



## Delivery Notification for: Small Archive Technical Mapping Support Tool

### Deliverable D13.3

#### “Preservation needs of small archives”

DOCUMENT IDENTIFIER	PS_WP13_ITI_D13.3_tech_mapping_v3.01
DATE	09 January 2008
ABSTRACT	This report describes a Web-based information resource developed by IT Innovation to specifically address the needs of small archives when they plan and cost their digitisation and storage projects.
KEYWORDS	Small archives, technical media mapping
WORKPACKAGE / TASK	WP13
AUTHOR, COMPANY	IT Innovation Centre
NATURE	P : Prototype
DISSEMINATION	PU : Public
INTERNAL REVIEWERS	Jean-Hugues Chenot, INA

#### DOCUMENT HISTORY

Release	Date	Reason of change	Status	Distribution
1.0	02/03/2007	First Draft	Living	Confidential
2.0	12/03/2007	Second Draft	Living	Confidential
3.0	07/01/2008	Third Draft	Living	Confidential
3.01	9/01/2008	Approved	Complete	Public

## Table of Contents

<a href="#">1. Document Scope.....</a>	<a href="#">2</a>
<a href="#">2. Summary of Deliverable.....</a>	<a href="#">2</a>
<a href="#">3. Notification of Delivery.....</a>	<a href="#">3</a>
<a href="#">4. Screenshots and Examples.....</a>	<a href="#">5</a>
<a href="#">5. Original Motivation.....</a>	<a href="#">12</a>
<a href="#">6. Approach.....</a>	<a href="#">16</a>

# 1.Document Scope

---

This report describes a Web-based information resource developed by IT Innovation to specifically address the needs of small archives when they plan and cost their digitisation and storage projects.

The work follows-on from D13.1, which contained cost models and processes for large archives who want to plan large-scale long-term digitisation and storage projects, and D13.2, which made these models available as simple to use on-line calculators.

This report describes a tool that helps small archives find the information they need so that they too can use and benefit from these calculators and cost models.

# 2.Summary of Deliverable

---

Small archives often have many problems when considering the use of digital technology for the long-term preservation of and access to their audiovisual assets.

Small archives would like to use the various PrestoSpace costing and planning tools when considering what to do with their holdings. Unfortunately, small archives (and some big ones too) often face some basic problems of knowing what they have and what they should do with it, in particular how to move their content into digital form so they can continue to use it. Addressing these problems is the first step on the road to using the rest of the tools, processes, and best practice that PrestoSpace has to offer.

This report describes a structured online database of information about audio and video carriers; how this resource is accessible through a simple web-based search and navigation interface; and how it can be used to find information on carriers, players, problems, treatments, risks, digitisation, digital formats, costs and other relevant information needed to answer the question 'what have I got and what should I do with it'.

## 3. Notification of Delivery

This report is an update to the previous version, which was released for the 3<sup>rd</sup> PrestoSpace Annual Review.

The Web-based resource that this report describes is publicly accessible over the Internet and is open to anyone in the AV archiving community to use free of charge and for whatever purpose they wish.

The resource is available at: <http://prestospace.it-innovation.soton.ac.uk>

The resource will be publicised at the final user meeting in Rome and on various emailing lists as appropriate, e.g. AMIA.

This report is marked confidential rather than public because it isn't intended for use outside the PrestoSpace consortium. The report describes the rationale and process used for developing D13.3, the inputs provided by the consortium partners, and the versions that have been issued to the consortium and the public. Therefore it provides important information for the consortium and project Reviewers, but isn't appropriate for public consumption.

Following the 3<sup>rd</sup> Annual Review, the following changes have been made to the Web-based resource. These address the concerns that the Reviewers rightly had at that time about the previous version being incomplete and not publicly accessible.

- The database has been enhanced with a wider range of carriers, in particular information on a range of audio formats. There are now almost 60 different carriers in the database. Both ORF and the BBC have contributed information on the audio formats now in the database.
- The user interface has been improved to allow easier searching and navigation of the carriers in the database. Additions include a new 'front page' that provides ways to enter the database when looking for formats, e.g. by allowing the user to see the entries that are Analogue/Digital, Audio/Video, Cassette/Reel/Disc/Other, or in use for different decades, e.g. 1970-1980.
- Entries in the database link to pages on the PrestoSpace wiki to allow users of the resource to make their own comments/corrections/additions, for example to provide links to other web-based resources. The wiki pages link back to the database to allow wiki users to go straight to the database.
- The look and feel of the user interface is now consistent with other PrestoSpace websites, e.g. by use of the same colour scheme and logo.
- An extended list of synonyms allow users to find items using a wider range of terms, e.g. '1/2 inch', '0.5 inch', and 'half inch' will all match the same database entries for tapes of this width. 'DCC' will match 'Digital Compact Cassette' and 'M2' will match 'MII'. Synonyms are also included for the audio formats now in the database. For example, 'Vertical cut' will match 'Edison'

and 'Pathe' discs, 'Grammophone' and 'phonograph' will both match 'record', and 'LP' will match '33rpm' and 'single' will match '45rpm'.

- Both imperial and metric units are supported for format sizes, e.g. 30cm/12inch discs.
- More navigation links have been added to the details page for each media format to allow easy navigation to other formats of the same size or type etc.
- The advanced searching is easier to use and has an extended set of search fields.
- The detailed information on each format has been organised into a set of 'tabs' that are easy to switch between (e.g. further images, surveys, problems, players). This contrasts with previous versions where the user was expected the user to scroll through a long list of information.

