# Community-Based Survey and Oral Archive Infrastructure in the *Archivio Vi.Vo.* Project

**Silvia Calamai**
Siena University
Arezzo, Italy
silvia.calamai
@unisi.it

**Niccolò Pretto**
University of Padova
Padova, Italy
niccolo.pretto
@dei.unipd.it

**Maria Francesca Stamuli**
Ministry of Culture and
Tourism, Florence, Italy
mariafrancesca.stamuli
@beniculturali.it

**Duccio Piccardi**
Siena University
Arezzo, Italy
duccio.piccardi
@unisi.it

**Giovanni Candeo**
ILC, CNR
Pisa, Italy
giovanni.candeo
@ilc.cnr.it

**Silvia Bianchi**
Siena University
Arezzo, Italy
bianchi.silvia
@unisi.it

**Monica Monachini**
ILC, CNR
Pisa, Italy
monica.monachini
@ilc.cnr.it

## Abstract

Audio and audiovisual archives are at the crossroads of different fields of knowledge, yet they require common solutions for both their long-term preservation and their description, availability, use and reuse. *Archivio Vi.Vo.* is an Italian project financed by the Region of Tuscany, aiming to: (i) explore methods for long-term preservation and secure access to oral sources, and (ii) develop an infrastructure under the CLARIN-IT umbrella offering several services for scholars from different domains interested in oral sources. This paper describes the project's infrastructure and its methodology through a case study on Caterina Bueno's audio archive.

## 1 Introduction

Audio and audiovisual archives[1] are scattered all over the Italian peninsula, from researchers' private houses to universities and research centres, from cultural institutions (e.g., Istituti per la Resistenza) to State institutions, such as State Archives and Libraries. The pilot survey made by Galatà and Calamai in 2018 has emphasised the status of precariousness, instability, and insecurity that affects audio and audiovisual archives available at different communities of Italian researchers. Almost half of the resources listed in the survey (49.6%) were barely accessible. Only 9.2% of the resources was accessible and available, 4.6% was partially accessible, 35.1% was available upon request, 1.5% was available upon request but only for selected parts. As for the resources which were declared to be accessible, the access policies were as follows: only 9.2% of these resources was freely accessible online (with no authentication), 7.6% was accessible online via authentication, and 29% was accessible onsite (i.e. where the resources are physically stored). As for the long-term maintenance and preservation, the answer receiving the highest number of responses was no-one (43%), followed by reference institutes, such as associations, foundations, libraries and their archives (17%), reference universities (16%), and the owners/individuals themselves (15%)[2]. Several research projects in recent years have aimed to disseminate audio and audiovisual archives, which have been collected over the years by both researches and amateur fieldworkers. Some Italian[3] examples are, among others, *Grammo-foni. Le soffitte della voce*, also referred to as Gra.fo (Calamai and Biliotti, 2017), *Voci, parole e testi della Campania*[4], *I granai della*

---

[1] Archives containing speech and sounds are differently labelled according to the fields of expertise of the researchers (audio archives, oral archives, speech archives). Among linguistics, the label *corpus/corpora* is also used. In the present paper, the label "oral archives" is considered a synonym of "audio archives".

[2] Further details available in (Galatà and Calamai, 2019)

[3] A detailed review of similar projects from other countries goes beyond the scope of this paper. We limit ourselves to indicating the Telemeta (Fillon et al., 2014) project and the Endangered Languages Archive (ELAR - www.elararchive.org), which share some similarities with the work presented here. At the following URL oralhistory.eu/collections/clarin-eric an extensive list of European oral archives and related research projects can be found.

[4] www.archivicampani.unina.it/archivi_campani_dev Last visited February 4th, 2021

*memoria*[5], and *Circolo Gianni Bosio Audio Archives*[6]. Nevertheless, fragmentation and lack of common and shared standards are often the features of certain initiatives, whose duration over time crucially appears to be dependent on the duration of external funding, if any. Moreover, a researcher working with audio archives is not necessarily competent in long-term preservation of audio data and data management. Co-incidentally, not all the research projects dealing with audio archives receive financing for all the different professional profiles involved in their preservation, managing and valorisation.

Given this picture, it appears urgent and no longer postponable to provide an infrastructure offering: i) a long-term preservation service for audio archives, ii) a shared set of metadata compliant with the main international standards and FAIR principles, and iii) an access interface which takes into account the peculiarities of the audio modality and which is able to support researchers in different disciplines. This paper presents how the *Archivio Vi.Vo.* project tackles these problems, illustrating the overall methodology adopted and infrastructure that is being developed.

