

Project no. 034762

DigitalPreservationEurope (DPE)

Instrument: Coordination Action

Thematic Priority: IST-2005-2.5.10

Access to and preservation of cultural and scientific resources

D. 6.1 Market and Technology Trends Analysis

Due date of deliverable: 30 September 2006

Actual submission date: 30 March 2007

v. 1

Start Date of Project: 01 April 2006

Duration: 36 Months

Name of Organisation of Lead Contractor for this Deliverable:
Národní knihovna České republiky

Names of Partners Engaged in this Deliverable: HATII, University of Glasgow; Technische Universität Wien, Statsbiblioteket, Nationaal Archief van Nederland, Národní knihovna České republiky, Ministero per I Beni e Le Attività Culturali, Fondazione Rinascimento Digitale, Vilnius University Faculty of Communication, FernUniversität Hagen

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)

Dissemination Level: **PU (Public)**



Legal Notices

The DPE Dissemination Plan is licensed under a Creative Commons Attribution - Non-Commercial - Share-Alike 2.0 License.

© in the collective work - DigitalPreservationEurope (which in the context of these notices shall mean one or more of the consortium consisting of HATII at the University of Glasgow, FernUniversität Hagen, Fondazione Rinascimento Digitale, Ministero Per I Beni E Le Attività Culturali, Národní knihovna České republiky, Nationaal Archief van Nederland, Statsbiblioteket, Technische Universität Wien, Vilnius University Faculty of Communication, and the staff and agents of these parties involved in the work of DigitalPreservationEurope), 2006.

© in the individual reports – the organisations as indicated in the catalogue entry below unless otherwise stated.

DigitalPreservationEurope (DPE) confirms that the owners of copyright in the individual instalments have given permission for their work to be licensed under the Creative Commons license.

Catalogue Entry

Title	Market and Technology Trends Analysis
Creator	Národní knihovna České republiky
Subject	Information Technology; Science; Technology-Philosophy; Computer Science; Digital Preservation; Digital Records; Science and the Humanities; Market Analysis
Description	The objective of this deliverable is market analysis based on the consultation of the main stakeholders on the benefits (access to joint materials and training, licensing of products, etc.) and technology trends analysis providing the main target groups with information on technological solutions available for digital preservation.
Publisher	HATII at the University of Glasgow, includes FernUniversität Hagen, Fondazione Rinascimento Digitale, Ministero Per I Beni E Le Attività Culturali, Národní knihovna České republiky, Nationaal Archief van Nederland, Statsbiblioteket, Technische Universität Wien, Vilnius University Faculty of Communication.
Contributor	HATII at the University of Glasgow
Contributor	Ministero Per I Beni E Le Attività Culturali
Contributor	Fondazione Rinascimento Digitale
Contributor	Vilnius University Faculty of Communication
Contributor	FernUniversität Hagen
Contributor	Nationaal Archief van Nederland
Date	30 January 2007
Type	Text
Format	Adobe Portable Document Format v.1.3
Resource Identifier	ISBN 978-1-906242-02-2
Language	English
Rights	© DigitalPreservationEurope Partners

Citation Guidelines

Digital Preservation Europe, (January 2007), Market and Technology Trends Analysis, DPE-D6.1, Retrieved <date>, from <http://www.digitalpreservationeurope.eu/publications/dpe-market-analysis.pdf>

Document Version Control

Version	Date	Change Made (and if appropriate reason for change)	Initials of Commentator(s) or Author(s)
0.1	30/01/2007	Initial Version of D6.1	Bohdana Stoklasová Jan Hutař Anna Nerglová Adolf Knoll
0.2	30/03/2007	Version with added Chapters	Bohdana Stoklasová Jan Hutař Anna Nerglová Adolf Knoll
0.3	10/04/2007	Version with comments included from the DPE team	Bohdana Stoklasová
0.4	14/04/2007	Version with comments included from the DPE team	Jan Hutař
1.0	19/10/2007	Changes for re-submission	Jan Hutař Anna Nerglová

Document Quality Control

Version QA'd	Date	Recommendations (and, if appropriate, reason for change)	Initials of Quality Assurance Person

Document Change Commentator or Author and Reviewer List

Author Initials	Name of Author/Reviewer/Responsible QA Person	Institution

Approved by with DPE (signature) (S. Ross)	Date

Accepted by at European Commission (signature)	Date

1. EXECUTIVE SUMMARY	7
1.1 Purpose	7
1.2 Structure	7
2. INTRODUCTION	9
2.1 WP 6 Sustainability and Exploitation: broader context	9
2.2 The Methodology developed	9
3. PART I: DPE – DESCRIPTION AND ITS TARGET AUDIENCE.....	10
3.1 Objectives	10
3.2 List of project participants/partners	11
3.3 Target audiences.....	13
4. PART II: ANALYSIS OF PRESENT AND FUTURE GOALS AND NEEDS OF MAIN DPE TARGET GROUPS IN THE AREA OF DIGITAL PRESERVATION.....	16
4.1 Survey on long-term preservation in European National Libraries	17
4.2 Survey on long-term preservation in Archives	25
4.3 Survey on long-term preservation in Research Institutions	31
4.4 Survey on long-term preservation in ICT Companies and Media	38
4.5 Survey on long-term preservation in “Others”	45
4.6 Comparison of all survey results	52
5. PART III: TECHNOLOGICAL SOLUTIONS AVAILABLE FOR DIGITAL PRESERVATION.....	54
5.1 List and characteristics of systems mostly used for digital preservation.....	54
5.1.1 Commercial.....	54
5.1.2 Open Source based.....	61
6. PART IV: CONCLUSIONS AND RECOMMENDATIONS	84

1. EXECUTIVE SUMMARY

1.1 Purpose

This report presents the **Market and Technology Trends Analysis** to be used and distributed by DPE, DigitalPreservationEurope, a project funded by the Sixth Framework Programme, Priority IST-2005-2.5.10, action line: Access to and preservation of cultural and scientific resources.

This deliverable has two main objectives:

- **Market analysis** based on experience and knowledge of all the contributors and the consultation of the main stakeholders on their needs and plans so that the outputs of the DPE project meet their present and future demands.
- **Technology trends analysis** providing the main DPE target groups with information on technological solutions available for digital preservation.

1.2 Structure

The document is divided into four main parts, each providing information and knowledge necessary for answering two main questions:

- What are the present and future goals and needs of the main DPE target groups in the area of digital preservation, and how can the DPE outputs best meet these? This part of the analysis should be an important background and stimulation for digital preservation technology vendors and developers.
- What are the technological solutions, both commercial and Open Source, applicable and already applied for digital preservation in different institutions belonging to the main DPE target group that meet minimal functional requirements?

I. DPE – description and its target audiences

A definition of the main DPE target audiences is necessary in order to structure DPE outputs into groups appropriate to their particular needs and to prepare surveys with a two-fold purpose:

- To get necessary feedback for them.
- To inform them about DPE outputs available now and in the future.

II. Analysis of present and future goals and needs of the main DPE target groups in the area of digital preservation and finding how the DPE outputs can best meet these.

This analysis is based on an extensive survey distributed to the main DPE target groups. The choice of these groups is based on DPE targets, as listed in the DPE Dissemination plan (D 7.1).

- National Libraries
- Archives
- Industry (ICT companies; Media)
- Research Institutions (including universities)
- Others (Non-governmental institutions and organisations; Related projects, coalitions and initiatives; Governmental institutions and local authorities)

III. Technological solutions available for digital preservation

List and characteristics of systems mostly used for digital preservation

- Commercial
- Open Source based

IV. Conclusions and recommendations

SWOT Analysis aggregating information derived from the Market and technology trends analysis.

- Strengths
- Weaknesses
- Opportunities
- Threats

2. INTRODUCTION

2.1 WP 6 Sustainability and Exploitation: broader context

This deliverable presents the first results of *WP 6 Sustainability and Exploitation, Task 6.1 Market and Technology Assessment*. It is intended that the report, which is in the version for re-submission, be completed by all the contributors (in the way referred to in specific chapters).

This work package will ensure that the outcomes of DigitalPreservationEurope are sustainable and that its outputs are fit for purpose and utilised by their intended audiences.

Ensuring that the outputs of DigitalPreservationEurope are of high quality, are recognised and valued by the intended stakeholders, and are evaluated, developed and enhanced in the light of user feedback will also be important to the impact and uptake of outputs from the project and its future sustainability.

Task 6.1 Market and Technology Assessment focuses on market analysis and the systematic consultation of stakeholders on the benefits (access to joint materials and training, licensing of products, etc.) and the scale of individual national contributions or individual institutional contributions to joint collaboration on a European basis.

2.2 The Methodology developed

After conducting preliminary desk-based research, multi-step analysis was chosen as the method most suitable for the purpose of the project.

Step 1: Target audience analysis based on experience and knowledge of all the contributors.

Step 2: Market analysis based on experience and knowledge of all the contributors.

Step 3: Market analysis based on the consultation of the main stakeholders on their needs and plans so that the outputs of the DPE project meet their present and future demands.

Step 4: Technology trends analysis based on experience and knowledge of all the contributors.

Step 5: Technology trends analysis based on experience and evaluation of different digital repositories.

All the information acquired during this five-step analysis was structured and evaluated, and the synthesis is presented in the concluding SWOT Analysis.

3. PART I: DPE – DESCRIPTION AND ITS TARGET AUDIENCE

DigitalPreservationEurope (DPE) fosters collaboration and synergies between many existing national initiatives across the European Research Area. *DPE* addresses the need to improve coordination, cooperation and consistency in current activities to secure effective preservation of digital materials. *DPE*'s project partners lead work to:

- raise the profile of digital preservation;
- promote the ability of Member States acting together to add value to digital preservation activities across Europe;
- use cross-sectoral cooperation to avoid redundancy and duplication of effort;
- ensure auditable and certificated standards for digital preservation processes are selected and introduced;
- facilitate skills development through training packages;
- enable relevant research coordination and exchange;
- develop and promote a research agenda roadmap;
- help both citizens and specialist professionals recognise the central role that digital preservation plays in their lives and work.

DPE's success will help to secure a shared knowledge base of the processes, synergy of activity, systems and techniques needed for the long-term management of digital material.

3.1 Objectives

DigitalPreservationEurope (DPE) has three main objectives, each of them with one or more sub goals.

1. *To create a coherent platform for proactive cooperation, collaboration, exchange and dissemination of research results and experience in the preservation of digital objects.*

- a. To identify and raise awareness of sources on the issues surrounding the curation and preservation of digital objects across the broad spectrum of national and regional cultural and scientific heritage activity in Europe.
- b. To contribute to the elimination of the duplication of effort of research activities by researchers at different institutions and to enable identification, collection and sharing of knowledge and expertise.
- c. To create a conduit between the research community and practitioner community that will foster the collaborative approaches to preservation needs.
- d. To stimulate and coordinate further research on digital preservation in key areas and encourage the development of standards where gaps and opportunities have been identified. This will include promoting and developing research agendas.

2. *To increase prevalence of preservation services and their viability and accountability.*

- a. To support the development of a European-wide approach to the audit and certification of digital repositories as an essential stage in creating content management and delivery services and to repository federation.
- b. To stimulate ICT companies and software developers to incorporate some of the curation and preservation thinking into newer generations of software.
- c. To relate the digital preservation research agenda more directly to the development of exploitable product opportunities and to develop links with the industrial sectors.

3. *To improve awareness, skills and available resources.*

- a. To examine core issues that will deliver essential guidelines, methods and tools to enable preservation action with European public and private sectors.
- b. To implement a suite of training seminars based on best practice and to identify where and what further practitioner training and staff development initiatives might be undertaken.

3.2 List of project participants/partners

Role	No.	Name	Short name	Country
Co	1	HATII, University of Glasgow	GU	UK
CR	2	Technische Universität Wien	TUW	AT
CR	3	Statsbiblioteket, Arhus	SB	DK
CR	4	Nationaal Archief van Nederland	NANETH	NL
CR	5	Národní knihovna České republiky	NKP	CZ
CR	6	Ministero per i beni e le attività culturali	MIBAC	IT
CR	7	Fondazione Rinascimento Digitale	FRD	IT

CR	8	Vilnius University, Faculty of Communication	VUFC	LT
CR	9	Fern Universität Hagen	HUF	DE

3.3 Target audiences

DPE targets all the stakeholders producing (creators), storing (curators) and using (users) digital data from the cultural, educational, industrial and public sectors ranging from memory institutions, schools and universities, private companies, private foundations, government agencies and local authorities, mass media and publishers to individual citizens whose production of digital data has recently been increasing dramatically along with the fast development of technologies supporting leisure activities, such as digital photography and digital video in particular.

DPE should recognise the breadth and the incredible variety of its potential targets. At the same time it should differentiate between communities that are particularly aware of the importance of digital preservation issues and do not need so many awareness-raising activities, and communities that are less aware of the strategic importance of digital preservation and to whom DPE awareness-raising activities need to be addressed primarily.

Already aware to some extent:

- Non-governmental institutions and organisations (e.g. UNESCO, IFLA, ICA, ICOM, ICRC or other foundations, societies, associations)
- Related projects, coalitions, organisations and initiatives (e.g. CASPAR, PLANETS, DELOS, DPC, DCC, PADI)

Low awareness:

- ALM
- Research Institutions (public and private, commercial and non-profit: e.g. Universities, Open Archives community)
- Governmental institutions and local authorities (e.g. EU institutions, e-government community, e-health community)
- ICT companies
- Media

DPE will address producers and curators of digital data and all those who can contribute to promoting digital preservation and to the safeguarding and long-term archiving of digital data. An important part of DPE dissemination will be focused on those who can - by their decisions and/or funding - influence digital preservation in a decisive way. These are governments, ministries, local authorities, etc. (key decision makers and funding bodies). Of course, both corporate and individual users will also be addressed.

In the next 12 months, DPE will prioritise and focus awareness-raising activities on ALM, Research institutions and Governmental institutions and local authorities. The

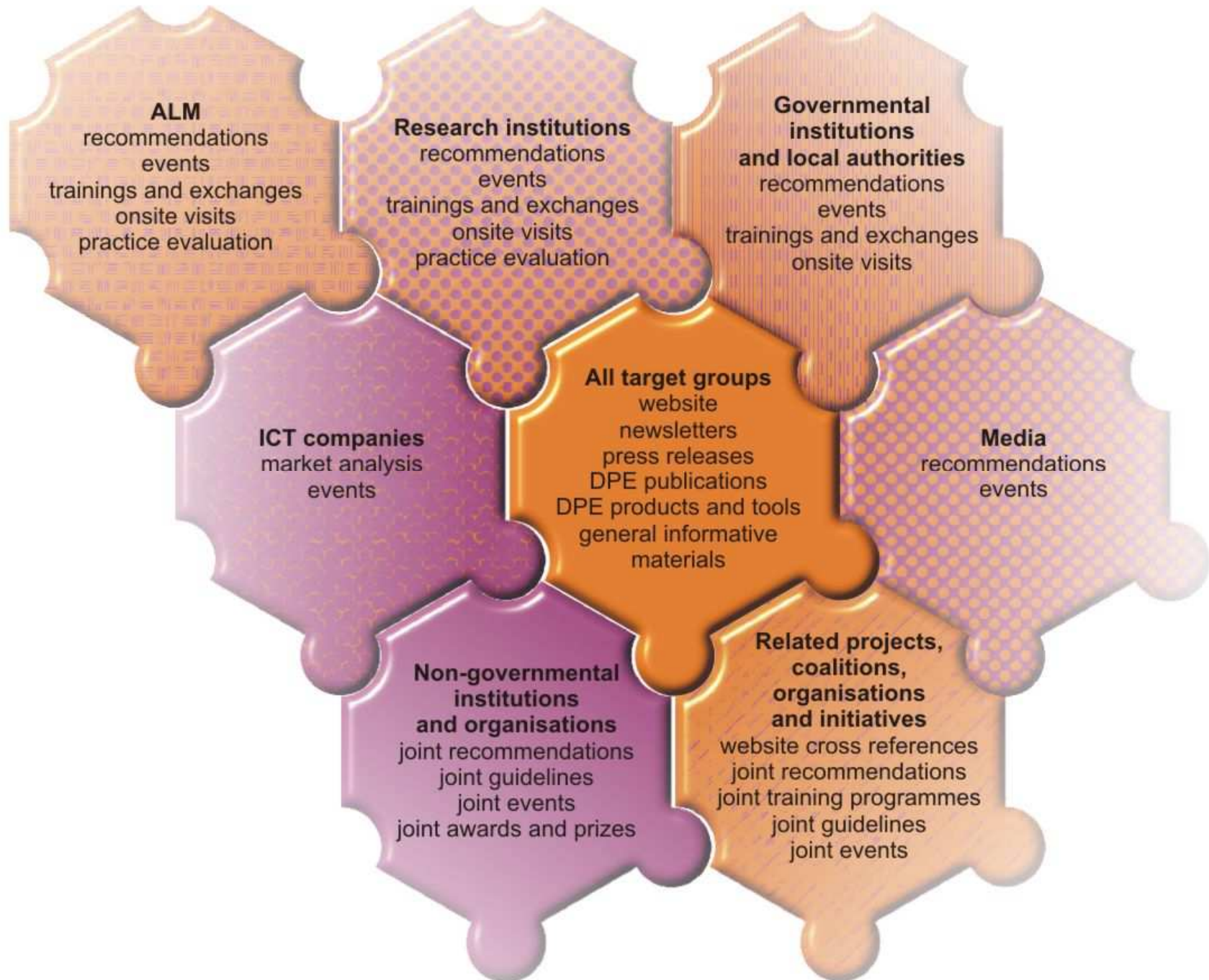
reason for prioritising is based on the assumption that the 'preservation mandate' is considerably stronger for these communities than for the others.

Almost all these stakeholders are active in many areas and they have a combined role of creators/curators/users. However, DPE will address them through its differentiated outputs focusing (in both content and form) on their main mission and needs. DPE aims to address all the stakeholders in the most appropriate and effective way, and thus very different events and materials have to be prepared for, for example, memory institutions (much deeper and more problem-oriented) and government and ministries responsible for their funding (more general, attractive, warning and persuading).

DPE also targets all the players who are potentially responsible for increasing awareness of digital preservation, especially mass media, and those responsible for the management of programmes from which the digital preservation solution and applications can be funded. We are aware that addressing ICT companies and media is a difficult task, as all the DPE partners come mainly from the memory institutions sector.

The DPE Dissemination Model presents different target groups together with services offered to them.

DPE Dissemination Model:



It is obvious from the model that most DPE services and outputs are focused on the following groups:

- **Memory institutions** (libraries, museums and archives)
- **Research institutions** (public and private, commercial and non-profit: e.g. Universities, Open Archives community)
- **Industry** (ICT companies and Media)

The first two groups are building, or will in the future build, large digital repositories, and digital preservation is important for them; however, their approach to digital preservation, especially in the long term, could be different.

ICT companies and media are building repositories too; their solutions could be very helpful and inspirational for the whole digital preservation community.

- **Memory institutions** (libraries, museums and archives) were for several centuries trusted repositories for traditional documents. Nowadays, they face a new challenge of becoming trusted digital repositories also. They belong to the category of **Digital Repositories** – broadly defined as ‘organisations that intend to maintain information for access and use’. The reason for addressing archives and libraries separately is based on the assumption that the ‘preservation mandate’ is considerably stronger for these communities and that development in archives and libraries is quite different.
- **Research institutions** belong more to the category of **Institutional Repositories** - broadly defined as a set of services by which an institution makes its intellectual output available to the wider community.
- **Industry** (ICT companies and Media) has a special status within DPE and digital preservation also. They have their own digital repositories, but a slightly different purpose – not to store cultural or educational documents for users, but rather their own data exclusively for their own use. They are that part of the digital preservation community that helps to implement and bring to life technical solutions for digital preservation problem and issues.

The survey on present and future needs and expectations described in the following chapter focused specifically on these groups.

4. PART II: ANALYSIS OF PRESENT AND FUTURE GOALS AND NEEDS OF MAIN DPE TARGET GROUPS IN THE AREA OF DIGITAL PRESERVATION

This analysis is based on the same survey distributed to important target groups:

- European national libraries (survey distributed to all of them)
- National archives (not limited only to Europe)
- Research institutions (not limited only to Europe)
- ICT companies and Media (Industry) (not limited only to Europe)

The survey was designed at the NKP. We decided to address libraries and archives directly as they are the institutions with the strongest 'preservation mandate' and we (NKP as a national library) have lots of established contacts in this memory institution area. We had no serious problems getting a sufficient number of surveys completed by research institutions (mainly universities). This helped us to achieve a high percentage return of completed questionnaires.

We had to use different channels to address research institutions, but we were quite successful and received a reasonable number of responses to our survey.

It was much more difficult to find a way to address major technology and telecommunications players and media. We received fewer responses, but nevertheless some significant ones, for example from ExLibris, Siemens, Sony, IBM, BBC, Ask.com, Daydream Ltd etc.

We did not address institutions in the 'Others' category directly, but we did receive some responses, in fact sufficient to have an appropriately helpful analysis.

4.1 Survey on long-term preservation in European national libraries

Under the umbrella of the Digital Preservation Europe (DPE) Project, the NKP, as one of the DPE partners, carried out a representative survey focusing on recent developments and plans concerning long-term preservation of digital documents in European national libraries. This report represents the results of this survey. Having fresh information about all current (or planned) activities is very important in terms of achieving the DPE objectives, which can sustain the coordination of European activities and knowledge sharing in this field.

NKP distributed 55 questionnaires both by electronic e-mail and in printed form by mail. The deadline for returning completed questionnaires was 8 September 2006. The response rate was low in the first round probably because of holiday time, therefore NKP decided to conduct a second round and extend the deadline to 8 October 2006. In the end, NKP received 36 completed (65%) questionnaires. An evaluation of the answers received was carried out for 36 libraries, or more precisely for 35 libraries, because the National Library of Andorra was not able to answer the questions at that time, as they stated in their response to the questionnaire.

The questionnaire containing seven simple questions was accompanied by an 'explanatory letter' describing DPE project aims etc. The first five questions were intended to find out how important long-term preservation was for libraries in general and what stage they were at in the area of building digital repositories, if there was a will or need to cooperate and, in the case of a positive answer, with whom. The last two questions dealt with DPE project issues.

The vast majority of libraries/respondents addressed considered the long-term preservation of digital documents to be one of their key strategic priorities. Only 2 of the 36 libraries replied that digital document preservation was not a priority for them at the moment. Four libraries did not include this topic among their key strategic priorities; however, they stated that the topic was one of their priorities for the future. The National Library of Andorra stated that they were not able even to think about this topic area and they did not continue completing the survey.

More than half of the libraries addressed did not have a trusted repository for the long-term preservation of their digital documents. The majority of libraries seem to be in the planning phase, which should result in the building of trusted repositories in the near future. Only a few libraries stated that they already had a trusted repository in operation.

Note: The question asked precisely if they have or if they plan to have a so-called 'trusted repository', but it was quite obvious from the answers that this question was not properly understood. Many libraries reported having a trusted repository, without knowing what the term 'trusted repository' meant and they considered it just as any type of repository.

Digital preservation is too big an issue for individual institutions to address independently, which is the main reason why institutions try to cooperate. All the libraries answered that they are cooperating on this issue with other memory institutions (for example, other libraries, museums, archives). Almost two-thirds of the libraries cooperate with research institutions, digital document producers and SW developers/vendors or IT companies. In a few cases they indicated cooperation with other institutions such as policy bodies or specialised archives.

The building and operation/maintenance of digital repositories is a very complex and expensive business. It is clear that some kind of cooperation between institutions is very likely in this field. Just seven libraries responded that their digital repository had been created and is now operated exclusively for their own needs. Almost all the libraries cooperate or plan to cooperate with other institutions in the country, mainly with other libraries and archives.

The system used for repository management is of key importance for its performance and reliability. It is absolutely necessary to guarantee its quality, integrity and ability to keep large amounts of digital documents permanently safe on the professional level and provide adequate access conditions/controls to the documents now and also in the future. Overall, most of the European libraries do not have a trusted repository in operation at the moment, so it might have been very difficult for them to say which software system for repository management would be the best for the library and which one will be finally chosen for implementation. Four libraries would like to use or are already using commercial systems (independently/stand-alone or in combination with other tools). A number of libraries would like to use Open Source systems, very often in combination with other commercial or proprietary in-house systems. Seven libraries decided to develop their own in-house system for repository administration.

