, audio/wav.

---

2. See http://jhove.openpreservation.org, accessed 15 October 2018.
3. MIME codes are unique identifiers for digital files, according to their nature and format. MIME stands for "Multipurpose Internet Mail Extensions".

Encoding. Type of encoding of the wave file. Typically: PCM audio in integer format.

Profile. Typically, WAVEFORMATEX, Broadcast Wave Version 1.

Status. Validity status check.

Expected value. Well-Formed and valid.

Number of tracks (channels). Should match that of the source carrier.

Bit-depth. Number of bits per sample. Typically, 24.

Sample rate. The sample rate is typically 96kHz, except D/D transfers for which it matches the sample rate of the source document, and particularly endangered and obsolete carriers for which it can be higher (see Bressan and Canazza [2013], section 3.5.1).

Byte order. Typically: LITTLE_ENDIAN.

First sample offset. Offset of the first byte of sample data.

Compression. Eventual compression applied to the audio data. Should be: none.

***Equalisation curve applied during signal extraction.** For this field, several options are available but not "unknown"—unlike in the equivalent for the source document. We may not be sure what curve to apply, and we should gather information to determine the best guess, which could also be "none". However, "unknown" is not permitted because the operator will make a choice about this item and it should be declared.

***Noise Reduction System applied during signal extraction.** Same as previous, "unknown" is not permitted. Typical value: "none".

***Tape transfer rate applied during signal extraction.** This is the speed at which the tape has actually been read during signal extraction. It should match the equivalent field in the description of the source document.

Max. amplitude (normalised to 1). Max amplitude peak (positive or negative) reached by the waveform. Normalising to 1 means that all the values assumed by the waveform samples are comprised between 0 (digital silence) and 1. Example: 0.484299.

RMS amplitude (normalised to 1). Root Mean Square (RMS) amplitude value. Example: 0.022442.

Checksum MD5. A string indicating the checksum signature of the audio object. See explanation below. Example: 5b4c565c337c2cb625007e716e0d5f1c.

Checksum CRC32. A string indicating the checksum signature of the audio object. See explanation below. Example. 49651b13.

Checksum SHA-1. A string indicating the checksum signature of the audio object. See explanation below. Example: af944772c8f6cf82751ab60ab4fa5d1cd28eb3ea.

The terms "checksum" and "message digest" are commonly used interchangeably. However, the term "checksum" is more correctly used for the product of a cyclical redundancy check (CRC32), whereas the term "message digest" refers to the result of a cryptographic hash function (MD5, SHA-1). Both are alphanumeric strings of predetermined length, calculated from a big chunk of data, for the purpose of detecting errors, which may have been introduced during its transmission or storage. They have several applications; in this context, they are mainly used to check data integrity over time, by comparing the string in this data sheet with the string recalculated at the time of the verification. This avoids susceptibility to a single point of failure in the system. According to the preservation best practices expressed in Casey and Gordon (2007), the checksums are stored within the preservation copy (in the descriptive sheet and in an XML file) as well as in the database, with backup copies in multiple physical locations. In the setup adopted by the author, automatic routines checks every file for integrity every month and stored the results in a special module of the database. A mail notification is sent if a mismatch is detected.

(a) Reel replacement. (b) Splice. (c) Debris on magnetic head.

Fig. 4. Example of a critical moment in reel replacement.

Another value stored in the database but not shown here: Mean amplitude (normalised to 1).

4.3.4. Photographic documentation

The photographic documentation does not only include a thumbnail of the pictures that were taken before and during the visual inspection of the source document. It can also include 1) pictures that were taken after the process of carrier optimisation (e.g., dust removal, leader tape replacement, reel replacement), and 2) relevant visual information emerging during the process of signal extraction (e.g., splices, pictures of the reading device, e.g., in case of debris on the magnetic heads; see Figure 4).

The comment in the subsection specifies the number of pictures enclosed in the preservation copy (e.g., "This preservation copy contains 12 images"). This should match with the number of thumbnails. Thus it is possible to verify that none of the pictures was removed (accidentally or intentionally), nor replaced or counterfeited.

The comment also redirects the user to "the archive database" for "more information about each image". As mentioned in subsection 2.1, the data in this sheet is a subset of the overall data maintained in the preservation database. Table 1 summarises the complete metadata set stored in the database about each picture.

