Digital Folklore Preserved for the Future

A Case Study of the Use of Planets in Preserving the Digital Collection of the Danish Folklore Archives, The Royal Library, Denmark

planets

Executive Summary

The Danish Folklore Archives are a newly added branch of the Royal Library in Copenhagen, which is the national library of Denmark. The Folklore Archives hold a rich collection of manuscripts, images and sound recordings of intangible culture, such as songs and stories, primarily from Denmark but also from other parts of the world. They are engaged in digitising their collection, both to provide preservation copies of fragile material and to enable greater access. This case study details their experience in planning for the long term preservation of their digital collections.

In order to plan the transition of the digitised collection to the digital long-term preservation store of the Royal Library, it was necessary to obtain a profile of the digital objects and identify whether any preservation actions were needed before transferring them to the preservation store. Plato, the preservation planning tool developed by the Planets project, was chosen to support and guide them in this work. A project team was established, consisting of an IT analyst from the Digital Preservation team at the Royal Library and a sound engineer from the Folklore Archives.

The team has used Plato to provide automated descriptions of the technical characteristics of their digital audio collection (characterisation) and to help them specify their preservation requirements for this material. They have also used Plato to characterise their digital image collection and have completed the preservation plan for this.

They found Plato's structured workflow very useful and it guided them through the preservation decision making process. Plato was particularly helpful for specifying their preservation policy requirements, and it has helped to identify a number of issues regarding the digital materials, for example the sound quality of older digitised sound files, and the potential loss of original creation date from image files.

Plato has provided the team with an evidence base that underlies their preservation planning decisions, and also detailed documentation of the preservation plans and the decision making process.

The Folklore Archives and the Digital Preservation Team at the Royal Library believe Plato was indispensable in providing a profile of the digital collection and helping produce preservation plans and they regard Plato to be the obvious preservation planning tool in the years to come.

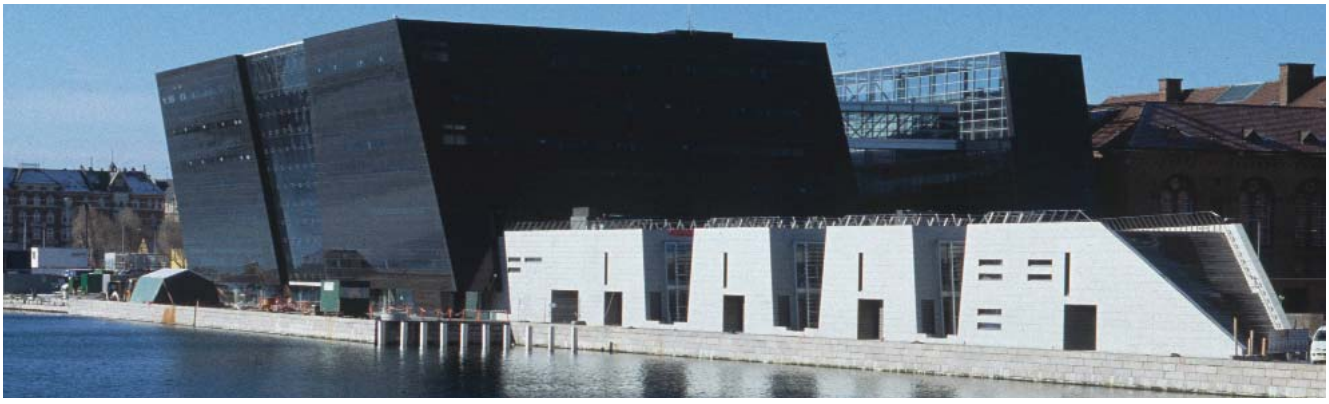
Authors
Lynne Chivers, The British Library
Marianne Holm Pedersen, Danish Folklore Archives, The Royal Library
Tue Hejlskov Larsen, The Royal Library
Emmerik Warburg, Danish Folklore Archives, The Royal Library

This case study is part of a series of case studies on the application of Planets in major European libraries and archives.

They are all available via the Planets website, **www.planets-project.eu**

The Royal Library, Denmark and The Danish Folklore Archives

DET KONGELIGE BIBLIOTEK



The Danish Folklore Archives, Copenhagen. Photographic Studio, Royal Library

The Royal Library in Copenhagen is the National Library of Denmark and administers the national cultural heritage of both Danish and foreign origin in terms of published works (books, periodicals, newspapers, leaflets), manuscripts, documents, maps, pictures, photographs and music in conventional or digital form[1]. The Royal Library provides access to the collections for the purpose of research and study, while at the same time making sure that the collections are preserved for posterity.

In 2008, the Danish Folklore Archives[2] became part of The Royal Library. The collection of the Folklore Archives comprises three archives and a library, and includes more than 100 years of songs, music, stories, and other intangible cultural materials. The collections are primarily from Denmark, but also contain material from the rest of the world.

Sources in the manuscript archive include, among other things, field records, letters, and diaries regarding topics such as belief and magic; customs, rites, and ways of life, fairy tales, legends and other stories. The manuscript archive also holds a large collection of ballads, songs and song games, including more than 10,000 tunes written in 73 handwritten notebooks.

The sound archive includes more than 3,300 quarter inch reel-to-reel analogue tapes of field recordings from Denmark and the Faroe Islands. These tapes were recorded by researchers from the Folklore Archives. The sound archive also holds 855 wax cylinders, and a large variety of other recordings from Denmark and other countries. Finally, the image archive holds photographic negatives, pictures, paintings and much more.

Peder Jensen, Storyteller. Photo Olaf Sand Kristensen; Danish Folklore Archives

Postcard from 1918 depicting two pixies on a farm; Danish Folklore Archives

[1] http://www.kb.dk/en
[2] http://www.dafos.dk/

The Danish Folklore Archives are used by a broad audience consisting of researchers, students, and representatives from local historical archives all over Denmark and around the world, and also members of the general public interested in intangible culture. Researchers using the archives come from such different fields as History, Ethnology, Danish, Anthropology, Musicology, and the Classical Music Colleges and Conservatoires. Their work has covered varied topics such as Greenlandic drum dances, street cries, children's musical games[3] and folk music, for example a recent book about Nordic perspectives on traditional singing (Lene Halskov Hansen, Astrid Nora Ressem og Ingrid Åkesson (eds) 2009).

