

268 Are you Ready?

# The International Journal of Digital Curation

## Issue 1, Volume 6 | 2011

### Are you Ready? Assessing Whether Organisations are Prepared for Digital Preservation

Pauline Sinclair, James Duckworth, Lewis Jardine, Ann Keen, Robert Sharpe,  
Tessella

Clive Billenness, Adam Farquhar, Jane Humphreys,  
The British Library

#### Abstract

In early 2009 the Planets project undertook a survey of national libraries, archives, and other content-holding organisations in Europe to better understand the organisations' digital preservation activities and needs, and to ensure that Planets' technology and services are designed to meet them. Over 200 responses were received including a cross-section of major libraries and archives especially in Europe. The results provide a snapshot of organisations' readiness to preserve digital collections for the future. The survey revealed a high level of awareness of the challenges of digital preservation within organisations. Findings indicated that approximately half of those organisations surveyed have taken measures to develop digital preservation policies and to budget for it, while a majority have incorporated digital preservation into their organisational planning. Organisations predict that within a decade they will need to store large quantities of data in a wide range of formats from a variety of sources; three quarters of them are looking to invest in a solution within the next two years. However, the findings also point to varying degrees of readiness. Organisations with a digital preservation policy are significantly further advanced in their work to preserve digital collections for the long-term than others.<sup>1</sup>

---

<sup>1</sup> This paper is based on the paper given by the authors at iPRES 2009; received January 2010, published March 2011.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. ISSN: 1746-8256 The IJDC is published by UKOLN at the University of Bath and is a publication of the Digital Curation Centre.





## Introduction

In the last few years, digital preservation has developed from a theoretical discipline to one where real solutions are starting to be developed and implemented. While more research is still needed, practical steps are now possible. But, how ready are libraries, archives, and related organisations to begin taking those practical steps?

In early 2009, the Preservation and Long-term Access through Networked Services (Planets)<sup>2</sup> project (Farquhar & Hockx-Yu, 2007) conducted an on-line survey to assess the state of readiness of archives, libraries, and other organisations interested in digital preservation. The survey aimed to understand the state of digital preservation, and the digital preservation needs, of European organisations that create or hold digital content (Sinclair & Jardine, 2009).

Previous studies have provided snapshots of the state of digital preservation. In 2005, the Digital Preservation Coalition (Waller & Sharpe, 2006) surveyed 104 organisations in the UK. These included memory institutions, government departments, research institutions, and companies in a range of sectors. The results showed that there was considerable confusion about how to address the problem of digital preservation. While 41% of respondents said there was a need to keep digital information alive for 50 years or more, and 52% said they had a high level of commitment to digital preservation, just 18% had a strategy in place, and 20% had funding. Half (55%) were unclear about roles and responsibilities, half (55%) had not yet assessed the volumes of material they needed to preserve, and half (49%) did not know the life spans of digital data. The same proportion (50%) stated that they printed out hard copies of digital information as a means to preserve it. The study revealed the scale of the problem (with a growing volume of digital information of increasing value), and the lack of good solutions in place in organisations. It concluded that despite the high levels of awareness: “the level of implementation of digital preservation solutions is significantly lower than would be expected given the awareness and commitment that were measured.”

In 2006 and 2007, Digital Preservation Europe (DPE<sup>3</sup>; 2007) surveyed 172 national libraries, archives, research institutions, Information and Communication Technology (ICT) and media companies, and other organisations in Europe. The results similarly pointed to high levels of awareness. Seventy-seven percent of respondents considered long-term preservation to be a key strategic priority. However, only one-third (35%) had implemented a trusted repository. The findings demonstrated that where respondents had digital preservation systems in place, these were a mix of open-source, commercial, and software developed in-house. Organisations also considered cooperation across organisations to be important to digital preservation.

The Planets survey aimed to build on the earlier surveys and determine how awareness has grown, how far organisations are along the path to implementing a full digital preservation solution, what unfulfilled requirements organisations have, and what barriers exist that hinder the adoption of solutions.

