International Association of Sound and Audiovisual Archives

Internationale Vereinigung der Schall- und audiovisuellen Archive

Association Internationale d'Archives Sonores et Audiovisuelles

Asociación Internacional de Archivos Sonoros y Audiovisuales

iasa

Task Force to establish selection criteria of analogue and digital audio contents for transfer to data formats for preservation purposes

Majella Breen
Gila Flam
Isabelle Giannattasio
Per Holst
Pio Pellizzari
Dietrich Schüller

October 2003

Contents

1. Foreword

The Task Force on Selection for Digital Transfer was commissioned by the IASA Executive Board in February 2000 to examine the issues underlying the process of setting priorities for the digital transfer of analogue and digital audio content, and to deliver a statement of principles for use by sound archives in their planning for digitisation. The members of the Task Force were drawn from IASA's Cataloguing and Documentation, Discography, and Technical Committees, and its National Archives and Radio Sound Archives Sections,.

The Board provided Terms of Reference which included the following:

> Sound archives in all sectors are addressing the challenge of digital collection management using mass storage technology. The capital cost of implementing this technology is significant, but in the medium and longer term the continuing costs of transferring holdings to the new environment are greater. This effect is more pronounced for the larger archives. It is thus necessary to select, to make choices and set priorities which will determine for each institution the parts of its holdings which will be transferred to the new environment at each stage. Success in attracting investment and funding for this activity will depend in part on the ability to demonstrate adherence to internationally agreed guidelines, particularly if they address issues of potential duplication of effort.
>
> This Task Force will examine the issues underlying the process of setting priorities for digital transfer. It will analyse the various criteria which can be applied in the institutional, national, and international context, and will identify strategies for co-operation and co-ordination to avoid duplication of expenditure where institutions have overlapping holdings. It will deliver a statement of principles which can be used by different kinds and sizes of sound archive in planning and setting priorities for digitisation. The issues to be examined will include the following:
>
> - Cultural, scientific, or academic significance of content
> - fragility of existing analogue carriers
> - primary institutional responsibilities
> - technical obsolescence of existing analogue platforms
> - present and future level of demand for use and access
> - restrictions on archival activity arising from intellectual property law
> - the resource required to generate metadata to support the digitised recordings

There is no universally applicable formula to be applied to the complex set of issues addressed by the Task Force, but there are clear principles based around institutional objectives and the intrinsic nature of audiovisual materials. To some extent, this document remains a work-in-progress, and it is presented here as a Working Paper, rather than a formal policy document, but it is hoped that the presentation of these issues here will provide professional colleagues responsible for the care of audiovisual collections with useful guidelines to formulating institutional policy in this increasingly important area of work.

Crispin Jewitt
IASA Executive Board

2. Introduction

Safeguarding of audio recordings in the long term is a task that requires highly complex strategic and technical measures combined with sufficient financial resources. Ethical and technical principles have been laid out by IASA in the document: IASA-TC 03 The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy. This document explains that preservation of audio content in the long term can only be achieved by successive migration of contents in the digital domain, and sets out consistent digital archiving principles. Digital Mass Storage Systems (DMSSs) have proved viable tools to fulfil these aims, providing at the same time new dimension of access, an attractive feature especially for broadcast and national archives. The great challenge of the present time for sound archives is the well organised transfer of their holdings, be they analogue or digital, into DMSSs, or, before these become affordable, into a pre-DMSS environment fulfilling digital preservation principles. It should be recognised that working in a digital preservation environment requires a custodial approach which is significantly different from the established role of the archivist working in the analogue world. The logistical and technical implications of digital preservation are considerable, and archivists must be aware these challenges can only be met by adequate funding.

The document presented here should assist sound archives in establishing a consistent strategy by advising on selection criteria for the prioritisation of transfer projects from different perspectives. A third document under preparation by the IASA Technical Committee will contain operational guidelines for the transfer of analogue and digital holdings into DMMSs.

In comparison to print media, audio recordings, like all audiovisual documents, require a higher degree of physical integrity than text documents. While the latter have a comparatively high degree of redundancy - a speck of mould, generally, does not render a text unreadable - audiovisual documents are analogue representations of physical phenomena (light, sound). Every detail of an audiovisual document is information and, therefore, greater care needs to be taken to maintain the physical integrity of the carrier.

Additionally, due to their chemical composition, which is sometimes highly complex, audio carriers are inherently unstable, comparing unfavourably with the average stability of analogue textual documents. With increasing development and sophistication, the technical platforms developed for audio carriers have ever shorter life expectancies. Additionally, being machine readable documents, the retrievability of the signals embedded in audio carriers is dependant on the availability of dedicated hardware, and sometimes software. Thus, even if carriers remain in playable condition, the retrieval of the recorded contents may be impossible, due to the unavailability of dedicated replay equipment. This obsolescence of formats and systems is, potentially, the great danger inherent in some of the modern digital carriers, while carrier instability is the more general problem with more traditional media.

Consequently, audio archivists are facing considerable problems in attempting to preserve the original carriers placed in their care, because in the mid- to long- term there is a major risk that carrier degradation combined with playback obsolescence will defeat the efforts of archivists to ensure the survival of the content in their care. Around 1990 it became clear that the only viable method of preserving audio contents in the long term is by transfer into the digital domain, and subsequent migration to new formats whenever the need arises. Automation was recognised as a key strategy in optimising the associated workload. This led to the concept of Digital Mass Storage Systems (DMSS), which automatically monitor the data integrity of their contents, copy endangered carriers to new ones before they become unreadable, and migrate the whole archive into a new system once the old system is threatened with becoming obsolete. DMSSs have been introduced over the last ten years and have become state-of-the-art in audio archiving.

