



***Puzzling over digital preservation – Identifying traditional and new skills needed for digital preservation***

**Thomas Bähr**  
German National Library of Science and Technology

**Michelle Lindlar**  
German National Library of Medicine

**Sven Vlaeminck**  
German National Library of Economics

**Meeting:**

**217 — Education for digital curation — Education and Training Section with Preservation and Conservation, Information Technology; co-sponsored by ICA Section for Archival Education and Training**

**Abstract:**

*Digital preservation is a task requiring library and information science as well as information technology skills. It simultaneously utilizes traditional library skills and requires knowledge from information technology that goes far beyond the traditional roles of library and archive staff. But where does one start when implementing a digital preservation program? What knowledge is needed? What tasks can be covered by existing personnel? Where can one acquire expert knowledge needed? What information resources exist? Can a scalable approach be implemented to gain necessary skills?*

*The paper is based upon a gap analysis conducted by the Leibniz Library Network for Research Information “Goportis”. It describes necessary know-how identified, ranging from digital curation skills needed to evaluate digital data carriers to specialist digital preservation knowledge of file formats needed to describe information with the goal of sustaining accessibility over long-term. It shows how central tasks of digital preservation like process description and preservation planning require expert knowledge of traditional librarian and information technology skills as well as new knowledge which is described as digital preservation skills.*

**Introduction**

Digital preservation is not a single task but encompasses all strategies and measures taken to recover and maintain access to digital information. These strategies and measures are by no

means restricted to activities undertaken by information technology experts. Necessary tasks to ensure the longevity of a digital data carrier start as soon as that carrier enters the collection and encompass all fields of daily library work. Good practise in digital preservation thus requires good practise in the overlaying task – digital curation. The paradigm shift which libraries are facing results in a need for new skills for content experts and technology experts alike. Libraries that hold a stewardship for digital cultural heritage are furthermore required to implement a trustworthy digital repository, which calls for specialized digital preservation skills.

But what kind of staff is involved in digital preservation activities? A study conducted by the U.S. based Digital Preservation Outreach and Education (DPOE) initiative in 2010 showed that out of 481 libraries which took part in the survey, 27% had paid full-time or part-time digital preservation staff available, 21% relied on external vendors or volunteers and 14% stated that no one was responsible for digital preservation at their institution. With 38% the majority of the libraries stated that various staff was assigned to digital preservation tasks as needed.<sup>1</sup>

Symbolically speaking digital preservation can be looked at as a big puzzle. Solitary skills seen by themselves make as little sense as a random group of unconnected puzzle pieces. Only when we combine all and put them into perspective, the picture becomes clear.

## **The Goportis digital preservation project**

Goportis – the Leibniz Library Network for Research Information – consists of the three German national subject libraries: the German National Library of Science and Technology (TIB), the German National Library of Medicine (ZB MED) and the German National Library of Economics (ZBW). The three Goportis partners are currently conducting a digital preservation pilot project to evaluate organisational, technological and institutional needs for a collaboratively operated digital preservation system. The pilot project follows two approaches. One part of the project is aimed at fostering the awareness of digital preservation throughout the three institutions. This is done through various activities, including presentations on digital preservation open to the entire library personnel, small group workshops and student projects which are supervised by the digital preservation experts. The other part of the project is devoted to a test implementation of a digital archive and involves digital preservation experts, content experts and information technology experts.

## **Gap analysis of skills needed**

The experience gained in both parts of the Goportis project is used for a gap analysis of know-how needed for the Goportis digital preservation program. Since the overall personnel of the libraries covers a wide range in regards to age and experience-on-the-job, no direct correlation can be made between the gaps identified and the state of vocational, college and university level education in library and information science and information technology.<sup>2</sup> However, since the experiences made at the three libraries are almost identical, a careful assumption can be made as to existing gaps in digital-preservation related know-how in

---

<sup>1</sup> (DPOE 2010).

<sup>2</sup> Some exceptions do exist, e.g. in form of supervised bachelor projects. These are however limited experiences and allow no generalized statement about digital preservation related coursework at the bachelor level.

general. This is further underlined by experience that Goportis has made through its work in national and international competence networks.

The results are presented in an approach which aims to be scalable to institutions of different size and with different needs and requirements. The findings are outlined in four basic rules: “Know what you have”, “Know how to handle your digital materials”, “Know what comes ahead” and “Know how to keep it safe”. The sections describe work that can aid a smaller institution in starting a digital preservation program with the existing staff. They give brief examples of tasks and required skills, summarize Goportis’ experience in a short gap description and list helpful resources.

