

Alessandro Russo, Niccolò Pretto, Antonio Rodà and Sergio Canazza

## L'informatica per la gestione e la conservazione di informazioni acustiche (musica e voce).

The lack of a standard preservation methodology is an important issue that digital speech archives have to face nowadays. Information that needs to be stored requires the implementation of a methodological framework in order to prevent cultural losses. The Centro di Sonologia Computazionale (CSC) of the University of Padova, cooperating with several digital archives (both speech and music) implemented a scientific methodology for the active preservation of audio documents. The framework is still a work in progress and the methodology is continuously improved with new additions, such as neural networks and machine learning-based techniques for detecting the right equalization curves or recognizing discontinuities on the tape. This contribution presents the abovementioned methodology, as well as several informatics tools for the study, the access and the preservation of audio documents.

Audio documents, Digitization, Preservation, Oral archives

### *1. Introduzione*

Negli ultimi anni è emersa in maniera sempre più evidente la necessità di individuare una metodologia condivisa per la digitalizzazione e la conservazione dei documenti sonori. Numerosi archivi, infatti, affrontano il problema adattando i metodi tradizionali utilizzati per la gestione del materiale cartaceo, i quali, seppure ormai ampiamente consolidati e idonei nel caso di beni librari, mostrano i loro limiti e la loro inadeguatezza nel caso di documenti complessi e di difficile lettura quali i documenti sonori. Infatti, oltre al segnale audio che costituisce l'informazione primaria, un documento sonoro è costituito da un insieme di numerosi elementi, che vanno dal supporto fisico a tutta una serie di informazioni aggiuntive (scatola, note dell'autore, segni sul nastro, ecc.) di grande importanza e che, in quanto tali, devono essere preservate per ottenere una copia digitale che rispecchi fedelmente l'originale.

Gli ultimi 20 anni sono stati interessati da una vera e propria "corsa alla digitalizzazione", in quanto, grazie al progresso tecnologico e all'introduzione di nuovi formati digitali, per la prima volta è stato possibile vedere nelle tecniche di salvaguardia digitali una soluzione efficace per la conservazione di svariate tipologie di beni culturali (IFLA/UNESCO, 2000). Questo era precedentemente impossibile a causa delle limitazioni offerte dalla tecnologia esistente, che

comportavano un'inevitabile e non trascurabile perdita di qualità nel passaggio da analogico a digitale, ragione per la quale la digitalizzazione era utilizzata solo ai fini dell'accesso, ma non della conservazione a lungo termine dei materiali (Smith et al., 1999).

Questo processo è nato dall'esigenza di mettere rapidamente in sicurezza e salvaguardare un patrimonio ad alto rischio, soprattutto a causa di problematiche inerenti al degrado dei supporti analogici, all'obsolescenza dei macchinari utilizzati per riprodurre i contenuti e a un'inevitabile perdita di conoscenze e competenze relative al loro utilizzo, sempre più ristrette a una piccola cerchia di appassionati e tecnici del settore (Davies, 2007).

In questo processo, qualora non si tenga conto dell'importanza del documento sonoro nella sua interezza, si rischia di andare incontro a una enorme perdita di informazioni. I laboratori del Centro di Sonologia Computazionale (CSC) del Dipartimento di Ingegneria dell'Informazione (DEI) dell'Università di Padova, si occupano da decenni di salvaguardia e digitalizzazione di documenti sonori di interesse storico, artistico e culturale, principalmente musica elettronica ed elettro-acustica, ma anche registrazioni di parlato ed etnomusicologia. Nella messa a punto di una metodologia per la conservazione dei documenti sonori, una grande attenzione è stata data alla raccolta dei metadati e ai supporti analogici, i quali, specialmente nel caso della musica elettronica, devono essere considerati parte integrante dell'opera stessa e come tali devono essere tutelati e conservati al pari dell'informazione audio registrata al loro interno. I documenti sonori si differenziano dai beni culturali realizzati attraverso altre forme di espressione artistica, quali per esempio la pittura e la scultura, dove l'oggetto di interesse non può essere separato dalla sua espressione fisica (Brandi, 1988). Questo processo, invece, può essere effettuato nel caso dei documenti sonori, estraendo l'informazione audio e riversandola su un altro supporto. Per questo motivo, spesso gli interventi di digitalizzazione non tengono conto del ruolo di grande importanza che rivestono i supporti fisici che, da un punto di vista strettamente musicologico, oltre a costituire un'importante testimonianza tecnologica dell'epoca, possono fornire informazioni importantissime per capire la genesi e il processo creativo che ha portato alla realizzazione delle diverse composizioni, soprattutto attraverso lo studio delle annotazioni che il compositore può aver preso direttamente sul supporto (frequente nel settore della musica elettronica) e sulla sua custodia. Anche quando si ha a che fare con registrazioni che contengono parlato, sul supporto o sulla custodia in cui è conservato il documento originale è spesso possibile trovare indicazioni guida inerenti alla configurazione dei canali e alla corretta velocità di lettura e curva di equalizzazione e, di conseguenza, rivestono un ruolo di fondamentale importanza. In numerosi casi, inoltre, i nastri magnetici sono copie uniche o possono contenere i materiali preparatori utilizzati per la composizione di una determinata opera ed è quindi possibile, attraverso il loro studio, estrarre