Such systems enable the long-term preservation of content, and also provide, as a by-product, a powerful new means of playback access. For the first time in the history of many collections, contents would be easily accessed from the desk of radio journalist or from the visitor's booth of library clients.

Digital Mass Storage Systems are equally important as target repositories for the safeguarding of analogue material, still the greater part of many sound archives' holdings, as well as for digital holdings, generally kept on consumer formats of questionable carrier and/or format stability. Consequently, this paper equally deals with analogue and digital carriers.

Because the transfer of holdings to DMSSs is an elaborate and time consuming process, it is crucial to establish meaningful hierarchies within the sound collections to be transferred. Audio archives can rarely afford "factory style" transfers as envisaged by relatively well-resourced radio archives. This approach may be feasible if the originals are of uniform character, and of high technical quality. Most heritage collections, because of their heterogeneous nature, do not lend themselves to such semi-automated transfers, nor can the respective institution usually afford the large investment needed for such kind of quick and total transfer, because of the considerable costs of equipment and personnel. Because conventional, operator monitored transfer is extremely time consuming (one operator needs at least 3 hours for the transfer of one hour of original material) and because budgets of heritage institution are notoriously low, transfer of holdings into a DMSS will last years, if not decades. A clearly defined hierarchy of priorities for digitising is imperative to avoid, for example, stable materials being transferred first, while in the meantime unstable materials deteriorate to the point where they become irretrievable.

The prioritisation of archive material for digitisation will depend on the statutory obligations and business objectives of the archives. National sound archives will normally have legal deposit responsibilities, and must together with research archives comply with the needs of scientists and researchers. The significant selection criteria for broadcast archives are the possibility of reuse of the material in programme production. Generally, archivists are reminded that according to archival principles neither digitised originals nor non selected original materials should be destroyed. The selection may thus involve a policy decision on keeping the originals for later consultation, or offering them to other archives for further preservation.

3. Technical Selection Criteria

3.1 Technical aspects relevant for prioritisation of transfer projects

Technical aspects play a prominent role in establishing a hierarchy of transfer for carriers that have been selected on the basis of other criteria.

In its current version, IASA-TC 03[1] considers digital technology within the general context of audio archiving philosophy. The document specifies the principles of digital archiving, specifically the stringent quality control of the digital holdings, and the continuous further monitoring of data integrity, before migration becomes imperative. It also deals with the transfer of analogue holdings into the digital domain, the problems related to signal extraction, i.e. the adequate retrieval of the content from their original analogue carriers, the choice of digital resolution and target formats, and the principle of unmodified and linear (i.e. uncompressed) transfer to digital. As transfer technology still improves, and ever higher digital resolutions become available at ever decreasing prices, IASA-TC 03 recommends an unhurried strategy for the digitisation of analogue documents. Priority should be given, however, to those analogue and digital documents which are at immediate risk, and/or which are in regular demand. The document includes a list of inherently unstable carriers.

This chapter builds upon IASA-TC 03, explaining the relevance of technical aspects for prioritisation of transfer projects, and assisting archivists to assess their collections accordingly, specifically their degree of degradation, taking format obsolescence into account.

3.2 Original versus replicated carriers

Originals of all kinds of carriers have higher priority than replicated items. While originals are irreplaceable, there is always a chance that other, even better preserved copies are held by collections elsewhere. The chances may be low for antique replicated cylinders, but very high for recent LPs and CDs. It is recommended that every effort is made to find the best copies available. In the case of the routine transfer of entire collections of originals produced by the phonographic industry, co-operation should be considered between audiovisual archives in order to avoid unnecessary expenses by the multiple, labour intensive transfer of such collections into DMSSs. It may be cheaper to arrange a meaningful division of labour, and to exchange digitised copies, observing of course the related legal aspects including the payment of licence fees, wherever applicable.

3.3 Degradation of Carriers and its Assessment

The most important, and the most difficult assessment is to determine the state of deterioration of a given carrier, or group of carriers in order to arrange for a timely transfer. This section attempts to give a broad overview of the specific situation of various types of carriers, in order to assist archivists in drafting a meaningful transfer strategy. It should be stated, however, that such risk assessment requires a high degree of expertise. Moreover, in many cases it is still impossible (even for experts) to assess the life expectancy of a given carrier, because no valid methods for their evaluation have yet been developed.

3.3.1 Cylinders

Original cylinders consist of wax, replicated cylinders either of wax or of a celluloid (nitrate cellulose) tube covering a plaster core. Both types of materials are considered highly endangered, as wax is extremely fragile and specifically prone to fungus growth under humid storage conditions, while celluloid is endangered by increasing embrittlement with age. Therefore, because of the risk of further deterioration, cylinders should always be of top priority in digitisation programmes.

[1] IASA Technical Committee, The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy (IASA TC-03), Version 2, 2001.

3.3.2 Coarse grooved discs (78rpm records)

There are significant differences in the risk of deterioration between replicated coarse grooved discs, so called shellac records, and so called "instantaneous", or “direct-cut” discs, original recordings, which may consist of manifold materials.