The last section briefly describes the role of the digital preservation specialist and the vital role that designated digital preservation personnel – as opposed to “various staff assigned to the process as needed”, to paraphrase the DPOE study – plays in a trustworthy digital repository.

### **“Know what you have”**

The first step needed in regards to digital preservation sounds rather simple. You need to know what your holdings hold. The truth to this is in the details. The equivalent in detail of whether a publication is a serial or a monograph and whether a printed publication is folio, quarto or octavo size needs to be respectively applied to digital data.

The simplest differentiation is that between offline and online data.<sup>3</sup> While offline refers to data carriers that are kept on the shelves – like CD-ROMs or floppy disks – online refers to data accessible through the internet or through closed networks – like e-publications kept in repositories. Another common differentiation is that between born-digital and digitized material, indicating whether an object has an original analogue master or not.

When establishing a good first overview of the digital collections, a meaningful grouping of items to a collection level needs to be made. This is usually done on an intellectual collection level. In a next step the collections needs to be surveyed. Surveying for the means of preservation is a traditional library task and is applied to analogue material as well. Surveying digital data, however, requires a new definition of criteria to be captured.

More specifically, different approaches to surveying offline and online materials need to be defined. For offline collections special attention should be given to capturing details pertaining to the carrier type. More detail captured allows for a better evaluation of materials potentially at risk.

This shall be briefly explained using the example of CDs. It does not suffice to state that a carrier is a “CD”. A differentiation between a commercially pressed CD-ROM, CD-R and CD-RW needs to be made. Furthermore it is advisable to capture the dye for CD-Rs, as the longevity of CD-R media heavily depends upon the materials used in the production process, such as the organic dye<sup>4</sup>. For CD-Rs and CD-RWs the manufacturer and maximum capacity of the carrier might be captured as well.

---

<sup>3</sup> A variety of different terms exist for online and offline media, e.g. handheld, transportable, managed.

<sup>4</sup> (Brown 2003)

The survey of offline data can be done using a random sample of the respective collection. It should include a readability test and a rough overview of formats contained on the carriers for at least a relevant subset of the sample.

The survey of online data does not need to capture the same detail in regards to the carrier.<sup>5</sup> Here, a random sample of file formats and file sizes should be captured.<sup>6</sup>

In addition to the survey a process analysis should be conducted for offline and online collections alike. The process analysis should capture the entire business process including curation steps of the material, legal constrictions, technical and procedural constraints and available documentation. A process analysis of this kind is somewhat of a new concept to library and information science. It might be comparable to business process descriptions done by the library management or to documentation of technological processes, e.g. within a repository, done by the information technology department.

#### *Gaps identified:*

While “offline”, “online” and “digitized” are terms understood by virtually all, the term “born digital” often required further explanation. This implies that a common vocabulary in describing digital data needs to be fostered and actively used. Whereas this gap is an important one to bridge in order to enhance the common understanding of a library’s digital holdings and preservation needs, the gaps that the tasks survey and process analysis form are more tangible ones.

Knowledge of digital data carriers appears to be restricted to the information technology department, even though the classical survey work of analogue collections forms a central librarianship task. This gap appears to carry on into the current student generation. Projects and workshops conducted with bachelor students of library science at one of the Goportis partners showed that the students had an understanding of basic digital preservation principles but were lacking an understanding of data carriers and file formats. Obsolete data carriers like LaserDiscs or zip discs could only be identified by information technology and digital preservation specialists.

Process analysis in the form as described above is also a relatively new task for the library and information science field. Knowledge exists in personnel familiar with project management procedures.

An additional gap is the lack of information sources regarding survey work and process analysis for digital collections.

#### *Information sources identified:*

As mentioned before, little literature and fundamental information exist on process analysis and survey work.

---

<sup>5</sup> Carrier for managed online systems as well as the backup procedure existing for them will be captured as part of the process analysis.

<sup>6</sup> This should be done in accordance to the process analysis. If the process analysis does, for instance, show that the business process of the repository limits the digital objects contained therein to a certain selection of file formats, this could be checked. This can be done automatically through a file analysis.

CALIPR<sup>7</sup> is an online tool which assists in selecting random samples from a digital collection in the stacks. CALIPR was written by the California Preservation Program and is free to use. Indiana University at Bloomington conducted a Media Preservation survey which covered the University's audio, video and film collections. Despite the fact that the focus of the survey was on audiovisual material, the main survey questions asked can serve as a basis for the development of an institutional process analysis template.<sup>8</sup>

A couple of further reading sources can be given in regards to optical media like CD and DVD; however, general information regarding other data carriers is scarce, which especially poses a problem when it comes to identifying an unknown obsolete data carrier.