## 2    The *Archivio Vi.Vo.* Project

### 2.1    The Project

In 2019, the Region of Tuscany decided to support a project entitled *Archivio Vi.Vo.*, which aims to catalogue and disseminate oral archives. The partners involved are: Siena University (Silvia Calamai), CNR-ILC & CLARIN-IT (Monica Monachini), Soprintendenza Archivistica e Bibliografica della Toscana (Maria Francesca Stamuli) and Unione dei Comuni del Casentino (Pierangelo Bonazzoli). In order to reach the above-mentioned ambitious objectives, *Archivio Vi.Vo.* has concentrated most of its efforts on the design and development of an architecture, hosted by CLARIN-IT, the Italian consortium of the CLARIN research infrastructure, which could be used by several other projects concerning audio archives. A crucial step towards this aim concerns the definition of the metadata set(s) used to describe the data. This set has to be compliant with international archival standards, such as ISAD(G) and ISAAR, as well as several others chosen from different disciplines (cf. Section 4.1).

### 2.2    The Case Study

The architecture (cf. Section 3) is in the process of being validated on a specific audio archive, namely Caterina Bueno's audio archive, which appears to be rather challenging, for the following reasons: (i) it has a complex archival history, (ii) it is in a very poor conservation condition, and (iii) it contains highly heterogeneous audio material.

Caterina Bueno (San Domenico di Fiesole, IT, 2nd April 1943 – Florence, IT, 16th July 2007) was an Italian ethnomusicologist and singer (Giorgi et al., 2013). Her work as a researcher has been held in high regard for its cultural value, as it brought together many folk songs from Tuscan and central Italy that had been orally passed down from one generation to the next until the 20th century (when this centuries-old tradition started to vanish). Her work as a singer was always oriented towards research. At the age of twenty, she started travelling through the Tuscan countryside and villages recording Tuscan peasants, artisans, common men and women singing any kind of folk songs: lullabies, *ottave* (rhyming stanzas sung during improvised contrasts between poets), *stornelli* (monostrophic songs), narrative songs, social and political songs, and much more. These were the same songs that she sang in her performances, making them well-known and appreciated both in Italy and abroad in the second half of the 20th century, when she was at the pinnacle of her career. Caterina Bueno's sound archive is composed of about 476 analogue carriers (audio open-reel tapes and compact cassettes), corresponding to more than 700 hours of recordings, and it was digitised during the previous Gra.fo project. The analogue recordings were located with two different owners: part of them were stored at Caterina's heirs' house, while the rest was kept by the former culture counsellor of the Italian Municipality of San Marcello Pistoiese, in the *Montagna Pistoiese*, where a multimedia library was supposed to be set up. Unfortunately, disagreements and misunderstandings between the two parties have so far left the archive fragmented and inaccessible to the community. Both owners, independently, have turned to Silvia Calamai for the reassembly of the

---

[5] `www.granaidellamemoria.it` Last visited February 4th, 2021
[6] `www.circologiannibosio.it/archivio.php` Last visited February 4th, 2021

whole archive in the digital domain, so as to respect the artist's wishes. After being digitised, the carriers were returned to their owners.

In several cases, the original carriers were devoid of all the contextual information (place and date of recordings, speakers involved in the recordings). In other cases, the open-reel tapes were recorded at different speeds and using different track-head configurations, thus making the digitisation process and the creation of access copies rather complex. In this respect, Caterina Bueno's audio archive represents an extreme test case where different levels of complexity call into question different professional profiles and skills. In addition, such an archive may be of interest to different fields of research; from this point of view it would have to meet different needs according to different types of users.

## 3   User Needs in Oral Archives

In order to figure out the different needs and the different requests of those using audio archives in their research, an online survey was launched. The aim of the questionnaire was to gather together exploratory data on i) the profiles of Italian oral archive users; ii) their research routines (e.g., the type of document they are interested in in their searching strategies); iii) the features that they would like to add to an oral archive infrastructure. Our intent is to tailor the development process of the *Archivio Vi.Vo.* infrastructure (see below, Section 5) to the needs of the target population of users. The 56 responses will be thoroughly analysed here.

*Materials*. The questionnaire was anonymous and written in Italian. At the beginning, the participant was informed of the context and purposes of the survey; the expected completion time was explicitly stated and set at approximately 2/3 minutes. Then, the participant was given the questionnaire, which consisted of 11 items and could be divided into 3 thematic sections.

i) The first explored the respondents' profiles by asking them to state their main field of interest ("Linguistics", "Oral history", "Sociology", "Anthropology", "Ethnomusicology", "Sound and music computing", "Other"), number of years of experience ("0-5", "6-10", "10+"), age and gender.

ii) The second group of items pertained to their research routines, asking them to state their driving motivation for the use of oral archives ("Study and research", "Work - for Institutions or Archives", "Work - of different nature", "Leisure and Hobby", "Other"), their frequency of use of online oral archives (on a 1-4 Likert scale from "Never" to "Always"), the oral archives which were most familiar (open-ended), the frequency of their searches for specific type of documents ("Audio files", "Summaries of audio files", "Catalog cards", "Video files", "Other", on a 1-4 Likert scale from "Never" to "Always"), the frequency of their use of specific searching strategies ("By author", "By keywords", "By topic", "By genre", "By dialect/language", on a 1-4 Likert scale from "Never" to "Always") and their levels of perceived usefulness of specific searching strategies ("By genre", "By topic", "By keyword", "By language/dialect", "By abstract", "Other", on a 1-4 Likert scale from "Useless" to "Very useful").

iii) The last section consisted of a single open-ended question: the participants were directly asked to provide their suggestions for the development of a digital oral archive.

