



## Zpráva ze zahraniční služební cesty

Jméno a příjmení účastníka cesty	<b>Tomáš Svoboda</b>
Pracoviště – dle organizační struktury	Úsek informačních technologií (název v době SC)
Pracoviště – zařazení	ředitel
Důvod cesty	<b>Účast na konferenci „Aligning National Approches to Digital Preservation“ , účast v ESDI Roundtable</b>
Místo – město	<b>Tallinn</b>
Místo – země	<b>Estonsko</b>
Datum (od-do)	<b>24. - 27.5. 2011</b>
Podrobný časový harmonogram	24.5. – přelet do Tallinu – registrace na konferenci ANDP  Středa 25.5. Konference ANDP Čtvrtek 26.5. ESDI RoundTable
Spolucestující z NK	<b>Mgr. Andrea Fojtů, PhDr. Bohdana Stoklasová</b>
Finanční zajištění	<b>IOP NDK</b>
Cíle cesty	Přítomnost na konferenci s mezinárodní účastí, získání nových informací a odborných kontaktů pro oblast dlouhodobé ochrany digitálního dokumentů a povinného elektronického výtisku.
Plnění cílů cesty (konkrétně)	<p>V souladu se stanovenými cíli cesty jsem se účastnil konference a následujícího kulatého stolu.</p> <p>Konference byla věnována mezinárodní koordinaci "alignment" v záležitostech dlouhodobé ochrany digitálních informací.</p> <p>Byly přednášeny nesčetné příklady svědčící o tom, jak je všeobecně zanedbávána péče o digitální data, přestože na jejich vytvoření jsou investovány nemalé prostředky.</p> <p>Jako podnětný zaznívá "celostní přístup", podle kterého dlouhodobá ochrana dat je předvídána už v době vzniku těchto dat a prolíná se celým jejich životním cyklem. To se projeví například v přehledné struktuře připravené pro archivaci. Naproti tomu stojí klasický/zastaralý model OAIS, který bere data jako hotová a jejich ochranu pojímá jako izolovaný proces.</p> <p>Vystoupení Dr.Stoklasové akcentovalo publicitu projektu NDK a Národné knihovny ČR.</p> <p>Následný round-table byl věnován ekonomické soběstačnosti ochraňujících procesů.</p> <p>Vynikající logistické zajištění ze strany NK Estonska.</p>
Program a další podrobnější informace	Uloženy v UIT
Přivezené materiály	Uloženy v UIT

Datum předložení zprávy	29.8.2011
Podpis předkladatele zprávy	
Podpis nadřízeného	
Vloženo na Intranet	
Přijato v mezinárodním oddělení	

## Příloha 1: Program konference Aligning National Approches to Digital Preservation včetně odkazů na prezentace

### Conference Presentations

- [Program](#)
- [Keynotes](#)
- [Panel Chairs](#)
- [Panelists](#)
- [Resources](#)
- [Volume](#)
- [Committee](#)



### Keynote Presentations

#### Day 1

- [Laura Campbell](#), *Exploring What We Can Do Together: Strategic Alignments for International Collaboration.* ([PDF](#))

#### Day 2

- [Gunnar Sahlin](#), *International and National Collaboration in the Digital Age.* ([PDF](#))

### Panel Presentations

#### Technical Alignment

The Technical Panel, Chaired by [Michael Seadle](#) from Berlin School of Library and Information Science, Humboldt-Universität zu Berlin, explored two key issues for national alignment: infrastructure and testing. Infrastructure includes the hardware and software necessary for managing digital archiving systems as well as the communication protocols for sharing resources across the internet. Testing involves reproducible experiments using, if possible, real data to show whether software and hardware perform under conditions that reflect a reasonable hypothesis about the future.

- [Chair, Michael Seadle](#), *Technical Alignment: The Role of Testing.* ([PDF](#))
- [Andreas Rauber](#), *Technical Alignment.* ([PDF](#))
- [Adam Rusbridge](#), *Digital Preservation Infrastructure via UK LOCKSS Alliance.* ([PDF](#))
- Sabine Schrimpf, *DNB Contribution to the Tallinn Technical Alignment Panel.* ([PDF](#))

#### Organizational Alignment

The Organizational Panel, Chaired by [Inge Angevaare](#) from Netherlands Coalition for Digital Preservation, covered three sets of key issues. First of all, why is cooperation and collaboration so essential for dealing with digital preservation issues? This was followed by a review of types of cooperation and collaboration that have been developed around the world. Lastly, was a look at the enormous organizational challenge of securing long-term access to born-digital objects.

- [Chair, Inge Angevaare](#), Michelle Gallinger, David Giaretta, and [Martin Halbert](#), *Organizing Digital Preservation on an International Scale*. ([PDF](#))

### Standards Alignment

The Standards Panel, Chaired by [Raivo Ruusalepp](#) of Tallinn University, looked at collaboration opportunities in four key areas where standardisation already has a long-standing history of cooperation between different domains. First, metadata standards - their development and use across all stakeholders in digital preservation. Second technical standards – starting from the standardization of file format information to technical standards that should be adopted across the board of digital preservation tools and software. Third, approaching a standard for the digital preservation repository audit and how the current five separate approaches to repository audit compare with this. Finally, looking more specifically into information security issues in memory institutions and how these are approached in the light of cyber-defence measures. These four topics offered a wide range of areas for alignment and further collaboration that were explored in the break-out session.

- [Chair, Raivo Ruusalepp](#), [Matthew Woollard](#), [Christopher \(Cal\) Lee](#), and [Bram van der Werf](#), *Standards Alignment*. ([PDF](#))
- [Christopher \(Cal\) Lee](#), *Contextual Information as a Point of Alignment in Digital Preservation*. ([PDF](#))
- [Matthew Woollard](#), *Standards Based Approach to Preservation Planning*. ([PDF](#))

### Legal Alignment

The Legal Panel, Chaired by [Adrienne Muir](#) from Loughborough University, broadly covered legal and contractual issues associated with acquiring digital content for the purpose of long term preservation, copyright issues in digital preservation and contractual and governance issues in cooperative digital preservation. The implications of misalignment between approaches were identified as were examples where issues are being addressed in law and in practice.

- [Chair, Adrienne Muir](#), *Legal Deposit and Web Archiving*. ([PDF](#))
- [Dwayne Buttler](#), *It's a Big World After All*. ([PDF](#))
- [Wilma Mossink](#), *Digital Preservation and Access to Europe's Cultural History*. ([PDF](#))

### Education Alignment

The Education Panel, Chaired by [Joy Davidson](#) from HATII, University of Glasgow, reviewed recent developments in embedding data management and curation skills in information technology, library and information science, and research-based postgraduate courses in various national contexts. The panel also investigated means of joining up formal education with professional development training opportunities more coherently. The potential of professional internships as a means of improving communication and understanding between disciplines was also explored. A key aim of this panel was to identify what level of complementarity is needed across various disciplines to most effectively and efficiently support the entire data curation lifecycle.

- [Chair, Joy Davidson](#), *Digital Curation Centre*. ([PDF](#))
- [George Coulbourne](#), *Continuing Education*. ([PDF](#))
- [Sheila Corral](#), *Education Alignment: Informations Science Perspective*. ([PDF](#))
- [Andreas Rauber](#), no presentation.

### Economic Alignment

The Economics Panel, Chaired by [Maurizio Lunghi](#) of Fondazione Rinascimento Digitale, aimed to lead a critical discussion of sustainable strategies in the preservation of cultural heritage. Specifically, this panel gave an overview of multi-institutional experiences and approaches in defining directive elements for efficient workflows in digital repositories management. Additional issues discussed included a clear definition of the objectives, and roles & responsibilities for the user community and service providers for the sake of establishing credible policy. As well as the choice of suitable business models, and cost analyses as essential components of a sustainable workflow for the life of digital repositories.

- [Chair, Maurizio Lunghi, Economic Alignment. \(PDF\)](#)
- [Neil Grindley, Economic Alignment. \(PDF\)](#)
- [Bohdana Stoklasova, Czech National Digital Library and Digital Preservation. \(PDF\)](#)
- [Aaron Trehub, Sustainable Preservation in North America: ADPNet & Friends. \(PDF\)](#)



#### Closing Remarks & Synthesis

- Cliff Lynch, transcript forthcoming.

## Příloha 2: Program ESDI Roundtable

The ESDI Roundtable

- [Program](#)
- [Keynotes](#)
- [Panel Chairs](#)
- [Panelists](#)
- [Resources](#)
- [Volume](#)
- [Committee](#)

**A Roundtable Meeting on the Economic Sustainability of Digital Information  
(The ESDI Roundtable)  
Thursday 26<sup>th</sup> May, 2011  
Cupola Hall, National Library of Estonia, Tallinn, Estonia**



*'By attending to the value of digital information, providing incentives to preserve these digital assets, and ensuring allocation of roles and responsibilities among stakeholders that share a common interest in valuable digital assets, we can continue to build high the shared body of knowledge that will enable all of us to see farther.'* [Sustainable Economics for a Digital Planet \(Feb 2010\)](#)

These were the concluding remarks in the Final Report of the *Blue Ribbon Task Force on Sustainable Digital Preservation and Access* and they underline the importance of what is at stake when we talk about the sustainability of digital information. The effective passing on of knowledge from one generation to the next underpins all aspects of society and enables new learning and new research to take place. The work of the Blue Ribbon Task Force (BRTF) was a significant and novel addition to the literature on Digital Preservation and was

the first systematic attempt to focus not just on the cost of managing information over time, but on the economic framework that is required to allow that to happen.

The purpose of this roundtable meeting is not only to provide an international forum to discuss the implications of the BRTF conclusions and recommendations, but also to focus on new work in this area, and to hear from a range of participants about the various national actions that are being taken to ensure an economically sustainable digital future.

The meeting will feature a presentation and discussion of work that is currently being funded by JISC and OCLC Research to produce a reference model that will help decision-makers understand the economic context surrounding digital lifecycle management, and inform the development of viable economic sustainability strategies. The reference model will be based on the findings and recommendations of the BRTF Final Report, but aims to translate the report's conclusions into a practically-oriented tool for economic decision-making. The goal is to produce a reference model that underpins the economic aspect of lifecycle digital planning and management much as the OAIS Reference Model has underpinned planning and management for technical/workflow issues.

Other initiatives that may be introduced at the meeting include work being contemplated by LIBER (the Association of European Research Libraries) and some of the sustainability work by Ithaka S+R.

The two main objectives of the meeting are:

1. To present and receive feedback on the draft economic reference model and to discuss next steps
2. To hear from a range of participants about national initiatives and reactions to the challenge of economically sustaining digital resources

In common with the conference to which this meeting is attached, the underlying agenda is to coordinate and align national approaches and to share what effective knowledge we have in this field for the benefit of all.

Whilst the event is free to attend, places at this meeting are very limited and some will be filled by targeted invitation. However, anyone with a specific interest in this area is welcome to register their wish to participate, and if places are available, we will include everyone we can. Please note that this will be a roundtable meeting and all participants should come along expecting to take an active role in the discussion.