Replicated shellac records consist mainly of mineral substances bound together by an organic binder. Unless stored under irregular conditions, they have proved to be fairly stable over many decades. Seen from that perspective they are on the low end of the priority scale.

Pre WW I replicated discs, however, may consist of more critical materials, which suggest a higher ranking for prioritisation.

Instantaneous discs are much more unstable. The most widely used variety is the so called "lacquer" disc. It consist of a glass or metal core, covered by a layer of lacquer, carrying the information, mainly consisting of nitrate cellulose, sometimes acetate cellulose, which gave this group of records the generic name "acetate discs". Mainly due to high levels of humidity and temperature, these lacquer coatings shrink, and become brittle with age until they crackle, without significant previous warning. Once the surface is crackled, generally the signal cannot be retrieved. Therefore, all lacquer discs should be transferred without delay, even if they seem in excellent condition - because of this susceptibility to sudden deterioration.

All other kinds of instantaneous disc should be accorded a high priority for transfer because of uncertainty about their stability (and because of their unique content).

3.3.3 Microgroove discs (“vinyls”, LPs)

Replicated microgroove discs, as produced since around 1948, consist of a PVC/PVA (polyvinyl chloride - polyvinyl acetate) co-polymer. Generally, they have proved stable so far. According to present knowledge, no systemic degradation of these materials is expected in the near future. It is necessary however, to inspect collections for the possible plasticizer exudation of earlier records, and of plasticizer migration which sometimes happens when such discs have been kept in inadequate plastic materials. Such items should be transferred immediately, but the great majority of vinyl discs can be ranked at the lower end of the priority list.

3.3.4 Magnetic tape

Magnetic tape mainly consists of two layers, the base film and the magnetic layer which carries the information. The base film consists, in historical sequence, of acetate cellulose (produced until the mid-1960s), PVC (produced mainly in Germany between 1943 and 1972) and polyester terephtalate, generally called polyester, which has been in use since the late 1950s. While PVC and polyester have proved stable materials and deterioration is not anticipated in the near future, acetate cellulose is severely at risk. As already explained for the lacquer coatings of instantaneous discs, this material shrinks and becomes brittle with age. The degradation is caused by hydrolysis, a chemical decomposition process under the influence of water, omnipresent in the form of humidity in the air. The higher the levels of humidity and temperature have been, the higher the deterioration. Additionally, acetic acid is produced in the course of hydrolysis, which reacts as a catalyst that accelerates further deterioration. This process mainly affects stocks of films made from acetate cellulose, and is known as the "vinegar syndrome" which sometimes render films irretrievable. With audio tape such disastrous consequences have not yet been observed. However, acetate tape stocks should be considered at risk and should be ranked at the upper end of priority.

Acetate tapes can be identified by their brand and type (list forthcoming from the IASA TC), but also by relative simple investigation. All standard play, most long play, and some double play acetate tapes have a translucent wind when held against a light source. This is not the case with PVC or polyester tapes. Acetate tapes are also fairly stiff, sometimes brittle, and break without previously stretching when mechanically stressed. Aged acetate tapes can also be identified by their warped edges or polygonal appearance when hanging down from a reel.

Generally, only standard play tapes open reel tapes (SP, 52 μm total thickness) should be trusted to be mechanically stable. Long play tapes (LP, 35 μm thickness), double play tapes (DP, 26 μm thickness) and the rarely used triple play tapes (TP, 18 μm thickness) are vulnerable at increasing degree. Audio cassette tapes are mechanically even more vulnerable: C 60, C 90 and C 120 cassettes have a thickness of 18, 12 and 9 μm respectively. R-Dat tapes have a thickness of 12 μm. The lesser the mechanical stability, the greater the chance that the tape suffers from inadequate winding, which is one of the most underrated risks for magnetic tapes. Open reels and cassette tape are equally affected, and prolonged storage of badly wound tapes causes irreversible deformations, which may lead to severe replay problems, specifically with thin tapes and high density recordings, e.g. R-Dat.

The magnetic layer usually consists of two parts [or components]: the magnetic pigment itself, that takes up and holds the information, and the binder that glues the magnetic particles together and onto the substrate. The following magnetic particles are in use: γ-Fe_2 0_3, used for all audio open reel tapes, and for IEC type I audio cassettes; CrO_2, and chromium doped particles, as used for IEC type II audio cassettes and the earlier video cassettes, some of which are also used for digital audio recording, e.g. U-matic (PCM 1600/10/30), Betamax (PCM F1), VHS (ADAT). More recent magnetic pigments known as metal particle (MP) have a core of pure, metallic (non oxidised) iron with an inert ceramic or mineral layer to protect them from oxidisation. These are used in IEC type IV audio cassettes, R-DAT cassettes, and video cassettes used for digital audio formats, e.g. DTRS. The stability of γ-Fe_2 0_3 magnetic particles is not questioned, although CrO_2 and chromium doped particles are less stable magnetically. Some MP tapes are threatened by oxidation, or corrosion of the particles. The MP coatings at greatest risk are those manufactured in the late 1980s and early 1990s prior to improvements in passivation techniques. A more recent development is the use of thin, vapour-deposited metal coatings known as metal-evaporative (ME). Some ME tapes suffered catastrophic failure as a result of the metal layer flaking away from the polymer base-film.