*Procedure*. The questionnaire was imported into Google Forms and distributed online via 8 mailing lists of the major Italian associations and their pertinent fields of research, such as Associazione Italiana Scienze della Voce[7] (AISV), Associazione Italiana di Storia Orale[8] (AISO), Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD)[9]. The survey was available from June 18th, 2020 to August 27th, 2020.

*Analysis*. Descriptive statistics of the demographic features of our respondents will be provided. Then, bar plots will be generated in R (R Core Team, 2021; Lüdecke, 2021) in order to present a compact visualisation of the Likert responses of section ii. An exploratory correlation matrix will also be built (cor-

---

[7]`www.aisv.it` Last visited February 4th, 2021
[8]`www.aisoitalia.org` Last visited February 4th, 2021
[9]`www.aiucd.it` Last visited February 4th, 2021

relation package: (Makowski et al., 2020); method set to "auto") with the aim of underlining significant patterns in the research routines of the respondents. Given the high number of tests (105), type I errors were controlled for with the Holm method; only the significant results will be reported and discussed here. Moreover, with the aim of proving that the involvement of researchers with diverse disciplinary backgrounds is not a mere exercise in academic demography, we run a second batch of point-biserial correlations between a dummy variable representing the expressed expertise in either oral history or linguistics (1 vs. 0, decided by coin flip) and each of the 15 series of numerical responses provided in the second thematic section of the questionnaire. This procedure will highlight the needs of the users with specific research interests as a reflection of their work routines. We chose to compare oral historians with linguists because of their representativeness in our pool of responses (see below); moreover, we may safely advance some preliminary hypothesis, such as the preference on the part of linguists for searching strategies by language or dialect. A Holm adjustment of p values was applied to this analytical batch as well. Lastly, a qualitative commentary on the answers to section iii will be discussed.

*Results: Demographics (i).* 32 female (57.1%) and 22 male (39.3%) respondents took part in our survey. Two participants did not specify their gender. Their age ranged from 27 to 74 (mean: 47.6, st. dev.: 13.7). Oral historians (22: 39.3%) and linguists (11: 19.6%) formed the most numerous groups of respondents. Sound and music engineers (6: 10.7%), ethnomusicologists (4: 7.1%) and anthropologists (3: 5.4%) followed with moderate numbers of participants. The remaining group of 10 interviewees was extremely fragmented, presenting one individual per category (sociologist, psychologist, archaeologist, mixed competences etc.), proving the importance of oral archives in several disciplines. Twenty-seven participants (48.2%) stated they had more than 10 years of experience in their respective field, followed by 19 (33.9%) with 0-5 years of experience and 10 (17.9%) with 6-10 years of experience.
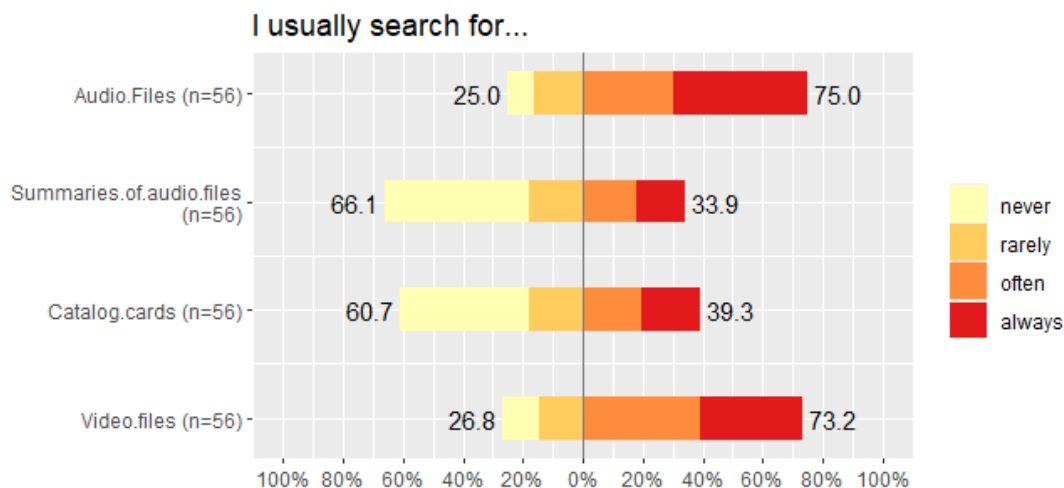


