PRESTO – Preservation Technologies for European Broadcast Archives

IST-1999-20013

Process analysis for audio

DOCUMENT IDENTIFIER PRESTO-W4-ACS-167b.05 Process analysis for audio

DATE 30/04/2002

ABSTRACT This document provides a detailed definition of the workflow for optimised audio archive preservation, also incorporating key-links developed by PRESTO.

AUTHOR, COMPANY Massimo Sarachino, Stefano Grego (ACS)

EDITOR ACS

WORKPACKAGE/TASK WP4, D4.2

DOCUMENT HISTORY

Release	Date	Reason of change	Status	Distribution
1.0	30/04/2002	First release	Issue	Public

Table of Contents

CHAPTER 1 INTRODUCTION 3
1.1 Purpose and Scope of the Specification 3
1.2 Applicable and Reference Documents 3
1.3 Glossary 3

CHAPTER 2 REFERENCE PRESERVATION PROCESS 7
2.1 Main Objectives 7
2.2 Composition and Transfer 9
2.2.1 Composition 9
2.2.2 Transfer 10
2.3 Massive Transcription 12
2.3.1 Transcription Set-Up 13
2.3.2 Digitisation 13
2.3.3 Validation 16
2.3.4 Metadata and Documentation 17
2.3.5 Digital Processing 17
2.3.6 Transcription Completion 18
2.4 Expert Transcription 19
2.5 Archive Update 20
2.5.1 Transcription Batch Completion 20
2.5.2 Digital Media Store 20
2.6 Off Line Digital Processing 22
2.6.1 Metadata and Documentation 22
2.6.2 Features Extraction 22
2.6.3 Digital-to-Digital Processing 22

CHAPTER 3 TECHNOLOGY AND COSTS 23
3.1 Equipment 23
3.2 Sizing 23
3.3 Costs 23

Figures and Tables

Figure 1: Reference Preservation Flow 7
Figure 2: Flow Control Station context 9
Figure 3: Transcription Station context 12
Figure 4: Transcription flow 14

Chapter 1 Introduction

1.1 Purpose and Scope of the Specification

The workflow for the preservation of audio materials, as documented in PRESTO-W2-BBC-001218, “Archive Preservation and Exploitation Requirements”, has been analysed to develop a reference transfer process that makes the overall process as efficient as possible. This work is a standard task in 'business process re-engineering', here applied to the technology and workflow of audio preservation.

This document provides a detailed definition of the workflow for optimised audio archive preservation, also incorporating key-links developed by PRESTO and specified in PRESTO-W3-ACS-167b.04, “Key Links Systems Specifications”.

1.2 Applicable and Reference Documents

[R1] Contract IST-1999-20013, PRESTO, Annex 1 - “Description of Work”

[R2] PRESTO-T11-JRS-20001006, “Quality Assurance Plan”

[R3] PRESTO-W2-BBC-001218, “Archive Preservation and Exploitation Requirements”

[R4] PRESTO-WP3-INA-001218, “Existing and emerging technology”

[R5] EBU Technical Review, Special Supplement August 1998, “EBU / SMPTE Task Force for Harmonized Standards for the Exchange of Programme Material as Bit streams, Final Report: Analyses and Results”

[R6] PRESTO-W3-ACS-167b.04, “Key Links Systems Specifications”

1.3 Glossary

The following list of specialized terms and acronyms concerning television production, post-production, broadcasting, telecommunications and computer industries, mainly derives from EBU Technical Review, Special Supplement August 1998, “EBU / SMPTE Task Force for Harmonized Standards for the Exchange of Programme Material as Bit streams, Final Report: Analyses and Results”.

Glossary	
-A-	
A / D	Analogue-to-digital conversion.

Glossary	
ADAT	Alesys Digital Audio Tape recorder.
Analogue signal	An (audio) signal, one of whose characteristic quantities follows continuously the variations of another physical quantity representing information.
Analogue transmission	A type of transmission in which a continuously variable signal encodes an infinite number of values for the information being sent (compare with "digital").
Asset	An Asset is any material that can be exploited by a broadcaster or service provider. An asset could therefore be a complete programme file, or it could be a part of a programme, individual sound, images etc.
-B-	
Backbone	The top level in a hierarchical network.
Broadcast	In general terms, a transmission sent simultaneously to more than one recipient. There is a version of broadcasting used on the Internet known as multicast. In multicast, each transmission is assigned its own Internet Protocol (IP) multicast address, allowing clients to filter incoming data for specific packets of interest.
Broadcaster	(Service Provider) An organization which assembles a sequence of events or programmes, based upon a schedule, to be delivered to the viewer.
Buffer	An area of storage that provides an uninterrupted flow of data between two computing devices.
BWF	Broadcast Wave File. The EBU has defined a file format, which contains the minimum information that is considered necessary for all broadcast applications. The basic information, together with the audio data, is organized as “Broadcast Wave Format” (BWF) files. From these files, using an object-oriented approach, a higher-level descriptor can be used to reference other files containing more complex sets of information, which can be assembled for the different specialized kinds of applications.
-C-	
CBR	Constant bit rate. A type of traffic that requires a continuous, specific amount of bandwidth (e.g. digital information such as video and digitised voice).
CD	Compact Disc, an audio media capable to store sounds in digital form using an optical coding on a polycarbonate disk.
Compression	The process of reducing the number of bits required to represent information, by removing redundancy. In the case of information content such as video and audio, it is usually necessary to extend this process by removing, in addition, any information that is not redundant but is considered less important. Compression techniques that are used include: blanking suppression, DPCM, sub-Nyquist sampling, transform coding, statistical coding, sub-band coding, vector coding, run length coding, variable length coding, fractal coding, and wavelet coding.
Content provider	A person or company delivering broadcast Content.
-D-	
D/A	Digital-to-analogue conversion.
DAB	Digital Audio Broadcasting. The new coming standard for radio that defines in a digital broadcasting channel not only audio, but also various formats of information.
DAT	Digital Audio Tape. The first consumer standard that makes possible recording audio in digital form. A long battle to protect music authors copyright has delayed the DAT market introduction causing its commercial flops.
DB	DataBase, a software component capable to store information’s providing tool for complex search operations.
DBS	Digital Broadcasting System.