#### **Draft Programme (subject to confirmation)**

09:00 - 09:30 Assemble

09:30 - 09:45 Introduction

- **Neil Grindley (JISC)**

09:45 - 10:15 Presentation

- **Chris Rusbridge/Brian Lavoie (OCLC):** *The Draft Reference Model for the Economic Sustainability of Digital Information*

10:15 - 11:00 Discussion on next steps for the Reference Model

11:00 - 11:30 Break

11:30 - 12:00 Presentation

- **Wouter Schallier (LIBER):** *Work on the readiness of European stakeholders to support economic sustainability*

12:00 - 12:30 Discussion on international responses to Digital Preservation Sustainability Issues

12:30 - 13:30 Lunch

13:30 - 14:00 Presentation

- **Ithaka:** *Work on sustainability*

14:00 - 14:30 Discussion on future work and emerging ideas

14:30 - 15:00 Rapporteur feedback/capture actions

15:00 Finish





# The International Journal of Digital Curation

Issue x, Volume x | 2xxx

## Economic Alignment

Maurizio Lunghi,  
Chair,  
Fondazione Rinascimento Digitale

Neil Grindley,  
Panel Speaker,  
Joint Information Systems Committee

Bohdana Stoklasová,  
Panel Speaker,  
National Library of Czech Republic

Aaron Trehub,  
Panel Speaker,  
Auburn University, ADPNet

Christin Mollenhauer,  
Fondazione Rinascimento Digitale

### Abstract

The Economics Panel, chaired by Maurizio Lunghi of Fondazione Rinascimento Digitale, aims to lead a critical discussion about the nature of the economic issues that define and inhibit effective national and international progress in relation to preserving digital cultural heritage materials. Specifically, this panel will present and reflect on multi-institutional experiences and approaches in defining directive elements for efficient workflows and policies in managed digital information environments. Related issues will also be introduced including: service/user relationships; roles and responsibilities throughout the various communities; the choice of suitable business models; and cost analyses as essential components of defining sustainable economic approaches to preservation. In keeping with the aims of the conference, the panel will conclude by considering what a blueprint for success in this area may resemble.

---

## Introduction

### Context of the paper

The ‘Aligning National Approaches to Digital Preservation’ conference<sup>1</sup> brings together digital preservation experts in Tallinn in May 2011 to launch an action plan for supporting digital preservation strategies for our collective cultural heritage. The event focuses on six strategic sections: organizational, technical, legal, educational, standards and lastly, economic, which is the topic of this paper. It is implicit in the conference title that successful strategic approaches will be accomplished by national and international alignment, and it is from this perspective that the issues will be examined.

The four economic panelists bring diverse but complementary approaches to the discussion:

- Chair, Maurizio Lunghi (Fondazione Rinascimento Digitale)
- Neil Grindley (Joint Information Systems Committee)
- Bohdana Stoklasova (National Library of Czech Republic)
- Aaron Trehub (Auburn University, ADPNet)

The panel session is followed by a roundtable and open discussion which will elicit additional views and opinions from delegates. This input will feed into a new post-conference paper, which in turn will be one of the inputs to the conference ‘action plan’ deliverable.

### Structure of this report

This report focuses on the core aims of the ‘Aligning National Approaches to Digital Preservation’ conference and attempts to respond to questions posed by the organisers to the panel participants. The first question was to consider (from the economic perspective) the most important alignment accomplishments that have taken place in the digital preservation field. The second was to examine the current challenges and gaps that represent barriers to establishing sustainable digital preservation activities. And the third asked where the panel thought the digital preservation community should aim to be in five years time and what would success in this area look like?

These generic questions conveniently encompass many of the issues that have a bearing on sustainable economic digital preservation strategy and action. There are a number of questions and issues, however, that relate more specifically to the topic of this panel and they include: the nature of costs and business models; the effectiveness and demand for services; strategies for selection and appraisal; requirements for partnership and training; and the general need for clarity around roles and responsibilities. The report is prefaced with some general remarks about digital preservation and its value proposition.

### Digital preservation

---

<sup>1</sup> For further information please visit <http://www.educopia.org/events/ANADP/>.

---

The long-term preservation of digital materials is an issue that has global relevance and it has become more widely understood over the last decade or so that engagement with preservation is an unavoidable corollary to the creation and use of nearly all forms of digital content. To the casual observer, however, over a five year period of watching information technology (IT) related activity within an organisation, there may not be much observable difference between an organisation that demonstrates the usual responsible approach to ensuring that their IT infrastructure is robust, fit-for-purpose, and responsive to user requirements; and one that (from the outset) pursues a determined strategy of planning and implementing digital preservation as an additional expense to maintaining and delivering the IT function.

It is clear then that five years—in digital preservation terms—is not that long and it is probably the single most critical reason that making any kind of business case or economic argument for preservation is a difficult proposition. Individuals, organisations and businesses are usually highly motivated to principally think about the issues and challenges they are likely to face in the next phase of planning, which generally means several years. So the first and most important concept to argue is that digital information *does have implicit enduring value*; or alternatively, that it can be used to create entities that will have value. Whilst a case can—and should more often—be made for some of the short-term benefits of preservation, it is this long-term value proposition that underpins all other arguments and evidence for engagement in this area.

Digital preservation often looks across to its equivalent in the physical realm and cites the maintenance of manuscripts over centuries as proof of the impact and worth of caring about the integrity, complexity, intricacy and context of materials produced by human endeavour. The starting point for this paper, and presumably for all those attending the conference, is that digital preservation is an important activity that will enable subsequent generations to make choices and exploit opportunities that they would otherwise be unable to take advantage of. It is ultimately these human outcomes, rather than technical or bureaucratic ones, that make the economic and every other case for digital preservation.

Putting the ideological view to one side and given the understandable focus of most people on short-term goals, acceptance among a wide range of information professionals of the importance of digital preservation as an essential and embedded part of their daily work is always going to be a challenge. Given that issues span technical, legal, educational, organisational and of course economic categories, there is an innate complexity to tackling digital preservation that many find a disincentive to engagement. For the minority that find this complexity stimulating, digital preservation continues to present rewarding intellectual opportunities. For the vast majority, continuing ‘access to’ or future ‘use of’ the preserved materials will always be the principle motivation for continuing to fund preservation activity. This level of interest from the user community is crucial. Preservation, whether physical or digital, is going to seem like wasted investment without any current or future usage intention. If the demand for access to preserved digital objects and their permanent storage is well-articulated, then economic sustainability becomes far more likely. If those arguments originate from across the community, and even across national boundaries, then so much the better.

---

The difficulty of assigning accurate value to digital information is a global problem and sharing that problem is a good mitigation measure. Whilst it may be possible in hindsight to judge that people made errors of judgement in assigning substantial resources to preserving material that was subsequently never used or was considered of negligible value, it will be a compelling defence to cite community, national or international precedent as proof of good faith.

## **Economic Alignment**

### **Core Approaches**

#### **Most important alignment accomplishments**

The first core task of the panel is to consider where progress has already been made either nationally or internationally to help ameliorate problems relating to the economics of preservation. The topic itself encapsulates a lot of complexity in that there are various perspectives that need to be factored into any discussion of what constitutes economic issues in this field. The focus could conceivably be on the cost of maintaining digital material over time; the budget strategies of organisations obliged to engage with preservation; the economic framework in which preservation may effectively occur; or the type and extent of funding required for effective preservation to flourish. This paper takes the view that all of these perspectives are valid areas for discussion, though some have been subject to more development and attention than others in terms of the amount of alignment that may have occurred.

The first of those options, the work that has been done on the lifecycle cost of information management, is arguably the most widely understood interpretation of any question about the ‘economics of preservation’ and probably makes the most immediate sense to the non-specialist who may be concerned to know whether preservation constitutes a ‘nice-to-have’, but dispensable layer of assurance, or whether it is an information management necessity. Obviously knowing the cost of preservation does not necessarily determine this issue, but it may focus the enquirer’s mind on how seriously he or she needs to contemplate the question.

The cost of preservation has recently been the focus of various phases of the LIFE project<sup>2</sup> undertaken by the British Library and University College London, who developed and refined a lifecycle model that primarily relates to materials that may be found in a (digital) library context (e.g. text and images), and developed a complex spreadsheet tool to help with calculating the cost over time of storing, managing and preserving that material. This work has also been picked up and further developed by Danish National organisations<sup>3</sup> and an online version of the costing tool is being developed and piloted by the Humanities Advanced Technology & Information Institute (HATII) at the University of Glasgow in collaboration with the Open Planets Foundation (OPF).

Further detailed work looking at the long term cost of preserving materials, in this instance research data, was carried out in two phases of reporting by the Keeping Research Data Safe (KRDS) project.<sup>4</sup> As well as relying on new research in

---

<sup>2</sup> Lifecycle Information for e-Literature: <http://www.life.ac.uk/>

<sup>3</sup> Danish National Archives, Royal Library Denmark, State and University Library: [eprints.ucl.ac.uk/9313/1/9313.ppt](http://eprints.ucl.ac.uk/9313/1/9313.ppt)

<sup>4</sup> <http://www.beagrie.com/krds.php>

---

collaboration with data centres to assess the real costs of keeping data over long periods, it drew on both the LIFE Project modelling work and the Cost Estimation Tool (CET) developed by NASA; as well as other resources such as: the TRAC (Transparent Approach to Costing) Model; the Open Archival Information System (OAIS) Reference Model; and the Digital Curation Centre (DCC) Lifecycle Model, in order to create an effective generic framework to discover the cost of managing research data.

More generally, the cost of digital preservation figured prominently in the *eSpida* Project at the University of Glasgow, an initiative aimed at ‘mak[ing] business cases for proposals that may not necessarily offer immediate financial benefit to an organisation, but rather bring benefit in more intangible spheres’—itself a pretty good characterization of digital preservation.<sup>5</sup>

So the stewardship costs of keeping digital material over time has been demonstrably tackled by various projects, both recently and in the past, and it seems appropriate to declare that some alignment around this work, and the initiatives of other organisations and projects on this topic, has taken place. If not around the precise cost of various preservation tasks, then at least around some of the digital lifecycle information models on which they are based. These models are themselves significant as the digital equivalent to earlier examples from the realm of archival practice and records management, the former a discipline that goes back hundreds of years and the latter an activity that emerged in response to the burgeoning amount of documentation being produced during the middle of the 20<sup>th</sup> century. There is plenty of evidence to suggest that the lifecycle of information and its management is well understood by now, and there is also reason to believe that digital preservation and curation components have been successfully described and categorised. Whether every component in a diagram such as the DCC Lifecycle model<sup>6</sup> is understood and implementable (or even practical to contemplate) for many organisations is another question, but there does appear to be some alignment and agreement about the nature of, and the relationship between, preservation tasks.

Slightly more contentious, particularly beyond the edges of the broad preservation community, is the notion that there is alignment around the principle of selection and appraisal. This is a deeply significant point in relation to the economics of preservation as the amount of material that one chooses to keep does, of course, have an impact on the infrastructure that one needs to manage it with.

It seems justified, however, to state that at least amongst communities who have spent time thinking hard about the consequences of information management policies (economic and otherwise), there is alignment about the value—indeed the necessity—of selecting and appraising digital information, in effect, assigning value to it and prioritising some data as more valuable than others. What there is less alignment about is the practicality and processes for actually carrying out appraisal routines. There is, however, a ‘gaps and challenges’ section later in this paper to accommodate a discussion of this sort.

---

<sup>5</sup> <http://www.gla.ac.uk/espida/>. See Currall, James, and Peter McKinney. 2006. “Investing in Value: A Perspective on Digital Preservation.” *D-Lib Magazine* 12 (4). doi:10.1045/april2006-mckinney.