Figure 1: Bar plot of the perceived frequency of searching for specific document types

*Results: Research routines (ii).* Most of the respondents (42: 75%) refer to oral archives because of research and study activities. Eight individuals (14.3%) work for institutions pertaining to archives, while the remaining 6 made fragmentary references to the other categories outlined above. The distribution of the responses concerning the perceived frequency of use of online oral archives forms a left-skewed bell, with 13 (23.2%) "never", 24 (42.9%) "rarely", 13 (23.2%) "often" and 6 (10.7%) "always". The majority of our interviewees (27: 48.2%) work (or worked) on a single online oral archive, while 10 of them (17.9%) mentioned more than one resource, and 19 (33.9%) failed to specify the name of specific archives. Curiously, Youtube was referred to as an oral archive repository in two responses. Figures 1 and 2 show the rate of responses to the questions about the perceived frequency of searching for specific document type and using specific search strategies. The plots clearly show that our participants search for

audio/video files more frequently than they do for summaries or catalog cards. The preferred searching strategies are by keywords and topic. Other materials they searched for included transcriptions (3 occurrences), complementary materials, metadata, time stamps, preprocessed data for linguistic analysis, etc. Figure 3 illustrates the visualisation of the perceived usefulness of specific searching strategies.
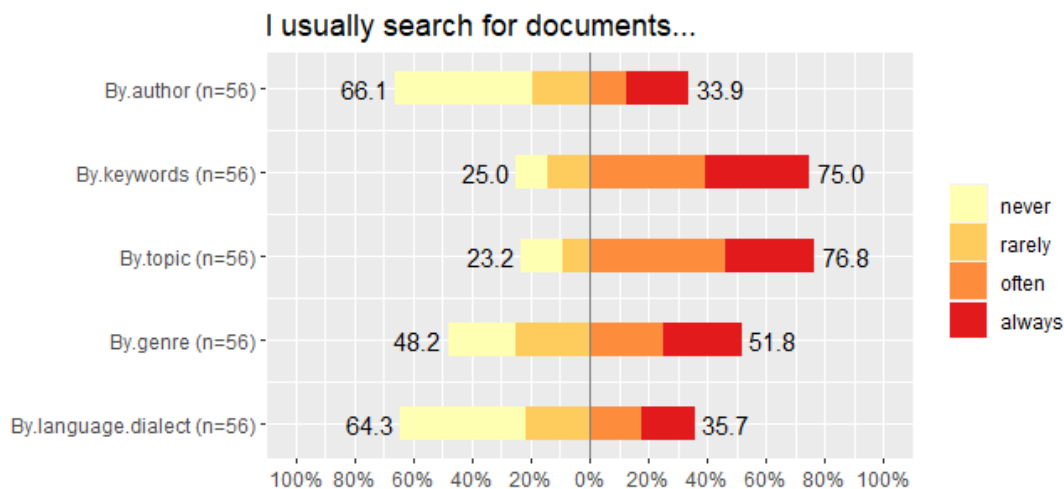


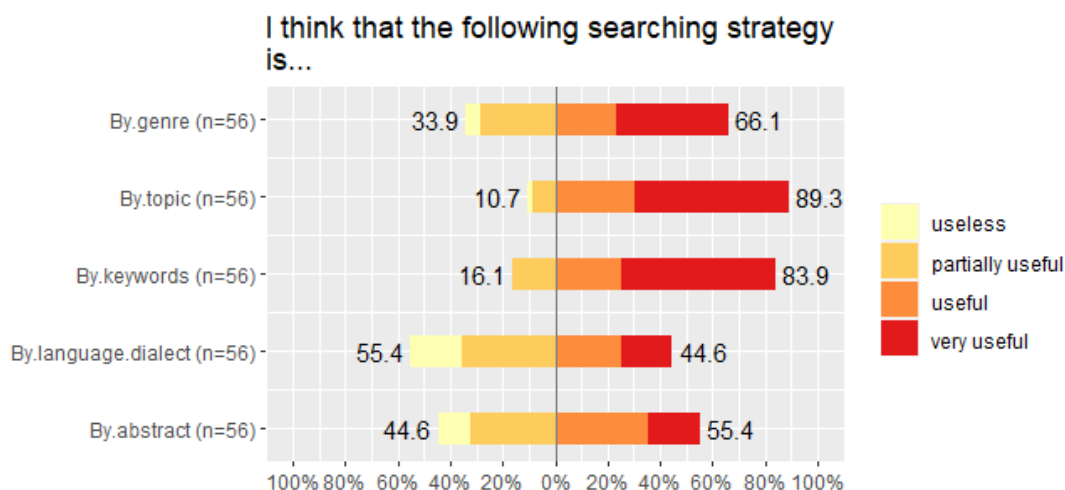Figure 2: Bar plot of the perceived frequency of specific searching strategies