---

<sup>2</sup> Planets: <http://www.planets-project.eu/>.

<sup>3</sup> Digital Preservation Europe: <http://www.digitalpreservationeurope.eu/>.



## Method

The Planets survey of long-term management of digital information was conducted in February and March 2009, in the form of an on-line questionnaire. The survey was targeted at organisations and individuals with an interest in retaining and accessing digital content in the long term.

Invitations to participate in the survey were sent to around 2,000 individuals, whose role could encompass the long-term maintenance of digital information in libraries, archives, and other organisations across Europe. These individuals were selected from a number of sources. Over half were individuals known to Tessella<sup>4</sup> (the technology and consultancy company who undertook the research on behalf of Planets) as having an interest in digital archiving. The majority of the remainder were individuals who had registered to receive updates of Planets' activities, and the final group was people who were contacted personally by members of the Planets Scientific Board and Executive Steering Committee and invited to participate. Follow-up telephone calls were made to 120 of these individuals to encourage them to take part in the survey. In particular, individuals in the 96 national archives and libraries in Europe, as listed on the website of the United Nations Educational, Scientific and Cultural Organization (UNESCO)<sup>5</sup>, were targeted.

As well, initial announcements about the survey were placed on approximately 30 mailing lists related to digital preservation, and followed up by two reminders during the lifetime of the survey. The lists included international digital preservation mailing lists such as Preserving Access to Digital Information (PADI)<sup>6</sup>, and specialist mailing lists targeting sub-sections of the digital preservation community, such as research institutes, government, and film and sound archives.

In addition, the survey was publicised through intermediary organisations and projects in European Union (EU) member countries. Digital Preservation Europe, the Digital Curation Centre (DCC)<sup>7</sup>, and the Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR)<sup>8</sup>, Sustaining Heritage Access through Multivalent Archiving (SHAMAN)<sup>9</sup>, and PReservation Organizations using Tools in AGent Environments (PROTAGE)<sup>10</sup> digital preservation projects were all asked to cascade notices on Planets' behalf, and the foundation Conference of European National Libraries (CENL)<sup>11</sup>, International Council on Archives (ICA)<sup>12</sup>, and the Association of European Research Libraries (LIBER)<sup>13</sup> were asked to disseminate the message to their members. Finally, a news item about the survey and inviting participation was placed on Planets' website.

<sup>4</sup> Tessella: <http://www.tessella.com/>.

<sup>5</sup> UNESCO: <http://www.unesco.org.uk/>.

Archives list: <http://www.unesco-ci.org/cgi-bin/portals/archives/page.cgi?d=1>.

Libraries list: <http://www.unesco-ci.org/cgi-bin/portals/libraries/page.cgi?d=1>.

<sup>6</sup> PADI: <http://www.nla.gov.au/padi/>.

<sup>7</sup> DCC: <http://www.dcc.ac.uk/>.

<sup>8</sup> CASPAR: <http://www.casparpreserves.eu/index.html>.

<sup>9</sup> SHAMAN: <http://shaman-ip.eu/shaman/>.

<sup>10</sup> PROTAGE: <http://www.protage.eu/>.

<sup>11</sup> CENL: <http://web3.nlib.ee/cenl/index.php>.

<sup>12</sup> ICA: <http://www.ica.org/>.

<sup>13</sup> LIBER: <http://www.libereurope.eu/>.

Respondents were promised confidentiality and anonymity in the introduction to the survey. The survey comprised 29 questions which took up to half an hour to complete, therefore, it is not surprising that not everyone completed all questions.

## Results

Two hundred and six responses were received before the survey closed.

### *Distribution of Responses*

**Countries.** Fifty-six percent (115) of responses were from EU countries, and 11% (23) of responses were from European countries outside the EU. Sixteen percent (33) came from Canada and the USA. Just 3% (6) of responses came from the rest of the world. Fourteen percent (29) did not disclose their country. Ten or more responses were received from: the UK (54), USA (26), Germany (16), Switzerland (15), and Netherlands (10).