<sup>6</sup> <http://www.dcc.ac.uk/resources/curation-lifecycle-model>

---

Another highly visible area of alignment that must surely result in enhanced economic sustainability for preservation is the amount of community building and national and international collaborations that occur, not only as a result of the numerous seminars, workshops and conferences that take place around the world, but also from the open flow of information that generally occurs between preservation practitioners, many of whom are based within public and non-profit institutions such as universities, libraries and archives. Whilst it would be banal to spell out the benefits of cooperation and discussion between theorists and practitioners in any given field, the exchange of experience and good and bad practice; wide participation in advocacy and awareness raising; and the development of common terminology and common approaches, have all been key components of establishing digital preservation as a sub-discipline. International cooperation has not simply been a by-product or an extension of the peer-review process: it has been critical for the establishment of practice and policy in a field where many onlookers are still waiting to hear and understand what a convincing and robust long-term business-case for preservation looks and sounds like.

As well as forming useful contacts and becoming more closely acquainted with the concerns of peer practitioners, attending and participating in meetings is a way of accelerating the learning and training process for staff who are developing knowledge in the field. This is of very practical economic benefit to organisations that might otherwise have to contemplate expensive training and staff development. National and international preservation related conferences, workshops, seminars, symposia, etc., are numerous, occasionally free, and increasingly focused on communicating and delivering practical preservation outcomes.

In addition to standalone or annual events such as iPres (International Conference on the Preservation of Digital Objects), and IDCC (International Digital Curation Conference), funded projects have made an enormous contribution to aligning policy, strategy and practice in the field, not only through dissemination meetings funded as part of project workplans, but also through their associated reports and deliverables. One of the outstanding contributions in this area has been made by the European Commission which has funded a number of major European projects which continue to collectively have a massive impact on preservation. These include the following:

- ERPANET – Electronic resource preservation and access network
- DPE – Digital preservation Europe
- PLANETS – Preservation and long-term access through networked services
- CASPAR – Cultural, artistic and scientific knowledge for preservation, access and retrieval
- KEEP – Keeping emulation environments portable
- PrestoSpace – Preservation towards storage and access. Standardised Practices for Audiovisual Contents in Europe.
- PARSE Insight – Permanent access to the records of science in Europe
- APARSEN – Alliance for Permanent Access to the Records of Science Network

These are all ambitious multi-partner institutional undertakings where many participants from all over Europe (and in some cases beyond) have been given an opportunity to hone or develop their skills in an emerging area. Whilst it isn't training as such, there will almost certainly have been ample requirement for many participants



---

to learn fast ‘on the job’, and this accelerant factor, bringing people up to speed within finite deadlines, is of broad economic benefit.

Similar work is being carried out in the United States under the auspices of the National Digital Information Infrastructure and Preservation Program (NDIIPP) of the Library of Congress.<sup>7</sup> The NDIIPP’s mission is ‘to develop a national strategy to collect, preserve and make available significant digital content, especially information that is created in digital form only, for current and future generations,’ and to that end it has focused on three areas: capturing, preserving, and making available digital content; building a nationwide network of preservation partners; and supporting the development of a technical infrastructure of tools and services, including BagIt, Heritrix, iRODS, JHOVE, and Storage Resource Broker (SRB). Perhaps the NDIIPP’s most important accomplishment has been articulating a convincing case for the importance of long-term digital preservation, one that bears the imprimatur of the closest thing that the U.S. has to a national library. An endorsement by the Library of Congress carries weight for organisations working in related fields and the library has succeeded at least in making the argument that digital preservation ought to be a national priority. This can be seen, for example, on the Web site for the National Digital Stewardship Alliance (NDSA),<sup>8</sup> an outgrowth and extension of the NDIIPP.

In addition to the Library of Congress, the Institute of Museum and Library Services (IMLS)<sup>9</sup>, a federal funding agency, has also supported digital preservation initiatives in the United States, most notably the establishment of the Alabama Digital Preservation Network (ADPNet), a statewide LOCKSS-based network. Aaron Trehub will be discussing this initiative in an individual paper that focuses on achieving economic sustainability.

From a standing start, it is incredibly difficult to persuade library administrators and other senior managers to embrace the requirements of digital preservation and to get it embedded into organisational strategies and thought-processes. Bohdana Stoklasová is preparing an individual paper in the context of this conference that addresses some of the challenges of advocating preservation at these levels. She will argue that the gradual introduction of both effective technology and skilled personnel is a critical requirement but it is not cheap, and it is not easy to accomplish.

Once momentum is achieved though, and with the backing of powerful advocates, a great deal of progress can be made and partnerships can be brokered and usefully exploited. Returning again to North America, the Library of Congress, IMLS, and other funding organizations have supported efforts to define best practices and procedures for digital preservation. They have also supported the development of governance instruments (a crucial but often-overlooked precondition for creating economically sustainable and scalable preservation networks, especially among different kinds of institutions in different states, provinces, and countries), and have actually created functioning preservation networks. For example, the NDIIPP supported the creation of the Data-PASS network, the PeDALS project, and the MetaArchive Cooperative, the first Private LOCKSS Network (PLN) explicitly designed for the preservation of locally created digital content. For its part, the IMLS-

---

<sup>7</sup> <http://www.digitalpreservation.gov/>

<sup>8</sup> <http://www.digitalpreservation.gov/ndsas/>

<sup>9</sup> <http://www.imls.gov/index.shtm>



---

supported Alabama Digital Preservation Network was the first statewide PLN and served as the model for the Council of Prairie and Pacific University Libraries (COPPUL) PLN in western Canada. Indeed, the ADPNet-COPPUL relationship is an example of two self-sustaining DDP networks that are collaborating fruitfully across national borders. As such, they represent a working example of economic alignment and offer proof that it is possible to create affordable and sustainable preservation networks.

In the UK, the Joint Information Systems Committee (JISC) has been influential in funding innovation and building capability through preservation programmes and projects (most often based within UK universities) that have supported a wide range of activity including feasibility and scoping work, technical development, policy and legal studies, and network and partnership support. The Dutch National Library and the National Archives have been an influential force in the Netherlands driving preservation practice there and being influential around the world, as have their UK, Australian, New Zealand, German, and Danish counterparts (in association with those responsible for their core and capital funding). It is significant of course that this partial and arbitrary list exclusively describes publicly funded organisations and this goes some way to underpin the next point of alignment which is around the theme of ‘openness’. It is tempting to think that the natural tendency of all publicly funded organisations would surely be towards the open: i.e. open source (software); open access (content); open standards; and indeed open communities, where participants from all sectors are welcome and encouraged to join in the discussion and add value: but on reflection this is taking an influential core value of a group of mostly large and powerful institutions and extending it across a whole diverse community.

Intuitively, the adoption of ‘open’ approaches, in particular open source software in the context of technical preservation solutions, seems like a tactic designed to appeal to cash-strapped organisations with little by way of resources and funding to engage with the complexities of preservation. But as anyone who works with a range of software will state, ‘open source does not mean free’. Whilst the source code may be accessible and re-usable, there may be a cost for distribution, for support, or for particular licence conditions. Additionally, to actually implement, use and locally maintain the software in a way that is useful for one’s own organisation, there may well be significant costs down the line that are inherent to a community-owned resource. In some cases, it may be valid to argue that such costs would be usefully subsumed within a service-level agreement on a piece of proprietary software from a commercial vendor.

Whilst acknowledging that it may be attractive for some organisations to buy off-the-shelf proprietary solutions with generic interfaces and bundled support, there has nonetheless been great progress with establishing a swathe of open and free tools, toolkits, models, frameworks and guidance that have largely removed many of the financial barriers to engaging with preservation, at least up to a certain level. Resources such as: Archivematica (a comprehensive digital preservation system); California Digital Library Micro-Services; The Curator’s Workbench (UNC Chapel Hill pre-ingest tool); and HOPPLA (Vienna University of Technology) may indicate the path to further progress in this direction. Other tools, namely DRAMBORA (a risk audit tool), AIDA (Assessing Institutional Digital Assets), and DAF (the Data Asset Framework) are now being put together as an integrated suite to tackle sophisticated

---

work in the area of long-term data management practice (see the IDMP project – Integrated Data Management Planning Toolkit and Support based at the DCC). This approach emulates various projects over the years that have built on and integrated various stalwart preservation components such as: DROID and PRONOM produced by The National Archives (UK), JHOVE ( JSTOR and the Harvard University Library), and the NLNZ Metadata Extractor (National Library of New Zealand).

In terms of open standards, probably the most widely referenced and influential standard reference point in Digital Preservation is ISO 14721, better known as the Open Archival Information System (OAIS) reference model.<sup>10</sup> But perhaps where this has really come into its own and made an impact on both disciplinary and economic alignment is in the area of teaching and training. The OAIS model is accommodating as a framework for defining the inputs, processes and outputs of an eligible preservation system and it is this flexibility, combined with a collection of canonical terms and an easily reproduceable graphic which has earned it a ubiquitous role throughout the preservation literature and a place in almost every entry-level presentation ever given about the topic.

The economic impact of the OAIS diagram alone is interesting to contemplate, given that it has served as the backdrop for countless training sessions around the world and has absorbed many thousands of hours of student’s time and attention as it has appeared and been talked through in a variety of training and educational contexts.

### ***Challenges and Gaps***

The alignment accomplishments alluded to above signal that preservation, and more particularly economic positions in relation to preservation, have either purposefully been developed (e.g. cost models, dedicated preservation funding programmes, and institutional policy development), or have realised some oblique benefits from such initiatives as the tendency towards ‘openness’ in many public institutions, and also perhaps the general move towards e-only provision of resources, (a trend that is particularly apparent in the area of scholarly journals).

But despite the many advances there is still a great deal to be done to ensure that we have sustainable economic strategies for preservation. This is especially important precisely because digital preservation can be a financially demanding undertaking whose benefits may not be immediately apparent. A large number of digitised volumes is eye-catching proof of a project’s ‘success’; the substantial investment required to ensure their long-term availability is less likely to generate enthusiasm among decision-makers. As a result, long-term preservation is still not perceived as an indispensable part of digitisation projects and its cost is underestimated or entirely ignored in favour of digitising more materials. Therefore, a powerful argument for digital preservation and evidence of its economic sustainability is the ability to build effectively on previous and current investment.

### **Building on current investment**

Given the wide variety of institutions that should be concerned about digital

---

<sup>10</sup> ISO (International Standards Organisation) 14721:2003; originally proposed by the Consultative Committee for Space Data Systems

---

preservation and the differences among them in culture, policies, laws, regulations, and resource levels, it would be misleading to speak of economic alignment in terms of a single, uniform approach. What may work for one type of institution in a given country would not work at all for the same type of institution in another country. That said, there are general principles that are useful in designing economically sustainable digital preservation networks, and some of them can be discerned from working examples in North America and Europe. One task, therefore, might be to compile an international library of recommended practices that can be modified and applied to different situations, in other words, national resource pages writ large. Experience with the MetaArchive Cooperative, the Alabama Digital Preservation Network (ADPNet), and other LOCKSS-based networks in North America suggests that the following principles contribute to economic sustainability and can therefore promote economic alignment among otherwise very different networks:

1. Whenever possible, use open-source solutions (e.g. LOCKSS). Not necessarily because they cost less, although generally they do, but because they can be managed locally (an important consideration if one believes that cultural heritage organizations should retain control of the content they want to preserve, including access to it) and reduce dependence on third-party solutions.

2. Whenever possible, take advantage of existing administrative infrastructure. There is a corollary here: whenever possible, avoid creating new administrative infrastructure. ADPNet is part of the Network of Alabama Academic Libraries (NAAL), an existing state agency. The COPPUL PLN is part of the Council of Prairie and Pacific University Libraries (COPPUL), an existing consortium of academic libraries in western Canada. For various reasons the MetaArchive Cooperative chose to create a new administrative entity (the Educopia Institute in Atlanta, Georgia) to manage that network, but that decision was necessitated by the network's geographic dispersion across a number of states and the absence of a satisfactory existing administrative home. In the event, this arrangement has not impeded the network's growth.