*Further reading:*

California Preservation Program. CALIPR. Available online at <http://sunsite.berkeley.edu/CALIPR/index.html>, last checked May 22<sup>nd</sup> 2011.

CD-Info.com. The CD Information Center. Available online at <http://www.cd-info.com>, last checked May 22<sup>nd</sup> 2011.

Casey, Mike (2009). Indiana University Bloomington. Media Preservation Survey. A Report. Available online at [http://research.iub.edu/communications/media\\_preservation/iub\\_media\\_preservation\\_survey\\_FINALwww.pdf](http://research.iub.edu/communications/media_preservation/iub_media_preservation_survey_FINALwww.pdf), last checked May 22<sup>nd</sup> 2011.

Optical Storage Technology Association. Available online at <http://www.osta.org/>, last checked May 22<sup>nd</sup> 2011.

## **“Know how to handle your digital materials”**

Knowing what you have is the logical prerequisite for knowing how to handle the digital materials in your holdings. Proper handling of digital holdings can be divided into proper physical handling and proper semantic handling.

Proper physical handling describes proper storage and proper processing of carriers like optical and magnetic data carriers. It describes tasks conducted in every-day library work with the digital carriers. Since storage conditions of most digital offline media is similar to that of print materials in regards to temperature and humidity, proper storage can be easily achieved in most cases<sup>9</sup>. Bigger problems in proper processing include the application of adhesive labels, bending or flexing the data carrier while handling it or special requirements for transport.

Proper physical handling of the data carrier requires a good understanding of how the data carrier works. This knowledge is also needed for evaluating which data carrier is suited for short-term, mid-term or long-term storage. There is no single correct answer when choosing a data carrier. The evaluation process needs to consider business and workflow aspects like capacity required and cost per carrier, as well as preservation aspects like longevity and viability (i.e. available error-detection methods of the respective drive).

---

<sup>7</sup> <http://sunsite.berkeley.edu/CALIPR/index.html>

<sup>8</sup> (Casey 2009, Appendix A, p. 113 ff)

<sup>9</sup> (See Byers 2003 p. 16 and Van Bogart 1995)

Brown (2008) suggests a scorecard approach for “benchmarking” digital data carriers. The scale in the sample scorecard shown below ranges from 1 (worst) to 3 (best). Brown states that only data carriers with a score of 12 or higher should be considered.

Media	CD-R	DVD-R	Zip Disk	3.5" Magnetic Disk	DLT	DAT
Longevity	3	3	1	1	2	1
Capacity	2	2	1	1	3	3
Viability	2	2	1	1	3	3
Obsolescence	3	2	2	3	2	2
Cost	3	2	1	1	3	3
Susceptibility	3	3	1	1	3	2
<b>Total</b>	<b>16</b>	<b>14</b>	<b>7</b>	<b>8</b>	<b>16</b>	<b>14</b>

**Figure 1: Media Selection Scorecard**

Sample scorecard of data carrier evaluation according to Brown (Brown 2008)

Brown’s approach shows that the decision for a data carrier should be based on good evaluation criteria. This holds true for initial storage as well as a target carrier for media refreshment. Media refreshment – as well as all backup procedures – is a traditional information technology skill. The backup procedure chosen should be described in the process analysis mentioned in the previous chapter.

Proper semantic handling of digital data is achieved through good capture of metadata. Bibliographic metadata should be complete and ideally include information regarding the data carrier type. Studies like the TAPE study, which surveyed audiovisual holdings in European memory institutions, have shown that it can not always be assumed that bibliographical metadata for complex objects such as audiovisual material on offline data carriers exists.<sup>10</sup> Metadata should ideally be captured in de-facto standards to achieve sustainability of the metadata itself thus sustaining meaningful access to the object and increasing the flexibility in regards to transfer and exchange of objects. Especially bibliographic metadata for online digital objects should be available in Dublin Core to enhance the possibility of metadata exchange across repositories and digital archives. Online data of complex works should also be accompanied by structural metadata, i.e. in the form of METS. Besides bibliographic metadata, technical and administrative metadata is desirable. The de-facto standard for technical and preservation metadata is PREMIS.

Capturing and handling bibliographical metadata is a traditional library science skill. Technical and preservation metadata significantly widens the scope, capturing and classifying information more or less native to the information technology sector.