**Organisation types.** Forty-one percent (75) of responses represented libraries, and 30% (55) represented archives. Fifteen percent (28) were from government departments and the public sector. Seven percent (12) were from suppliers and vendors, and 4% (8) from commercial organisations. Three percent (6) were from museums. See Figure 1 for the full results and note that the total number of responses received for this question was 183, not 206.

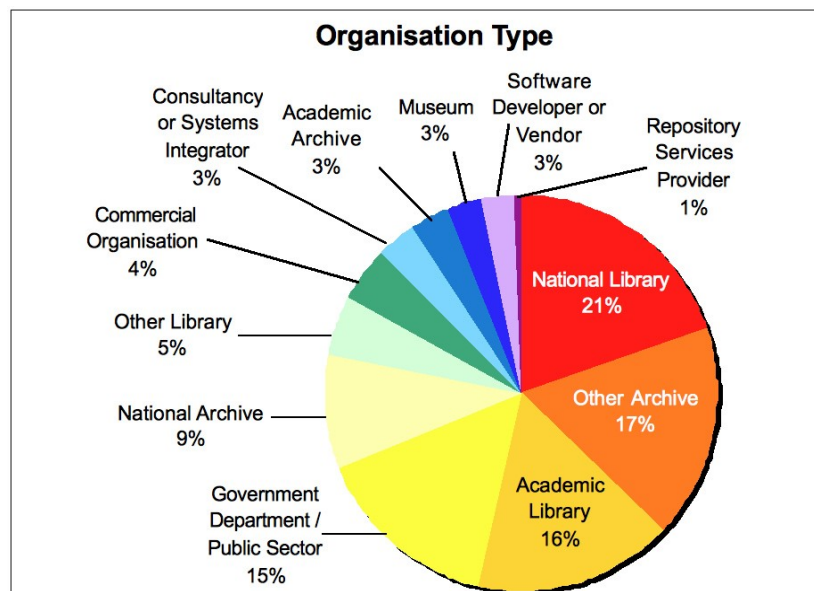


Figure 1. Survey Respondents by Organisation Type (183 Responses)

**Respondents' professions.** Respondents came from a wide range of professional backgrounds. Fifteen percent stated that they specialise in digital preservation. Twenty-two percent work in curation and records management, 16% work in preservation in general, and 16% work in Information Technology (IT). The remainder work in a variety of professions, including management, research, and those that produce digital information.

### Digital Information Requiring Preservation

The survey aimed to establish the volume and types of digital content that organisations need to hold now, the volume and types of digital content that they predict they will need to hold over the next 10 years, and the source systems this content will be derived from.

**Data volumes.** Respondents were presented with categories of data volumes (from < 1 terabyte (TB) to > 1 petabyte (PB)), and asked to indicate the volume of digital content they store now, and the volume they expect to store in 2, 5, and 10 years time (Figure 2, Table 1). Eighty-seven percent of respondents hold < 100 TB of content now (Figure 2). The median volume of content is < 20 TB (Table 1).

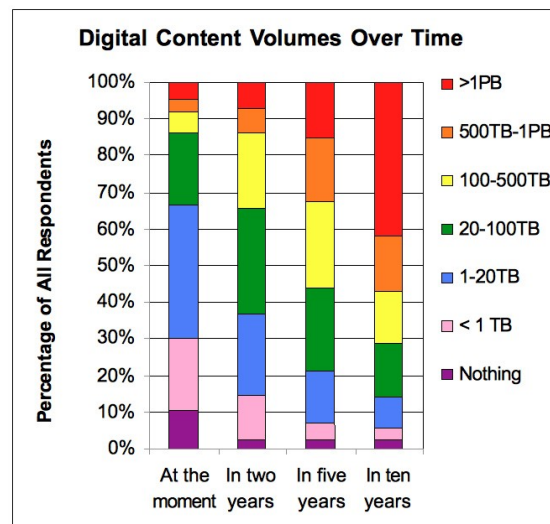


Figure 2. Volumes of Digital Content that Organisations Intend to Store over the Next 10 Years (129 Responses).