This information is automatically extracted from the image files and ingested in the database by a function of PSKit implementing ExifTool, a free open-source platform-independent command-line application for reading, writing and editing meta information in a wide variety of files[4].

4.3.5. Video documentation

Especially for open-reel tapes, video documentation is produced to capture defects, text and any event of interest that are revealed during replay. The camera is positioned close to the reading heads (see Figure 5) and the data stream is sent to a computer other than that receiving the audio stream (both are dedicated). The monitors are next to each other so the operator can easily monitor both at the same time. The audio included in the video comes from a secondary output of the playback device so it is perfectly synchronised. Besides documenting defects and relevant events in the course of the replay, the video can also be exploited in many applications for access (Bressan et al., 2017). Table 1 summarises the complete metadata set stored in the database about each picture.

4. See https://www.sno.phy.queensu.ca/phil/exif tool/, accessed 15 October 2018.

(a) Freeze frame. (b) Camera position. (c) Recording setup.

Fig. 5. Video recording during the digitisation of an open reel tape.

Table 1. Complete list of metadata stored in the preservation database for pictures (left) and videos (right). Fields marked with a star are not automatically extracted. The "mime" field contains a pointer to a controlled vocabulary in the database, where MIME codes are stored (see subsection 4.3.3).

PICTURES		VIDEO	
Attribute	**Example**	**Attribute**	**Example**
file_name*	MAGRN037_01.JPG	file_name*	MAGRN037.mov
size	4.1 MB	size	4.5 GB
mime	image/jpeg	mime	video/quicktime
exposuretime	1/45	framerate	25 fps
iso	100	imagesize	720x576
shutterspeed	1/30	bitdepth	24 bit
aperturevalue	3.1	compression	DV - PAL
brightnessvalue	3.16	audioformat	lpcm
meteringmode	Multi-segment	audiochannels	2
focallength	4.3 mm	audiobitspersample	16
colorspace	sRGB	audiosamplerate	48 kHz
imagesize	4288x3216	shutterspeed	1/120
encodingprocess	Baseline DCT, Huffman coding	recordingdevice*	Panasonic AG-DVX100BE
bitspersample	8	notes*	Default: none.
recordingdevice*	FUJIFILM FinePix S4200		
notes*	Default: none.		

4.4. Technical scheme of the transfer system

Information about the technical setup of the transfer system is summarised in the scheme included in this subsection. See the sample data sheet in the Appendix. The scheme ought to be read from the top down. The input is the source carrier. The left part of the scheme represents the flow of the audio data. The carrier is read with a compatible machine (type, brand, and model specified in the block). The audio signal is sent to an analog-to-digital and digital-to-analog (A/D-D/A) converter (brand and model specified in the block). The audio converted to digital is sent to a computer (brand, model, and basic features specified in the block), and there recorded. The same audio signal that is being recorded is sent back to the converter, converted to from digital to analog again and amplified on two loudspeakers. It is very important to point out that the operator supervising the process is constantly monitoring the audio stream, and in particular the audio stream after digitisation (that is why the signal is first sent to the computer and then back to the monitoring system), in order to detect eventual artefacts introduced by the equipment or the power grid.

The right part of the scheme represents the flow of contextual information, i.e., the video and photographic documentation. The type, brand, and model of the devices used to produce the documentation is specified in the block. Note: the sample data sheet in the Appendix only includes photo documentation. The documentation is sent to a computer (brand, model, and basic features specified in the block) for processing and validation. Finally, the preservation copy is assembled and sent to a storage system. The storage system is not necessarily that of the institution that owns the documents, but that of the facility in charge of content transfer. That means that the type, brand, and model of the storage system specified in the block correspond to the system where preservation copies are sent at the end of the working cycle, and kept there until the owner institution is ready to move to the next step in the preservation workflow.

Other important information included in the technical scheme is the alignment between analog and digital audio levels. The setup described in this article uses the EBU *Technical Recommendation R68-2000* for the alignment level in digital audio production equipment and in digital audio recorders. "It is recommended that in digital audio equipment coding levels for digital audio signals are set such that they correspond to an alignment level which is 18 dB below the maximum possible coding level of the digital system, irrespective of the total number of bits available" (EBU, 2000)[5].

If the alignment is constant (not adjusted, "optimized" for every carrier), with this information, it is possible to use the digital audio file to calculate the intensity of the magnetisation on the original tape (measured in Weber [Wb]).