*Gaps identified:*

As described in „Know what you have“, profound knowledge of different data carriers forms a gap for library professionals<sup>11</sup>. The situation in regards to handling data carriers is not quite

<sup>10</sup> (Klijn, de Lusenet 2008, p.37)

<sup>11</sup> This is viewed from a general point. For example Archivists specializing in AV-materials do have a good knowledge of the data carriers relevant to their field.

as drastic as it is the case of surveying data carriers. In general a good understanding of how to properly handle different digital media exists. Experience has however shown that especially younger colleagues can not necessarily identify older data carriers like floppy discs or zip discs and do not know how to handle them accordingly.

While profound knowledge of bibliographical and structural metadata exists amongst library practitioners, technical and preservation metadata introduces a new field to library and technology staff alike. An example for this is intrinsic information of a file captured as significant properties within a PREMIS file. While the information captured is new to the library science field, the fact that the information is captured in metadata is new to the information technology field.

#### *Information sources identified:*

As metadata is a core library and information science competency, most regional training centers for professionals-on-the-job should offer workshops and seminars on bibliographical and structural metadata, thus enabling library staff to refresh their knowledge or learn about newer standards. Workshops on technical and preservation metadata are rather rare, as this is a relatively new and highly specialized topic. The PREMIS editorial board announces workshops on its website and offers materials of past workshops and presentations for download.<sup>12</sup>

The websites of the Dublin Core Metadata Initiative<sup>13</sup>, of the PREMIS board<sup>14</sup> and of the METS Board<sup>15</sup> are good information sources for the respective metadata standards.

Furthermore Jenn Riley created an excellent visualization, accompanied by a glossary pamphlet, of today's metadata universe<sup>16</sup>.

While information on data carriers was already mentioned in the last chapter, the reading suggestions below contain two works specifically on care and handling of optical and magnetic media.

#### *Further reading:*

Byers, Fred R. (2003). Care and Handling of CDs and DVDs: A guide for Librarians and Archivists. CLIR Reports. Pub121. Available online at <http://www.clir.org/pubs/abstract/pub121abst.html>, last checked May 22<sup>nd</sup> 2011.

Brown, Adrian (2008). Care, Handling and Storage of Removable Media. Digital Preservation Guidance Note: 3. The National Archives. Available online at <http://www.nationalarchives.gov.uk/documents/removable-media-care.pdf>, last checked May 22<sup>nd</sup> 2011.

Caplan, Priscilla (2009). Understanding PREMIS. Library of Congress. Available online at <http://www.loc.gov/standards/premis/understanding-premis.pdf>, last checked May 22<sup>nd</sup> 2011.

*Please note that the PREMIS website also contains translations of this work into Spanish, Italian and German.*

---

<sup>12</sup> <http://www.loc.gov/standards/premis/tutorials.html>

<sup>13</sup> <http://dublincore.org/>

<sup>14</sup> <http://www.loc.gov/standards/premis/>

<sup>15</sup> <http://www.loc.gov/standards/mets/>

<sup>16</sup> <http://www.dlib.indiana.edu/~jenlrile/metadatamap/>

Van Bogart, John (1995). Magnetic Tape Storage and Handling. CLIR Reports. Pub54. Available online at <http://www.clir.org/pubs/reports/pub54/index.html>, last checked May 22<sup>nd</sup> 2011.

## **“Know what comes ahead”**

Knowing what is future-proof is an impossible goal to reach. Nevertheless it should be used as a guideline for all digital preservation activities. Integrating a sort of “permanent outlook” from a sustainability standpoint into digital curation activities will significantly improve the awareness of digital preservation and of risks that data can face.

Sustainability of access to digital information needs to be based on content determined factors as well as technology determined factors alike, thus requiring a profound understanding of both. Knowledge of content is certainly a traditional library and information science skill, whereas knowledge of technological possibilities and limitations are traditional computer science skills.

Content knowledge in the digital age goes beyond knowing the intellectual information contained in an object. Digital information has widened the scope and it can be said that content knowledge today includes knowledge of presentation and usability of a digital object. Today’s content specialist knows whether a digitized book within a library’s presentation platform can be full-text searched and zoomed to different levels as well as whether it can be exported into formats such as pdf and xml. The presentation and usability in the given example meet today’s user expectancies but as technology advances user expectancies are bound to change. Changing user expectancies should to be followed closely through a community watch as they play a key role in digital preservation strategies.<sup>17</sup>