Figure 3: Bar plot of the perceived usefulness of a specific searching strategy

The four categories shared by the plots in Figures 2 and 3 follow very similar distributions. The searching strategies which are perceived as very useful are by topic and keywords. Indeed, all the correlations between these four pairs of items are positive and significant (by keywords: $r(54) = .47$, $p = .026$; by topic: $r(54) = .5$, $p = .01$; by genre: $r(54) = .5$, $p = .01$; by language/dialect: $r(54) = .53$, $p = .003$). While these correlations are conceptually self-evident, they signal a certain level of internal coherence and that, presumably, the participants maintained a sufficient level of attention to the questionnaire. A number of searching strategies were also suggested by our interviewees, which focused mainly on the time and the place in which the original recording was performed. Other significant correlations unveil subtler nuances of data distribution. The search for oral materials by keywords goes hand in hand with the strategy by topic ($r(54) = .62$, $p = .001$), which may imply some level of conceptual or functional overlap between the two categories. By the same token, there seems to be some sort of connection between the search by language/dialect and by genre. The usefulness ratings of the two strategies are positively correlated ($r(54) = .51$, $p = .005$), and those who search by language/dialect more often also give higher usefulness

ratings to the strategy by genre (r(54) = .48, p = .016). The search for summaries of audio documents correlates with the use of keywords (r(54) = .5, p = .008) and topic (r(54) = .58, p < .001), stressing the need for an effective reduction to only a few focal concepts. Indeed, keywords are also used more frequently by those who search for video files more often (r(54) = .64, p < .001). Lastly, the reliance on written summarisations can also be noted at other levels of analysis. Overall, those who refer to online oral archives more often search for written summaries (r(54) = .49, p = .013) and catalog cards (r(54) = .46, p = .031) more frequently.

In our pool of responses, the language/dialect searching criterion lacked behind the other options both in terms of frequency of use and perceived usefulness. However, this population pattern may hide the preferences of specific subsets of respondents. A series of point-biserial correlations between the expressed expertise in oral history (22 observations) or linguistics (11 observations) and the numerical answers to the 15 questions presented in this section suggests that this is indeed the case. As we anticipated, the only correlations which survived the Holm adjustment concern the language/dialect searching criterion (frequency of use: r(31) = -.5, p = .045; perceived usefulness: r(31) = -.71, p < .001). Contrary to the general trend, these coefficients confirm the paramount importance of this strategy for the research routines of linguists. Overall, the presence of strong disciplinary specificity in the expressed preferences may suggest the need for the development of additional access strategies tailored around the research background of the user.

*Results: Suggestions (iii)*. Unfortunately, only less than half of the participants (26: 46.4%) submitted suggestions on the development of an online oral archive. In line with the results of the previous section, the most common recommendation (5 comments) concerned the development of a written counterpart to the audio documents. While a single participant asked for abstracts, four of them suggested the implementation of transcriptions, which should be aligned with the audio file and made searchable. Four participants were concerned with the quality of audio materials, which should be at least 24 bit and refreshable. A search engine by audio quality is also desirable. Others mentioned the need for supplementary materials, which could help in the search for relevant documents (e.g., by image). Lastly, other topics of discussion were granularity of metadata and terminology, cataloguing standards, accessibility, reusability and networking with other archives.

## 4 Building a Home for Audio Archives

### 4.1 Data and Metadata

In order to preserve and provide access to analogue audio recordings (e.g., the compact cassettes or the open-reel tapes), it is essential to digitise them. The result of the digitisation process is the digital *preservation copy*, which is composed of the audio content as well as other information about the carrier (such as the photo of the carrier itself, or its box)[10]. As the name suggests, the preservation copy is the "means" for safeguarding the content of the audio documents and it can be considered as the new digital master for long-term preservation. The preservation copy is only one element of the data workflow, since it is inextricably linked to the *archival unit*. In audio and audiovisual archives, it is defined as a set of data and documents pertaining to the very same communicative event, per unit of time and place. The archival unit is the outcome of a meticulous process involving listening, analysis and comparison. In the domain of oral archives, it is not infrequent that the content of a carrier needs to be re-organised. For example, an archival unit might be composed by content that is stored in several physical carriers (and, therefore, in several preservation copies), or vice versa, several archival units might be stored in the same physical carrier[11]. Given the absence of a one-to-one relationship between the physical carrier (i.e., compact-cassettes, open-reels) and the archival unit, the preservation copies are kept separately from the archival units (Mulè, 2003; Calamai et al., 2014; Stamuli, 2020).