The last part of the questionnaire focused on getting feedback about the DPE outputs most suited to the needs of different national libraries. The following outputs seem to be the most popular: conferences, seminars, workshops and websites. Other much appreciated dissemination media include guidelines, recommendations, training, tutorials and on-site visits. At the end of the wish list we can find press releases and associate partnerships. Newsletters and evaluations were also indicated as a suitable method of dissemination.

The very last question in the survey concerned possibilities to create so-called national competence centres responsible for digital document preservation at the national level, which would cooperate with other similar centres in Europe, as referred to in FP7. The majority of respondents indicated memory institutions (national libraries or archives) as being appropriate to play this role.

From all the answers it is obvious that all the issues and possible solutions concerning long-term preservation and access to digital documents are still pending, and we have no hesitation in saying that this is a 'hot' topic of today in all European national libraries. All the players in the field consider wide international cooperation to be extremely important for providing complex solutions, which is a positive sign.

Most libraries in Europe are just starting to address the digital preservation issue and are trying to find the most important information, good examples and partners. They are trying to avoid blind alleys and to avoid wasting time and resources. This is the right moment to foster cooperation among institutions at both national and international levels, and this is one of the main DPE goals.

We strongly believe that the answers received on the basis of this survey will help to increase the level of awareness about digital preservation in general and that the DPE project will provide more information and support for implementing trusted repositories for keeping European cultural heritage safe and widely accessible.

1. Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your library?

1 a) Yes

1 b) No

1 c) Not yet (please specify when it will be)

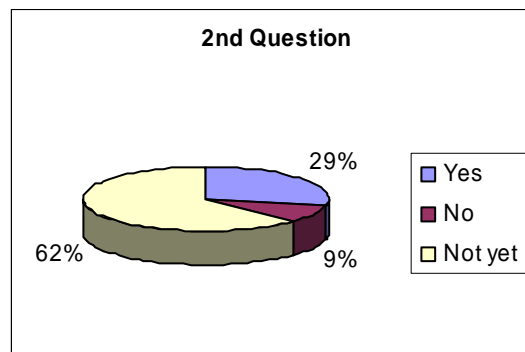


The majority of libraries/respondents addressed consider the long-term preservation of digital documents as one of their key strategic priorities. Just 2 of the 36 libraries answered that digital document preservation was not a priority for them at the moment. Four libraries do not regard this topic as one of their key strategies, but they count on addressing this topic in the future. The National Library of Andorra stated that they are not even able to think about this topic area at present and did not complete the survey.

2. Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?

<http://www.ndk.cz/dokumenty/rlgnara-repositorieschecklist.pdf>

- 2 a) Yes
- 2 b) No
- 2 c) Not yet (please specify when you plan to have it)

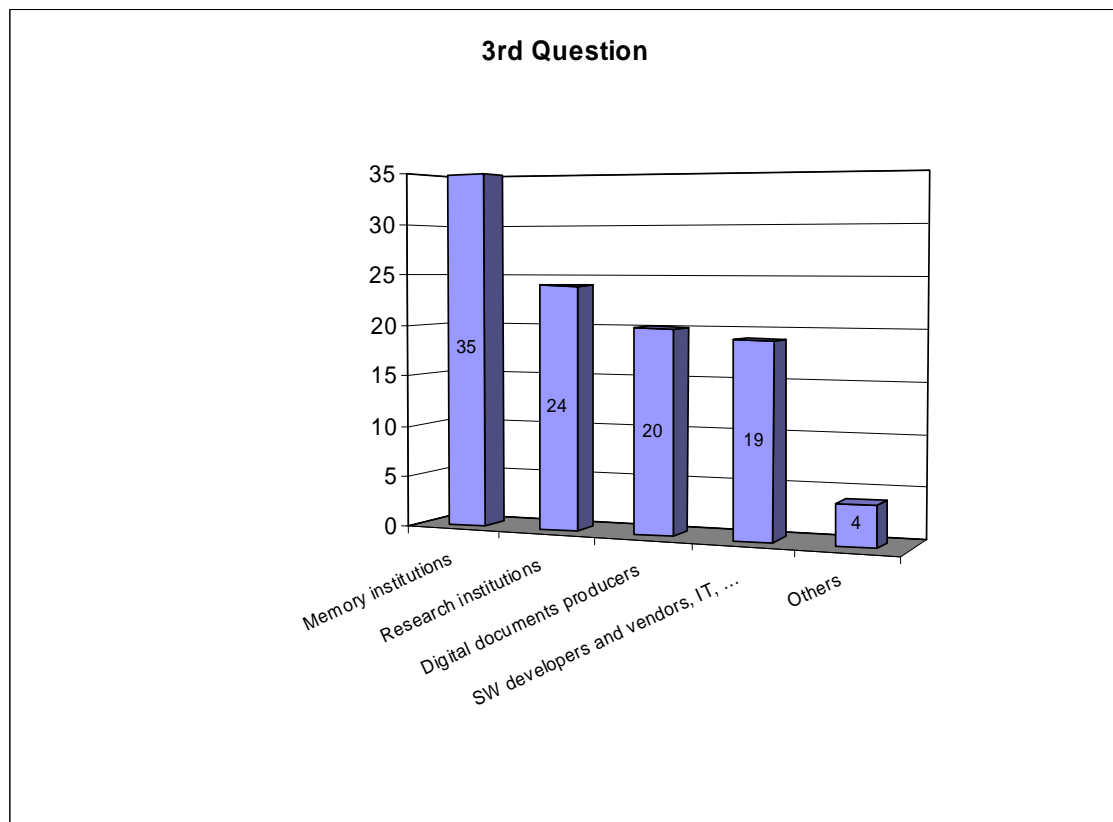


The question of whether libraries have (or will have) a trusted digital repository was answered in most cases (22 libraries) to the effect that they don't have one yet but are planning to create such a repository in the near future. Ten libraries stated that they already have a repository conforming to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories (<http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf>). Only three libraries neither have this kind of repository nor are they planning to have it in the future, and this in spite of considering the long-term preservation of digital documents as their priority (from the previous question).

Note: The question asked precisely whether they have or plan to have a so-called 'trusted repository', but it was quite obvious from the answers that this question was not properly understood. Many libraries reported having a trusted repository, without knowing what the term 'trusted repository' means and they considered it just as any type of repository.

3. Digital preservation is too big an issue for individual institutions to address independently. Your library will cooperate in this area with:

- 3 a) Memory institutions (libraries, museums, archives etc.)
- 3 b) Research institutions (schools, universities etc.)
- 3 c) Digital document producers (publishers, broadcasting etc.)
- 3 d) SW developers and vendors, IT, computer science
- 3 e) Others (please specify)

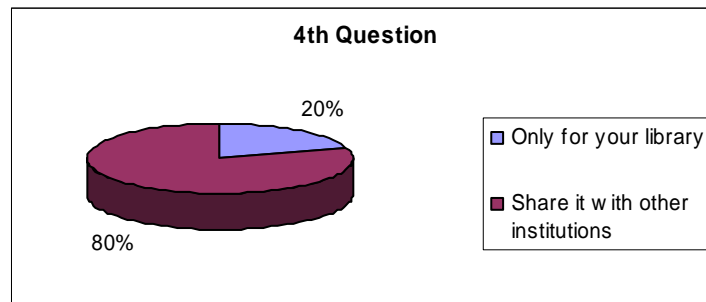


Digital preservation is a major issue and complex problem, which is the main reason why institutions cooperate with each other. Regarding this, there was a question in the survey questionnaire asking respondents to identify the types of institutions the libraries are (or will be) cooperating with on long-term preservation activities. Almost all the libraries (35) answered that they are cooperating on this issue with other memory institutions (for example, other libraries, museums, archives). Almost two-thirds of the libraries cooperate with Research institutions (24), Digital documents producers (20) and SW developers/vendors or IT companies (19). Other institutions such as policy bodies or specialised archives were mentioned by four libraries.

4. Building and operation of a trusted digital repository is a big and expensive business. You will create and operate the repository

4 a) Only for your library

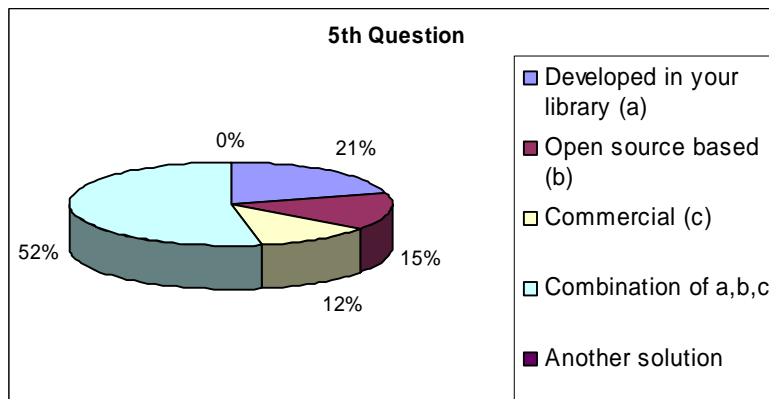
4 b) Share it with other institutions (please specify)



The building and operation/maintenance of digital repositories is a very complex and expensive business. It is clear that some kind of cooperation among institutions is very likely in this business. Just 7 libraries responded that their digital repository had already been created and is now operated exclusively for their own needs. The remainder of the 35 libraries (i.e. 28) cooperate with other national/local institutions, mainly national memory institutions - libraries and archives.

5. The system used for your digital repository is (will be)

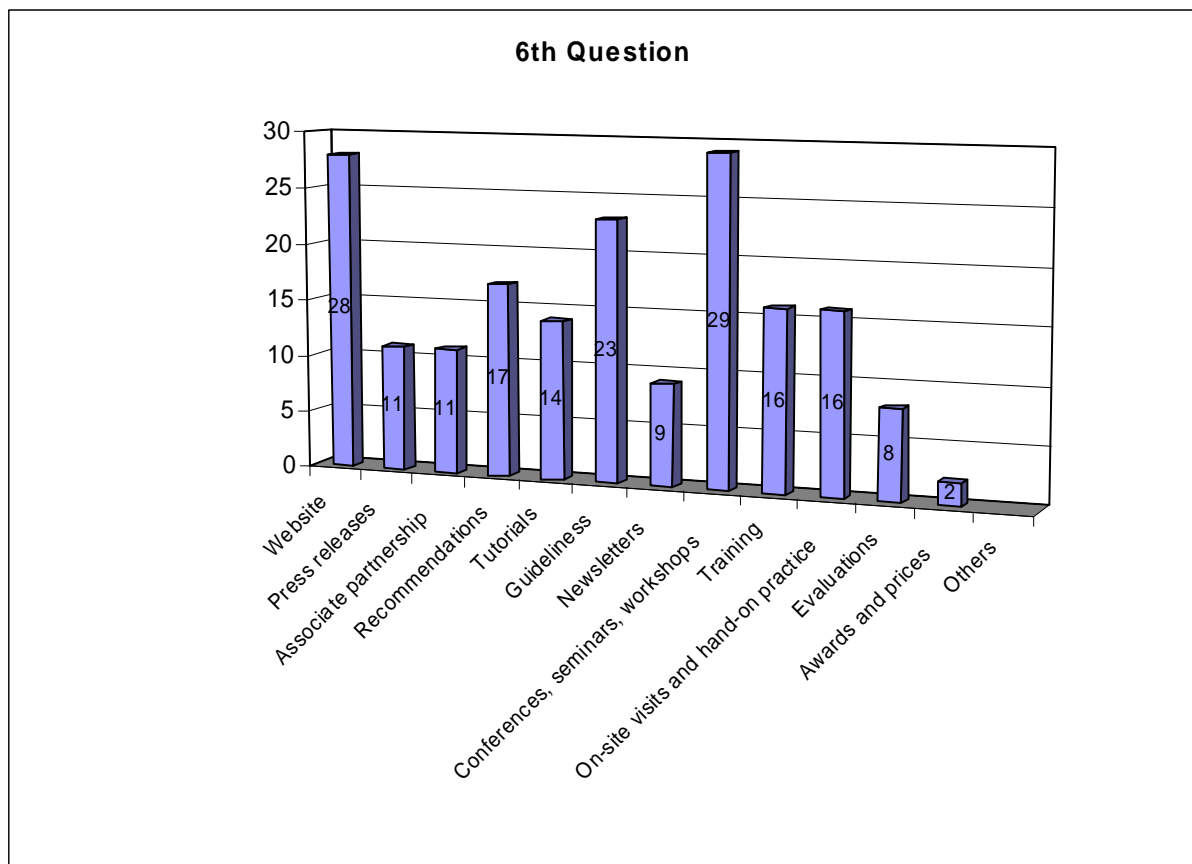
- 5 a) Developed in your library
- 5 b) Open Source based
- 5 c) Commercial
- 5 d) Combination of 5a), b), c) (please specify)
- 5 e) Another solution (please specify)



In this question libraries were asked to provide some detailed information about the system used for their digital repositories that is already in use or will be implemented in the near future. Four libraries would like to use, or are already using, commercial systems (independently/standalone or in combination with other tools). A number of libraries would like to use Open Source systems, very often in combination with other commercial or proprietary in-house systems. Seven libraries decided to develop their own in-house system for repository administration.

6. Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your library?

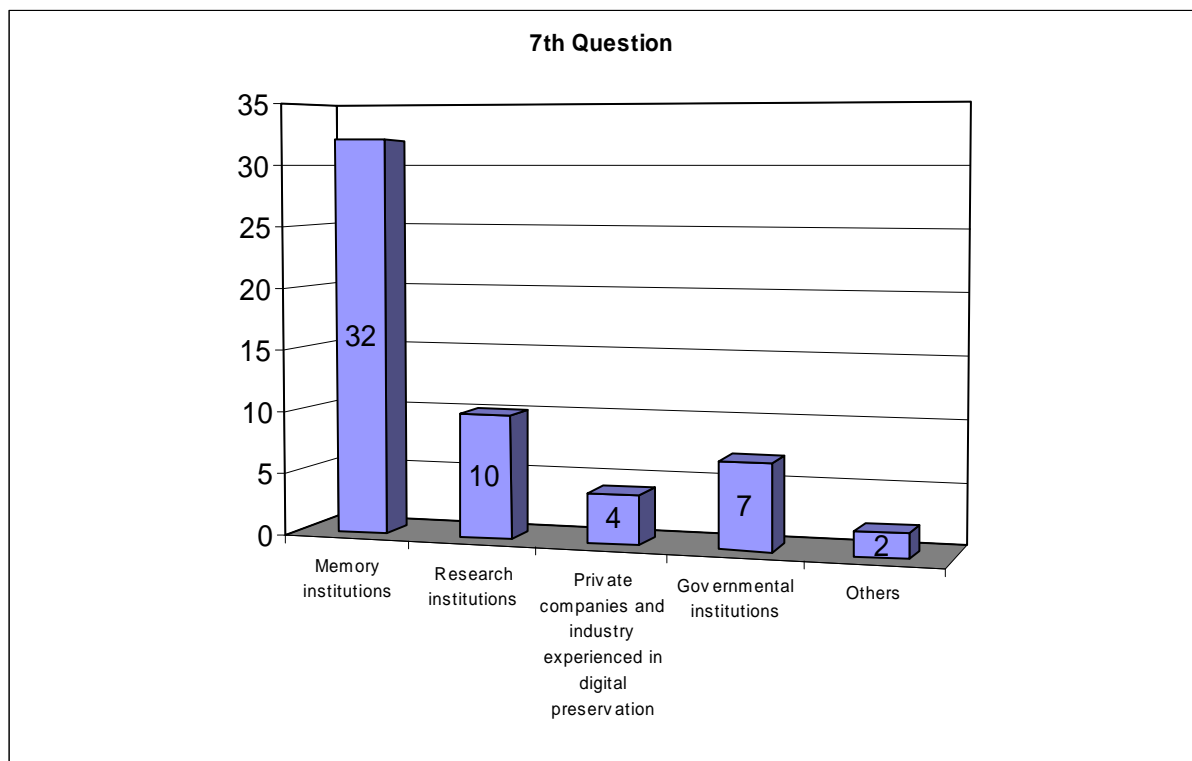
- 6 a) Website
- 6 b) Press releases
- 6 c) Associate partnership
- 6 d) Recommendations
- 6 e) Tutorials
- 6 f) Guidelines
- 6 g) Newsletters
- 6 h) Conferences, seminars, workshops
- 6 i) Training
- 6 j) On-site visits and hands-on practice
- 6 k) Evaluations
- 6 l) Awards and prizes
- 6 m) Others – not listed here but desired (please specify)



In the sixth question libraries were required to decide which of the outputs listed in the model of DPE dissemination they consider to be most relevant for them. The most frequent answers were: option 6h) Conferences, seminars and workshops (29 times) and option 6a) Websites (28 times). Other frequent answers were: Guidelines (23 times), Recommendations (17), On-site visits and hand-on practice (16), Training (also 16) and Tutorials (14). Similarly, 11 libraries consider as most relevant for them options 6b) Press releases and 6c) Associate partnership. As suitable methods of dissemination, option 6g) Newsletters (9) and 6k) Evaluations (8) were also indicated. On the other hand, option 6l) Awards and prizes does not seem to be considered as a relevant method of dissemination (2).

7. In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background to become fully operational and trusted national competence centres?

- 7 a) Memory institutions (libraries, museums, archives etc.)
- 7 b) Research institutions (schools, universities etc.)
- 7 c) Private companies and industry experienced in digital preservation
- 7 d) Governmental institutions
- 7 e) Others (please specify)



The last question in the survey concerned possibilities for creating so-called national competence centres responsible for digital document preservation at the national level, which would cooperate with other similar centres in Europe, as referred to in FP7. The question

was which kind of institution should play this role in each country. The overall majority of respondents (32) selected memory institutions as appropriate. On this level it should be, for instance, national libraries, national archives. This answer also included in 10 cases Research institutions, which were never listed alone but always with other institutions. Governmental institutions were mentioned in seven cases, again in six of these answers in combination with memory institutions. Governmental institutions were mentioned only once (Portugal) as appropriate on their own, not in combination with other institutions. Private companies experienced in the field of digital preservation were listed three times, always in combination with other institutions. In one case a private company was given as the best candidate for the National competence centre position (Czech Republic). The National Library of Serbia considers as appropriate the existing National Centre for Digitisation, while the Helsinki University Library selected the Finnish IT centre for Science.

4.2 Survey on long-term preservation in Archives

This is a report on the second survey addressing Archives. It represents the results of the survey that was conducted on behalf of the DPE project during Spring 2007 and Autumn 2007 (second round). The aim of this survey was to describe the status of long-term preservation of digital documents in archives. It was expected to receive responses from European archives, but because of the distribution method we received many answers from non-EU archives and these are also included.

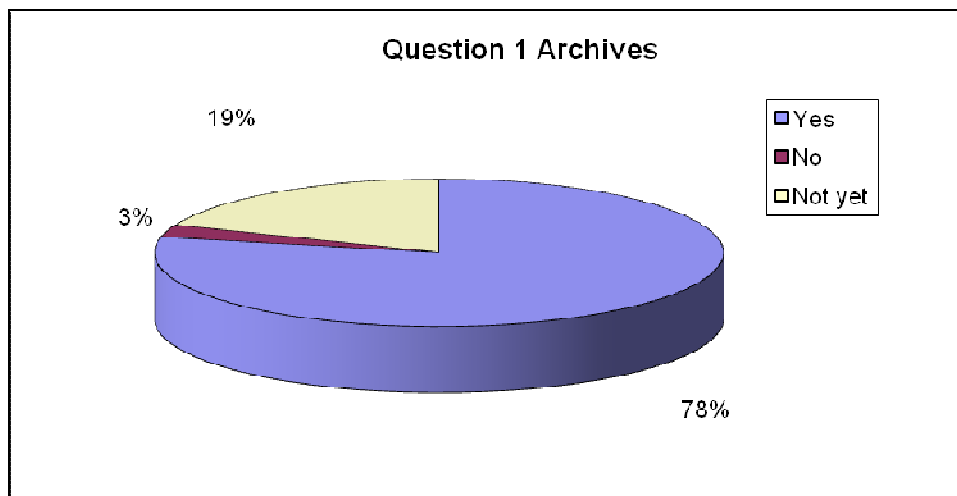
The announcement about the online survey questionnaire was distributed via the following mailing lists:

ICA-L (International Council on Archives list)
Dutch listserv: DIVAkoepel discussionlist
ERECs-L@LISTSERV.ALBANY.EDU
'interoperability@jiscmail.ac.uk'
'aus-archivists@archivists.org.au'
'ICARN-LIST@JISCMail.AC.UK'
JISC-REPOSITORIES@JISCMail.AC.UK
REPOSITORY-AUDIT-CERTIFICATION@JISCMail.AC.UK
DIGITAL-PRESERVATION@JISCMail.AC.UK
CERL mailing list

In the end, we received 37 completed questionnaires.

1. Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your institution?

- 1 a) Yes
- 1 b) No
- 1 c) Not yet (please specify when it will be)



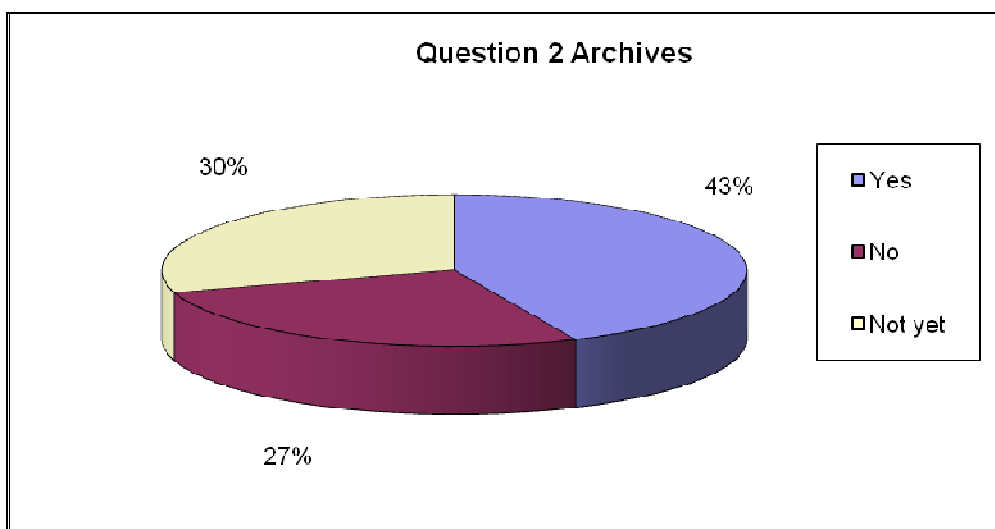
The vast majority of respondents addressed consider the long-term preservation of digital documents as their key strategic priority. From 37 answers, 29 were positive (78%) and just 1 negative (3%). Seven institutions (19%) stated that this topic is not one of their strategic priorities, even if they count on addressing this issue in the future.

2. Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?
<http://www.ndk.cz/dokumenty/rlgnara-repositorieschecklist.pdf>

2 a) Yes

2 b) No

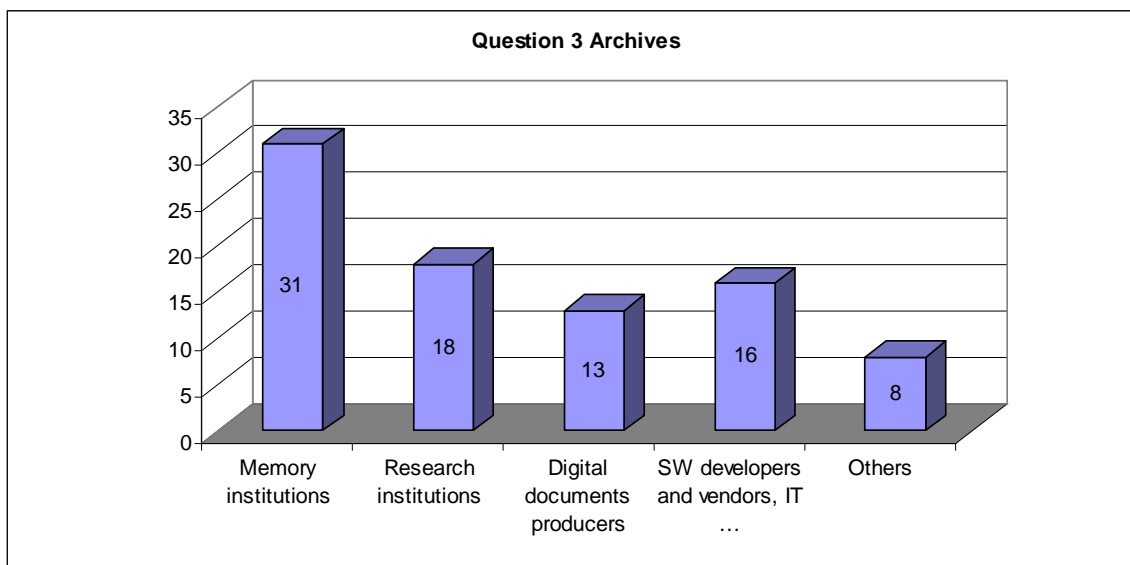
2 c) Not yet (please specify when you plan to have it)



The question whether respondents have (or will have) a trusted digital repository received the answer “yes” from more than one third of institutions (43%). Ten archives (27%) stated that they don't at present have a repository conforming to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories, but they are planning to have one in the near future (in 2-5 years' time). More or less one third of institutions (30%) do not yet have a trusted repository.