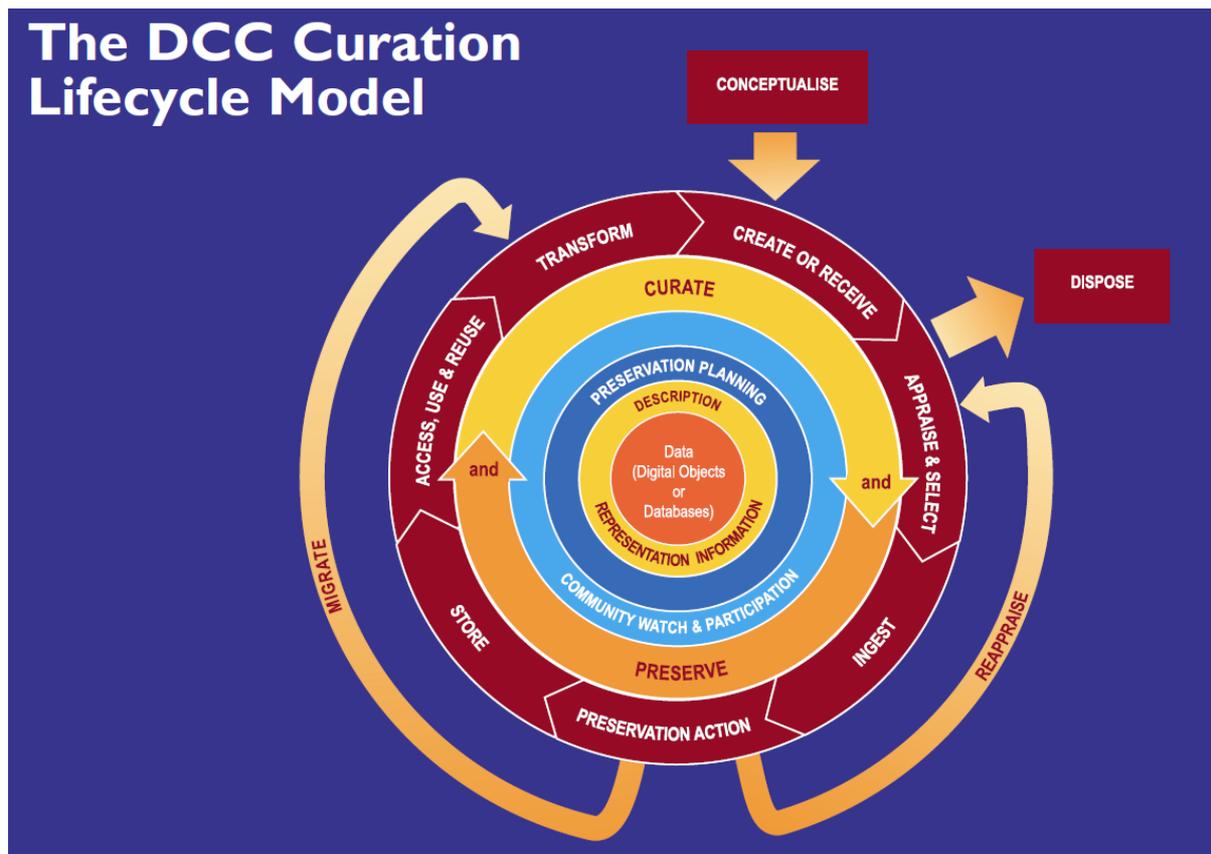
Technological knowledge applies to technical sustainability of digital frameworks and formats. This covers aspects like whether a format is bound to certain licenses, whether it is supported by a limited or a wider range of software, the issue of lossless or lossy data reduction and possible limitations due to digital rights management (DRM).

Content experts, information technology experts and digital preservation experts need to work hand in hand in the development of digital preservation strategies. Decisions like a possible normalization, i.e. a limitation to a certain set of file formats for a collection, need to consider users’ expectancies as well as technological constraints.

---

<sup>17</sup> In digital preservation practise user needs in access and preservation needs are often addressed through separate access and preservation representations of an object. Nevertheless the business process needs to define user expectancies closely to ensure that expected access copies can be derived from the preservation copy, if needed. Moreover separate copies are not feasible to all institutions e.g. due to limitations in storage space.

# The DCC Curation Lifecycle Model



Source: DCC. The DCC Curation Lifecycle Model. Available online at <http://www.dcc.ac.uk/sites/default/files/DCCLifecycle.pdf>, last checked May 22<sup>nd</sup>, 2011.

## Gaps identified:

Profound knowledge of content related criteria and technology related criteria exists. However, applying this knowledge in the context of digital preservation is not always integrated into digital curation practise of content experts and information technology experts alike. The understanding of digital preservation needs need to be fostered and should be integrated into digital projects from the start. Technology Watch and User Expectancy Watch form new fields of activity for libraries.