The greatest problem with magnetic tapes is the material which binds the magnetic pigments to the substrate. Generally, traditional binding materials have a good to fair reputation of stability. From the mid-1970s onward, however, new polyester polyurethane binders (PEU) have been used, which, to various degree, are prone to hydrolysis. Water present in humidity of the air reacts with the binder, which leads to its chemical transformation, accompanied by a different physical performance. Binders loose their binding properties, which lead to a loss of pigments. In the course of the replay process, these pigment particles are deposited on tapes guides and replay heads swiftly impairing the quality of the replayed signal. This phenomenon is called "Sticky tape/sticky shed syndrome" and is often accompanied by a squeal in the replay process, caused by undue friction of affected tapes in the tape guides. In severe cases this friction may even lead to the break down of the tape travel. Sometimes, massive oxide shedding and even a total peeling-off of the magnetic layers can be observed .

There are only a few types of tape with conventional chemical binder technology which have not been affected to date by this unfavourable process, mainly broadcast studio tapes of German origin. A list of these tapes is under preparation by the IASA Technical Committee.

It is yet unclear whether binder degradation is the problem of a limited number of ill-designed or ill-produced tapes, or whether sooner or later all magnetic tapes will affected by this phenomenon. The development of methods to predict life expectancy (LE) of magnetic particle binder is in its infancy, and considerable research is needed before a valid methodology will be available.

Consequently, most of the tapes produced after the mid-1970s should be suspected of being inherently unstable. Before efficient and easily applicable LE tests become available, utmost vigilance is necessary to find potentially affected stocks by labour intensive individual tape inspections. The IASA Technical Committee hopes that the problem will be solved by a continuation of the systematic co-operation between sound archives and manufacturers that has recently been established under the auspicies of UNESCO. The disclosure of potentially risky types of tapes by the manufacturers would be a substantial help.

Apart from the potential effects on the LE of magnetic tapes arising from the specific parameters described above, the storage history of archival holdings has also to be taken into account. Heavily used materials, as well as irregular storage condition over the lifetime of specific holdings, both call for a higher ranking in transfer projects.

In summary, only general and indirect conclusions can be drawn from parameters related to the physical and chemical conditions of magnetic tape carriers as to the inherent risk of their becoming irretrievable.

Exceptionally, the digital R-Dat format does allow for an objective assessment of its state of preservation. Based on the availability of specific replay equipment in association with dedicated software, the monitoring of data integrity of a given tape can be assessed objectively. Monitoring in regular intervals allows assessment of the slope of degradation of a given carrier, this in turn allows for a timely transfer onto a new digital carrier before uncorrectable errors occur as a consequence of further deterioration (cf. also the respective notes on CD integrity).

3.3.5 Optical Carriers

Compact Discs are the most widely used optical carriers. Replicated CDs were first introduced in 1982. These consist of a polycarbonate body which carries the information on its upper surface in form of a helical track of "pits" and "lands" embossed in the process of production by injection moulding. This upper surface is covered with a reflective layer, generally of aluminium. The reflective layer is protected by a layer of varnish, also carrying the label information.

Before the stability of the various components is discussed, it is important to remember that data integrity of CDs, like all other digital media, is objectively measurable by special CD players and suitable software, which should be standard equipment of all sound archives holding and generating CDs. According to digital archival principles (cf. IASA-TC 03, § 11), every CD must be free of uncorrectable errors. It is important to understand, that the CD-Audio format (CD-A) provides for error concealment (interpolation), if the level of (true) error correction is surpassed. The CD still produces an audible signal in such cases, which, however, contains interpolations and is thus not exactly the original signal. Consequently, it is advisable that newly acquired replicated CDs as well as internally generated CD-Rs are tested for full error correction, and such tests be repeated in regular intervals. Transfers onto a new medium becomes imperative well before the threshold for full error correction is surpassed. It must be noted that under likewise identical conditions CD-Rs which start with a low level of correctable errors will have a longer lifespan before they fail.

Regarding stability, the polycarbonate substrate of the CD has proved to be fairly stable. Considerable problems have been observed, however, with its predecessors, the analogue video discs that are composed of the same materials. Instances of crazing have occurred, rendering the polycarbonate opaque to the reading laser beam. Whether this problem has been totally overcome with current CDs, or only retarded, is unknown. The reflective layer, generally of aluminium, is prone to oxidation. Therefore, this layer is covered by a protective layer of varnish. This varnish is the most delicate part of a CD. Instability in this layer, which was frequently observed in the early years of the CD, can cause the reflective layer to deteriorate, rendering discs unreadable. Deterioration in the protective layercan result from mechanical scratches, the use of inappropriate, "bleeding" dyes for the label information, and by chemical degradation due to ageing or improper storage. The migration of adhesives from labels traditionally used in libraries to identify the object and its owner, can have destructive effects. For all these reasons, specifically older parts of CD stocks, and all labelled CDs must be considered at risk.