Coloured arrows in the scheme distinguish between analog data (blue), digital data (red), physical embedding (green), and touch free procedures (orange).

Conclusion

This article has presented a detailed description of a data sheet for audio preservation. The data sheet includes information on the original carrier, the preservation copy and the process of content transfer. A complete example of the data sheet is enclosed in the Appendix. The data sheet served as example to make a case for the importance of supporting the choice of audio preservation metadata with clear documentation on the conceptual framework applied to the preservation workflow. In particular, some key concepts have been discussed, 1) the difference between content transfer and content description, 2) the conceptual model of the audio document, 3) the status of the audio document (cultural heritage or not?), and 4) high- vs. low-profile resources. The full documentation of

5. *Excellent*: Never used (for commercial documents only).
Good: The carrier does not require treatment before playback (no corruptions or negligible).
Fair: The carrier requires treatment before playback (serious corruptions).
Very bad: The carrier cannot be played, regardless of treatment.

the database schema associated with the framework and data sheet described in this article is available upon request to the author. The article aims to contribute to the development of a full theory of information preservation (Dubin, Futrelle, and Plutchak, 2006; Brock-Nannestad, 1997).

English Abstract
This article presents the analytical description of a data sheet for audio preservation. The data sheet is normally enclosed in an audio preservation copy and contains information about the original carrier, the preservation copy and the process of content transfer. The data serves as example to make a case for the importance of a clear conceptual framework that is made explicit through open documentation. In particular, some key concepts for the design of the framework are discussed: the difference between content transfer and content description, the importance of developing a good conceptual model for the audio document, the status of the audio document (cultural heritage or not?), and high- vs. low-profile resources. The article aims to contribute to the development of a full theory of information preservation, while at the same time providing a complete example of data sheet ready to use.

French Abstract
Cet article présente la description analytique d'une feuille de données pour la conservation des documents audio. La feuille de données est jointe à l'intérieur d'un exemplaire du document destiné à la conservation, et contient les informations sur sa provenance, la copie de conservation, ainsi que les modalités de transfert du contenu. Les informations servent d'exemple pour faire cas de l'importance de la clarté et de la précision de la structure conceptuelle à travers une documentation ouverte et accessible. Certains concepts clés concernant la forme de la structure sont discutés: la différence entre transfert de contenu et description de contenu, l'importance de développer un bon modèle conceptuel pour les documents audio, le status du document audio lui-même (d'héritage culturel ou non), les ressources plus ou moins connues.

L'article vise à contribuer au développement d'une théorie complète sur la conservation de l'information, et en même temps à proposer un exemple complet de feuille de données prête à l'emploi.

German Abstract
Dieser Beitrag präsentiert die analytische Beschreibung eines Datenmusters für die Bestandserhaltung von Audiomedien. Das Datenmuster ist normalerweise in eine Sicherungskopie des Audiomediums integriert und beinhaltet Informationen über das ursprüngliche Format, die Sicherungskopie und den Umwandlungsprozess des Inhaltes. Die Daten dienen als Beispiel für die These der Wichtigkeit klarer konzeptioneller Richtlinien, die durch offene Dokumentation zugänglich gemacht werden. Der Artikel erörtert einige Voraussetzungen für die Anlage von Richtlinien: die Unterscheidung zwischen dem Umwandlungsprozess und der Beschreibung des Inhaltes, die Wichtigkeit, ein gutes konzeptionelles Modell für das Audio-Dokument zu entwickeln, die Bedeutung des Tondokumentes (kulturelles Erbe oder nicht?) und die Abgrenzung des Bestandes bezüglich seiner Bedeutung bzw. Wertigkeit. Der Aufsatz soll einen Beitrag zur Entwicklung einer vollständigen Theorie zur Bewahrung von Informationen leisten, gleichzeitig aber auch ein vollständiges und nachnutzbares Datenschema zur Verfügung stellen.

References

AES. 1997 (r2008). AES recommended practice for audio preservation and restoration – Storage and handling – Storage of polyester-base magnetic tape. New York, NY: AES.
Brandi, Cesare. 2005. Theory of Restoration. Arte e restauro. Florence (Italy): Nardini.
Bressan, Federica. 2018. “A framework for the description of age-related symptoms in audio media: Definition and implementation”. Studies in Conservation under review.