Glossary	
Device	A unit of hardware, for example a videotape machine or a server.
Digital signal	A discretely-timed signal in which information is represented by a number of well defined discrete values that one of its characteristic quantities may take in time.
Digital transmission	The transmission of digital signals by means of a channel or channels that may assume, in time, any one of a defined set of discrete states.
DLT	Digital Linear Tape. A magnetic tape cartridge capable to store in the 7000 version 35 Gbytes of non-compressed data. This king of media is a standard de-facto in the multimedia archive based on robot system. DLT is the most compact media today available on the market and the one with the longest media life and number of readout.
DSP	Digital signal processor.
DVD	Digital Versatile (Video) Disk.
-E-	
-F-	
File	An organized collection of related records, accessible from a storage device via an assigned address. The relationship between the records and the file may be that of common purpose, format or data source, and the records may or may not be sequenced.
-G-	
-H-	
-I-	
-J-	
-K-	
Kbit/s	Kilobits per second. A digital transmission speed expressed in thousand of bits per second.
-L-	
LAN	Local Area Network. A network dispersed over a relatively limited area and connected by a communications link that enables each device on the network to interact with any other.
Link	Any physical connection on a network between two separate devices, such as an ATM switch and its associated end point or end station.
-M-	
Mbit/s	Megabits per second. A digital transmission speed expressed in millions of bits per second.
Metadata	Data describing other data.
-N-	
Network	In computing, a data communications system that interconnects a group of computers and associated devices at the same or different sites. In broadcasting, a collection of MPEG-2 Transport Stream multiplexes that are transmitted on a single delivery system, e.g. all the digital channels on a specific satellite or cable system.
NFS	Network File System. Communication standard between computers on a LAN introduced by Sun that allows a remote computer to see a storage unit of a remote computer as a local one.

Glossary	
-O-	
OPAC	On-line Public Access Catalogue
-P-	
PC	Personal Computer.
-Q-	
-R-	
RAID	Redundant Array of Independent Disks. A means of constructing a server by interconnecting several hard disk units such that the data is distributed across all of them. If individual hard disks fail, the remainder can continue working and the defective unit can be replaced, usually without taking the server out of service.
RAM	Random access memory. RAM is semiconductor-based memory within a personal computer or other hardware device that can be rapidly read from and written to by a computer's microprocessor or other devices. It does not generally retain information when the computer is turned off.
Resource	A unit of functionality provided by the host for use by a module. A resource defines a set of objects exchanged between the module and the host by which the module uses the resource. An example of a resource is a piece of static data, such as a dialog box, that can be used by more than one application or in more than one place within an application. Alternatively, it is any part of a computer or network, such as a disk drive, printer or memory, that can be used by a program or process.
-S-	
Server	A computer or other device connected to a network to provide a particular service (e.g. print server, fax server, play out server) to client devices connected to the network.
Service	A set of elementary streams offered to the user as a programme. They are related by a common synchronization. They are made of different data, e.g. video, audio, subtitles and other data. Alternatively, it is a sequence of programmes under the control of a broadcaster, which can be broadcast as part of a schedule.
Streaming	A collection of data sent over a data channel in a sequential fashion. The bytes are typically sent in small packets, which are reassembled into a contiguous stream of data. Alternatively, it is the process of sending such small packets of data.
-T-	
-U-	
-V-	
VBR	Variable Bit-Rate. A type of traffic that, when sent over a network is tolerant of delays and changes in the amount of bandwidth it is allocated (e.g. data applications).
-W-	
-X-	
-Y-	

Chapter 2 Reference Preservation Process

The *Reference Preservation Process* is implemented by a chain consisting of more steps:

- Composition: identify and assemble materials
- Digitisation: create a digital master copy plus low data rate version
- New media creation: create a new archive item
- Archive update (and update metadata)

At each step, a number of actions have to be performed both at media processing and at metadata processing level.

The main activities implied by the preservation chain are sketched in Figure 1 and.

