

The Planets Approach to Migration Tools

Eld Zierau, Royal Library of Demark, Copenhagen; Caroline van Wijk, National Library of the Netherlands, The Hague

Abstract

This paper discusses the Planets approach to migration tool development. The approach consists of enhancing existing migration tools rather than developing tools from scratch. This pragmatic approach is based on the Planets view of the current situation for migration tools and two claims. The first claim is that the market will cover the required tools for commonly used formats. The second claim is that in the long term less tools will be required due to growing use of archiving standard formats. The Planets view on the current situation, the scope of tool development and the claims stated are, however, open for discussion and re-evaluation.

Introduction

The Planets project (Preservation and Long-term Access through NETworked Services [1]) is funded by the Information Society Technologies (IST) R&D Programme of the European Union. The Planets consortium consists of national archives, national libraries, research institutes and technology companies with experience in the field of digital preservation. The project aims to provide a framework that will enable organisations to ensure permanent access to their digital collections, through definition, evaluation and execution of preservation actions.

In the Planets project, a preservation action is defined as ‘a non-destructive action that creates new data from existing data in the archive, with the intent of preserving or increasing access to information stored in the archive’. Planets research into preservation action options ranges from migration and emulation strategies to other emerging technologies. Migration is defined as modification of the digital objects to ensure permanent access to these objects. This paper focuses on the Planets approach to migration tools.

Until now, the project’s experiences have led to a decision to base developments of migration strategies on enhancement of existing tools. The applicability of this approach at present and in the future will be argued for in this paper. First the current situation concerning migration tools is outlined as it serves as the context for the approach. Then two claims that form the arguments for the pragmatic approach to migration tools in the Planets project will be discussed. As quality is an important part of the Planets approach, challenges and further work concerning migration tool quality for available tools will also be discussed as part of the approach description. These sections will be followed by a discussion on the arguments that form the basis for the approach and a conclusion.

Current situation

The current situation concerning migration tools forms the background of the two claims stated later on in this paper. To define the scope of migration tool development for digital preservation purposes, the Planets working group on migration

tools for objects (tools working group) has delineated two groups of tools: required tools for migration in archival institutions and migration tools available on the market in the form of commercial and non-commercial tools. Figure 1 illustrates a general view of this perspective that indicates three areas to be investigated.

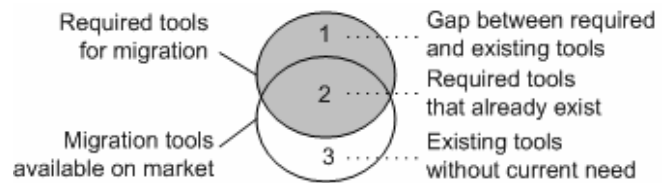


Figure 1: Scope of migration tool development

The first area shows which tools should be developed, if any. The second area indicates which tools should be investigated to see if they fit the requirements. The third area is not of interest to this paper. Therefore, the focus will be on the grey area, which represents area one and two.

At the start of the Planets project, the only source to identify this grey area was a survey among the institutions involved in Planets. The survey includes a question of whether the partners had special requirements concerning migration tools, if any. There were few responses, which is probably due to the fact that not many institutions have executed or have needed to execute preservation actions at the moment.

At present, a preservation action tool gap analysis is the focus of ongoing work in the Planets project. This gap analysis aims to clarify which specific transformations are considered necessary and therefore important by the target audience of the Planets project end products. Hence, discovery of tools in area one is covered by the gap analysis.

Peer institutions of the Planets partners in the UK, Denmark and the Netherlands have participated in surveys on archival file formats as part of the gap analysis process. These surveys will be held in more countries during the project. Preliminary trends found in the gap analysis are displayed below.

Table 1: Archival formats used by types of institution

Institution	Archives	Libraries	Museums	Other
1. Format	TIFF	JPEG	JPEG /TIFF	TIFF
2. Format	JPEG	TIFF	PDF	JPEG /XML
3. Format	PDF /XML	PDF	HTML /MP3	PDF

The table is based on the number of institutions using the format, not how much data they have stored in the format.

