PRESTO – Preservation Technologies for European Broadcast Archives
IST-1999-20013
Archive Preservation and Exploitation Requirements

DOCUMENT IDENTIFIER	PRESTO-W2-BBC-001218 Archive Preservation and Exploitation Requirements
DATE	06.06.2001
ABSTRACT	A survey has been made of the holdings and preservation requirements of ten major European public service broadcast archives. The survey also covered 'the business they do and how they do it': the contribution made by the archives to their companies. The archives also gave their plans for future services as media moves from 'tapes on shelves' to mass storage and electronic distribution. This report summarises the findings.
AUTHOR, COMPANY	Richard Wright, Adrian Williams BBC Contributions from e-vod, NTEC and ACS
INTERNAL REVIEWER	
WORKPACKAGE / TASK	WP2

DOCUMENT HISTORY

Release	Date	Reason of change	Status	Distribution
1.0	19.01.2001	First draft; BBC text	Draft	Confidential
1.1	02.02.2001	Final draft, all contributions	Draft	Confidential
1.2	04.05.2001	Complete	Final	Public
1.3	06.06.2001	Complete	Final	Public

Table of Contents

1. Executive Summary

A survey was made of the holdings and preservation status of ten major broadcast archives. These archives represent a significant portion of total European broadcast archives, including some of the largest individual collections.

Holdings
The survey found about 1 million hours of film, 1.6 million hours of video recordings, and 2 million hours of audio recordings in the ten archives. It is estimated that the total European holdings of broadcast material are probably **ten times** larger:

- 10 million hours of film
- 20 million hours of video
- 20 million hours of audio

Access
The content is unavailable to the general public and often unavailable even to national archives and educational institutions. Much of the content is unique, e.g. master material that cannot be allowed to circulate generally, and all of the content has rights issues. A major goal of preservation work for broadcast archives must be to find joint solutions to preservation and access problems: preservation for access

Preservation Status

> **Obsolescence:** At least 2/3 of the material in archives cannot easily be used in its existing form
> **Deterioration:** Approximately 1/3 of the material has one form or another of deterioration
> **Fragile media:** Roughly ¼ of the material cannot be released for access because the media are too easily damaged

Cost per use
The true cost of an asset is total lifecycle cost. True benefit is related to the number of times that asset is used over the lifecycle. **Archive preservation strategy should aim at the "lowest cost per use" over the life cycle of the new media, NOT at the lowest transfer cost.**

Actual costs

Bringing together information from all ten archives allows the following overall preservation costs to be identified: (from Table 4.1.1)

	Like-for-like preservation lifecycle cost/hr	Usage over life-cycle	Cost per use (€)	Digitisation preservation lifecycle cost/hr	Usage over life-cycle	Cost per use (€)
Film	2000	5	400	3000	10	300
Video	200	5	40	300	10	30
Audio	120	5	24	180	10	18

These costs, for the year 2000 in Euros, exclude much significant detail, and do not represent the actual costs for any particular archive or project. Digitisation and mass

storage is about 50% more expensive, but is expected to double the usage of an asset.

Funding

The aim of preservation work is to retain for the future, as cost effectively as possible, the portion of existing broadcast archives that will contribute most to future usage. The conclusion from current overall archive usage figures is that **the value of an item must be more than four times the preservation cost** in order to be financially justified on a commercial basis.

Commercial basis

For most broadcast archive material, this condition can easily be met**,** as **one minute of sold or re-used archive material will pay for preservation of one hour of archive material.**

Heritage basis

For material that cannot pass the ‘commercial economics’ criterion as outlined above, there should be a safety net of assessment for cultural/historical value, and a separate funding mechanism.

2. Introduction

2.1. The Audio-visual Heritage of the 20th Century

The Twentieth Century was the first century with a record of its significant events – the sounds and moving images – on film, audio and video media. A major repository for this record is the collective broadcast archives, particularly the archives of major broadcasters. In Europe, the main broadcasters are publicly supported, leading to archives that have a role not only in the business of broadcasting, but also fulfilling a public service requirement to support wider educational, cultural and heritage purposes.

This record is now largely at risk, as the bulk of these recordings reach the stage at which they are deteriorating, or on an obsolete format, or both. Broadcasting was never developed as a mechanism to create and hold permanent audio-visual history. The consequence is that these archives have arisen to support broadcasting, and have no business model or funding model specifically designed to support preservation.

2.2. Purpose of the Document

This document presents the findings of a detailed study of the preservation status of European broadcast archives. The purpose of the study is to understand current preservation strategy, technology and cost – as a basis for developing, within EC Project Presto, a comprehensive and cost-effective preservation strategy, and developing key-link technology for reducing preservation costs.

2.3. Scope of the project

EC Project PRESTO has the aim of increasing the efficiency of the technical work needed to preserve broadcast archives. The efficiencies are of two sorts: reducing costs, and developing strategies to ensure that archives are not simply physically preserved, but preserved in ways which maximise their future benefit. To get the size, shape and urgency of the problem, a survey (see Appendix 1) has been made of the holdings and preservation requirements of ten major European public service broadcast archives (see Appendix 2).

The survey also covered 'the business they do and how they do it': the contribution made by the archives to their companies. The archives gave their plans for future services as media moves from 'tapes on shelves' to mass storage and electronic distribution. This report summarises the findings.

The survey covered the following areas:

The present
1- what broadcast archives do: their place in the business
2- what they hold
3- current preservation technology
Processes and costs

The future
4- new services from new holdings
5- requirements for preservation
Processes and technology

2.4. Methodology

Orientation documents and the Archive Survey (see Appendix II) were distributed among the Broadcast Archive Users Group (BAUG), a committee representing the major archives across Europe, to gain a clear perspective of the challenges facing Project PRESTO. The results were collated and are examined in this report. Appendix 1 details the list of BAUG.

3. Holdings

3.1. Content

The ten archives in the survey represent a significant portion of total European broadcast archives, including some of the largest individual collections, but total European holdings of broadcast material are probably ten times larger. The survey found about 1 million hours of film, 1.6 million hours of video recordings, and 2 million hours of audio recordings in the ten archives.

The content covers the entire century, as broadcast archives include bought-in film and even wax cylinder material from before the development of the broadcasting industry. The record of news, current events, sports, culture and entertainment covered in radio archives dates from the 1920s (recorded originally on shellac discs), and there are film recordings of television output from 1936 onward.

Access to this content is notoriously difficult, because almost all the material is on 'professional' formats (film, broadcast-standard videotape) which need special players, certainly unavailable to the general public and often unavailable even to national archives and educational institutions. Also much of the content is unique, master material which cannot be allowed to circulate generally. **A major goal of preservation work for broadcast archives must be to find joint solutions to preservation and access problems: preservation for access.**

3.2. Audio

Table 3.2.1 Number of Audio Archive Items by Format

Format	No. of Items
Shellac & Vinyl	863,459
1/4" tapes	3,162,032
Cassettes	39,946
DAT	132,487
CD	175,422
Minidisc	2,104
Tandberg QIC cartridges	15,000
	Total Sum: 4,390,450

3.3. Video

Table 3.3.1 Number of Video Archive Items by Format

Format	No. of Items	Format	No. of Items
2" tape	198,585	D3	286,663
1" tape	935,993	Digital BETACAM	395,099
3/4" UMATIC	943,446	D5	37,000
M2	355,000	DVC pro	19,430
1/2" BETACAM & BETACAM SP	1,986,135	DVCAM	73
VHS	657,510	CD-ROM	37,575
D1	1,090	DVD-R	496
D2	142,549		
			Total Sum: 5,996,644

3.4. Film

Table 3.4.1 Number of Film Archive Items by Format

Format	No. of Items	Format	No. of Items
35 mm negative	208,180	16 mm print	479,915
16 mm negative	148,445	35 mm SEPMAG	241,040
16 mm reversal (all)	442,000	16 mm SEPMAG	266,162
35 mm print	322,950	Intermediates (Interpos/neg., CRI)	6,714
			Total Sum: 2,812,021

4. Preservation

4.1. Status

- Obsolescence: At least 2/3 of the material in archives cannot easily be used in its existing form, because the medium is too specialised (film) or obsolete (2" videotape) to allow easy access. For audio, this includes the massive holdings on ¼" open-reel tape.
- Deterioration: Approximately 1/3 of the material has one form or another of deterioration:
- Dye fading on colour film
- Weakening of binders/adhesives holding magnetic particles to the polymer tapes used for both video and audio recordings; this problem is expected to become acute for ¼" tape in audio archives over the next decade, and is already a major problem for 1" and U-Matic material in video archives.
- Decay of the polymer material itself, becoming brittle or non-uniform
- Degradation (to the point of failure) of tape splices; this is a problem for the majority of TV news material from the 1960's and 1970's, which was 'direct recorded' onto film, spliced by journalists, and stored in newsrooms.
- Formation of acetic acid from acetate tape (vinegar syndrome). This acid, primarily from film sound tracks, can then attack both film and video (or even audio) material if held in the same room.
- Nitrate film: all archives have segregated this material, which is extremely combustible and can be explosive, but not all archives have copied the material onto safety film.
- Fragile media: A large part of the holdings cannot be released for access because the media are too easily damaged. E.g. film negatives; film prints except for access by qualified professionals; all shellac and vinyl audio recordings.

4.1.1. Uncertainties in cost determination

There are two main problems with determining costs of preservation:

1. Uncharged costs (work that is known to be part of preservation, but with costs that cannot be isolated) – the business models and accounting methods in individual archives do not reveal the costs of known factors in archive work and preservation work. The costs of finding, transporting, labelling and quality checking of material selected for preservation may be inaccessible. Even storage and media costs may be uncharged, or simply 'part of the overhead'. Often staff labour rates, and outsourced piecework rates are the only information easily available.
2. Unconsidered costs (work that is not considered part of preservation) – the true cost of preservation must be based on 'total cost of ownership' of various media assets, for the life of the asset – including further replacement costs at the end of that life. Most preservation projects identify the "transfer cost" to turn an item on one format (carrier) into an equivalent item on new, up-to-date media. However total digital asset cost has at least seven aspects (adapted from Hendley, 1998):

    1. Item creation – the "transfer cost"
    2. Item selection and evaluation – selecting/locating what to preserve
    3. Item management – what is needed to make the item usable, including:
        Item documentation – all associated metadata costs
        Item validation – all quality control costs
        Item storage – all physical accommodation costs, including temperature and humidity control if required
    4. Resource disclosure – creation of formal cataloguing, or other method to find the item
    5. Item use – costs associated with 'issuing' the item
    6. Item preservation – cost of maintaining the item in good condition, including recopying as necessary if item is subject to 'wear in use'

7. Rights management – ensuring the item is used legally, and costs associated with collecting information to minimise rights clearance costs at time of item usage.

Where like-for-like preservation work is being undertaken, as in the transfer of individual items from an old tape format to a new one, it may be possible to ignore both uncharged and unconsidered costs, because all seven aspects of unconsidered costs might apply equally to both the old and new item. Yet even here, cost types 2, 6 and 7 could be significant, especially if selection is difficult, if the new item has different storage requirements than did the old, or if an object of the whole project is to increase utilization and so needs to improve the rights management.

When an entirely new technology is being considered, as in conversion from rooms full of ¼" tapes to a large data-tape robot with networked access. Then all the layers are important in order to make a proper comparison of alternative preservation strategies (such as conversion to CD with physical storage and delivery, instead of data-tape/robot with electronic delivery).