3. Digital preservation is too big an issue for individual institutions to address independently. Your institution will cooperate in this area with ...

- 3 a) Memory institutions (libraries, museums, archives etc.)
- 3 b) Research institutions (schools, universities etc.)
- 3 c) Digital document producers (publishers, broadcasting etc.)
- 3 d) SW developers and vendors, IT, computer science
- 3 e) Others (please specify)

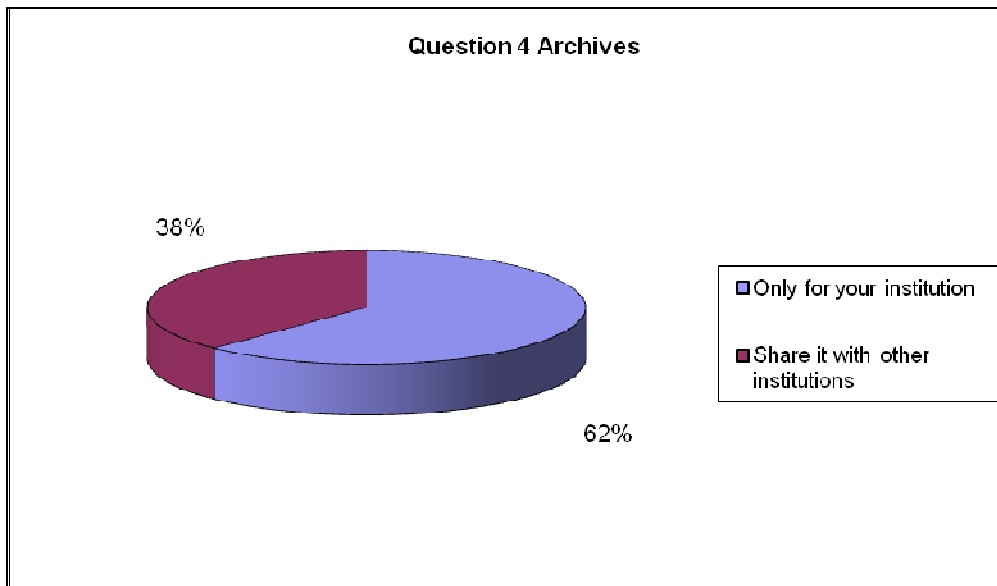


Digital preservation is a major issue and complex problem, which is the main reason why institutions try to cooperate. The survey contained a question trying to identify the types of institutions the respondents are (or will be) cooperating with in long-term preservation activities. The majority of the institutions answered that they cooperate on this issue with other memory institutions (31), for example other libraries, museums, archives. 18 institutions will cooperate with research institutions. About one third stated that they cooperate with Digital document producers (13) or SW developers/vendors and IT companies (16). Other institutions such as policy bodies or specialised archives were mentioned by 8 respondents.

4. Building and operation of a trusted digital repository is a big and expensive business. You will create and operate the repository ...

- 4 a) Only for your archive

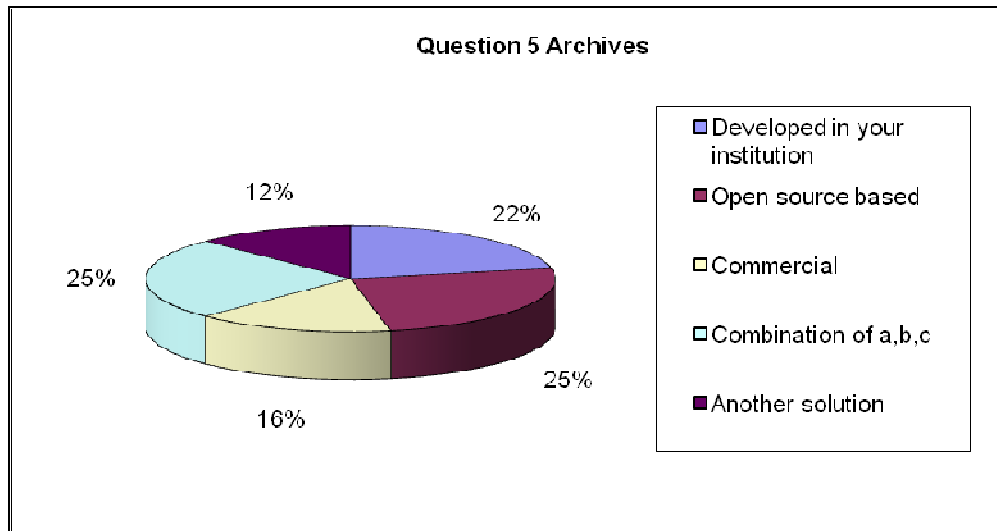
4 b) Share it with other institutions (please specify)



The building and operation/maintenance of digital repositories is a very complex and expensive business. It is clear that some kind of cooperation between institutions is very likely in this business. This tendency was much more obvious in the previous survey carried out with national libraries. In the case of archives the situation is slightly different. More than half of the respondents (62%) stated that their digital repository had already been created and is now operated exclusively for their own needs. The remaining institutions (38%) cooperate with other national/local institutions.

5. The system used for your digital library is (will be) ...

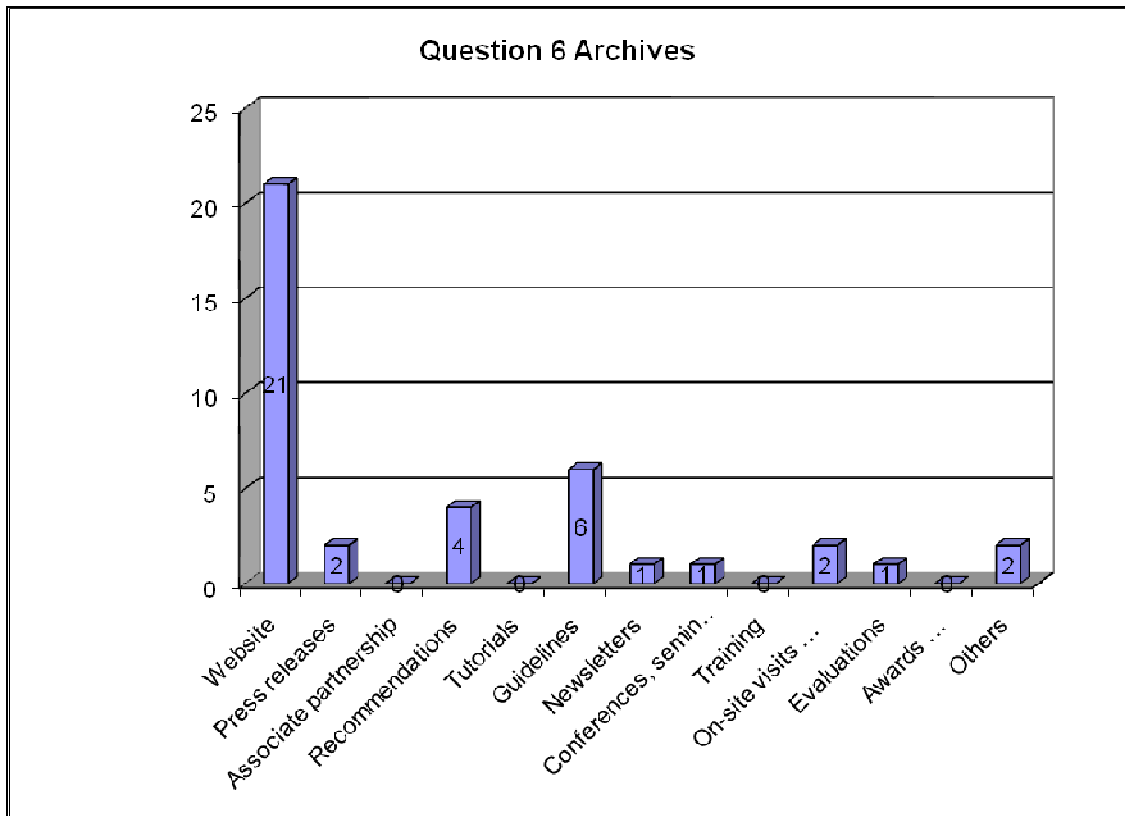
- 5 a) Developed in your archive
- 5 b) Open Source based
- 5 c) Commercial
- 5 d) Combination of 5a), b), c) (please specify)
- 5 e) Another solution (please specify)



In this question institutions were asked to provide some more detailed information about the system used for their digital repositories that is already in use or will be implemented in the near future. The majority of respondents (15) stated that they would like to implement some kind of Open Source system. Almost half of those interested in the Open Source system mentioned the possibility of using it in combination with a commercial system. 12 institutions (24%) answered that they plan to develop their own in-house system, 5 of which would like to develop it exclusively for their own institution.

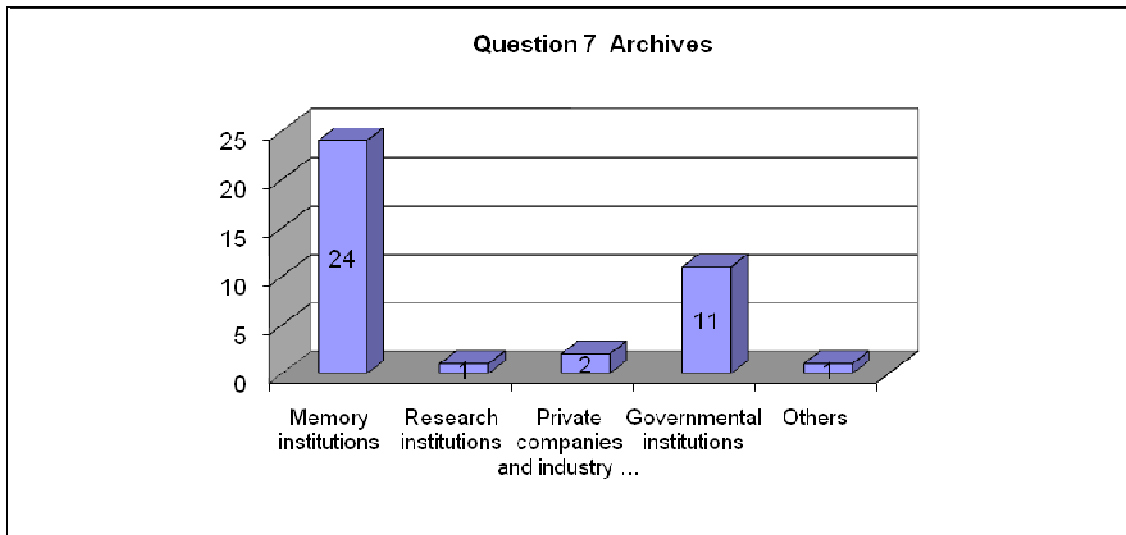
6. Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your archive?

- 6 a) Website
- 6 b) Press releases
- 6 c) Associate partnership
- 6 d) Recommendations
- 6 e) Tutorials
- 6 f) Guidelines
- 6 g) Newsletters
- 6 h) Conferences, seminars, workshops
- 6 i) Training
- 6 j) On-site visits and hands-on practice
- 6 k) Evaluations
- 6 l) Awards and prizes
- 6 m) Others – not listed here but desired (please specify)



In this question our respondents were asked to decide which of the outputs listed in the model of DPE dissemination they consider to be most relevant for them. The most frequent answers were: Websites (21) and Guidelines (6). Other frequent choices were Recommendations (4), Press releases (2), On-site visits (2), Others (2). Conferences, seminars, workshops had only 1 vote. No respondent selected the option Awards and Prizes and surprisingly there were no responses for Tutorials, Training and Associate partnership.

- 7. In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background for becoming fully operational and trusted national competence centres?**



The last question in the DPE survey concerned possibilities to create so-called national competence centres responsible for digital document preservation at the national level, which would cooperate with other similar centres in Europe, as referred to in FP7. The question concerned which kind of institution should play this role in each country. The overall majority of respondents (24) named memory institutions as appropriate, as also in the previous survey conducted with National Libraries. Governmental institutions were mentioned 11 times, Private companies and industry experienced in digital preservation twice. Both Research institutions and Others had 1 vote.

4.3 Survey on long-term preservation in research institutions

This is a report on the survey covering Research institutions (public and private, commercial and non-profit: e.g. Universities, Open Archives community). It represents the results of the survey, which was conducted on behalf of the DPE project during Spring 2007 and Autumn 2007 (second round). The aim of this survey was to describe the status of long-term preservation of digital documents in research institutions.

The questionnaire was distributed via e-mail lists:

JISC-REPOSITORIES@JISCMail.AC.UK

REPOSITORY-AUDIT-CERTIFICATION@JISCMail.AC.UK

DIGITAL-PRESERVATION@JISCMail.AC.UK

CERL mailing list

'diglib@infoserv.inist.fr'

'sigmed-l@asis.org'

'sigd-l@asis.org'

'eurchap@asis.org'

'asis-l@asis.org'

'dbworld@cs.wisc.edu'

JISC-METADATA@JISCMail.AC.UK

DLM forum

Nestor mailing list

DINI mailing list

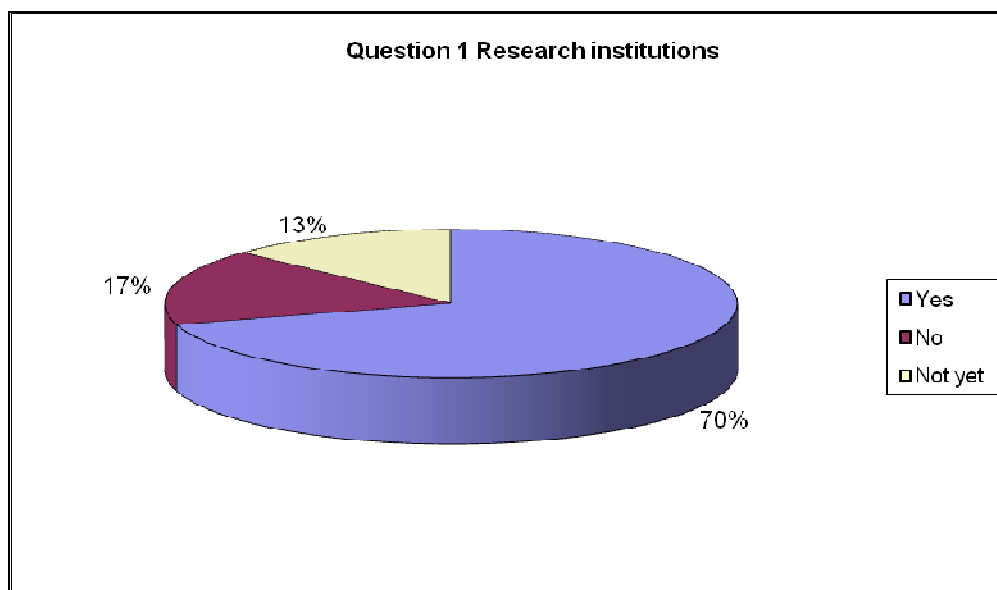
Because of the distribution method, we received some responses from outside the EU. We decided to include these in the report. In total, we received 54 completed questionnaires.

1. Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your institution?

1 a) Yes

1 b) No

1 c) Not yet (please specify when it will be)

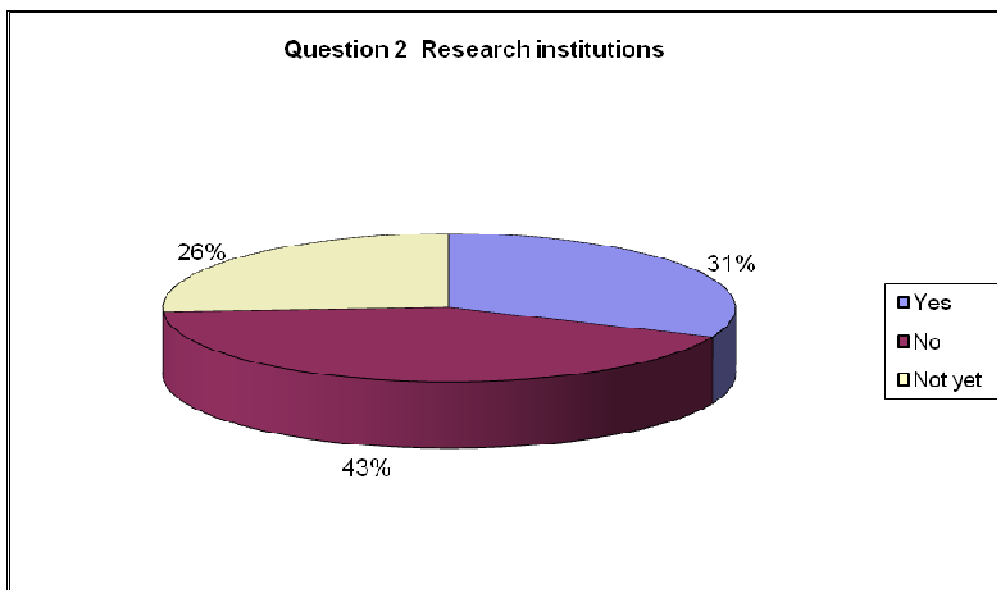


The majority of respondents addressed consider long-term preservation of digital documents as their key strategic priority. From 54 answers 38 were positive (70%) and 9 negative (17%). Seven institutions (13%) stated that this topic is not one of their strategic priorities, even if they count on addressing this issue in the future.

2. Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?

<http://www.ndk.cz/dokumenty/rlqnara-repositorieschecklist.pdf>

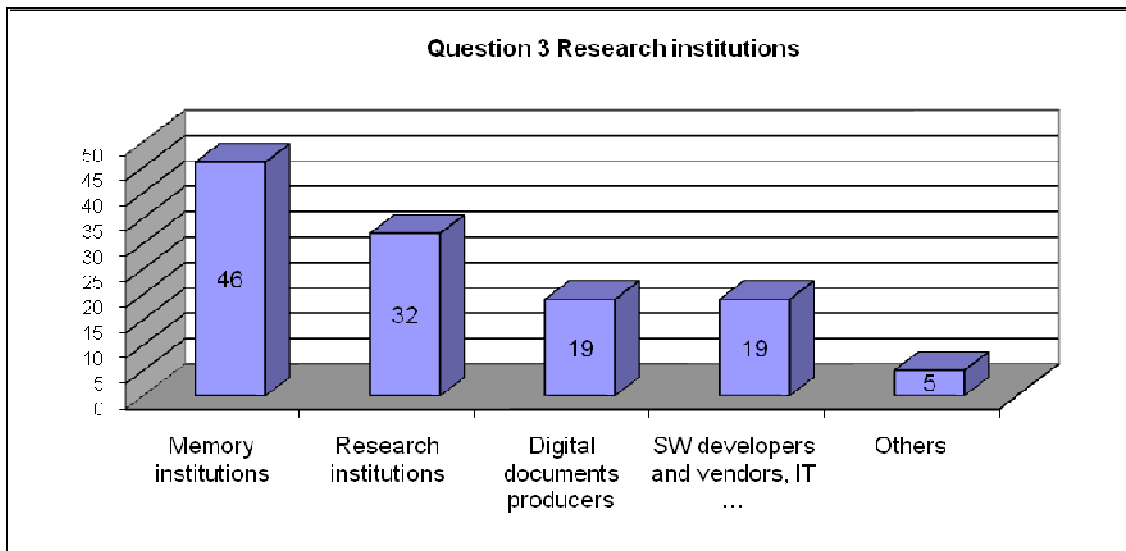
- 2 a) Yes
- 2 b) No
- 2 c) Not yet (please specify when you plan to have it)



The question whether respondents have (or will have) a trusted digital repository was answered “yes” by 17 institutions (31%). 23 institutions (43%) stated that they do not have a repository conforming to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories, but they are planning to have one in the near future (in 2-5 years’ time). Less than one third of institutions (26%) do not yet have a trusted repository.

3. Digital preservation is too big an issue for individual institutions to address independently. Your institution will cooperate in this area with ...

- 3 a) Memory institutions (libraries, museums, archives etc.)
- 3 b) Research institutions (schools, universities etc.)
- 3 c) Digital document producers (publishers, broadcasting etc.)
- 3 d) SW developers and vendors, IT, computer science
- 3 e) Others (please specify)

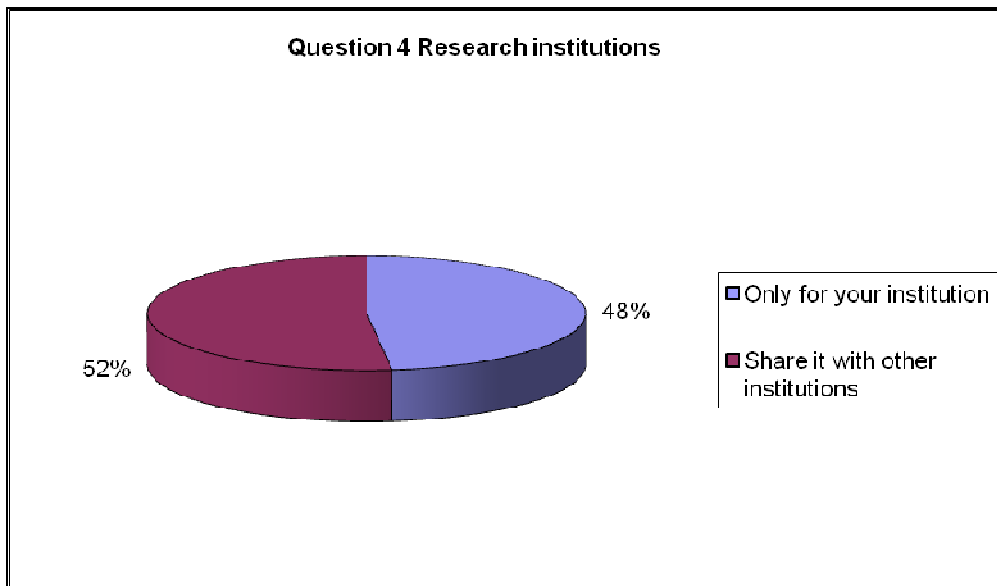


Digital preservation is a major issue and complex problem, which is the main reason why institutions try to cooperate. The survey contained a question trying to identify the types of institutions the respondents are (or will be) cooperating with in long-term preservation activities. The majority of the institutions answered that they cooperate on this issue with other memory institutions (46), for example other libraries, museums, archives. 32 institutions will cooperate with research institutions. 19 of them stated that they cooperate with Digital document producers and SW developers/vendors and IT companies. Other institutions such as policy bodies or specialised archives were mentioned by 5 respondents.

4. Building and operation of a trusted digital repository is a big and expensive business. You will create and operate the repository ...