Bressan, Federica, and Sergio Canazza. 2013. "A Systemic Approach to the Preservation of Audio Documents: Methodology and Software Tools". *Journal of Electrical and Computer Engineering* 2013 (Article ID 489515): 21.
Bressan, Federica, Sergio Canazza, Carlo Fantozzi, and Niccolo Pretto. 2017. "Tape music archives: From preservation to access". *International Journal on Digital Libraries* 18 (3): 233–249.
Brock-Nannestad, George. 1997. "Applying the Concept of Operational Conservation Theory to Problems of Audio Restoration and Archiving Practice". In *Proceedings of the 103rd AES Convention*, New York, NY, 1–15.
Brock-Nannestad, George. 2000. "The Rationale Behind Operational Conservation Theory". In *Conservation Without Limits*, IIC Nordic Group XV Congress, ed. by Riitta Koskivirta, Helsinki, FI, 23–26 August, 21–33.
Brock-Nannestad, George. 2001. "What Are The Sources Of The Noises We Remove?" In *20th AES International Conference on Archiving, Restoration, and New Methods of Recording*, 5–7 October, 175–180.
Casey, Michael. 2008. *FACET (Field Audio Collection Evaluation Tool) – Format Characteristics and Preservation Problems*. Version 1.0. Technical Report. Bloomington: Indiana University.
Casey, Mick, and Bruce Gordon. 2007. *Sound Directions: Best Practices for Audio Preservation*. Harvard University and Indiana University.
CASPAR Consortium. 2008. *Report on OAIS – Access model*. Technical Report. Centre national de la recherche scientifique (CNRS) and Universite de Technologie de Compiegne (UTC).
Dubin, David, Joe Futrelle, and Joel Plutchak. 2006. "Metadata Enrichment for Digital Preservation". In *Proceedings of Extreme Markup Languages 2006 Conference*, 1–9.
Duranti, Luciana. 2012. *InterPARES3 – TEAM Canada Final Report*. Technical Report. University of British Columbia.
EBU. 2000. *EBU Technical Recommendation R68-2000*, "Alignment Level in Digital Audio Production Equipment and in Digital Audio Recorders". European Broadcasting Union.
Edmonson, Ray. 2002. *Memory of the World: General Guidelines to Safeguard Documentary Heritage*. Paris (France): UNESCO.
Edmonson, Ray. 2016. *Audiovisual Archiving: Philosophy and Principles*. Bangkok: UNESCO.
Factor, Michael, Ealan Henis, Dalit Naor, Simona Rabinovici-Cohen, Petra Reshef, Shahar Ronen, Giovanni Michetti, and Maria Guercio. 2009. "Authenticity and Provenance in Long Term Digital Preservation: Modeling and Implementation in Preservation Aware Storage". In *First Workshop on the Theory and Practice of Provenance*, San Francisco, CA, 23 February.
Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
Gigliozzi, Giuseppe. 2003. *Introduzione all'uso del computer negli studi letterari*. Milano, Italy: Bruno Mondadori.
Hedstrom, Margaret. 2002. *It's About Time: Research Challenges in Digital Archiving and Long-term Preservation. Final Report on a Workshop on Research Challenges in Digital Archiving*. Technical Report. Washington, DC: Library of Congress.
IASA (International Association of Sound and Audiovisual Archives). 1999. *The IASA Cataloguing Rules*. Hungary: IASA Editorial Group.
IASA-TC 04. 2004. *Guidelines on the Production and Preservation of Digital Objects*. IASA Technical Committee.

IFLA – Audiovisual and Multimedia Section. 2002. *Guidelines for Digitization Projects: For Collections and Holdings in the Public Domain, Particularly those Held by Libraries and Archives*. Technical Report. Paris (France): International Federation of Library Associations and Institutions (IFLA).

Mimoso, Joao Manuel. 2009. "Cesare Brandi's Theory of Restoration and azulejos". In *Proceedings of the International Seminar on Conservation of Glazed Ceramics*, Lisbon, Portugal.

Orlandi, T. 1997. "Discipline Umanistiche e Informatica. II Problema della Formalizzazione, Chap. II, Testo Critico e il Supporto Magnetico", 159–172. Roma, Italy: Accademia Nazionale dei Lincei.

Valentini, Francesca. 2007. "Cesare Brandi's Theory of Restoration: Some Principles Discussed in Relation with the Conservation of Contemporary Art". HAWK Hildesheim / Holzminden / Göttingen. doi:10.5165/HAWK-HHG/378.