However even the 7-layer model doesn't include what may be the biggest factor in comparing technologies – new service or business opportunities offered by, for example, electronic delivery. Therefore an important part of this report is detailed examples (Chapter 5), from technology leaders in audio, film and video, of both the required technologies and the business opportunities provided by pioneering examples of best use of new technology in storage and delivery for these media.

4.1.2. Cost per use

The true cost of an asset is the total lifecycle cost. True benefit is related to the number of times that the asset is used over the lifecycle. Not every use has equal benefit: not every test-drive means a sale to a car dealer. But on average, more test-drives mean more sales – and more media issued from the archive means more benefit to the broadcaster and to the wider public service. Therefore a simple way to combine transfer cost, life cycle cost, and the significance of new service opportunities, is to translate those new opportunities into a predicted rate of item usage. Options for preservation can then be compared, in monetary terms, on a "cost per use" basis. A significant conclusion of this report is that archive preservation strategy should aim at the "lowest cost per use" over the life cycle of the new media, NOT at the lowest transfer cost.

Actual costs: bringing together information from all ten archives allows the following overall preservation costs to be identified:

Table 4.1.1 Comparison of costs and usage between traditional and digital systems

	Like-for-like preservation lifecycle cost/hr	Usage over life-cycle	Cost per use (€)	Digitisation preservation lifecycle cost/hr	Usage over life-cycle	Cost per use (€)
Film	2000	5	400	3000	10	300
Video	200	5	40	300	10	30
Audio	120	5	24	180	10	18

These costs, for the year 2000 in Euros, exclude much significant detail, and do not represent the actual costs for any particular archive or project, because of the issue of uncharged costs that vary from archive to archive. There are many assumptions in the above numbers, as described in detail in the report. These assumptions may not be at all valid in particular cases, but it is the opinion of the project that these cost and usage estimates numbers do represent the general situation.

4.1.3. Breakdown of Reported Costs

Costs for individual media types, averaged over questionnaire results from all respondents, were as follows.

Table 4.1.2 Average costs for preservation of film formats

Film Format	Average Cost / h (€)
35 mm negative	1994
16 mm negative	1685
16 mm reversal (all)	3011
16 mm reversal (scotch edited)	2850
35 mm print	1900
16 mm print	1872
35 mm SEPMAG	2415
16 mm SEPMAG	373

Table 4.1.3 Average costs for preservation of video formats

Video Format	Average Cost / h (€)
2” all formats	176
1” all formats	357
¾” U-Matic all formats	344
½” BETACAM and SP	230
½” MII	230
VHS	100

Table 4.1.4 Average costs for preservation of audio formats

Audio Format	Average Cost / h (€)
Shellac and vinyl, all formats	299
¼” tapes	238
DAT	195
Mini disc	
CD	70
Cassettes	

4.1.4. General conclusion on costs/benefits

Digitisation and mass storage is about 50% more expensive, but is expected to double the usage of an asset. The difficult situation is the transition period, where archive users do not have access to electronic delivery, or a delivery infrastructure of sufficient bandwidth does not exist. In such a case, some archives are doing BOTH like-for-like and mass digitization, at about double the cost of like-for-like – which is still just about cost effective, will become more cost effective as the need for like-for-like diminishes, and keeps the door open for the new service/business opportunities.

4.2. Preservation Funding

If the 'lowest cost per use' principle is accepted, then the aim of preservation work is to retain for the future, as cost effectively as possible, that portion of existing broadcast archives that will contribute most to future usage. This model ignores intrinsic value of individual items (the archive 'crown jewels'), such as the only footage of a key historical event. We believe this approach is justified, for two reasons:

1. Intrinsic worth is very difficult to assess on a shot-by-shot basis, especially when dealing with millions of hours of material
2. Key footage/recordings are generally known to archivists, and constitute a small portion of holdings. An archive preservation strategy cannot be based on minority issues

When 'crown' jewel items are excluded, archive preservation decisions should be made on gross patterns of usage for particular large categories of recordings. Therefore it is important to have information on past use. It is equally important to have an estimate of the value of future use. If the expected future use has a value greater than the preservation cost, it is financially viable. If not, then it goes to the 'back of the queue', awaiting technology to lower the preservation costs, or new reasons to increase the estimate of future use. One such 'new reason' would be an expansion of the access to archive material, for business, commercial and educational or heritage purposes.

The value of archive material is very difficult to assess, because usage can be 'in house' or for commercial sale, and because footage rights are sold for specific purposes, which are unknown at the time of making preservation decisions. If a future user develops a commercial project that requires archive material, then footage, which for years has produced no revenue. Suddenly becomes worth 1000 Euro per minute or more. Again the only solution is overall statistics.

About 20 to 30% of the archive content is issued to users each year. Estimating shelf life at 20 years, this means that each item is used from 4 to 6 times, hence the figure of 5 uses/item in the costs tables, above. However the estimate is that about 75% of this is viewing (browse) material, and then when broadcast quality material is issued, about 75% of that material is not actually used in programme making. This means that a given archive item has a probability of 1/16 of ending up actually shown/heard in a new programme every time it is issued. The total likelihood of usage is 5-uses/item times 1/16 probability of a usage ending up in a new programme, meaning a 5/16 chance of future use. Taking the slightly more pessimistic figure of 4/16 (= ¼), the value of an item (the price it can be sold for, must be more than four times the preservation cost in order to financially justify the preservation costs.

For most broadcast archive material, this condition can be met, and so preservation work has been funded and is currently being funded. With new technology (mass storage and electronic delivery), which our figures lead us to believe will roughly double archive use, the expectation of use is ½ instead of ¼, and so the value only needs to be double the preservation cost.

It is more effective to preserve the bulk of current content rather than a piecemeal approach because of costs inherent in monitoring usage, selecting material and the inevitable high price at time of re-use. The content that will, inevitably, be reused will pay for a great deal of unused material, and will prevent the loss (forever) of material that may deemed today to be of low value, but which in the future may have uses and values which we cannot begin to predict. As a rough rule-of-thumb, the high value of material when actually used means that one minute of sold or re-used archive material will pay for one hour of preservation.

There may however still be material that is narrowly defined in scope and interest, and for which statistics indicate very low reuse. The preservation of this material cannot be justified on pure economics, in terms of future use. This situation highlights the problem of having an essentially commercial model for the preservation of broadcast archives. For material that cannot pass the 'commercial economics' criterion as outlined above, there should be a safety net of assessment for cultural/historical value, and a separate funding mechanism and model, in order to ensure that value in the wider sense is not irretrievably lost during archive preservation.

4.3. Preservation Processes and Requirements

4.3.1. Current Projects

The following charts summarise some of the projects in progress within the archives covered by the survey. These charts only cover dedicated projects, not the volumes of the 'conversion on demand' work being done as day-to-day activity in the archives.

Table 4.3.1 Number of dedicated film preservation projects

FILM	BBC		INA	NAA		ORF	SVT	YLE	
	Per year	remain	Per year	per year	remain	per year	per year	per year	remain
35 mm negative	start in 2003		20 h	40 h	160 h		50 it.	60 h	675 h
16 mm negative				25 h	285 h			70 h	5.500 h
16 mm reversal (all)	12.000 it	24.000 it	1.600 h						
16 mm reversal (scotch edited)			400 h						
35 mm print	start in 2003					On req.			
16 mm print						On req.			
35 mm SEPMAG	21.000 it	42.000 it						60 h	675 h
16 mm SEPMAG			1.200 h	70 h	150 h		5 it	70 h	5.500 h

Table 4.3.2 Number of dedicated video preservation projects

VIDEO	BBC		RAI		INA	NAA		ORF	SVT	YLE	
	per year	remain	per year	remain	per year	per year	remain	per year	per year	per year	remain
2" PAL	11.500 it.	finished	Systematic transfer plan started 1998		5.000 h		finished	Finished	3500 it.	600 hr	1200 hr
2" B/W							150 h				Finished
2" SECAM											
2" other Standard											
1" B PAL			No systematic transfer plan, but individual preservation process		5.000 h	3.000 h	7.500 h	on req.	~200 it.		
1" C PAL	9.000 it.	18.000 it.									
1" B SECAM											
1" C SECAM											
¾" Umatic LB PAL			Systematic transfer plan is starting		12.000 h	675 h	2.900 h	On req.			
3/4" Umatic H PAL	5.750 it.	11.500 it.									
3/4" Umatic SP PAL											
3/4" Umatic LB SECAM											
3/4" Umatic H SECAM											
1/2" BETACAM											
BETACAM SP	start '06										

Table 4.3.3 Number of dedicated audio projects

AUDIO	BBC		INA		RAI		NAA		ORF	
	per year	remain	per year	remain	per year	remain	per year	Remain	per year	remain
Shellac	10.000 it.	20.000 it.					500 h (incl. CD)	4.500h		
78 RPM vinyl			300 h							
LP vinyl									On req.	
45 RPM vinyl										
1/4" tapes	5.600 h (Radio 1 archive)		8.000 h		50.000 it.		367 h	1465 h	100 h	
DAT			1.500 h							
Mini disc										
CD					automated transfer system under test					
Cassettes								350 h		
Other							750 h	3.000 h		

4.3.2. Current Processes

The tables (above) do not give an impression of the scale of the effort, time, organisation and equipment required to handle tens of thousands of items. Nor do the tables give any indication of the complexity of the overall process. The process is key, because efficient technology can still be used inefficiently. The following lists give detail for conventional processes for 'on demand' transfers for film and videotape, as actually used in one of the surveyed archives.

Transfer process for film (news):

1. A film box is chosen in the numerical order that also - more or less - is parallel to date order.

2. A first check is made concerning material in relation to catalogue and the physical status is examined.

3. In cases where film and magnetic tape originally are kept in the same cardboard box, they are separated and get new labels.

4. All boxes get a barcode label.

5. The material is registered in the stock control system (Mathis) with automated transfer of information to the library database.

6. Some notations are made in the old manual stock-control systems.

7. The processed boxes are labelled with a small orange "V" meaning the material has browsing material on video cassette and is transferred to either Beta SP (from the beginning) or today's Digital Beta. This is to make sure films are used for browsing and unnecessary scanning.

8. The length in meters is checked.

9. The film is sent to (external) laboratory for cleaning.

10. Laboratory invoice is registered for check up purposes - both budget and work process.

11. After return of film, film and magnetic tape is sent to scanning (in-house facilities) to make the Beta and a parallel browsing VHS.

12. Film and videocassettes arrive from scanning together with technical report.

13. Bar code and other labels for the cassettes are printed and the material is registered in the Mathis-system.

14. The Beta cassette is sent to the automated robot storage system.

15. The card catalogue information is manually transferred into the library database.

16. A spot test of the transfer is made via the browsing cassette especially for items containing both black/white and colour sequences.

Transfer Process for Video:

1. The preservation of a video item is primarily initiated by a request from a new production or a rerun purpose. The material has thus been located through a catalogue research in the library database where you also find the necessary technical information such as format and browsing material. Maybe there has also been a picture browsing with a time-coded videocassette if you search for a certain insert.

2. Via the library database - or if need be for a deeper investigation of the technical status, the material handling system, Mathis is consulted - you determine whether you need to transfer the material to your present production format or current broadcast format.