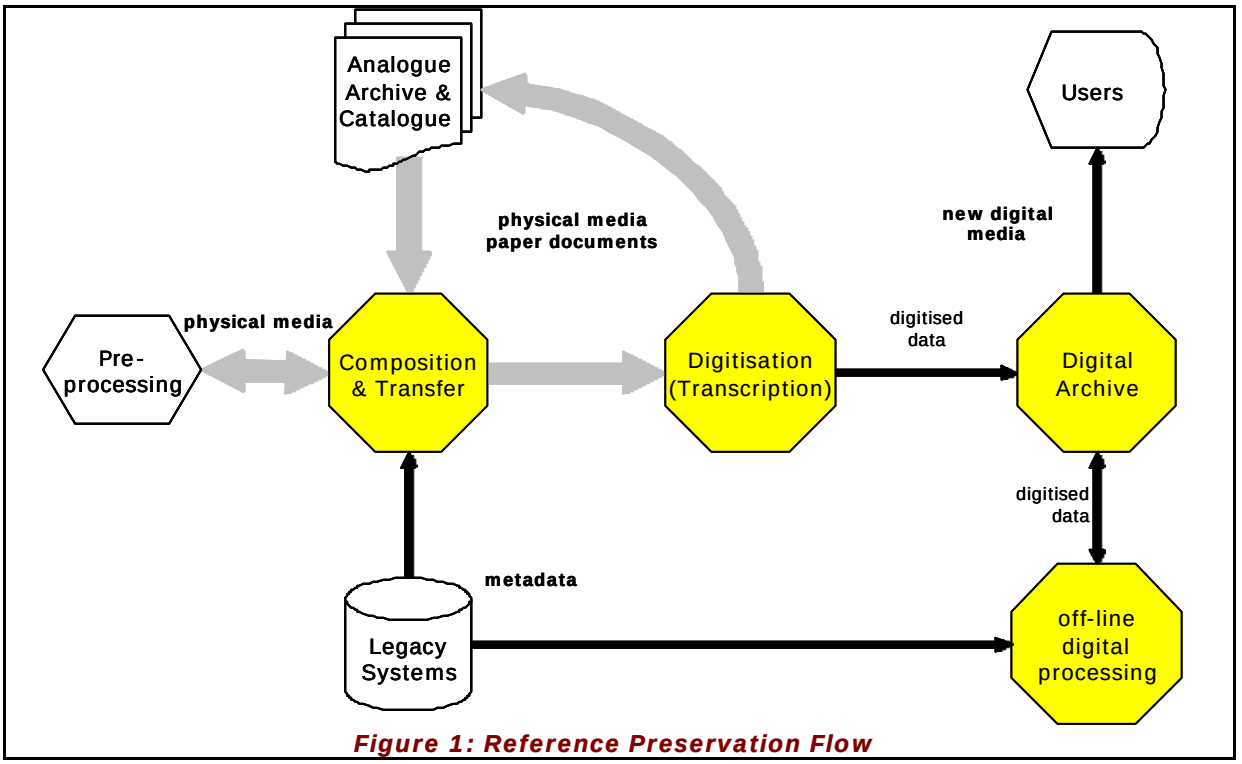

Figure 1: Reference Preservation Flow

2.1 Main Objectives

An optimal process for analog archive preservation through digitisation of audio media and paper documents should be aimed to reach more main targets.

Reduction of labour required. Work force is the main cost drive of the preservation process. Intensive human intervention is required during data collection, since analog archives are typically manually operated. Also quality verification is a huge time consuming task, since an operator should physically listen to the whole audio to ensure that transcription is performed properly, w.r.t. the correct behaviour of transcription chain and the maintenance status of analog media. The *Reference Preservation Process* dramatically reduces the latter effort, due its capability of automatic quality verification and reporting, thus allowing to free operator from listening each audio media for its whole duration.

Check of media and catalogue consistency. A typical problem of not automated archives is the possibility of mismatching between catalogue information and actual media and documents (e.g., media placed in the wrong shelf, empty boxes …). The *Composition* phase, where audio media and associated data are collected and prepared for conversion, is also used to perform this check, to reconstruct archive internal coherence.

Provide information for restoration. Preservation and restoration are better carried out by different processes. Restoration is generally limited to a subset of the media (and you don't know what, until digitisation is performed). Yet, thanks to its real time quality monitoring, the *Reference Preservation Process* allows to detect media needing restoring and provide precise input information to that activity.

2.2 Composition and Transfer

Composition and transfer concerns selection of audio data to be digitised and their transfer towards the Transcription (digitisation) Stations.

Strong procedures for configuration control should be enforced in this phase, possibly by means of a specific computed *Flow Control Station* system; the workflow of the operations required to handle materials back and from the archive should ensure that new media are correctly identified and that analogue material is correctly returned to the archive after processing.

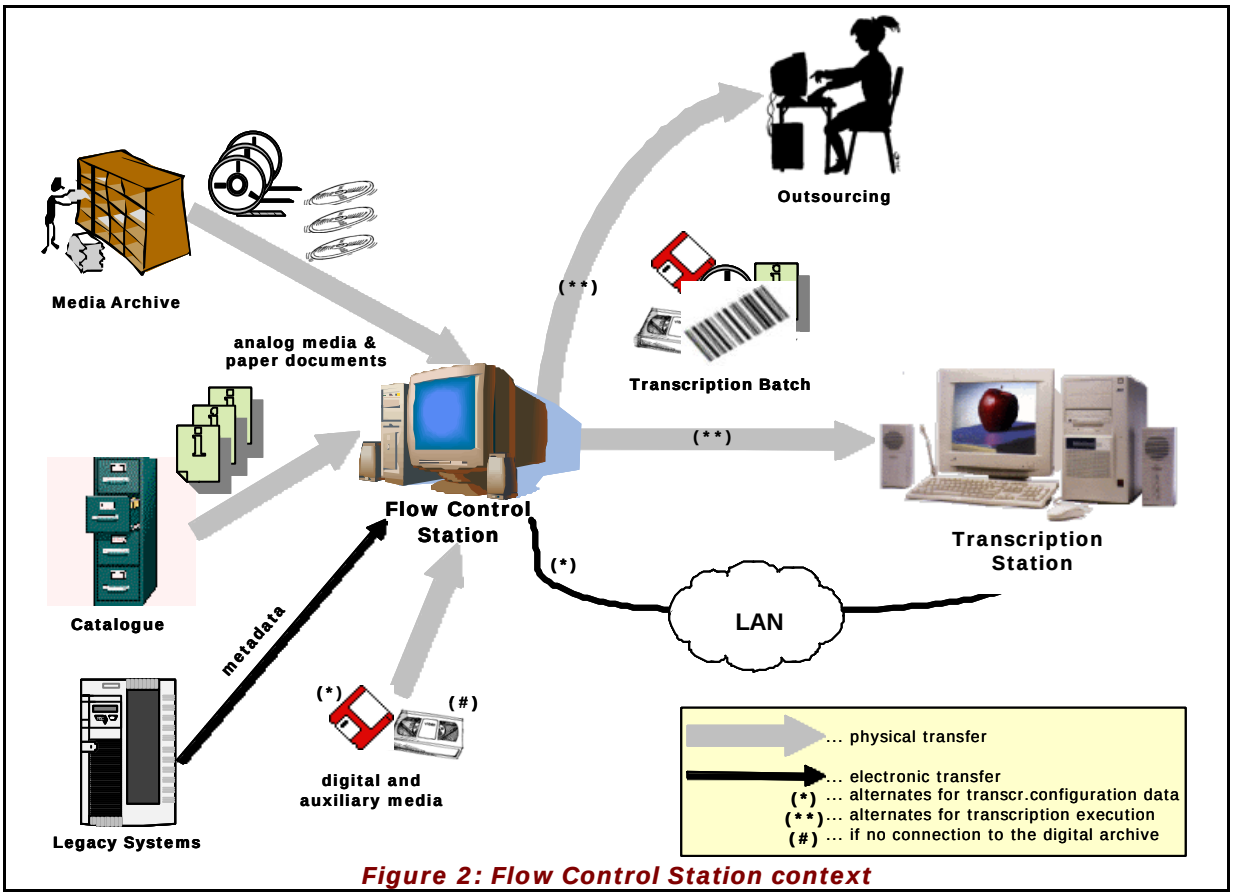

Figure 2: Flow Control Station context

2.2.1 Composition

Composition consists in collecting audio media and associated data and preparing for conversion. The archive operator has to:

- Identify material needing preservation; a database has to coordinate this activity, to avoid repeated transcriptions of the same audio and transcription errors. Policies

adopted for selecting media to be transcribed vary case by case, and can't be generalised.

- Gather basic item:
    - Extract metadata from existing documents;
    - Pick media from archive.
- Gather associated items. In most cases, when an audio archive has to be digitised, also a huge catalogue exists. The catalogue can have different levels of complexity and structures, ranging from collections of paper cards, only human readable, to electronic catalogues, based on database and full-text retrieval systems. All available information should be collected in this phase, to avoid repeating the effort of searching information in subsequent stages.
    - Get any available additional information;
    - Gather audio media and any physical item relevant to digitising.
- Check of media and catalogue consistency. This step allows fixing of mismatching between catalogue information and actual media and documents (e.g., media placed in the wrong shelf, empty boxes …).
    - Ensure that metadata are complete and accurate;
    - Ensure that media matches metadata.

2.2.2 Transfer

Transfer phase aims to setting up transcription sets of audio source analogue media, to be retrieved from their storage places and delivered to the *Transcription Station*. At transcription end, the audio sources will be returned to the analogue archive.

In order to ensure that material returns correctly to the archive after processing was successfully completed and that materials are not mixed up, all operations should be logged into a suitable database and materials to be digitised should be unambiguously identified. To this aim, all labels for old and new physical media are prepared: each media is bar code labelled in order to prevent media loosing or mixing.

More paths are possible to allow analogue contents flowing from analog to digital archive, e.g., depending from availability of network connection or depending whether the digitisation is performed by internal resources or externally (outsourcing).

The definition of a *Transcription Batch*, a set of material ready to be digitised (analog media coming from the archive and paper documentation), data cartridge to store the digitised data (if needed) and processing information, allows both to identify all physical media involved in the transcription and to provide a controlled workflow to material transfer. Following information should be associated to the Batch (and stored in the *Flow Control Station*), as a minimum:

- A unique Transcription Batch identifier, automatically associated to date and time
- For each audio analog source to be processed:

  - Media identifier in the analog archive
  - Type of media (tape, vinyl, …)

- Information concerning storage of digitised data: e.g., in the case data cartridge are used for temporary storage before data ingestion into digital archive, then their number, type (DLT, LTO, …) and intended usage (e.g., master copy, browse quality, streaming quality …) should be given.

At *Transcription Batch* definition completion, the *Flow Control Station* should print a number of labels containing Transcription Batch bar code, Media identifier in the analog archive (only for source media, in a human readable form) and (only for destination media) the intended usage. Archive operator will place labels on:

- Each analog source media
- Each digital destination media (if present)
- Each media box (source and destination).

In order to enforce as much as possible automatic checking and flow control, all the information gathered by the *Flow Control Station* should be transferred to the *Transcription Station*, as a *Transcription Configuration File* (TCF).

The Transcription Batch is then ready for dispatching. According to the specific broadcaster's layout of the preservation systems, following alternative paths are possible:

- Data are transferred by network;
- Data are written into one of the destination media;
- Data are written into a dedicated storage media (e.g., a floppy disc or equivalent). In this case, a bar code label should be printed also for this additional resource.