ARCHIVE DOCUMENT
cc_MAGRN037

Note: This document should be located in the *root* directory of the preservation copy `cc_MAGRN037`.

Contents

1 List of documents enclosed in this preservation copy

This section provides a list of the documents enclosed in the preservation copy `cc_MAGRN037`, with the corresponding file extension and relative path.

Type of document	File extension	Path
This data sheet	pdf	./
Audio files	wav	./audio
Checksums of the audio files	xml	./metadata/checksum
Images of the source document	jpeg	./metadata/images
Video of the source document	mov	./metadata/video

2 Description of the source document

This section contains a general description of the source document, followed by a detailed assessment of its state of preservation.

2.1 General information

Archive	Centro per il Dialetto Romagnolo
Collection	Tullia Magrini
Label for preservation	MAGRN037
Signature in the original archive	EMILIA ROMAGNA 41; BRUSCHI 5
Carrier type	Open-reel tape
Archiving procedure	Open-reel
Accompanying material	See enclosed images of the source document
Text and notes	See enclosed images of the source document
Container type	Plastic
Container brand	Basf
Carrier brand	Basf
Carrier model	DP 26
Tape width	1/4 inch
Reel flange diameter	5 in (12 cm)
Type of recording	Analog
Recording technique	Magnetical
Equalisation curve	unknown
Noise Reduction System	unknown
Number of sides	1

Number of tracks (channels) for each side	2
Signal type	Stereo
Tape transfer rate	7 1/2 ips (19 cm/s)
Take-up reel wind type before treatment	Heads-out
Take-up reel wind type after treatment	Heads-out
Notes on the original carrier	none
Notes on the signal	See Fig. 1 for track configuration. The recording begins at speed 1 7/8 ips, and switches to 7 1/2 ips at minute 7:07 of the file MAGRN037_vel178.

2.2 Assessment of the state of preservation

The overall state of preservation of the archive document `MAGRN037` is: Good[1]
The tape has been tested against the following symptoms and signs of degradation. Positive evaluations are marked in the check-box, and the descriptions in the right column include the severity code on a 5-stepped scale from zero to four; if the check-box is clear, the code is assumed to be zero.

☒ Pack: Tape pack/Wind quality	Uneven pack without blocking nor damaged edges. (Code 2)
☐ Pack: Blocking	The tape shows no signs of blocking.
☐ Pack: Leafing	Leafing is completely absent.
☐ Pack: Loose wind	The tape wind is not loose.
☐ Pack: Windowing	Windowing is completely absent.
☐ Pack: Spoking	Spoking is completely absent.
☐ Dirt, dust, oil	Dirt, dust or oil are completely absent.
☐ Other particulates	Other particulates are completely absent.
☐ Liquid stains	Liquid stains are completely absent.
☐ Splices	Splices are absent or in perfect condition.
☐ Vinegar odor	No sign of vinegar odor.
☐ Mold	No sign of mold.
☐ Pests	Pests are completely absent.
☐ Other bio[logical contamination]	Other bio are completely absent.
☐ Powder, crystals	Powder, crystals are completely absent.

[1] *Excellent*: Never used (for commercial documents only).
Good: The carrier does not require treatment before playback (no corruptions or negligible).
Fair: The carrier requires treatment before playback (serious corruptions).
Very bad: The carrier cannot be played, regardless of treatment.

☐ Tears, breaks	Tears, breaks are completely absent.
☐ Brittle, curling	Brittle, curling is completely absent.
☐ Folds, cinching	Folds, cinching are completely absent.
☐ Cupping	Cupping is completely absent.
☐ Edge damage	Edge damage is completely absent.
☐ Backcoat shedding	No evidence of backcoat shedding.
☐ Magnetic coating shedding	No evidence of magnetic coating shedding.
☐ SBS-SSS	No evidence of SBS-SSS.
☐ Bleeding	No sign of bleeding.
☐ Curvature	No sign of curvature.
☐ Embossing	No sign of embossing.
☐ Interlayer adhesion	No sign of interlayer adhesion.
☐ Kink/Wrinkle	Kinking/Wrinkling is completely absent.
☐ Magnetic losses	Magnetic losses are completely absent.
☐ Manufacturing surface defects	No sign of manufacturing surface defects.
☐ Gummy deposit	Gummy deposit is completely absent.
☐ SBS-UP	No evidence of SBS-UP.
☐ Print-through	No sign of print-through effect.