---

[10]An extended description of the preservation copy is available in (Bressan and Canazza, 2013)

[11]Several other kinds of transformation could be performed, but their description goes beyond the scope of this paper.

This approach leads to a very complex set of metadata, organised along three different layers:

(i) metadata for the description of the preservation copy,

(ii) metadata for the description and managing of oral sources as items of an (audio) archive (archival unit),

(iii) metadata expressing the relationship between the preservation copy metadata and the digital archive metadata.

In *Archivio Vi.Vo.*, a customised set of metadata has been defined for (i), inspired by other international standards for audio material description, in particular the one proposed by the Association of Sound and Audiovisual Archives (IASA Technical Committee, 2009). The project adopted ISAD(G) and ISAAR standards for the archival units (ii), encoding the information about archival material with Encoded Archival Description (EAD) and Encoded Archival Context (EAC) standard data models. One of the main challenges is to make these metadata structures interoperable with the CLARIN VLO infrastructure component which is part of CLARIN's Component Metadata Infrastructure and can cope with many different metadata descriptions, as long as they are implemented through (or converted to) the Component Metadata framework. The metadata structure for expressing the relationship between the preservation copy and the archival unit (iii) is based on the methodology described below.

### 4.2 From Preservation Copies to Archival Units

The methodology formalised and adopted in *Archivio Vi.Vo.* is composed of several steps. All the operations performed during these steps and the information inserted by audio technicians, researchers and/or cataloguers are stored and duly described through a set of appropriate metadata, thus maintaining the relation between preservation copies and archival units. The methodology starts with the creation of the preservation copy of a carrier. This phase has proved to be very delicate and time-consuming.

Digitised audio recording is often barely accessible, due to, e.g., different speeds, configurations or digitisation errors (Pretto et al., 2020). In these cases, if necessary, researchers or audio technicians recur to *clip*[12] in order to separate parts with different speeds, channels with different recordings or recordings in different directions. In *Archivio Vi.Vo.* a *clip* is defined as a duplicate of an audio segment extracted from a preservation copy. One or more clips can be extracted from a preservation copy. In some cases, the clips are the result of a restoration operation, which is necessary for accessing the sound content. The process of creating (and restoring) the clips cannot modify the preservation copy. The resulting clips will be accessible and allow the single researcher/cataloguer to listen, analyse and describe their contents. If some parts of the very same clip belong to different events (and, therefore, to different archival units, see Section 4.1), they will be segmented accordingly and new sub-clips will be created (archival unit clips). In some cases, some archival unit clips, derived from different preservation copies, would be part of the same event, and therefore they will also be part of the same archival unit. During the analysis, the researcher/cataloguer may decide to discard some clips in case she/he believes the content is not related to the archive or not relevant for the users (e.g., several minutes of silence without recording due to empty tape at the end of the digitised recording). It is essential that these choices be formalised and preserved in the metadata maintaining the history of the documents. Since the questionnaire (see Section 3) outlines the need for written summarisation for some of the participants, the *Archivio Vi.Vo.* platform includes a complex abstract (lt. *regesto*) field. In the project, the abstract of an archival unit is divided into several segments related to the different parts that compose a single event. These segments are recognised and described by researchers and/or cataloguers during their analysis. Each segment is characterised by two temporal instants (the beginning and the end of the segment, respectively) and the description. The segments' length could be equal to or smaller than an archival unit clip (in the second case, the audio file will not be trimmed). As soon as all the archival unit clips are put into order, and all the missing metadata required by ISAD(G) is added, the archival unit will be created and available through the access interface.

---

[12]The concept of *clip* is commonly used to indicate data of either video or audio that has been clipped out (copied) from a larger carrier such as a reel or a video tape.

# 5 The Infrastructure

As for the infrastructure, the CLARIN-IT national data centre hosted in Pisa, ILC4CLARIN (Monachini and Frontini, 2016), will implement new experimental approaches to preservation, management and access to audio data and metadata. The experimental activity aims to adopt the model and the high-performance computing and archiving services of the new GARR network infrastructure, built along the Cloud paradigm[13]. The project will also exploit the federated identity service of the CLARIN infrastructure, in order to manage users' access. A robust system for managing authentication is essential for audio and audiovisual archives because of the frequent privacy, ownership, and copyright issues concerning their content (Kelli et al., 2019; Kelli et al., 2020). Several classes of users are taken into consideration, each of them with different access grants.