	Data Volumes Now	Data Volumes in 2019
Mean <sup>14</sup>	150 TB	1.0 PB
Median	1-20 TB	500 TB – 1 PB
Mode	1-20 TB	> 1PB

Table 1. Average Volumes of Digital Content that Organisations Store Now, and Intend to Store in 10 Years' Time (129 Responses).

In 10 years' time, 70% of respondents expect to hold > 100 TB (Figure 2). The median volume of content held is expected to be > 500 TB (see Table 1). Forty-two percent of respondents' organisations expect to hold > 1 PB of data in 10 years' time. Ten percent store nothing now; this is expected to fall to 2% in two years' time (see Figure 2).

National archives and national libraries hold the largest volumes of data: a mean<sup>14</sup> of 190 TB in 2009 (200 TB and 180 TB respectively). They expect to hold a mean of 1.4 PB of digital information in 2019 (data not shown).

<sup>14</sup> The mean was calculated using the mid-point of each band of data volumes, and a value of 2 PB for the > 1 PB band; it is given to 2 significant figures.

**Types of digital information.** Over 80% of organisations indicated that they currently have a need to preserve documents and images, and this rises to > 95% in 10 years time. Within 10 years, > 70% of organisations expect to need to preserve video, audio, databases, websites, and email. Almost half (49%) of organisations already have a need to preserve databases, and by 2019, 85% expect to need to preserve them (see Table 2).

Digital Information Type	2009	2019
Documents	81%	99%
Images	80%	95%
Databases	49%	85%
Audio	49%	80%
Websites	47%	79%
Video	51%	76%
E-mails	39%	66%
GIS	23%	52%
Scientific Data	22%	51%
Software	21%	49%
eBooks	21%	49%
eJournals	25%	49%
ISO or disc images	20%	40%

Table 2. Types of Digital Information that Organisations Currently, or Expect to in the Future, Preserve (138 Responses). The percentage of responding organisations that indicated a need to preserve each digital information type is shown.

In 2019, the percentage of libraries storing websites (99%), eBooks (81%), and eJournals (81%) is significantly higher (at the 99% confidence level) than the average of all types of organisations (data not shown).

**Source systems.** Organisations receive content from a range of source systems. Those used by more than half of the respondents were: file systems (77%), document scanning programmes (58%), the internet (55%), electronic document management systems (55%), email systems (54%), and media digitisation programmes (54%). The survey showed that niche, or domain-specific, source systems are used by far fewer organisations; computer-assisted design (CAD) is used by 29%, and lab systems by 18% of respondents' organisations.

Libraries concentrate on archiving the internet (77%), and media digitisation programmes (75%), whereas archives have more of a focus on the systems used to manage organisations: email (64%), Electronic Document Management Systems (EDMS) (66%), Electronic Records Management System (ERMS) (52%).

### **Digital Preservation Readiness**

**Awareness of digital preservation.** Ninety-three percent of respondents stated that their organisation is aware of the challenges presented by digital preservation. Twenty-four percent of respondents currently have a solution in place or planned, and over half (52%) of surveyed organisations are actively seeking or working on a digital preservation solution.

**Digital preservation policies.** Nearly half (48%) of the organisations surveyed have a policy for the long-term management of digital information, where long-term is defined as greater than five years (see Figure 3). This varies by organisation; 64% of archives, and 43% of libraries, have a digital preservation policy. However, only one-quarter (27%) of government departments, and the public sector in general, have a digital preservation policy in place. A high proportion of commercial organisations (88%), and suppliers and vendors (60%), have digital preservation policies, although these results should be treated with caution, due to the small size (eight commercial organisations, and ten suppliers and vendors), and potentially unrepresentative nature, of the sample.