Recordable CDs also consist of a polycarbonate substrate, carrying on its upper surface a helical groove filled with an organic dye that carries the information content. Above this is the reflective layer, originally of gold, more recently of silver. This is in turn covered by a protective layer of varnish which is generally of more sturdy character than that of replicated discs, allowing the use

of felt pens and also the application of printed labels with specials adhesives for identification. Various dyes have been used so far, to which various degrees of stability have been ascribed. While some manufacturers claim a life time of their products for 100 years and more, experience tells that many CD-Rs have failed only after few years.. Beyond the chemical and physical decay processes, an important element of LE is the rate of correctable errors produced during recording. As explained above, under otherwise equal conditions, the lower the error rate of CD-Rs, the longer the LE; the higher the error rate, the shorter the LE. Currently, however, CD-Rs are undergoing further development in order to comply with endeavours for ever increased recording speeds, leading to problems in the compatibility between disc writers and blank discs. Such incompatibility inevitably leads to increased error rates, negatively influencing the LE of these CDs from the outset. Whenever CD-Rs are used, compatibility between writers and blanks has to be carefully explored, the error status of each CD-R has to assessed and recorded for further monitoring.

Consequently, in assessing the state of preservation especially of CD-R stocks, the measurement of data integrity is imperative. Transfer must be organised well before uncorrectable errors occur. It is difficult to recommend a certain level which must not be surpassed. A flat gradient of deterioration may allow for a higher level of correctable errors before taking action, while a steeper gradient may call for earlier or even immediate action.

DVDs, replicated as well as recordable, are of the same construction as CDs. Because of their higher data density (by factor of 7 as compared to CDs), the factors determining the stability of CDs will have a relatively more aggravating influence on DVD stability. As with CDs, a serious risk assessment can only be based on frequent monitoring of the data integrity of these media.

Replicated MiniDiscs (MDs) also function like CDs. Recordable MiniDics function on the basis of magneto-optical recording, which has been successfully used for data recording in various formats, and has proved fairly stable, at least in the medium term.

In concluding the section on carrier degradation it should be emphasised that substantial enhancement of preservation activity will be a key element in the establishment of successful transfer strategies. Prioritisation of transfer projects should be complemented by strategies to prolong the life time of lower ranked parts of the collection by optimising storage conditions.

3.4. Obsolescence of Replay Equipment and Associated Software

As mentioned earlier, all audio carriers are machine readable documents. A meaningful strategy in the ranking of transfer projects has also to take the availability of replay equipment into account, specifically the threat of sophisticated equipment becoming obsolete resulting in major problems with the future retrieval of content, even where carriers are still in good condition.

3.4.1 Formats already obsolete

All mechanical carriers formats are already obsolete. However, this does not constitute a major threat to their continuing playback. Replay equipment designed by experts is available for the playback of cylinders and coarse grooved discs.

Another group of obsolete audio formats are early digital audio formats employing analogue video formats as target carriers. The most prominent of these are: Sony PCM F1 (701, 601 and 501) using Betamax recorders, and Sony PCM 1600/10/30, using U-matic recorders. Betamax was widely used by smaller recording studios as well as research and heritage institutions, and U-matic was the standard for CD-mastering. Betamax players can only be found on the second hand market. U-matic machines are out of production by now, but spare parts continue to be available for the time being.

It may also be noted that most early digital formats of the DASH and PD families are practically dead. These have been employed by the recording industry and it may be assumed that sufficient numbers of replay machines, including spare parts, are available in these institutions.

3.4.2 Formats about to become obsolete

The quarter inch analogue magnetic tape format is currently about to become obsolete through the progressive withdrawal of manufacturers from the production of new equipment. It is of utmost importance that sound archives immediately assess their need of new machines, and spare parts for their existing equipment, in order to assure the future orderly and complete transfer of available stocks. It should also be noted that replay equipment must match the recordings in terms of speed, track format, and equalisation. This poses considerable problems for many heritage collections, as no modern replay machines are available for low speeds such as 2.38 and 4.76 cm/s.

3.4.3 Other formats

R-Dat has been widely utilised in smaller studios and radio stations. It has also been widely used as a digital target format in digitisation projects of endangered analogue carriers. It remains the format of choice for location recording and field work across all scholarly disciplines for the time being. However, because of the spread of digital audio workstations, and the increased use of CD-Rs or computer tape like DLT as target formats, R-Dat has recently lost its predominant position. Consequently, it is a wise precaution to consider the forthcoming unavailability of new equipment and spare parts and to act accordingly.

With the spread of multi-channel digital workstations recording formats such as ADAT or DTRS may become obsolete very soon.

With format or platform obsolescence, service engineers may also become obsolete; another important factor to be taken into account when establishing a hierarchy of transfer projects. Unlike obsolete mechanical formats, obsolescence of all other formats is associated with a severe problem regarding service and repair of modern equipment, specifically digital cassette formats. Expert knowledge and skill is required, often associated with specific reference tools and dedicated equipment.

Presently, the MiniDisc as such seems not to be threatened by obsolescence. It is a data reduced format, however, and its data reduction process (ATRAC) is proprietary and unpublished. Since the introduction of MDs, ATRAC has been further developed and is presently in its fourth version. It is unknown, whether it will be further developed and whether future versions will be fully backwards compatible, capable of replaying contents encoded with earlier versions with full original quality, or of replaying at all.

Given this situation it is recommended that unique contents from recordable Minidiscs should be transferred without delay in a pseudo-linearised version as recommended by IASA-TC 03, § 10. It should be noted, however, that the many replicated MiniDiscs are parallel editions to CDs, and therefore are not necessarily items for which early preservation transfer is a high priority.