After the first EC review, a new version of the tool was reviewed by the BBC in October which prompted a further revision by IT Innovation in November to create a candidate release for the consortium. The consortium was invited to review the tool at the start of December 2007 (hosted as a web-accessible service at IT Innovation). Further iterations of the user interface and additions to the content were done in December 2007 and January 2008, in particular by incorporating content from ORF. A final version of the tool is now publicly accessible.

IT Innovation will continue to host and support the online database for at least the next 18 months. The database of content will be made available to community via the PrestoSpace website. The software used to run the online tool is available as open source.

## 4.Screenshots and Examples

A set of screen shots are provided below that show the information available through the tool and how this information can be accessed.

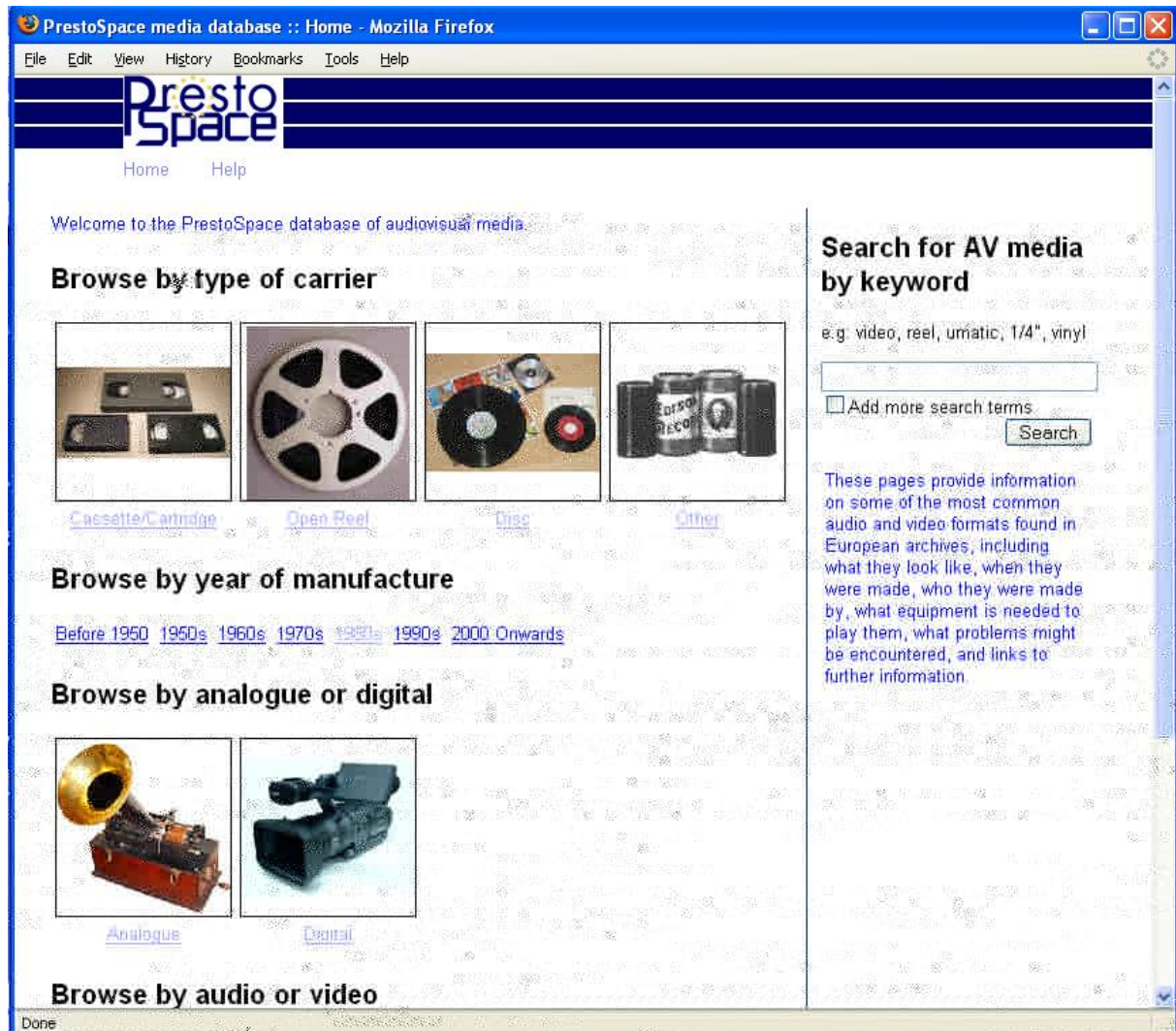
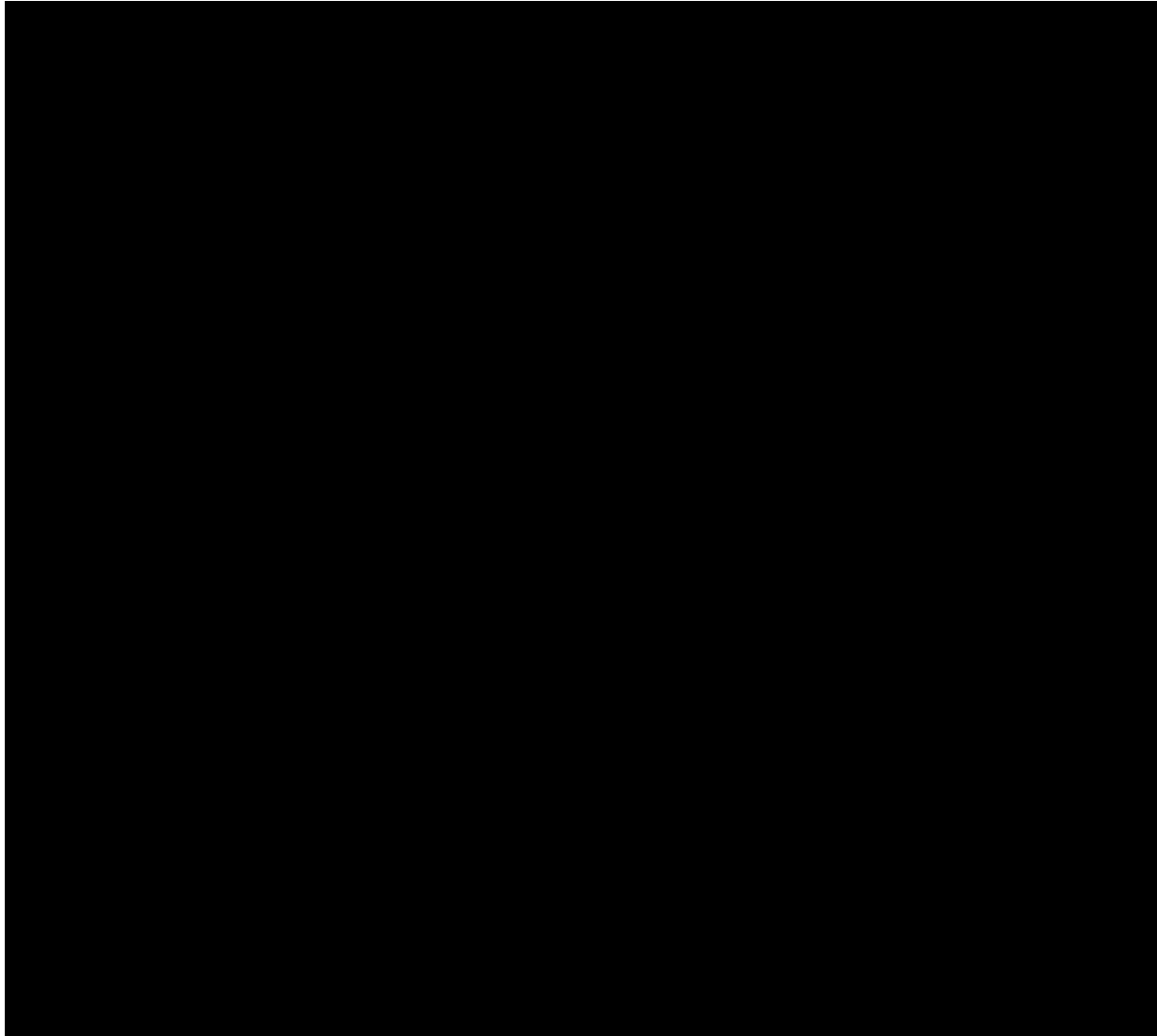