3. Aim for a lightweight administrative structure. Administering a preservation network costs time and money. ADPNet and the COPPUL PLN each have two committees: a steering committee for policy and a technical committee for hardware and software issues. ADPNet has monthly conference calls; the COPPUL group meets via Skype every other week. The idea is to make digital preservation a routine, low-maintenance, and integral part of an institution's information-management activities.

4. Finally, delegate as much decision-making power as possible to the individual member institutions. They know their digital collections best, and are best able to set preservation priorities.

Whichever approach one chooses, it is advisable to keep it as simple and cheap as possible. Simplicity contributes to economic sustainability; complexity undermines it. This maxim rings true across a whole spectrum of activity as there is a great deal of anecdotal evidence to suggest that preservation is a hard sell because of the perceived complexity of its processes. It is true that without recourse to technical effort and knowledge a non-specialist will struggle to gracefully embed current preservation tools into a local infrastructure, let alone be able to wrestle with the complexities of

---

developing an emulated environment for legacy software to run in. But these issues are a distraction from the fact that all the principle preservation issues, certainly at a managerial level, are almost exclusively non-technical.

What is required is clear and attractive advocacy material that focuses on the issue of what decisions are required *to effectively deal with content*. At some stage, someone in the institution will have to take responsibility for technology choices but those decisions will be inordinately easier, and will be taken more effectively and probably more economically, if they are working from a clear specification, with clear policy guidelines, and in the context of a considered, responsible and joined-up set of information policies.

Many organisations are conservative and suffer from inertia in the face of emerging and disruptive technologies, but memory institutions are notoriously prone to both of these responses. In a risk-averse atmosphere, trust is a valuable commodity and audit and certification of preservation environments and processes can be helpful as assurance for organisations to make investments they would otherwise be nervous of making. The development of standards is a form of assurance and as mentioned above, the OAIS model sits alongside other ISO entities (such as ISO 15489:2001 for records management) to offer a useful formal framework to build on. But although there are emerging models such as: the DINI (Deutsche Initiative für Netzwerkinformation = German Initiative for Networked Information) framework and DINI-Certificate; the TRAC (Trustworthy Repositories Audit and Certification): Criteria and Checklist; the DANS (Data Archiving and Networked Services) Data Seal of Approval; and the DRAMBORA (Digital Repositories Audit Method Based on Risk Assessment) audit tool<sup>11</sup>, the community is waiting for an ISO (or equivalent) approved process for preservation certification that is both effective and affordable, with the audit component delivered by a trusted and sustainable agency. There are indications, however, from the Alliance for Permanent Access that something is on the way,<sup>12</sup> and if this is the case then it will represent significant progress.

Reducing complexity and streamlining preservation functions down to the level and scale that is appropriate for the task in hand is a general requirement and this applies also to metadata. The PREMIS metadata dictionary is a comprehensive and exhaustive catalogue of nearly all of the fields that one might need for the purposes of preservation and is one of the standard works of reference in the field. Combined with various work that examined the significant properties of information (e.g. the CEDARS and INSPECT projects in the United Kingdom<sup>13</sup>), and work on representation information (carried out in the context of the PLANETS and CASPAR projects amongst others), there has been a great deal of progress made with understanding what technical, content and administrative data may be required to effectively describe digital material for long-term preservation purposes. But a gap remains between understanding the ideal metadata requirements for digital objects and choosing to implement that metadata in real-world situations. That gap is to do with time and resource and is therefore another economic issue.

---

<sup>11</sup> <http://www.dini.de/>; <http://www.dini.de/dini-zertifikat/>;  
[http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf); <http://www.dans.knaw.nl/>;  
<http://www.repositoryaudit.eu/>

<sup>12</sup> <http://www.alliancepermanentaccess.org/archives/479>

<sup>13</sup> <http://www.ukoln.ac.uk/services/elib/projects/cedars/>; <http://www.significantproperties.org.uk/>

Metadata is currently laborious to comprehensively and effectively assign to digital objects in a manner that is likely to satisfy all of their potential future use scenarios. Manual tasks, or even semi-automated tasks, of this nature will not scale up to the level that most organisations require. Whilst metadata extractors such as DROID, JHOVE and the NLNZ Metadata Extractor can harvest useful information, work is still required to work out ways of either automatically extracting or intelligently tagging objects such that they align with institutional policies around value and retention. Automated ways of managing digital objects require machine-readable protocols which in turn require reliably and persistently identified components. There are different views on the best identifier sets for all sorts of purposes, including file formats, subject classification terms, organisational identity, researcher identity, and so on and so forth, but the overarching issue once again is one of trust—which in turn often depends on prospects for sustainability, which ultimately leads back to economics.

### **Learning from failed initiatives**

It is important to build on success in designing economically sustainable digital preservation programmes. It is equally important to learn from unsuccessful initiatives. For example, the Australian PADI project was a digital preservation subject gateway set up and maintained by the National Library of Australia (NLA) from the mid-1990s until late 2010. The project was discontinued in that year, primarily because of business decisions about resourcing. ‘Subject interest, expertise and enthusiasm are vitally important but not sufficient,’ one of the project participants later observed. ‘Ongoing sustainability of a service like PADI over a long period probably also requires some dedicated discretionary budget funds, not just a few dedicated individuals. It also requires some available expertise in the means of communication, not just the content.’ Another important element contributing to sustainability is sharing the ‘ownership’ of a programme among a number of institutions and building community engagement in it, even at the expense of managerial efficiency. Again, the fate of the PADI project illustrates the dangers of concentrating ownership in one institution: ‘The other significant development that came with, and contributed to the growing success of PADI as an information gateway, was a local decision against collaboration, taking control of PADI away from a diverse committee of organisations, and investing it in one institution.[...] A case of making it much more easy to manage, but closing off local commitment to its survival and usefulness.’<sup>14</sup> This lesson has been taken to heart by the Private LOCKSS Networks in North America, whose governance policies were designed to ensure that management of the networks rotates among the participating institutions.

### **Services and business models**

As stated in various ways previously, preservation is not necessarily an activity that easily lends itself to being configured for delivery as a business practice or commercial enterprise. One of the conclusions of the influential final report of the Blue Ribbon Task Force on the Sustainability of Digital Preservation and Access (BRTF)<sup>15</sup> stated that devising strategy for preserving digital materials was made difficult by four

<sup>14</sup> Personal communication between Maurizio Lunghi and Colin Webb, January 2011.

<sup>15</sup> <http://brtf.sdsc.edu/>



---

inherent factors:

- i. Long time horizons
- ii. Diffused stakeholders
- iii. Misaligned or weak incentives
- iv. Lack of clarity about roles and responsibilities among stakeholders

And this concisely explains why the demand for preservation services is relatively weak, and consequently why the list of commercial vendors queuing up to supply those services is fairly short. Tessella have had success, principally (in this area) with their Safety Deposit Box (SDB) system, which is in use in major national archives around the world and has recently been implemented by the Church of Latter Day Saints to deal with their not inconsiderable ingest challenge for the Family Search archive. OCLC launched their Digital Archive Service in 2008 and have been marketing it to state libraries and archives, especially those that are already using CONTENTdm, another OCLC product, to manage their digital collections. Ex Libris have a digital preservation product called Rosetta and are building up their customer base. They are pursuing an interesting collaboration model with the National Library of New Zealand who are taking the view that by working with a commercial vendor, this offers the best chance for creating and sustaining some of the core services that a preservation system will require, including a file format registry that will sit at the heart of the product and supply an identification function.

It is clear that handled in the right way and if setup as a mutually beneficial partnership, relationships between vendors and public-sector bodies can bring enormous benefits to client organisations in terms of economic efficiencies and clarity of business processes. There is a strong argument for saying that organisations should play to their strengths. Taking a slightly different approach, it is possible to engage with technology but only on terms that are advantageous to one's own organisation. In banking, telecommunications, health-care and most other sectors of society, organisations set out their principles and mission; and then establish their rights, values and basic rules. Then they define the components, functionalities and workflow; and the models and terms of specific services; and then invite competitive tenders to bid for aspects of the work. An example from Italy is the Magazzini Digitali ('Digital Stacks') project, in which the Ministry of Culture has set up the global architecture and functions of a trusted digital depository (complete with ingest rules and selection criteria for long term preservation) and has the put out a call to tender to private companies.<sup>16</sup>

At Auburn University, they have given clear specifications to external vendors who have then carried out the actual digitisation of the objects, but when it comes to the digital preservation function, they stop short of 'entrusting such a crucial part of our mission to an external vendor'. They take the view that the primary responsibility for ensuring the long-term preservation of the human record in digital form ought to rest with public institutions or alliances of public institutions. That view is shared by other research universities in the United States. That said, acknowledge that there may be room to explore the topic with some major commercial players in the digitisation field (e.g. Google). Conversations of this nature being undertaken by the HathiTrust

---

<sup>16</sup> <http://www.rinascimento-digitale.it/projects-digitalstacks.phtml>

and the Digital Public Library of America will be worth monitoring.

In common with the broadly accepted view that preservation is an international concern and should be tackled using broadly collaborative working methods, preservation is also increasingly being viewed as a process and a workflow that need not be dealt with by an end-to-end local process. The cost efficiencies and the accelerated development processes that accompany collaborative working can enhance the preservation workflow and can relieve institutions of training and technical overheads that they may not be equipped to meet.

Disaggregated services for preservation were much in vogue several years ago (service-oriented architectures), but the focus has now moved onto the potential for cloud services to offer preservation and curation capacity using elastic storage and computing provision. ‘Trust’ remains an issue for organisations contemplating cloud services and whilst one could imagine most services, e.g. replication, hashing, identification, characterisation, validation, ingest, migration, verification, authentication, etc., being offered as some form of service, these would need to be underpinned by the type of trusted certification processes mentioned previously (e.g. TRAC, DINI-Certificate).

Fig. 1 (see Appendix) offers one possible representation of a method for dividing preservation workflow components between a content owner and one or a number of service providers. The light boxes represent actions that have to necessarily devolve to the content owner and the black boxes are those that it might be possible (in some circumstances) to ask an external party to be responsible for. The boxes with graded colour from black to grey represent actions which would probably require close collaboration between content and service provider. These actions may involve matters of judgement, relating to ongoing policy decisions, or changeable degrees of openness and security depending on the nature of the content. This micro-services approach has become widely referenced over the last few years and there are interesting economic considerations associated with it. There is doubtless considerable balancing to be done between the complexity of keeping track of multiple services and potential service providers, against the cost efficiencies and/or flexibility that this approach could potentially offer an organisation.

When faced with hard economic choices about service provision, organisations may inevitably run through a fairly universal set of questions:

- Is this something that we really need?
- How much will it cost?
- How much money have we got?
- How much of what we’ve got do we want to spend on this?
- Can we get someone else to pay for it?

And in the unlucky event that the answers to all of those questions is somewhat unsatisfactory, the final question becomes:

- How can we adapt what we already have to do what we need to do?

This is a somewhat long-winded way of illustrating that most organisations are

generally forced to make very pragmatic decisions, but in terms of gaps and challenges, it follows that the clearer the arguments are for the value of digital materials, the easier it will be to win the argument about funding. This is true irrespective of whether the chosen solution is an entirely outsourced approach (let's pay someone else to do this for us), or an entirely self-managed one (let's do this ourselves, or with a group of like-minded institutions). In either and all cases, the goal should be the same: to codify long-term digital preservation in institutional (or consortial, or national) policy, and incorporate it into an institution's regular workflow.