The results of an ongoing gap analysis will provide a possible basis for defining required transformations in the (near) future and will initiate the development of new migration tools, if needed. The two areas of interest are described in more detail in the subsequent paragraphs.

Area one: Gap between required and existing tools

Area one represents the migration tools that are required for executing digital preservation actions, but are not available on the market. Results from the gap analysis will indicate which tools fit into this area. At present, the tools working group assesses that tools in this area deal with file formats that are very complex and specialized, e.g. formats with external dependencies and formats requiring many to many transformations [2]. Preservation action tools for digital material in such formats will be developed for individual institutions as each institution will have very specialized requirements for these tools.

Examples of these ‘tool gaps’ are preservation action tools for databases, scientific data and websites. These types of digital material are or have been the subject of dedicated research projects. Long-term usability of scientific data sets such as social sciences data, geospatial data and research data from archaeology projects, has been the subject of projects by respectively Data Archiving Networked Service (DANS) [3], Archaeology Data Service (ADS) [4] and the National Geospatial Digital Archive (NGDA) initiative [5] among other projects. The Swiss National Archives are currently working on a migration strategy for databases [6] in the context of the Planets project.

For now, the tools working group concludes that due to budget and time considerations it has not been viable to investigate complex and specialised formats in the Planets project other than the tools for databases. However, if the Planets gap analysis indicates apparent gaps, the tools working group will aim to develop these required migration tools.

Area two: Required tools that already exist

Area two represents the migration tools that are required for executing preservation actions and that are available on the market.

Based on the knowledge of each project partner’s major needs, the tools working group made a rough estimate of what tools generally would be required and scanned the market for available tools. It appears that up to now all partners in the tools working group have been able to execute most preservation actions with existing tools. These actions have ranged from normalisation of text-based documents to an archiving standard such as PDF/A [7] to normalisation of images to preferred image format standards such as TIFF [8].

Looking to other cultural heritage institutions outside the Planets project, it appears that migration tool gaps have been filled by development of non-commercial tools. The most common case is tools that normalise non-standard formats to a preferred standard format. For instance, the National Archives of Australia (NAA) and the National Library of Medicine (NLM) have defined specific requirements for their migration actions. These requirements have led to the development of tools such as Xena [9] and MyMorph [10]. Xena normalises deposited digital material to preferred open formats for NAA. NLM has developed MyMorph, which can convert a range of formats to PDF in a web service architecture.

The challenge for the Planets project will be to include these required existing tools in the Planets system.

Encountered problems with currently available tools have mostly been in the area of the quality of the conversion. The paragraph *Planets approach to migration tools* will elaborate on the challenges encountered. These problems led the tools working group to conclude that important issues to look at for tools in area two are quality of the existing tools and how to perform quality assurance automatically along with basic challenges on how to incorporate existing tools in the Planets system.

Summary

The current situation of migration tools has been outlined in three areas of tool development for the Planets project. In area one, the gap analysis will identify apparent gaps in the available tool range. Furthermore, the tools working group assesses that development will be needed for complex formats e.g. databases.

In area two, quality assurance of the existing tools, how to do this quality assurance automatically and how to incorporate existing tools in the Planets system have been defined as important issues for tool development. The third area is out of scope for tool development by the Planets project.

Claims

The current approach to migration tools is based on two claims. The first claim is that ‘The market will cover the tools for widely used formats’. This claim is the basis for the tools working group’s choice to concentrate on looking at quality of the relevant existing tools. The second claim states that ‘Only very few migration tools will be in demand on the long-term’. This claim is the basis for trusting that the applicability of the current approach will also hold in the future.

The following sections describe the claims in more detail and contain examples that support the claims. It must be noted that the statistics and research in this area are not comprehensive enough to make it more than a claim of a tendency.

Claim 1: Tools for popular formats will be provided

The claim states that the market will provide tools for most of the commonly used formats. Arguments and examples that support this claim are described in the following paragraphs.

Commercial interest

Widely used file formats offer an interesting selling area for commercial tool development. Next to this, commercial companies like Microsoft take part in Planets to explore what the requirements are for their Office suites concerning digital preservation. They will provide Planets with wrapped conversion tools for older Microsoft Office formats to Open XML formats and open standard formats like ODF. Taking part in Planets will offer Microsoft the chance to tune in to a growing market.