The recorded sound materials have been published on a number of records and audio CD's, for example "Music in Bahrain. Traditional Music of the Arabian Gulf" by Poul Rovsing Olsen (2002). This book contains three CD's of recordings from the Folklore Archives. These recordings contain unique material of Arab folk singers, and The Danish Folklore Archives are the only place in the world where the original recording of this music can be found.

The children's musical game "Når vi sejler op og ned" ("When we sail up and down"). Photo Per Melsvig; Danish Folklore Archives

"Music in Bahrain. Traditional Music of the Arabian Gulf" by Poul Rovsing Olsen (2002). Photo Photographic Studio, Royal Library.

The collections of the Danish Folklore Archives are of international importance:

«The Danish Folklore Archives have the largest collection of mediaeval ballads across the Nordic countries, including a great many 19th Century manuscripts. The Nordic ballads are all the same genre, so it is not meaningful to research them from a national point of view. In addition to manuscripts and publications, the Archives have a very large and varied Sound Archive of recorded music. I am looking forward to hearing the results of the digitisation of wax cylinder recordings. Their collections are very valuable materials for research internationally and it is vital they are preserved.»

Velle Espeland, Head of the Norwegian Folk and Popular Song Archives[4]

«We meet regularly with colleagues from the Danish Folklore Archives and refer our researchers there. Categories of music go beyond their country of origin and are influenced by others. The collections of the Danish Folklore Archives are so much a part of telling of our history. It is important not only to preserve the collections, but to have experts who can explain them and make them available. Digitising the often fragile collections helps to preserve them and also makes access to the material easier for readers and for publication. Preservation of the digital copies is therefore very important.»

Dan Lundberg, Director of the Centre for Swedish Folk Music and Jazz Research, National Collections of Music in Sweden[5]

[3] http://www.dafos.dk/brug-arkivet/kilder-online/boernetraditioner.aspx

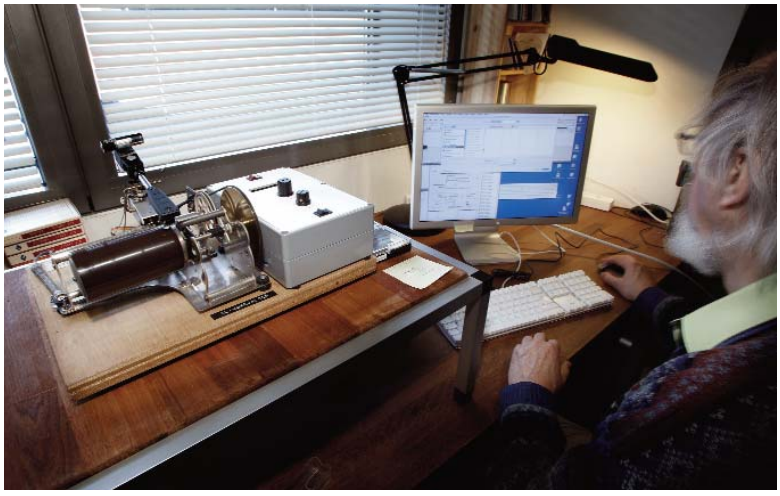
The Digital Collection

None of the digital collection of the Danish Folklore Archives is born-digital, and staff are engaged in digitising the archives, including analogue audio and video items and images. Digitisation began in order to create preservation copies, and now the Folklore Archives are also using digitisation to improve access[6]. For example, most audio items on tape are only available on site, and if they are not digitised staff have to go to the storage area, find the appropriate tape, put it in the machine and play it for the user. Digitisation allows the user to click on a file and listen to the recording – or many recordings at once. It is very likely that in future more files will also be accessible via the Folklore Archives' website.

The digital audio collection currently consists of approximately 7,710 files in wav format. The image collection contains 400 images in .jpg and .tiff formats. The video collection so far only holds approximately 100 files. There are many more items in the archive still not in digital form, and so the digitisation process at the Folklore Archives will continue over the coming years.

[4] http://www.visearkivet.no/sider/in_english.html
[5] http://www.visarkiv.se/en/index.htm
[6] Samples of images, sound and text are available on the Danish Folklore Archives website at http://www.dafos.dk/guiden/#

Selection of media from the Danish Folklore Archives.
Photo Photographic Studio, Royal Library.



Digitisation of wax cylinders at the Danish Folklore Archives
Photo Photographic Studio, Royal Library

Emmerik Warburg, Sound Engineer, in the sound studio at the Danish Folklore Archives. Photo Photographic Studio, Royal Library

The Need for Preservation Planning

As a newly added branch of The Royal Library, the Danish Folklore Archives needed to plan the transition of their collection into the framework of the Royal Library. Eventually, the digital collection will be moved to the digital long-term preservation archive of the Royal Library. First they had to get an overview of the collection's digital objects and identify whether all could be transferred to the repository in their current state, or if further actions had to be taken first. These actions could include extraction of metadata from the digital objects, addition of metadata, migration to formats suitable for long term preservation, for example.

To do this, it was necessary to determine exactly what was held in the digital collection. This was to be done by producing a profile of the digital collection and characterising the digital files. In other words, by discovering the formats of the digital objects in the collection, validating their compliance to those formats, and then determining the format-specific significant properties of those objects. It would then be possible to produce a preservation plan for the digital collection.

The Folklore Archives were particularly keen to characterise their collection so that they could be confident of little or no information loss in future migrations. It was already known that during previous migrations of their sound files, some unexpected corruption had occurred, in one case losing some technical metadata embedded in file headers and in another rendering some files unusable (see fig. 1).