If the ideal is to embed preservation practice into the core institutional function so effectively that the cost of it simply disappears, then there is also a pressing need to acknowledge and understand all of the steps short of that ideal. Practically speaking, all organisations (except those for whom preservation *is* the core mission) are probably going to find themselves somewhere down the rungs of that particular ladder for the immediate future. There remain large challenges and gaps in both defining the business case and the business models for preservation but interesting work has emerged in recent years to start classifying and examining possible options. The BRTF report lists five 'common funding models for digital preservation'.<sup>17</sup> Ithaka S+R have also done some very useful work to produce case-studies of sustaining digital resources.<sup>18</sup>

### **Roles and responsibilities**

When considering what we might refer to as the preservation interrogatives: the 'who', 'what', 'where', 'when' and 'how' of digital preservation (see Appendix, Fig. 2), the question 'who' is probably the most interesting (and sometimes the most intractable) question for many people, focusing as it does on the human aspect and drilling down into the detail of who is actually responsible for preserving material.

There is currently uncertainty within many institutions about who ought to take responsibility for the long-term stewardship of digital content and this is also reflected at the macro-level where funding bodies, government agencies, institutions and individuals are looking from one to the other trying to work out their moral and financial responsibilities vis-à-vis content that is of interest to them.

In terms of building *capability* to preserve, this could be characterised as a problem that funders are interested in. Organisations such as JISC, the Library of Congress, IMLS, Deutsche Forschungsgemeinschaft (DFG), SURF, and various other agencies that support research and innovation have a vested interest in ensuring that the communities that they support have the tools and frameworks and infrastructure that they need to manage the information that they produce in the course of their largely education-related activities.

When focusing on the *capacity* to preserve, this is arguably more likely to devolve to institutions and organisations whose responsibility it is to deal with the logistics of having staff in place with the right skills to do the work that the institution requires of them.

<sup>17</sup> BRTF, p. 45

<sup>18</sup> Maron, Nancy L., Kirby Smith, K., & Loy, Matthew. (2009). Sustaining Digital Resources: An On-the-Ground View of Projects Today, Ithaka Case Studies in Sustainability, Strategic Content Alliance. Available at <http://www.ithaka.org/ithaka-s-r/research/ithaka-case-studies-in-sustainability>



---

When it comes to the *sustainability* of both of the above, then this is where the gap or challenge becomes identifiable. A funder may be able to commission the creation of a useful tool or resource but ensuring that the transferral of that capability into the institution actually happens is an uncertain proposition, particularly given that short-term ‘soft’ funding may often result in the loss of staff at the end of a project, (thereby also affecting the organisation’s *capacity* to preserve). But these are not extraordinary problems. Staff come and go all the time. Perhaps the answer to these problems lies at the sectoral level and needs to be addressed in terms of a culture-shift towards a more realistic evaluation of our collective requirements from archived digital culture.

As already stated above, preservation issues for the majority of people revolve around non-technical issues and when focusing on roles and responsibilities in this domain, the discussion at some point needs to drill down to a granular level, and ultimately requires someone to take some form of position on the nature and the value of the content in question. In any discussion of the economics of preservation, ‘value’ is an interesting word: different from ‘cost’; and not as practical as ‘benefit’. But if we can establish who regards the content as valuable, then we may arrive at a better understanding of who the potential beneficiaries of the preservation process are. We may then be able to find out if anyone is likely to benefit from that preserved content without contributing to the cost of its preservation, which is of course a ubiquitous scenario in a digital world where instant global access to a dazzling universe of material has become not only common but expected. This is what the BRTF report (and the language of economics) calls the ‘free rider’ problem.

In some contexts, universal permanent access is not only a convenient by-product of digitised material finding its way into an open preservation environment, but is the intended and funded outcome. Legal deposit arrangements in various countries are the logical conclusion to the information as ‘public goods’ arrangement, where taxes pay for comprehensive tranches of material to be made available in perpetuity (sometimes under particular access conditions) by trusted public repositories. But in many other contexts and for the vast majority of institutions and organisations, this is an irrelevance. These bodies have budgets to balance and priorities to define and are very conscious of the opportunity costs of assigning precious resources to an enterprise as currently ill-defined as long-term preservation. At some point, the question will be asked, ‘who is going to pay for this?’ Should the creator pay? Should the user pay? Should responsibility fall to the institution? Or is it a public problem?

Perhaps one way to examine this problem is to take a step back and look at the creation or acquisition process and work through the decisions that are involved at the instigation of this whole process. In some instances, the case for acquiring a digital file is straightforward. Where the original object is unique or at risk, there is a clear justification for creating a surrogate and this also indicates ownership and interest in the digital file. As a surrogate, the physical object and its digital manifestation are related. In cases where a physical object needs to be copiously used by a great variety of people, there is also a clear justification for digitisation, although given that the original is probably sturdy and common, the subsequent stewardship issues begin to get murky when questions are asked about the point of storing something that can be easily accessed in a number of other ways.

The following represent four selection criteria elements that might help inform policy-making:

- Are we allowed to preserve it? (who owns it?)
- Is there someone (right now) that wants to use it?
- Can we carry on making it accessible? (will it be technically possible?)
- How interesting is the information? (will someone want it in the future?)

As stated before, selection is an absolutely key part of effective preservation practice, particularly as we exist in a period where analogue material is likely to be with us for some time to come whilst the amount of new digital material requiring storage grows all the time.

It may be possible for some organisations to settle on fairly loose or general policies towards responsibility for material, such as forming the view that any decision to ingest material into a given preservation environment implies the acceptance of responsibility, and therefore the acceptance of ongoing cost. Other general statements of this nature may be applicable also, but there is a potential problem with this approach in that the stewardship of digital material and collections is not a static and tidy problem. As digital objects progress through a lifecycle, their value—like any investment—may rise and fall. Perhaps what is needed is some low-overhead administrative (or even just conceptual) way to keep track of three vital pieces of information that will assist content owners with the ongoing challenge of appraisal, which can be defined as the iterative selection process that ideally takes place at various points subsequent to the initial selection decision.

The role of *creator* of the digital object/collection/dataset is fairly clear and should often be reflected in the metadata associated with an object, or will be known to those managing the environment that the object is destined to be stored in. This is often a key piece of information for a great variety of reasons but may also be important for appraisal purposes. What is less obvious, and not by any means likely to be the same as the *creator* is the identity of a person who might be referred to as the *principal keeper*. This would refer to someone who has appropriate authority and is interested in knowing that the object(s) in question are supposed to be residing in the preservation environment. The third piece of information that might be useful to know is who the *principal user* is? This would refer to someone who had self-identified themselves as a person who was interested in the object(s) in question and who had a vested interest in seeing that they continued to be stored safely.

In many environments, one suspects that these designations would not make much sense as two, or perhaps all three, of the designations would be the same person. But in other cases—particularly perhaps where special collections of digital material were stored for long periods of time (at some expense) and the original motivations for archiving the material had become unclear—designations of this type may be helpful in determining ongoing value.

In order for this proposal to be valuable, refinements would need to be introduced whereby the identity of the keeper or the user would be passed on as appropriate to new incumbents or to others taking on the mantle of research or teaching in that area

---

(if that was the use case). An action would be triggered however if at any point the keeper or the user identity became blank—that is, if one or other of those roles became vacant in relation to an object. This would alert the host organisation to the fact that either somebody thought that the object was no longer worth storing, or that the file was no longer worth using, either of which represents a strong case for disposal.

A number of other refinements (e.g. designated community alerts and automatic retention periods) could be introduced as safeguards but the point would be to try and tackle the problem of unmanaged persistence.

### **What Will Success Look Like? A Five-Year Forecast**

The members of the Economic Alignment panel provided somewhat different answers to this question, the last of the questions posed to the six panels. There were several common themes and areas of agreement, however.

There was general agreement that one measure of success is that analogue and digital documents are considered and treated equally in any preservation regime. It was pointed out that the current practice of acquiring, cataloguing, protecting and making available predominantly or even only analogue materials while postponing similar treatment for digital content entails possibly irrecoverable losses to the corpus of cultural heritage materials and important research resources.

Of course, given limited resources, selection and prioritization will have to be applied to both types of resources. This will require fundamental changes in strategic planning and organisation at many institutions. *Cultural memory institutions are by nature conservative, and transforming them will be far from easy. However, the panel members agreed that these changes will be necessary to achieve success in this field.* When normal practice within an organisation automatically factors in the whole life-cycle costs of acquiring or creating a digital collection (including the opportunity costs) and the institution has a clear view of the likely short, medium and long-term benefits of doing so, then it might be possible to claim that the role of digital preservation is as innately understood within an organisation as (analogue) archival practice or records management. Fortunately, as the body of this paper shows, there is a growing number of successful transformations underway. Taken together, they suggest that momentum is building in the right direction.

In that connection, there was also agreement that success in this area will begin with an institutional recognition that long-term digital preservation is a high-priority activity that requires an ongoing commitment of time and resources. This will involve having policies which are broadly meaningful across institutions and model governance instruments that can be adapted to reflect local conditions and practices. It will mean that staff are trained in basic preservation competencies, ranging from digitization best practices and optimal file organization to writing instructions for digital preservation software (e.g. LOCKSS manifests and plugins). It will mean that digital preservation is embedded into the institutional way of behaving and operating (i.e. linked to policy and workflow measures) and embodied in an optimum balanced budget from the start.

Finally, it will include an internationally agreed-upon glossary of terms and

principles, perhaps as part of an internationally supported online resource (or competence) centre.

In short, we will have achieved success when long-term digital preservation becomes a routine and economically sustainable activity and a generally accepted part of the mission of cultural memory organizations—that is, when most institutions have incorporated the long-term stewardship of digital materials into their day-to-day operations, preferably with some degree of mutual assistance and coordination. This may happen as a result of national policy and government mandates, or because of a series of local initiatives. The main thing is that it happens—and in a sustainable way, with long-term institutional commitment, public understanding and support, budget lines, and dedicated personnel.

To that end, the panel offers the following recommendations and guiding principles:

- Digital preservation should be an integral part of all of projects dealing with the digitisation of analogue documents and/or the acquisition of born-digital documents having to do with the national cultural heritage.
- Digital preservation is not a luxury. Ensuring adequate protection for digital content should be just as much a part of an institution’s workflow as protecting analogue materials from water, fire or careless handling.
- More broadly, digital preservation should also figure in national public policy. The recipients of public funding (libraries, museums, archives) should be required to include digital preservation in their activities, build and share a knowledge base, and pool resources to develop or add to preservation tools and services. The recent requirement by the National Science Foundation (NSF) in the United States that recipients of NSF funding engage in long-term research data management is just one example of this.
- Sufficient funding should be dedicated exclusively to digital preservation. Large-scale publicly funded digitization initiatives that do not also include a budget and a clearly defined strategy for digital preservation are disasters waiting to happen and an unwise use of public monies.

As was pointed out at the beginning of this paper, digital preservation is a relatively new area of activity for most cultural memory organizations. It is all the more important, therefore, to share experiences, tools, and successful approaches across institutions and countries. It is hoped that the ‘Aligning National Approches to Digital Preservation’ conference will do precisely that.