2.3 Massive Transcription

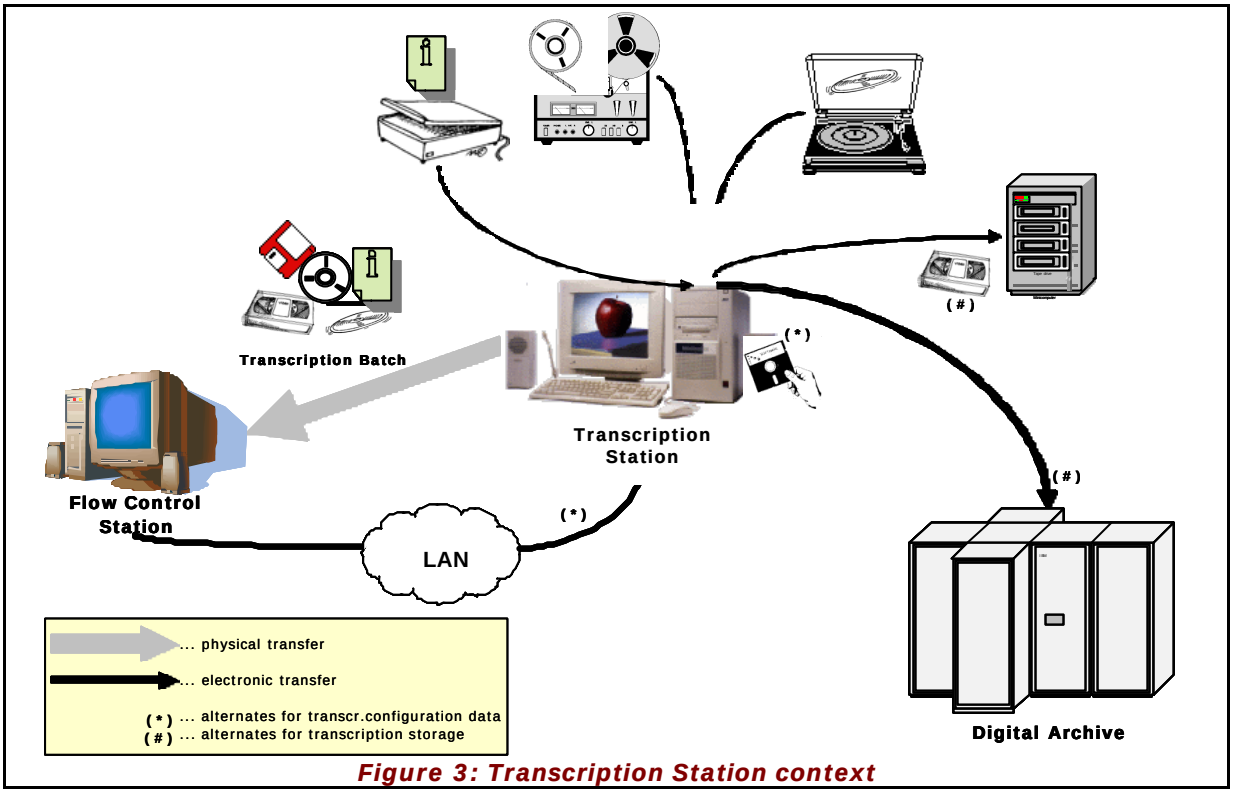

Figure 3: Transcription Station context

Solutions not targeted to massive processing provide an overall archive transcription duration estimated in 1.5÷2 times the whole of analogue media duration. This is too expensive for the most of cases. The duration of the process can be reduced by:

- Using computers with several audio input channel to perform concurrent recording
- Moving the signal restoration to another phase.
- Performing automatic detection of critical segments of the sound file

An effective and efficient transcription chain should be able to optimise costs of the process, duration, quality of the result and original materials safety. The massive transcription approach is aimed to maximize performances and is characterized by its capability of automatic metadata collection, especially concerning the quality of the performed transcription process, so that operator has to check only critical segments of already digitised sound files, considering critical saturated and mute segments. It allows for the parallel execution of playback and digitisation of more channels.

It is then possible to define two separate transcription approaches PRESTO-W2-BBC-001218, “Archive Preservation and Exploitation Requirements”: one specially optimised in terms of speed and cost reduction (*Massive Transcription*), and a second involving a high degree of specific expertise for endangered valuable materials that privileges accuracy against speed (*Expert Transcription*).

According to the most usual preservation status of considered archives, the massive transcription system should allow the processing of most of the audio items, operated by

low-level staff, or even committed to third parties. Only a few items require expert intervention.

The following paragraphs describe the massive approach.

2.3.1 Transcription Set-Up

A new transcription starts when a *Transcription Batch* is available at the *Transcription Station*.

By using the station bar code reader, the operator should enter the *Transcription Batch* identifier, and then the *Transcription Station* should process the *Transcription Configuration File* (TCF). According to the overall layout, more scenarios can be envisaged:

- Data were transferred by network: the *Transcription Batch* identifier is used to retrieve the proper *TCF* file from the temporary storage area;

- Data were written either into one of the destination media or into a dedicated storage media (e.g., a floppy disc): the *Transcription Station* bar code reader is used to check media label, then the media in loaded into the proper reader and data are retrieved.

Once the *TCF* has been read, the *Transcription Station* bar code reader is used to check the labels of all media involved in the transcription.

Also playback and recording equipment should be checked when a new transcription is required, for availability, readiness and, as far as possible, for calibration verification (e.g., trimming and conditioning of the playback devices to follow the media peculiarities, like azimuth alignment, selection of recording format -mono, stereo- for tape players and pickup type, rotating speed, equalization for turntables). Also the regular execution of cleaning procedure for analog and digital tape drives should be checked before starting the transcription and enforced by the Station controller.

To support the on-demand approach, the *Reference Preservation Chain* allows authorised users to set-up the transcription without the need of a TCF (*Transcription Configuration File*) coming from the *Flow Control Station*. Relevant information should be directly typed at the *Transcription Station* console.

2.3.2 Digitisation

2.3.2.1 Playback

Vinyl material is generally not in critical condition, with exception for older 78 RPM. Even so, media must be inspected before transcription, and, if needed, mechanically treated in order to:

- Remove dirt, dust and grease from the surface

- Discharge electrostatic charges
- In some cases, to flatten the disc surfaces if deformed (for instance, after years of incorrect off-vertical positioning or exposure to heat).