Tools used in practice

Institutions dealing with digital preservation all have their specific requirements for the preservation tools that they use. Restrictions like budget, resources and time also differ. In the following examples, institutions have successfully applied commercial software to preservation actions in very differing contexts.

Image migration at the State and University Library in Denmark (SB) forms an example of the market providing required tools. The library had a collection of a 5000 TIFF Images that were stored in the old-style JPEG [11] system that became hard to access and with no known migration tools. The problem was the way the images were embedded in TIFF, not because of the old-style JPEG. Rather than being an embedded file-stream, old-style JPEG has several essential chunks of information in separate TIFF fields that reference each other. The tools working group wrapped the open source program suite ImageMagick and software LibTiff to a web service interface to fit the Planets architecture. The images could then be extracted.

Another example of a filled tool gap was also found at SB (actually the only example found where the market did not cover the gap initially). The library had tapes of Danish Radio programs delivered on DAT-tapes in the period from 1998-2001 with about 60 TB of uncompressed data. The tapes were written with equipment from the American company Eventide. In 2001 the SB discovered that some of the tapes were no longer readable. In an attempt to rescue the data of healthy tapes, they copied the tapes and tried to read from the new tape. This attempt failed, since the copy was not well formed for the access tool. Consequently, the new strategy was to migrate the data. There were no commercial tools that could assist in migration of the data, but luckily the Eventide Company was able to make a tool used for migration into WAV format, which is one of the de facto standards at SB.

The Austrian National Library (ONB) prefers to archive text-based documents in the PDF format. Content suppliers are asked to deposit their material in the PDF format. As a result of this restriction only 10% of the documents present in the ONB digital archive are in MS Office format. These documents are converted to PDF/A by ONB. The small amount of documents to be converted and the budget limitations have been important requirements that influenced the choice for an Adobe migration tool.

At the National Library of the Netherlands (KB-NL), several tools are being tested to normalise MS Office formats to PDF/A, next to archiving the original MS Office document. In KB-NL's case, retaining the layout is one of the most important requirements for a normalisation tool. Currently, two commercial tools are serious contenders.

Cost and benefit

Institutions first gathered information about existing tools that may fulfil required preservation actions in most of the examples above. If an institution were to choose between using an existing migration tool or develop one itself, the cost will have to be weighed against the benefits gained. Developing a new tool from scratch can only be justified if it introduces new relevant functionality or features compared to existing tools.

In the example of the DAT-tapes at the SB, there was a strong benefit of a quick solution, since the tapes were deteriorating. On the other hand the Eventide Company may well have been the only organisation with skills and knowledge of the format to develop the migration tool.

The cost-benefit approach is also applied in the MIXED project [12] by DANS, the Dutch institution that works on permanent usability of scientific data. The MIXED project mostly uses existing software to perform specific file format conversions

and extraction of information from binary formats to an interchange format (XML). Development resources within the project are focused on building a framework in which these existing tools can be deployed.

Claim 2: Fewer tools in demand in the long term

The second claim discussed in this paper states that the range of migration tools used for digital preservation actions will be small in the long run. The basis of this claim can be found in a forecast: a tendency exists among libraries and archives to move towards a few standard archiving formats. As each institution has specific requirements for the preservation of their digital collections, it is deemed impossible to define one perfect file format for digital preservation. However, the range of formats used for archiving will tend to be smaller than it is now, and consequently the range of tools required for preservation will tend to be smaller. The second claim provides the basis for confidence in the future applicability of the current approach to migration tools. This section will discuss the examples and arguments that support this claim.

Many libraries, archives and documentation centres – among which the institutions working on migration in the Planets project - issue preferred format guidelines for suppliers on file formats they consider suitable for archiving purposes. The guidelines cover a range of types of objects such as digital still images, audio and video based information and text based documents. The way the format guidelines are implemented differs according to the nature and means of the institution.