4 a) Only for your institution

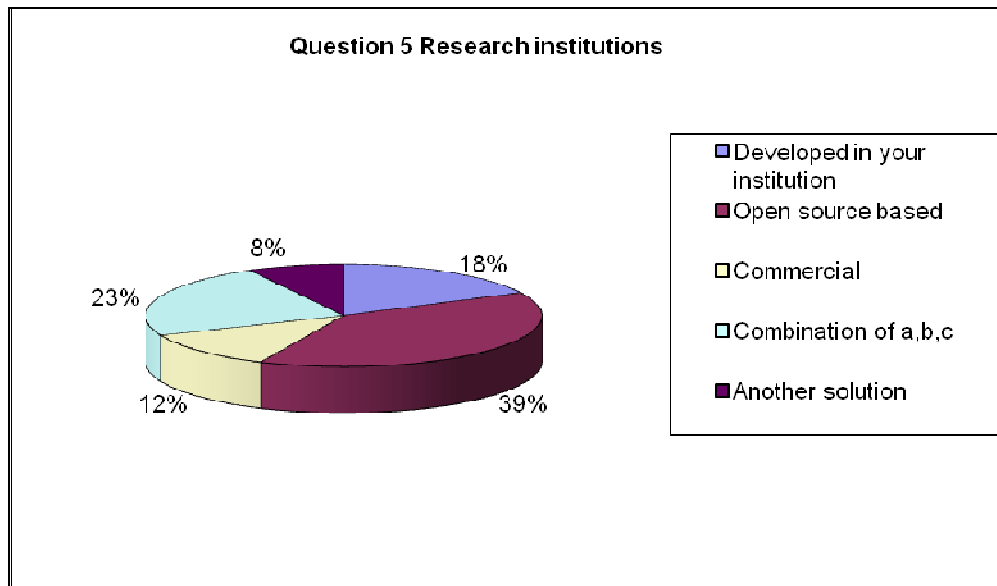
4 b) Share it with other institutions (please specify)



Building and operation/maintenance of digital repositories is a very complex and expensive business. Although some kind of cooperation between institutions is very likely in this business (especially in libraries), in the case of research institutions the situation is slightly different. 48% of respondents stated that their digital repository had been created and is now operated exclusively for their own needs. The remaining institutions (52%) cooperate with other national/local institutions.

5. The system used for your digital library is (will be) ...

- 5 a) Developed in your institution
- 5 b) Open Source based
- 5 c) Commercial
- 5 d) Combination of 5a), b), c) (please specify)
- 5 e) Another solution (please specify)

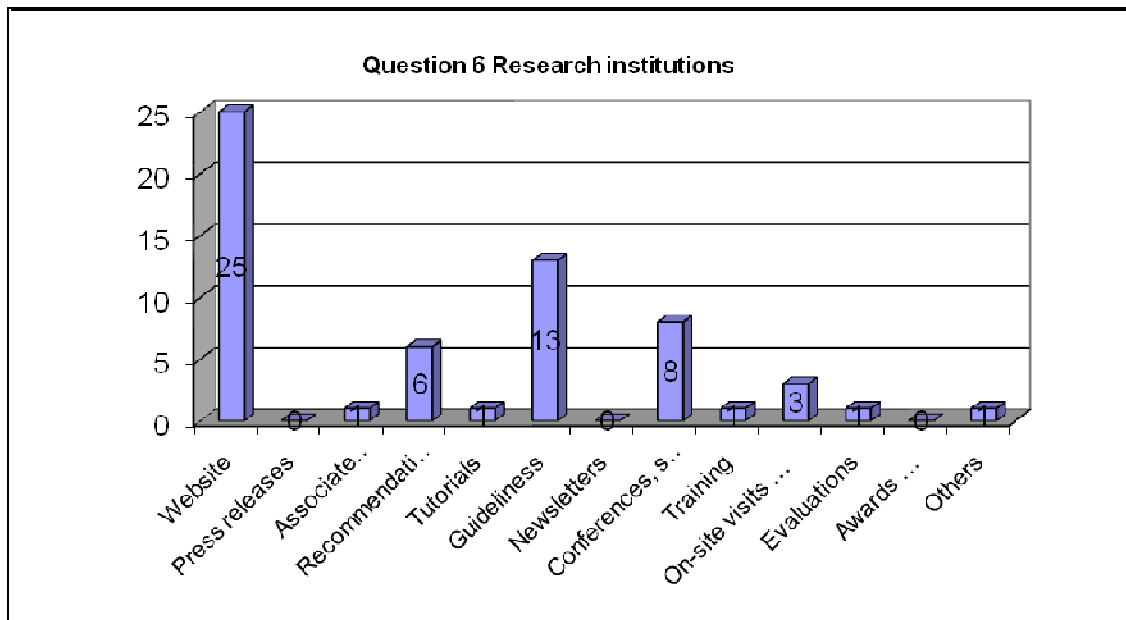


In this question institutions were asked to provide some more detailed information about the digital repository system that is already in use or will be implemented in the near future. The majority of respondents (20) stated that they would like to implement some kind of Open Source system. In seven cases those interested in the Open Source system added the possibility of using it in combination with a commercial system. 15 institutions answered that they plan to develop their own in-house system, 6 of which would like to develop it exclusively for their own institution.

6. Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your institution?

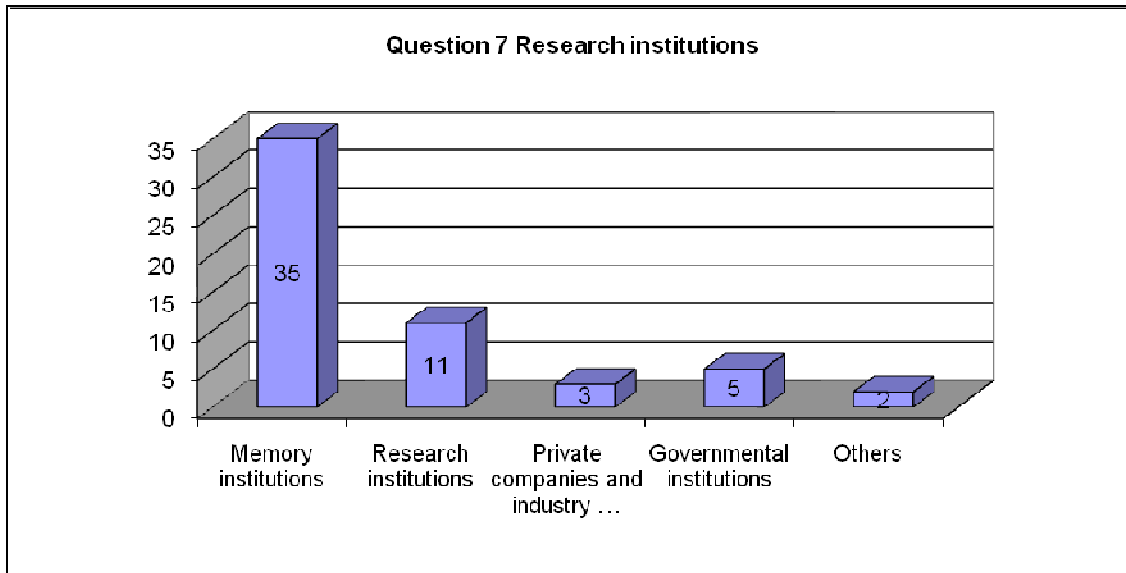
- 6 a) Website
- 6 b) Press releases
- 6 c) Associate partnership
- 6 d) Recommendations
- 6 e) Tutorials
- 6 f) Guidelines
- 6 g) Newsletters
- 6 h) Conferences, seminars, workshops
- 6 i) Training
- 6 j) On-site visits and hands-on practice
- 6 k) Evaluations
- 6 l) Awards and prizes

6 m) Others – not listed here but desired (please specify)



In this question our respondents were asked to decide which of the outputs listed in the model of DPE dissemination they consider to be most relevant for them. The most frequent answers were: Websites (25) and Guidelines (13). Other frequent choices were Conferences, seminars, workshops (8), Recommendations (6), On-site visits (3). There was only one vote for Others, Tutorials, Associate partnership, Training and Evaluations. No respondent selected the option Awards and prizes and, surprisingly, there were none for Newsletters and Press releases.

7. In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background for becoming fully operational and trusted national competence centres?



The last question in the DPE survey concerned possibilities to create so-called national competence centres responsible for digital document preservation at the national level, which would be cooperating with other similar centres in Europe, as referred to in FP7. The question was which kind of institution should play this role in each country. The overall majority of respondents (35) named memory institutions as appropriate. Research institutions were mentioned 11 times, Governmental institutions 5 times, Private companies and industry 3 times and Others 2 times.

4.4 Survey on long-term preservation in ICT companies and Media

This is a report on the survey covering Industry (ICT companies and Media). It represents the results of the survey conducted on behalf of the DPE project during Autumn 2007. The aim of this survey was to describe the status of long-term preservation of digital documents in this area.

The questionnaire was distributed via the following e-mail lists:

'sigmed-l@asis.org'

'sigdl-l@asis.org'

'eurchap@asis.org'

'asis-l@asis.org'

'dbworld@cs.wisc.edu'

JISC-METADATA@JISCMAIL.AC.UK

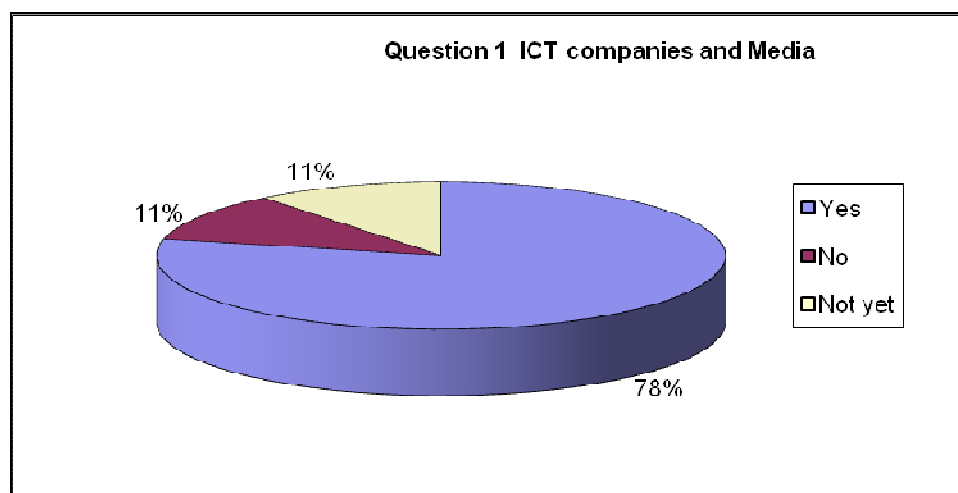
JISC-REPOSITORIES@JISCMAIL.AC.UK
REPOSITORY-AUDIT-CERTIFICATION@JISCMAIL.AC.UK
DIGITAL-PRESERVATION@JISCMAIL.AC.UK
community@delos.info
DLM forum
Nestor mailing list
DINI mailing list

Besides using mailing lists, the survey was also sent to specific individuals (email addresses) with the kind request to complete the survey questionnaire. We used email addresses of people registered in the DPE user community list database, and FRD (Fondazione Rinascimento Digitale - Florence, Italy) provided us with relevant contacts from the participant list of the 'Cultural Heritage on line' conference (December 2006). This means of addressing people directly with a kind request was much more successful in achieving a reasonable number of responses.

Because of the distribution method, we received some responses from outside the EU. We decided to include these in the report. In total, we received 28 completed questionnaires, which is not a large number, but we have to consider how difficult it is to address this area.

1. Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your institution?

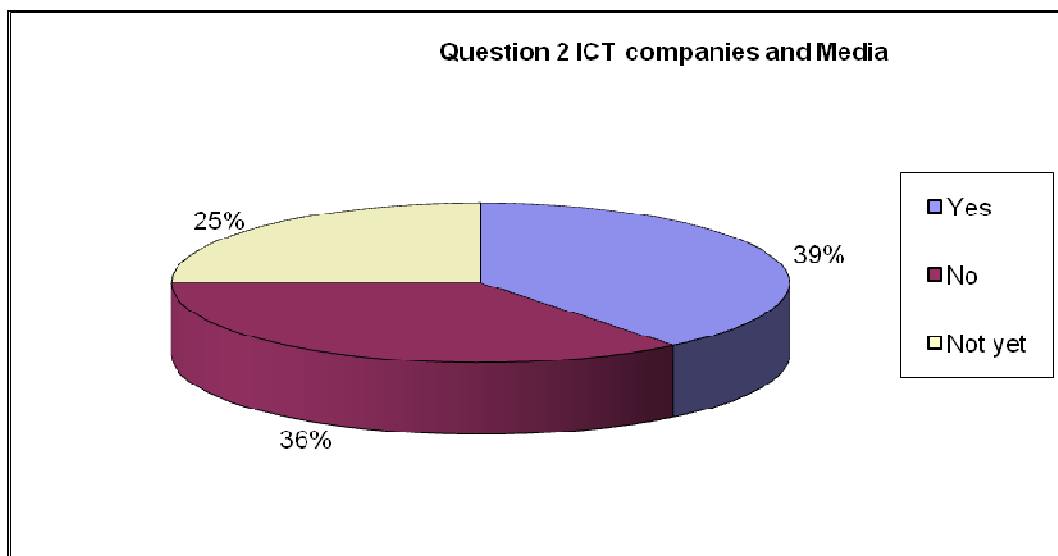
- 1 a) Yes
- 1 b) No
- 1 c) Not yet (please specify when it will be)



The vast majority of respondents considered the long-term preservation of digital documents as a key strategic priority. From 28 answers, 22 were positive (78%) and 3 negative (11%). Three institutions (11%) stated that this topic is not one of their strategic priorities, even if they count on addressing this issue in the future.

2. Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?
<http://www.ndk.cz/dokumenty/rlgnara-repositorieschecklist.pdf>

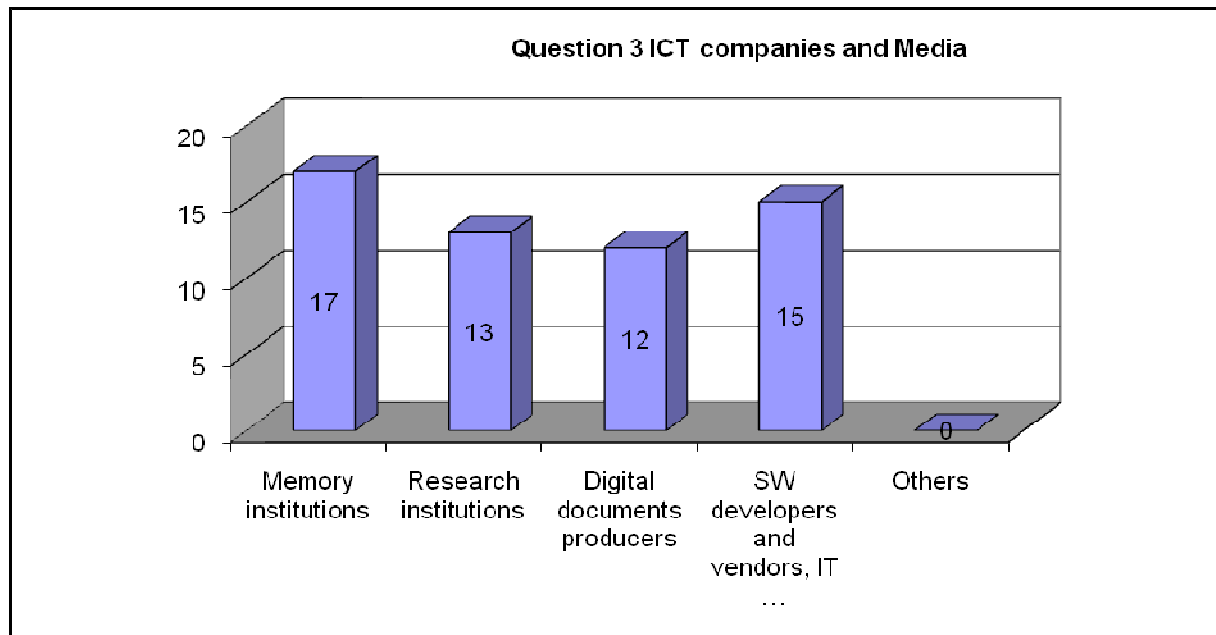
- a) Yes
- 2 b) No
- 2 c) Not yet (please specify when you plan to have it)



The question whether respondents have (or will have) a trusted digital repository was answered “yes” by 11 institutions (39%). Seven institutions (25%) stated that they don't have a repository conforming to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories, but are planning to have one in the near future (in 2-5 years' time). More than one third of institutions (10) do not yet have a trusted repository and are not even planning to have one.

3. Digital preservation is too big an issue for individual institutions to address independently. Your institution will cooperate in this area with ...

- 3 a) Memory institutions (libraries, museums, archives etc.)
- 3 b) Research institutions (schools, universities etc.)
- 3 c) Digital document producers (publishers, broadcasting etc.)
- 3 d) SW developers and vendors, IT, computer science
- 3 e) Others (please specify)

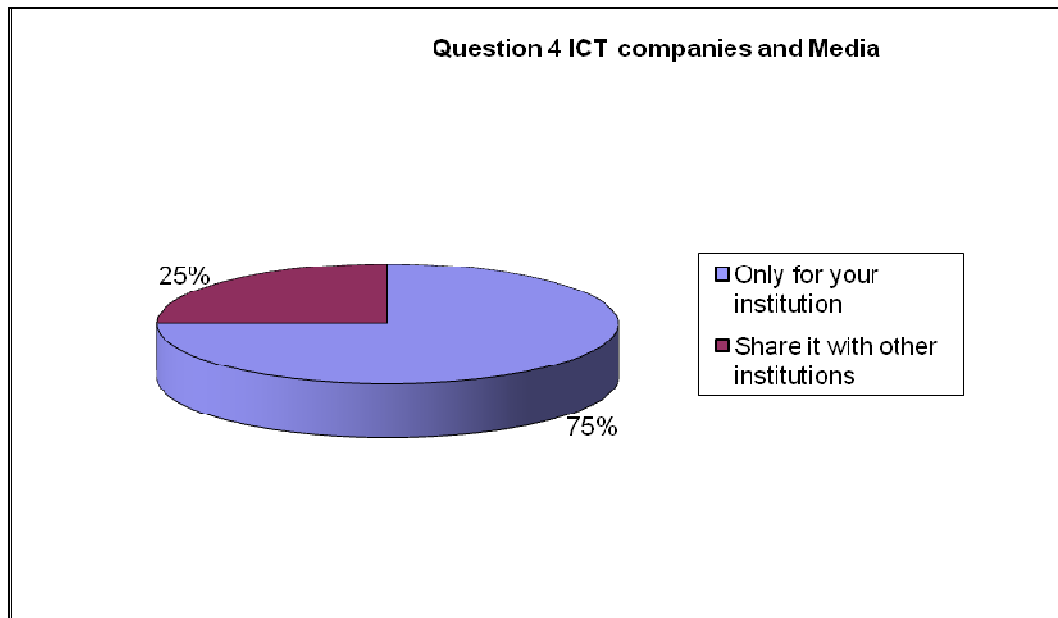


Digital preservation is a very big issue and complex problem, which is the main reason why institutions are trying to cooperate. The survey contained a question trying to identify the types of institutions the respondents are (or will be) cooperating with in long-term preservation activities. The majority of the institutions answered that they cooperate on this issue with other memory institutions (17), for example other libraries, museums, archives. Not surprisingly, the second group mentioned were SW developers and vendors (15). 13 institutions cooperate with research institutions, 12 of which stated that they cooperate with Digital document producers. Other institutions such as policy bodies or specialised archives were not selected by a single respondent.

4. Building and operation of a trusted digital repository is a big and expensive business. You will create and operate the repository ...

4 a) Only for your institution

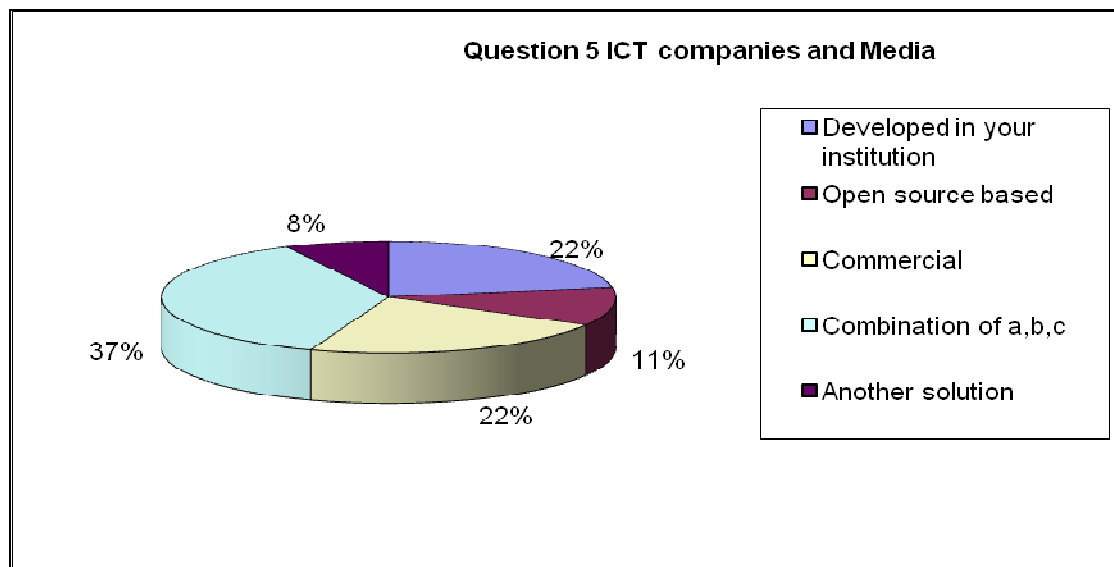
4 b) Share it with other institutions (please specify)



Building and operation/maintenance of digital repositories is a very complex and expensive business. Although some kind of cooperation between institutions is very likely in this field (especially in libraries), in the case of ICT companies and Media it is completely different. Twenty-one (75%) respondents stated that their digital repository had been created and is now operated exclusively for their own needs. The remaining institutions (7) cooperate with other national/local institutions. This is not surprising as the majority of these companies are private companies, saving and concealing their crucial data to keep business secrets and know-how.

5. The system used for your digital library is (will be) ...

- 5 a) Developed in your institution
- 5 b) Open Source based
- 5 c) Commercial
- 5 d) Combination of 5a), b), c) (please specify)
- 5 e) Another solution (please specify)



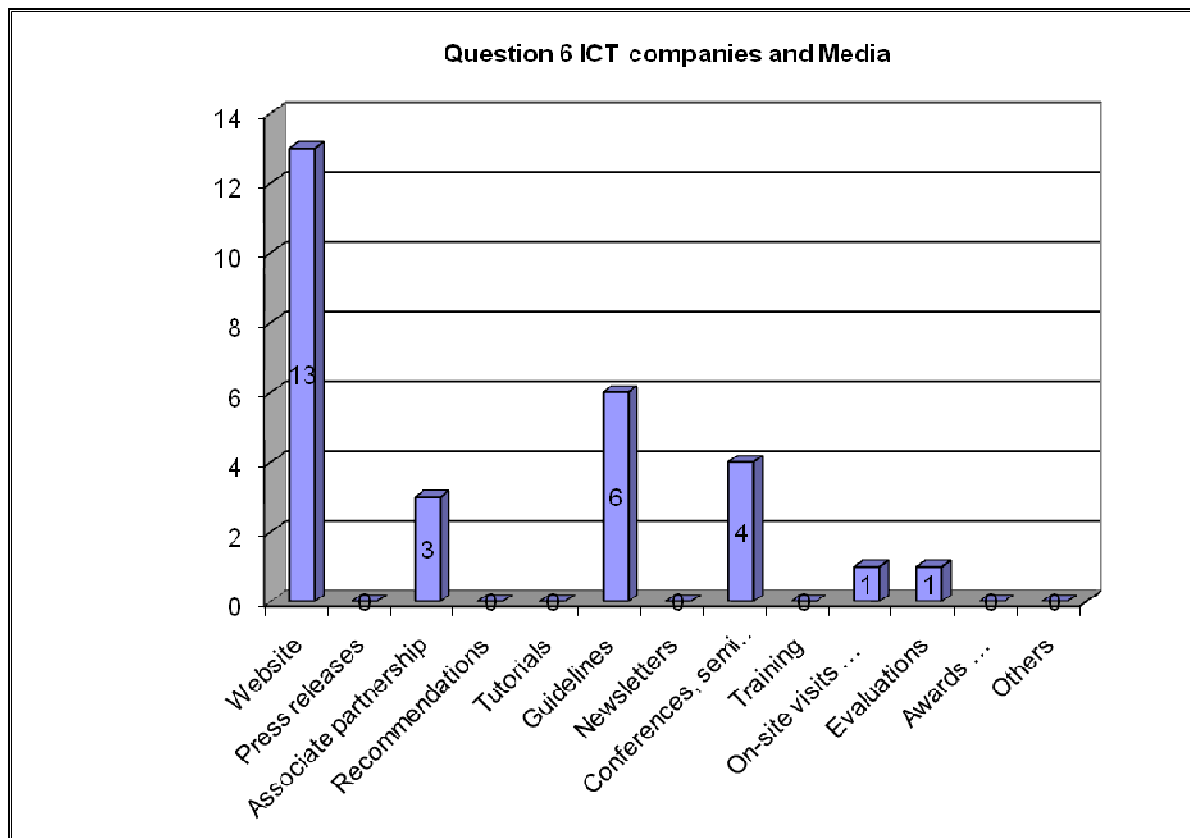
In this question, institutions were asked to provide some more detailed information about the system used for their digital repositories that is already in use or will be implemented in the near future. Twelve respondents stated that they would like to implement some kind of Open Source system. In five cases those interested in an Open Source system added the possibility of using it in combination with a commercial system. Thirteen institutions answered that they plan to develop their own in-house system. Eleven would like to have a commercial system, 5 of them in combination with Open Source SW and/or SW developed in their institution.

6. Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your company?

- 6 a) Website
- 6 b) Press releases
- 6 c) Associate partnership
- 6 d) Recommendations
- 6 e) Tutorials
- 6 f) Guidelines
- 6 g) Newsletters
- 6 h) Conferences, seminars, workshops
- 6 i) Training
- 6 j) On-site visits and hands-on practice
- 6 k) Evaluations

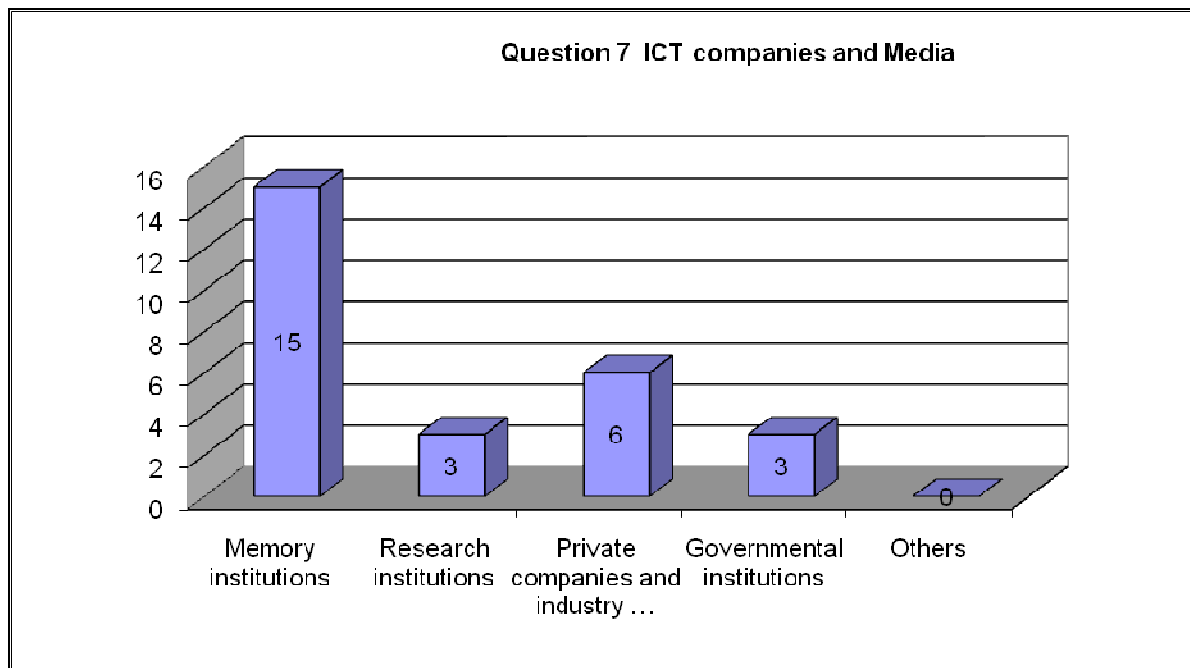
6 l) Awards and prizes

6 m) Others – not listed here but desired (please specify)



In this question our respondents were asked to decide which of the outputs listed in the model of DPE dissemination they consider to be most relevant for them. The most frequent answers were: Websites (13) and Guidelines (6). Other frequent choices were Conferences, seminars, workshops (4) and Associate partnership (3). There was only one vote for On-site visits and Evaluations. Other possibilities are considered as not appropriate (0 votes).

7. In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background for becoming fully operational and trusted national competence centres?



The last question in the DPE survey concerned possibilities to create so-called national competence centres responsible for digital document preservation at the national level, which would be cooperating with other similar centres in Europe, as mentioned in FP7. The question was which kind of institution should play this role in each country. The vast majority of respondents (15) named memory institutions as appropriate. Private companies and industry were mentioned 6 times, Research institutions and Governmental institutions both 3 times.

4.5 Survey on long-term preservation in 'Others'

By 'Others' we mean Non-governmental institutions and organisations (e.g. IFLA, ICA, UNESCO etc.); Related projects, coalitions, organisations and initiatives (e.g. CASPAR, PLANETS, PADI etc.) and Governmental institutions and local authorities – according to the DPE Dissemination plan targets. The number of responses to the DPE questionnaire received was not high, because we did not address this area directly, and our answers are only a by-product of the above surveys conducted during Autumn 2007.

The questionnaire was distributed via the following e-mail lists:

'sigmed-l@asis.org'

'sigdl-l@asis.org'

'eurchap@asis.org'

'asis-l@asis.org'

'dbworld@cs.wisc.edu'

JISC-METADATA@JISCMail.AC.UK

JISC-REPOSITORIES@JISCMail.AC.UK

REPOSITORY-AUDIT-CERTIFICATION@JISCMail.AC.UK

DIGITAL-PRESERVATION@JISCMail.AC.UK

community@delos.info

DLM forum

Nestor mailing list

DINI mailing list

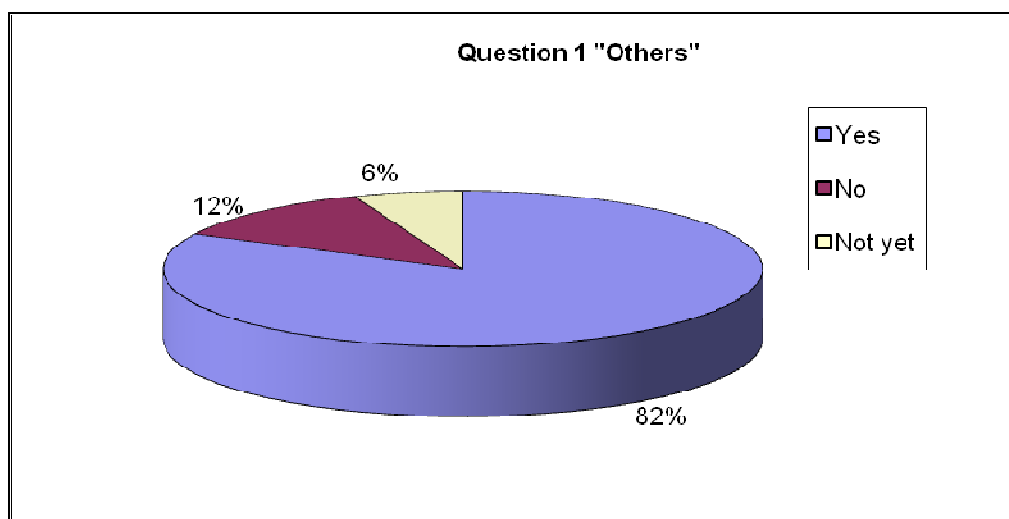
Because of the distribution method, we received some responses from outside the EU. We decided to include them in the report. In total, we have 17 completed questionnaires from this area.

1. Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your institution?

1 a) Yes

1 b) No

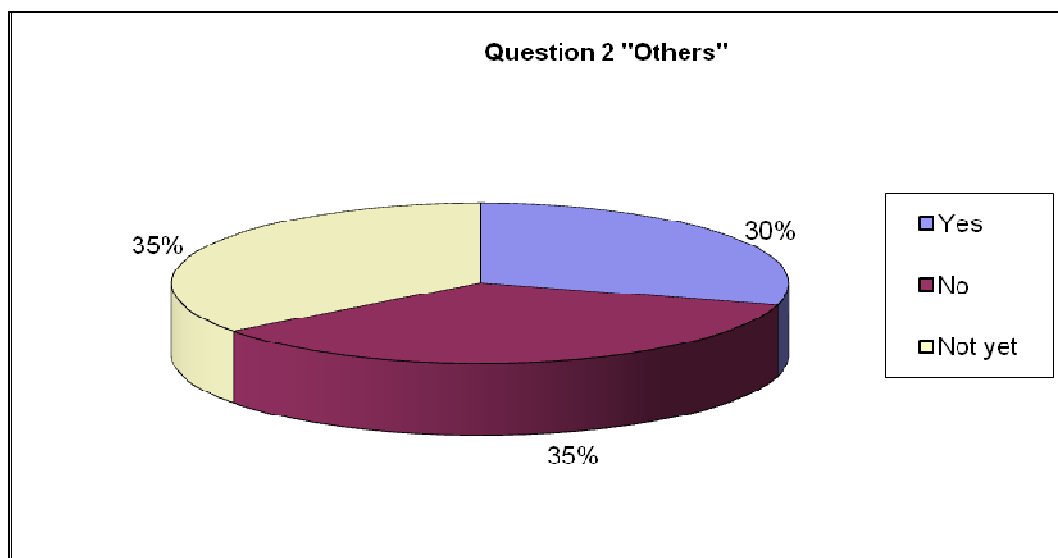
1 c) Not yet (please specify when it will be)



The vast majority of respondents consider the long-term preservation of digital documents as their key strategic priority. From 17 answers, 14 were positive (82%) and 2 negative (12%). One institution (6%) stated that this topic is not one of their strategic priorities, even if they counted on addressing the issue in the future.

2. Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?
<http://www.ndk.cz/dokumenty/rlgnara-repositorieschecklist.pdf>

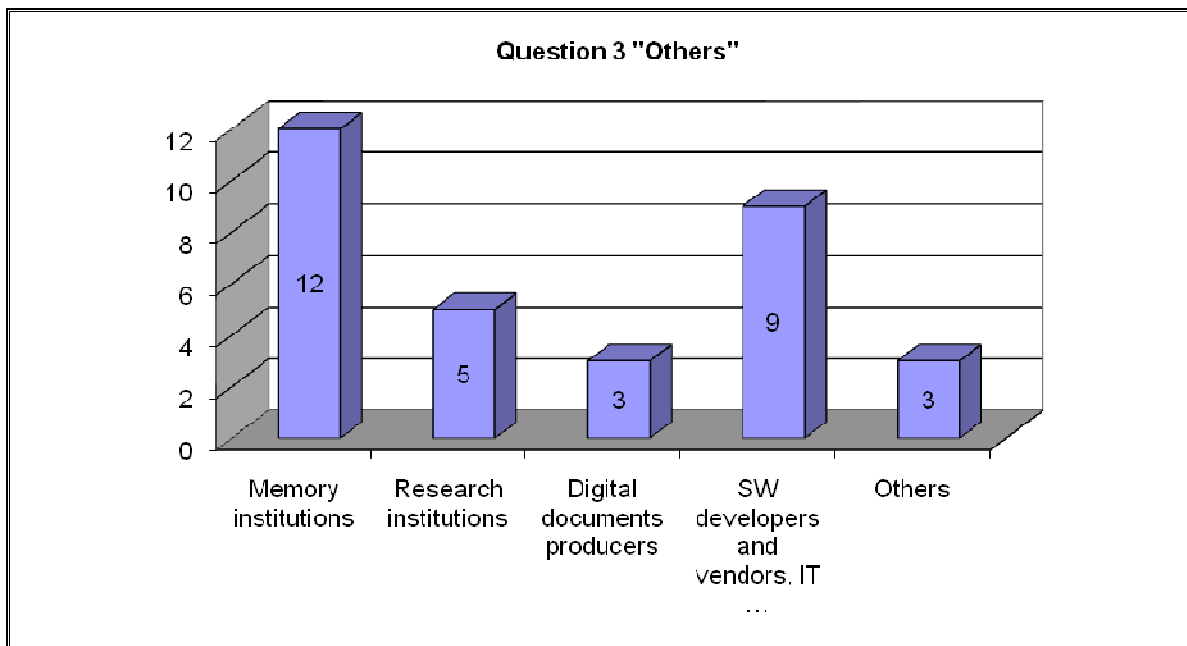
- a) Yes
- 2 b) No
- 2 c) Not yet (please specify when you plan to have it)



The question whether respondents have (or will have) a trusted digital repository was answered "yes" by 5 institutions (30%). Six institutions (35%) stated that they don't have a repository conforming to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories, but they are planning to have one in the near future (in 2-5 years' time). Similarly, 5 institutions don't have a trusted repository yet and are not even planning to have one.

3. Digital preservation is too big an issue for individual institutions to address independently. Your institution will cooperate in this area with ...

- 3 a) Memory institutions (libraries, museums, archives etc.)
- 3 b) Research institutions (schools, universities etc.)
- 3 c) Digital document producers (publishers, broadcasting etc.)
- 3 d) SW developers and vendors, IT, computer science
- 3 e) Others (please specify)

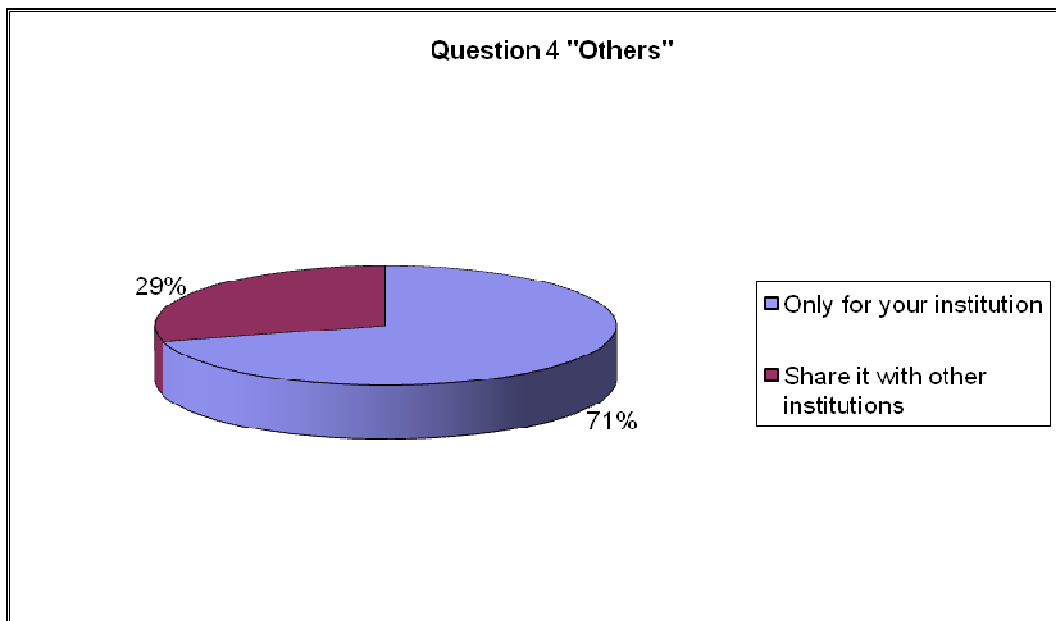


Digital preservation is a very big issue and complex problem, which is the main reason why institutions are trying to cooperate. The survey contained a question trying to identify the types of institutions the respondents are (or will be) cooperating with in long-term preservation activities. The majority of the institutions answered that they cooperate on this issue with other memory institutions (12), for example other libraries, museums, archives. The second good candidate for cooperation was SW developers and vendors (9). Five institutions will cooperate with research institutions, 3 of them stating that they cooperate with Digital document producers and the same number with Others (policy bodies etc.).

4. Building and operation of a trusted digital repository is a big and expensive business. You will create and operate the repository ...

4 a) Only for your institution

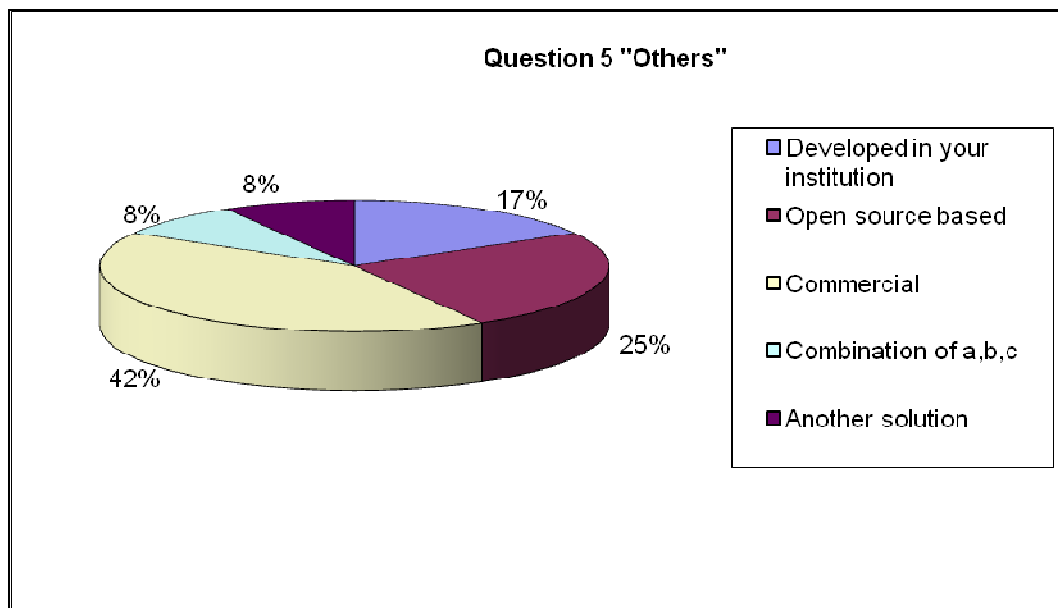
4 b) Share it with other institutions (please specify)



Building and operation/maintenance of digital repositories is a very complex and expensive business. Although some kind of cooperation between institutions is very likely in this field (especially in libraries), in the case of this particular area it is different. Twelve (71%) respondents stated that their digital repository was already created and is now operated exclusively for their own needs. The remaining institutions (5) cooperate with other national/local institutions.

5. The system used for your digital library is (will be) ...

- 5 a) Developed in your institution
- 5 b) Open Source based
- 5 c) Commercial
- 5 d) Combination of 5a), b), c) (please specify)
- 5 e) Another solution (please specify)



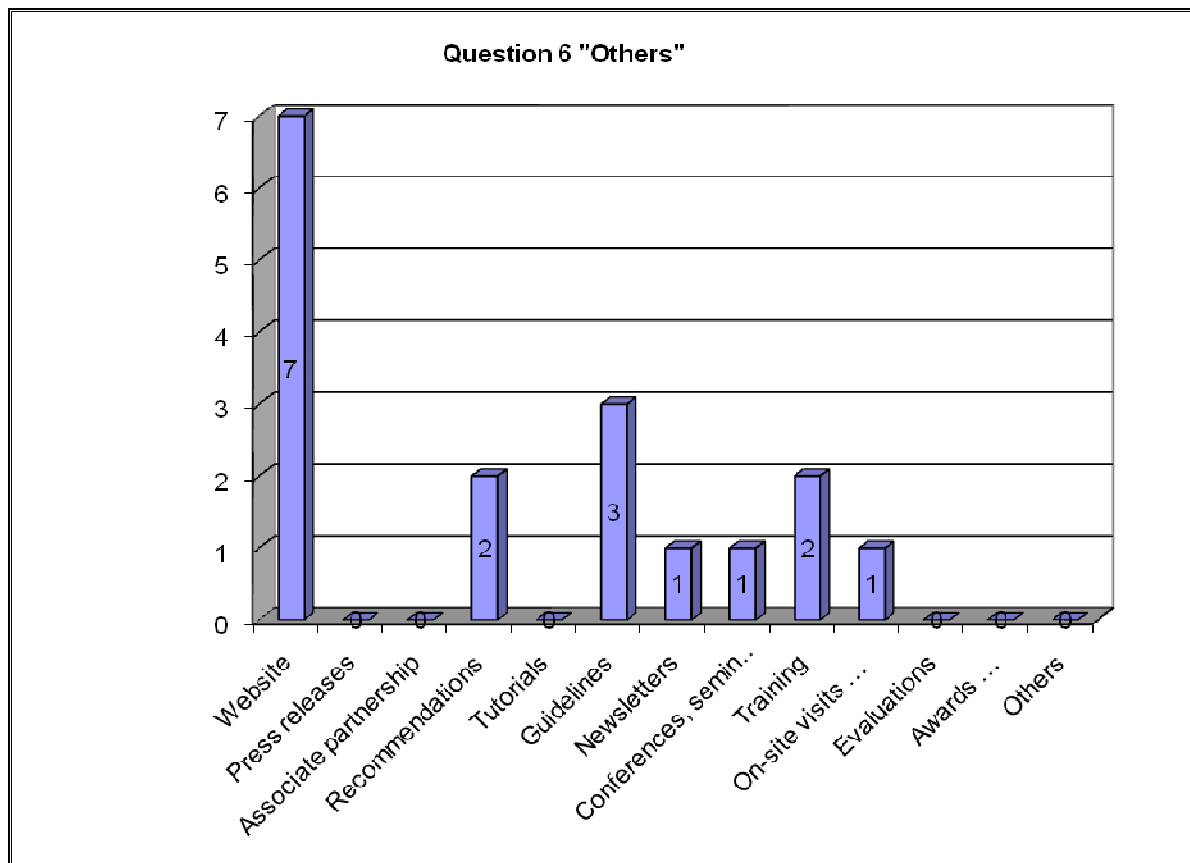
In this question institutions were asked to provide some more detailed information about the system used for their digital repositories that is already in use or will be implemented in the near future. Four respondents stated that they would like to implement some kind of Open Source system. In one case those interested in an Open Source system added the possibility of using it in combination with a commercial system. Six institutions answered that they plan to use a commercial system.

6. Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your institution?

- 6 a) Website
- 6 b) Press releases
- 6 c) Associate partnership
- 6 d) Recommendations
- 6 e) Tutorials
- 6 f) Guidelines
- 6 g) Newsletters
- 6 h) Conferences, seminars, workshops
- 6 i) Training
- 6 j) On-site visits and hands-on practice
- 6 k) Evaluations

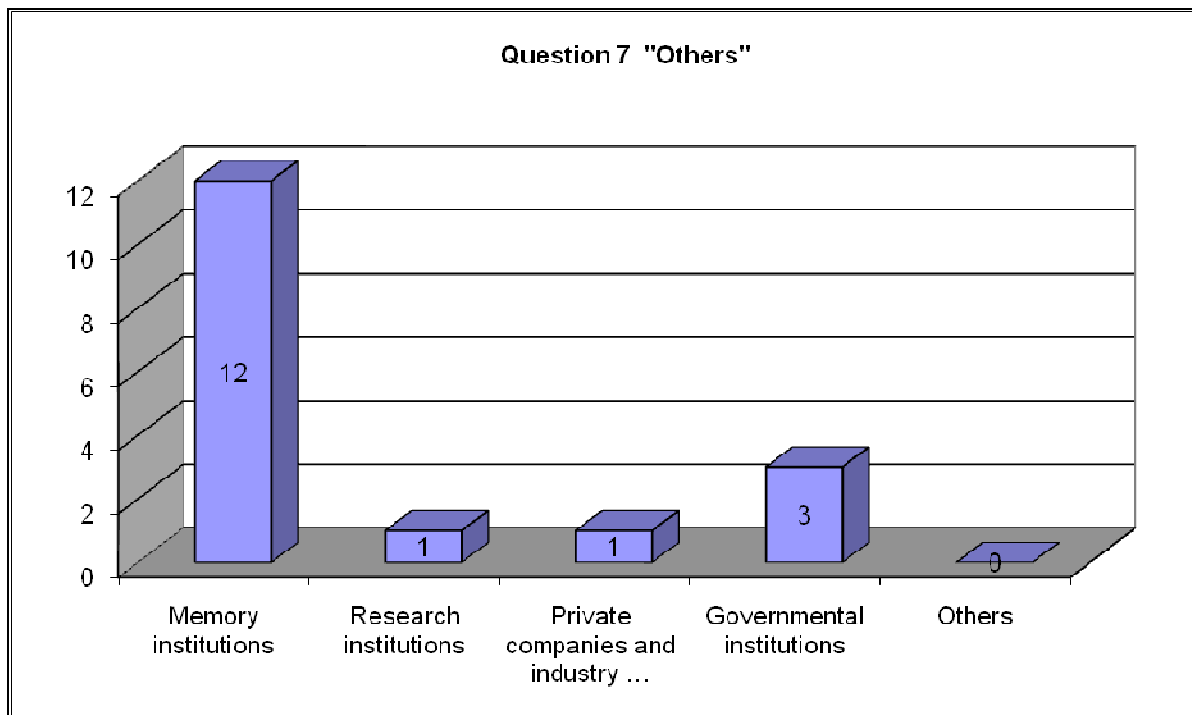
6 l) Awards and prizes

6 m) Others – not listed here but desired (please specify)



In this question our respondents were asked to decide which of the outputs listed in the model of DPE dissemination they consider to be most relevant for them. The most frequent answers were: Websites (7) and Guidelines (3). Other frequent choices were Recommendations and Training (both 2), then Conferences, seminars, workshops, Newsletters and On-site visits all with one vote. The remainder did not receive any score.