3. You order the transfer at production facilities

4. The technicians decides during the transfer process about different actions that might be needed - from vacuum cleaning to "baking" to editing, if the transfer has to be made in sequences depending on mechanical clips etc for old tapes.

5. A VHS browsing copy parallel to the new format is made.

6. New tape is run through for quality check.
7. New technical information is fed into the material handling system by the technician, which automatically updates the library catalogue system. A barcode label is printed out for the new copy.
8. The new material is ready to be lent to the customer for a new production or a rerun.

4.3.3. Process analysis

In order to understand the requirements of efficient preservation processes, a detailed analysis was made of the relationship between existing archive processes, and the extra needed for mass-transfer preservation work. The results are shown on the following figures. The whole process has been broken down into stages with colour codes, indicating where current archive processes are in place (green light), where new steps are needed (red light), and amber light for those tasks which may be possible within current archive capabilities.

PRESTO generic preservation process

The overall process has four stages:

Composition – collecting the media and associated data and preparing for conversion

Table 4.3.4 Break down of Composition phase of generic preservation process

Stage	Identify	Get Basic Items	Get assoc. items	Dispatch
		Legacy System "Issue"	New Process	
Metadata Processing	Material needing processing	Extract metadata from existing documents relative to digitising	Get any additional information	Prepare all labels for old and new physical media
Media Processing		Pick media from archive	Program docs. + physical items relevant to digitising	Dispatch media

Digitisation – converting the existing analogue signal to a digital representation; not all transfers involve digitisation, but any transfers which do NOT make a digital output are now obsolete, and any which start from a digital version are very much simpler than for analogue originals (in principle, a digital-to-digital transfer can be totally automated)

Table 4.3.5 Breakdown of Digitisation phase of preservation process

Stage	Check	Digitise media	Digitise associated media	Process digitised media	Format physical media	Coding of associated versions, e.g. "browse"
	Some steps may be the same as for standard archive "format conversions", but a new total process is needed for efficient mass transfers, and creation of media and metadata elements for new technology (e.g. server storage and network access).					
Metadata Processing	Ensure metadata complete and accurate	Automatic or manual quality check (of digitisation) during process	Capture of entire documents or essential new data	Create shot detection, key frames, speech recognition, MPEG-7 data	Arrange metadata to fit standard file format	Transfer metadata to any subsidiary media
Media Processing	Ensure media matches Metadata	Transfer content form original media to server or directly to new media	Scanning or other processing of any associated media or documents	(media must be on or written to a server); new data stored	Media on server made into "image" of final output (if any)	Create images or media for all subsidiary data

New Media Creation – production of the new item(s) which will replace the original media

Table 4.3.6 Breakdown of New Media Creation phase of preservation process

Stage	Label	Write	Label	Dispatch
	Electronic metadata goes with media			
Metadata Processing	Ensure all required identifying names or numbers are in the metadata	Metadata written to same media as audio and video	Produce paperwork	To accompany new media
Media Processing	Create electronic image of new media	Burn or record new media item	Physical print of Identifying data onto media item and onto packaging	Media grouped in batches, packaged for transport

Update of the Physical Archive – the re-introduction of the new media, whether a discrete item or just a data file in a mass-storage system – into the physical (shelves) and logical (catalogue) control of the archive

Table 4.3.7 Breakdown of the Update of physical Archive process

Stage	Check	Accession	Store	Update Metadata	Final	Remove old media
		Legacy system "Accesssion"	Standard Function	New, possibly very labour intensive		
Metadata Processing	Ensure metadata complete and accurate	Update legacy system to use new media identifiers	Update 'site map' and signs for physical archive	Import new metadata from media; produce shot / item level descriptions & time	Last chance to ensure that new media matches catalogue	Delete all stock numbers for replaced media items
Media Processing	Ensure media matches metadata	Scan barcodes inot stock control system	New media placed in physical archive		Last chance to ensure that new media matches original media	Dispose of Old Media

4.3.4. Process Requirements

Broadcast archives already perform all the steps necessary to preserve individual items. It is however common for a item to be available only on 'the wrong format', unusable by an archive customer. The problem is that low volume item-by-item transfers, never motivate broadcasters to set up really efficient mass transfer processes. It is however essential that funding is committed to develop a 'production line' approach, streamlining every possibly aspect with automation, elimination of manual handling, and ensuring an efficient workflow.

Piecemeal transfers are characterised by lack of automation, and manual intervention at every stage. Typically only the technical transfer cost is considered, not all the associated manual steps because they are seen as standard archive work done by existing staff, in a manner similar or identical to other daily tasks of searching, retrieving, issuing and acquiring of material.

What is missing is an overall perspective of the preservation work. Efficient preservation work needs a defined, well-engineered process, not just lots of piecemeal transfers. Attention must be given to the following elements

Process 1) End-to-end process – a view of all the work involved, to capture all possible efficiencies (to identify unnecessary steps, and find when information needed at one point could be created without manual intervention at another point)

Process 2) Identification of costs – to identify cost-attractors: steps which would benefit from investment (in technology, physical plant, or moving of staff or equipment) in order to reduce costs

Process 3) Development of cost-effective approaches to batch processing – investment in systems or equipment which more than pays for itself for volume work. Examples are

Process 4) Automatic quality-checking equipment

Process 5) Software for overall control of a preservation end-to-end process

Process 6) Purpose-built preservation areas to optimise equipment and streamline movement of material

Process 7) Investment in in-house rather than contracted facilities

Process 8) Links to the wider business process – Preservation of a broadcast archive is, in most cases, a part of the overall information and media flow of a broadcaster. An immediate consequence is that preservation work should deliver an output (a media format or storage and delivery system) which is useful to the business, which the piecemeal approach will accomplish. However there are wider issues such as programme documentation, subtitle information, rights clearance, creation of new metadata to improve the catalogue which are not considered part of preservation work in the narrow sense, but which need examining at from an end-to-end perspective of the entire business. At the business level, it may become cost effective to include work within a preservation project which pays off to another part of the business. An example would be digitisation of programme documentation (scripts, running orders, contracts) as part of preservation work, in order to increase usability of the new media, and decrease costs of activities such as rights clearance.

Process 9) The most important single issue, which does not fit into a like-for-like, item-based piecemeal approach, is the concept of using preservation work to create the future archive. The archive embodied in the technology that will be needed in the next decade, rather than in the particular tape format needed in the current year. The difficulty is that it is hard to calculate strict cost-benefit evaluations of future requirements, and hence difficult to develop a business case for justifying the transformation of archive technology in terms of lowest cost. Long term decision making requires managerial vision, a commitment to the eventual development of “tapeless working” in TV and Radio. At the minimum level, this vision consists of accepting that electronic access would double the use of archive (see table 3.1). **It is vital to the future of broadcasting and broadcast archives to have an increased appreciation of the value, not just the cost, of new technology.**

5. Preservation Technology and Requirements

The survey of broadcast archives revealed the following gaps in current archive technical practice, and the technology for preservation.

5.1. Security Copy

Over half of the audio (up to 80% in some cases) and film holdings (100% in some cases), and about 1/3 of the videotape has only one master version. In most cases the master is not circulated, because there are "viewing copies" such as film prints and VHS tapes – but it still means that when the full-quality material is needed, the unique archive item is taken off the shelf and exposed to a degree of risk. Restricting circulation to qualified professional users controls this risk. The result is a low-access, high-risk approach. The risk could be reduced – virtually eliminated -- with **use of datatape masters and backups**, with no necessity ever to compromise the original, and with a cheap method for creating off-site copies using the conventional backup procedures of the data-processing industry. **The costs and benefits of datatape masters, over at least a 20-year item lifetime, need to be assessed within an overall context of moving from media-based to mass-storage based processes during that time period.** This will require detailed modelling of preservation and production workflow, present and anticipated.

5.2. Videotape Transport / Playback

The advances in videotape players (VTRs) have not been introduced into players for old formats, because manufacturers have no market for such players. The problem is made much worse by having to use old media, where there are various problems assorted with the ageing of the media itself. What is really needed are players that work **better** than the latest VTRs, because of the media degradation. Particular problems frequently encountered are:

1. **Head clogging**: magnetic-coating particles coming free from backing tape and building up on rollers, guides and heads of VTR until various forms of signal loss or tape damage occur
2. **dropout**: loss of signal pickup for one or more lines or even for entire frames
3. **Sticktion**: stick/slip contact causing jerking of tape
4. inconsistent **Tension** adjustment (incompatibility between record and playback)
5. **Scratches** – damaged tape

There is therefore a need for using modern technology to assist the playback of old video formats. Four areas should be considered:

Technology 1) It is particularly important to have the best possible electronic processing, to ensure reliable and stable signal detection, to the greatest degree possible using modern methods. This will minimise the dropout, and hence minimise the transfer errors. This approach is better than relying entirely upon dropout correction, which is not perfect and is essentially a restoration technique, not a preservation technique.

Technology 2) It would also be helpful to improve the mechanics of old VTRs, but this is far more difficult than improving the electronics and may not be feasible or cost effective.

Technology 3) A third aspect of improving videotape playback is improved methods of conditioning of the videotapes. PRESTO should evaluate the costs and effectiveness of the various cleaning and baking regimes currently in use.

Technology 4) Dropout correction should be re-examined with regard to digitisation, because if a video signal is being captured to a server there are opportunities for signal correction which are not possible in a standard transfer (e.g. using information from adjacent frames and from preceding as well as following lines)

5.3. Film Transport

There is also a need for using modern technology to assist the playback of certain film formats. All the major archives had large collections of news material from the 1960's and 1970's that is on film, manually spliced by journalists rather than professionally spliced. About 500,000 items of this or similar type were found in the survey. This material is now in a deteriorated state, which prevents preservation work until the material is manually stripped, cleaned and re-spliced. This handling stage needs research, to remove the high-cost manual effort.

Two areas need particular attention:

Technology 5) Auto-splice: a cost-effective way to re-splice the film, to allow conventional telecine transfer for preservation

Technology 6) Alternative handling: a method of film handling and scanning which is unaffected by splices

5.4. Quality Control

There was general agreement in the survey that better quality control was vital to preservation work, and that existing methods were inadequate on the following grounds:

- manual checking is labour intensive: too costly and time consuming
- despite investment in manual checking, quality is not adequately insured – errors get through
- existing automation of error detection is primitive and faulty

Three particular areas of desired improvement were cited:

5.4.1. Playback monitoring

Technology 7) Multiple signals: it is necessary for cost efficiency for preservation equipment operators to work on more than one item simultaneously – usually to a peak of four audio or video transfers run by one operator. Although one operator can monitor four video screens, it is not possible to do continuous monitoring of multiple audio signals. There should be automatic methods to identify basic problems with audio signals, to back up the 'spot checks' performed by the operator. This will allow the operator to concentrate on properly setting up a transfer, ensuring initial quality, and then relying upon automation to warn of signal dropout or gross problems during the course of the transfer.

Technology 8) Signal capture: while video is easier than audio in terms of multi-signal monitoring, it has far more problems with signal dropout or transport errors (head clogging) during playback. The gross errors will appear on the monitors, and the trivial errors (single-line dropout) can be immediately compensated in hardware. In between these extremes, the operator will not know how much brief loss of signal has occurred. Tracking errors and other small distortions will also not be noted until they reach quite high levels. Improved monitoring of playback, ideally linked to improved playback itself (section 5.2), is needed to ensure quality of video playback.