informazioni di grande interesse musicologico, analizzando per esempio i diversi segmenti di nastro giuntati insieme per realizzare l'opera finale.

Generalmente, durante la digitalizzazione di documenti sonori analogici contenenti registrazioni di parlato e/o musica tradizionale ed etnografica, ci si trova a fronteggiare problematiche inerenti a una minor qualità dei materiali utilizzati e delle registrazioni audio. Al contrario di quanto avveniva nel settore della musica elettro-acustica, infatti, spesso l'informazione sonora è registrata da non professionisti, con una strumentazione che deve necessariamente scendere a compromessi sulla qualità in favore della portabilità. Le registrazioni sul campo, inoltre, a prescindere che si tratti di cassette o di bobine aperte, non essendo effettuate in ambienti trattati acusticamente, sono spesso di qualità inferiore rispetto a quelle realizzate in studio di registrazione. I magnetofoni portatili, quali per esempio i popolari modelli a transistor G257 prodotti da Geloso, registravano a velocità più basse rispetto ai modelli da studio più costosi e questo, se da un lato permetteva una maggior durata del nastro, dall'altro comportava una rilevante perdita di qualità generale.

Questo patrimonio, preziosissimo per gli studi linguistici e dialettali, riveste una grande valenza storico-culturale e negli ultimi anni ha visto un graduale interessamento da parte delle biblioteche digitali internazionali. Data la loro complessità e la loro eterogeneità, gli archivi sonori richiedono la messa a punto di strategie mirate e metodologie di conservazione specifiche. In questo settore, lo sviluppo e l'utilizzo di *tool* informatici per la conservazione e l'accesso ai documenti sonori possono essere di grande aiuto per agevolare lo studioso nelle diverse fasi del processo di analisi del materiale digitalizzato.

In questo contributo viene presentata la metodologia di conservazione messa a punto dal CSC in tutte le sue fasi (*Sezione 2*) e alcuni strumenti informatici per la conservazione, l'accesso e l'analisi dei documenti sonori (*Sezione 3*).

## *2. Metodologia di conservazione*

Sebbene negli ultimi anni la comunità archivistica abbia aperto un grande dibattito sulle modalità di intervento più efficaci da adottare durante la digitalizzazione e la conservazione dei documenti sonori, la qualità dell'audio digitalizzato resta l'aspetto più importante e costituisce un punto fermo e universalmente condiviso. Dalla qualità della conversione A/D, infatti, dipende la continuità tra il documento originale e le copie digitali, che devono essere in grado di ricrearlo il più fedelmente possibile. In questa sezione, gli autori introdurranno la metodologia messa a punto all'interno del CSC sulla base della decennale esperienza di collaborazione con le realtà archivistiche italiane e internazionali, con particolare attenzione sull'importanza che rivestono i metadati e i supporti analogici nella digitalizzazione dei documenti sonori contenenti registrazioni di parlato e musica popolare.