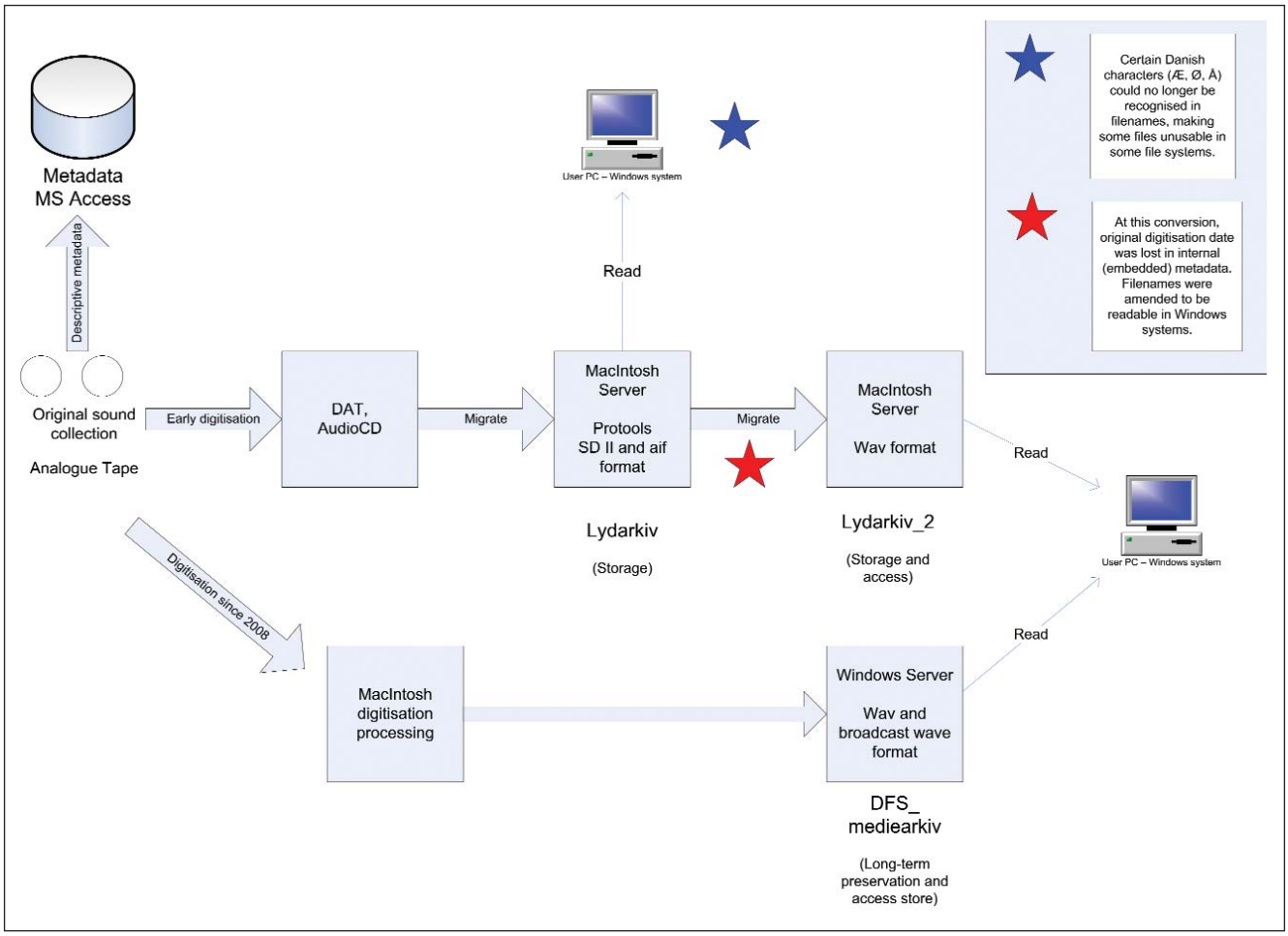


Figure 1 The digitisation of recorded sound on tape at the Danish Folklore Archives

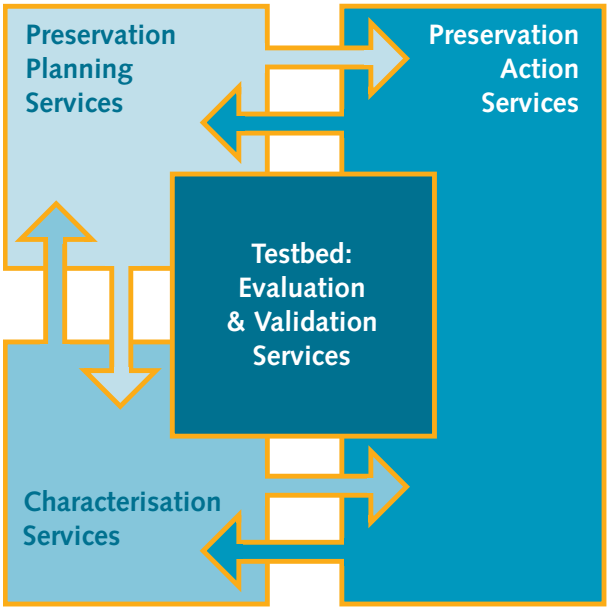
The Solution

As the Royal Library had limited experience of working with audio materials and the Folklore Archives had limited knowledge of digital preservation, a tool was needed to help structure the work and facilitate the input of all the relevant expertise. Plato[7], developed by the Planets project, was the only preservation planning tool that supported the detailed planning for specific objects which could meet their needs.

The Royal Library is a partner in Planets, a European joint-venture project for research and development in the field of digital preservation, which has produced a framework and set of practical tools and services to enable institutions to manage and access digital collections for the long-term[8]. The Folklore Archives' requirements were just the sort of needs that Planets tools and services were developed to satisfy, and it was natural that the Royal Library should wish to use the Planets products in a practical preservation situation. Members of the Folklore Archives had already attended a Planets training course and were confident that the Planets characterisation and planning tools would help them in their task.

A project team was established, consisting of an IT analyst from the Digital Preservation team at The Royal Library and a sound engineer from Folklore Archives, and work began in January 2010.

What is Planets?



Planets makes it possible to:

- Define your preservation policies and goals
- Assess the preservation needs of your organisation, collection and users
- Identify areas where preservation of your collections does not meet your policy requirements
- Build, evaluate and execute plans to address any problem areas
- Analyse and verify the results
- Document the decisions made and actions taken

Working with Plato

«The process of using a preservation planning tool may be compared to working with a project planning tool. You need to know the objectives in order to get a useful output. Likewise, you need to know what you want to achieve in the preservation planning process. Working with such a tool may raise questions that you may not be able to answer. In our case cooperation was necessary between the Royal Library, which had the necessary IT and preservation experience, and the Folklore Archives' knowledge about the content of the digital material.»

Birgit N. Henriksen, Head of Department, Digital Preservation, The Royal Library, Denmark

[7] http://www.planets-project.eu/software/

[8] For more detailed information about Planets tools and services see http://www.planets-project.eu/docs/comms/PLANETS_BROCHURE.pdf and http://www.planets-project.eu/docs/comms/PLANETS_PRODUCT_SPECIFICATION.pdf

The Plato planning tool guides you through the process of building a preservation plan and workflow. With it, one can identify preservation actions (e.g. migration), evaluate them and select the best option. It produces detailed documentation of the preservation plan and the decision-making process.