5.4.2. Error logging

Many aspects of signal playback my have problems. Signal restoration projects have identified dozens of artefacts in archive signal: blotches, scratches, flicker, grain, noise, fading. Many of these artefacts can be automatically detected, with varying degrees of difficulty and success.

Technology 9) Although preservation work is not the same as restoration, it is desirable when doing preservation work to capture any useful information that may be needed for subsequent handling of the item.

Information about artefacts is useful in three ways:

- Quality: the type and number of problems encountered during the transfer process

- re-use value: the ease with which the material can be re-used in new production without further processing
- restoration: if restoration work is to be undertaken, an error log will say how much work is required, and can also be used to greatly speed the restoration work – because the log is a list of exactly where and what the problems are (artefact identification can be the major part time-factor in a restoration process)

In audio, there is already a proposed EBU standard for such an error log.[1]

5.4.3. New signal checking:

Technology 10) The key aspect of media preservation is to create a new item that is as close as possible in signal content to the original. If the original is digitised, the possibility exists to directly compare the two items to ensure quality. This comparison depends for its validity upon the quality of the playback (Section 5.2), but if these problems can be handled to an assured level of quality, then further automation can be used to ensure the new archive item does indeed match the original.

Technology 11) A related problem is ensuring the match between multiple new items, which occur when various formats are being created. The main example is creation of a master copy and various lower-data rate versions for edit or browse access.

Technology 12) Part of the quality control is to ensure that the signals "match" in terms of content. This is non-trivial: the digital files will certainly not match, as the codings differ. Yet the content (the sound / image represented) must match for these various versions to be of any use.

Other issues in managing multiple new items include:

- ensuring exact time registration,
- metadata maintenance across all versions
    - identifiers (which will not be identical but need to conform to an overall system)
    - time code
    - other embedded metadata
    - links to external metadata (catalogue)

[1] EBU (European Broadcasting Union) Specification: Tech 3285 Supplement 2: The Broadcast Wave Format: Capturing Report – http://www.ebu.ch/pmc_bwf.html

5.5. Audio Preservation

This section provides user's requirements applicable to audio preservation, according to the information provided by the PRESTO "Preservation Requirements Questionnaire".

1.1.1. Transcription models

Based on the survey, the following media are considered as relevant to the audio preservation and then analyzed in this document, due the relevance of their presence in the analyzed archives:

Media
¼"
78 RPM vinyl
45 RPM vinyl
33 RPM vinyl
Compact Cassette
CD
DAT

Obsolete media such as wax cylinders, shellacs and others have been excluded, since their contents have already been subject to transcription during past years, generally towards ¼" tapes.

An effective and efficient transcription chain should be able to optimise the following parameters: cost of the process, duration, and quality of the result and security of the original materials. The first two parameters require automation and scalability, while the others need accuracy and care. According to RAI experience, of the above types of media, only 78 rpm vinyl materials and a small percentage of 33 rpm and 1/4" tapes are in such condition to require manual handling and continuous human monitoring during transcription. Therefore, it seems logical to define two separate transcription models: one specially optimised in terms of speed and cost reduction, and a second involving a high degree of specific expertise for endangered valuable materials that privileges accuracy versus speed.

Massive Transcription

A specific system aimed to maximise performances can be conceived to process most of the audio items. Such a system could be operated by low-level staff, or even committed to third parties.

A "massive transcription" system is characterised by its capability of automatic metadata collection, especially concerning the quality of the performed transcription process. It should allow for the parallel execution of playback and digitisation of up to 8 channels and must be easily replicated to obtain the desired transcription rate.

Expert Transcription

Expert human intervention is required for some part of the archive, concerning brittle and/or valuable materials. Broadcasters should deploy transcription systems aimed at ensuring that such media are handled carefully and safely.

Such activities could be performed either on demand or based on 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.

Technical metadata extraction needs in this case to be supervised by an experienced operator, thus implying the need of the whole media "manned" playback. Therefore, the resulting productivity is substantially lower than in the previous case.

The following paragraphs provide a set of requirements, generally applicable to both models. Special cases are pointed out in the text.

1.1.2. Source Media handling

Material handling

The workflow of the operations required to handle materials from the archive to the transcription stations (and vice versa) shall ensure that the material is correctly returned to the archive after the processing was successfully completed.

It should also ensure that materials are not mixed up (e.g., media “a” placed into the box of media “b”).

This can be achieved by logging all the operations into a suitable database and by bar code labelling the outgoing batches of media. Such information will be used to automatically compile dispatch notes and receipts, thus allowing identification of responsible staff in case of media loss.

For all the media considered, the appropriate technologies for bar code labelling of media and the media box/container should be identified.

Outsourcing

Massive transcription only: transcription activities could also be committed to third parties (outsourcing). In this case, material preparation process must release transcription batches including both source (original) and target media.

Manipulation

Due to the media fragility, they shall be subjected to the minimum manipulation possible.

1.1.3. Source Media Playback Machines

Accuracy

Media playback has to be performed as accurately as possible. To satisfy this requirement, only devices delivering outstanding playback accuracy shall be adopted.

Media modification

Transcription workflow shall ensure that no permanent modification is caused to the audio media, unless specifically needed, as in the case spliced tapes. Accordingly, adopted playback machines shall be selected among those models that are known to cause the minimal wear and tear to the supports. Adopted playback machines shall have any device or appliance able to bring unrecoverable modification to the media removed. As an example, for ¼” tapes, recording function shall be disabled in playback machines and cutter for mechanical editing shall be removed.

Playback machines maintenance

Deep checks of playback machines shall be performed at regular time intervals, especially for analog devices, ensuring that mechanical alignment, levels alignment and wear are satisfactory.

Transcription traceability

The description of the transcription chain has to be attached to the report of the transcription of each media, including the settings of all the adjustable parameters (i.e., equalization curve).

Transcription of CD and DAT

Concerning CDs and DATs (material already in digital format), the need for transcription is mainly given by the exigency to increase their usability by means of a network based distribution. In other words, instead of an analog to digital conversion (which is not needed) aimed to preserve original contents, we are dealing with a digital-to-digital media migration, aimed to allow automatic media management and usage (playback, delivery) and to reduce the room needed for physical storage.

CD transcriptions adopt computer based CD readers, extracting data from the media at a speed well beyond the actual duration of the contents (10x is a realistic value for the whole CD), providing track isolation and error correction while reading. An off-the-shelf CD reader would provide the best performance, both in terms of data transfer rate and playback accuracy. A similar approach can be adopted for the transcription of DAT.

1.1.4. Digitisation

Encoding and Format

Sampling frequency (time resolution)

Analogue to digital conversion of media must be performed by adopting sampling rates suitable to ensure accurate reproduction of the whole audio frequency range of the original media.
It has to be taken into account that production needs demand for a single sample rate for all the available media. Different data encoding should require preliminary data conversion for editing and broadcasting.

Samples depth (amplitude resolution)

The digital encoding of audio contents must be performed by adopting a sample depth (number of bits per sample) suitable to render the original contents dynamic.
As it is not infrequent to find materials recorded with unusual reference levels, it can be convenient to apply A/D conversion supporting a sample depth sufficiently higher than the estimated dynamic range of the media, so that the transcription can be performed without prior level adjustment on a media basis.

Loss less data compression

In order to ensure accuracy to the original storage of digitized media, it is mandatory to avoid any type of compression leading to lost of information (e.g., mpeg compression). Loss less compression (i.e., totally reversible) algorithms could be applied, provided that compression/decompression times are compatible with production and broadcasting needs and that the decompression method be fully documented and possibly standardised.

Audio file format

BWF (EBU Broadcast Wave Format)[2] format is required for audio files disc storage. The EBU is also working on standards and recommendations[3] for best practice in audio archive preservation, as is IASA (International Association of Sound Archives[4]).

Vinyl

(33 and 45 RPM)

Vinyl material is generally not in critical condition, with the exception of older 78 RPM. The need for digitisation resides mainly in the fruition problems connected to the support: good quality turntables are not very common anymore and the supports can be damaged during normal operation.

---

[2] EBU Standard N22: The Broadcast Wave Format: A format for the interchange of audio data files in broadcasting. http://www.ebu.ch/pmc_bwf.html

[3] EBU R105: Digitisation of Programme material in Radio Archives. http://www.ebu.ch/pmc_r105-2001.pdf

[4] IASA (International Association of Sound Archives) TC03: The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy. http://www.llgc.org.uk/iasa/iasa0013.htm

Before transcription, the media must be inspected 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).

The playback of discs must be carried out using broadcast quality turntables. Even if few early recordings of the 50' were made in the absence of a standard practice, the vast majority of the materials is recorded using the RIAA emphasis. Therefore, in absence of specific information, the audio preamplifier must be equipped with the appropriate RIAA de-emphasis circuit. It is important that the stylus be replaced at regular intervals following the manufacturer's prescriptions.

(78 RPM)
The transcription of these materials presents several critical issues, including:

- fragility of the media
- heterogeneous emphasis characteristics
- heterogeneous cutting parameters
- scarceness of information
- difficulty in finding on the market suitable turntables and styluses.

Therefore, it is probably convenient to digitise this material according to the Expert transcription model.

Given the great deal of variation in the characteristics of the emphasis curves adopted by disc manufacturers in absence of a common practice or standard, it is probably convenient to digitise all the material without applying any de-emphasis. It will then be possible to apply de-emphasis on the digital domain, with the advantage of non-destructive editing, and the possibility of testing different solutions on selected items without impacting on the massive process of digitisation.
Even if the effective turning speed during recording could have been significantly different from the nominal value (of up to 10% according to some estimation), the transcription will be performed at exactly 78 RPM as speed corrections can be easily applied to the digital signal.
A subtler problem is represented by the choice of the most appropriate stylus. During the years, up to 6 different types of stylus have been adopted as specified in the following table:

periodo	Conical truncated stylus	Elliptical truncated stylus
Pre 1920	.0040''	.0040'' x .0012''
1920-1939	.0035''	.0035'' x .0012''
1939-1966	.0028''	.0028'' x .0009'

The period indications are purely indicative, as some disc manufacturers can have used different configurations.

Quality Control

Technical metadata collection

Massive transcription systems shall adopt specific-to-the-media algorithms for collecting technical metadata.
Also **expert transcription** systems should extract as far as possible technical metadata collections in order to support manual checking.

Automatic quality check of audio transcription chains

In principle, automatic quality check should produce a report from which a classification of the recording can be made. The classification should be based on the following aspects:

- Acceptability of the transcription, that is adherence of the digitised signal to the reproduced one
- Estimation of the overall quality of the transcribed signal

If the first test fails, the transcription cannot be considered correct and needs to be repeated, possibly after some maintenance intervention. The second test is more difficult to interpret, as we cannot compare the reproduced signal to what is actually written on the media. Therefore, there is a need for an inference method that allows us to draw some conclusion, followed by appropriate actions, from the analysis of the measurements logged by the automatic check system.

Categories of errors:

- Systematic errors
- Media defects
- Reproduction errors

Systematic errors: errors that can be systematically observed irrespectively from the media reproduced. They are generally caused by a fault in some part of the chain, e.g., dirty heads, wrong head alignment. Possible solution: substitute/repair faulty equipment.