Figure 1 This is the 'front page' for the online resource. The user can search the system by either entering a free text query into the 'Search' box, or by using a series of 'drop down' lists to 'narrow down' the search. Alternatively, the front page provides various navigation starting points in the form of thumbnail images and categories, e.g. allowing the user to look at formats in the database according to whether they are analogue or digital. Note how the front page is now PrestoSpace branded and contains a short explanation of what the tool is for.



**Figure 2 . This screenshot shows the typical results of a search, in this case for all the ‘discs’ in the database. The user is presented with a set of ‘thumbnails’ of the matching formats. Each ‘match’ has some summary information under it providing the name and size of the media and the period it was manufactured for. The user can click on a thumbnail to get further information. The search terms used to get to the results are maintained on the right hand side to allow the user to modify their search or add further terms.**



Figure 3 Clicking on a result, in this case '10inch 78rpm record', brings up detailed information on a media format. The first part of the details page is shown above. This shows a high resolution image of the particular format and provides a link to the PrestoSpace wiki where the user can find further information from other users or provide comments of their own.

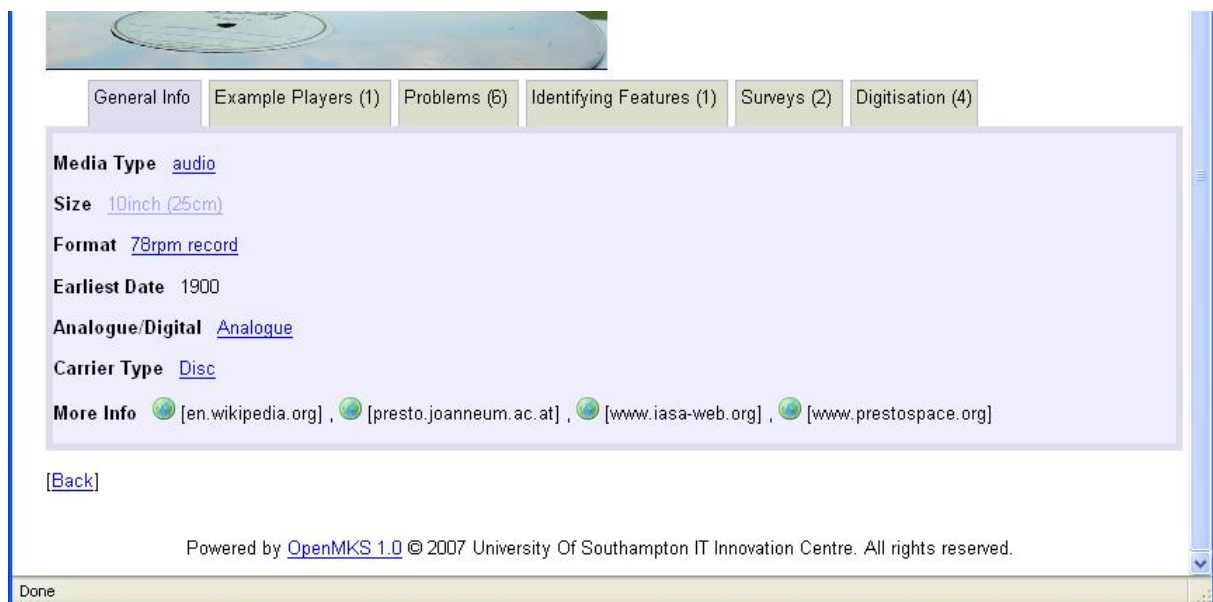
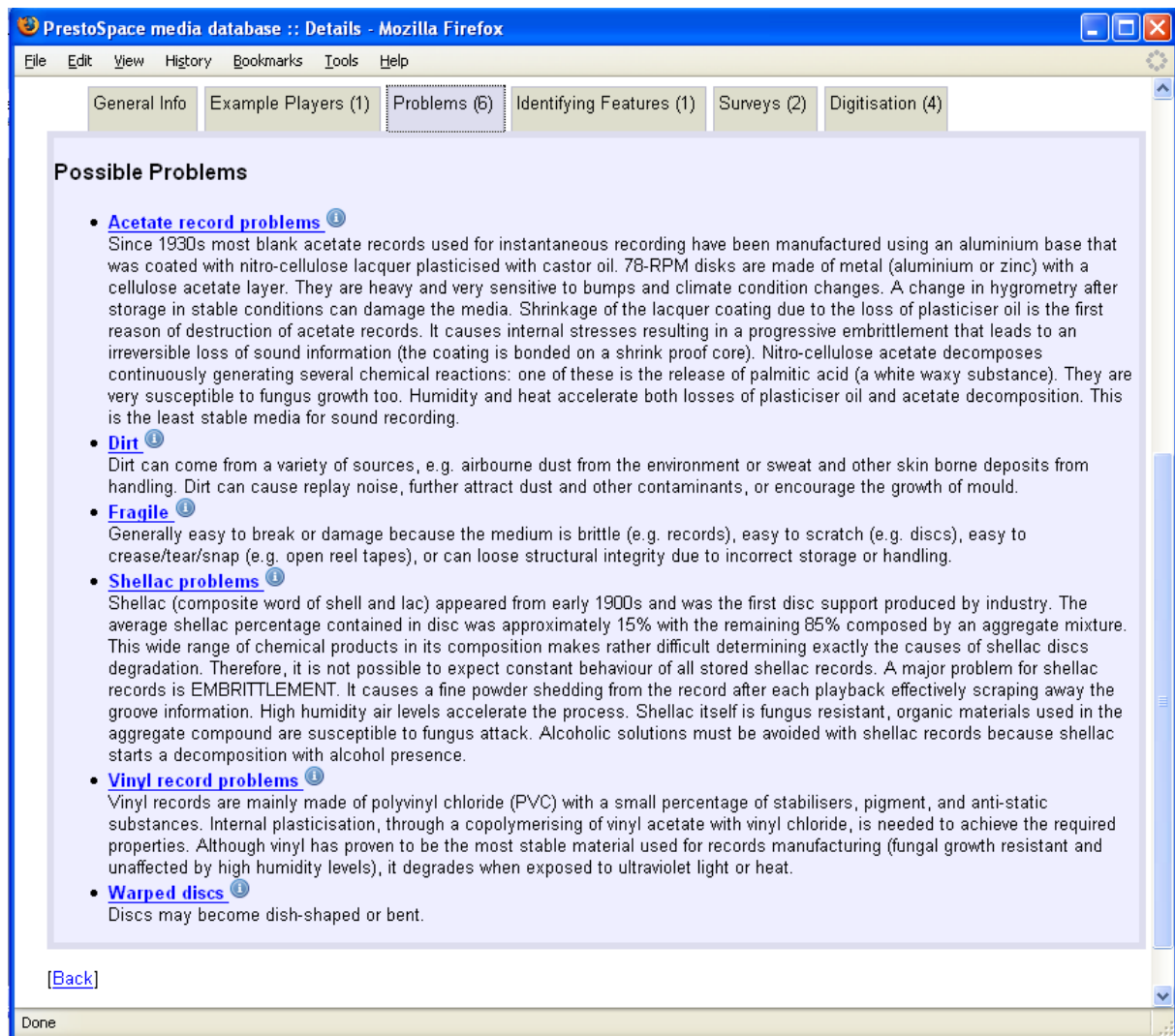
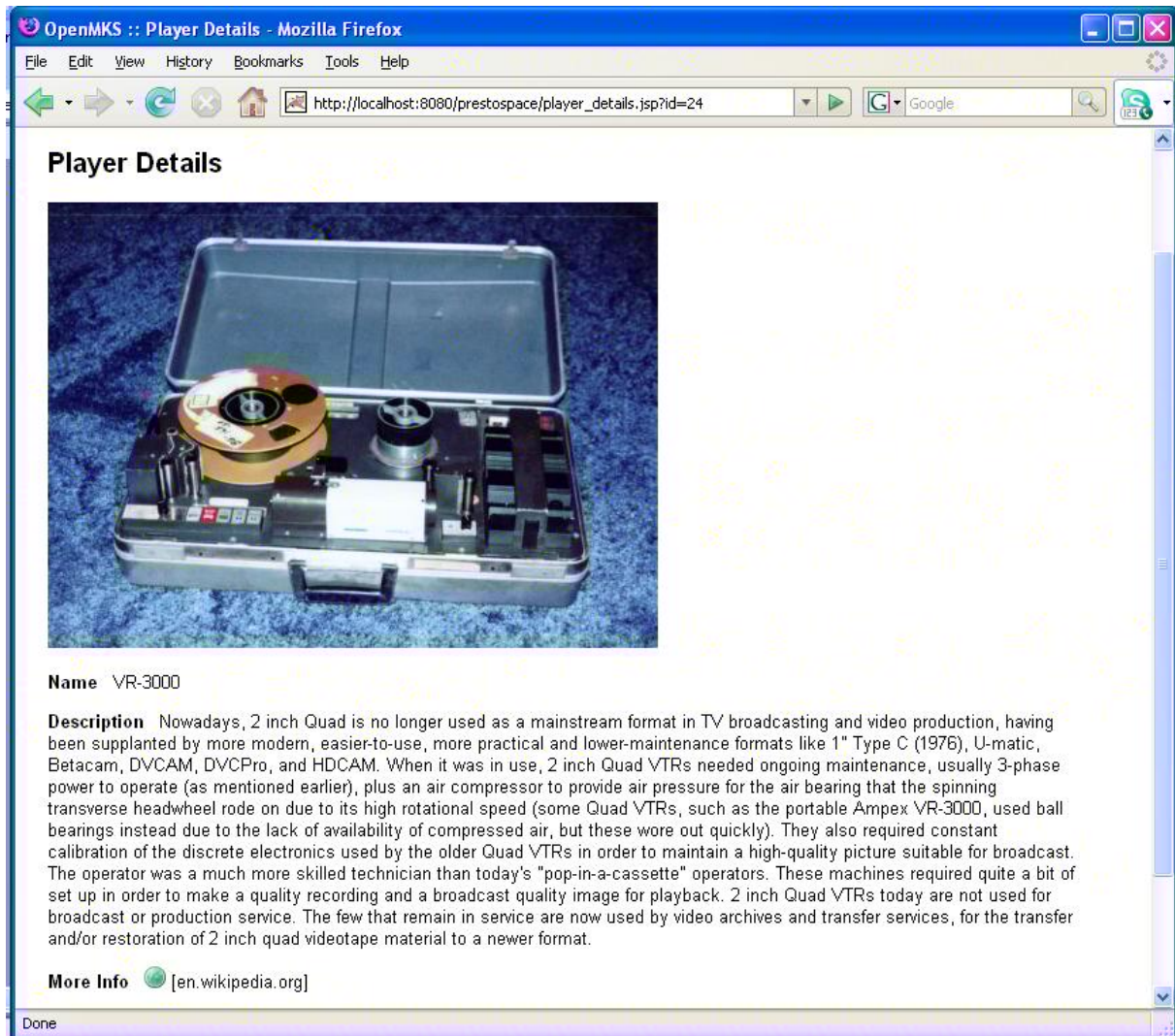