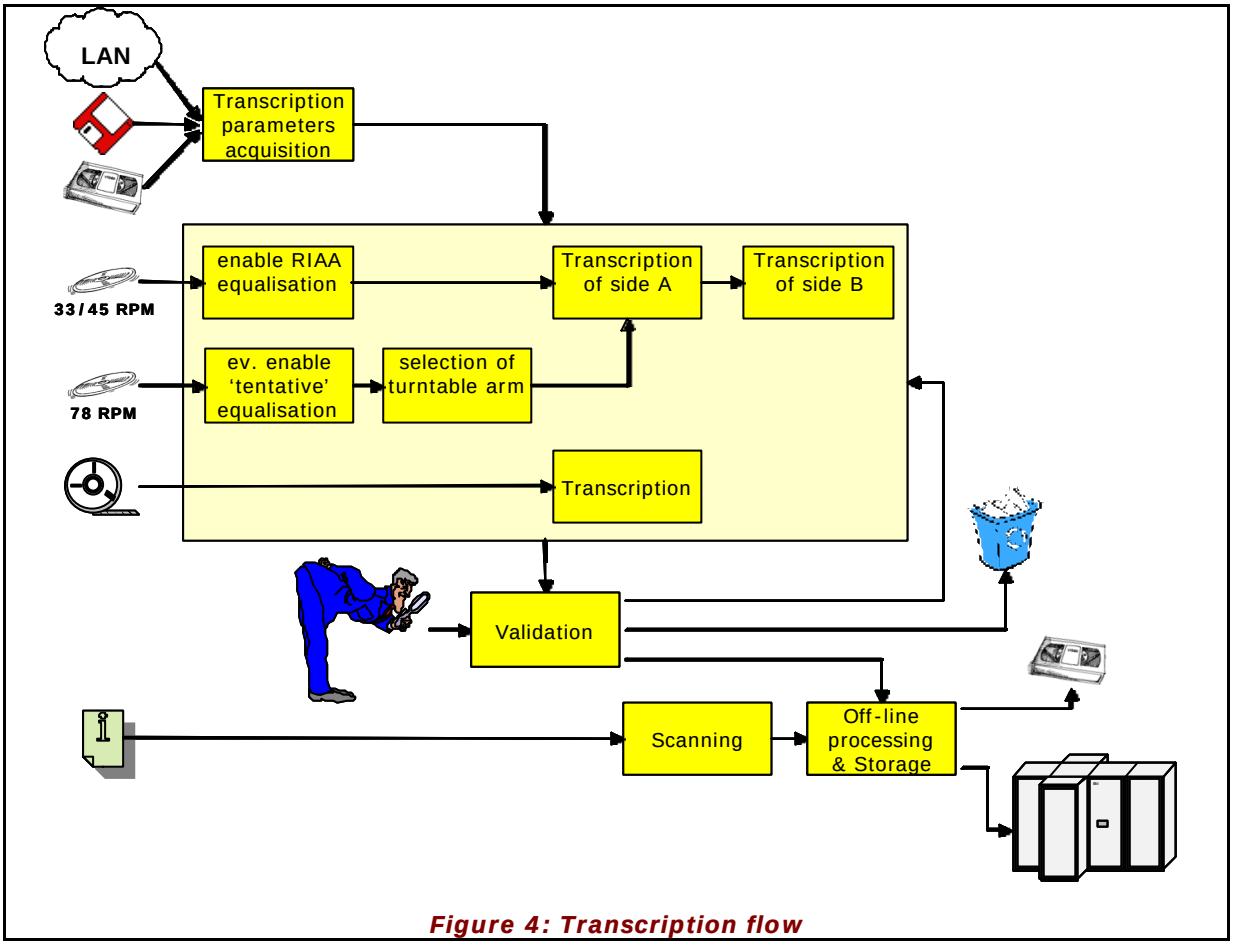

Figure 4: Transcription flow

After set-up and conditioning, operator can start transcription. This step requires some human intervention, since the operator is asked to:

- Select a single source media to be digitised, and type the corresponding identifier at the *Transcription Station*;
- Take the media and physically mount it in the proper driver;
- Place the media and/or reader at audio beginning (by playing back and forth tapes or by placing arm turntable just at the initial groove for vinyl's);
- Start recording and then reproduction (if the playback device is not controlled by the recording equipment).

Playback devices used in the massive transfer chains should be optimised for maximum speed of operation of the overall system. The audio *Reference Model* assumes the availability of devices remotely controlled (through a serial interface) for tape and vinyl playback.

While remotely controlled tape recorders are available, a specific development was required in PRESTO to achieve remote control of turntable (start/stop) and arm (up/down) PRESTO-W3-ACS-167b.04, “Key Links Systems Specifications”. A two arms configuration has also been adopted, to improve the efficiency of 78 RPM records transcription chains.

78 RPM records were produced in absence of standard, so that pick up geometry, reproduction speed and equalization applied can vary and information on recording parameters is inadequate. While speed and equalisation can be corrected by off-line digital processing, errors in pickup geometry result in the loss of signal details that cannot be recovered. Even if the most suitable stylus size for an LP can be determined from the production period (so that items to be digitised can be grouped into transcription batches accordingly), the selection between conical truncated and elliptical truncated stylus depends on the record surface conditions: it is convenient to fit two tone arms on a single turntable, so that two styli configurations can be tried on the fly, without any reconditioning of the chain between the transcription of an item and the next.

The *Transcription Station* should allow for parallel acquisition from both arms, deferring the selection of the best record to a subsequent step.

The massive approach doesn’t consider the transcription operator have time for detecting neither the proper 78 RPM equalisation nor the exact reproduction speed. Moreover, some music expertise is required for those jobs, while the transcription operator should have a more technical profile. If needed, the transcription operator should be enabled to select (among a limited number of possibilities) a ‘tentative’ equalisation to be applied only to the browse quality, just to allow a preliminary fruition of the digitised audio.

2.3.2.2 Digitisation and Real-Time Quality Monitoring

The digitisation step is aimed to converting the existing analogue signal into a digital representation. It also applies to digital-to-digital transfer, e.g., to migrate from obsolescent formats or to provide a format more suitable for automated handling. A major aim of audio preservation from analogue play back to digital recording is to do the transfer keeping the original quality and ensuring that new recording is free of errors due to the digitisation. An optimal quality controlled transfer is required.

Digitisation and restoring are carried out in different processes. While digitisation concerns the whole archive, restoring is limited to a subset of the produced digital media (and you don’t know exactly what, until digitisation is performed).

Detection of audio portions needing corrective intervention is normally performed by listening to the recorded audio. A dramatic improvement can be achieved by adopting automatic detection of defective intervals; metadata (as produced by specific QC software) are collected during digitisation process. So digitisation provides information to restoring execution.

The key to a cost-effective preservation is to provide as much as possible automatic detection (and visual presentation) of the level of quality for each single media during its digitisation, without the need of operator listening to audio for the whole duration of media: this way, the operator is able to execute more transcriptions at the same time: real time listening can be reduced to audio samples, only when the graphical interface signals an anomaly for the specific media. The importance of introducing a reliable quality monitoring module into the digitisation chain is that for each false alarm, the expensive

restoring chain is activated with no real need; for each false rejection, a corrupted signal not corresponding to the contents of the original media is erroneously added to the digital archive, with no chance to realize the damage until the digital media file will be played in the future.