The team decided to develop three plans in Plato: audio, image and video. Each plan therefore would cover a reasonably homogeneous set of digital objects. At the time that work began, Plato was still undergoing testing, and so the version of Plato used was not the current publicly-released version[9].

Most effort was concentrated on the audio and image plans. This was because tools to characterise the video formats in use in the Folklore Archives' collection, for example MP4(H264) and Quicktime(.mov), were not supported within Plato at that time.

The Plato Preservation Planning Process

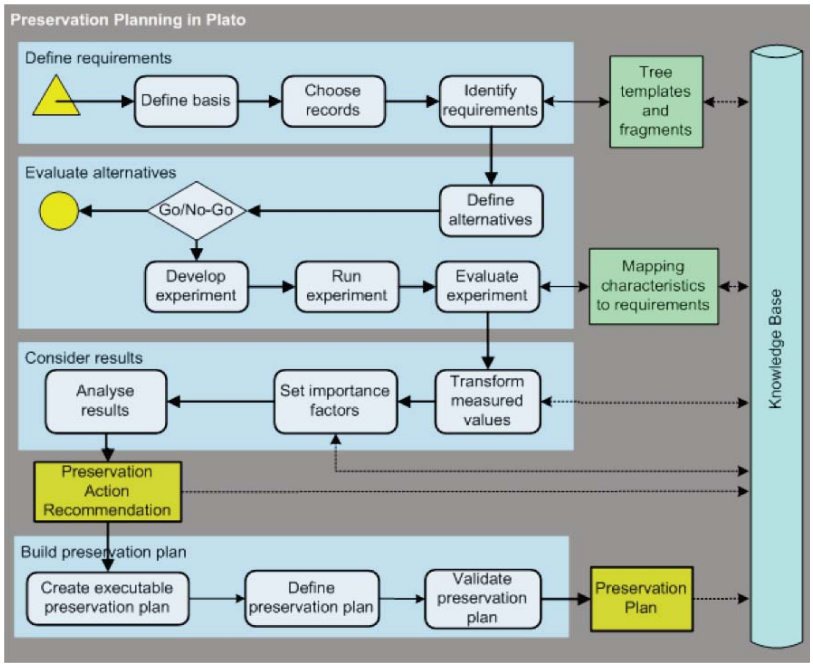
There are four high-level phases in the Plato preservation planning process:

1 **Define requirements**
Specific goals and characteristics are defined for a collection of digital objects, and representative samples are identified.

2 **Evaluate alternatives**
Different preservation actions are tested and evaluated against the defined requirements.

3 **Consider results**
Recommendations are made as to which preservation tool with support the identified requirements for the particular collection of digital objects.

4 **Build preservation plan**
Specify a series of concrete steps or actions, along with organisational responsibilities, rules and conditions for executing the preservation action on the collection in the scope of the plan.



[9] At time of writing, the current version of Plato is release 2.1

The Audio Plan

The team began by working on the preservation plan for the audio collection. The first phase of the Plato preservation planning process, Define requirements, is where constraints and influences on possible preservation actions are documented, a representative sample of digital objects is selected and characterised, and a preservation policy with requirements is defined. During this phase, the team found Plato worked very well. It supported the definition and characterisation process and was well structured, with a logical progression of questions.

An essential part of creating a new preservation plan is defining the policy. The team began this and then used Droid and Jhove[10] to characterise the sample objects they had identified. Then, in order to specify requirements it was necessary to go back and complete the policy, in line with Plato's process.

In Plato, requirements are specified via an objective tree, and this can be done in Plato itself or uploaded to Plato from a mind map objective tree. The team used the template mind map objective tree provided by Planets and the Freemind[11] tool to edit the tree to their requirements, and uploaded it to Plato.