7. In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background for becoming fully operational and trusted national competence centres?



The last question in the DPE survey concerned possibilities to create so-called national competence centres responsible for digital document preservation at the national level, which would be cooperating with other similar centres in Europe, as mentioned in FP7. The question was which kind of institution should play this role in each country. The overall majority of respondents (12) named memory institutions as appropriate. Governmental institutions were selected 3 times. Research institutions and Private companies and industry were both mentioned one time.

4.6 Comparison of all survey results

The answers to some questions are very similar for both memory and research institutions. There are also some differences. Some of these could have been predicted, while others are rather surprising.

1. Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your institution?

We received positive answers from 83% of national libraries, 78% of archives, 70% of research institutions, 78% of ICT companies and media and 82% of 'Others'. This means that long-term preservation is a key strategic priority for all targeted institutions without any measure of doubt.

2. Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?

29% of national libraries answered yes, 9% of them answered no and 62% answered not yet. The 62% 'not yet' is the highest value of all targets. All the others answered 'not yet' in around 30% of their answers. This could show that libraries are more aware of the importance of having a trusted digital repository. 43% of Archives, 31% of Research institutions, 39% of ICT companies and Media and 29% of 'Others' stated they have a trusted digital repository.

3. Digital preservation is too big an issue for individual institutions to address independently. Your institution will cooperate in this area with ...

Memory institutions (libraries, museums, archives etc.) are the most popular institutions for cooperation for all our target groups. They are followed in second place by research institutions (except for ICT companies and Media and 'Others'). Digital document producers are listed in third place for Libraries and Research institutions. SW developers and vendors are important for Archives and, not surprisingly, especially for ICT companies and Media (second place) and for 'Others' (third place). In all the charts it is obvious that it is very important and comfortable to cooperate with institutions from the same area as the institution seeking cooperation.

4. Building and operation of a trusted digital repository is a big and expensive business. You will create and operate the repository ...

This question was also oriented towards cooperation, with a special focus on sharing the digital repository. The answers were very different. Only 20% of national libraries plan to create and operate a repository exclusively for themselves, while 38% of Archives plan to have one just for their own institution. It is not surprising that the situation is completely different for Research institutions, ICT companies and 'Others', where 48%, 75% and 71%, respectively, of these want to have one exclusively for their own use. Their aim is not public access, but just to store their data safely. The difference could also lie in the long experience of sharing repositories of traditional documents among libraries and coordinated digitisation programmes focused on preservation of traditional documents, coordinated archiving of the Web etc.

5. The system used for your digital library is (will be) ...

It was not surprising that national libraries, having fewer programmers and research resources than research institutions, plan combined solutions with a relatively high percentage (52%) opting for commercial systems. Research institutions rely mainly (38%) on Open Source solutions, which was also expected. All charts except Libraries are more or less the same. About 20% opt for systems developed in their own institution. Only 11% of ICT companies and Media would like to have an Open Source system, the reason being more than clear. On the other hand, Research institutions were 38% for Open Source systems,

'Others' and Archives 25% and 24%. The outcome for Commercial systems is interesting. For ICT companies and Media and for 'Others', it is one of the most favoured options.

Note: From another part of this analysis it is obvious that the Open Source area is fairly well established and documented. On the other hand, there is scarcely any choice in commercial systems offering preservation functionality now or in the near future. This situation of almost no competition and limited market results in astronomical prices of the products and their maintenance.

6. Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your institution?

The DPE website is the favourite output for all our targets. Only Libraries ranked Conferences, seminars and workshops at the same level of importance. Conferences, seminars and workshops are important outputs for ICT companies and Media and for Research institutions. On the other hand, Archives do not seem to be interested in Conferences, seminars and workshops. All DPE targets are interested in Guidelines and Recommendations. Newsletters are not popular at all (except for libraries). Only ICT companies and Media showed any interest in Associate partnership as a DPE output.

7. In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background for becoming fully operational and trusted national competence centres?

Memory institutions are the leading candidates mentioned by all the institutions. Research institutions were ranked second by Libraries and by Research institutions themselves. Governmental institutions are significant for almost all DPE targets, and especially for Archives and Libraries. Only ICT companies notably stated that private companies and industry could be a good candidate to become a competence centre.

5. PART III: TECHNOLOGICAL SOLUTIONS AVAILABLE FOR DIGITAL PRESERVATION

5.1 List and characteristics of systems mostly used for digital preservation

The DPE project deals mainly with digital preservation and thus only those systems having an operating preservation functionality or promising this for the near future were included.

5.1.1 Commercial

- DIAS (Digital Information Archiving System)
- DPS (Digital Preservation System)

DIAS (Digital Information Archiving System)

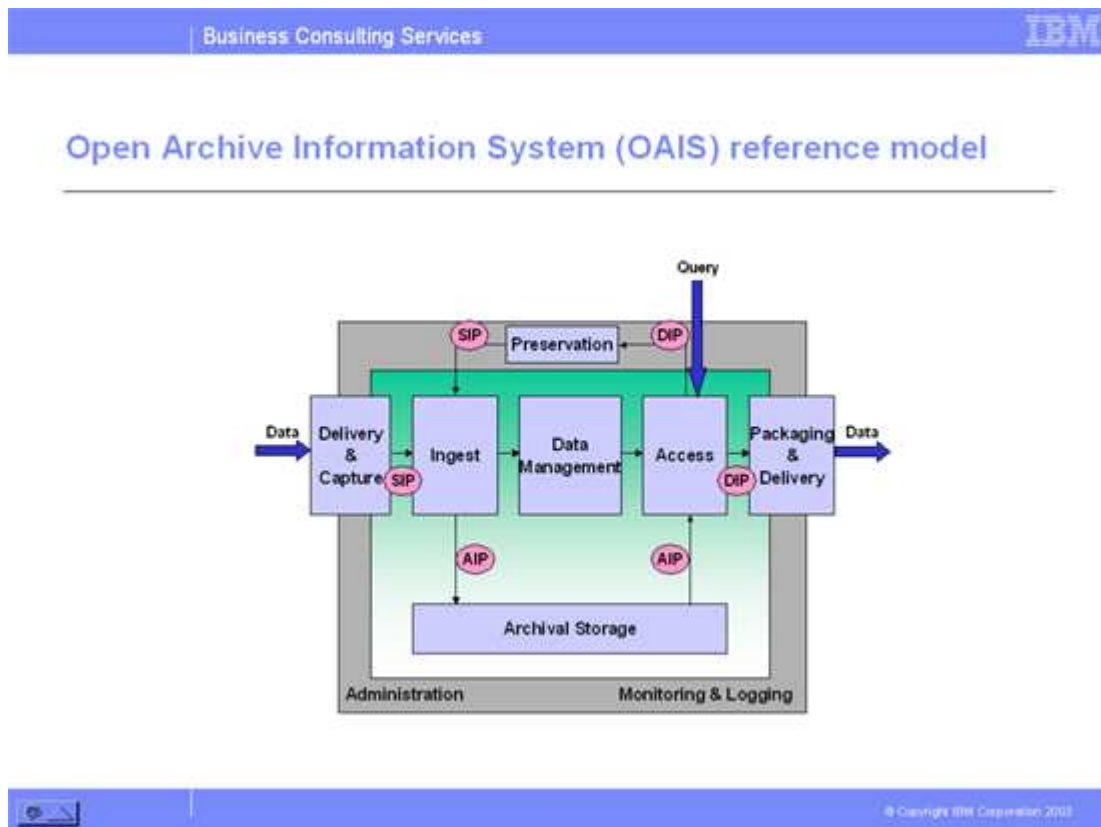
Vendor/producer

- IBM

URL: <http://www-5.ibm.com/nl/dias/>

Description

The DIAS (Digital Information Archiving System) solution provides a flexible and scalable open deposit library solution for storing and retrieving massive amounts of electronic documents and multimedia files. It conforms to the ISO Reference OAIS standard and supports physical and logical digital preservation.



Technology and Value Proposition

The DIAS solution allows manual as well as automated ingest of digital information (assets) into the system. Once the asset has been successfully stored it will be maintained and preserved. The preservation functionality gives signals for stored assets that must be converted or migrated to keep them available for use. Stored assets can be accessed either via a web-based interface (for assets with standard file types) or via a specific work environment on a Reference Workstation.

Current DIAS-Core Highlights

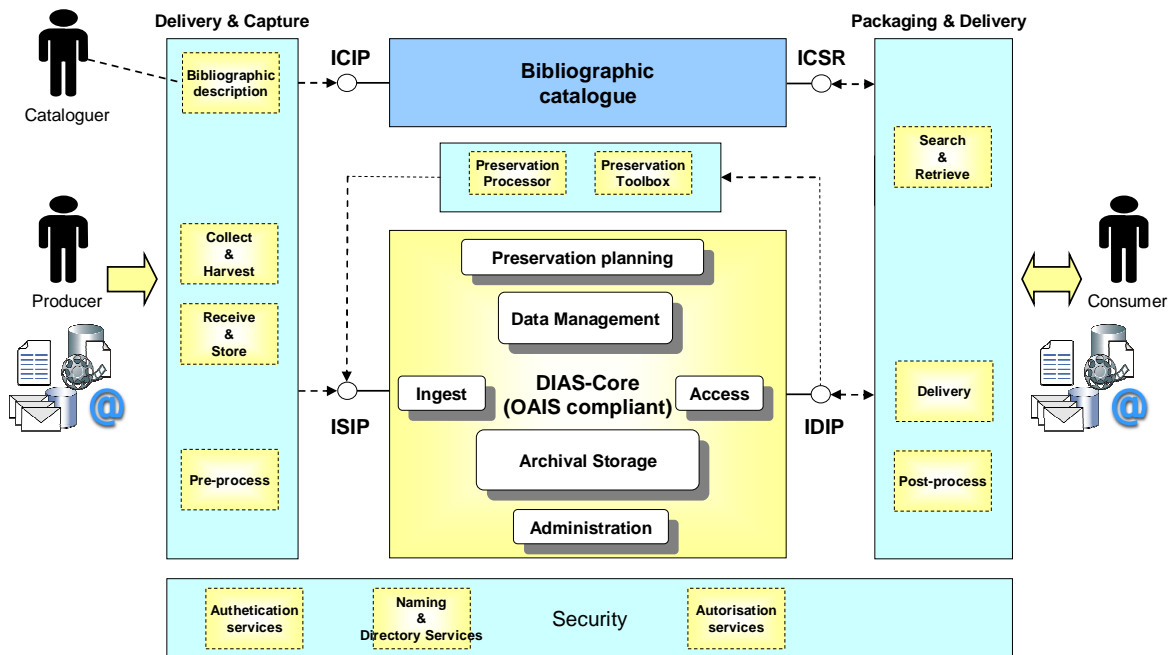
DIAS-Core provides the following functions and features in the order from business to technology:

- OAIS Compliant open archiving solution
- URN (Uniform Resources Name: RFC2141) indexed digital archive
- Multi-organisation support, allowing different organisations to share a common DIAS system
- Support for DIAS-METS v1.0 format based on METS v1.4 and LMER v1.2
- Support for migration of Assets
- What goes in comes out when needed
- Browser-based retrieval and access for supported file types
- Standard and custom reports in a web-based reporting environment
- Supports many storage media and devices through Tivoli Storage Manager
- Prepared and focused for Long Term Preservation (LTP), e.g. Preservation Layer Model (PLM) and Universal Virtual Computer (UVC) concepts
- Physical preservation
- Load balancing on retrievals
- Distributed AIP object storage
- Can integrate with standard security concepts based on LDAP standard
- Backup and restore through proven Tivoli Storage Manager solution
- Open, scalable, flexible solution built on open industry standards like J2EE and XML

Target Customers

Any organisation that has the need to store and keep available digital information over many years such as Deposit Libraries, National Archives, Governmental Institutions, Pharmaceuticals, Banks and Insurance Companies.

Digital Information Archiving System (DIAS)



2

Platform/runs on

- Server hardware and software that can run IBM Content Manager (e.g. IBM/AIX or Sun/Solaris)
- PC/Windows systems for manual building for clients
- Web-based clients for access
- In addition, special PC/Windows systems for the Reference Workstation work environment.

Required Services

IBM Global Services will discuss potential consulting, implementation, customisation and maintenance services with interested customers.

Some Installations, References and Projects

- DIAS v. 1. National Library of the Netherlands <http://www.kb.nl/dnp/e-depot/dm/dias-en.html>
- DIAS v. 2. German National Library. http://kopal.langzeitarchivierung.de/index_software.php.en
- Göttingen State and University Library

DPS (Digital Preservation System)

Vendor/producer

- Ex Libris

Description

The DPS is a preservation solution for digital objects. The system conforms to the OAIS standard recognised by ISO and supports many of the standards in the library environment (METS, PREMIS, MARC, DC, OAI-PMH etc.). The system is designed to support the acquiring, validation, ingest, storage, management, preservation and dissemination of different types of digital objects. The system is designed to support the E-legal deposit requirements; the loading of SIPs can be done directly by the producers or by library staff, and the system supports loading of SIPs individually or in bulk depending on circumstances.

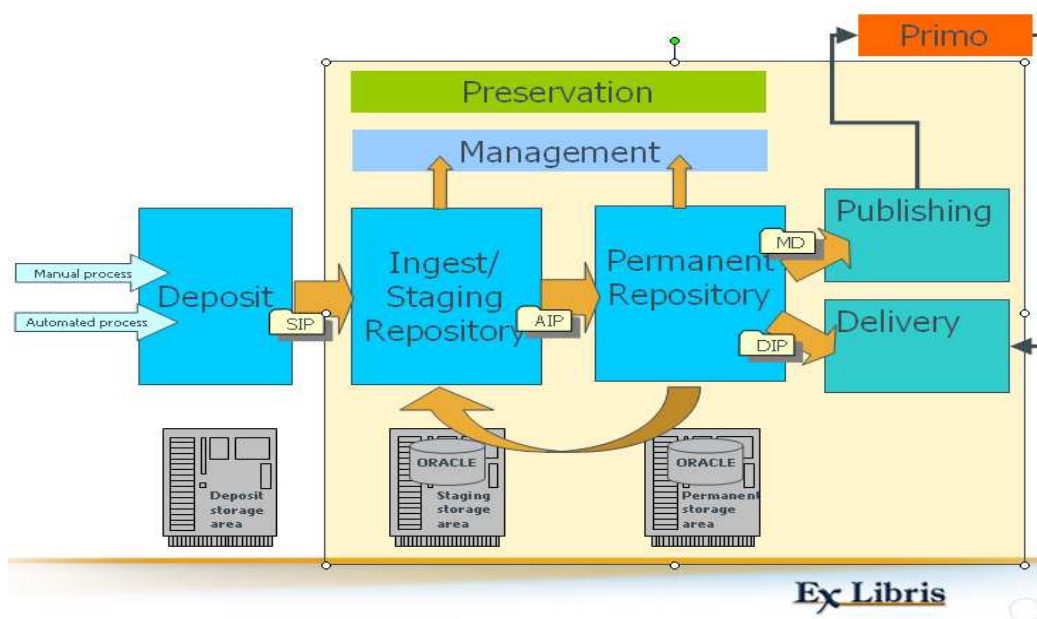


Fig.1: DPS architecture.

Below is a high-level description of the DPS modules (described from the workflow perspective)

Deposit

- A module for either external producers (publishers) or internal producers (library staff) to upload and save SIPs that should be ingested into the preservation repository. The module supports multi workflows, so different workflows can be attached to different producers and they can be automatic, semi-automatic or manual, depending on the pre-agreement and negotiations with the specific producer. At this stage in the process, the SIPs loaded are not ingested into the repository and are stored in an external deposit area until the producer decides to 'Submit' the material.

- The module also supports the management of producers and producer members including contact information, attached workflow, SIP formats and the relation between producer and producer members.

Staging/Ingest Repository

- A working area for materials that have been submitted but have not yet been approved for storage into the permanent repository. This temporary repository includes both automatic and manual processes to ensure the data streams' (i.e. files) integrity (virus, checksum, well formed etc.) as well as the structure of the files in the submitted SIP (relationships between data streams that are part of one intellectual entity). At this point in the process SIPs can still be rejected and changes to structure and format can apply. In the OAIS terminology this is where the SIP becomes an AIP.
- Furthermore, the staging repository is designed to serve as the working environment for the actual execution of preservation migration actions on AIPs that are extracted from the permanent repository using the Preservation module (see below).
- It is important to mention that, while some of the actual process (e.g. virus checks, format migration) should be executed using third-party tools, the staging repository is designed to allow a seamless working environment for the embedding of such tools, i.e. designated UI tools and APIs to support integration of tools.

Permanent Repository

- The Permanent Repository is responsible for the long-term storage of AIPs. The AIPs are comprised of a Data Stream, such as an image or text file stored in the file system, and associated metadata (for example, descriptive, technical, preservation, rights) stored in the Oracle Database. The Repository provides a flexible data model that can support Intellectual Entities, Representation and file stream levels.
- In addition, the Repository provides a set of administrative services such as export, index, access rights to the Repository, and embedded tools to ensure the integrity and quality of the stored AIPs. The DPS routinely supports virus and integrity checks to ensure the objects in the permanent repository have not been altered.
- The interaction between the permanent Repository and the other components of the DPS as well as with those of local or third-party systems is enabled via a standard Web Services layer.

Preservation Module

- A module to support preservation planning and preservation actions. This module will facilitate three main workflow/tasks
 - a) Risk analysis
 - b) Sample data migration
 - c) Preservation action.
- The actual preservation actions will be performed outside the preservation module, either within the staging/ingest repository (see above) or potentially externally (exporting the files out of the system to other systems or third party). In both scenarios the migrated data (new AIPs) will be reingested into the permanent repository via the staging/ingest repository.

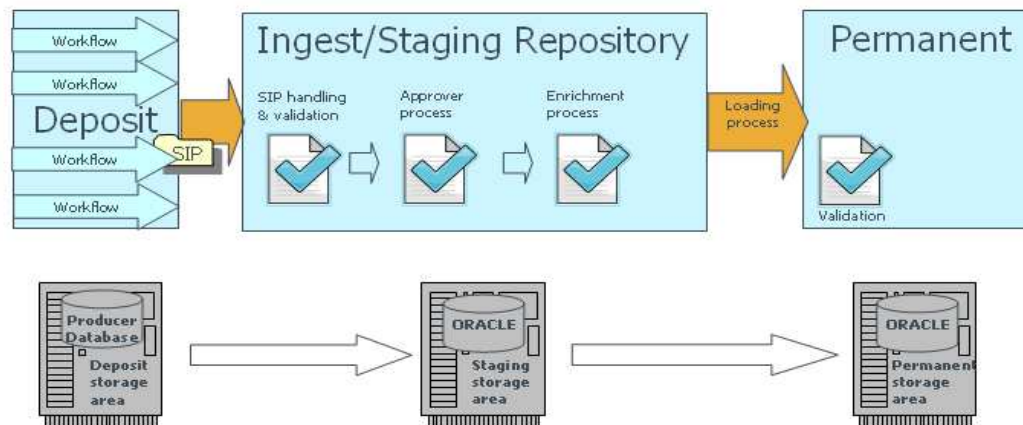



Fig 2: Data workflow.

Management

- The DPS includes a set of tools and interfaces to search and manage the objects both in the permanent repository and in the staging repository. The tools include a web interface to search objects, a web interface to perform maintenance jobs and a powerful client module, the Meditor, which provides the tools needed to edit the metadata of the objects stored in the Repository. Any action on an object will be logged and maintained for searching and managing.

Publishing

- A component to enable externalising information from the DPS permanent repository to external tools like Resource Discovery environments (search engines). The Publishing component supports either creating a replication of a set on the disk or using OAI-PMH to provide the information.

Delivery

- This module is responsible for disseminating the information stored in the repository (building DIPs). The DPS Delivery module is responsible for two main functions:
 - Controlling the access rights – checking the specific delivery request based on user location, attributes and the access rights defined for the specific AIP.

- Delivery roles – assuming the user is allowed to view the AIP, the delivery module decides on the best delivery method including activation of Pre Delivery processors that build the DIP (this can be different from the AIP).
- The Delivery module is designed to allow the embedding of different access rights, and the delivery of business roles as well as viewers.

Scalability

The system provides a scalable solution and is built on a distributed architecture to provide a robust system with no single point of failure. Any hardware part of the system can be duplicated on one or more machines.

5.1.2 Open Source based

- CDS Invenio
- DSpace
- EPrints
- Fedora
- Greenstone
- LOCKSS

CDS Invenio¹

URL: <http://cdsware.cern.ch/invenio/index.html>

Vendor/Producer

- CERN, the European Organisation for Nuclear Research

Current version

- v0.92.1, released 2007-02-20

Description

Developed by CERN, the European Organisation for Nuclear Research, based in Geneva, CDS Invenio (CERN Document Server Software) is designed to run an electronic preprint server, online library catalogue, or a document system on the Web. CDS Invenio (formerly CDSware), the integrated digital library system, is a suite of applications that provides the framework and tools for building and managing an autonomous digital library server. Its flexibility and performance make it a comprehensive solution for the management of document repositories of moderate to large size.²

¹ Formerly CDSware. As of 1 July 2006, CDSware's new name is CDS Invenio.

² <http://cdsware.cern.ch/invenio/index.html>

At CERN, CDS Invenio manages over 500 collections of data, consisting of over 800,000 bibliographic records (including 360,000 full-text documents) covering preprints, articles, books, journals, photographs and more. Besides CERN, CDS Invenio is currently installed and in use by over a dozen scientific institutions worldwide. CDS Invenio is suitable for those institutions looking for a robust system and willing to pay for the support done by developers.

Availability

- Free, Open Source software distributed under the GNU General Public Licence
- Download location: <http://cdsware.cern.ch/invenio/download.html>

Platform/runs on

- Unix-like operating system. The main development and production platform for CDS Invenio at CERN is Debian GNU/Linux, but we actively develop also on FreeBSD and Mac OS X. CDSware runs on an Apache/Python web application server.

Programming language

- Almost entirely written in the Python programming language, with some *ad hoc* modules and functionalities developed in PHP or Common Lisp.³

Database

- MySQL

Interoperability

- OAI compliant
- OAI-PMH 2.0
- MARC 21 metadata standard, its XML derivative MARCXML to store and process bibliographic metadata
- Not Z39.50 protocol compliant

Support

- Free email support at cds.support@cern.ch or through mailing list: project-cdsware-users@cern.ch
- Paid technical support is also available.