## 2.1 Preparazione del supporto

La prima fase della metodologia proposta è incentrata sullo studio preliminare del documento sonoro e l'ottimizzazione del supporto, passi fondamentali per poter procedere con l'acquisizione del segnale audio. In questa fase è necessario procedere con una dettagliata documentazione fotografica, utile a testimoniare le condizioni del supporto, l'eventuale presenza di sintomi di degrado chimico-fisico in corso e a tenere una traccia digitale delle note dell'autore e delle informazioni aggiuntive presenti sulla scatola del documento. Eventuali documenti cartacei trovati all'interno delle scatole vengono sottoposti a scansione e a successiva archiviazione. Una prima ispezione visiva del supporto, a volte, può essere sufficiente per identificare l'insorgere di particolari sindromi di degrado, ma qualora la situazione fosse particolarmente di difficile lettura o fossero necessarie informazioni aggiuntive sui materiali e sulle loro condizioni, è possibile procedere con un'ispezione approfondita tramite tecniche di analisi chimica. Il supporto viene poi ottimizzato con appositi trattamenti, che vanno dalla "cottura" in forno nel caso dell'individuazione della *Sticky Shed-Soft Binder Syndrome SSS-SBS* (Bressan, Bertani, Furlan, Simionato, & Canazza, 2016) al rifacimento di giunte in cattive condizioni con adesivi specifici. Prima di procedere con la seconda fase, il nastro viene svolto più volte a bassa velocità per ovviare ai problemi derivanti da un cattivo avvolgimento [Figura 1] o a leggeri fenomeni di magnetizzazione locale.

*Figura 1 – Esempio di bobina soggetta a cattivo avvolgimento.*

## 2.2 Acquisizione del segnale

Una volta ottimizzato il supporto analogico è possibile procedere con l'acquisizione del segnale audio. La strumentazione, prima di proseguire con le operazioni di digitalizzazione, deve essere accuratamente calibrata e i parametri impostati correttamente, affinché non si introducano errori in fase di digitalizzazione. Il segnale audio viene convertito in digitale tramite l'utilizzo di convertitori A/D e software professionali per l'acquisizione e l'elaborazione audio. Eventuali dispositivi secondari, quali per esempio sistemi per la riduzione del rumore, possono essere introdotti in questo punto della catena tra il dispositivo di lettura e i convertitori A/D. La componente sonora viene digitalizzata secondo gli standard proposti dalla International Association of Sound and Audiovisual Archives, IASA (Bradley, 2009), così da ottenere copie digitali ad alta fedeltà. Durante l'intera acquisizione, il segnale audio viene monitorato attraverso un sistema di diffusori acustici per permettere all'operatore di annotare eventuali eventi significativi (rumori, disturbi, cambi di velocità, ecc.) che possono interessare la registrazione audio, facendo particolare attenzione a discernere tra quelli propri del documento originale e quelli che potrebbero essere accidentalmente introdotti dalla catena di digitalizzazione (rumori digitali e/o

analogici) o da errori dell'operatore. In particolare, i principali disturbi rilevabili durante la riproduzione di un supporto magnetico sono: (a) rumori locali quali *click*, *pop* e attenuazioni del segnale dovute a giunte in cattive condizioni o a porzioni di nastro degradate, (b) rumori estesi quali fenomeni di *hum*, rumori di fondo o distorsioni, (c) rumori e alterazioni del segnale prodotti da dispositivi elettronici durante la registrazione del nastro, quali distorsioni microfoniche, rumori elettrici e feedback e (d) degrado del segnale audio dovuto al malfunzionamento del sistema di registrazione (Brock-Nannestad, 2001). Nel caso di registrazioni di parlato, spesso effettuate da operatori non professionisti del settore, quali linguisti, antropologi ecc., sono molto frequenti i nastri che presentano più velocità di lettura, in quanto rallentare la velocità di registrazione permetteva, a scapito della qualità generale, di ottenere tempi di acquisizione più lunghi, spesso necessari nel caso di interviste, conferenze ecc., o nel caso in cui lo studioso non disponesse di nastri magnetici di riserva a portata di mano. Prima di procedere con l'acquisizione del segnale audio, è necessario individuare la velocità di lettura e la curva di equalizzazione corretta. A questo proposito, mentre la maggior parte dei magnetofoni da studio danno la possibilità di registrare alle differenti velocità di 9,5, 19 e 38 cm/s (in alcuni casi addirittura fino a 76 cm/s), i magnetofoni portatili spesso presentano velocità molto più basse, che scendono dai 4,75 cm/s ai 2,38 cm/s, spesso alternate nel corso dello stesso nastro. In molti di questi casi, anche la configurazione dei canali audio (*mono*, *dual mono*, stereo, quadrafonia, doppio stereo, ecc.) veniva sfruttata per registrare sullo stesso nastro il maggior numero di informazioni possibili e non è raro trovare documenti sonori nei quali il segnale audio è stato registrato con configurazioni, velocità e direzioni variabili all'interno dello stesso nastro e/o canale [Figura 2].