Figure 4 The lower part of the details page for a particular media format provides a series of tabs for further information. The general info tab shown above provides summary information. Note how most of this information is linked. When the user clicks on one of these links they are presented with other entries in the database with the same characteristics. This allows the user to navigate the database.




**Figure 5** The 'problems' tab provides more information on the specific type of problems that can occur for the particular media format. Each problem is linked to other entries in the database that also have the same problems. Each problem also has a small 'i' circle next to it that links to PrestoSpace reports or other web resources that provide more details.





**Player Details**



**Name** VR-3000

**Description** Nowadays, 2 inch Quad is no longer used as a mainstream format in TV broadcasting and video production, having been supplanted by more modern, easier-to-use, more practical and lower-maintenance formats like 1" Type C (1976), U-matic, Betacam, DVCAM, DVCPro, and HDCAM. When it was in use, 2 inch Quad VTRs needed ongoing maintenance, usually 3-phase power to operate (as mentioned earlier), plus an air compressor to provide air pressure for the air bearing that the spinning transverse headwheel rode on due to its high rotational speed (some Quad VTRs, such as the portable Ampex VR-3000, used ball bearings instead due to the lack of availability of compressed air, but these wore out quickly). They also required constant calibration of the discrete electronics used by the older Quad VTRs in order to maintain a high-quality picture suitable for broadcast. The operator was a much more skilled technician than today's "pop-in-a-cassette" operators. These machines required quite a bit of set up in order to make a quality recording and a broadcast quality image for playback. 2 inch Quad VTRs today are not used for broadcast or production service. The few that remain in service are now used by video archives and transfer services, for the transfer and/or restoration of 2 inch quad videotape material to a newer format.

**More Info** [\[en.wikipedia.org\]](http://en.wikipedia.org)

Done


**Figure 6** Some media formats have information on example players (in this case a VR-3000 player for Quadruplex video tape).



The screenshot shows a Mozilla Firefox browser window displaying the PrestoSpace media database. The page title is "1/2inch (13mm) Betacam". Below the title, there is a link to the PrestoSpace Wiki. A photograph of a Betacam cassette tape is shown on a green grid. Below the photo is a link to the "Image Source". There are five tabs: "General Info", "Problems (6)", "Identifying Features (6)", "Surveys (2)", and "Digitisation (1)". The "Identifying Features" tab is selected, showing a list of identifying features for Betacam tapes.