3 Description of the preservation copy

This section contains a general description of the preservation copy `cc_MAGRN037`, followed by detailed descriptions of the multimedia files representing the contextual information.

3.1 General information

Signature	cc_MAGRN037
Provenance (Archive)	Centro per il Dialetto Romagnolo
Creation date of this preservation copy	2015-04-14
Venue of re-mediation	Centro di Sonologia Computazionale (CSC), Dept. of Information Engineering, University of Padova
Supervisor	Federica Bressan

3.2 Track configuration

Figure 1 shows the number of tracks in which the tape is divided (in its width), the channel configuration and the recording direction(s). When the direction is not specified by arrows, it is

intended to be left-to-right.
Note: speed changes are not marked in Figure 1. Please refer to the Notes about the signal in Section 2.

Figure 1: Scheme of this tape configuration.

Speed changes are also expressed in audio files names. For more details on each audio file see the next subsection. Complete list of audio files in this preservation copy:
`MAGRN037_vel712.wav`
`MAGRN037_vel178_chL.wav`
`MAGRN037_vel178_chR.wav`

3.3 Description of the audio files

This preservation copy contains 3 audio file(s). Each file begins with a print of the noise floor of the active system (timing: 00:00:00 – 00:00:05 sec).

3.3.1 Audio file 1

File name	MAGRN037_vel712.wav
Duration (hh:mm:ss)	00:32:40
Size	1.1 GB
File extension	wav
Format (MIME Type)	audio/wav
Encoding	PCM audio in integer format
Profile	WAVEFORMATEX, Broadcast Wave Version 1
Status	Well-Formed and valid
Number of tracks (channels)	2
Bitdepth	24 bit
Sample rate	96 kHz
Byte order	LITTLE_ENDIAN
First sample offset	46 byte
Compression	none
Equalization curve applied during signal extraction	CCIR

Noise Reduction System applied during signal extraction	none
Tape transfer rate applied during signal extraction	7 1/2 ips (19 cm/s)
Max. amplitude (normalised to 1)	0.484299
RMS amplitude (normalised to 1)	0.022442
Checksum MD5	5b4c565c337c2cb625007e716e0d5f1c
Checksum CRC32	49651b13
Checksum SHA-1	e0ee8862706b6d51c073b34959de85377a16d6aa

3.3.2 Audio file 2

File name	MAGRN037_vel178_chL.wav
Duration (hh:mm:ss)	02:10:42
Size	2.1 GB
File extension	wav
Format (MIME Type)	audio/wav
Encoding	PCM audio in integer format
Profile	WAVEFORMATEX, Broadcast Wave Version 1
Status	Well-Formed and valid
Number of tracks (channels)	1
Bitdepth	24 bit
Sample rate	96 kHz
Byte order	LITTLE_ENDIAN
First sample offset	46 byte
Compression	none
Equalization curve applied during signal extraction	CCIR
Noise Reduction System applied during signal extraction	none
Tape transfer rate applied during signal extraction	7 1/2 ips (19 cm/s)
Max. amplitude (normalised to 1)	0.577166
RMS amplitude (normalised to 1)	0.023776
Checksum MD5	6f45326a70069fb0c77599b11c93e1b4

Checksum CRC32	ac522870
Checksum SHA-1	29173c44ab3981bcf4aa37db04b75b9dd9c8c9ed

3.3.3 Audio file 3

File name	MAGRN037_vel178_chR.wav
Duration (hh:mm:ss)	02:10:42
Size	2.1 GB
File extension	wav
Format (MIME Type)	audio/wav
Encoding	PCM audio in integer format
Profile	WAVEFORMATEX, Broadcast Wave Version 1
Status	Well-Formed and valid
Number of tracks (channels)	1
Bitdepth	24 bit
Sample rate	96 kHz
Byte order	LITTLE_ENDIAN
First sample offset	46 byte
Compression	none
Equalization curve applied during signal extraction	CCIR
Noise Reduction System applied during signal extraction	none
Tape transfer rate applied during signal extraction	7 1/2 ips (19 cm/s)
Max. amplitude (normalised to 1)	0.577193
RMS amplitude (normalised to 1)	0.023452
Checksum MD5	b213b03ee460a00a889d2a78948b6f5e
Checksum CRC32	df15c2fa
Checksum SHA-1	af944772c8f6cf82751ab60ab4fa5d1cd28eb3ea

3.4 Photographic documentation

This preservation copy contains 12 image files.
For more information about each image, see the archive database.
Figure 2 shows a preview of the images.