Figura 2 – Esempio di segnale audio registrato a velocità e direzione variabili.

Monitorare l'acquisizione del segnale audio durante l'intero svolgimento del nastro è, pertanto, fondamentale per creare copie conservative che rispecchino fedelmente il documento di origine. L'acquisizione del segnale e il monitoraggio audio possono essere accompagnati da una documentazione video che interessa l'intero svolgimento del nastro, così da permettere all'operatore di analizzare gli eventi significativi isolando i singoli fotogrammi e confrontando la componente sonora con quella visiva. Dopo ogni fase, i dati acquisiti vengono controllati e validati, prima di procedere con l'archiviazione degli stessi e del supporto originale.

### 2.3 Analisi dei dati, copie conservative e copie di accesso

L'ultima fase della metodologia è centrata sull'analisi dei metadati estratti e raccolti durante l'intero processo, fondamentali per completare il quadro generale del documento sonoro insieme alle informazioni contestuali, quali foto, video

ecc., utili per minimizzare la perdita di informazioni nel passaggio da analogico a digitale e ottenere copie conservative che rispecchino il più fedelmente possibile quello che è il documento sonoro d'origine. A questo proposito è fondamentale distinguere il concetto di copia conservativa da quello di copia d'accesso.

Le copie destinate all'archiviazione e alla conservazione contengono i file audio ad alta qualità digitalizzati a 96 kHz, 24 bit in formato non compresso (.wav) e senza alcuna elaborazione secondaria. Come definito da IASA (Bradley, 2009), la copia conservativa di un documento sonoro è "l'artefatto creato per essere archiviato e conservato". Da questa definizione si evince come l'utilizzo e l'accesso alla stessa si limiti a pochi casi eccezionali, essendo principalmente destinata alla conservazione. Il documento sonoro deve essere ricreato digitalmente nella sua interezza e, per questo motivo, le copie conservative devono contenere anche le informazioni contestuali, quali documentazione fotografica, documentazione video ecc. Per quanto riguarda i metadati, la metodologia messa a punto dal CSC prevede l'inclusione nelle copie conservative di schede riassuntive (.pdf) compilate grazie all'utilizzo di un software realizzato appositamente per l'acquisizione e la gestione dei metadati inerenti al documento di origine, sia dal punto di vista del supporto fisico (tipologia, marca, sintomi di degrado chimico/fisico in corso ecc.), sia da quello dei file digitali ottenuti nelle fasi dell'intero processo, quali audio, video e documentazione fotografica (formato di acquisizione, frequenza di campionamento, durata, dimensione del file, ecc.). Inoltre, per quanto riguarda i sintomi di degrado eventualmente riscontrati sul supporto di origine, le schede prevedono una valutazione da 0 a 5 delle condizioni dello stesso. Qualora fosse necessario procedere con elaborazioni quali processi di restauro digitale, filtraggi, enfasi di alcune frequenze ai fini estetici, o semplicemente fosse richiesto di accedere velocemente e facilmente ai contenuti, è possibile creare, partendo dalle copie conservative (che vengono archiviate come tali e non vengono pertanto modificate), corrispondenti copie destinate all'accesso. Queste copie sono generalmente realizzate utilizzando formati compressi (per es. mp3) di qualità e dimensioni inferiori e il loro utilizzo agevola la gestione e l'accesso ai dati da parte dello studioso. Qualora le finalità di utilizzo non richiedano la compressione dei file, possono essere realizzate delle copie d'accesso ad alta qualità da utilizzare, per esempio, per processare i segnali audio con algoritmi di restauro digitale finalizzati a migliorare l'intelligibilità dei fonemi nelle registrazioni di parlato.