Imposing standards

There is a big difference between archives and libraries, since archives in most countries can impose standard formats on the external producers of digital material that will be deposited, since these producers are institutions under government regulations. The institutions can either impose requirements on deposited objects, or migrate digital objects to the preferred formats themselves.

In all Nordic countries the national archives have made requirements for the formats in which digital material is deposited. These do however vary in strictness, where the Danish National Archives has some of the strictest requirements. One reason for the variation is the time limit for deposit of material which is varying from 5 years (e.g. in Denmark) to 25-30 years (e.g. in Norway).

The National Archive of the Netherlands has developed clear guidelines for archiving digital material created by governmental organisations. Digital private archives donated by public figures will be converted to preferred standard file formats for archiving if necessary. The main restriction for depositing digital materials is that the material should be deposited in de jure open standard formats like PDF/A.

Archiving guidelines

For most libraries it is not possible to impose standards. There are, however, several examples of attempts to make standards for delivered material, in form of guidelines to producers, or negotiate formats with digital object producers.

The KB-NL provides a PDF guideline for publishers [13]. The guideline is similar to the PDF/A standard, but is also readable for non-technical persons and summarizes and explains the KB-NL's preferences for specific configuration of the PDF.

In Belgium, the eDAVID expert centre [14] offers guidelines for digital preservation, including a list of preferred file formats suitable for preservation. Their target audiences are governmental institutions, but all guidelines are also very useful for other digital document producers such as commercial companies.

Another example of file format guidelines for preservation can be found at the Florida Digital Archive. Their website lists documentation that specifies the 'Recognized File Formats', 'Guidelines for Creating Archival Quality PDF Files' and 'Recommended Data File Formats'. These documents [15] specify what the Archive will attempt to preserve, and how the archival object producer can assist in forming the archival object in a way that makes it more probable to be preserved in the long term.

At the Royal Library of Denmark (KB-DK), the delivery of material from artists and photographers in digital form is still fairly new; however there are already examples where photos are delivered together with the digital version, and where the digital version is more original than the photos or physical slides themselves. Therefore, work has started on preparing guidelines for the producers to how they should deliver the material in order to ensure the best preservation of their work. The guidelines are also of interest to the artists themselves, if they want their work to survive; it will not just be the interest of easing the work at the digital preservation department.

An example of negotiation with producers about formats in which material is deposited can be found at the KB-DK. Deposit Geospatial data for Geographical Information Systems consisting of geographical data where the images corrected according to information of geographical coordinates is delivered to KB-DK by COWI. This Danish company generates aerial and satellite imagery for the Danish government. When the delivery was initiated, COWI suggested delivery in Earth Resource Mapping's Enhanced Compression Wavelet format (ECW [16]) and JPEG, which was not acceptable for KB-DK, because of the quality of these formats. Also the JPEG2000 [17] format was offered, but since the preferred preservation format at KB-DK is TIFF it has been arranged that the data be delivered in GEOTIFF which is TIFF 6.0 with header meeting requirements for GEOTIFF standard [18].

Normalisation

Examples of normalisation of proprietary, closed formats to formats more suitable for digital preservation are discussed in the section of the first claim where the activities of the KB-DK, ONB and KB-NL are mentioned. More examples of normalisation from less suitable file formats for preservation to 'standard' preservation formats can be found in the *Current situation, Area two: Required tools that already exist* section that describes normalisation at NAA and NLM. DANS normalises scientific data to an intermediate format to decouple the data from (proprietary) applications needed to use the data.

Restrictions to tool candidates

There is a tendency that only formats, that are well known and have been on the market for some time, will be subject for use as a preservation format. If a new format arises it will probably be fair to assume that both commercial and non-commercial products will come along before that format is accepted as a well-established archiving format.

In most assessments of preferred file formats for archiving purposes there will be parameters of how well tools support the format, how big a user community is, and if the format is (or will be) standardized. Only very few archival institutions can afford to be the sole user and steward of a special format. It would be a very expensive burden and they would at the same time refrain themselves from sharing experiences and ideas with other users. Examples of such considerations are found in the next section.