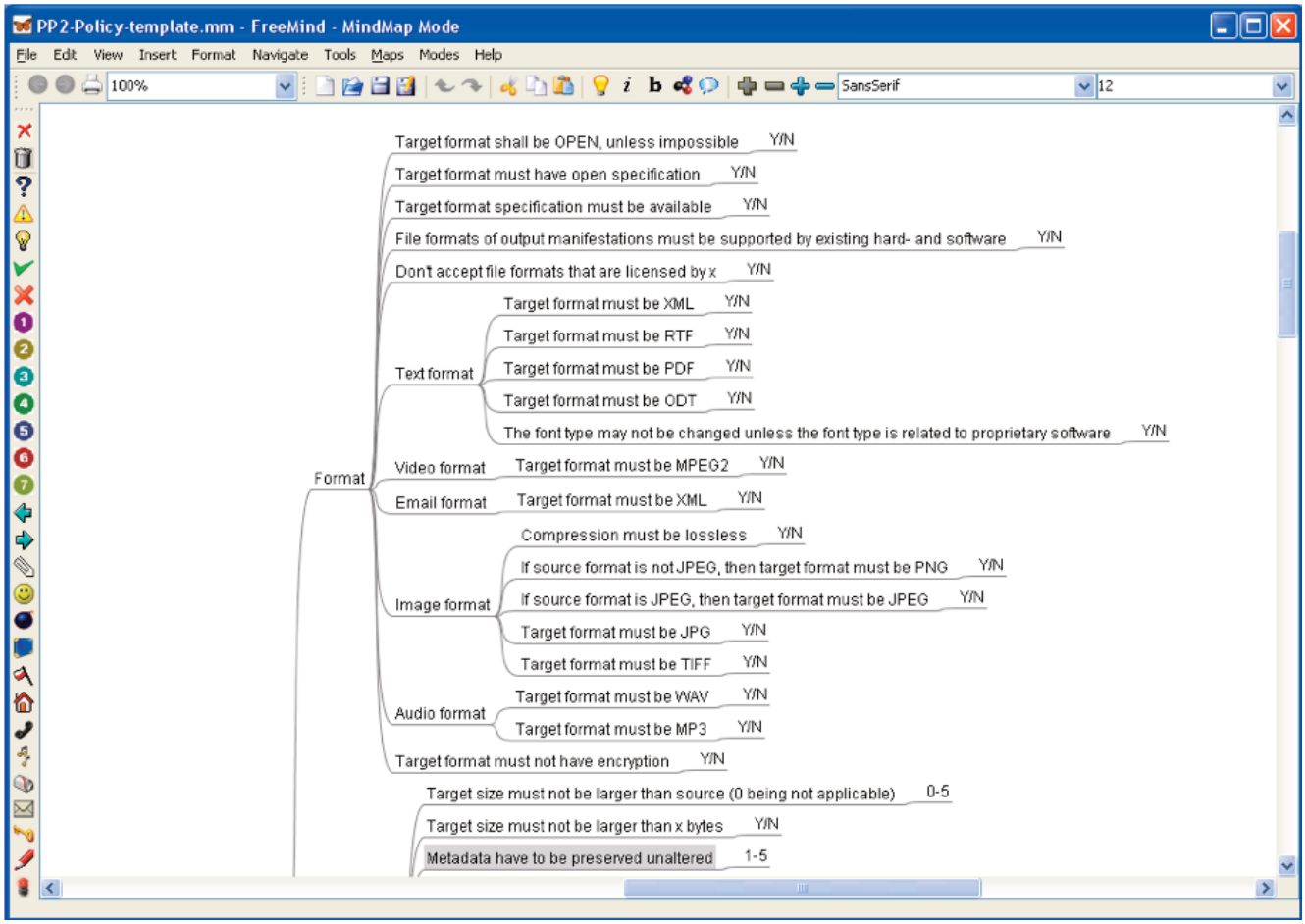


Figure 2 Screen shot showing part of the template objective tree in Freemind tool.

[10] http://sourceforge.net/projects/droid
http://sourceforge.net/apps/mediawiki/droid
http://hul.harvard.edu/jhove

[11] http://freemind.sourceforge.net/wiki/index.php/Main_Page

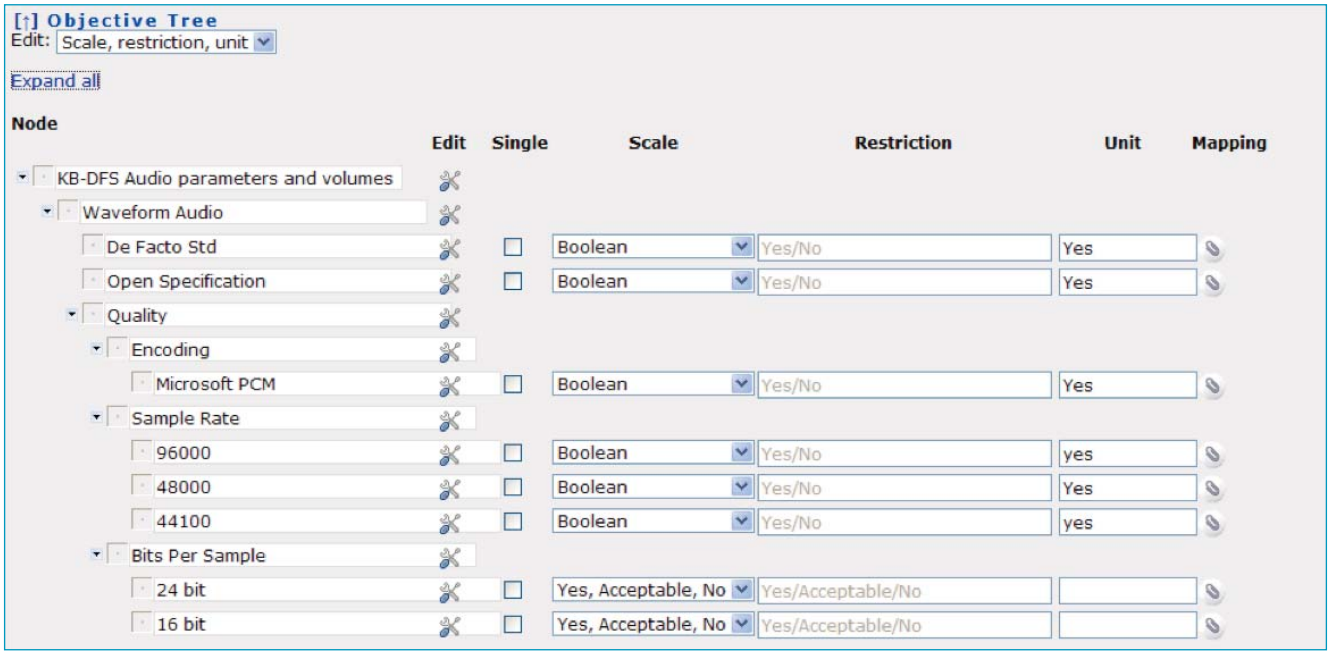


Figure 3 Screen shot of Plato showing part of the requirements objective tree defined by the team for audio files.

At the time of writing, work on the audio plan is continuing. The next task for the team is to analyse the WAV format and its embedded metadata to ensure all the important information is preserved.

The Image Plan

The team completed a preservation plan for image files. Again, the template objective tree provided by Plato was used to help define the policy and requirements, and sample .jpeg and .tiff files were uploaded and characterised. It was specified as a requirement that the preservation format for images must be TIFF.

The team selected the Jhove tool to characterise the sample image files and used Exiftools outside Plato to identify which software version had created them.