### *3. Strumenti informatici*

L'ingegneria dell'informazione può essere di grande aiuto sia agli archivi digitali, sia allo studioso, attraverso l'utilizzo di strumenti informatici sviluppati *ad hoc* per l'accesso, la conservazione e l'analisi dei contenuti digitalizzati. Anche per quanto riguarda gli studi sulle registrazioni di parlato esistono diversi software molto utili per facilitare l'analisi del contenuto dei documenti sonori. In particolare, gli

studiosi possono avvalersi di strumenti dedicati alla trascrizione automatica del parlato, tra cui, per esempio, Nuance Dragon, alla trascrizione fonetica quali Kaldi (Povey et al., 2011) e i diversi *tool* disponibili all'interno del progetto CLARIN, ecc. o, semplicemente, alla gestione e l'inserimento dei dati con possibilità di visualizzare le forme d'onda delle registrazioni sonore, come ELAN (Polly & Lenard, 2006; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

Oltre a questi tradizionali ambiti di analisi, relativi al contenuto sonoro, le nuove metodologie e tecnologie informatiche possono essere applicate anche allo studio del supporto. Quest'ultimo coinvolge una serie di analisi, i cui risultati sono preziosi per supportare i ricercatori nel loro lavoro. In questa sezione vengono presentati alcuni *tool* per la conservazione e l'analisi dei documenti sonori, attualmente in fase di sviluppo presso i laboratori del CSC, che coinvolgono sia il segnale audio, sia le informazioni contestuali.

Il video del nastro che scorre sulle testine del magnetofono, incluso nelle copie conservative (vedi Sezione 2.3), è un'importante fonte di informazioni, che richiede però un'attenta analisi che necessita di molto tempo e attenzione. Gli eventi di interesse possono essere di varia natura: corrottele, segni sul nastro, giunte, ecc. Per ridurre i tempi di analisi e il carico di lavoro degli studiosi, è in fase di sviluppo uno strumento per il riconoscimento automatico di punti di interesse sul video e della loro classificazione basata rispettivamente su tecniche di *computer vision* e reti neurali (Pretto, Fantozzi, Micheloni, Burini, & Canazza, 2019). L'obiettivo ultimo è l'integrazione di questi algoritmi all'interno di un *tool* per l'analisi, nella quale i punti di interesse estratti automaticamente siano facilmente accessibili e integrabili con l'analisi dei contenuti [Figura 3].

*Figura 3 – Punto di discontinuità (giunta) sulla superficie del nastro riconosciuto e classificato automaticamente grazie alle tecniche di computer vision e reti neurali*

Tecniche di *machine learning* possono essere utilizzate per il riconoscimento automatico delle curve di equalizzazione (Micheloni, Pretto, & Canazza, 2017). Durante la registrazione del nastro magnetico, infatti, viene applicata una curva di pre-enfasi che deve essere compensata in fase di lettura applicando una curva di post-enfasi inversa per ottenere l'audio originale come era stato registrato. Riversare un nastro applicando una curva di equalizzazione non corretta comporta un'inevitabile alterazione del segnale, compromettendo l'integrità e la validità della copia conservativa. Sebbene nel corso degli anni siano state introdotte numerose curve di equalizzazione differenti, le più utilizzate sono la IEC/CCIR (Europa) e la NAB (America). Per questo motivo, sono stati condotti degli studi preliminari per il riconoscimento automatico di queste due curve di equalizzazione (Micheloni et al., 2017). Il *dataset* creato artificialmente era composto da campioni di silenzio e rumore bianco prima registrati in un nastro

magnetico e poi digitalizzati. La lettura è stata effettuata sia utilizzando uno standard di equalizzazione corretto, sia quello sbagliata. Utilizzando algoritmi di *clustering* e di classificazione, lo studio preliminare ha evidenziato la possibilità di discernere le curve di pre- e post-enfasi utilizzate. I risultati ottenuti sono stati incoraggianti e quindi ulteriori test su campioni reali sono stati effettuati (Pretto et al., 2018). Una volta realizzati i *dataset*, affinché i risultati potessero essere confrontabili con quelli ottenuti dall'analisi dei campioni di rumore bianco dei test preliminari, sono stati utilizzati i profili estratti da sezioni di silenzio all'interno del nastro. Sono stati così identificati tre principali profili di rumore: quello estratto da sezioni di silenzio nel mezzo di una registrazione (es. oratore che fa una pausa tra una parola e l'altra) (a), rumore dovuto alla testina di registrazione senza l'introduzione di uno specifico input sonoro (b) e profilo di rumore estratto da sezioni di nastro vergine (c). Dai risultati si è visto che i campioni appartenenti alle tipologie (a) e (b) possono essere utilizzate per identificare la curva di pre-enfasi. Grazie ai campioni appartenenti alla tipologia (c) è stato possibile identificare la curva di post-enfasi.