Media defects: the surface of the media is damaged. Simply replaying the media cannot solve these problems, e.g., tape drop outs, disk scratches. Possible solution: mechanical conditioning of the media.

Reproduction errors: isolated, transient errors in some part of the chain, where replaying the media would potentially correct the error condition, e.g., spike or interference with other equipment, edge condition of the media. Possible solution: repeat the transcription.

The classification of the transcription errors according to the above schema can be based on the observation of the artefacts and parameters listed in the following non-exhaustive list:

Types of artefacts to detect:

- local artefacts
    - clicks
    - scratches
    - drop outs
    - saturation
- global artefacts
    - broadband noise
    - distortion
    - wow & flutter
    - speed fluctuations

Parameters that can be usefully logged:

- peak signal level (overall and for each channel)
- noise level (during silence)
- RMS signal level (overall and for each channel)
- bandwidth
- silence duration
- noise duration (not classified as silence)
- stereo/mono - phase correlation

In particular, for types of media where a reduced basic quality has to be expected, statistics on the normative values for the above parameters can be utilised by the classification module in order to reduce the occurrence of false problem detection.
The transcription process should be organised to ensure material of similar period and brands of media are transcribed in one batch.

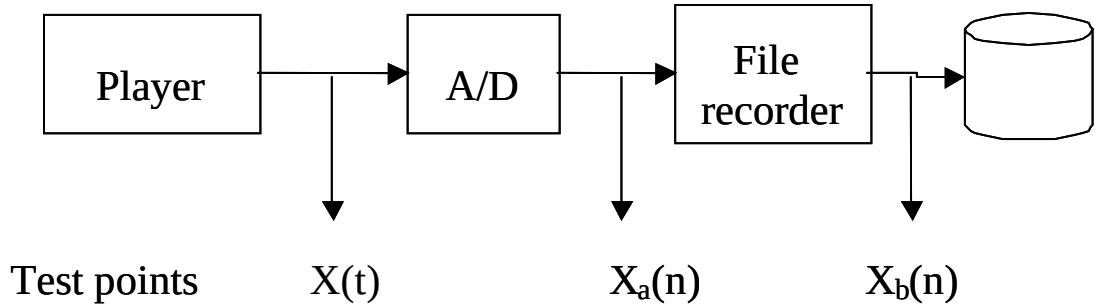

Figure 5.1 Diagram of Basic Audio Transcription

The diagram above shows a basic reference chain for audio transcription of analogue media like 1/4" tapes and vinyl records. Three possible observation points are identified:
X(t) is the played back signal, that could differ from the signal stored on the media in case of play back problems. Unfortunately, as the stored signal cannot be directly observed, problems on the play back can only be inferred but not detected by direct comparison.
X_a(n) is the digitised version of **X(t)**. A direct comparison between **X(t)** and **X_a(n)** is not possible as they are defined in two incompatible domains. It can be expected, however, that no major problems should occur during conversion that require a comparison to be detected. A comparison of the instantaneous measures of peak signal levels and phase correlation at **X(t)** and **X_a(n)** can give sufficient information about possible misalignments between the analogue and digital domain. Therefore, we can, as a first approximation, consider **X_a(n)** as a faithful digital representation of **X(t)**, after checking that some kind of artefacts, typical of the A/D process, are not observed on **X_a(n)** (e.g., clipping, digital clicks, muting).
X_b(n) is the recorded version of **X_a(n)**, and should be, in absence of errors, bit to bit equal to the latter.

¼" Technical metadata

The main defects that can be automatically detected on ¼" tapes are:

- **Decomposition of Pigment Binding Materials of Magnetic Tapes**. The main risk to magnetic tapes results from the instability of the chemical binders. A considerable number of audiotapes, especially those produced during the 1970's and 80's, are suffering from pigment binder hydrolysis. Atmospheric moisture is absorbed by the pigment binder causing the polymer to hydrolyse and loose its binding properties. Tapes of this kind deposit a smear of magnetic particles onto the replay heads. This clogs the heads and swiftly makes the tape unreadable. This problem has been found especially in hot and humid areas where many tapes do not last longer than a few years. [5]
  This problem can be detected by monitoring signal mutes during playback
- **Tape improperly rewound.** Monitoring signal fall and peaks during playback can detect this problem.

[5] UNESCO web site: Safeguarding Documentary Heritage.
http://www.unesco.org/webworld/virtual_exhibit/safeguarding/expo01.html

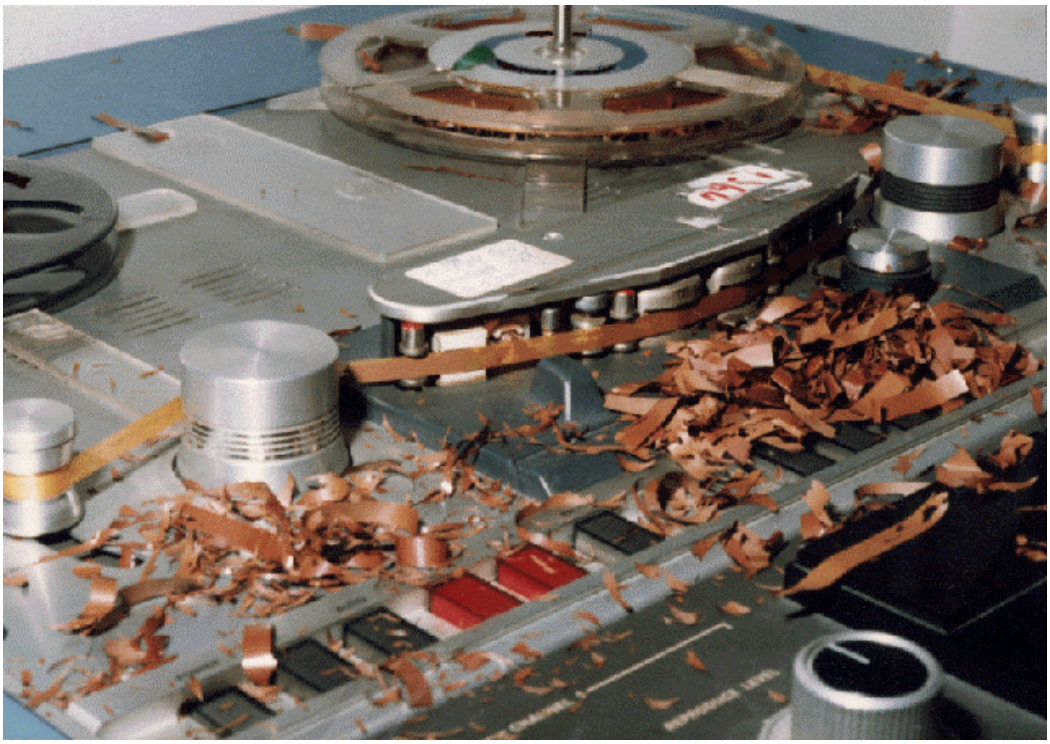

Figure 5.2. Decomposition of Pigment Binding Materials of Magnetic Tapes (Photo: Phonogrammarchiv, Austrian Academy of Sciences)

Vinyl Technical metadata

Vinyl is considered to be the most stable media ever used for storage of analogue sound recordings. Nevertheless, several problems can adversely affect the reproduction. Some of them can be removed or minimised by means of a proper mechanical conditioning of the disc prior to or during the reproduction, others, caused by irreversible damage of the surface, can only be removed, or at least concealed, applying some kind of electronic manipulation to the reproduced signal. This process, called restoration, is generally very time consuming and requires specialised skills not only on the general technical matter of vinyl restoration, but also on the specific content. Restoration is therefore, a process that cannot be successfully applied blindly to a massive automatic (or semi-automatic) transcription chain of vinyl materials. It is important that during transcription a report be generated that summarises the results of the transcription. This report allows the technical quality of the material to be evaluated without listening to the digitised signal. Material can thus be chosen for broadcasting or for restoration. This information, even if not 100% accurate, is important to plan the activity required for re-use of content leading – which is after all the final goal of the transcription process.

Metadata

Metadata collection – the content

Prior to transcription, all existing content-related metadata should be collected; examples are data extracted from a legacy data base or kept in printed form. This material must be accessible to the transcription operator.

During transcription, all the available information must be collected, translated into a suitable digital format and attached to the transcription for subsequent ingestion into the new digital archive. In any case, all the transcribed material, including essence and metadata sets, shall be uniquely identified according to an established standard convention.[6]

---

[6] SMPTE 330m-2000: SMPTE STANDARD for television – Unique Material Identifier (UMID). http://www.smpte-ra.org/s330mex.html

This step is vital to maintain the relationship between the digitised essence and the original media and between essence and metadata.
This policy also simplifies the workflow when the transcription is performed by third parties.

Metadata collection – the media

During transcription, it is necessary to collect and organise as much information as possible about media and related folder/box, that is relevant to the subsequent processing phases.
Optical scanning should be used, optionally followed by OCR, to efficiently capture information from media and box surfaces.

Metadata storage

Even though BWF can be used to convey a limited set of metadata, it is preferable to store metadata in dedicated files, in order to avoid the limitations implied by BWF, that was originally conceived to hold only essence and vital metadata. XML- based formats are far more flexible and efficient for metadata storage if it is necessary to perform updates to a database. The coherence between essence and metadata should be enforced by the use of unique identifiers and optionally by means of standard file containers as Windows ZIP or Unix tar.

Automation

Process automation shall be enforced at the maximum extent, to save significant resources in terms of man-power.
Process automation shall be constrained by the requirements specifying careful handling of the media.
Process automation shall also be constrained by the requirements concerning unique media and contents identification.

1.1.5. Preservation Media

Backup copy

Transcription execution requires huge amount of human and technical resources. In order to reduce risks due to future non readability of media used to store audio files, it is required that at least two copies are contextually produced, regardless of the adopted technology.

Preservation media

Preservation shall be on magnetic or optical media, with a minimum capacity of 30 GBytes. It is desirable to use linear tapes, since helical scan tapes are weaker. Optical media need to be very carefully written, using media and media-writers that are compatible, and that have been checked for low error-rate output. Various standards exist in the information technology and archive industry for best practice in data archiving, as cited in the major study on TV Preservation Status prepared by the US Library of Congress.[7]

[7] Television and Video Preservation 1997: A report on the current state of American television and video preservation: report of the Librarian of Congress. October 1997. ISBN 0-8444-0946-4. http://lcweb.loc.gov/film/tvstudy.html

6. Cost drivers

Individual high-cost elements of preservation activity will be described below (section 6.2). Archives were also questioned about overall cost factors. The following factors were the most common:

6.1. Manual handling

The main manual elements are operator time, time to find material, packaging and transport and tracking of material.

Efficient preservation requires an assembly-line process: preservation work on-site, in a dedicated 'preservation suite', with fulltime staff doing shift work to get maximum return on capital investment.

In addition, the true cost must be considered of all staff time on all associated processes of material identification and movement. **The clear need is to eliminate or automate ALL manual handling** except the essential component of the skilled technician monitoring the transfer process.

An important contribution to elimination of manual handling is use of mass storage. Although video files remain too large for storage in full quality on servers, audio should be stored on a server to eliminate all physical steps between digitisation and production of the final output(s) of the preservation process. The IT control system described in Section 4.4.2 should include media as well as metadata storage and management. This principle also applies to the browse-quality video files generated in film and video preservation work.