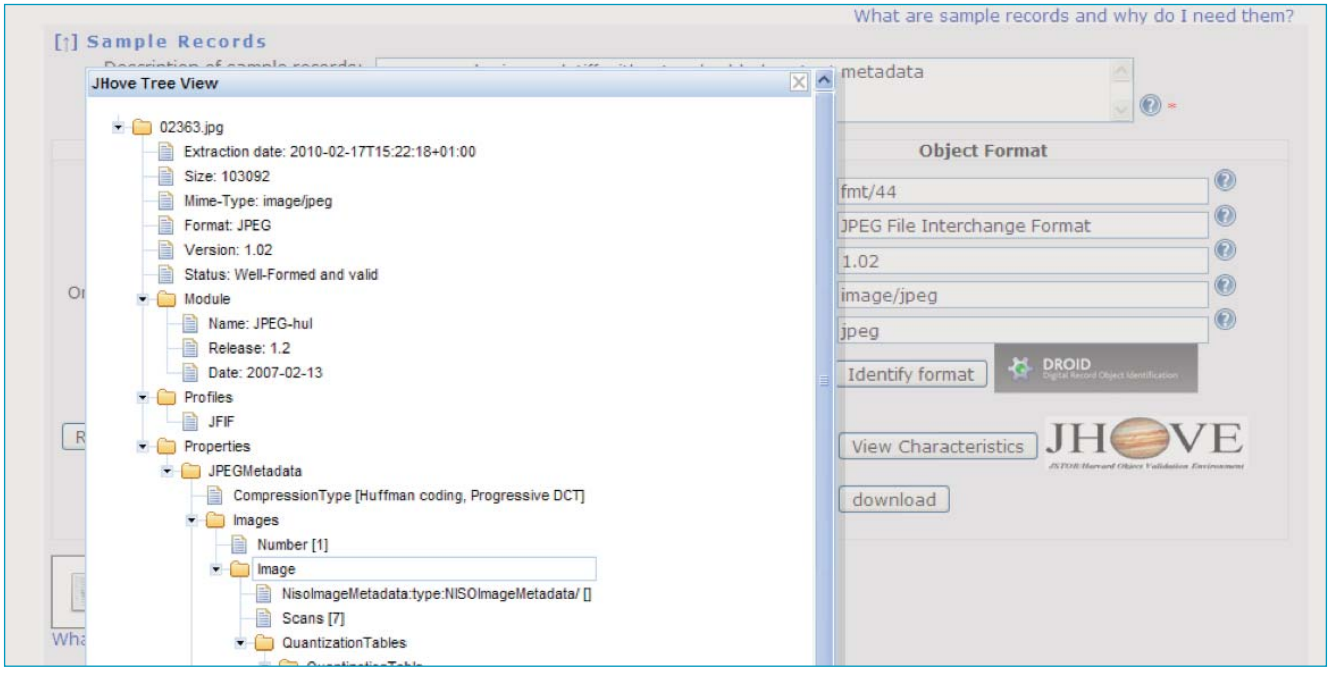


Figure 4 Screen shot of Plato showing part of the Jhove characterisation of a sample .jpg file for the Image plan.

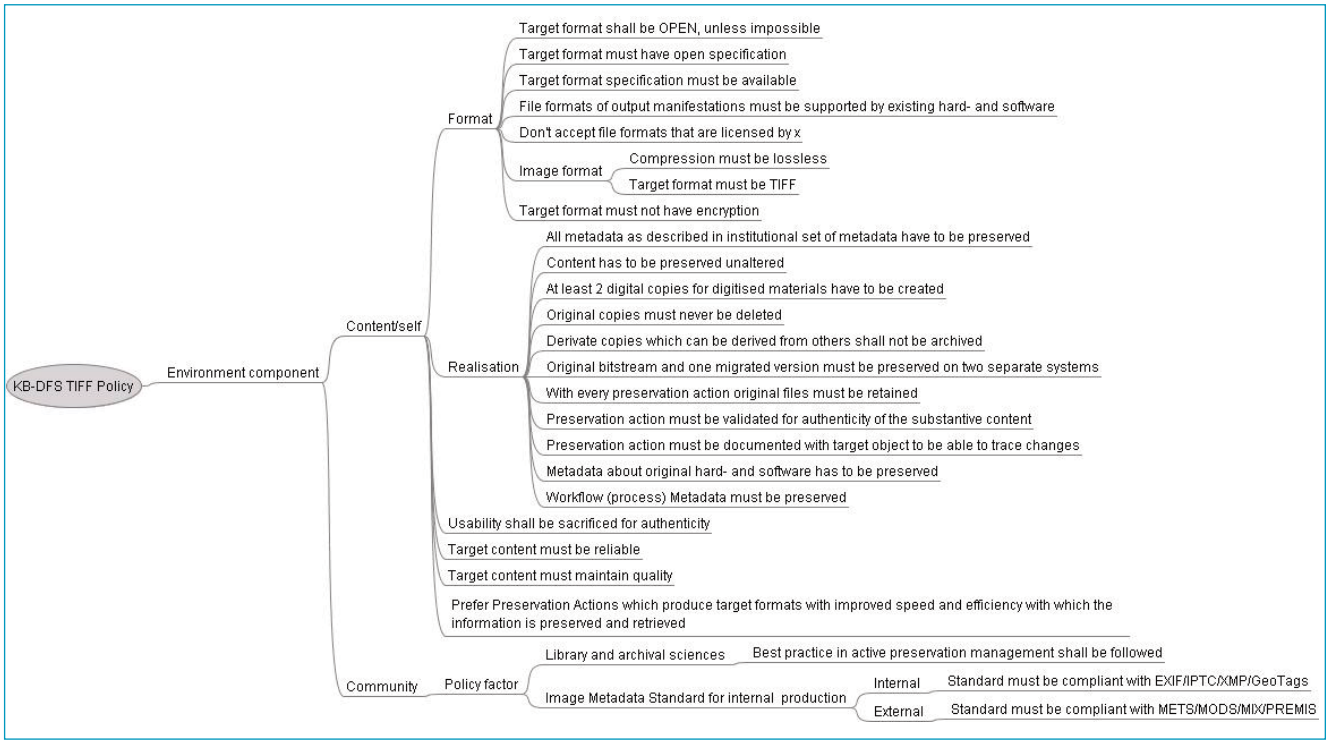


Figure 5 Mind map of the requirements objective tree which was uploaded to Plato for the Image plan