## 1/2inch (13mm) Betacam

[Visit the PrestoSpace Wiki to leave comments or see what others have written about this media](#)



[Image Source](#)

General Info Problems (6) **Identifying Features (6)** Surveys (2) Digitisation (1)

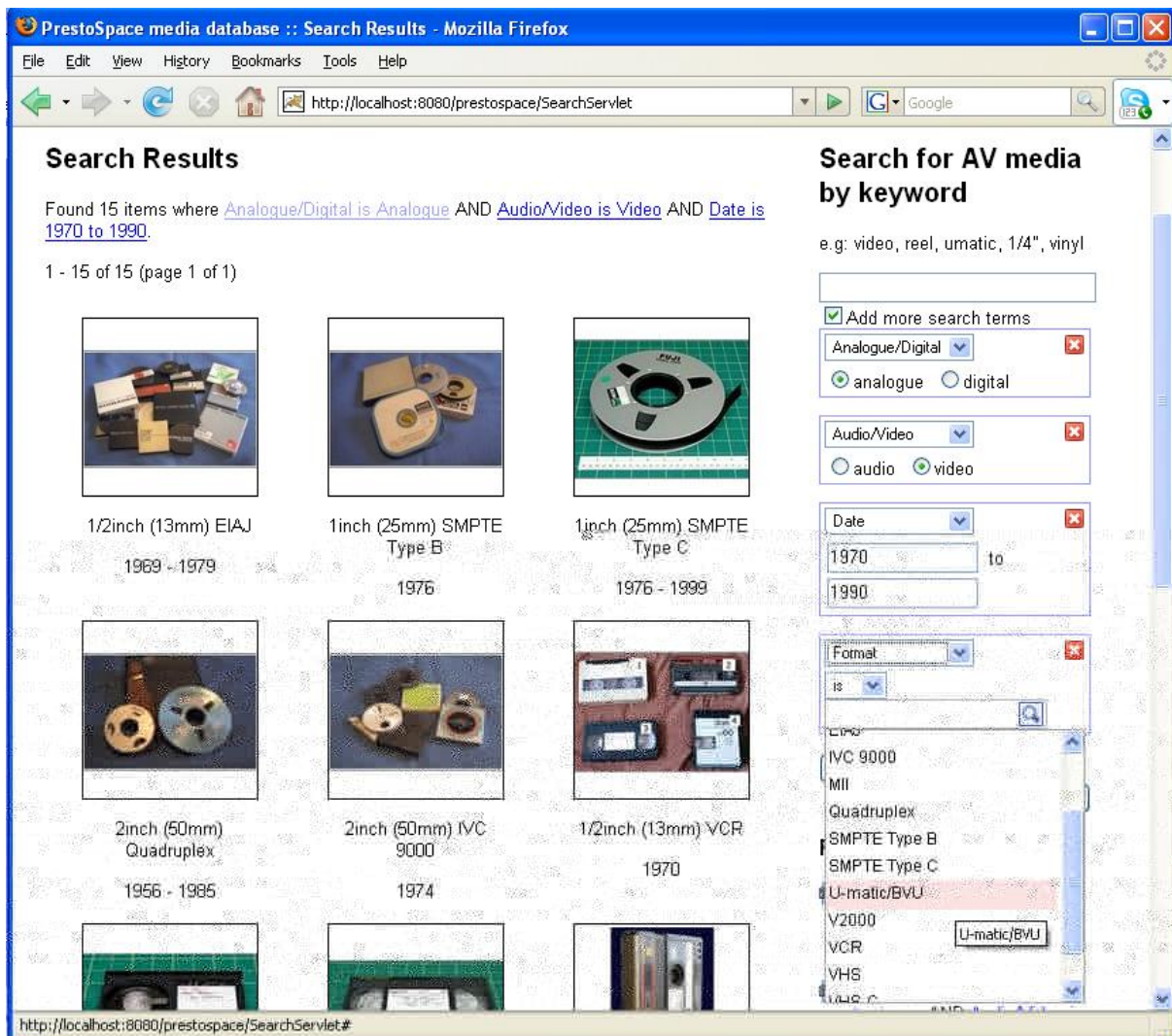
### Identifying Features

- Cassette dimensions: Large cassettes are 9 5/16" x 5 11/16" x 1. They are used in recording and editing decks, and will record up to 194 minutes in length. Small cassettes measure 6 1/8" x 3 3/4" x 15/16" and are for use in cameras, recording up to 62 minutes in length.
- Tape container: Most common tape containers are hard plastic snap-closure boxes. Large cassette cases are 10 5/8" x 6 3/8" x 1 1/4". Small cassette cases are 6 3/4" x 4 3/8" x 1 3/16".
- Tape variations and/or identifying features: Tape cassettes are generally labeled in the upper right-hand corner as Betacam or BetacamSP. Tape length is often written on the anti-static cover (door) that retracts to reveal the tape.
- Sony cassettes are dark blue with a black anti-static cover
- Maxell cassettes are dark gray with purple anti-static cover.
- small Betacam and BetacamSP cassette tapes are the same size as Betamax tapes, so look carefully for a distinguishing logo.

[\[Back\]](#)

Done

Figure 7 Some of the media formats include details of how to identify them (in this case, half inch Betacam tapes).



**Figure 8** If a user wants to search the database in a structured way then they can use the advanced search facility to set a series of search terms (shown on the right hand side). These terms can be added one at a time during the searching process until the user has narrowed down their search enough to find exactly what they are looking for. In the screenshot above, the user has first defined that they are looking for analogue formats, then that they are looking for video, and then that they are only interested in media manufactured between 1970 and 1990. The results are shown on the left. The user can then further narrow the search, e.g. by as shown in the screenshot where the user is about to choose U-matic.

PrestoSpace media database :: Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://localhost:8080/prestospace/SearchServlet




PrestoSpace

Home Go to PrestoSpace Home Page Help


### Search Results

Found 4 items for [analogue](#) AND [tape](#) AND [cassette](#) AND [audio](#) .