## Information sources identified:

DCC (Digital Curation Centre) is the national digital curation center of the UK. A number of excellent resources regarding a holistic approach to digital curation – including preservation – can be found on the DCC’s website.<sup>18</sup> Furthermore the DCC conducts digital curation trainings. As part of the DCC’s “Train the Trainer” effort materials for the courses can be found online.<sup>19</sup>

Another UK initiative, the Digital Preservation Coalition (DPC) is known for its excellent series of technology watch reports.<sup>20</sup>

## Further reading:

<sup>18</sup> <http://www.dcc.ac.uk>

<sup>19</sup> <http://www.dcc.ac.uk/training/train-the-trainer>

<sup>20</sup> <http://www.dpconline.org>

DCC. Digital Curation 101. Community Watch & Participation. Online available at <http://www.dcc.ac.uk/sites/default/files/DC%20101%20Community%20Watch.pdf>, last checked May 22<sup>nd</sup>, 2011.

DCC. Digital Curation 101. Curate and Preserve. Online available at <http://www.dcc.ac.uk/sites/default/files/DC%20101%20Curation%20and%20Preservation.pdf>, last checked May 22<sup>nd</sup>, 2011.

Library of Congress. Sustainability of Digital Formats – Planning for Library of Congress Collections. Online available at <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>, last checked May 22<sup>nd</sup>, 2011.

Judtih Rog, Caroline van Wijk. Evaluating File Formats for Long-term Preservation. Online available at [http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/KB\\_file\\_format\\_evaluation\\_method\\_27022008.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/KB_file_format_evaluation_method_27022008.pdf) last checked May 22<sup>nd</sup> 2011.

The National Archives. The technical registry PRONOM. Online available at <http://www.nationalarchives.gov.uk/PRONOM> last checked May 22<sup>nd</sup> 2011.

## **“Know how to keep it safe”**

While the previous points dealt with necessary “foundation” work for digital preservation practise, the information gathered as part of tasks like a collection survey and a process description needs to be evaluated and used to formulate strategies, plans and action.

All information gathered, including the process analysis, survey results and information on the type of metadata captured needs to be incorporated into a risk assessment report. Risks need to be identified and ranked in order to adequately plan prioritization in action. Risks can stem from a variety of areas – for instance technology based as in the case of media deterioration or business process based, as in the case of legal constraints. This makes risk assessment a complex procedure, which requires expert knowledge from library professionals, information technology professionals and also the management sector. The process should be guided by digital preservation experts. The same holds true for preservation planning<sup>21</sup>:

“A preservation plan defines a series of preservation actions to be taken by a responsible institution due to an identified risk for a given set of digital objects or records (called collection). The Preservation Plan takes into account the preservation policies, legal obligations, organisational and technical constraints, user requirements and preservation goals and describes the preservation context, the evaluated preservation strategies and the resulting decision for one strategy, including the reasoning for the decision.”

Preservation strategies and action are always based on risk assessment and preservation planning. The implementation of strategies and actions, ranging from the planning and building of a digital archive to migration of a set of files, requires a profound knowledge of the relevant library and technology processes involved but also an excellent understanding of digital preservation practises.

---

<sup>21</sup> (Becker et al. 2009)

### *Gaps identified:*

Preservation Planning and preservation action builds upon the information covered in the previous chapters. It furthermore requires a good understanding of digital preservation practise. This introduces the new role of the preservation specialist, which is briefly covered in the next chapter. Nevertheless the library and information technology experts involved in the preservation processes need to be familiar with digital preservation fundamentals like standards and subject specific vocabulary. The development of strategies in the digital preservation process needs to be based on processes which integrate different aspects and fields. The ability to “think-outside-the-box” is one of the biggest assets for the successful development of strategies.

### *Information sources identified:*

National competence centers like nestor<sup>22</sup>, NDIPP<sup>23</sup> and the DPC<sup>24</sup> offer training programs and workshops targeted at different groups, ranging from “Getting started in Digital Preservation” (DPC) to “Collaboration, Cooperation and Challenges in Digital Preservation” (nestor). Furthermore competence centers offer practitioner meetings like the “nestor Praxistage” or the NDIPP “Best Practises Exchange”.

Since no fixed schema for risk assessment exists, knowledge needs to be gained through best practises. An excellent information source for a risk assessment process is the report “Risk Assessment 2007” by the British Library Digital Preservation Team.<sup>25</sup>

Extensive work in preservation planning has been done by the EU planets project, which ran from 2006 – 2010. Information and training material can be found on the Plato page, the home of planets preservation planning tool: <http://www.ifs.tuwien.ac.at/dp/plato/intro.html> Further training material from the planets project covering topics such as “How to understand files” can be found on the “training materials” section of the project’s homepage.<sup>26</sup>

### *Further reading:*

DPC (2009). DPC Jargon Buster. Online available at [http://www.dpconline.org/component/docman/doc\\_download/379-dpc-jargon-buster-july-2009](http://www.dpconline.org/component/docman/doc_download/379-dpc-jargon-buster-july-2009), last checked May 22<sup>nd</sup> 2011.