## Conclusion

In a 2004 article whose title was inspired by American poet Wallace Stevens’ ‘Thirteen Ways of Looking at a Blackbird’, Brian Lavoie and Lorcan Dempsey recognized that digital preservation is ‘an economic process, in the sense of matching limited means with ambitious objectives.’<sup>19</sup> They were right on both counts: the means are limited and the objectives are indeed ambitious. As this paper shows, however, an impressive—one might even say ‘ambitious’—amount of work has already been done

<sup>19</sup> Lavoie, Brian and Lorcan Dempsey. 2004. “Thirteen Ways of Looking At: Digital Preservation”. *D-Lib Magazine* 10 (7/8). [doi:10.1045/july2004-lavoie](https://doi.org/10.1045/july2004-lavoie)

---

in Europe, North America and elsewhere on identifying the costs of digital preservation and devising tools, techniques, and procedures for absorbing those costs into ongoing preservation programmes. Moreover, this work has been accomplished in large part by realizing economies through collaboration among institutions. Despite their different origins, missions, and management structures, the preservation initiatives identified in the body of this paper—the Digital Stacks project in Italy, the MetaArchive Cooperative and the Alabama Digital Preservation Network in the United States, the COPPUL PLN in Canada, the Digital Curation Centre and the Open Planets Foundation in the United Kingdom, and so forth—prove that it is possible to take advantage of accumulated experience and community effort to build working, economically sustainable digital preservation networks across states, provinces, and even countries. In Lavoie’s and Dempsey’s words, digital preservation ‘is an ongoing, long-term commitment, often shared, and cooperatively met, by many stakeholders.’ The task facing us now is to continue and extend the collaborative work that has been done.

---

## Acknowledgements

Body Text Indent 3 In consectetur, diam at volutpat elementum, libero lectus pulvinar sem, vel sollicitudin quam libero ac felis. Aenean in augue in leo fermentum accumsan.

### References [use Style: "Reference-list-item"]

[book][please leave the categories before your refs]Borgman, C.L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, MA: MIT Press.

[proceedings]Borgman, C.L., Wallis, J.C., & Enyedy, N. (2006). Building digital libraries for scientific data: An exploratory study of data practices in habitat ecology. *10th European Conference on Digital Libraries. Alicante, Spain*: Berlin: Springer.

[journal article]Currall, J., & McKinney, P. (2006). Investing in Value: A Perspective on Digital Preservation. *D-Lib Magazine* 12 (4). doi:10.1045/april2006-mckinney

[report]Digital Curation Centre. (2005). *Digital curation and preservation: Defining the research agenda for the next decade*. In Report of the Warwick Workshop, November 7-8, 2005. Digital Curation Centre: Warwick, UK. Retrieved July 27, 2007, from [http://www.dcc.ac.uk/events/warwick\\_2005/Warwick\\_Workshop\\_report.pdf](http://www.dcc.ac.uk/events/warwick_2005/Warwick_Workshop_report.pdf)

[proceedings]Esanu, J., Davidson, J., Ross, S., & Anderson, W. (2004). Selection, appraisal, and retention of digital scientific data: Highlights of an ERANET/CODATA workshop. In *Data Science Journal, 2004, 3*. Retrieved July 30, 2007, from <http://www.jstage.jst.go.jp/article/dsj/3/0/227/pdf>

[standard]ISO/IEC Guide 73. (2002). *Risk management. Vocabulary. Guidelines for use in standards*.

[journal article]Lavoie, B., and & Dempsey, L. (2004). Thirteen ways of looking at: digital preservation. *D-Lib Magazine* 10 (7/8). doi:10.1045/july2004-lavoie

[report]Maron, N.L., Kirby Smith, K., & Loy, M. (2009). Sustaining Digital Resources: An On-the-Ground View of Projects Today, Ithaka Case Studies in Sustainability, Strategic Content Alliance. Retrieved April 28, 2010, from <http://www.ithaka.org/ithaka-s-r/research/ithaka-case-studies-in-sustainability>

[book]Mayer-Schönberger, V. (2009), *Delete: The Virtue of Forgetting in the Digital Age*. Princeton, NJ: Princeton University Press.

[journal article]Mayernik, M.S., Wallis, J.C., & Borgman, C.L. (In press). Adding context to content: The CENS Deployment Center. *Journal of the American Society for Information Science & Technology*. Milwaukee, WI: Information



---

Today.

[personal communication] Moore, R., (personal communication, 2008)

[report]National Science Foundation. (2003). *Report of the National Science Foundation Blue-Ribbon Advisory Panel on cyberinfrastructure*. Retrieved September 30, 2006, from [http://www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=cise051203](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=cise051203)

[Internet journal]Payette, S., Staples, T., & Wayland, R. (2003, April). The Fedora Project: An Open-source digital object repository management system. *D-Lib Magazine* 9,(4). Retrieved September 30, 2006, from <http://www.dlib.org/dlib/april03/staples/04staples.html>

[blog post]Rusbridge, C. (2008, July, 29) Re: “Digital preservation” term considered harmful? [Web log message from *Digital Curation Blog*]. Retrieved September 7, 2009, from <http://digitalcuration.blogspot.com/2008/07/digital-preservation-term-considered.html>

[proceedings]Santini, M. (2004a). A shallow approach to syntactic feature extraction for genre classification. *Proceedings of the 7th Annual Colloquium of the UK Special Interest Group for Computational Linguistics*.

[report]Santini, M. (2004b). *State-of-the-art on automatic genre identification*. (Technical Report ITRI-04-03). University of Brighton, UK, Information Technology Research Institute (ITRI).

[book]Skinner, K., & Schultz, M., eds. (2010). *A Guide to Distributed Digital Preservation*. Atlanta, GA: Educopia Institute. Retrieved April 28, 2011, from <http://www.metaarchive.org/GDDP>

[Internet journal]Smith, M., Barton, M., Bass, M., Branschofsky, M., McClellan, G., Stuve, D., et al. (2003, January). DSpace: An open source dynamic digital repository. *D-Lib Magazine* 9,(1). Retrieved September 30, 2006, from <http://www.dlib.org/dlib/january03/smith/01smith.html>

[book]Witten, I. H., & Frank, E. (2005). *Data mining: Practical machine learning tools and techniques*. (2nd ed.). San Francisco, CA: Morgan Kaufmann.

## Appendix: Figures

Fig.1. Exemplar preservation service framework

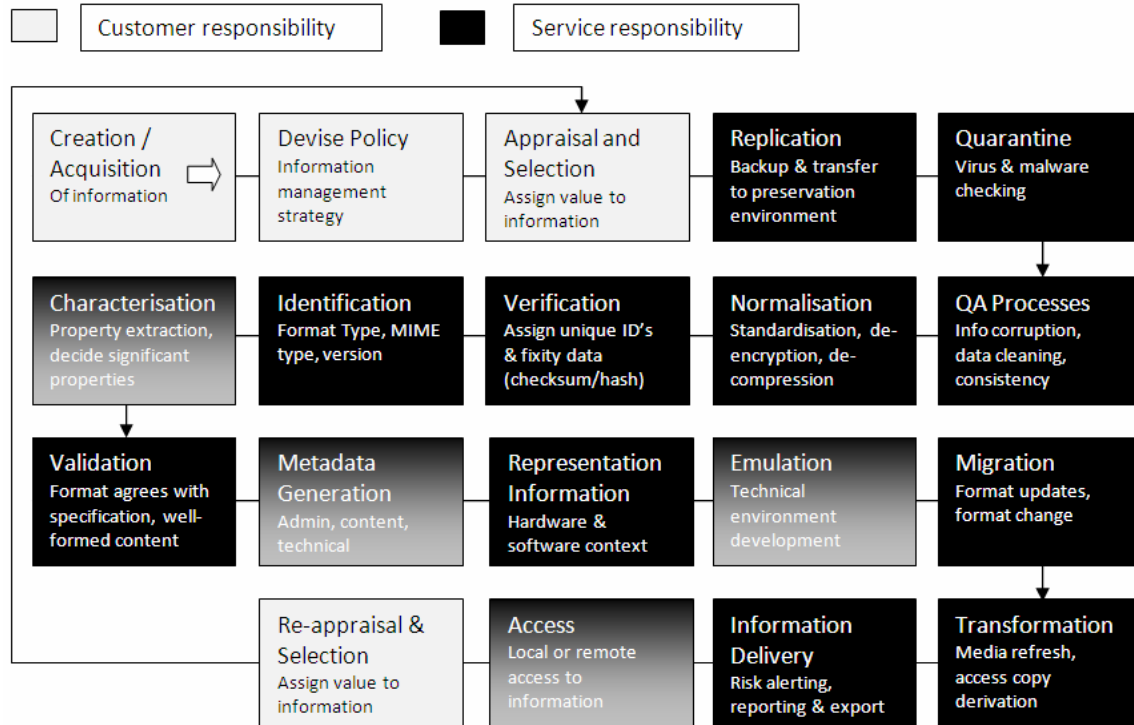


Fig. 2. Preservation Interrogatives

Preservation interrogatives	
<b>Who</b>	Identify the key players involved with long-term preservation of the targeted content.
<b>What</b>	Select, describe, and characterize the collection and content.
<b>Where</b>	Decide the locations of all the copies of the content and metadata.
<b>When</b>	Decide the targeted preservation timeframe and impact of loss.
<b>How</b>	Decide how the key content management and preservation tasks will occur.



# The International Journal of Digital Curation

Issue x, Volume x | 2xxx

## Czech National Digital Library: The Economic, Strategic and International Aspects of Digital Preservation

<Jan> <Hutar>,

<Head of Digital Preservation Department>,

<National Library of the Czech Republic>

<Marek> <Melichar>,

<Digital Preservation Department>,

<National Library of the Czech Republic>

<Bohdana> <Stoklasova>,

<Director of Library Collections and Services>,

<National Library of the Czech Republic>

### Abstract

The article proceeds from the experience with digital preservation gained by the authors over several years in the National Library of the CR (NL CR) as well as through participation in international projects. The first part of the article deals with the development of digital preservation at the NL CR in connection with the digitisation of analogue documents and archiving of born-digital documents. After a short description of the main projects related to digital preservation, the primary accent will be placed on the strategic-economic and international aspects of digital preservation in a memory institution financed by the state. At a time of budget cuts, it is not easy for the new area of digital preservation, for many hard to understand, to reach one of the foremost places in the list of strategic priorities and attain the necessary financial and personnel provisions. In the area of digital preservation, the NL CR has in a relatively short period of time moved from nothing to a very good level chiefly thanks to the utilisation of the results of international projects and the best practices of foreign national libraries and other institutions. In the article, the most useful projects, products and practical experience acquired during visits of workplaces abroad are evaluated, categorised and described in the form of recommendations for digital preservation beginners.

---

## Introduction

This article describes how the National Library of the CR (NL CR) began its digital preservation efforts. As a memory institution, the NL CR dealt with the preservation of paper documents for several centuries but turning its attention to the preservation of digital documents has not been easy. We believe that a number of national libraries or other memory institutions have had or will soon have similar experience. In the following pages, we will reflect upon the fundamental effects of digital preservation's introduction on various library processes, on financing and on staffing.

### Digitisation, Harvesting and Born-Digital Documents Acquisition at the NL CR

The NL CR started to learn about the digital preservation rather long after the institutions' infrastructure housed many terabytes of digital data. The data were generated mainly in the digitisation of paper documents. The NL CR digitised all of its card catalogues (ca 5,000,000 cards) in the middle of 1990s. Later, the digitisation moved to the library documents – both historical and modern. The routine digitisation of manuscripts began in 1996 and the digitisation of endangered newspapers (at the beginning based on microfilms) began in 2000. In the same year, the NL CR started to harvest and archive the Czech internet. The NL CR created three digital libraries – Manuscriptorium<sup>20</sup> (historical collections), Kramerius<sup>21</sup> (modern books and periodicals) and WebArchiv<sup>22</sup> (archiving the Czech web) and built a unique national access point, the Union Information Gateway<sup>23</sup>.