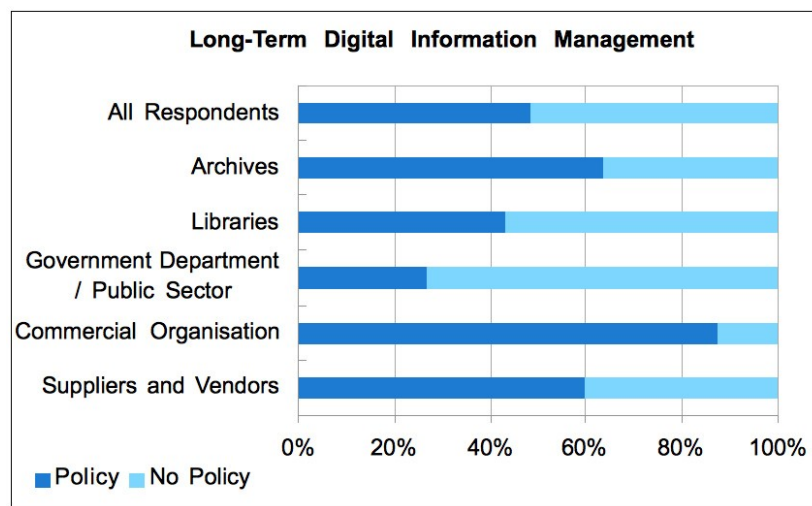


Figure 3. Organisations with Long-Term Digital Information Management Policies (161 Responses).

**Digital preservation's inclusion in organisations' general planning.** When respondents were asked about the inclusion of digital preservation in their organisations' operational, financial, and business continuity planning, it was found that 76% of respondents' organisations include it in their operational planning, 71% in their business continuity planning, and 62% in their financial planning. Ninety-three percent of national archives include digital preservation in their operational and business continuity plans, and 86% in their financial plans. All surveyed national libraries include digital preservation in their operational planning, 82% include it in their business continuity planning, and 91% include it in their financial planning. Seventy-one percent of government departments, and the public sector in general, include digital preservation in their operational planning, 53% in their business continuity planning, and 44% in their financial planning.

**Budgets for digital preservation.** Almost half of respondents (47%) said that their organisation has a budget for digital preservation. Sixty-one percent of archives, 50% of libraries, and 26% of government departments, and the public sector in general, have a budget for digital preservation.

Fewer European organisations (45%) have a budget for digital preservation than North American organisations (59%). North American budgets are fairly evenly split

between revenue and capital: 15% capital only, 22% revenue only, and 22% both capital and revenue. In contrast, European budgets are more likely to be capital budgets: 24% capital only, 3% revenue only, and 17% both capital and revenue.

**Timescales for investment.** The majority (77%) of organisations plan to invest in a solution in the next 2 years. One third (32%) of organisations are currently investing in a digital preservation solution, and two-fifths (45%) are looking to make an investment in the next 6 months to 2 years. One-fifth (23%) do not plan to invest for over 2 years.

### ***Digital Preservation Implementations***

**Implementation phases.** Respondents were asked to describe the stage that their organisation was at in working towards a digital preservation solution. They were allowed to select > 1 option from the 6 options presented to them, resulting in the total percentage exceeding 100%. Eighty-five percent of organisations stated that they are working towards a solution, or have one in place. The remaining 15% of respondents have no plans to deal with the long-term management of digital content. Of those working towards a solution, 27% are assessing their needs using consultancy, and 22% with a prototype; 13% are tendering for a solution; 48% have a long term solution in development, and 7% already have one in place.

Many of the respondents were at more than one stage in working towards a long-term solution. For example, of those who already have a long-term solution in place, 18% are assessing their needs and requirements, and 32% are looking to improve or extend their current solution.

**Solution implementation.** Respondents were asked about how they expect to implement their solution (respondents were allowed to select more than one answer, resulting in the total exceeding 100%), who they expect to implement it, and whether they use, or plan to use, open source or proprietary software. Two-thirds (64%) of organisations are integrating components into a custom solution, with the remainder evenly split between developing a custom solution (33%), and using an off-the-shelf package (32%). Respondents are combining these approaches, with half (50%) of those developing a custom solution also integrating components into that solution, and two-fifths (40%) of those using an off-the-shelf package also integrating components into a custom solution. Approximately one-tenth (11%) of organisations are developing a bespoke or custom solution from scratch, that is, without using existing components or off-the-shelf software packages.