Moving towards standards

Several user communities have been able to reach common cross-country standards for archival formats. Examples are the PDF/A, the TIFF 6.0 and the WARC [19] format. These file formats have been defined by a user community which includes institutions that deal with digital preservation including libraries and archives.

Movement towards using a few formats for digital preservation is also due to requirements such as cost reductions next to technical requirements. KB-NL has used TIFF as a preferred format for archiving digital still master images. Due to a forecast of costs for storage of TIFFs, a report [20] was made in order to analyse whether using another format for archiving digital master images could reduce the costs. The parameters for the analysis were; required storage capacity, image quality, long-term sustainability and functionality such as adoption rate on the consumer market and browser support.

KB-DK have also started the same considerations, and started an analysis. However, the requirements are not completely the same. Very important factors here are also how widely the format is used, both with respect to tool support and sharing of knowledge with other institutions. Another difference is that KB-DK has bit-preservation on the material in question, which means that robustness of the format with respect to bit errors has little weight, whereas a solution to solving special geographic data in TIFF as described in the next section has much weight.

Although institutions have specific requirements, most guidelines on file formats suitable for digital preservation or lists of mandatory deposit formats mention the same group of well known and often de jure or de facto (open) standard formats. Therefore, issuing guidelines and imposing specific archiving formats on suppliers of digital material will enforce the tendency to move to fewer standard archiving formats.

Summary

In claim one it is argued that institutions have been able to apply existing tools, commercial and non-commercial, for required digital preservation actions. The examples given are certainly not exhaustive. However, literature on concrete examples of preservation actions is scarce. One reason could be that not many cases of format obsolescence in archived digital collections have occurred yet and, therefore, preservation actions have not been required much. If institutions have had to save the usability of digital collections, in most cases the cost and benefit analysis into either buying an existing tool or developing a tool from scratch has favoured the former.

In the second claim it is argued that there is a tendency to move to fewer archival formats and which are (mostly) supported by the market. There will be different formats according to varying demands for different institutions. Furthermore, there will be

different means to implement limitation to preferred formats for archival information. The means in form of imposing or guiding to preferred formats or normalization will be defined according to varying missions and opportunities for the different institutions.

Planets approach to migration tools

Next to having the required tools available, one of the biggest challenges for the tools working group is to measure the quality of a migration action using the tools. This challenge includes defining what is meant by 'acceptable quality'. Furthermore, it includes automation of quality assurance in order to handle migration of large collections. The following paragraphs describe issues that will be explored further during the project.

Acceptable quality

Defining acceptable quality has many aspects and is dependent on the collection policy of an institution. Firstly, it will in many cases have to be based on a cost-benefit analysis. Migrations that are not subject to some form of information loss are rare. One example is the migration from MS Office format to PDF/A at ONB; here some of the quality requests were compromised due to budget considerations in choosing the conversion tool.

Secondly, questions on whether the loss is acceptable will rely on the type of material. An example is the land charges register in Denmark, which has been digitised by the Agency for the Courts of Denmark primarily for presentation purposes in the format PDF. The originals were supposed to be transferred to the National Archives, but are now awaiting a decision of destruction. This decision will depend on quality assurance of the digital images, when they are transformed into preservation format TIFF. For example, it must be verified that handwritten notes are still readable. Another example is digitisation of deteriorating negatives at KB-DK where the quality must be high enough to preserve the original significant characteristics, which for instance means that it must zoom in on particular parts of the image [21].

Automated quality control in Planets

The tools working group will concentrate on how computer-assisted QA can be of assistance in migration of large collections. In the Planets project, several working groups are working on Comparative Quality Assurance (QA) that compares the source and destination object of a transformation based on various metrics rather than on strict equality. There are three different approaches to different measures of quality in Planets which each has their relevance to the digital preservation process.

The Planets characterization working groups have developed the XC*L technology [22] which supports Comparative QA, where aspects from a source format can be compared to aspects of the target format. It includes a language for extraction of characteristics (XCEL) and another language for description of characteristics (XC DL). The XC*L languages are still under development and only operate on limited formats and there still is a challenge in defining the descriptions in order to achieve measurements that reflect the definition of acceptable quality.