6.2. Metadata handling

Finding, converting, quality checking and updating of metadata is seen as taking up to 40% of the effort of preservation work. Much of this could be automated if standardised metadata were available. Even though such is not the case, experience has shown that use of a dedicated IT system to move metadata and control the 'assembly line' can be a significant savings – certainly cutting metadata and associated handling costs by 50%.

Archives need a generic IT system, capable of being adapted to local requirements, as the control centre of preservation work. Important requirements of this overall system are:

- Import/export of metadata – the ability to take in whatever exists in legacy catalogues, and other data sources, and then produce new metadata, reflecting the updated status of the media, and export back to the legacy (or other) system
- Support checking and updating of metadata, including automation wherever possible: data updating (from old format information to new), checking for missing or incomplete data, data format conversion as needed.
- Conversion to standardised metadata (such as Broadcast Wave Format descriptive header, Dublin Core 'discovery metadata', EBU P/META proposed standards, and alignment of technical metadata with SMPTE standards) where needed. This step is especially relevant to preservation projects that generate a datafile format, as all datafiles need header information, creation and attachment of a unique identifier. For efficient asset management, most large projects will also require that the file links to extended metadata in a conventional database.

7. Failures

Many archives identified media and equipment failures as significant cost drivers. However efficient the 'assembly line', if it doesn't actually work then for an item, the process reverts to extensive manual intervention, time delay, and loss of all efficiencies introduced by way of capital investment, specialist equipment and process design. There are two implications:

> **Minimisation of failure** – once a process is efficient, a failure becomes far more expensive, in proportion. This needs to be considered in process design, as the cost of special equipment may be justified most clearly in terms of the prevented cost of re-handling of failures. Because archive material in need of preservation has much higher failure rates than relatively newer material, there is a strong case for special equipment and special procedures (cleaning, temperature treatment) to prevent transfer failure. Attention must also be given to frequency of equipment maintenance, which may have to be increased when dealing with older material – material which sheds more-than-average amounts of oxide.
>
> **Handling procedure** – failures are a fact of life and need to be built into the overall system workflow. In particular, metadata should still be handled efficiently. Process design must ensure that a failure does not impede overall workflow.

There are two basic types of failure: the item, and the player.

> **Item failure**: an item can have a fault that requires special attention, or the media player itself can malfunction (usually because of the effect of one or more problem items, generally tape with poor adhesion which causes head clog and related problems). Failure management is a particular issue for transfer suites with one operator handling multiple transfers (often four). The process should deal with a failure of one item without impeding processing of the remaining items. **The process control software should automate handling of failures just as it automates the normal procedure**: isolate the failed item from process control workflow, introduce a substitute item as quickly as possible, and deliver the failed item with full metadata to whatever element of the total workflow has been set up to deal with failures, and allow the "rescued failure" to be re-introduced into the workflow. What should NOT happen is that three (or more) transfer systems are idle while an operator makes repeated trials on a failure.
>
> **Equipment failure:** If a player is clogged or otherwise needs attention, the whole transfer of all items is compromised as the operator cannot both clean the faulty machine and monitor operation of other machines running in parallel. It may be cost effective to have an extra technician available, rather than stopping the other three or more transfers (or letting them run unmonitored) while the one technician deals with the fault. Total cost effectiveness depends upon the number of transfer suites in operation, and frequency of faults.

Analysis should also be made of costs and benefits of spare equipment. For very old formats such as 2" videotape it is difficult to get sufficient equipment for multi-station transfers, but for more recent formats it may be cost effective to have spare equipment available. This option would allow quickest return to working and maximum throughput.

7.1. Rights

Rights are seen as a major cost factor by most archives. There are several categories of rights cost for any item in an archive.

Identification:

- Whether an archive owns the rights
- Who else own rights, in addition to or instead of rights owned by the archive

Specification:

- What rights the archive owns
- What each additional rights owner actually owns
- Specifying what usage can be made or the material, under these 'archive rights', and at what cost

Clearance:

- Specifying the actual fees to be paid to each rights owner (or representative body) for a specific proposed usage

Most archives have to perform rights identification, and any uncertainties introduce extra work and cost for a preservation programme – as most preservation budgets exclude work on un-owned material.

Specification of rights is important to minimise clearance costs, and to support commercial exploitation, especially the prospect of new forms of commerce, specifically internet-based archive access and sales. These "specification rights costs" should be seen as an investment in improving the ease-of-use of archive material. However a strict business case for rights investment can only be made if there is a genuine prospect of extra sales. There is a 'chicken or the egg' problem here, with investment in rights information depending upon sales, and sales depending upon the better information.

The solution to the 'who invests in rights?' question is found at the overall business level. Typically the archive cannot make the business case, because the archive only gets 'service costs' – remuneration for cost of storage, cataloguing, and retrieval. They do not get a 'slice of the pie' of commercial fees from external customers, and they do not charge a commercial rate to internal customers.

At the overall broadcast business level, there is potential for improving access, improving programme making, increasing revenue, and decreasing rights clearance costs. These advances would come about through improved archive 'rights metadata', **adding specification information** to the existing (usually) identification data. One clear opportunity for adding this information is during preservation work, when the existing metadata is being accessed, checked and augmented.

Broadcasters depend upon the rights clearance process, which is slow and expensive if the archive has no specification data. The question to be examined is **the degree to which rights clearance effort could be reduced if effort went instead into adding rights specification data** to the archive metadata. If the specification data included costs for 'standard uses', which could possibly be done by a straightforward classification and coding system, then archive users could get an immediate estimate of usage costs for a particular item and a standard use. The formal clearance procedure would still need to be followed, in case a particular item has any unusual rights restrictions, but with full specification data the clearance procedure would be shorter and cheaper. The result would be clearer information to archive users, encouraging archive use – paid for by efficiencies in the clearance procedure based upon clear rights specifications. It does require vision and commitment at the overall business level to seize this opportunity.

8. Key Link Technology requirements

In addition to the various user requirement (or desires) listed in the previous sections, the following additional 'key links' were identified. Some of these relate to items previously mentioned, but are repeated here because they are issues repeatedly identified by the archives:

8.1. Missing or Unnecessary Links

8.1.1. Handling of A/B rolls

This is a film format for which alternating shots are on two separate rolls, which allowed for fade/dissolve effects between shots when making the final print – in the days before video editing and digital special effects. There is no equipment used in scanning / preservation / telecine conversion which specifically caters for this format.

8.1.2. Metadata flow

A very common problem is data that cannot electronically move from one system to another, but requires manual retyping. Standardisation of metadata and use of a generic 'process control system (1.6.5.2) are two main ways to eliminate data re-entry.

8.1.3. Signal monitoring

The manual effort in monitoring has already been identified as labour intensive, but archives specifically identified a 'missing link' of automated monitoring. Archivists did NOT want to use spot checks and statistical methods of quality control. They want every audio or video signal to be monitored during the entire transfer process, but monitored by automatic rather than by manual means.

8.2. Expensive Links

Overall cost drivers were discussed in Section 6. Specific 'key link' cost drivers were identified as being:

8.2.1. Realtime transfers

New media (CD-ROM, some videotape) can be written to at faster than real time, up to 4x. The ability to playback old material faster than realtime would have an immediate impact on preservation costs. This links to automation of monitoring, because human-based monitoring is difficult or impossible at faster than real time, but an automated monitor would have no such constraint.

8.2.2. Missing documentation

This situation causes expense because if legacy metadata is not there, essentially the cataloguing process has to be performed in its entirety if the standard of the catalogue is to be maintained. Storage of material (audio or browse video) on a server would allow the cataloguing staff to perform this task, rather than the technicians doing the transfer. The overall control programme should treat the missing metadata as a 'failure', just like missing media if a tape head clogs, and should supervise and automate the process of rectifying the metadata.

8.3. Opportunity Links

Preservation projects provided two specific new opportunities, which were identified by virtually all archives:

8.3.1. Online media

During preservation, compressed video and audio, and possibly full-quality audio, should be produced at the same time as producing the full-quality item on videotape or datatape (or CD, DVD or datatape in the case of audio). This idea requires management of multiple versions of the same inherent item, but broadcast archives have always dealt with this concept – as in film negatives and prints, and with

video on professional formats and on VHS for viewing. Online media is seen to have multiple efficiencies.

- Easier, quicker, shorter documentation if cataloguers can have video at their PC
- electronic distribution to archive users (at least to key users in the first instance)
- less distribution of unwanted media if users can preview
- rapid programme production with frame-accurate desktop selection of material
- provides the technical basis for electronic commerce for archive material

8.3.2. Enhanced metadata

While material is being digitised and held (at least temporarily) on a server, a variety of additional stages of processing are possible – at lowest possible cost since the material handling, digitisation and server storage will all take place anyway. Some of these stages are necessary to fully exploit online media, specifically:

- Introduction of proper timecode into the catalogue and viewing copies, which is needed to make a frame-accurate pre-edit from the viewing material
- Shot-level description / indexing – which is needed for item rather than programme-based retrieval. Archive users will not want to (or be able to, with current datarate restrictions) download entire programmes for the purpose of evaluating specific shots
- Keyframe extraction – 'pictures in the catalogue'. The use of keyframes as part of the catalogue eliminates the need for descriptive text, providing better information to the user, at less effort to the archive

9. Current Business of the Archive

9.1. Inputs

Archives exist as a safe place to hold the valuable content produced by broadcasters, for the purposes of re-use, research, and for cultural heritage (this last function is commonly shared with national cultural archives).

The following table gives the volume of material taken into the archives in the survey, every year. In most cases, the input represents the most of the TV output, selected radio output, and selected non-broadcast material (rushes, stockshots). It can be readily seen that the volume of material flowing into broadcast archives is large – especially in relation to the volume of preservation work currently undertaken (as shown in Section 4.4).

Two conclusions may be drawn:

- A huge amount of very valuable material comes into these archives every year
- The current scale of preservation work is not in any way keeping up with the scale of the input

Table 9.1.1 Survey volume of annual intake to archives

Archive:	YLE	NAA	ORF	INA	RAI	BBC	TRT	SVT	SWR
Input									
Audio - own material	7500	11400	10000	25000	9500	38000	1200	0	25000
Commercial recordings		11000						0	
video	16700	12500	75000	35000	80 000	105000	4000	9600	30000
film	marginal	670	350		little	5000		2400	300
items or hours?	hours	hours	items	hours	items	items	items	items	items

9.2. Outputs

Broadcast archives are heavily used – up to 25% of holdings are circulated EVERY YEAR. The survey showed that about 75% of the TV circulation was in browse quality (VHS tapes), indicating that about 3 items are "browsed" for every one that is seriously considered for broadcast re-use. It is difficult to verify how many of the high-quality items issued from the archives actually end up being broadcast (at least in part). The estimate from the archivists was that at least one in four of the broadcast quality items circulated would be transmitted, though often just a small clip would be used, not the whole 'shelf-item'.