DPC (2009). Digital preservation handbook. Online available at <http://www.dpconline.org/advice/preservationhandbook>, last checked May 22<sup>nd</sup> 2011.

nestor (2008). Into the Archive – A guide for the information transfer to a digital repository. Online available at [http://files.d-nb.de/nestor/materialien/nestor\\_mat\\_10\\_en.pdf](http://files.d-nb.de/nestor/materialien/nestor_mat_10_en.pdf), last checked May 22<sup>nd</sup> 2011.

---

<sup>22</sup> Nestor is the network of expertise for the long-term storage and accessibility of digital resources in Germany <http://www.langzeitarchivierung.de>

<sup>23</sup> NDIPP is the National Digital Information Infrastructure and Preservation Program at the Library of Congress <http://www.digitalpreservation.gov/>

<sup>24</sup> DPC is the UK based Digital Preservation Coalition <http://www.dpconline.org/>

<sup>25</sup> <http://www.bl.uk/aboutus/stratpolprog/ccare/introduction/digital/riskassessment.pdf>.

<sup>26</sup> <http://www.planets-project.eu/training-materials/>

British Library – Digital Preservation Team. Risk assessment 2007. Online available at <http://www.bl.uk/aboutus/stratpolprog/ccare/introduction/digital/riskassessment.pdf>, last checked May 22nd 2011.

## The role of the digital preservation specialist

The DPOE study cited in the beginning of the paper showed that a significant amount of the libraries surveyed do not employ special digital preservation staff. Instead the digital preservation work is assigned to various staff as needed.

Institutions that hold a stewardship to digital cultural heritage need to implement sustainable digital preservation strategies. Organizational infrastructure plays a vital role in a trusted digital preservation process. The U.S. Center for Research Libraries' "Trustworthy Repositories Audit & Certification: Criteria and Checklist" describes staffing requirements as follows<sup>27</sup>:

“Staffing for the repository must be adequate for the scope and mission of the archiving program. The repository should be able to demonstrate an effort to determine the appropriate number and level of staff that corresponds to requirements and commitments. (These requirements are related to the core functionality covered by a certification process. Of particular interest to repository certification is whether the organization has appropriate staff to support activities related to the long-term *preservation* of the data.) The accumulated commitments of the repository can be identified in deposit agreements, service contracts, licenses, mission statements, work plans, priorities, goals, and objectives. Understaffing or a mismatch between commitments and staffing indicates that the repository cannot fulfill its agreements and requirements.”

Smaller institutions that do not hold a stewardship for digital cultural heritage might be able to stem tasks like surveying the holdings without full-time digital preservation staff, thus possibly selecting and preparing data to be handed to national institutions conducting the digital preservation for them. For larger institutions conducting the actual preservation in a trustworthy digital preservation repository, digital preservation specialists are an indispensable asset. Their tasks include the planning and steering of all relevant processes described in the previous chapters as well as further relevant technical and business analyst tasks.. They need to be familiar with best-practises in digital preservation and follow developments closely.

### *Gaps identified:*

Digital preservation experts need to have a sound understanding of library processes and technological processes alike. College and university level courses in digital preservation have only recently been established.

As digital preservation is a global problem, active involvement in the national and international community is of immense importance. The community fosters the development of strategies and tools and is an important platform for necessary knowledge exchange. The fast pace at which technological exchange takes place results in a constant evaluation of the strategies implemented. In this regard, the digital preservation specialist must be committed to “lifelong learning”.

---

<sup>27</sup> (US. Center for Research Libraries 2007, p. 11)

### *Information sources identified:*

Digital preservation is a relatively young and very active field. The best information source is the community itself. Workshops conducted by national competence centers or national and international projects with a digital preservation focus foster active exchange between practitioners. DPE (Digital Preservation Europe) maintains registries of training materials, competence centers, projects and online resources. These form a good starting for both general as well as specific information sources about digital preservation.<sup>28</sup>

A vital part of the information exchange takes place online, as in the case of the Wiki and Blog provided by the Open Planets Foundation (OPF).<sup>29</sup>

### *Further reading:*

U.S. Center for Research Libraries (2007). Trustworthy Repositories Audit & Certification: Criteria and Checklist. Available online at [http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf), last checked May 22<sup>nd</sup> 2011.

nestor (2008). Catalogue of Criteria for Trusted Digital Repositories. Version 2. nestor-materials 8. Available online at [http://files.d-nb.de/nestor/materialien/nestor\\_mat\\_08\\_eng.pdf](http://files.d-nb.de/nestor/materialien/nestor_mat_08_eng.pdf), last checked May 22<sup>nd</sup> 2011.