However, in recent years, the budget available for digitisation has continually declined. The solution was found in two projects: The National Digital Library project financed mainly from the EU Structural Funds, which will digitise and store primarily modern documents (issued after 1800) and the Google Books project, which will focus on digitisation of early prints. These projects will result in ca 500,000 volumes or 100,000,000 pages of digitised analogue documents by 2020.

In 2011, the NL CR also launched a pilot project aimed at the acquisition and processing of born-digital documents and of the digital preprints of printed documents. The pilot programme relies on voluntary cooperation with the publishers; the obligation to deposit electronic publications should be anchored in new legislation.

As the amounts of data in the digital repository of the NL CR have been growing, its administration has required ever more attention and finances. The library management decided first to stabilise the HW infrastructure of the repository, which improved at least the bitstream level preservation of the data. More sophisticated logical preservation, the data and metadata management should be fully serviced by

---

<sup>20</sup> <http://www.manuscriptorium.com>

<sup>21</sup> [http://kramerius-info.nkp.cz/welcome/view?set\\_language=en](http://kramerius-info.nkp.cz/welcome/view?set_language=en)

<sup>22</sup> <http://en.webarchiv.cz>

<sup>23</sup> [http://info.jib.cz/welcome-to-uniform-information-gateway-uir?set\\_language=en](http://info.jib.cz/welcome-to-uniform-information-gateway-uir?set_language=en)

---

the Long-term preservation system solution, which the NL CR should build in the National Digital Library project.

### **National Digital Library Project**

In February 2010, the NL CR applied along with the Moravian Library (ML) as its partner for the ‘Creation of the National Digital Library’ (NDL, Czech initials: NDK) project<sup>24</sup>. The project was submitted within the Call 07 of the Integrated Operational Programme ‘Electronisation of Public Administration’. In June 2010, the project was approved. It is one of the corner stones of the eCulture concept, through which the sector of culture significantly contributes to the fulfilment of the aims of Smart Administration.

The budget of the project is ca EUR 12 million, of which 85% comes from the ERDF structural fund and 15% from the Czech Republic state budget.

The NL CR and ML are deposit libraries for more than 200 years. In their collections is the majority documents published in the CR (*Bohemica* in the narrow sense of the word), a great number of documents related to the CR published abroad (*Bohemica* in the broad sense of the word) and abundant historical collections. The libraries cherish an extensive and unique material of singular cultural and factual value.

The NDL project has three main aims:

1. The digitisation of a significant part of the *Bohemica* of the 19th–21st centuries, i.e. books issued in the Czech Republic, written in Czech or discussing the Czech Republic. By the end of 2019, we will have digitised in total more than 50 million pages, approximately 300,000 volumes. The digitisation will continue beyond the scope of the project (2014) and also beyond the mandatory sustainability of the project.

2. Building a reliable digital repository for the long-term preservation of digital documents. The system will provide an environment for the management and preservation of documents digitised in previous years and ingest also the digital documents created during the NDL project.

3. Provide a single point of access to all these digital documents, in user-friendly interface with advanced personalisation options. The system should overlay digitised documents, online scholarly journal databases and all other information resources.

## **Strategic and Economic Aspects of Digital Preservation**

---

<sup>24</sup> [http://www.ndk.cz/narodni-dk/podrobnejsi-popis-projektu/ndk-context/view?set\\_language=en](http://www.ndk.cz/narodni-dk/podrobnejsi-popis-projektu/ndk-context/view?set_language=en)

---

From the brief description above, it is clear that the NDL project is rather complex. What is more, in the course of the project, the NL CR has to cope with two even larger projects: the reconstructions of the historical building of the Klementinum in the centre of Prague and the construction of a new building on Prague's periphery, which should house major part of the technology and staff of the NDL project. The NDL project will affect many activities, and sometimes will require profound transformations of the processes.

The context and expectations of the stakeholders are naturally changing too. The users expect more off-site services, and are not ready to bridge the traditional library barriers. The users expect single place of access, Google type indexes. The NL CR has to fulfil also coordination functions – for the digitisation project, the NL CR has to publish standards and requirements on the quality of the metadata. During the NDL project, also other institutions will start new larger digitisation projects financed in the regions.

In the following part of the text, we will explain how the NDL project affected and will yet affect the NL CR's budgeting, staff decisions, and also organisation issues. This will set our digital preservation efforts into appropriate context.

### ***Financing and staffing***

The funds from the EU Structural Funds for the NDL project cover the expected expenditures only partially. In the area of human resources, the project must be heavily subsidised from the internal resources of the NL CR (as well as the partner ML). The project will have to integrate so far individually managed organisation units and will absorb several projects, which existed before rather independently.

The project will integrate the departments which were involved in in-house digitisation, in management of external digitisation and administration of digital libraries as well as the web archiving department and only recently established digital preservation department. Besides, as the NDL project will build digitisation infrastructure in-house in the NL CR and ML, the IT team will have to be strengthened too.

Both Manuscriptorium and Kramerius were independent projects based on cooperation with external entities supported by small teams on the part of the NL CR. The teams were constituted at a time when the NL CR had enough staff available and their creation did not require any reductions in other departments.

The Web Archiving Department was created in a different atmosphere – the decision to begin with this new activity required reorganisation, brought new requirements on the library processes and required new definitions of the library document. Web archiving brought heterogeneous activities into the institution which had until then dealt exclusively with traditional library analogue documents. Besides the technical part of the harvesting, which remains even for many librarians blurred, other aspects required attention too: first the legal conditions of this project and then also unusual curatorial processes (selecting, acquiring, describing and archiving of the documents is rather different in a web archive than in the rest of the Library). This all caused some misunderstandings and tensions around this department.

---

The Digital Preservation Division emerged slowly. The embryo was one full-time equivalent with a single employee. Later during the participation of the NL CR in the Digital Preservation Europe project<sup>25</sup> in 2006–2009 the NL CR has invested all of the financial means acquired from the DPE project into building the digital preservation team. The preparation of the NDL project began in 2008. Without the existence of a high-quality digital preservation team, the NL CR would not have been capable of preparing and submitting the project with one part focused on building a trusted digital repository.

The departments involved in the development of the Kramerius digital library system, the departments running current in-house digitisation, the departments managing our current digital content in the archive and digital libraries, the web archive department and a substantial part of the IT department augmented by several new employees form the new NDL team. Also other departments like the cataloguing department and administration of the library catalogue will be substantially affected by both the Google Books project and the NDL. All of this will change the long time established balance in the staff structure. Less pure librarian skills will be required and a more technically oriented system librarians or even only partly library-oriented technical experts will be needed.

The NDL project as other projects dealing with digital data require skills and an organisation culture which traditionally did not exist in the pre-digital library. In the scope of the NDL, the NL CR and ML will have to cooperate with large numbers of suppliers, project manager and administrators from external commercial institutions. This will bring the need to accommodate to the standard project management processes. The librarians tend to be over-sensitive about the standards and metadata, which may cause friction in dealing with external subjects.

### *Strategic planning*

The strategic aspects are very closely connected with the economic and staffing aspects. A widespread myth is reducing digital preservation to the purchase of HW and basic SW, pushing digital preservation somewhere into the area of IT. Digital preservation, however, influences various processes in the institution fundamentally, pervading them and requiring deep changes of the organisational structure as well as the strategies of the institution's direction.

Limiting digital preservation first to the work of one organisational unit and gradually pushing this topic to entire organisation is not the ideal approach. Precisely this was the NL CR's experience: going from one singular department of digital preservation to whole institutions devoted to digital curation. This process requires a number of small organisational and budgetary changes. A better approach would be that of the 'enlightened ruler' in the top management, who would set the digital preservation as one of the main strategic priorities. And then would steer the entire institution in this direction. However, the 'enlightened ruler' approach is seldom possible in practice within the traditionally directed memory institutions. The NL CR's experience is that at the moment the digital preservation has started to influence the basic processes of the institution, the well-worn routines can be very resistant to

---

<sup>25</sup> <http://www.digitalpreservationeurope.eu>

change and the qualification of the existing employees may be hard to improve. Without explicit support in the strategic documents, the organisational inertia may result in attempts to weaken or dismiss the whole area of digital preservation, as happened several times at the NL CR. With the instalment of the new general director of the NL CR, digital preservation has become one of the main priorities of the institution, but this does not mean complete victory yet.

Digital preservation is a very financially demanding area, but the benefits are not visible at first sight. The number of digitised volumes is eye-catching evidence of the project's success, but the investment into the preservation of metadata or building a whole preservation system is hard to sustain. The future savings of finances or documents are not visible immediately. Digital preservation will therefore be vulnerable for some time to funding cuts as the digital documents are to technological and environmental changes.

Hardly any national library today can venture to deposit millions of paper books or periodicals forming the national cultural heritage in spaces unprotected from water, fire or the entry of unauthorised persons. A number of national libraries have begun large digitisation projects and collect born-digital documents without having instruments for ensuring the long-term preservation of the digital content. Long-term preservation has not been perceived so far as an indispensable part of the digitisation projects and when setting project budgets it is usually underestimated or entirely ignored in order to digitise more documents.

## International Context

The National Library of the CR joined the DigitalPreservationEurope (DPE) project in April 2006. The library management approved the participation in this project with the aim of developing staff qualifications and learning from the emerging European digital preservation community. The main target of DPE was to 'raise the profile of digital preservation' (DPE, 2006), which was exactly what was needed. One of the main benefits of the involvement in the DPE was personal contact, the possibility to ask and to see what others were doing. In addition, there was a unique chance to organise a week-long WePreserve training in Prague<sup>26</sup> with people from other European projects like PLANETS<sup>27</sup>, CASPAR<sup>28</sup>, Nestor<sup>29</sup>, DRAMBORA<sup>30</sup> and JISC<sup>31</sup>, presenting the basic concepts of digital preservation. Some 25 librarians and archivists from across the country's culture heritage institutions could profit from this workshop. Following the DRAMBORA training in Prague<sup>32</sup>, it had the same impact for the NL CR and for the memory-institution community in the Czech Republic.

During the DPE project, the NL CR published a number of DPE papers in Czech translation<sup>33</sup>, chiefly the document called PLATTER<sup>34</sup>, co-authored by one of the NL CR's staff. This was the first document in the Czech language to explain the digital

<sup>26</sup> <http://www.wepreserve.eu/events/prague-2008/>

<sup>27</sup> [www.planets-project.eu](http://www.planets-project.eu)

<sup>28</sup> <http://www.casparpreserves.eu/>

<sup>29</sup> <http://www.langzeitarchivierung.de/>

<sup>30</sup> <http://www.repositoryaudit.eu/>

<sup>31</sup> <http://www.jisc.ac.uk/>

<sup>32</sup> <http://www.repositoryaudit.eu/training/prague-2008/>

<sup>33</sup> <http://www.digitalpreservationeurope.eu/publications/briefs/>

<sup>34</sup> <http://www.digitalpreservationeurope.eu/platter/>



---

preservation issues to a wider audience. The impact of this document even surpassed the culture heritage institutions. PLATTER (Planning Tool for Trusted Electronic Repositories) is one of the main outcomes of the DigitalPreservationEurope project.

Both PLATTER and DRAMBORA were presented in a one-day Archiving Digital Documents<sup>35</sup> workshop, part of the largest conference in the field of culture heritage institutions<sup>36</sup>.