Obsolescence is not foreseen in the nearer future for the (analogue) compact cassette, the CD and its subformats, and, of course, the DVD.

3.5 Summary

In establishing a meaningful hierarchy of transfer projects it is necessary to take into account the nature of the document (unique or replicated), the threat of carrier degradation, and the threats of obsolescence. It will often be very difficult to make a proper quantitative assessment while properly balancing degradation against obsolescence. Because of new experiences and results of research related to carrier degradation on the one hand, and the development of the market which determines obsolescence on the other, sound archives need to be on constant alert and prepared to adapt their strategy to new situations as they arise.

4. Selection for Digitisation in Broadcasting Sound Archives

4.1 Considerations and constraints

Many broadcasting organisations are embarking on projects to digitise parts or all of their audio collections. The extent and scale of projects varies a great deal depending on available resources. However, even in cases whereby there are plans to digitise the whole collection, priority to digitise is given to selected parts of the collection. The following seeks to identify considerations and constraints and to propose guidelines for the setting of priorities.

Cost is the most obvious constraint. Funding for digitisation may be limited to a project comprising a small percentage of the whole collection. Public funding and co-operation with national or cultural institutions should be considered. Some broadcast archives are part of a national archive and some are not, and the role of the archive will influence selection criteria .

Broadcasting organisations need access to archive material for re-use in transmission and increasingly for re-use and re-purposing in other digital projects (DAB, Internet, etc.). An example of re-purposing would be using audio underneath a still image on a web page. Unique material and more broadly material which is of value from a heritage perspective needs consideration. Other considerations are the management of Metadata and copyright restrictions. There will, of course, always be a case for making exceptions when selecting.

4.2 Cultural, scientific, or academic significance of content

Most radio archives hold material that represents historical, social, political and cultural aspects of life in this century and the last. Material unique to an organisation is of particular interest. Other recordings may be available elsewhere, commercially or in international collections. Co-operation on the digitising of overlapping holdings is recommended. Priority should be given to digitising holdings unavailable anywhere else.

4.3 Primary Institutional Responsibilities

Responsibilities and roles depend on whether the broadcast archive is part of a national archive or not. Apart from a broader national role most broadcast archives serve to provide material for re-use in transmission and other means of output. It is recommended that selections from current output are as inclusive as possible. Some material receives more frequent use while other parts of a collection remain dormant. It is recommended that an analysis of current requests be undertaken to determine what material is in high demand. Once identified, frequently used material can be given priority for digitising. While it makes sense to identify a body of material and to then systematically digitise, consideration should also be given to the possibility of digitising on demand.

Output to other media such as DAB and Internet may now be as important as conventional re-transmission. Re-purposing increases the opportunities for re-use of material. Whereas in the past a radio play may have had to wait for a re-schedule, an audio clip may now be associated with another digital object (an image, for example) and used to illustrate the voice of a well-known actor on the Internet. New Media may be more or less of a priority depending on the organisation.

4.4 Technical Considerations

Many broadcast archives have stock on carriers now degrading and unstable. IASA document ref. IASA –TC 03 provides a list of unstable analogue and digital carriers and this issue is dealt with in some detail in chapter 3 of this paper.

4.5 Metadata

In the digital domain there is a heavy dependence on correct labelling for the purpose of future retrieval. Serious consideration must be given to descriptive metadata. The textual description

(if it exists in the first place), may need to be enhanced and must be associated with the digitised audio. Records Management, while an essential part of any digitising project, is a labour intensive and therefore costly process. Even the most basic description is better that nothing. However, in exceptional cases carriers in danger may have to be transferred without Metadata.

4.6 Copyright

Digitising facilitates easier and quicker access and the exchange of material. Many broadcasting organisations have legacy contracts drawn up for the purpose of transmission only. These restrictions may severely limit the potential to exploit material via new media. Organisations may wish to digitise material for which they have rights, sooner than material with copyright restrictions attached. Consideration needs to be given to devising new and more inclusive contracts.

Differences in national copyright laws are likely to cause hindrance to the exploitation of exchange material. Co-operation in this field and the introduction of international legislation aimed at facilitating access should be encouraged.

4.7 Representative Selection

Harald v. Hielmcrone[2] outlined the use of Formal, Structural and Content based criteria in order to achieve representative selection. Periods of time (e.g. by decade) are an example of a formal criterion. The programme unit is an example of a structural criterion. Genre is an important content-based criterion that needs consideration in representative selection. Von Hielmcrone suggests eight main genres: Fact; Fiction; Entertainment; Public Access; Juvenile programmes; Local and regional programmes; Presentation; Advertisements.

2 . Harald v. Hielmcrone, Selection criteria for archiving radio and television programmes: The Danish Experience, State and University Library, Aarhus, Denmark, 2002.

5. Selection for Digitisation in National Sound Archives

5.1 National Archives

A sound archive is defined as a National Archive when, by law or in fact, the institution has the responsibility of collecting and preserving the national production.

National archives should consider the selection criteria for the digitisation of sound carriers in three steps.
The first and urgent step is based on two criteria: 1) institutional assignment and responsibilities; 2) technical emergency.
A second stage should determine the intrinsic cultural value of the contents, according to : 1) their cultural, scientific or academic significance for the institution ; 2) their rarity.
A third stage would evaluate the benefits and costs of making the digitised content accessible.