Sixty-nine percent of respondents indicated that they expect to use an in-house team to implement their solution. Forty-six percent said they expect to use a third-party development team, and 21% expect to use a third party system integrator. Forty-five percent of respondents are using more than one type of implementer. Of those using an in-house software team, 34% also expect to use a third-party development team, and 12% a third-party system integrator.

Over half (57%) of respondents currently use a mixture of open source and proprietary software, with the rest of the responses evenly split between open-source only (13%), proprietary only (14%), and undecided (16%). When looking towards the



future, the proportion of respondents who have not yet decided what type of software they will use increases to 25%, the proportion using proprietary-only software decreases to 2%, and the other proportions remain essentially unchanged (at 59% and 14%).

**Control over formats.** Twenty-seven percent of respondents indicated they had complete control over the format of content in their digital archives. Two-fifths (42%) work with content providers to influence the formats that they will accept, and one-third (31%) said they have little or no control and are obliged to accept the formats provided to them.

Thirty-eight percent of archives have complete control over formats, compared with 13% of libraries, and 45% of libraries having no control, compared with 27% of archives. This difference is even more marked when just national libraries and national archives are compared: 14 times more (56% vs. 4%) national archives than national libraries state that they can completely control the formats of the content they receive.

### **Digital Repositories**

**Important capabilities for a digital archive.** Respondents were asked to rate from 1 to 5 how important they thought various capabilities of a long-term digital information management system were. The ratings scale was: 1 = not applicable, 2 = least important, 5 = critical. So, any capability rated  $\geq 3$  is deemed important. The mean ratings assigned by respondents are given in Table 3, ordered by the mean rating from highest to lowest.

<b>Capability</b>	<b>Mean Rating</b>
Maintains authenticity, reliability, and integrity of records	3.8
Checks records have not been damaged	3.5
Plans the preservation of content to deal with technical obsolescence	3.4
Complies with established data or digital information management standards	3.4
Ensures records are accessible for up to 50 years	3.4
Performs migrations to deal with technical obsolescence	3.4
Is able to store many different types of content	3.3
Handles a wide variety of file formats	3.3
Ensures records are accessible for more than 50 years	3.3
Adheres to metadata standards	3.2
Retrieves content by description	3.0
Characterises records by extracting technical metadata	3.0
Integrates with content delivery systems	2.7
Retrieves content using full text	2.7
Supports emulation to deal with technical obsolescence	2.5
Integrates with content producing and holding systems	2.5
Checks for duplicate items	2.3

Table 3. The Important Capabilities for a Digital Archive to Have, as Rated by the Survey Respondents (135 responses).

For archives, the three key capabilities are (with their mean ratings): maintains authenticity, reliability, and integrity of records (3.8); ensures records are accessible for more than 50 years (3.5); and plans the preservation of content to deal with technical obsolescence (3.5). For libraries, the three key capabilities are: maintains authenticity, reliability, and integrity of records (3.8); is able to store many different types of content (3.7); and checks records have not been damaged (3.5). For government departments and the public sector in general, the three key capabilities are: maintains authenticity, reliability, and integrity of records (3.8); plans the preservation of content to deal with technical obsolescence (3.6), and complies with established data or digital information management standards (3.6).

**Scalability of digital archives.** Respondents were asked to rate the importance of scalability for digital archives, using a scale from 1 (not important) to 5 (critical). The mean ratings assigned to each aspect of scalability were: 3.8 for scalable to large volumes of data (petabytes of content), 3.7 for scalable to high ingest rates (millions of objects per year), and 3.1 for scalable to high access rates (hundreds of objects per second). Significantly (at the 95% confidence level), more national libraries (73%) rate scalability of content as critical than national archives (27%).

**Metadata standards.** The survey investigated the metadata standards used by organisations to describe stored digital objects. Dublin Core was the most popular standard, with 51% of respondents already using it, and 18% planning to use it. Machine-Readable Cataloguing (MARC) came next, with 34% already using it, and 5% planning to, followed by General International Standard Archival Description (ISAD(G)), with 28% already using it, and 10% planning to (see Figure 4).