To assist in development of migration tools a QA Framework has been developed in the Planets migration tools work group. The QA Framework is intended for developers working on transformation tools who want to track the progress. Here numeric

mathematically defined measures can be defined to be repeatedly calculated on different versions of a migration tool.

In the Planets project, a benchmark Test-bed application has been developed which permits experiments to be performed by running preservation services on test data. This Test-bed application can be used for comparing the quality of several migration tools on the same test data, running the tools under the same circumstances.

Summary

Quantifying quality is a very difficult and broad spectrum exercise that can only be computer assisted to some extent. In Planets, the research into 'quality' and migration tools includes computer assisted quality assurance in the form of comparative QA for source and target formats, a quality track tool for migration tool developers and defining quality of tools by experimenting in a benchmark Test-bed application.

Discussion

The Planets approach to migration tools mainly focuses on discovery of, quality assurance of and incorporation into the Planets architecture of existing tools. It is a pragmatic approach based on the tools working group's view on the current situation concerning migration tools and the claims stated in this paper. The definition of the scope of tool development and the claims are supported in this paper by examples. However, the tools working group is aware of possible arguments against the view on the current situation for migration tools expressed in this paper, and therefore also against the claims and scope of tool development. In the next paragraphs, arguments that can be made against the view on the current situation and the chosen approach are shortly discussed.

Definition of tool development areas

The definition of the tool areas described in this paper is open for discussion, since it can be questioned whether a tool actually is the *required* tool in case it does not meet a predefined quality standard or in case it does not meet specific user requirements defined by an institution.

A case of special user requirements for migrations is the GEOTIFF data at KB-DK. Since GEOTIFF is TIFF 6.0 with geographical header data, the tools used for this kind of TIFF must be able to interpret these data following the GEOTIFF standard. Consequently, migration of the TIFFs is *not* only a matter of migration of the image. The special user requirement in this case is that the geographical header data is migrated in a way that enables the data to be interpreted in the same way as before the migration. At KB-DK there have been several migration attempts with unacceptable results such as placement of a Danish city on the American continent! As discussed before, it can be arranged for KB-DK to have the data delivered in a JPEG2000 format with the valid geographical data, but it is still an issue how to migrate the old deliveries that are already in the care of the KB-DK.

Gap analysis

Whether or not the claim that the market will provide tools for most commonly used file formats will hold, partly depends on the (intermediate) results of the gap analysis. And even with these

results, it remains to be seen whether the analysis based on sampling comes close to the actual worldwide situation.

One exception has already been mentioned and consists of the 'known' gaps in tools that handle specific, very complex formats including formats with external dependencies and formats requiring many to many transformations. Tools for such formats are and will most likely not be catered for by the commercial market due to their niche existence. Unless an institution that uses such a format will decide to develop a tool, it remains a gap.

Quality control

There are a lot of remaining challenges in defining requirements for and measuring of quality of the migration tools. Decisions on whether something is good enough is as a starting point a subjective decision by persons who are to evaluate material's heritage value and which aspects of the material are important to preserve.

Defining what is important will have to rely on procedures for evaluation and will be part of the individual institution's preservation planning process. This includes translating subjective statements like 'readability of all handwriting notes must be preserved' into any QA computer system.

Conclusion

To conclude, this paper is intended to explain the arguments for the Planets approach to migration tools but also to start a discussion on the merits of this approach.

The Planets project's approach to migration tools is focusing on reusing and enhancing existing migration tools rather than trying to redevelop migration tools from scratch. By taking on this approach, the tools working group has chosen to take a pragmatic stance to how it sees the current situation for migration tools and what it expects to happen in the (near) future.

However, as pointed out in this paper, this current situation, the scope of tool development and ideas on quality control are issues open to differing opinions. The Planets migration tool work group hopes that this paper will invite others to express varying opinions on the claims in this paper to the Planets project so that its horizon can be broadened and the approach to migration tools can be adjusted where necessary.