## **Conclusion**

The tasks briefly described in this paper ranged from necessary “basis” work for digital preservation like surveying the digital holdings to specialized work like preservation planning. Preservation is a core competency of librarians, however, gaps still exist when it comes to being able to accept that role fully in the digital realm.

The gaps outlined show where concrete knowledge is lacking. One problem which is hard to capture as part of a gap analysis is the ability to develop sustainable strategies. The importance of this needs to be stressed, as the focus in digital preservation does not rely on “ready-made” fixed concepts but rather on the application of all factors in an immersed approach.

It is not possible to draw a direct correlation between Goportis’ experiences and the state of current vocational and university programs for library & information science. What can be said is that digital preservation schooling for practitioners-on-the-job is currently mainly conducted either by memory institutions who are directly involved in digital preservation work, by funded projects or by the national and international competence centers for digital preservation. Classical preservation workshops, on the other hand, or also conducted by library union offices, broader library associations and commercial training institutes. This implies that there is still significant room for growth of digital preservation awareness.

---

<sup>28</sup> <http://www.digitalpreservationeurope.eu/registries/materials/>

<sup>29</sup> <http://www.openplanetsfoundation.org>

The international DigCurV project is currently analyzing the state of vocational training in digital preservation. It's goal is "to establish a curriculum framework for vocational training in digital curation launched today" (DigCurV 2011).

Goportis' experience showed that a step-by-step process involving all areas of the library that deal with digital material will help in raising the awareness of the dangers that digital material face. Collaboration and community participation are the key factors in furthering understanding and knowledge of digital preservation – within the library as well as outside the institution, on national and international levels.

## Works Cited

Becker, Christoph et. al (2009). Systematic planning for digital preservation: evaluating potential strategies and building preservation plans. International Journal of Digital Libraries (IJDL). December 2009. Online accessible at <http://www.ifs.tuwien.ac.at/~becker/pubs/becker-ijdl2009.pdf> , last checked May 22<sup>nd</sup> 2011.

Brown, Adrian (2003). Selecting Storage Media for Long-Term Preservation. Digital Preservation Guidance Note: 2. The National Archives. Online accessible at [http://www.nationalarchives.gov.uk/documents/selecting\\_storage\\_media.pdf](http://www.nationalarchives.gov.uk/documents/selecting_storage_media.pdf) last checked May 22<sup>nd</sup> 2011.

Byers, Fred R. (2003). Care and Handling of CDs and DVDs: A guide for Librarians and Archivists. CLIR Reports. Pub121. Available online at <http://www.clir.org/pubs/abstract/pub121abst.html> , last checked May 22<sup>nd</sup> 2011.

Casey, Mike (2009). Indiana University Bloomington. Media Preservation Survey. A Report. Available online at [http://research.iub.edu/communications/media\\_preservation/iub\\_media\\_preservation\\_survey\\_FINALwww.pdf](http://research.iub.edu/communications/media_preservation/iub_media_preservation_survey_FINALwww.pdf), last checked May 22<sup>nd</sup> 2011.

DigCurV (2011). About. Digital Curator Vocational Education Europe Website. Available online at <http://www.digcur-education.org/eng/About>, last checked May 22<sup>nd</sup> 2011.

DPOE (2010). 2010 DPOE Training Needs Assessment Survey. Cross-Tabulated questions. Available online at [http://www.digitalpreservation.gov/education/documents/DPOE2010Survey\\_CrossTabs.pdf](http://www.digitalpreservation.gov/education/documents/DPOE2010Survey_CrossTabs.pdf), last checked May 22<sup>nd</sup> 2011.

U.S. Center for Research Libraries (2007). Trustworthy Repositories Audit & Certification: Criteria and Checklist. Available online at [http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf), last checked May 22<sup>nd</sup> 2011.

Van Bogart, John (1995). Magnetic Tape Storage and Handling. CLIR Reports. Pub54. Available online at <http://www.clir.org/pubs/reports/pub54/index.html>, last checked May 22<sup>nd</sup> 2011.

Klijn, Edwin; Lusenet, Yola de (2008). Tracking the reel world. A survey of audiovisual collections in Europe. European Commission on Preservation and Access.