During digitisation, following information should be made available to the operator by quality control:

- *The behaviour of the transcription chain, errors due to the playback machine*. Mainly for what concerning analogue devices, housekeeping data have to be collected, to monitor their correct functioning during playback.
- *Errors due to original medium*. Detection of clicks, crackles, hum, dropouts, splices, distortion…
- *Punctual parameters*, like energy, bandwidth, saturation, peak signal level, silence duration, noise duration, phase correlation;
- *Global quality of the captured audio*. Average signal to noise ratio, mean levels, max peak etc…

2.3.2.3 Digitisation Completion

A transcription always starts by operator command, while it can be terminated either manually or automatically. The automatic stop criterion is given by the detection of a long (configurable) silence. In case of automatic stop, the Transcription Station will always ask confirmation to the operator.

On the basis of parameters computed by the real time quality monitoring, a sort of go/no go response should be provided by the *Reference Preservation Chain*, informing the user whether restoration should be applied to the media.

2.3.3 Validation

Based on real-time quality monitoring results and sample listening performed during digitalisation, the transcription operator can be in charge of providing a first level validation. Depending on the single broadcaster's process requirements, the operator can be asked to provide one of the following:

a) A simple evaluation concerning the transcription execution (in other word, whether the transcription chain worked at its best, independently from the quality of the digitised signal). In case of a negative evaluation, the digitisation should be repeated after proper conditioning of the chain.

b) Some evaluation about the satisfactory quality of the resulting digitised media and whether restoration applied to the analog source could lead to better results.

c) Depending on the adopted policies, the operator could also be authorised to reject the transcription of the single media, so that no further processing is performed and resulting digitised data are not sent to the digital archive.

d) In the case of 78 RPM vinyl's, if both arms have been used, the operator could also select which channel is the "master" (i.e., which is the source for further digital processing).

The *Reference Preservation Chain* includes the a) and b) possibilities as the baseline for the *massive transcription* approach: actually, the most valuable part of the transcription process is given by the automatic support to the detection of audio intervals needing to be restored. Remaining decisions are reserved to the ingestion into digital archive stage, in order to achieve high volumes in transcription and to avoid that critical decisions are taken by low-level staff.

2.3.4 Metadata and Documentation

This step concerns activities related to on-line metadata preparation:

- Technical metadata: media with embedded metadata (timecode is the main example)
- Preservation metadata, concerning information about the original material format and characteristics, and the preservation (digitisation / conversion) process, i.e., information for understanding the new 'artefact' and for any further processing (e.g., signal restoration).
- Documentation, mainly related to the media contents description, including media original box/ folders or additional enhanced metadata (indexing …). A transcription session generally includes acquisition via optical scanner of paper documents associated to the audio data (e.g., covers and booklets for LP). Once acquired, they are added to the transcription files.

2.3.5 Digital Processing

While the digitisation stage, requiring the availability of the source media and human operator attendance, is an 'on-line' phase, the processing of digitised media can be performed in a subsequent time, as an 'off-line' phase. Some of the activities can be executed automatically (e.g., lossless compression and multiple formats generation), while other require the intervention of staff with specific skills, differing from what required for digitisation (artistic metadata and features extraction).

The *Reference Preservation Chain* assumes the transcription process only including lossless compression and multiple formats generation (if required). This means digitised master data are stored locally to the Transcription Station only until such processing is not completed; any other processing is performed by accessing the digital archive.

2.3.5.1 Multiple Formats Generation

If required, coding of associated versions (master, browse …) is automatically performed, according to the broadcaster's needs.

2.3.5.2 Lossless compression for audio

If required, lossless compression of the master copy is automatically performed.

An ad-hoc development has been performed in PRESTO to develop a FLAC compliant (http://flac.sourceforge.net/) lossless compression utility to be applied to 24bit, 48KHz BWF files PRESTO-W3-ACS-167b.04, “Key Links Systems Specifications”.

2.3.6 Transcription Completion

For each source analog media, the *Transcription Station* should provide the following output data files:

- Digitised master copy (24bit, 48KHz) as a BWF (Broadcast Wave Format) file (after lossless compression, if applicable);
- A number of audio files for the compressed formats (mp3 …);
- A number of image files for scanned documents;
- An .xml file containing housekeeping metadata;
- An .xml file containing the quality analysis log and summary results

The housekeeping file should include technical metadata related to the transcription process (date and time, identifiers for transcription operator and used transcription physical chain …). It should also include the list of all the files created by the transcription process.

Depending on the overall architecture, the files could be either available on file system, or stored into digital tape cartridge, or directly sent to the digital archive. If digital media have been produced, their identifiers should be included among housekeeping data.

When the transcription batch has been completely processed, a further file should be produced, to summarize results for all the source analog media belonging to the batch.

If data have to be stored on tape cartridge (e.g., DLT, LTO …), tapes should be written now. In such a case, it is strongly recommended that two copies of the digitised data are produced, at least of the audio digital master.

Then, the *Transcription Batch* is sent back to the physical archive, and digital data are made available for long-term storage.

2.4 Expert Transcription

Expert human intervention is required for some part of the archive, concerning fragile and/or valuable materials. Technical metadata extraction needs in this case to be supervised by experienced operators, listening to the whole media. Therefore, the resulting productivity is substantially lower. While the massive approach allows a single operator to transcribe 5÷8 analogue media at the time, an expert approach (without the help of on-line quality monitoring) only allows 1 analog media at the time to be processed.

Broadcasters should deploy transcription systems aimed at ensuring that such media are handled carefully and safely. Expert transcription could be performed either on demand or based on a continuous processing, depending on the total amount of items that need to be processed with this chain and by the availability of a priori knowledge about the media conditions.

The main technical differences between massive and expert approaches reside in workflow and audio manipulation tools. Workflow constraints need to be released, since the operator should be able to repeat material conditioning and digitisation until a satisfactory result is reached (unless restoration is required). Whilst the level of analysis offered by the *Transcription Station* on-line and off-line quality monitoring is almost as complete and accurate as required, additional software tools should allow the operator to manipulate digitised audio in order to remove the most common defects.

2.5 Archive Update

2.5.1 Transcription Batch Completion

After transcription completion, a *Transcription Batch* is returned to the physical archive.