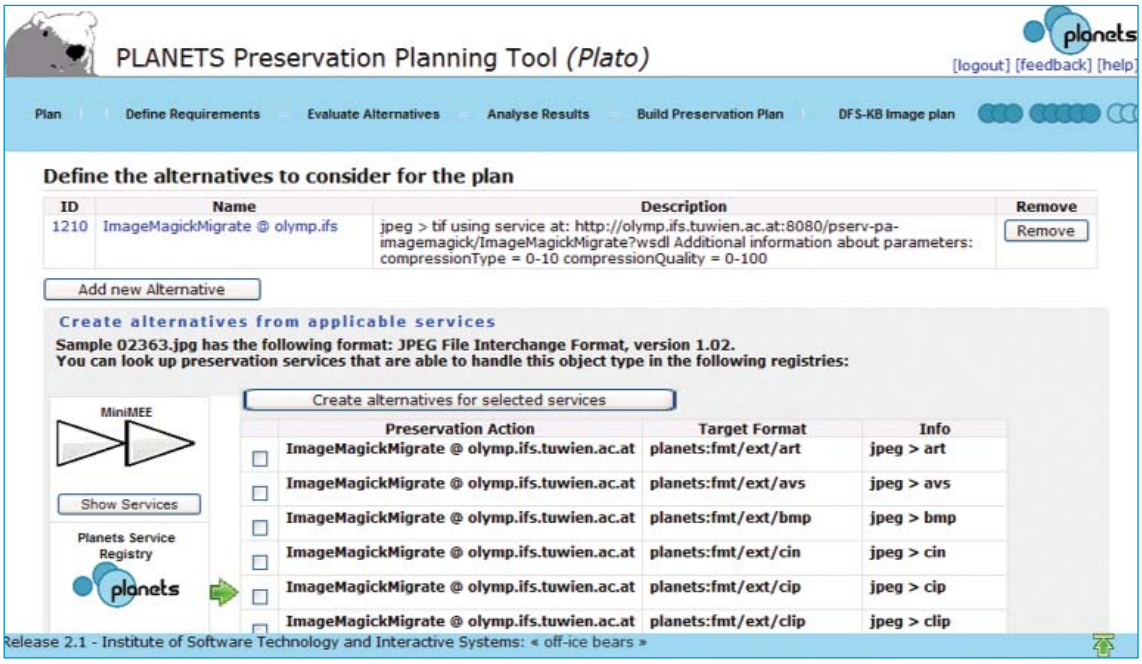
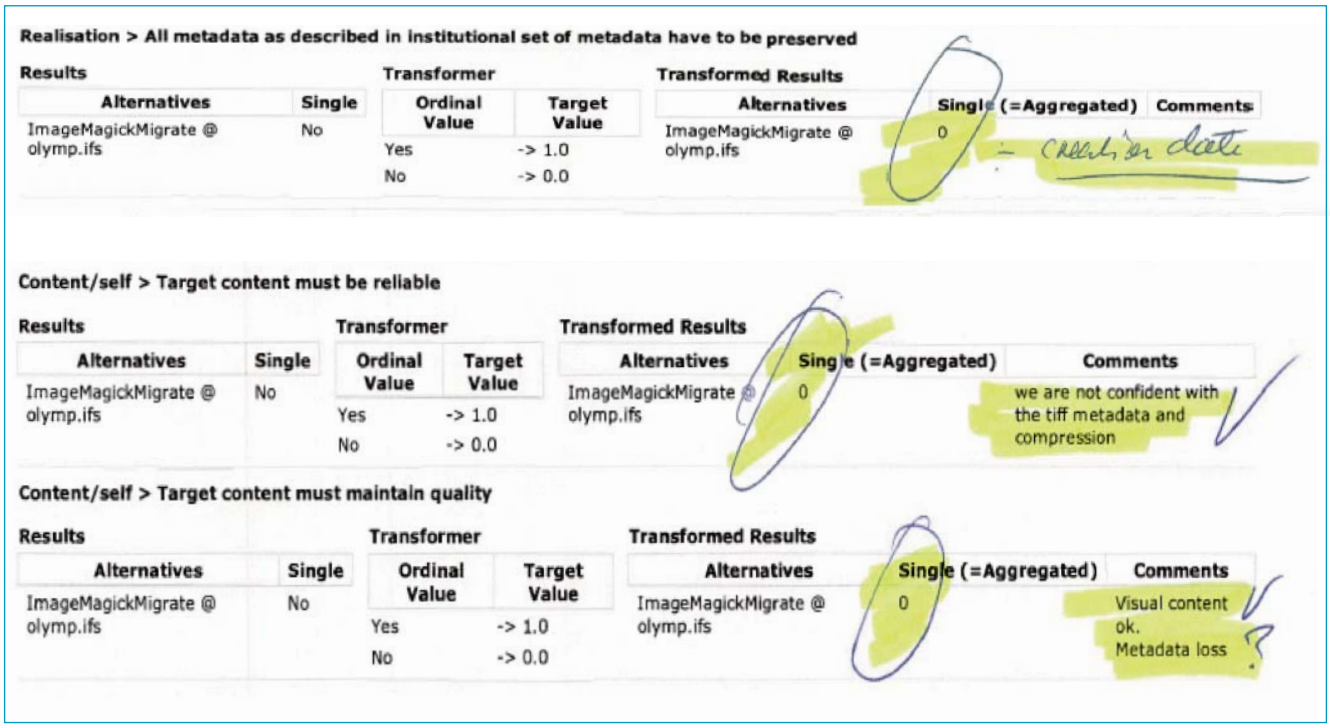


Figure 6 Screen shot of Plato showing selection of alternative migration/emulation tools for the Folklore Archives' Image plan.

During the Evaluate alternatives phase, ImageMagick[12] was the only suitable tool offered by Plato for .jpg to .tiff migration. Plato stated explicitly that the migration from .jpg to .tiff evaluated did not meet their requirements. Two of the central requirements specified for images were that metadata should not be lost in the migration process and that files should remain uncompressed. The Plato evaluation report highlighted that these particular requirements were not fulfilled by the migration tool tested (see figure 7) and recommended that the team go back and revise some of the requirements.

[12] http://www.imagemagick.org/script/index.php



Realisation > All metadata as described in institutional set of metadata have to be preserved

Results

Alternatives	Single
ImageMagickMigrate @ olymp.ifs	No

Transformer

Ordinal Value	Target Value
Yes	-> 1.0
No	-> 0.0

Transformed Results

Alternatives	Single (=Aggregated)	Comments
ImageMagickMigrate @ olymp.ifs	0	= check for date

Content/self > Target content must be reliable

Results

Alternatives	Single
ImageMagickMigrate @ olymp.ifs	No

Transformer

Ordinal Value	Target Value
Yes	-> 1.0
No	-> 0.0

Transformed Results

Alternatives	Single (=Aggregated)	Comments
ImageMagickMigrate @ olymp.ifs	0	we are not confident with the tiff metadata and compression

Content/self > Target content must maintain quality

Results

Alternatives	Single
ImageMagickMigrate @ olymp.ifs	No

Transformer

Ordinal Value	Target Value
Yes	-> 1.0
No	-> 0.0

Transformed Results

Alternatives	Single (=Aggregated)	Comments
ImageMagickMigrate @ olymp.ifs	0	Visual content ok. Metadata loss ?

Figure 7 Parts of the evaluation report for Image plan produced by Plato, printed and showing handwritten comments by the team.

The team members reviewed the Plato plan and their own observations made during the planning process. They identified that the issue over compression of the .tiff file was a result of their not understanding how to use the migration parameters in Plato migration services and this was easily resolved.

The metadata lost in the migration process was identified as the original creation/modification date of the .jpg file, because the new .tiff had its own, new creation date embedded. As the original creation date of the image file was important and needed to be saved somewhere, the project team have proposed either to move the creation date to another embedded metadata field within the .tiff file or extract it to a migration process metadata container.