1 - 4 of 4 (page 1 of 1)

		
0.15inch Compact Cassette	16mm Tefifon Cassette	0.15inch Micro cassette
1963	1948 - 1965	1968


1/4inch (6mm) Elcaset
1976 - 1990

http://www.prestospace.org/?\_sourcePage=/results.jsp

### Search for AV media by keyword

e.g: video, reel, umatic, 1/4", vinyl

analogue tape cassette audio

Add more search terms

Search

### Previous Searches

- [digital AND tape AND cassette AND audio](#)
- [analogue AND tape AND cassette AND audio](#)
- [analogue AND tape AND cassette](#)
- [where Analogue/Digital is Analogue AND Audio/Video is Video AND Date is 1970 to 1990](#)
- [where Analogue/Digital is Analogue AND Audio/Video is Video AND Date is 1970 to 1990 AND Format is U-matic/BVU](#)
- [reel](#)
- [where Problems is Acetate](#)

**Figure 9** An alternative to structured searching is to put a series of key words into the free text box. In this case, the terms 'analogue', 'audio', 'cassette' and 'tape' match 4 entries in the database shown on the left. Note how this screenshot also shows the previous searches done by the user so it is easy for them to go back to searches they have done before, either in whole or in part.

## 5. Original Motivation

This section contains a description of what motivated the development of the web-based resource described above.

From a series of meetings involving small audiovisual and non-broadcast archives, as well as from the results of various surveys into archive holdings, it rapidly became clear to IT Innovation that small archives don't always have a good understanding of the technical aspects of their holdings (carriers, formats, age, condition etc.), or what the options are for moving this content into the digital domain (e.g. codecs, bit rates, storage options), and then how to store and manage this digital content in the long-term (disks or tapes, shelves or mass storage).

At the STAG Small Archives workshop in London, a 'wailing wall' was set up to allow small archives to voice their questions/issues. Posted items included: "As a non

expert, how am I supposed to know what software/ hardware to buy? Everyone seems to recommend different things!", "What is the best audio storage if '.wav' files are not an option? (lack of IT support etc.)", and "Why can't we transfer once to a digital medium that will last for centuries".

At the PrestoSpace user meeting for small archives, questions raised by the participants included "what digital formats should I use for audio?", "what should I do with umatic – compressed or uncompressed?", "what are the problems with 2 inch?", "how should I store my tapes".

These are just a few examples that show the general level of understanding and awareness that small archives have about their holdings and what to do with them.

Awareness is one of the biggest problems for small archives and the results of the survey by the TAPE project, for example as reviewed in the 2005 PrestoSpace Annual report on AV collections <sup>1</sup> show that small archives do not know what they have and what problems they face.

As with so many things, the devil is in the detail, and whilst a small archives might be able to identify that they have some 1" tapes from the early 1970s and they might recognise that the tape player is now obsolete, they may not have the expertise to know that there's a big difference in practical terms between having 1" tape that's type A, B, C or IVC<sup>2</sup>. Likewise for audio tape, it's easy to identify an item as being 1/4" open reel, but much harder to work out the tape speed, number of tracks, equalisation, use of noise reduction etc., all of which that all impact on how to transfer the tape. For digital formats it's easy to say that uncompressed or lossless encoding is best, but for video/film there are tradeoffs with storage volume that may mean that some degree of compression is necessary. This can have an impact on quality or the ease with which content can be edited processed, e.g. the use of intra-frame compression (DV) that allows easy editing using AVI files or the use of inter-frame compression (MPEG2) that takes up less space on a tape, but needs conversion before editing. Different types of encoding may cope better or worse with things like film grain, and some encoding types are more suitable for delivery (e.g. over the web) than they are for archiving.

The answers to these questions are needed before a small archive can use some of the other PrestoSpace tools for predicting the costs of a digitisation and storage projects<sup>3</sup>. For example, carrier type and digitisation bit rate are both needed as inputs to the storage ready-reckoner, and probability of playability problems and the need for web delivery are also needed as inputs to the preservation cost calculator.

---

<sup>1</sup> PrestoSpace Second Annual Report on European AV collections "Deliverable D22.6 (2005)

Preservation Status: Annual Report on Preservation Issues for European Audiovisual Collections

<sup>2</sup> 1 inch IVC has few available machines, often tapes are 'sticky', and there is very low compatibility between machines, so this format is problematic. Likewise, 1 inch A also has few available machines and sticky tapes are often impossible to preserve (the likelihood of stickiness varies between manufacturers). On the other hand, 1 inch B has good compatibility between machines and good quality machines are readily available. There are no major problems with tape condition, except from Agfa. Likewise, 1 inch C is also generally easy to migrate. See PrestoSpace Deliverable D5.3 for more details

<sup>3</sup> <http://prestospace-sam.ssl.co.uk/>

Of course, there are already plenty of information sources that can help answer these questions. These include:

- PrestoSpace deliverables, e.g. the annual reports on AV collections, and specific reports such as D5.3 which describes difficult media (almost all media is difficult sooner or later!) and some of the problems that occur when dealing with it.
- Surveys, for example the IASA 1995 and 2003<sup>4</sup> “Survey of Endangered Audiovisual Carriers” that cover AV archives across the world and include statistics on the types of media held and the condition it is in (or more accurately is perceived to be in).
- Guidelines and reports, e.g. the IASA TC-04 guidelines for the production and preservation of digital audio objects that describes in detail a range of analogue carriers, how they can be digitised and the appropriate options for their digital surrogates.
- Websites, identification guides and factsheets, including specific guides such as the VidiPax guide to video formats<sup>5</sup> and the Texas Tape Identification and Assessment Guide<sup>6</sup>, and more general resources including the PrestoSpace wiki<sup>7</sup>, the AMIA factsheets<sup>8</sup>, and even a large body of information on wikipedia.