The DPE partners were also asked to test a first version of the DRAMBORA, tool developed by the DPE and Digital Curation Centre<sup>37</sup>. At the NL CR, the DRAMBORA self-audit took place in the summer of 2007 while a second audit was run in the repository of Charles University in Prague. The audit outcomes were used in the negotiations about the future repository budget, and the first steps aimed at mitigating the principal risks identified were initiated. The outcomes were also widely published across Czech memory institutions, and DRAMBORA was adopted by several other institutions. The National Technical Library for example uses DRAMBORA annually for the audit of the repository of grey literature.

Besides the DPE project, the NL CR participated in other projects related to digital preservation like Living Web Archives<sup>38</sup> (LIWA, 2008–2010) and in less formal projects and cooperation like the project called the ‘LTP Group’, initiated by the National Library of the Netherlands. In this project, the NL CR could follow the experience of a number of other European national libraries, better understand the tenants of the OAIS mapping and its practical implementations and see the achievements of other institutions in a less formal environment. This project led the NL CR to reconsider its Long-Term Preservation system requirements and to realise that wider integration of the electronic and traditional library processes will be needed.

From the beginning, the NL CR followed the development in the CASPAR and PLANETS projects, using in its day-to-day process the PLANETS tools – the PLANETS testbed and mainly PLATO for preservation planning.

Other tools like JHOVE and DROID are used in its current workflows as well. Adding technical metadata and performing format characterisation and validation of all of the data coming to the repository is the necessary first step in building a repository with a long-term preservation mission.

As the logical next step, the NL CR tried to establish partnerships with universities and other institutions. The first intention was to spread knowledge about the digital preservation issues. Endeavours were made to attract interest in digital preservation at conferences and inter-institutional working groups in the Czech Republic. Secondly, the NL CR endeavoured to involve more experts in certain fields (specifically: file formats and metadata) from universities. A strong connection to Charles University in Prague and Masaryk University in Brno was established. The NL CR’s staff holds regular courses at Charles University’s Institute of Information Studies and Librarianship at the Faculty of Arts as well as proposing and supervising relevant thesis topics. In this way, the NL CR can profit from the work of young professionals, get the tools developed and find motivated new employees.

---

<sup>35</sup> <http://skip.nkp.cz/akcArch09.htm#work>

<sup>36</sup> <http://skip.nkp.cz/akcArch10.htm>

<sup>37</sup> <http://www.dcc.ac.uk/>

<sup>38</sup> <http://www.liwa-project.eu/>

---

## Onsite Visits and NDL Planning

As the planning of the NDL project proceeded, the leading team visited several institutions with digital preservation experience. Reading articles and reports was useful, but personal visits and chat with the staff and managers proved able to provide even more. The National Library of the Netherlands (The Hague), New Zealand National Library (Wellington), German National Library (Berlin) and Wellcome Trust Library (London) were selected for visits with the aim of discussing their digital preservation experience and strategies. It was possible to see running systems and heard a great deal of experience and future plans. All of this information was used in the preparation of the requirements for the future Long-Term Preservation system, which was planned as one of three main parts of the NDL project (NDK, 2011).

As a second step, the NL CR also tried to conduct a market survey based on initial long-term preservation system requirements. The aim was to receive feedback from potential commercial suppliers. The NL CR does not have a strong IT development team, so the NDL Long-Term Preservation system was from the beginning planned as a chance to purchase a commercial solution. Two rounds of RFI with IBM, ExLibris and later on also with Tessella have taken place in 2008 and 2009. All of the companies were very welcoming, readily presented their systems and provided access to all of the relevant documentation.

Finally in 2010, the NL CR ran a small ‘proof of concept’ project. The staff had more or less only theoretical knowledge, which was not enough. The team needed to see what the commercial systems really could do and how complex it was to set the systems up, configure them, manage and ingest data into them. The three above-mentioned companies were asked for cooperation. Only Ex Libris and Tessella agreed to run through proof of concept and let the NL CR’s staff understand their systems much better. The NL CR has invested into the developers’ work and used the API of both systems, performed the necessary transformations and ingested data into the systems. We realised that Rosetta and SDB bring different approaches to building a complete solution in the area of digital preservation and that both solutions have strong points. Thanks to the ‘proof of concept’, the NL CR was able to better specify the staff skills and other requirements on the organisational structure needed to run one of the systems in a real life setting.

Besides the ‘proof of concept’, the NL CR became active in a digital preservation community and made use of the newly developed tools. Many of the individual preservation tools are freely available, but until recently no free complete digital preservation repository solutions existed. This has changed. The preservation department have experimented with Archivematica<sup>39</sup>, seen DPSP<sup>40</sup> from Australian National Archive and monitors the news and improvements of other tools like Mopseus<sup>41</sup>, RODA<sup>42</sup>, ePrints<sup>43</sup> and Fedora with its preservation extensions.

The knowledge on the recent achievements and innovations in the digital preservation area comes from personal contacts, tracking relevant websites, projects

---

<sup>39</sup> <http://archivematica.org>

<sup>40</sup> <http://dpsp.sourceforge.net/>

<sup>41</sup> <http://www.ifs.tuwien.ac.at/dp/ipres2010/papers/gavrilis-34.pdf>

<sup>42</sup> <http://www.fedora-commons.org/about/examples/roda>

<sup>43</sup> <http://www.eprints.org/software/>



and benefit is drawn also from a relevant mailing list subscription<sup>44</sup>.

## Conclusions and Recommendations

### Strategic and Economic Conclusions and Recommendations

- Digital preservation should be an inseparable component of all of the projects dealing with the digitisation of analogue documents and/or the acquisition of born-digital documents forming the national cultural heritage.
- Digital preservation is not a luxury which can be postponed until later or even entirely jettisoned. Ensuring adequate protection for digital documents should be just as natural as protecting the space for the deposition of analogue documents from water, fire or intrusion of undesired persons.
- There are several substantial differences between the securing of analogue and digital documents: digital documents are more vulnerable than analogue documents – digital preservation has not only a physical but also a logical level.
- Whereas the preservation of analogue documents is locally limited to the areas of their deposition and movement, the idea that digital preservation takes place somewhere on the grounds of IT and begins and ends with the procurement of suitable HW and SW is mistaken.
- Invest in the HW and IT staff, but do not dismiss the project management part. Strategies, preservation plans, setting-up processes, documentation writing are time-consuming. The stakeholders have to acknowledge that the digital preservation is not solely an IT issue; it is also an issue of management and financing. Often meet people as well as institutions' representatives in different countries still believe the back-up policy is a sufficient means of long-term preservation.
- Digital preservation affects an institution very complexly and creates the need for a transformation of the routine approaches and organisational changes.
- The unpreparedness of the institution for relatively fundamental changes could become a more serious hurdle for digital preservation than a lack of financial means for investment.
- Preservation policy is more important than one would think at the beginning of one's digital preservation efforts. The preservation policy should meet the needs of your projects, workflows, data types and should be in line with the strategy of the institution as a whole. Advocating the digital preservation related budget items always runs more smoothly when the goals are explicitly stated in strategic documents. This may be common practice in some countries, but it is less usual in others.
- From the wider perspective, the strategy and coordination of digital preservation on a national level are crucial as well. The recipients of public funding (libraries, museums, archives) should be impelled to concentrate on digital preservation, share the knowledge base and develop tools among institutions.
- Some funding should be clearly focused on archiving and preserving only the

---

<sup>44</sup> for example [diglib@infoserv.inist.fr](mailto:diglib@infoserv.inist.fr); [jisc-repositories@jiscmail.ac.uk](mailto:jisc-repositories@jiscmail.ac.uk); [digital-preservation@jiscmail.ac.uk](mailto:digital-preservation@jiscmail.ac.uk) etc.

---

digital data. When all of the national programmes and funding schemes ‘produce’ just digital data without relying on a clearly defined national strategy for digital preservation with resources committed to this area, it is a disaster and a waste of money.

- Considering that it is a relatively new area with which a number of memory institutions are only now beginning, it is exceptionally important to share experience and the results achieved on the international level.

### **Practical Conclusions and Recommendations**

- Starting with digital preservation now will take less time than a couple years ago.
- Tools, systems and experience in the form of papers, case studies and reports are available and their number is increasing. It is ever harder for a novice to become acquainted with the field of digital preservation. Although it is still a research area, the number of the ‘best practices’ found around the world for novices to study constantly increases.
- Do not hesitate to arrange an onsite visit. The addressed colleagues are usually keen to share their knowledge.
- Become part of the digital preservation community – follow listservs, blogs, conference proceedings and relevant project outcomes.
- Contact commercial producers. The NL CR’s experience was extremely positive; they provided the library with access to various testing sites, documentation, training materials and presentations.
- Develop a committed IT staff. Both developers and managers for the IT part of your digital preservation project are vital.
- Test and start using the available tools for format validation, metadata extractors, etc.
- A profound planning of a digital documents’ lifecycle is not a waste of time and money. The term ‘data curation’ (Harvey, 2010), which covers the whole lifecycle of a digital document, is virtually unknown to Czech libraries.
- The systems themselves are inefficient if the data do not flow in well-designated workflows. It is fairly easier to find money for building one system than running a rationally (inter-)connected workflow.

---

## References

- [website] DPE (2006). *About DPE*. Retrieved on 16 March 2011, from <http://www.digitalpreservationeurope.eu/about/>
- [website] NDK (2011). *National Digital Library Project*. Retrieved on 16 March 2011, from [http://www.ndk.cz/narodni-dk/national-digital-library/view?set\\_language=en](http://www.ndk.cz/narodni-dk/national-digital-library/view?set_language=en)
- [book] Harvey, R. (2010). *Digital Curation. A How-To-Do-It Manual*. Neal-Schuman Publishers, New York.
- [proceedings] Hutař, J., Fojtů, A., Pavlásková, E. (2008). DRAMBORA – tool for internal audit of digital repositories (version 2) and relevant information from the carried out audits. In *Inforum 2008, 28–30 May 2008 VŠE Prague*. Retrieved on 10 April, 2011, from <http://www.inforum.cz/archiv/inforum2008/en/proceedings/45/>
- [proceedings] STOKLASOVÁ, B., HUTAŘ, J., KRBEC, P., (2007). Preservation of digital cultural heritage in Europe and in Czech Republic. In *Communication of Memory in Archives, Libraries and Museums: The Interaction of Science, Policy and Practices*, Vilnius, 4–5 October 2007, Vilnius: Vilnius University Press, 2008, pp. 213–230 (ISBN 978-9955-33-314-2)
- [proceedings] STOKLASOVÁ, B., HUTAŘ, J., KRBEC, P., (2008) Long-term preservation and accessing of digital documents in national and international context. Paper presented at the UNICA Scholarly Communication Seminar: Partnership in Academic Excellence, Prague, 15–16 May 2008, from [http://www.ulb.ac.be/unica/docs/Sch-com-2008-Bohdana\\_Stoklasova.ppt](http://www.ulb.ac.be/unica/docs/Sch-com-2008-Bohdana_Stoklasova.ppt)
- [proceedings] STOKLASOVÁ, B., (2007). Strategy and Cooperation on Long-Term Preservation in the Czech Republic. Paper presented at The Challenge: Long-term Preservation, Frankfurt, April 2007, from [http://lza.ddb.de/eu2007/modules.php?op=modload&name=PagEd&file=index&page\\_id=45](http://lza.ddb.de/eu2007/modules.php?op=modload&name=PagEd&file=index&page_id=45)