Gli autori, inoltre, hanno sviluppato algoritmi innovativi per il restauro audio (Canazza, 2012) utilizzando la tecnologia VST proprietaria di Steinberg, tra i quali CREAK (Canazza REstoration Audio Extended Kalman filter) (Canazza, De Poli, & Mian, 2010) per il restauro di segnali audio estratti dai dischi in gommalacca con cui venivano realizzati i primi 78 giri, CMSR (Canazza-Mian Suppression Rule), algoritmo dedicato al restauro dei segnali audio registrati su cilindri di cera e PAR (Perceptual Audio Restoration), un sistema di *de-hisser* che combina modelli psicoacustici con algoritmi funzionanti nel dominio della frequenza per il restauro di registrazioni su bobina e audiocassetta. Esistono, inoltre, numerosi software commerciali finalizzati al restauro audio digitale (Izotope RX, Adobe Audition, ecc.). Generalmente il restauro viene effettuato per ragioni estetiche, quali per esempio la rimozione di rumori indesiderati (pop, hum, hiss ecc.), attenuazione di frequenze risonanti ecc., per esempio al fine di realizzare una copia destinata all'uso commerciale del documento sonoro. Tuttavia, il restauro digitale può essere molto utile anche nel settore degli studi sul parlato, in particolare qualora si abbia a che fare con registrazioni che presentano problematiche relative alla difficile intelligibilità della voce registrata. In questi casi, infatti, le stesse tecniche di riconoscimento automatico dei fonemi possono mostrare i loro limiti e risultare inefficaci. In (Rodà, Russo, Canazza, & Così, 2018) gli autori hanno evidenziato come, nel caso di sorgenti sonore altamente compromesse e di difficile lettura, per esempio per l'eccessiva presenza di rumore di fondo, gli algoritmi di restauro audio possano essere utilizzati per migliorarne l'intelligibilità. In particolare, sono stati messi a confronto i risultati ottenuti filtrando campioni audio di difficile intelligibilità della durata di 8s con strumenti di restauro sviluppati nel settore della ricerca, tra i quali CREAK (Canazza REstoration Audio-extended Kalman Filter), CMSR (Canazza-Mian Suppression Rule) e PAR (Perceptual Audio Restoration) e alcuni algoritmi di



de-noise e de-hum inclusi nel software commerciale iZotope RX. Analizzando i campioni audio con lo strumento per il riconoscimento automatico dei fonemi Kaldi, prima e dopo il restauro digitale, è stato possibile osservare in quasi tutti i casi una diminuzione del PER (Phoneme Error Rate), con una riduzione massima del 9% nel caso dei file processati con i filtri percettivi PAR. In Tabella 1 è riportato un esempio delle variazioni del PER in un campione audio processato con i diversi algoritmi di restauro. Come è possibile osservare, fanno eccezione i risultati ottenuti applicando il filtro CREAK, per i quali non si sono osservate differenze sostanziali, in quanto questo algoritmo è maggiormente indicato nel caso di rumori impulsivi puntuali.

Tabella 1 – Esempio di variazione del PER in un campione audio prima e dopo il restauro digitale (DDD = Denoise+Dehum+Denoise, DD = Denoise+Dehum, EKF = ExtendedKalmanFilter, P = Perceptual, H = High, M = Medium, S = Small)

|      | <b>PER%</b> |
|------|-------------|
| B    | 68,42       |
| BDDD | 64,47       |
| BDD  | 61,84       |
| BEKF | 68,42       |
| BPH  | 64,47       |
| BPM  | 65,79       |
| BPS  | 65,79       |