The infrastructure consists of two different parts. The first one provides a data and metadata entry interface for archivists, archive owners or, in general, researchers. The system is highly complex; it must be able to manage various international standards and several kinds of specific functionalities. Considering the complexity of the project, the infrastructure could hardly be developed from scratch. Therefore, as a first step, ten types of archival software were evaluated on the basis of several features and technologies (standards, programming languages, frameworks, DBMS, license, etc.). The software selected was the open-source platform xDams[14]. Three main characteristics influenced the adoption of the software: (i) the completeness of its coverage of standards, (ii) its extensible no-sql database as well as (iii) the open-source license. The second part of the infrastructure consists of an access interface that can support researchers of different disciplines in discovering and studying audio or audiovisual documents.

In order to study the interaction with the software, two mock-ups were developed for studying and testing the interfaces for inserting and cataloguing the digitised documents (Figure 4a) and for accessing their content (Figure 4b), respectively. The two mockups have been developed with the Vue.js and Bootstrap frameworks, respectively, Web Audio API, as well as Peak.js and Audiowaveform, two libraries developed by the BBC[15].
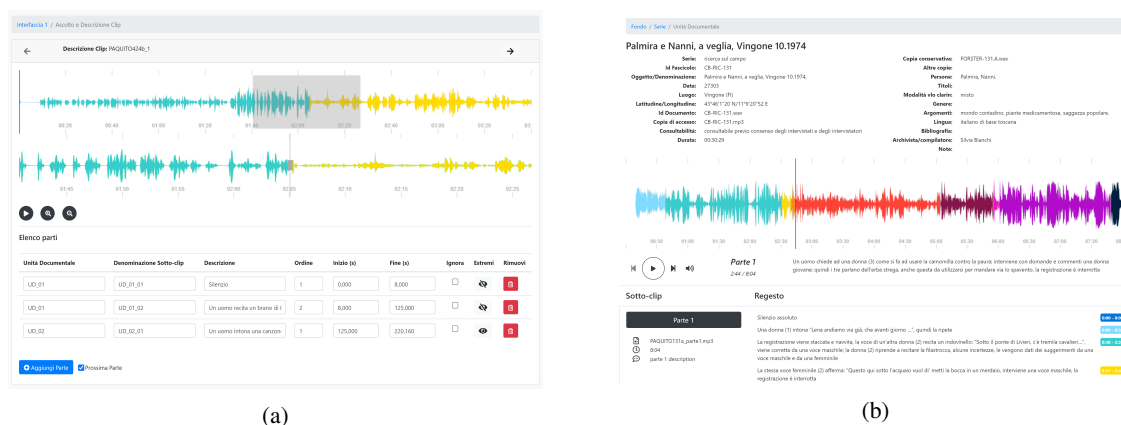


(a)    (b)

Figure 4: (a) section of the interface for creating archival units from preservation copies: the clips derived from the duplicate of the preservation copy are segmented, described, ordered and assigned to an archival unit; (b) the interface for accessing the archival units' contents.

The first mock-up interface is intended to provide the researchers and/or cataloguers with a web tool for listening to and describing the audio content of the clips extracted from the preservation copy. In addition, the user should be able to assign the described audio segments to different archival units, if necessary. The interface is composed of two main elements: the audio player and the section for the segmentation of the audio content. The audio player is complemented by the visualisation of the generated waveform using Peak.js and Web Audio API. The waveform is displayed both in its entirety and in a zoomed view,

---

[13]`cloud.garr.it` Last visited February 4th, 2021

[14]`www.xdams.org` Last visited February 4th, 2021

[15]`github.com/bbc/peaks.js` and `github.com/bbc/audiowaveform` Last visited February 4th, 2021

giving the user a very accurate time interval selection. In the section below the player, the data binding of Vue.js combined with the Segment API of Peak.js is exploited to allow the user to create, describe and sort the segments and assign them to the correct archival units.

The second mock-up interface aims to provide the user with an access tool that takes into the account the available contextual information in order to facilitate an adequate analysis of the archival unit. This interface can be divided into different areas: an upper area that displays the metadata of the archival unit, a middle area that contains the audio player with the waveform of the audio clip that is playing, and a bottom area containing the clips with their metadata, along with the abstract. In this latter area, the clips are grouped into Bootstrap card components. On the left side of the card, one can view the metadata and play/pause the clip. On the right side, there are the segments that compose each archival unit clip. The user can interact with this list by clicking on the different segments which are bound to the temporal instant of the audio clip. In order to facilitate user interaction with the interface, the segments are colour-coded with the time intervals to which they correspond in the waveform view. In this mockup interface, unlike the previous one, the waveform visualisation is created using a previously generated DAT file of the audio file in question (using Audiowaveform). This design choice was made because using Peak.js with Web Audio API requires downloading the entire audio file for waveform generation and this would result in a major slowdown in the web application. The use of pre-generated waveform files is very beneficial in our case since it is not necessary to load the complete audio files in the initial page loading. This allows the user to interact and view the waveform almost instantly (when the page is loaded, only the metadata of the audio files are loaded), taking advantage of audio streaming and thus avoiding long wait times due to the download of large audio files.