Example site

- CERN document server: <http://cdsweb.cern.ch/>

More features

- Full-text search

³ PEPE, A. *et al.*, CERN Document Server Software: the integrated digital library, 2005. 6 pages. Presented at the ELPUB 2005 conference, Heverlee (Belgium), 8-10 June 2005. Available at: <http://cdsware.cern.ch/invenio/doc/elpub2005.pdf>

- Extensibility: API available
- Powerful search engine with Google-like syntax
- User personalisation, including document baskets and email notification alerts
- The software complements other librarians' tools such as Aleph 500, with which CDS Invenio can synchronise⁴
- Navigable collection tree
- Powerful search engine
 1. Specially designed indexes to provide Google-like search speeds for repositories of up to 1,500,000 records
 2. Customisable simple and advanced search interfaces
 3. Combined metadata, full-text and citation search in one go
 4. Results clustering by collection
- Flexible metadata
 1. Standard metadata format (MARC)
 2. Handling articles, books, theses, photos, videos, museum objects and more
 3. Customisable display and linking rules
- User personalisation
- Multiple output formats
 1. HTML
 2. XML
 3. MARC
 4. OAI

Some Installations, References and Projects

For more, see <http://cdsware.cern.ch/invenio/demo.html>

- [CERN Document Server](#) - CERN, Geneva, Switzerland.
- [MeIND](#) - HBZ NRW, Cologne, Germany - Metadata on Internet Documents (MeIND) is an OAI service provider carrying all kinds of subjects from different data providers in Germany (e.g. Die Deutsche Bibliothek, OPUS, DuetT).
- [EPFL Infoscience](#) - EPFL, Lausanne, Switzerland - The scientific information portal of the École Polytechnique Fédérale de Lausanne (EPFL).
- [RERO DOC](#) - RERO, Martigny, Switzerland - Digital library of RERO, the Library Network of Western Switzerland.
- [CAB UNIME](#) - University of Messina, Italy - Search portal of the Centro di Ateneo per le Biblioteche (CAB) of the University of Messina. Containing more than 200,000 records.
- [PADIS](#) - Università La Sapienza, Rome, Italy - Pubblicazioni Aperte Digitali Interateneo Sapienza (PADIS) is an open archive repository containing PhD theses of Università La Sapienza.
- [Aristotle University of Thessaloniki](#) - Aristotle University, Thessaloniki, Greece
- [FYNU UCL Document Server](#) - Université catholique de Louvain, Belgium
- [HBZ Digitalisierte Drucke Portal](#) - HBZ NRW, Cologne, Germany
- [SYSDOC](#) - Systems Competence Area, Research Center COM, Technical University of Denmark

⁴ CAFFARO, Jerome, *Improving the formatting tool of CDS Invenio*. Master's Thesis, 2006. 88 pages. Available at: http://hci.epfl.ch/website/publications/2006/Caffaro_master_thesis.pdf

- [Dipòsit Digital de Documents \(DDD\)](#) - *Universitat Autònoma de Barcelona, Spain*
- [RomDoc](#) - *UPB-CTTPI, Bucharest, Romania.*
- [EELA Document Server](#) - *CIEMAT, Spain*
- [Documents Consorzio Cometa](#) - *INFN, Italy*
- [Documents TriGrid VL](#) - *INFN, Italy*
- [Healthgrids Knowledge Base](#) - Within the framework of the SHARE project funded by the European Commission, a Knowledge Base on healthgrids has been developed and made available to the worldwide community.

DSpace

URL: <http://www.dspace.org/>

Vendor/producer

- Massachusetts Institute of Technology and Hewlett- Packard Company

Current version

- 7 December 2006: DSpace 1.4.1 released

Description

The DSpace digital repository system was designed to capture, store, index, preserve and provide access to institutional digital research materials. It can accept all forms of digital materials, ranging from text, images and datasets to websites, multimedia, video and audio files. DSpace can be used in a variety of ways, including as an institutional repository, e-learning objects or e-theses repository, an electronic records management system, a digital asset management system, and a digital preservation system.

Originally developed by MIT (Massachusetts Institute of Technology) and Hewlett-Packard, further development is ongoing by the DSpace registered community of users (also known as the DSpace Federation). As the requirements of communities might vary, DSpace allows the workflow and other policy-related aspects of the system to be customised to serve the content, authorisation and intellectual property issues of each. Supporting this type of distributed content administration, coupled with integrated tools to support digital preservation planning, makes DSpace well suited to the realities of managing a repository in a large institutional setting in terms of its feature set. DSpace is also focused on the problem of long-term preservation of deposited research material.

It is suitable for large and complex organisations that anticipate material submissions from many different departments (so-called communities) since DSpace's architecture mimics the structure of the organisation that uses DSpace. This supports the implementation of workflows that can be customised for specific departments or other institutional entities.

Availability

- DSpace uses the BSD license. It has a more open community in developing the software compared with the other repository systems evaluated.
- Download location <http://sourceforge.net/projects/dspace/>

Platform/runs on⁵

- Unix-like OS
- Java 1.4 or higher
- Apache Ant 1.5 or higher
- Relational Database Management System (RDBMS) / PostgreSQL 7.3 or higher / Oracle 9i or higher
- Jakarta Tomcat 4.x or higher, or something equivalent
- Lucene (indexing)
- Runs on Windows XP as well

Programming language

- Java

Database

PostgreSQL/Oracle

Interoperability

Supports:

- OAI-PMH 2.0
- METS
- Dublin Core
- REST and SOAP web services
- SRU/SRW
- LDAP authentication
- OpenURL
- Creative Commons
- Z39.50 protocol compliant – no

Security

- DSpace supports SSL
- Configurable infrastructure for authentication in DSpace that currently supports web UI or LDAP authentication
- Supports different groups and roles
- A web UI also allows you to edit the permission and policies.

Support

- DSpace documentation is reasonable
- There is a fair amount of activity and community for DSpace
 - A public bug tracker and patches are available in the Sourceforge project space
 - There is a free mailing list for DSpace: http://sourceforge.net/mail/?group_id=19984
 - Wiki pages for DSpace are available at this address: <http://wiki.dspace.org/>

⁵ More details can be seen here: <http://DSpace.org/technology/system-docs/install.html>

- DSpace also has an IRC channel at freenode.net #DSpace
- DSpace User Conferences are held every year

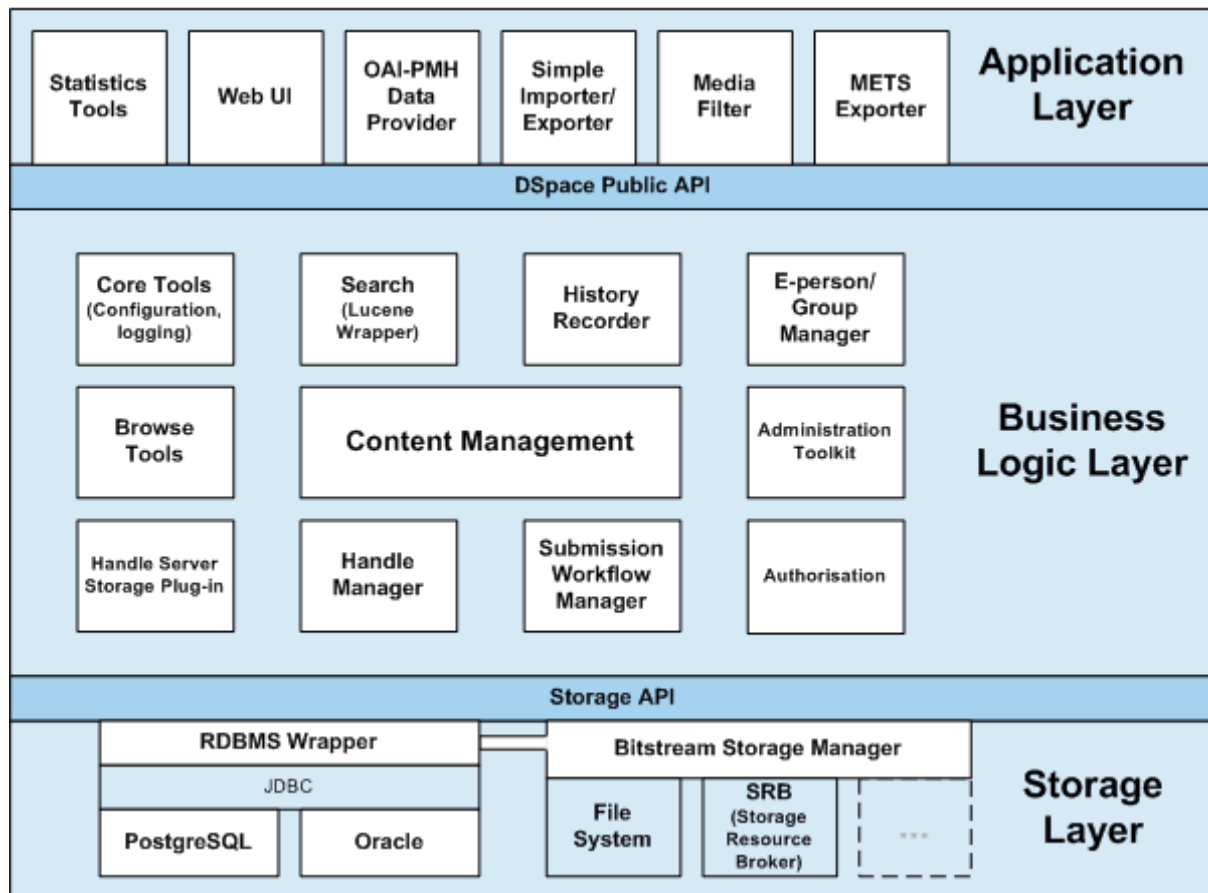
More features

- DSpace has a user interface to add new metadata and namespaces. The database layout that stores the metadata supports the addition of new metadata. DSpace supports the addition of different workflows to a collection that holds different digital objects
- METS is used to maintain links between item components
- Uses the CNRI (Corporation for National Research Initiatives) Handle System to provide unique and persistent identifiers for every item stored
- Supports localising the UI; it even has downloadable language packs. Storage and metadata can be done in unicode
- DSpace also takes a first step towards archiving websites. It is capable of storing self-contained, non-dynamic HTML documents.

Preservation

One of the earliest repository systems to tackle the issue of preservation, DSpace captures details of the specific file formats users submit and maintains a bitstream format for each bitstream in the system. System administrators can maintain a registry of known bitstream formats and the preservation service level available for each format type; however, if the format of the bitstream is unknown, the system will not be able to reliably support preservation and future access or re-use of the file contents. Most implementations maintain lists of 'supported' and 'unsupported' file formats.⁶

⁶ PENNOCK, Maureen, *Dspace digital repository software*, Bath: DCC, 2006. Available at: <http://www.dcc.ac.uk/resource/technology-watch/dspace/>



DSpace System Architecture overview⁷

Some Installations, References and Projects

For more, see: <http://wiki.dspace.org/index.php/DspaceProjects>

- [DSpace/ARK Integration Project](#) - adding support for alternative identifier schemes in DSpace. Project partners include NYU, UCLA, UCSD and the California Digital Library (CDL).
- [DspaceSrbIntegration](#) - San Diego Supercomputer Center's Storage
- <https://digitallibrary.sissa.it> - SISSA Digital Library
- The Australian National University (ANU), <http://www.apsr.edu.au/> - Australian Partnership for Sustainable repositories (APS) project, <http://dspace.anu.edu.au> institutional repository at the ANU based on DSpace - DS
- DSpace @ Cambridge - [DigitalPreservationToolsAndStrategies](#)
- [The Tapir](#) - Edinburgh University Library's E-Theses tools.
- oMEO Green Project at the University of Kansas, populating the <http://kuscholarworks.ku.edu> - ScholarWorks repository.
- CAS single sign-on solution for DSpace by Technical University of Denmark, <http://www.cvt.dk/>

⁷ TANSLEY, Robert *et al.*, *DSpace version 1.4.1 beta 1 documentation*. [s.l.] : [s.n], 2006. 157 pages. Available at: http://wiki.dspace.org/static_files/b/be/DSpaceStandard141beta1.pdf

- 'Manakin' XML UI project @ <http://www.tamu.edu/> - Texas A&M University <http://www.cs.tamu.edu/> Computer Science (with others)
- [Digital Object Catalog](#) - A digital asset management system for digital surrogates of Washington Research Library Consortium (WRLC) special collections material, using DSpace as the core repository - contact gourley@wrlc.org
- <http://www.inf.aber.ac.uk/bridge/> - JISC-funded 'Repository Bridge' project
- [PLEDGE: PoLicy Enforcement in Data Grid Environments](#) - Developing Scalable Data Management Infrastructure in a Data Grid-Enabled Digital Library System
- <http://bib3.ulb.ac.be/DI/DISpace/> - DISpace is an internal project of the Université Libre de Bruxelles (Belgium) aiming at setting up an institutional repository which will contain the complete academic bibliography of the University, and which will function as the official deposit of all scientific output of the Institution
- [Symlink DSpace](#) is the project of the Institute of Computer Science at Masaryk University in Brno, Czech Republic. The main purpose of this project is to modify the original DSpace to work with large files (e.g. video files) in a more efficient way
- DRUM - Digital depository at the University of Maryland - <https://drum.umd.edu> - Faculty may now authenticate against the campus LDAP directory, and are granted submission privileges based on their campus profile
- Columbia University
- Cornell University
- Massachusetts Institute of Technology
- Ohio State University
- University of Cambridge
- University of Rochester
- University of Toronto
- University of Washington

EPrints

URL: <http://www.eprints.org/software/>

Vendor/producer

- Developed by the School of Electronics and Computer Science of the University of Southampton/JISC

Current version

- 24 January released v. 3.0

Description

The EPrints software has probably the largest – and most broadly distributed – installed base of any of the repository software systems. Developed at the University of Southampton, the first version of the system was publicly released in late 2000. It was designed as repository software for e-prints, electronic versions of research articles, in either pre-print or post-print versions (or both). The project was originally sponsored by CogPrints, but is now supported by JISC, as part of the Open Citation Project, and by NSF. The system can be readily modified to meet local requirements. EPrints is a good candidate for many institutions as it is

the least complex of the three systems, and hence has the lowest skill level barrier of the three to implement and maintain. EPrints has the widest installed base, a significant factor in that it goes a long way to ensuring its longevity as a fully supported system. The Code base for Eprints is uniform and well documented, making it easier to work on for low-level customisation.

EPrints is already established as the easiest and fastest way to set up repositories of open access research literature, scientific data, theses, reports and multimedia.

Availability

- EPrints uses GNU Public License (GPL) and University of Southampton holds the copyright. This means that it is Open Source but some code contributions may not be accepted.

Platform/runs on

- Unix - developed on Redhat Linux (both Fedora Core and Enterprise), but it is used on any number of Linux distributions, and other UNIX-like systems including OSX
- Apache with mod_perl
- PHP
- MySQL
- Various perl modules

Programming language

- Perl

Support

- Eprint's documentation and code consistency are very good. As EPrints has been developed primarily by a single author, the code has a consistent structure and standard
- EPrints documentation is of good quality
- Relative to the other communities, its size is small, but it is currently active
- A Wiki page is available to the public at this address: http://wiki.eprints.org/w/Main_Page
- A paid community member will have more channels for collaboration and such a support service is provided by the University of Southampton; more details can be found here: <http://www.EPrints.org/services/>
- Wiki.eprints.org
- Technical mailing list
- EPrints Service – who will do everything for you, at a price

Interoperability

- Supports: OAI-PMH 2.0
- Dublin Core 'out of the box'
- METS export through a modified version of the OAI exporter
- LDAP integration is possible
- Z39.50 protocol compliant – no

Security

- Supports SSL by reconfiguring Apache
- Server security is not as good as desired.

More features

- EPrints supports localisation of the UI through language strings
- Storage of metadata can be done in unicode
- The ability to configure a workflow
- EPrints is the best candidate for a self-configuring solution for institutions wanting to set up and host their own repository
- Optimised for Google Scholar
- Works with bibliography managers
- Works with desktop applications and new Web 2.0 services
- RSS feeds and email alerts keep you up to date
- A new '**autocompletion**' feature to assist in better quality metadata. This is preset for the author, journal and ISSN fields
- An **embargo option** for content that cannot immediately be made publicly available. The software will 'release' it on a specified date
- **Flexible workflows** which can be set according to different conditions, e.g. by content type or user
- **Additional content types** including video and sound

Preservation⁸

The features described here have been jointly developed with the [Preserv project](#), with coding on the METS and Creative Commons (CC) licensing components by Preserv. The features are designed to allow an EPrints repository to support preservation through a specialist service provider. The key actions covered include:

- Recording changes to a repository object by updating its 'preservation metadata' (History Module)
- Enabling the service provider to download all the files and metadata comprising an object (METS and DIDL export plugins)
- Notifying the service provider of any rights it has to copy and act on the content of an object (CC licensing).

Some Installations, References and Projects

For more, see: <http://www.eprints.org/software/archives/>

- [Archive of European Integration](#)
- [Australian National University EPrints2 Archive](#)
- [Bioline International EPrints Repository](#)
- [CogPrints Cognitive Sciences Eprint Archive](#)
- [DLR electronic library](#)
- [E-Lis: Research in Computing and Library and Information Science](#)

⁸ EPrints Wiki: preservation support [online]. Last modified 13:49, 29 January 2007. Available at: http://wiki.eprints.org/w/Preservation_Support

- [E-Prints Universidad Computense Madrid](#)
- [ePrints@OUDIR : Okayama University Digital Information Repository](#)
- [Glasgow ePrints Service](#)
- [Goteborg University: School of Economics and Commercial Law](#)
- [Indian Institute of Science, Bangalore, India](#)
- [Iowa Publications Online](#)
- [Lund University Institutional Archive: LU:research](#)
- [Open Research Online](#)
- [Organic Eprints](#)
- [PASCAL EPrints](#)
- [Policy Documentation Center](#)
- [Queensland University of Technology - ePrints Archive](#)
- [University College London Eprints](#)
- [University of Queensland ePrint Archive](#)
- [University of Southampton: Department of Electronics and Computer Science](#)
- [University of Southampton: e-Prints Soton](#)
- [University of Strathclyde](#)
- [University of Twente Repository](#)
- [Universita di Bologna AMS Acta](#)
- [Birkbeck College](#) (London LEAP Consortium)
- [British Library](#)
- [Imperial College](#)
- [Kings College](#) (London LEAP)
- [London School of Economics and Political Science](#) (London LEAP)
- [Royal Holloway](#) (London LEAP)
- [School of Oriental and African Studies](#) (London LEAP)
- [University of Birmingham](#)
- [University of Durham](#)
- [University of Glasgow](#)
- [University of Leeds](#) (White Rose Partnership)
- [University of Nottingham](#)
- [University of Oxford](#)
- [University of Sheffield](#) (White Rose Partnership)
- [University of York](#) (White Rose Partnership)
- [University College London](#) (London LEAP)

Fedora

URL: <http://www.fedora.info/>

Vendor/producer

- Cornell University Information Science and the University of Virginia Library; supported by generous grants from the Andrew W. Mellon Foundation

Current version

- Fedora 2.2 released on 19 January 2007

Description

Fedora is a general-purpose repository system developed jointly by Cornell University Information Science and the University of Virginia Library. Fedora began in 1997 as a DARPA and NSF funded research project at Cornell University, where the initial reference implementation was developed by Sandra Payette, Carl Lagoze and Naomi Dushay. The Fedora Project is devoted to the goal of providing Open Source repository software and related services to serve as the foundation for many types of information management systems.

The Fedora digital object repository management system is based on the Flexible Extensible Digital Object and Repository Architecture (Fedora). The system is designed to be a foundation upon which full-featured institutional repositories and other interoperable web-based digital libraries can be built. The current version of the software provides a repository that can handle one million objects efficiently (the Fedora community aims to test storage and retrieval of 20 million to 30 million objects).⁹

Fedora demonstrates the best scalability among the present systems, and stores multiple types of digital objects and collections particularly well. The funding for Fedora lasted until September 2007. It is a good choice for institutions that take it seriously.

Availability

- The Fedora software is available under the terms of the Educational Community License 1.0 (ECL).

Platform/runs on

- Unix
- Apache
- Java (requires Sun Java SDK 1.4.2 or above)
- Tested on MacOSX and Windows

Database

- MySQL/Oracle 8i
- Designed to be RDBMS-independent
- McKoi
- PostgreSQL

Programming Language

- Java

Interoperability

Supports:

- SOAP and REST web services

⁹ A Guide to Institutional Repository Software, v3, New York: Open Society Institute, 2004. 28 pages. Available at: http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf

- OAI-PMH 2.0
- METS
- MODS
- Dublin Core
- Bulk import and export scripts
- FOXML and METS formats
- Authentication through LDAP
- Z39.50 protocol compliant – no

Security

- Fedora supports SSL. It requires a data directory that is not accessible from the outside and can be secured fairly well
- Fedora only supports two types of access – Read and Management access; it was designed this way as a web service
- Authentication through LDAP

Technical support

- It has a strong development team and development roadmap
- Documentation and code consistency is very good; the quality of the code is high. Adding a new content type is supported, a new content type being defined by a new XSD document
- Its documentation can be found here: <http://www.fedora.info/documentation/>
- The size and activity level of the community is small compared with other Open Source projects. The developers from University of Virginia and Cornell University primarily undertake all the development
- Wiki – http://www.fedora.info/wiki/index.php/Main_Page
- Mailing list – <http://www.fedora.info/community/mailLists.shtml>
- Bug tracker
- FAQ – <http://www.fedora.info/resources/faq.shtml>

More features

- Repository access and management via web services
- Versioning
- Service-oriented architecture
- XML-based Ingest and Export
- XML-based Digital Object Storage
- Basic OAI Provider Interface
- RDF-based Resource Index with Search
- Security Architecture – XACML-based Policy Enforcement
- Server Command Line Utilities
- Repository Rebuilder Utility
- Backend Security for Remote Service Callbacks
- Tools such as Fez support localisation of UI. Currently Fez supports different languages in terms of different PHP templates, rather than localisation of strings.
- As foundation architecture with powerful API-based interoperability features, Fedora is highly flexible and powerful, and has proven itself with large networked repositories similar to those envisaged with the OARINZ project.

- With no set user interface, Fedora has true separation between the ‘backend’ and ‘front-end’. Fedora cannot offer a full repository service ‘out of the box’
- Unlike the other repositories reviewed, which only support download of digital objects, Fedora supports adding operations to a digital object. An example of a complex operation is the ability to zoom digital images, or get text from a digital document by using OCR software. Fedora addresses this problem as it has the ability to proxy the complex operations to different machines.

Preservation

Fedora repositories incorporate a number of features that facilitate the complex tasks associated with digital preservation. Internally all Fedora digital objects are represented in the file system as files in an open XML format. These XML files include data and metadata for the objects plus relationships to services and other objects. The entire structure of a Fedora repository can be rebuilt from the information in these files. In addition, Fedora repositories are compliant with the Reference Model for an Open Archival Information System (OAIS) due to their ability to ingest and disseminate Submission Information Packages (SIPs) and Dissemination Information Packages (DIPs) in standard container formats such as METS and MPEG-DIDL.¹⁰

Fedora service framework¹¹

As described earlier, a Fedora repository runs as a service within a web server. All the functionality of Fedora is exposed as a set of web service interfaces. While Fedora provides the set of core repository services listed earlier in this document, there are many other services that are beneficial companions to a repository. These include specialised ingest services, workflow services and preservation services. The Fedora Service Framework facilitates the integration of new services with the Fedora repository. It takes a service-oriented architecture approach to adding new functionality around a Fedora repository, allowing new services to be built around the core repository as stand-alone web applications that run independently of the Fedora repository.

The Fedora development team has developed an initial set of services – a directory ingest and OAI-PMH service – and will continue to develop new services in the future, especially services for workflow, preservation and search. New services will be part of the main Fedora distribution and will be kept up to date with new versions of the core Fedora repository distribution. Members of the Fedora community are also developing new services that will be shared through the Fedora website. Figure 1 illustrates the Fedora Repository Service in the context of the Fedora Service Framework with current and projected services and applications.