Per quanto riguarda l'accesso, al giorno d'oggi è necessario che le copie digitali siano accessibili da qualsiasi luogo e in qualsiasi momento. Tuttavia, data la complessità del caso, l'accesso ai documenti sonori può risultare incompleto, soprattutto nel caso in cui vengano adoperati i comuni lettori per la riproduzione dei contenuti multimediali. Con il loro utilizzo infatti non è possibile effettuare operazioni quali cambi di velocità, configurazione dei canali, ecc. proprie della strumentazione analogica, né accedere a tutto l'insieme di metadati e contenuti aggiuntivi che costituiscono il documento sonoro nella sua interezza. Il CSC ha realizzato delle interfacce di accesso utilizzabili da postazione fissa (Pretto & Canazza, 2014) (Canazza, Fantozzi, & Pretto, 2015) e da dispositivo mobile, che virtualizzano le apparecchiature analogiche per la riproduzione dei documenti sonori (Fantozzi, Bressan, Pretto, & Canazza, 2017).

L'utilità delle interfacce è stata valutata da professionisti del settore musicale di differente età, esperienza e competenza. Le interfacce sono state accolte positivamente dai professionisti intervistati, a prescindere dalla loro età e livello di esperienza. Tramite interfacce scheumorfiche, le applicazioni ne ricreano fedelmente l'aspetto e le funzionalità, così da preservarne l'esperienza di utilizzo. Per ogni documento sonoro, oltre al contenuto audio, è possibile accedere alle

informazioni contestuali, quali la documentazione fotografica e i contenuti video e alla scheda riassuntiva su cui sono riportati i relativi metadati. Futuri sviluppi prevedono di includere nelle interfacce di accesso gli strumenti per l'analisi automatica della superficie del nastro, al fine di consentire all'utente di avere accesso alla visualizzazione istantanea degli eventi che emergono durante la riproduzione del documento sonoro.

#### 4. Conclusioni

La grande quantità di registrazioni di parlato e musica popolare costituisce un bene di grande importanza storico culturale e, in quanto tale, deve essere salvaguardato e preservato. In molti casi, quando si ha a che fare con questa tipologia di archivi, i documenti sonori sono conservati in piccole collezioni private che spesso non seguono le linee guida internazionali per la salvaguardia dei documenti sonori (AES, 1997). Infatti, parametri quali temperatura e umidità, qualora non controllati, possono fungere da catalizzatori per l'insorgere di numerosi sintomi di degrado chimico/fisico e ridurre le aspettative di vita dei documenti archiviati. In questo scenario, la digitalizzazione costituisce una strategia efficace per salvaguardare l'enorme quantità di informazioni a rischio a causa delle problematiche relative all'obsolescenza dei supporti analogici. La messa a punto di una metodologia efficace e condivisa è un passo necessario per poter effettuare interventi di salvaguardia mirati e specifici, minimizzando la perdita di informazioni. La metodologia proposta dal CSC risponde efficacemente alla necessità di creare copie conservative digitali fedeli ai documenti sonori nella loro interezza, dove l'informazione audio costituisce il principale centro di importanza, ma grande attenzione è data anche ai supporti fisici e alla loro ottimizzazione, alla raccolta dei metadati e alle informazioni contestuali, quali documentazione fotografica e video. Le metodologie informatiche possono essere estremamente utili, sia per la digitalizzazione e la conservazione degli archivi sonori, sia per gli studiosi del settore, velocizzando e automatizzando processi altrimenti molto dispendiosi dal punto di vista dei costi e dei tempi. Tecniche basate su *machine learning* per il riconoscimento automatico dei punti di discontinuità sul nastro e delle curve di equalizzazione, lo sviluppo di nuove interfacce di accesso e l'utilizzo di algoritmi di restauro per migliorare l'intelligibilità automatica dei fonemi, costituiscono, infatti, solo alcuni esempi di come le metodologie informatiche trovino applicazione nel settore dei documenti sonori.

#### Bibliografia

AES. (1997). *AES recommended practice for audio preservation and restoration -- Storage and handling -- Storage of polyester-base magnetic tape*. New York, NY: AES.