## 6 Final Remarks

*Archivio Vi.Vo.* constitutes a pilot case study within CLARIN-IT to experiment with methods for long-term preservation and secure access to oral archives and offer targeted services for both specialists and the general public interested in these data. *Archivio Vi.Vo.* aims not only to develop an infrastructure for preservation, description and use of audio archives, but, more ambitiously, to define and develop a model for the management, protection and enhancement of intangible archival heritage which can be replicated, even outside the context of Tuscany.

## References

Federica Bressan and Sergio Canazza. 2013. A systemic approach to the preservation of audio documents: Methodology and software tools. *Journal of Electrical and Computer Engineering*, pages 1–21.

Silvia Calamai and Francesca Biliotti. 2017. The Gra.fo project: from collection to dissemination. *Umanistica Digitale*, 1:85–103.

Silvia Calamai, Francesca Biliotti, and Pier Marco Bertinetto. 2014. Fuzzy archives. what kind of an object is the documental unit of oral archives? In Marinos Ioannides, Nadia Magnenat-Thalmann, Eleanor Fink, Roko Žarnić, Alex-Yianing Yen, and Ewald Quak, editors, *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*, pages 777–785, Cham. Springer International Publishing.

Thomas Fillon, Joséphine Simonnot, Marie-France Mifune, Stéphanie Khoury, Guillaume Pellerin, Maxime Le Coz, Estelle Amy de la Bretèque, David Doukhan, Dominique Fourer, Jean-Luc Rouas, Julien Pinquier, Julie Mauclair, and Claude Barras. 2014. Telemeta: An open-source web framework for ethnomusicological audio archives management and automatic analysis. In *Proceedings of the 1st International Workshop on Digital Libraries for Musicology*, pages 1–8.

Vincenzo Galatà and Silvia Calamai. 2019. Looking for hidden speech archives in Italian institutions. In *Selected papers from the CLARIN Annual Conference 2018, Pisa, 8-10 October 2018*, number 159, pages 46–55. Linköping University Electronic Press.

Pamela Giorgi, Fabiana Spinelli, and Serena Masolini. 2013. *Caterina Bueno: inventario del fondo documentario*. Firenze: Consiglio Regionale della Toscana.

IASA Technical Committee. 2009. *Guidelines in the Production and Preservation of Digital Audio Objects: standards, recommended practices, and strategies. IASA TC-04*. International Association of Sound and Audiovisual Archives, 2nd edition.

Aleksei Kelli, Krister Lindén, Kadri Vider, Pawel Kamocki, Ramunas Birštonas, Silvia Calamai, Penny Labropoulou, Maria Gavriilidou, Pavel Stranák, et al. 2019. Processing personal data without the consent of the data subject for the development and use of language resources. In *Selected papers from the CLARIN Annual Conference 2018, Pisa, Italy, 8-10 October 2018*, pages 72–82. Linköping University Electronic Press.

Aleksei Kelli, Arvi Tavast, Krister Lindén, Kadri Vider, Ramūnas Birštonas, Penny Labropoulou, Irene Kull, Gaabriel Tavits, Age Värv, Pavel Straňák, et al. 2020. The impact of copyright and personal data laws on the creation and use of models for language technologies. In *Selected Papers from the CLARIN Annual Conference 2019, Leipzig, Germany, 30 September, 2019*, pages 53–65. Linköping University Electronic Press.

Daniel Lüdecke, 2021. *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.7.

Dominique Makowski, Mattan S. Ben-Shachar, Indrajeet Patil, and Daniel Lüdecke. 2020. Methods and algorithms for correlation analysis in r. *Journal of Open Source Software*, 5(51):2306.

Monica Monachini and Francesca Frontini. 2016. CLARIN, l'infrastruttura europea delle risorse linguistiche per le scienze umane e sociali e il suo network italiano CLARIN-IT. *IJCoL - Italian Journal of Computational Linguistics*, 2(2):11–30.

Antonella Mulè. 2003. Le fonti orali in archivio. un approccio archivistico alle fonti orali. *Archivi per la storia*, 16(1):111–129.

Niccolò Pretto, Alessandro Russo, Federica Bressan, Valentina Burini, Antonio Rodà, and Sergio Canazza. 2020. Active preservation of analogue audio documents: A summary of the last seven years of digitization at csc. In *Proceedings of the 17th Sound and Music Computing Conference, SMC20, Torino*, pages 394–398, Torino.

R Core Team, 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Maria Francesca Stamuli. 2020. Fonti orali, documenti e archivi: riflessioni e proposte per la nascita di un 'archivio vivo'. In Duccio Piccardi, Fabio Ardolino, and Silvia Calamai, editors, *Studi AISV 6*, pages 95–109, Milano. Officinaventuno.