Table 9.2.1 Circulation o material figures from archives

Archive	YLE	NAA	ORF	INA	RAI	BBC	TRT	SVT	SWR
Issues / year									
Video	74000	75000	265000	40000	370 000	600 000	15000	45000	Roughly 80000
Audio		15000			95 000	130 000			

9.3. Income and Costs

The following table gives the basic sources of funding for the archives – as an annual grant for the overall service of the archive, or Pay As You Use income (basically charges per item for archive services). Other income is mainly external sales, which is a small figure because a separate, commercial arm of the broadcaster typically handles the main sales of archive material

Table 9.3.1 Basic sources of funding for European archives

Archive	YLE	NAA	ORF	INA	RAI	BBC	TRT	SVT	SWR
3.3 Income & Cost									
annual grant	99	60	99	60	50	74		91	
other	1	5	1		10	5			
PAYU		35		40	40	21		9	

The following table breaks down the total archive budget into major categories, though from the results it is evident that these figures say more about the accounting practices in the broadcaster, than about the archive practices themselves.

Table 9.3.2 Breakdown of spending within archive budgets

Archive	YLE	NAA	ORF	INA	RAI	BBC	TRT	SVT	SWR
Cost breakdown as a percentage of total budget					(Radio)				No perm staff!
Facilities	22%		free	31%	39	16			50
Staff	56%		95%	51%	38	48		70	35 (temps ONLY)
Transport			1%	incl. in facilities	10				7
Media			free	3%	11				5
Overheads	22%		4%	15%	2	36			3

9.4. Physical Storage and Handling

The figures for archive input and output add up to, in most cases, several hundred (or more) tape movements per day. This can be viewed as a lot of physical handling which electronic mass storage and delivery could replace. Alternatively, the current practice could be viewed as a cost effective method of handling a bandwidth problem – by delivery van instead of by computer network. The following information was captured, relating to storage, handling and transport costs.

Table 9.4.1 Storage space, handling and transport figures

Archive	YLE	NAA	ORF	INA	RAI	BBC	TRT	SVT	SWR
Shelves, m	20000	38000	70000 in use	70000	31000	85000	3290	11300	
Annual cost /metre of shelf; Euro				38	90	24			
Staff working in stores	7	14	20	30		48		7	
Transport costs	320k €/yr.	50k €/yr.	1 person + car	243.5 excluding preserv'n transport	radio: 70k € / yr.	160k € for vans; 10k for press wk		<1% of total	7% of budget

The above table does not represent the total cost of physical handling, because it does not include acquisition costs, though all-electronic systems would still have an acquisition process. Acquisition should be seen as an area of increasing automation, as required information is "captured at source" and delivered to the archive along with the programme content.

10. Conclusions

The specific conclusions of this report are listed, with supporting evidence, in the individual sections of the report – and the major points are given in the executive summary.

One purpose of this study was to find out what PRESTO should do in order to reduce the cost of broadcast archive preservation. Another purpose was to collect evidence about current costs, to provide a baseline for comparison against the achievements of the PRESTO project.

The first purpose has been fulfilled: this report presents a range of process and key-link technology requirements. The second issue is less clear, because a major finding of the study is the difficulty in determining actual costs: actual cost of ownership, over a lifecycle, of an archive item; actual cost of preservation of an archive item. However ballpark costs have been established, and important criteria for establishing preservation costs have been defined:

> ***Cost per use*:** determining expected usage and judging a preservation investment in terms of its return each time an item (that otherwise would be lost or unusable) is accessed. This model is necessary to give a place for an economic judgement of investment into new archive technology – to decide whether (or not) to make an investment into a technology offering increased access, which could ultimately repay that investment through lower cost-per-use.
>
> ***Lifecycle costs (cost of ownership*):** determining cost for every aspect of the creating, storage, management, use, and maintenance and replacement of an item, and judging options in preservation work according to this total life cycle, NOT just according to minimum transfer cost.

The final major outcome of the study is the process analysis of preservation work, establishing clearly the fact that preservation work is NOT just a lot of single-item transfers. Single-item work, as accomplished every day in archives, does not scale up to preservation projects on two counts:

> ***Cost*:** most of the costs for single-item work are hidden: the staff and systems are already there, sometimes the equipment and other facilities are also there, and it can look like a transfer is 'practically free', or 'just the cost of the media', or 'just the transfer operator's time'. The total cost of the process is unknown (and ignored), and certainly the lifecycle cost of the new item is ignored.
>
> ***Efficiency*:** one-off transfers don't stretch the capacities of existing systems, which often only work with considerable manual input (for finding the material, labelling, checking, re-cataloguing). This is the work that goes on every day anyway, by the 'staff that are already there', and the odd format-transfer is just one more item in a days work. Preservation projects are inherently mass transfers, not one-off. It is in preservation work that the inefficiency of all the manual steps, and the inadequacy of standard computer systems, is made evident. **The only way to achieve efficiency in preservation work is through developing an assembly-line approach**, with every possible source of efficiency eliminated. It is also true that it is only preservation work that has a sufficient volume of identical material to justify such an assembly line, as the ordinary technical work of an archive is far too varied to offer this 'efficiency opportunity'.

The implementation stage of PRESTO will establish such assembly lines, optimised for audio, video and film, and combine them with key-link technical solutions to set a standard for best practice for broadcast archive preservation.

11. Appendix I – Survey Participants

SWR: Südwestrundfunk, Hans-Bredow-Straße, D-76530 Baden-Baden, Germany http://www.swr-online.de

NRK: Norwegian Broadcast Archive, Norway

SVT: Swedish TV Archive, SE-105, Oxenstiernsgatan 20,10 Stockholm, Sweden http://www.svt.se

YLE: Finnish Radio, PL 20, 00024 Yleisradio, Finland http://www.yle.fi

ORF: Österreichischer Rundfunk, Würzburggasse, 30,A-1136 Wien http://www.orf.at

NAA: Netherlands Audio-visual Archive, PO BOX 1060, NL – 1200, Hilversum, Holland

TRT: Turkish Broadcast Archive, PK 7606442, Yenisehir, Ankara, Turkey

BBC: British Broadcasting Corporation, Room G0O7, Broadcasting House, Portland Place, London W1A 1AA http://www.bbc.co.uk

RAI: RadioTelevisione Italiana S.P.A. Viale Mazzini 14, 00195, ROMA

INA: Insititut National De L'Audioviseul. 4, Avenue de L' Europe, 94366, Bry Sur Marne, France.

12. Appendix II – Presto Questionnaire

PRESENT STATE OF COLLECTIONS AND PLANS FOR PRESERVATION

Survey of Preservation Needs

- Archive Services
- Archive Holdings
- Preservation Technology

1. Scope and General Instructions

The present

- 1- what broadcast archives do: their place in the business
- 2- what they hold
- 3- current preservation technology
  - Processes and costs

The future

- 4- new services from new holdings
- 5- requirements for preservation
  - Processes and technology

What to do with this document: This document is the formal method for gathering information about the preservation of broadcast archives. It will be supported by interviews or visits as required. The purpose of this document is to specify the information sought, and provide a framework for its capture. The PRESTO project will 'talk you through' the document during the interview stage. **Please go through the document to understand** what we want to know, and **to collect** as much as you can of the required information. Please complete the questionnaire if possible. Fill in what you can, collect your own relevant documentation, and then the next step is a formal interview to fill in the gaps.

A PowerPoint presentation (Presto Questionnaire Overview) is available which summarises the purpose and methods of this phase of the PRESTO project. A summary of the findings will also be available as a PowerPoint presentation. Detailed findings will be collated, used to determine the activity of the PRESTO project and made publicly available. All information will be made available to all PRESTO partners. Company confidential information should not be divulged. Please discuss with the workpackage leaders (BBC and INA) any information that is important to the goals of PRESTO but needs to remain confidential.

Parts of this document reserved for answers are expandable; don't hesitate to give as many details or comments as you wish.

2. Organisation Details

Organisation	
Archive	
Year of foundation	
Address	
Phone	
Fax	
Email	
Who is responsible for the storage of archive collections on film?	*Name phone fax email*
Who is responsible for the storage of archive collections on video?	*Name phone fax email*
Other details about your organisation	

3. Services –

Preservation is about continuing – and expanding - the service provided by an archive to the broadcast industry. This section of the questionnaire documents the basic services provided by your archive.

3.1. Input: What comes into the archive, and why, and what effort is used to acquire and document

Please list the amount of new film, audio and video material entering the archive annually – please answer separately by physical format.

Format	**Hours or items acquisitioned annually**
Videotape (specify formats)	
Audio: ¼" tape	
CD	
Other	
Film: 16 mm	
35 mm	

Also please list the source of the material – by broadcaster if your archive serves more than one company, and by major category (eg news and current affairs, sport, arts and drama, light entertainment) if known. If you have both radio and television holdings, then please answer separately for each.

Please supply documentation of your input or selection or retention policy if available.

Also please give information about the legal basis for the work of your archive, with particular reference to any specification of the required content of the archive.

Finally, new archive content requires documentation. What percentage of the final documentation is provided to you -- with the content – as distinct from the **documentation which you create or reformat yourselves**? **How much effort** do you invest in documentation? **How long** between between time of broadcast and when material reaches the archive, and how long to acquisition and catalogue? (Please answer this any way you can!) Again, separate answers for radio and television if available, and any additional detail if available.

3.2. Outputs

Who uses the archive? how many people, who are they (by category) and where are they?

How much material goes out of the archive every year? Please answer separately by physical format**.**

How does the material get from the archive to the users?

Do you have lending copies, or do you make a new copy for the user? If you do both, please indicate relative volumes.

How do people access the archive? Catalogue at their workplace? Telephone call to the archive?

Please describe the catalogue – printed examples of catalogue entries would be very useful.

How long does it take to get archive material? How long to:
Do the research _____
See the material (in at least browse quality)_____
Acquire broadcast-quality material. _____

What information do you supply with the material – ownership, rights, transmission history, programme production information? Please give any available details.

Do you currently re use archives stored on old video and audio format or film for new broadcasting or insert in new programmes ?

Programmes for new broadcasting (number of hours or items per year)? From which formats?
Short extracts for new programmes (number of hours or items per year)? From which formats?

What video and audio tape formats do you deliver to your customers?

3.3. Income & Costs

Preservation work will depend upon succesful funding, which in most archives requires a business case justification. A major goal of PRESTO is to provide solid information about costs and benefits, to help 'make the case' for preservation. This goal requires some basic information about current funding models and costs.

Please give information in terms of total income rather than actual money, if possible. This allows combining results across archives.

Income breakdown (in % of total income)

- Annual grant _____
- Usage-based income _____
- Other _____

If available, please provide a breakdown of usage-based income according to main sources of income (main customers) or main service provided (as appropriate; both if possible)

Service or Customer	Percentage of Total Income

Cost breakdown (in % of total income)

- Facilities _____
- Staff _____
- Transport _____
- Media _____
- Overheads _____

Detailed Costs for Storage and Technical Processes

- How many meters/ feet of shelves are in your storage rooms?

- Do you have an estimate of the current complete cost of a linear meter of storage on the shelves? (including cost of the buildings, human costs…)

- What are the current storage conditions in your storage rooms? (temperature, relative humidity, air filtering…)

- Do you think that storage is a cost driver in the overall archive process ?

- Have you already considered solutions to improve storage costs?

- If yes, what are the considered solutions? (change location, reduce volume of cassettes , selection, elimination of redundancy, use of compression, use of data tapes versus videotapes, large automated libraries, ….)

3.4. Movements of tapes and films

- Do you often need to move tapes to other locations to be processed ?

- Do you have an idea of the average cost of transportation in a year?

- What is the daily frequency of movements of tapes (average number of tapes coming in and going out)

- How many people do you employ in your archive stores?