The recommendation by Plato was therefore accepted by the team.

The Outcome

The project team found it very useful that Plato provided a structured work process and took them in a logical order through the first steps of a long planning process. It helped them to gain an overview of the material in the collection and revealed a number of issues regarding the material that needed their attention. Plato also guided the team through the decision making process and helped to determine when and in what ways the preservation policies should be updated.

The team found Plato particularly useful for specifying their preservation policy requirements. The Plato preservation planning process was well structured and presented a logical progression of questions to be considered. The template provided in Plato was complex enough to handle the requirements for the audio and image collections and the team found it useful to have a large set of options to consider, and discard if not appropriate. They were able to define a policy plan within just one to two days.

The use of Plato highlighted specific issues in preserving each digital collection that the Folklore Archives have needed to address. During the Define requirements phase of the audio planning, Plato helped to identify an issue with the sound quality of previously digitised audio materials. Emmerik Warburg, sound engineer in the project team explains,

> "Plato forced us to focus in very specific ways on the internal metadata and description of the audio quality. The sound files in Lydarkiv (see figure 1), digitised between 1995 and 2008, had been created by the sound engineer using an analogue dynamic compressor to optimise the input dynamics to the analogue/digital converter. It became clear during our work with Plato that this had slightly altered the sound quality which could only be changed by re-digitising the tapes."

The project team now needs to examine the consequences of this finding in relation to both the digitisation of the collection of reel-to-reel tapes and the previous migration from Lydarkiv to Lydarkiv_2.

Plato also helped to identify an important issue for the team regarding image file migration, namely the loss of original file creation date.
Tue Larsen, the IT analyst on the team explains,

> "The most important thing we lost in the migration process was the original creation/modification date from the .jpg file. It was a very important discovery for us that we could lose this information – thanks to Plato for showing us!"

Plato produced an evidence base that underlies the preservation planning decisions taken by the team, and provided detailed documentation of the preservation plans and decision-making process.

«The Plato tool can really support digital preservation experts in preservation planning. The Plato report is an excellent tool to document the whole preservation planning process, and we now use Plato as the main site for documentation.»

Tue Larsen, IT Analyst, The Royal Library, Denmark

«Using Plato gave us indispensable assistance in achieving our aim of characterising and providing a profile of the digital collection in order to produce a preservation plan for it. We could not have carried out this work without the use of Plato, and we find that the concept is excellent. We regard Plato to be the obvious preservation planning tool in the years to come.»

Birgit N. Henriksen,
Head of Department, Digital Preservation, The Royal Library, Denmark

«Plato, and our cooperation with the digital preservation unit at the Royal Library, proved indispensable for developing a long term preservation plan in connection with our ongoing production of digital material. We would not have had the necessary resources and knowledge to carry out this work by ourselves. As a consequence of this work we will now be able to adjust our work flow and digitisation. For us, the use of Plato has been a success.»

Else Marie Kofod, Head of the Danish Folklore Archives, The Royal Library, Denmark

Conclusions

The Danish Folklore Archives hold a collection of international value, some of which is held on fragile media. To preserve it, and enable greater access, they are engaged in a programme of digitisation. However, the digital collection created by this programme also needs to be preserved. The Folklore Archives' team, together with digital preservation colleagues in the Royal Library, have successfully used Plato to help produce their preservation plans for the digital collections. They have found its contribution indispensable, and intend to continue to use Plato in the future.

References and Further Reading

Lene Halskov Hansen, Astrid Nora Ressem og Ingrid Åkesson (eds) 2009:
Tradisjonell sang som levende process – Nordiske studier i stabilitet og forandring, gjentagelse og variasjon. Norwegian Folk and Popular Song Archives Publication 2. Texts published by the Centre for Swedish Folk Music and Jazz Research 26.
Oslo: Novus Press.

Poul Rovsing Olsen "Music in Bahrain. Traditional Music of the Arabian Gulf" (2002), published by the Jutland Archaeological Society, Moesgaard

"Digital Preservation – the Planets Way"; summaries for technical/developer staff, Raymond van Diessen (IBM), Laura Molloy and Andrew McHugh (both HATII, University of Glasgow), 2010
http://www.planets-project.eu/training-materials/IntroductiontoDigitalPreservation-TechnicalSummary-Final.pdf

"Systematic planning for digital preservation: Evaluating potential strategies and building preservation plans." Becker, Kulovits, Guttenbrunner, Strodl, Rauber, Hofman. International Journal on Digital Libraries (IJDL), Dec 2009.
http://www.springerlink.com/content/801685k478136425/

An overview of Planets is given in the Planets brochure at
http://www.planets-project.eu/docs/comms/PLANETS_BROCHURE.pdf

A more in depth look at Planets tools and services may be found at
http://www.planetsproject.eu/docs/comms/PLANETS_PRODUCT_SPECIFICATION.pdf

Plato and its manual can be found here: **http://www.planets-project.eu/software**

Acknowledgements

The authors would like to thank all those who have contributed to this case study.

Front cover images
Danish Folklore Archives building in Copenhagen.
Credit Photographic Studio, Royal Library.

Percy Grainger assisting the great Danish folklore collector Evald Tang Kristensen.
Credit H.P Hansen; Danish Folklore Archives.

Planets (Preservation and Long-term Access through NETworked Services) is a four-year, €15 million project, co-funded by the European Commission under the Information Society Technologies (IST) priority of the 6th framework Programme (IST-033789).

The project has developed a suite of tools and services to support preservation of digital content for the long-term. Planets tools make it possible to define digital preservation goals and policies; understand the characteristics of a collection; build, evaluate and execute preservation plans, convert objects into up-to-date and accessible formats and run software on legacy operating systems. It offers an automated solution to support informed decision-making and justify actions taken.

Planets is coordinated by the British Library and has been delivered by a Consortium of 16 national archives, libraries, research institutions and leading IT companies.

Further Information

For more information about Planets visit: **http://www.planets-project.eu**

You can email your questions to us at:
info@planets-project.eu

Open Planets Foundation

The Open Planets Foundation (OPF) builds on the investment made in the Planets project. It will sustain the results of this investment and further develop and coordinate development of the capabilities that its members require. It will provide services, knowledge, methods and tools to its members and the broader community.

For further information visit: **http://www.openplanetsfoundation.org**