5.2 Basic criteria

5.2.1 Institutional assignments and responsibilities

The institution may be tasked with the preservation of national, regional or special heritage collections.

Such an institution has three ways of increasing its collections
- legal deposit: the types of documents to be collected are defined by the Law
- acquisitions: defined, or not yet defined, by a Documentary Policy
- acquiring the production of research activity, as a moral obligation if there is not a Research Archive, with a preservation policy.

The legal or statutory task of an archive is the main criterion of prioritisation. Legal deposit usually involves a commitment to the conservation of all documents received in this way.

NB: Some Legal deposit Law definitions can be more comprehensive than others: e.g. documents produced in the country only, or documents produced or sold in the country.

NB: Preservation of sound documents implies the production of copies (copies for access, copies for preservation). Ultimately, long term preservation can only be achieved by subsequent copying of contents from one generation of system to the next. Although preservation is generally legally prescribed, copying for the purpose of preservation is not always exempted from the provisions of copyright law.

5.2.2 Technical emergency

National Archives consider that technical reasons should be predominant when determining digitisation priorities. These issues are dealt with in some detail in chapter 3 of this paper.

5.3 Secondary criteria

5.3.1 Cultural, scientific or academic significance of content

In their everyday work, sound and audiovisual archivists usually acquire documents according to their own evaluation of the contents. A Documentary Acquisitions Policy is defined for this activity. Choices for digitisation follow similar judgements and will be formalized and justified in the same way. The Preservation Policy should accord with the Documentary Acquisitions Policy. Archival research collections have to be evaluated as a whole, according to the basis of archival management. The various contents of a collection that reflects a researcher's or an institution's activity (e.g. spoken and musical recordings, or different subjects) should not be divided in high and low priorities.

5.3.2 Rarity of the item

Equivalent recording in one's own collections

A number of questions should be asked. Has the recording been reissued and if so, in an analogue or a digital version? Who owns it? Should the re-issue be acquired and regarded as the preservation copy on the grounds that this would be cheaper than copying the original? Or should the original recording be regarded as irreplaceable by the reissue (e.g. because of the way it has been transferred?)

Same recording in other collections

Again, questions arise. Do other archives possess the same recording? Can users easily have access to it? There is a need for published or on-line catalogues of potential duplicate material, which clearly indicate the digitisation status of documents. There is also a need for a real culture of cooperation between archives[3].

5.4 Benefits and costs linked with access

Among all issues examined here, cost is the factor which illustrates the different scale of possibilities in the digital as opposed to the analogue domain. Digitisation and Internet facilities allow increased access from outside the archival institution, and has the potential to generate new services. Yet copyright severely restricts the spreading of digital files.

5.4.1 Benefits: current and potential users

As archives have the mission to work for an audience, current use has to be taken in consideration, as well as the potential demand enabled by digitisation. Current use may be low because of restrictive access policies, whereas digitisation allows much broader access off- or on-line. Accessibility is a factor which can generate new fields of research.

All of the following services can be off line, by delivering or publishing a CD-ROM, and they can now be on line, via the Internet:

- exchange of documents between national archives
- delivery of documents for research and education purpose
- delivery of documents to final users, for private use, on demand
- publishing from collections

All these services require regulation, by Law or by special agreements for copyright. This has to be cleared before considering criteria for digitization.

5.4.2 Costs

Documentation

Un-digitised, but already catalogued holdings can be rendered easily accessible to the user. Cataloguing is an important cost and extension of access to a wider public needs appropriate editorial work.

Restoration

Restoration can be necessary to make audible (thus accessible to a large audience) the contents once digitised. This operation can multiply the costs of digitisation.

---

[2] An example for a policy of collaboration between various sound archives:
The German/Swiss-German branch (D-CH) of IASA is working on a project called "Historical label discographies", a discography of labels of historical sound carriers [Firmendiskographie historischer Tonträger (FDHT)]. This interesting work requires the collaboration of international institutional members and includes, apart from various archives of institutes, archives of broadcasting companies, national archives, and also important private archives.
If such discography is completed with identifications of exemplars kept by members, it can be used to evaluate rarity of items. If technical description is provided, it can be used to select the best one.

5.5 Some concluding considerations

Digitisation planning should be liable to revisions, in order to take into account enrichments of the institution's collections. This dimension is important, for wider access made possible by digitisation can be an important factor of donations and deposits.

The Preservation Plan is not completed by digitisation.: It is necessary to continue to observe the development of digital formats, ageing process etc.

6. Selection for Digitisation in Research Archives

Research Archives are repositories of recorded sound materials that have been produced, collected and preserved as acoustical sources for a number of scholarly disciplines, e.g. anthropology and folklore; (ethno-)musicology, linguistics; social sciences, zoology, etc. At the core of such collections are original, unique recordings, made by scholars to serve the specific interests of their disciplines. Generally, such recordings are accompanied by documentation about their origin and their contents. In addition to these unique materials, research archives may also hold published recordings or original recordings of sister institutions for reference purposes. Special interest collections, e.g. Jazz or Popular Music Collections, also are part of the Research Archives group. They may consist of a mix of original recordings made by the respective collection, of original recordings of sister institutions, and of commercial recordings.