3.5. Cleaning and acclimatisation

- Do you use any cleaning process before transfers? How do you decide when to use a cleaning process?

- Which method or what kind of cleaning device do you use? (for film and video)

- What is the average time to clean each format tape or film ? _____

- Do you have an idea of costs of this step?

- Do you include in your process an acclimatisation time for tapes before transfer? How long?

4. Current status of the Archives

4.1. Elements of size of current holdings

FILM	Meters/feet	Shelf length	Cans/items	Age and Condition	Recovery
Total film holdings					
35mm negative					
16mm negative					
16 mm reversal (all)					
16 mm reversal edited with cellophane tape					
35mm print					
16mm print					
35mm SEPMAG					
16mm SEPMAG					
other					

VIDEO	Hours	Shelf length	Tapes/items	Age and Condition	Recovery
total videotape holdings					
2inch B/W					
2inch PAL					
2inch SECAM					
2inch (other standard)					
1inch B PAL					
1inch B SECAM					
1inch C PAL					
VIDEO continued	**Hours**	**Shelf length**	**Tapes/items**	**Age and Condition**	**Recovery**
1inch C SECAM					
1inch (other standard)					
¾" Umatic LB PAL					
¾"UmaticLB SECAM					
¾"Umatic H PAL					
¾" Umatic H SECAM					
¾" Umatic SP PAL					
½ inch BETACAM					
BETACAM SP					
½ inch MII					
VHS					

Others analogue					
D1					
D2					
D3					
AMPEX DCT					
D5					
Digital BETACAM					
DVCpro					
DVCAM					

AUDIO	**Hours**	**Shelf length**	**Tapes/items**	**Age and Condition**	**Recovery**
Total audio holdings					
wax cylinders					
shellac					
78 RPM vinyl					
LP vinyl					
¼ inch tapes					
cassettes					
DAT					
CD					
minidisc					
Others					

4.2. Content

- What are the main types of programme recorded on each format? Fiction, documentary, news, shots, edited masters, broadcast tapes… do you have ratios on each format (for example on 1" tapes x% of fiction, y% of documentaries…). Same question for audio.

- What is the average duration of each type of programme?

- Is there redundancy in the holdings? (does the same programme exist in more than one sample?) If it is the case is it on the same tape format or is the same programme on different formats. What is the ratio of programmes that are on more than one format?

Comments on recording status (free comments)

5. PRESERVATION PROCESS

5.1. Preservation process (general information)

- Do you have a systematic transfer plan for film or old video and audio formats?

Film holdings	Tapes or films Hours per year	Tapes or films hours remaining to transfer
35mm negative		
16mm negative		
16 mm reversal (all)		
16 mm reversal edited with scotch tape		
35mm print		
16mmprint		
35mm SEPMAG		
16mm SEPMAG		
other		
Analogue Video holdings		
2inchB/W		
2inch PAL		
2inch SECAM		
2inch (other standard)		
1inchB PAL		
1inch B SECAM		
1inch C PAL		
1inch C SECAM		
1inch (other standard)		
¾ inchUmatic LB PAL		
¾ inch Umatic LB SECAM		
¾ inch Umatic H PAL		
¾ inch Umatic H SECAM		
¾ inch Umatic SP PAL		
½ inch BETACAM		
BETACAM SP		
½ inch MII		
VHS		
Others		
Audio holdings		
¼ inch tapes		
78 RPM vinyl		
LP vinyl		
cassettes		
DAT		
CD		
minidisc		
Others		

- Do you use your own facilities for preservation? If not who does the job? Please give details.

- What is your policy (if any) for transferring film to video or data ?

- How do you manage old tapes already transferred onto digital format?

- When the preservation plan is completed, how do you think you will manage old tapes (transfer to another location cheaper, destruction,wait…)

- For each format what is the average cost of transfer (or expected cost) per hour of programme ? For each format give the main cost driver.

Film holdings	Average cost per hour	Main cost driver (with %)
35mm negative		
16mm negative		
16 mm reversal (all)		
16 mm reversal edited with scotch tape		
35mm print		
16mmprint		
35mm SEPMAG		
16mm SEPMAG		
other		
Analogue Video holdings		
2inch B/W		
2inch PAL		
2inch SECAM		
2inch (other standard)		
1inchB PAL		
1inch B SECAM		
1inch C PAL		
1inch C SECAM		
1inch (other standard)		
¾ inchUmatic LB PAL		
¾ inch Umatic LB SECAM		
¾ inch Umatic H PAL		
¾ inch Umatic H SECAM		
¾ inch Umatic SP PAL		
½ inch BETACAM		
BETACAM SP		
½ inch MII		
VHS		
Others		
Audio holdings		
¼ inch tapes		
78 RPM vinyl		
LP vinyl		
cassettes		
DAT		
CD		
minidisc		
Others		

- What are you currently considering for formats not already transferred ?

- Where do you estimate are the costs drivers (difficulties or bottlenecks that generate costs) in your overall preservation process ?

- How do you estimate are shared the costs in % between technical preservation work, quality control, metadata management, media & storage costs, media handling, programme management, other (Each line should add up to 100%)

	technical work	quality control	metadata	media & storage	media handling	programme management	capital & maintenance	other (specify)
Film								
2 Inch								
1 Inch								
3/4								
VHS								
other								

5.2. Preservation process (technical information and costs)

- For each format, give a description of your current transfer chain for formats whose preservation has already began ? (type of equipment, devices quality control chain, labelling and metadata recovery and generation…human needs)

- What is the capital cost of a typical transfer chain for each of these formats?

- What is the cost of in use of a typical transfer chain (including human costs ,maintenance costs)

5.3. Play back of old analogue tapes

- What are the trade marks and types of VTRs you use ?

- Why? (better playback than others , existing , same kind of device which made recording …)

- What are the major technical problems encountered while playing back old formats?

- What are the expected improvements in this area?

5.4. Processing

- What kind of equipment do you currently use to decode, synchronise, apply corrections, convert to digital, encode…

- Is there any specific reason to use them?

- What are the technical problems related to this process?

- What are the expected improvements in this area?

5.5. Digital format for preservation

- What are the digital formats in use in your organisation for

High end production	
News production	
broadcasting	
Archive preservation	
Archive delivery	
Viewing/browsing	

- Give more details on current digital format(s)you have chose for preservation?

- What are the main reasons to these choices ? cost , choice of broadcasters, changes nothing in the data base …?

- How many digital copies of the same programme do you generate? Simultaneously? Sequentially?

- Do you generate different files on different formats for different uses (archive, communication, browsing…)

5.6. Storage of digitally transferred programmes

- How do you store digitally transferred programmes (on the shelf, or in automated libraries eg robotics)?

- What are the costs of linear meter of digital tapes shelves you use?

5.7. Technical data (technical metadata)

- What are the original data or identification available with old tapes? (Number, technical sheet, standard, time codes, bar codes, labels..)

- What are the data that are essential or interesting to keep from old recordings? Specify for each format.

- What are the new data to generate and store with or related to new recordings?

5.8. Data held on the library database

- What are the library functions involved in the transfer process?

- Are there any discrepancies between catalogue and actual holdings?

- Give a list or a short description of these discrepancies?

5.9. Quality control

- How is your quality control performed today?

- Which parameters do you control during transfer?

- How do you control them?

- What parameters do you control after transfer on the digital version of the programme ?

- What is the average cost of quality control ?

- What should be improved in quality check?

- What are indications (data) that should be kept from quality control?

- Do you think that improving quality control is saving costs ?

- What do you think that should be improved in quality control (missing or time consuming or …)

6. Delivery and communication

6.1. General

- How many programmes do you deliver to your customers (sell or rebroadcast) by year?

- How many extracts do you deliver to your customers by year?

- What is the average extract duration (minimum, maximum?)

- What is the ratio (amount) of extracts from news and from long programs fiction entertainment, magazines…)

6.2. Technical

- Have you already considered a digital format for delivery (offline or online)?

- Have you already considered a digital format for browsing (offline or online)?

7. Rights

- Describe your position regarding rights? (owner of the rights, delegate, conditions of rights

- Do you have a software system or a data base to process the rights?

- Is your rights data base independent or connected or merged with the descriptive data base and the holdings data base?

- Is rights examination manual or automated?

- What is the average time to examine rights for fiction, documentary, entertainment, magazines, news, extracts on each type?

- What is the average cost of rights examination per programme, per extract?

- What is the average cost of rights per sold programme, per sold extract?

- Does rights examination slow down or prevent fast delivery of programmes?

- How many people do you employ for rights examination?

- Do you estimate that rights examination is a major cost driver?

- Do you estimate that the rights examination should be tied to the preservation process? If so, for what reason: in order to save money later, to allow new forms of trading, or what other reasons?

8. Missing Links in Technology and Workflow

This section lists the areas where PRESTO will identify costly or ineffective processes, and develop better alternatives. Please indicate which areas you already know to be a problem, and why. This section is not easy to reduce to simple answers, so these questions are included as a rehearsal of the information we hope to acquire during the interview process.

8.1. Metadata Workflow -- Key Issues:

- Use of a 'preservation IT system' to control workflow

Do you use or plan to use any dedicated software for preservation work?

- Import / export of metadata from legacy system(s)

Can you move data in and out of your existing catalogue efficiently? If not, why not? What would it take to allow automated movement of metadata?

- Standards of metadata required and supported
  - –Associated metadata - the archive catalogue
  - –Embedded metadata - in media files

Are you aware of existing standards? Do you implement them? Are metadata standards an expense or a savings?

- Updating or completing metadata during preservation

Do you allocate effort to correcting or completing documentation during preservation? Is this a costly area? Should or could you do more?

- Quality control

How do you know your catalogue and embedded metadata are correct? How much time or effort does it take to ensure they are correct? Can this area be improved?

- Adding new metadata
  - –Descriptive: item-level descriptions
  - –Technical: time code
  - –Usage: rights
  - –Additional items: key frames, associated documentation

Do you or will you add documentation during preservation? What sort and why? How important is this area?

8.2. Media Workflow -- Key Issues:

- Handling physical media
  - –Reducing number of times an item is handled
  - –Reducing/consolidating operations performed each time an item is handled
  - –Number of transcription stations, their cost and usage

- Handling electronic media
  - –Savings in physical handling time
  - –Storage savings
  - –Method/quality/speed of delivery to users

- Monitoring
  - –Number of signals or media items handled per operator

- Labelling
  - –Bar-coding of all physical media
  - –Automated printout of barcodes and all other labelling

Please give us your experience in any or all of these areas.

8.3. Key Links:

We have identified three areas where preservation process can be improved:

12.1.1. Missing or Unnecessary Links

- Manual steps that could be automated
- Steps that could be consolidated
- Steps that could be eliminated
- Workflow that could be streamlined if a 'key link' were made possible

In your experience, what steps in preservation can you identify that fall into this category ?

12.1.2. Expensive Links – the main 'cost attractors' of preservation

What is the most expensive single step or 'link' in preservation work, in your experience?.

12.1.3. Opportunity Links – key extra steps that may add cost, but would add value that would securely justify the cost

- Extra steps in preservation work, which allow new services
- Cost of the step during preservation work
- Cost of the step if undertaken as a separate task
- Benefit to the business of the output of the step

What new opportunity (or opportunities) do you hope to achieve through preservation work ?

13. Appendix III – List of Tables and Figures