Figure 2: Preview of the image files contained in this preservation copy.

4 Technical scheme of the transfer system

Figure 3 shows the technical scheme of the transfer system.

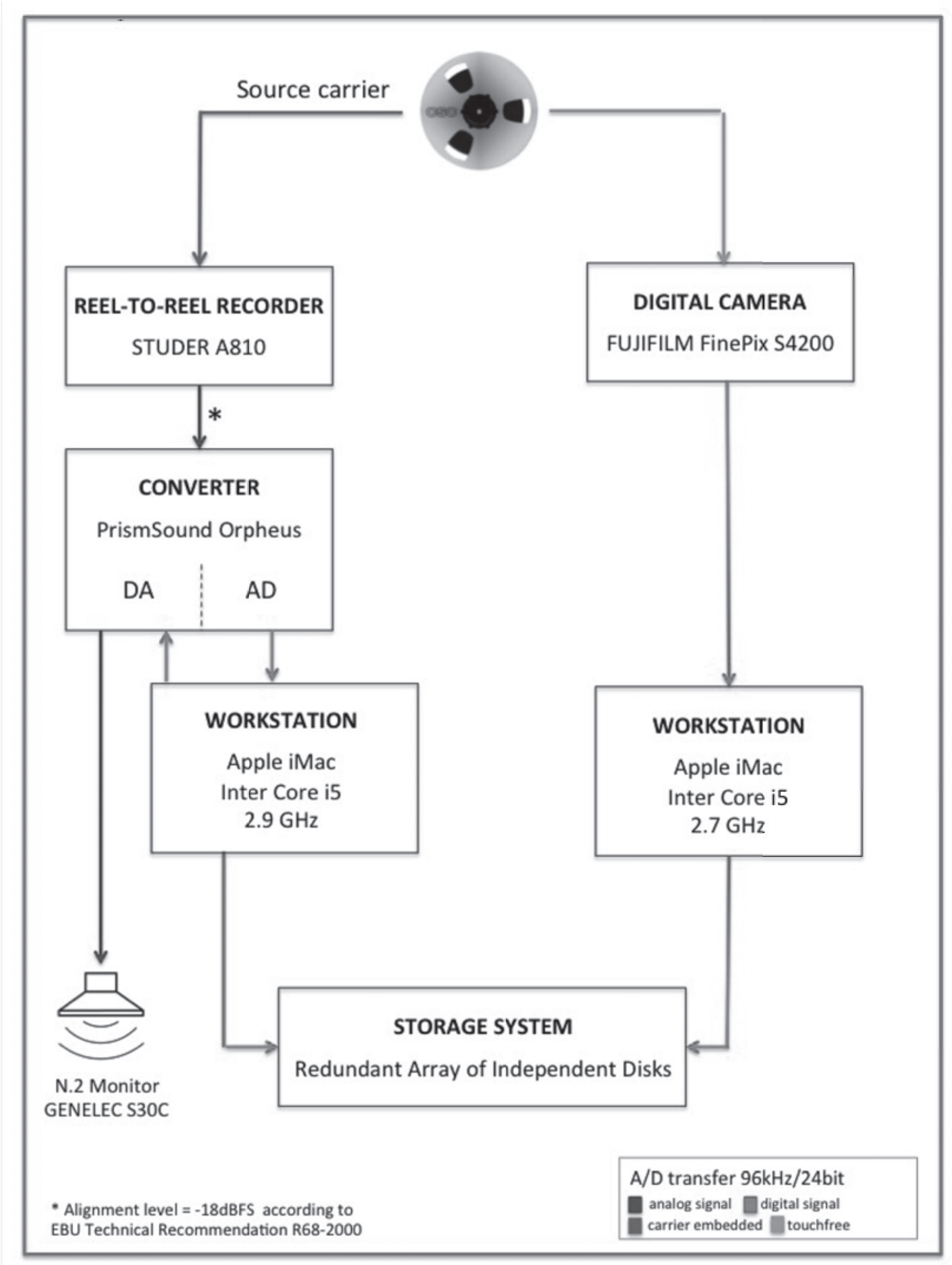


Figure 3: Technical scheme of the transfer system.