References

- [1] Information on Planets available at <http://www.planets-project.eu/>
- [2] Planets report on Report on tool and service approach http://www.planets-project.eu/docs/reports/Planets_PA4-D1_ReportOnToolAndServiceApproach-Final_Public.pdf
- [3] Information on DANS available at <http://www.dans.knaw.nl/en/>
- [4] T. Austin, J. Mitcham, Final Report Preservation and Management Strategies for Exceptionally Large Data Formats: 'Big Data', from http://ads.ahds.ac.uk/project/bigdata/final_report/bigdata_final_report_1.3.pdf - retrieved 1. April 2008
- [5] Information on NGDA available at <http://www.ngda.org/>
- [6] K. W. Ohnesorge, H. Thomas, Tutorial at IS&T Archiving 2008: A File Format for Archival of Relational Databases
- [7] Information on PDF/A available at <http://www.iso.org/iso/home.htm>

- [8] Information on TIFF available at <http://partners.adobe.com/public/developer/tiff/index.html>
- [9] Information on Xena available at <http://xena.sourceforge.net/>
- [10] F. L. Walker, G. R. Thoma, Image Preservation Through PDF/A, Proceedings IS&T 2005
- [11] Information on JPEG available at <http://www.jpeg.org/jpeg/>
Old-style JPEG is the (now deprecated) practice of use of JPEG compression for image data in a TIFF file.
- [12] D. Roorda, Whitepaper Migration to Intermediate XML for Electronic Data (MIXED): A strategy in digital preservation – A DANS software project, retrieved on 1 April 2008 from http://mixed.dans.knaw.nl/files/file/white_paper2.pdf
- [13] J. Rog, PDF Guidelines - Recommendations for the creation of PDF files for long-term preservation and access, from http://www.kb.nl/hrd/dd/dd_links_en_publicaties/PDF_Guidelines.pdf retrieved on 1st of April 2008
- [14] Information on eDavid available at <http://www.expertisecentrumdavid.be/eng/index.php>
- [15] Florida Digital Archive format information: <http://www.fcla.edu/digitalArchive/formatInfo.htm>
- [16] Information on ECW available at <http://www.ermapper.com/>
- [17] Information on JPEG2000 available at <http://www.jpeg.org/jpeg2000/>
- [18] Information on GEOTIFF available at <http://www.remotesensing.org/geotiff/spec/contents.html>
- [19] Information on WARC available at <http://archive-access.sourceforge.net/warc/>
- [20] R. Gillesse, J. Rog., A. Verheusen, Alternative File Formats for Storing Master Images of Digitalization Projects version 2.0, from http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/Alternative File Formats for Storing Masters 2 1.pdf - retrieved 1. April 2008
- [21] U. B. Kejsler, Pre print of Preservation Copying of Endangered Negative Collections, Proceedings IS&T Archiving 2008
- [22] V. Heydegger, J. Neumann, J. Schnasse, M. Thaller, PC/2 D1/D2 Architecture specification and first draft of the XCEL/XCDL languages, retrieved on 1 April 2008 from <http://lehre.hki.uni-koeln.de/planets/documents/deliverables/PlanetsPC2D1D2-end.pdf>

Authors Biography

Eld Zierau, Digital Preservation Specialist at the Royal Library of Denmark (KB-DK) is educated as a computer scientist (M.Sc.) at Aarhus University, Denmark in 1989. She joined Royal Library in May 2007. She has a long career behind her including 13 years in an IT consulting company as Chief consultant and some years in the Medicinal and Telecommunication industries. She has worked 6 years in a EU research project concerned with formal methods. Besides from that she has worked in various industries, various IT areas and at various organisational levels facing a lot of the same challenges that we also meet in long term preservation today. At KB-DK she works in the PLANETS migration subproject and the Danish web-archiving system.

Caroline van Wijk, Digital Preservation, the National Library of the Netherlands (KB-NL) has a BA degree in Art and an MA in Political Science. She finished a Java software engineer training in 2000. She then worked at a number of web development companies for well over four years before joining the KB-NL in 2004. At the KB-NL, she worked on the pilot project Tiff-archive as the technical project staff member until December 2005. She was the project leader of the migration project until 2008 and is now the project leader of the web archiving project. She also takes part in the European project PLANETS as a digital preservation researcher and work package leader.