The archive operator should check that all media are returned without damages and no item has been lost.

The *Flow Control Station* should support this verification by allowing the operator to read bar code labels of the returned media and checking the correspondence with the original batch materials. The operator is also required to verify that source media are actually present in their boxes and no material exchange occurred.

The transcription housekeeping data should be used to verify that all produced files are stored in the digital media.

The reservation information of the processed material should be updated according to the transcription results: analog media rejected or signalled for restoration should be properly marked for future processing (maybe, through the 'expert transcription' chain).

2.5.2 Digital Media Store

Digital media store consists of two main branches:

- Raw temporary storage, hosting data and metadata after digitisation, waiting for validation.

- Long-term digital media storage. On its own, this storage provides two main facilities: a Catalogue, storing metadata, accessed by users to select media to retrieve, and an Archive, actually storing the digitised media. The user access to the catalogue can be improved by hosting there also previews of the archived material, i.e., audio samples produced with a very low quality.

Upon satisfactory validation, media and metadata are ingested into the long -term storage system (actually, moved from raw storage). They can also be simply removed from the system if the validation fails.

End-users access is restricted to the long-term archive; they can access the digital media only after entries have been inserted into the Catalogue, through an OPAC facility.

On the other hand, local media and metadata management functions allow access at any time when required by preservation staff (technicians performing digitisation, cataloguing staff and others) and processes, to the proper set of data (either raw or long term).

The physical architecture of the digital media store depends on the size of the archive and expected performance. It can be based on some HMS (Hierarchical Storage Management) to manage automatically and transparently customer data across a storage hierarchy,

consisting of high performance disk, compressed disk, and tape libraries. Robotic libraries can also be used to fully automate media access.

The raw and the long-term stores can be either different physical archives or logical areas of the same storage system.

2.6 Off Line Digital Processing

2.6.1 Metadata and Documentation

Metadata can be extracted from existing documents (legacy system) or typed by the cataloguing staff.

When starting the preservation process, a huge amount of metadata is typically available, already created and entered by the archive or cataloguing staff when the original analogue material entered the archive. This information can't be lost or manually replicated in order to be present as metadata in the new digital archive.

In many cases organisations that want to introduce a content management system already have an existing documentation system or catalogue installed that is widely used within the organisation. Often these systems store metadata describing a large part or even all of the assets of the organisation. They need to be interfaced and often even be treated as the master database. In the case metadata are retrieved from external legacy systems, some translation and mapping effort will be probably required to properly import data. Data elements in a legacy system need to be mapped to elements of standard schema, such as the SMPTE metadata dictionary, or EBU P/META, or Dublin Core. Missing elements, such as a Universal Media Identifier (SMPTE UMID) need to be created, possibly using components from legacy identifiers.

New metadata can be gathered off-line (after media digitisation and storage): storage of material (audio browse) on a server would allow the cataloguing staff to perform this task, rather than the technicians doing the transfer.

2.6.2 Features Extraction

Additional enhanced metadata (contents-oriented) can be produced by accessing already stored digitised media, e.g.: time coding, indexing, meta-data related to acoustic contents as segments detection and classification, and automatic speech recognition.

Some of the features can be extracted automatically; nevertheless, features validation should always be performed by technician intervention, crucial for that phase.

2.6.3 Digital-to-Digital Processing

Depending on the broadcaster's adopted workflow, also digital-to-digital processing could be performed off-line, so that related technical metadata will be completed off-line too.

Chapter 3 Technology and Costs

3.1 Equipment

Thanks to the high performances offered by the current technology, both *Flow Control* and *Transcription* stations can be based on Intel (or AMD) personal computers in client server configuration, running Windows (client PC) and Linux (server PC) operating systems, configured with all devices and boards required to acquire, process and temporally store digitised audio, images and related metadata.

For what concerning playback devices, ¼" tape decks are going to be obsolete (as well as tapes). Even so, STUDER remote controlled tape recorders are typically already owned by broadcasters, or can be provided as refurbished devices. Vinyl turntables of broadcasting level and remotely controlled are difficult and/ or very expensive to provide, a specific development was performed in PRESTO to achieve remote control of turntable (start/stop) with double arm (up/down for each arm) PRESTO-W3-ACS-167b.04, "Key Links Systems Specifications".

The most suitable media for digitised audio, images and metadata consists of digital data cartridges (as DLT, LTO, …), allowing to safely store hundred GB of data at a reasonable cost. Even if the final storage of data is not strictly part of the reference preservation process, the usage of automatic storage systems is strongly recommended to achieve high usability of digital media: digital data cartridge allows the use of tape libraries.

3.2 Sizing

The digitised master copy (24 bit at 48 KHz per channel) requires about 1 Gbyte of data per stereo audio hour, plus what needed for lower quality copies (about 60 Mbytes per hour of 128kbps mp3) and scanned images. In this context, metadata occupation is negligible.

According to the number of transcription lines (i.e., the number of analogue media that can be processed at the time) the infrastructure (network and shared resources) must be properly sized.

The balancing among computation required to perform the transcription and processing capability of nowadays personal computer, leads to the definition of a Transcription Workflow architecture consisting of two PCs sharing a unique file system, cooperating in the digitisation of up to eight analog channels.

3.3 Costs

The main cost drive in transcription is the man labour. The below tables provide a tentative comparison between the Massive and Expert transcription approaches. Devices

costs are aligned to commercially available systems. Costs due to the procurement of playback devices are not included.

The study cases constraints are summarised in the **Common Data** box.

The **Massive Transcription** box shows how the cost of equipment is only a percentage of the labour cost, although the massive approach achieves an outstanding rate of 36 tapes (avg.) processed per shift and operator. The second major cost item is given by the digital storage media cost (only depending from the source archive size - two copies of each media are considered, i.e., original and backup).

The **Expert Transcription** box shows how dramatically labour costs increase, due to the limited number of tapes per shift processed by each operator.

The result of the exercise is that expert transcription cost is more than four times that of massive transcription.