These information sources provide information on:

- Carriers and their identification, e.g. what does a umatic cassette look like, or how can a nitrate and acetate film be told apart.
- Problems, e.g. what is sticky shed syndrome and what problems does it cause, or what is tape ‘popping’ and ‘stacking’.
- Treatments, e.g. what is ‘baking’ and when should it be used (or not), or how can wet gates help with scratch removal in film scanning.
- Storage, e.g. what temperature and humidity is suitable for video tape, or do you store tape ‘tail in or tail out’.
- Priorities, e.g. what are the most urgent carriers to transfer due to obsolescence of players, or what tape manufactures tend to have the most problems with head clogging.

---

<sup>4</sup> IASA Survey of Endangered Audiovisual Carriers, 2003 [http://portal.unesco.org/ci/en/ev.php-URL\\_ID=13437&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=13437&URL_DO=DO_TOPIC&URL_SECTION=201.html)

<sup>5</sup> VidiPax guides (audio, video) <http://www.vidipax.com/>

<sup>6</sup> Texas Commission on the Arts, Video Tape Identification and Assessment Guide <http://www.arts.state.tx.us/video/>

<sup>7</sup> <http://wiki.prestospace.org/>

<sup>8</sup> AMIA Video Tape Preservation Factsheet

<http://www.amianet.org/publication/resources/guidelines/videofacts/commonprobs.html>

- Players, e.g. what can you use to transfer a 1" C tape and how do you operate it.
- Digital formats, e.g. what file format should be used when transferring a low band umatic and at what bit rate.
- Digital storage, e.g. what is the break point between tape based and disk

the decisions that may be made on the basis of the content, e.g. the priorities and cost of a digitisation project.

## 6. Approach

---

Our approach has been to create a structured information resource that pulls together information from many of the existing sources and allows it to be searched and navigated in a consistent way, and most importantly in a way that can answer the type of questions that small archives ask.

In particular, we've actively avoided creating yet another big report, or a 'static' website with a lot of pages that are hard to maintain and navigate. Instead, by adopting a database approach with a search and navigation interface, it becomes much easier for users to find specific information and then allow these users to follow links out to external websites or reports. Importantly, since the data in the database is available to be checked by the PrestoSpace partners, our approach also allows us to offer a degree of quality assurance for the data and hence confidence in the contents of the database for the end users.

The database contains factual data (e.g. the dates when 1" C tapes were used), textual information (e.g. description of how to identify a 2" quad tape), and multimedia content (e.g. a photograph of a umatic cassette, a graph of how often problems occur for different carriers, or a video of how to operate a 1" playback machine).

The database is accompanied by a search and navigation tool (already existing from another European project) that allows the information resource to be accessed and explored from different view points that match the different questions that archive owners want to ask.

The information resource is targeted at assisting small archive owners and managers when they want to determine:

- what they have;
- whether it is at risk;
- what problems might exist when transferring it to digital form;
- what digital formats to use;
- how much digital storage they will need; and
- how long it will take.

The information that small archives will be able to extract from the resource will help them make more effective use of the SAM calculators by allowing them to make informed decisions on what parameters to use.

Below are some of the questions that we expect the resource to be able to answer. The key thing is to allow the user search/navigate/explore the information in various



ways to make it easy to answer the specific questions they have. These questions are many and varied. For example:

- An archive knows roughly the age of their video tapes, the number of hours they have, and that the video is cassette rather than open reel. They want to know the specific format and what digital file formats would be suitable for master copy encoding.
- An archive has a very heterogeneous collection and wants to know which formats/manufacturers are problematic for a particular age range (e.g. 1970 – 1980) so they can survey their collection to see if they have matching items that are at risk.
- An archive knows the specific format of their tapes (e.g. a collection of 1” C tapes that they have recently acquired) but not their age or what is needed to play them back. They want to know how old the tapes might be and the risk that they are not playable.
- An archive has a particular type of tape deck, but has lost the expertise to operate it. They want to find videos and technical manuals that show them what to do to get the best results.
- An archive wants to know how widely used a particular format was so they can assess whether they might be able to find, and cooperate with, other archives that have similar holdings.
- An archive knows that they have a collection of 1” tapes, but they do not know that 1” comes in different flavours (A,B,C) that can present very different problems. They want to know how to identify what flavour they have, what problems commonly occur for this flavour and what treatments might be necessary.
- An archive knows that their 1” tapes are mostly in good condition since they often play back without problems on an in-house machine. They want to know whether the collection is at risk due to technical obsolescence and the likelihood of being able to play back the tapes on other machines.
- An archive wants to know what encoding formats make sense when migrating a specific type of tape and what bit rates are sensible (e.g. 1” video could be encoded into MPEG2 at a minimum 50MBit/s for a master copy)
- An archive is considering using a common digital format when digitising a range of different analogue media and wants to know whether there would be any problems (e.g. use of mp3 at a certain bit rate is OK for ¼” audio tapes but not OK for 8 track professional tapes)
- An archive wants to know what type of problems occur for a particular type of item and how to spot them (e.g. for open reel tape you might come across popping, pack slip, mould, sticky shed etc.).

There are many other scenarios like these where the user wants to dip into what is a large information space in order to find out the answer to their questions.

The questions are likely to be asked in different contexts, for example:

- What parts of my archive are at risk?
- What is this box of stuff that someone’s just given me to look after?
- How much digital storage would I need if I migrated this collection?

The user may want to 'go in' via one part of the space (e.g. formats/manufacturers/age/volumes) but 'come out' in another (e.g. condition/treatments/obsolescence). This allows them to explore and use the links between these areas, e.g. so they can go from analogue carriers to appropriate digital formats/bit rates/storage technologies.