- Bradley, K. (2009). *LASA TC-04 Guidelines in the Production and Preservation of Digital Audio Objects: standards, recommended practices, and strategies: 2nd edition*. International Association of Sound and Audio Visual Archives.
- Brandi, C. (1988). Carta del Restauro 1972. *Teoría de la restauración*.  
[https://doi.org/10.1016/s0022-1139\(00\)00309-2](https://doi.org/10.1016/s0022-1139(00)00309-2)
- Bressan, F., Bertani, R., Furlan, C., Simionato, F., & Canazza, S. (2016). An ATR-FTIR and ESEM study on magnetic tapes for the assessment of the degradation of historical audio recordings. *Journal of Cultural Heritage*.  
<https://doi.org/10.1016/j.culher.2015.09.004>
- Brock-Nannestad, G. (2001). What are the sources of the noises we remove? *AES 20th International Conference, Archiving, Restoration, and New Methods of Recording*, 175–182. Budapest.
- Canazza, S. (2012). The digital curation of ethnic music audio archives: From preservation to restoration: Preserving a multicultural society. *International Journal on Digital Libraries*. <https://doi.org/10.1007/s00799-012-0088-x>
- Canazza, S., De Poli, G., & Mian, G. A. (2010). Restoration of audio documents by means of extended Kalman filter. *IEEE Transactions on Audio, Speech and Language Processing*. <https://doi.org/10.1109/TASL.2009.2030005>
- Canazza, S., Fantozzi, C., & Pretto, N. (2015). Accessing Tape Music Documents on Mobile Devices. *ACM Transactions on Multimedia Computing, Communications, and Applications*. <https://doi.org/10.1145/2808200>
- Davies, S. W. (2007). Knowledge: The Missing Element in Archiving and Restoration? *Audio Engineering Society Convention 122*.
- Fantozzi, C., Bressan, F., Pretto, N., & Canazza, S. (2017). Tape music archives: from preservation to access. *International Journal on Digital Libraries*, 18(3), 233–249.  
<https://doi.org/10.1007/s00799-017-0208-8>
- IFLA/UNESCO. (2000). *Safeguarding our Documentary Heritage / Conservation préventive du patrimoine documentaire*. Boston, Massachusetts, USA: UNESCO "Memory of the World" Programme, French Ministry of Culture and Communication.
- Micheloni, E., Pretto, N., & Canazza, S. (2017). A step toward AI tools for quality control and musicological analysis of digitized analogue recordings: Recognition of audio tape equalizations. *CEUR Workshop Proceedings*.
- Polly, L. R., & Lenard, Y. (2006). Elan. *The Modern Language Journal*.  
<https://doi.org/10.2307/327832>
- Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., ... Vesely,

- K. (2011). The kaldi speech recognition toolkit. *IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU*.
- Pretto, N., & Canazza, S. (2014). Rewind: simulazione di un'esperienza d'ascolto storicamente fedele di dischi fonografici digitalizzati. *Proceedings of XX CIM*.
- Pretto, N., Fantozzi, C., Micheloni, E., Burini, V., & Canazza, S. (2018). Computing Methodologies Supporting the Preservation of Electroacoustic Music from Analog Magnetic Tape. *Computer Music Journal*, 42(4).
- Rodà, A., Russo, A., Canazza, S., & Cosi, P. (2018). Audio documents restoration as a documentary source in the linguistic research comparison of instruments. In A. De Dominicis (A c. Di), "*Speech audio archives: preservation, restoration, annotation, aimed at supporting the linguistic analysis*". (pagg. 117–138). Roma: Accademia Nazionale dei Lincei.
- Smith, A., on Library, C., Resources., I., on Library, C., Resources., I., on Library, C., & Resources., I. (1999). *Why digitize? / by Abby Smith*. Council on Library and Information Resources Washington, D.C.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings Language Resources and Evaluation*. <https://doi.org/10.3758/BRM.41.3.591>

### *Affiliation*

Alessandro Russo – Dipartimento di Ingegneria dell'Informazione (DEI), Università di Padova.

[russoale@dei.unipd.it](mailto:russoale@dei.unipd.it)

Niccolò Pretto – Dipartimento di Ingegneria dell'Informazione (DEI), Università di Padova

[prettoni@dei.unipd.it](mailto:prettoni@dei.unipd.it)

Antonio Rodà - Dipartimento di Ingegneria dell'Informazione (DEI), Università di Padova

[roda@dei.unipd.it](mailto:roda@dei.unipd.it)

Sergio Canazza - Dipartimento di Ingegneria dell'Informazione (DEI),  
Università di Padova

[canazza@dei.unipd.it](mailto:canazza@dei.unipd.it)