Research Archives constitute the oldest and most widespread type of sound archives. Having started in 1899, their numbers grew in the 1950s, when magnetic recording became affordable and battery-powered tape recorders made it possible to easily produce field recordings everywhere in the world. The interest of the researchers generating such collections went well beyond the limits of their respective nations. In the course of the last fifty years research archives, specifically of wealthy countries, accumulated considerable stocks of sound recordings from manifold cultures from all over the world. Due to the global progress of western civilisation, they cover a considerable time span over a period of significant cultural change. Research Collections meanwhile constitute an important part of the audiovisually recorded cultural heritage, adding an important dimension of general cultural interest to the scholarly objective they originally produced and collected for. Thus, research archives have become by far the most significant repositories of acoustical sources related to the cultural and linguistic diversity of mankind.

In setting priorities for the digitisation of such collections, unique source materials should be in the first rank. Regarding criteria inherent to the contents of the collections, the principles as discussed under National Archives *(Ref. Chapter 5.3.1)* apply. It must be noted that the immediate research interest of the individual archive may be in conflict with the general cultural importance of a given collection. Research Archives must understand, however, that they have a moral obligation regarding the safeguarding of their holdings representing early and important historical testimonies of societies from all around the world; even if their immediate scholarly value may be less significant

Not all of these collections, however, are archives in the narrower sense, with an established specific preservation policy. Collections have have often simply been regarded as tools to advance the respective academic disciplines, with preservation issues accorded no priority in the policy of their parent institutions, and no specific budgets allocated to safeguard these holdings. Such policy worked more or less successfully until recently, because in general analogue tape stocks have been fairly stable, and tape players have continued to be available. This will change in the near future: Carriers will increasingly become unstable and irretrievable, and traditional tape replay equipment is about to become obsolete. Unless research archives, specifically the small and specialised collections attached to research institutions, establish a consistent preservation i.e. digitisation policy, there will be a significant loss of invaluable heritage materials. A large corpus of documents of cultural manifestations, mainly of orally transmitted cultures of all parts of the world will be lost in the medium term.

In view of this threat it is important that responsibility is taken to systematically ensure the further preservation and availability of these important materials. In the first instance this challenge should be adequately met by the respective institutions that have produced and kept these materials so far. This can either be done autonomously or by appropriate co-operation between several of such collections.

Where the research collection is for one reason or another unable to meet this preservation responsibility the respective National Archives should take care of these materials, as being part of the respective national production, irrespective of the contents and the cultural and geographical origin of such recordings.

7. Metadata

7.1 Metadata and digitisation

Metadata literally means "data about data." Any catalogue, card or online, contains metadata. But, today, the term is applied to the value-added information that information- specialists create to arrange, describe, track and otherwise enhance access to information objects.

Metadata is used to describe, in a standardized way, the minimum set of information that is necessary to locate a document. In addition, metadata provides a standard way to describe network-accessible material; metadata enable the user to make more precise queries, metadata help the search engines to present hits that are grouped by subject rather than a random mix.

Or in other words, metadata are data that describe the attributes of a resource; characterise its relationships; support its discovery, management, and effective use; and exist in an electronic environment. While this definition reflects the metadata ideal, in reality, most metadata schemes rarely accomplish all of these functions equally well. Therefore an authority control system must be understood and placed to each metadata scheme chosen by the institution.

An authority control system is based on four factors: controlled operating environment, trained provider, application of standards, and reference to authoritative lists. Therefore, a trained cataloguer of audiovisual material should be included in the selection and digitisation workflow.

The most common metadata systems today are the MARC (Machine Readable Cataloguing) bibliographic record and the DC (Dublin Core) Metadata Element Set. Both require adjustments for audiovisual collections. In the EBU Technical Review, Dublin Core Metadata Element Set is recommended. Depending on its concerns an institution may choose one scheme over another.

Metadata can be embedded within the document itself. Standard Generalized Markup Language (SGML) is one of the most common markup languages used for metadata, along with its related codes Hyper Text Markup Languages (HTML), and the newer Extensible Markup Language (XML).
Metadata can also be stored separately from the source (e.g. external catalogue) or separate but linked to the resource (e.g. linked with the digital object in a repository structure). Each strategy has particular benefits and disadvantages.

Information specialists along with archivists and cataloguers should create the metadata for audiovisual archives. Reference should be made to authoritative lists. Audiovisual archives should be part of the scholarly information culture and not the popular information culture.

In the metadata creation process, preservation metadata should be a key component in the preservation and management of the digital collection and must be designed to support future preservation strategies.

8. Conclusion

Due to the costs involved in digitisation it is imperative to find ways to reduce the economic burden on audiovisual archives. The avoidance of digitising duplicate collections will be an important strategy. The IASA website should be used to present information about duplicate collections in order to avoid unnecessary digitisation.

In the coming years guidelines for selection for digitisation will undoubtedly be further developed and expanded as more experience is gained from ongoing digitisation projects. Archives involved in such projects will have an important role to play sharing experiences concerning the selection process with archives facing the challenges of digitisation.

8. Conclusion

Due to the costs involved in digitisation it is imperative to find ways to reduce the economic burden on audiovisual archives. The avoidance of digitising duplicate collections will be an important strategy. The IASA website should be used to present information about duplicate collections in order to avoid unnecessary digitisation.

In the coming years guidelines for selection for digitisation will undoubtedly be further developed and expanded as more experience is gained from ongoing digitisation projects. Archives involved in such projects will have an important role to play sharing experiences concerning the selection process with archives facing the challenges of digitisation.