¹⁰ Fedora Open Source Repository Software [s.l.]: Fedora Development Team, 2005. 10 pages. Available at: <http://www.fedora.info/documents/WhitePaper/FedoraWhitePaper.pdf>

¹¹ Fedora Tutorial #1: Introduction to Fedora [s.l.]: [s.n.], 2005. 15 pages. Available at: <http://www.fedora.info/download/2.2/userdocs/tutorials/tutorial1.pdf>

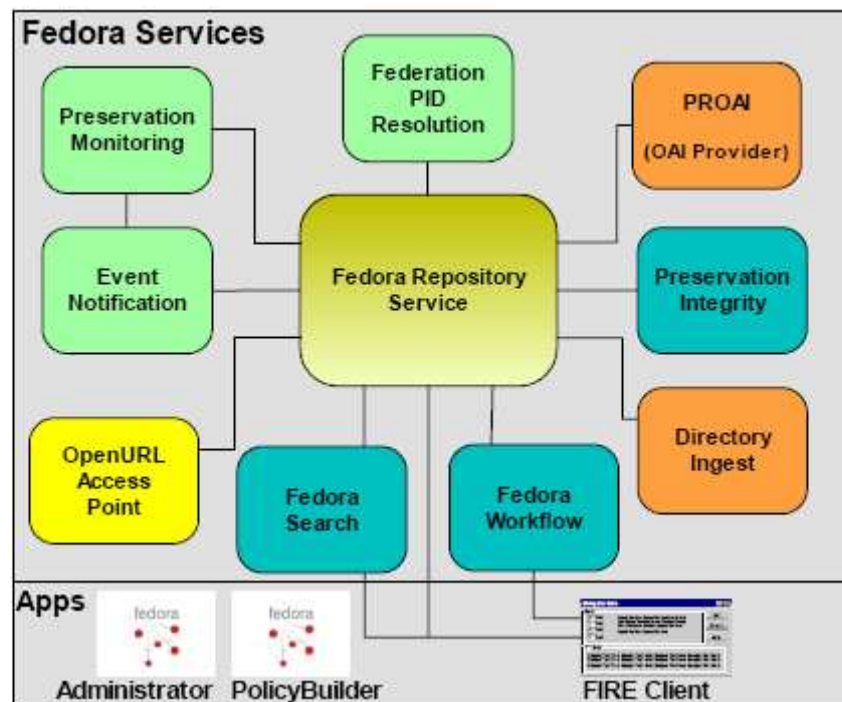


Figure 1: Fedora service framework.

Some Installations, References and Projects

For more, see: <http://www.fedora.info/community/>

- [AWI Homepage](#) - Alfred Wegener Institute for Polar and Marine Research
- [AGU Digital Archive -- Case Study of a Publisher's Use of Fedora](#) - American Geophysical Union
- [ARROW Project](#) - using Fedora-like Institutional repository
- [Digital Case](#) - Case Western Reserve University: Digital Case
- [DiPP Project Homepage](#) - Digital Peer Publishing (DiPP) - An Open Access initiative for eJournals
- [eSciDoc Project Homepage](#) - eSciDoc Project
- Glasgow Caledonian University
- Indiana University: Digital Library Program
- [MAMS Home Page](#) - Meta-Access Management System (MAMS)
- Llyfrgell Genedlaethol Cymru / National Library of Wales
- [Repository Gateway Page](#) - National Library of Estonia
- [NSDL: Creating a Network Overlay Architecture with Fedora](#) - National Science Digital Library (NSDL)
- New York University: The Humanities Computing Group
- [Northwestern University: Academic Technologies](#)
- [Northwestern University Archival Collections](#) - Northwestern University Library
- [Digital Resource Commons Homepage](#) - OhioLINK - Digital Resource Commons
- [Paradigm Project home page](#) - Oxford University Library Services
- [Technical University of Denmark](#)

- [Tufts University: The Digital Collections and Archives Department](#) - using Fedora-like Digital library collection system
- University of Athens, Libraries Computer Center
- [University of Virginia: Digital Library](#) - using Fedora-like Digital library collection system
- [VTLS, Inc.](#) - using Fedora-like Commercial content system
- Yale University: Electronic Records Archive
- The State and University Library, Aarhus, Denmark
- Rutgers University (New Jersey, USA) are using Fedora as the foundation for their digital preservation platform

Greenstone Digital Library Software

URL: <http://www.greenstone.org/cgi-bin/library>

Vendor/producer

- New Zealand Digital Library Project at the University of Waikato

Current version

- February 2007 v3.02 released
- This is a beta release: it contains (virtually) all the features needed for Greenstone2 compatibility, and has been extensively tested. For a production digital library we recommend using Greenstone 2.

Description

Developed by the New Zealand Digital Library Project at the University of Waikato, Greenstone is a suite of software for building and distributing digital library collections. Greenstone was developed and distributed in cooperation with UNESCO and the Human Info NGO.¹² The aim of the Greenstone software is to empower users, particularly in universities, libraries and other public service institutions, to build their own digital libraries.

Greenstone 3 is a complete redesign and reimplementations of the original Greenstone digital library software (Greenstone 2). It incorporates all the features of the existing system, and is backwards compatible: that is, it can build and run existing collections without modification. Written in Java, it is structured as a network of independent modules that communicate using XML; thus it runs in a distributed fashion and can be spread across different servers as necessary. This modular design increases the flexibility and extensibility of Greenstone.¹³

Availability

- Free multilingual, Open Source software

¹² *Greenstone Digital Library Software* [online]. Available at: <http://www.greenstone.org/cgi-bin/library?e=p-en-home-utfZz-8&a=p&p=home>

¹³ DON, Katherine, *Greenstone3 : A modular digital library*, Hamilton, New Zealand: University of Waikato, [s.a]. Available at: <http://www.greenstone.org/docs/greenstone3/manual.pdf>

- Distributed under the GNU General Public Licence

Platform/runs on

- GNU Linux
- Windows
- MacOSX
- C++ (runtime)
- Tested on Solaris and FreeBSD
- It is very easy to install. For the default Windows installation, absolutely no configuration is necessary, and end-users routinely install Greenstone on their personal laptops or workstations. Institutional users run it on their main web server, where it interoperates with standard web server software (e.g. Apache).

Programming Language

- PERL (building)
- GDBM, MG (indexing)

Interoperability

- Dublin Core (qualified and unqualified)
- RFC 1807
- NZGLS (New Zealand Government Locator Service)
- AGLS (Australian Government Locator Service)
- Support for Z39.50, both as a client and a server; support is not enabled by default, and recompilation is needed to enable it
- New metadata sets can be defined using Greenstone's Metadata Set Editor. 'Plug-ins' are used to ingest externally prepared metadata in different forms, and plug-ins exist for XML, MARC, CDS/ISIS, ProCite, BibTex, Refer, OAI, DSpace, METS
- Incorporates a server that can serve any collection over the Open Archives Protocol for Metadata Harvesting (OAI-PMH 2.0)
- Can harvest documents over OAI-PMH and include them in a collection.
- Any collection can be exported to METS (in the Greenstone METS Profile), can ingest documents in METS form. Greenstone uses METS in a very specific way – as an alternative archive format to Greenstone Archive format¹⁴
- Any collection can be exported to DSpace ready for DSpace's batch import program, and any DSpace collection can be imported into Greenstone.

Support

- Online support:
<http://www.greenstone.org/cgi-bin/library?e=p-en-home-utfZz-8&a=p&p=support>
- Technical email list: <https://list.scms.waikato.ac.nz/mailman/listinfo/greenstone-devel>
- User discussion list: <https://list.scms.waikato.ac.nz/mailman/listinfo/greenstone-users>
- Commercial support is available for a fee.

¹⁴ METS in Greenstone: GreenstoneWiki [online]. Last modified 22:27, 19 April 2006. Available at: http://greenstone.sourceforge.net/wiki/index.php/METS_in_Greenstone

- Training – UNESCO sponsored training courses, Digital Library conferences, tutorials (JC DL, ECDL, ICADL, ICDL, ALA Annual Conference)

More features

- Multilingual: The complete Greenstone interface, as well as all documentation, is available in **English, French, Spanish, Russian** and **Kazakh**. Over 25 additional language interfaces are available
- Includes a pre-built demonstration collection
- Offers an 'Export to CDROM' feature

Some Installations, References and Projects

For more, see:

<http://www.greenstone.org/cgi-bin/library?e=p-en-homepref-utfZz-8&a=p&p=examples>

- [Allen Park Veterans Administration Hospital Archives](#)
- [The Arafura Digital Archive](#)
- [Auburn University Libraries Digital Library](#)
- [Books from the Past / Llyfrau o'r Gorffennol](#)
- [Chopin Early Editions](#)
- [The Council of Independent Colleges Historic Campus Architecture Project](#)
- [The Cushing/Whitney Medical Digital Library](#)
- [Detroit Public Library: E. Azalia Hackley Collection](#)
- [iArchives](#)
- [Illinois Wesleyan University Argus Digital Collection](#)
- [Indian Institute of Management, Kozhikode](#)
- [Indian Institute of Science Publications Database](#)
- [Kazakhstan Human Rights Commission](#)
- [Kyrgyz Republic National Library](#)
- [Marshall Foundation Digital Library](#)
- [MOST Digital Library \(UNESCO\)](#)
- [Natural Sciences Digital Library, Vietnam National University, Ho Chi Minh City \(in Vietnamese\)](#)
- [NCSI Demonstration Collections](#)
- [New York Botanical Garden](#)
- [New Zealand Digital Library Project](#)
- [Peking University Digital Library](#)
- [Project Gutenberg](#)
- [Russian Greenstone Library](#)
- [State Library of Tasmania Sheet Music Collection](#)
- [Sudan Open Archive](#)
- [The United Nations Digital Library - Islamabad](#)
- [University of Applied Sciences, Stuttgart](#)
- [Washington Research Library Consortium Special Collections](#)

LOCKSS (Lots of Copies Keep Stuff Safe)

URL: <http://www.lockss.org/lockss/Home>

Vendor/producer

- The Stanford University LOCKSS Program team – LOCKSS Alliance (Mellon Foundation grant)

Current version

- LOCKSS platform CD 243

Description

LOCKSS (for 'Lots of Copies Keep Stuff Safe') is Open Source software that provides librarians with an easy and inexpensive way to collect, store, preserve and provide access to their own local copy of authorised content they purchase. Running on standard desktop hardware and requiring almost no technical administration, LOCKSS converts a personal computer into a digital preservation appliance, creating low-cost, persistent, accessible copies of e-journal content as it is published. Since pages in these appliances are never flushed, the local community's access to that content is safeguarded. Accuracy and completeness of LOCKSS appliances is assured through a robust and secure, peer-to-peer polling and reputation system.¹⁵

The LOCKSS technology has been undergoing increasingly severe testing since 1999. The alpha test ran through 2000, and an early beta version was successfully deployed to 50 libraries worldwide from 2000 to 2002. It ran at these sites with little operator intervention for nearly a year. From 2002 to mid 2004, the Stanford University LOCKSS Program team, with library staff from Emory University, Indiana University and the New York Public Library, addressed myriad questions surrounding collection development, collection management and collection access. The system was released into production in April 2004.

A library uses the LOCKSS software to turn a low-cost PC into a digital preservation appliance (a LOCKSS Box) that performs four functions:

- It collects newly published content from the target e-journals using a web crawler similar to those used by search engines.
- It continually compares the content it has collected with the same content collected by other LOCKSS Boxes, and repairs any differences.
- It acts as a web proxy or cache, providing browsers in the library's community with access to the publisher's content or the preserved content as appropriate.

¹⁵ http://www.lockss.org/lockss/About_LOCKSS

- It provides a web-based administrative interface that allows the library staff to target new journals for preservation, monitor the state of the journals being preserved, and control access to the preserved journals.

Availability

- LOCKSS is Open Source software
- http://sourceforge.net/project/downloading.php?group_id=47774&filename=lockssCD243.iso

Platform/runs on

- In order to improve security, the LOCKSS software runs from a CD rather than from the hard disk; it doesn't need any OS, just a PC.¹⁶
- The LOCKSS software distribution includes everything you need, including a specially configured OpenBSD operating system and the LOCKSS daemon. LOCKSS is not usually run alongside other applications in an operating system used for other purposes. If LOCKSS is installed on a machine that already has an operating system installed, that operating system and all existing data will be erased.
- The LOCKSS team provides the daemon packaged up on a CD image with a specially configured OpenBSD operating system. This CD boots and runs the daemon on a generic PC. It does not 'install' the daemon on the disk in the conventional way – the system in effect runs from the CD. Rebooting the system returns the system to a known state, because all software is reinitialised from read-only media.
- The LOCKSS team have taken great care to select the most secure available OS, to configure the system on the CD to minimise the risk of security breaches, and to provide mechanisms for updating the system to respond to any vulnerabilities that are detected. With our limited manpower, we are at present only able to support sites that run the daemon in this CD image environment.

Programming language

- Java

Interoperability

- OAIS compliant

Support

- lockss-support@lockss.org

How it works¹⁷

1. Collecting

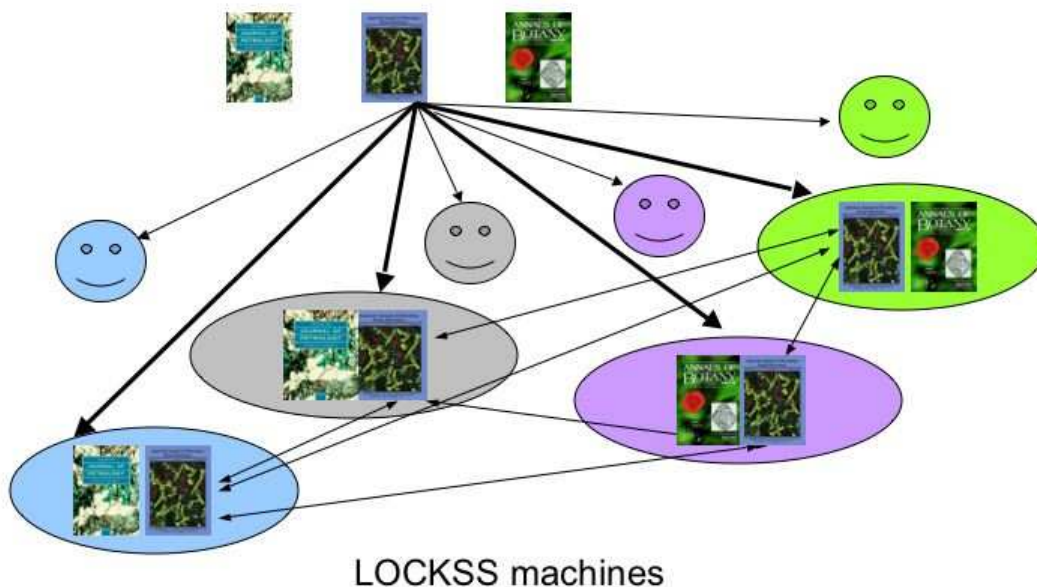
Before LOCKSS Boxes can preserve a journal, two things have to happen:

¹⁶ http://www.lockss.org/lockss/Platform_FAQ

¹⁷ http://www.lockss.org/lockss/How_It_Works

- The publisher has to give permission for the LOCKSS system to collect and preserve the journal. They do this by adding a page to the journal's website containing a permission statement, and links to the issues of the journal as they are published.
- The LOCKSS Box has to know where to find this page, how far to follow the chains of web links so that it doesn't crawl off the edge of the journal and try to collect the whole Web, some bibliographic information, and so on. In order to add new publishing platforms, the LOCKSS system provides a fill-in-the-blanks tool that a librarian or administrator can use to collect this information, and test that it is correct. The information is then saved in a file (a LOCKSS plugin) and added to the publisher's website or to some other plug-in repository, so that it is available to all LOCKSS systems.

Collection via Web Crawler

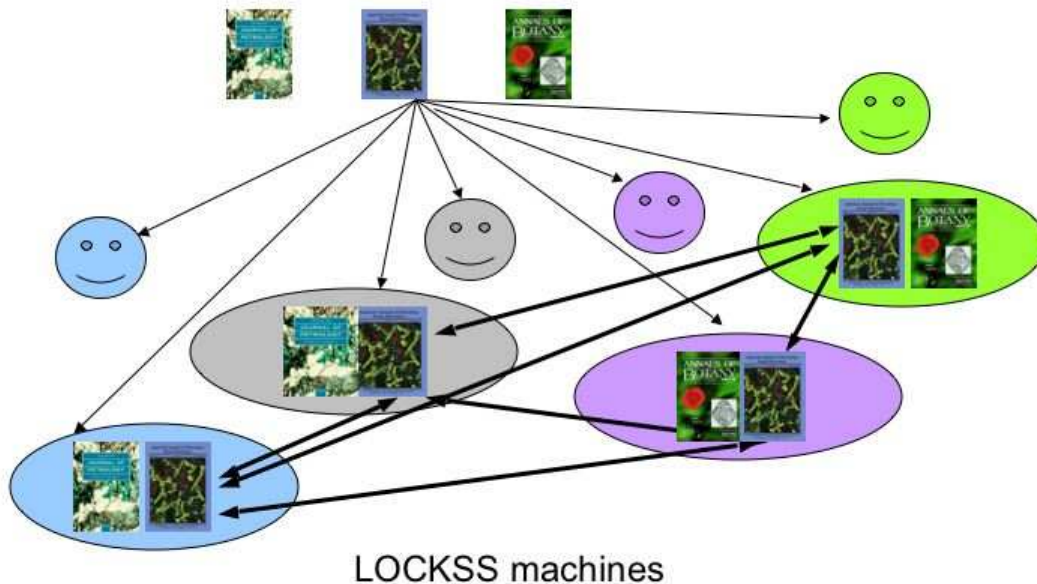


2. Preserving

The LOCKSS Boxes at libraries around the world use the Internet to audit, continually but very slowly, the content they are preserving. At intervals LOCKSS Boxes take part in polls, voting on the digest of some part of the content they have in common. If the content in one LOCKSS Box is damaged or incomplete that LOCKSS Box will lose the poll, and it can repair the content from other LOCKSS Boxes. This cooperation between the LOCKSS Boxes avoids the need to back them up individually. It also provides unambiguous reassurance that the system is performing its function and that the correct content will be available to readers when they try to access it. The more organisations that preserve given content, the stronger the guarantee they each get of continued access.

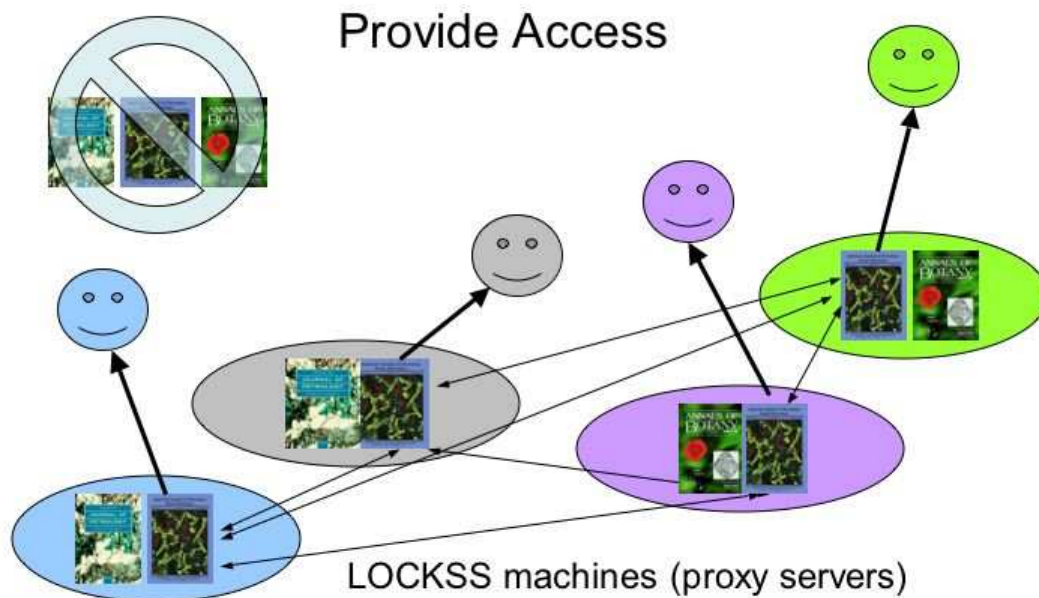
The LOCKSS system moves content forward in time through a process called format migration.

Preserve and audit content integrity



3. Providing Access

LOCKSS Boxes provide transparent access to the content they preserve. Institutions often run web proxies to allow off-campus users to access their journal subscriptions, and web caches to reduce the bandwidth cost of providing Web access to their community. Their LOCKSS Box integrates with these systems, intercepting requests from the community's browsers to the journals being preserved. When a request for a page from a preserved journal arrives, it is first forwarded to the publisher. If the publisher returns content, that is what the browser gets. Otherwise the browser gets the preserved copy.



4. Administering

Library staff administer their LOCKSS Box via a web user interface. The interface targets new content preservation, monitors the preservation of existing content, controls access to the appliance, and fulfils a wide variety of other functions.

Some Installations, References and Projects

For more, see: <http://www.lockss.org/lockss/Libraries>

- National Library of South Africa, Cape Town branch
- National Library of South Africa, Pretoria branch
- Chinese Academy of Sciences Library
- National University of Singapore
- Chinese University of Hong Kong
- National University of Singapore
- The British Library
- Cambridge University
- De Montfort University
- Imperial College
- John Rylands University Library
- Kings' College London
- London School of Economics and Political Science
- Loughborough University
- Middlesex University

- Oxford University
- UCL Library Services
- Humboldt-Universität zu Berlin
- University of Goettingen
- University of Amsterdam
- National Library of Portugal
- University of Edinburgh
- University of Glasgow
- Lund University
- University of Montreal
- Cornell University
- Colorado State University
- Harvard University
- Library of Congress
- Los Alamos National Laboratory
- OCLC Online Computer Library Center
- Stanford University

6. PART IV: CONCLUSIONS AND RECOMMENDATIONS

The SWOT Analysis aggregates information derived from the Market and technology trends analysis and experience of the project participants in the area of digital preservation.

STRENGTHS	OPPORTUNITIES
<ul style="list-style-type: none"> • Digital preservation becomes one of the main strategic priorities for both memory and research institutions – they are increasingly aware of the fact that digital resources and memory are fragile and that they are at risk. • Emerging platform for proactive cooperation under the umbrella of international projects and activities. • Emerging standards for digital preservation. • Emerging tools for certification of trusted digital repositories. • Good choice of Open Source software and established platform for cooperation for research and other institutions having enough research capacities to handle and develop Open Source solutions. 	<ul style="list-style-type: none"> • To raise awareness of curation and preservation of digital resources at governmental level. • To increase proactive cooperation under the umbrella of international projects and activities. • To stimulate the incorporation of preservation functionality in commercial software. • To promote an international approach to repository audit and certification.
WEAKNESSES	THREATS
<ul style="list-style-type: none"> • Digital preservation is not (with some exceptions) one of the main strategic priorities for governments and ministries – they are not sufficiently aware of the fact that digital resources and memory are fragile and 	<ul style="list-style-type: none"> • Digital resources not equipped with necessary metadata (not only descriptive, but also administrative, technical etc.) and not stored and made accessible via trusted digital repositories will disappear.

<p>that they are at risk.</p> <ul style="list-style-type: none"> • Missing national strategies and the lack of funding for digitisation of endangered paper documents and digital preservation. • Lack of practical experience with standards for digital preservation and tested guidelines useful in real life. • Lack of practical experience with certification of trusted digital repositories and tested guidelines useful in real life. • Almost no competition on the market and poor choice of commercial software resulting in extremely high prices for institutions having not enough research capacities to handle and develop Open Source solutions. 	<ul style="list-style-type: none"> • Lack of coordination in standards implementation in real life might result in proprietary solutions and interoperability problems.
--	--

There are strong EC strategic lines and recommendations concerning digital preservation. Several European countries have prepared their national strategies at governmental level and support funding and sustainability of institutions responsible for both memory institutions and research institutions responsible for digital preservation. However, there are many European countries where only certain institutions are aware of the importance of digital preservation and EC strategic lines and recommendations for this area. There is a gap between the European and institutional levels where no digital preservation strategy exists at governmental level. This situation needs to be addressed urgently.

The result of non-existent digital preservation strategies coordinating the activities of institutions responsible for digital preservation at national levels (and involved in international cooperation at both theoretical and practical levels) is a lack of recommendations concerning standards, systems, certification etc. This could result in proprietary solutions and interoperability problems in the future.

Project information

Project acronym:	DPE
Project full title:	DigitalPreservationEurope
Proposal/Contract no.:	IST-2006-034762

Project Officer: Manuela Speiser

Address:	INFSO-E3 Information Society and Media Directorate General Content - Learning and Cultural Heritage Postal mail: Bâtiment Jean Monnet (EUFO 1167) Rue Alcide De Gasperi / L-2920 Luxembourg Office address: EUROFORUM Building - EUFO 1167 10, rue Robert Stumper / L-2557 Gasperich / Luxembourg
Phone:	+352 4301 33632
Fax:	+352 4301 33190
Mobile:	
E-mail:	manuela.speiser@ec.europa.eu

Project Co-ordinator: Professor Seamus Ross

Address:	HATII, University of Glasgow
Phone:	+44 141 330 3635
Fax:	+ 44 141 330 2793
Mobile:	+44 797 435 7006
E-mail:	s.ross@hatii.arts.gla.ac.uk