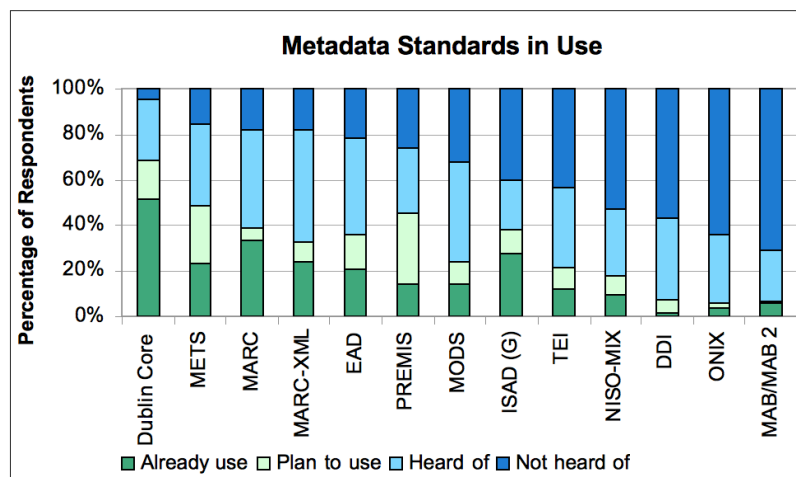


Figure 4. Metadata Standards Used by Respondents (134 Responses).

### **Policy and Implementation**

The overall results were further investigated by cross-correlating them with the information about which organisations have a digital preservation policy. Organisations with a digital preservation policy are less likely (3% vs. 11%) to have no experience, or be unaware of the challenges presented by, digital preservation, and nearly three times more likely (36% vs. 13%) to have a solution in place or planned. In addition, organisations with a policy are more likely to include digital preservation in their operational planning (92% vs. 60%), their business continuity (85% vs. 56%),



and financial planning (78% vs. 45%). Also, they are three times more likely to have a budget for digital preservation in place (72% vs. 23%).

Organisations with a policy are four times more likely (51% vs. 12%) to be investing in a solution now, and just 13% expect to wait longer than two years before investing, compared with 34% for those without a policy. Over three times (20% vs. 6%) as many organisations without a digital preservation policy as those with have no plans for the long-term management of digital information. Conversely, > 3 times (25% vs. 7%) as many organisations with a digital preservation policy as those without already have a long-term solution.

## Discussion

In 2005, there was widespread awareness within the information management community about the need to preserve digital content, but little action had been taken. Four years on, Planets' survey on long-term management of digital information indicates that significant strides have been made, in particular by those organisations that have established a digital preservation policy.


There was a relatively large response to the survey which included a cross-section of the major archives and libraries in Europe. The methods used to publicise the survey, and its inclusive nature, meant that although its primary target was European organisations, a fifth of responses were from outside Europe.

Digital preservation is not just a concern for archiving specialists in memory institutions such as archives and libraries. The ubiquity of digital information, and its importance in business, governmental, and private life, means that preservation of digital content is an issue that affects us all. Therefore, it is good to see a broad range of organisations responding to the survey, and in particular that some digital information producers are taking an interest in digital preservation.

Digital preservation is maturing as a discipline in its own right, so it is unsurprising that 15% of respondents specialise in this area. However, as demonstrated by respondents' roles, many of those involved in digital preservation still come from the more traditional backgrounds of preservation, curation, records management, and IT. Digital preservation is also drawing the attention of senior management; 11% of respondents were directors or heads of IT. The findings also indicated that producers of digital content (4% of respondents) are beginning to take an interest in the issue.

In contrast to the 2005 survey (Waller & Sharpe, [2006](#)), organisations now have a clear understanding of the volume of data they must archive. While the current storage needs of most organisations are quite modest, organisations predict a large increase in the volume of content over the next decade. At the same time, respondents need to preserve a wide range of types of digital information from a variety of sources. Almost all organisations expect to need to preserve digital objects not only in "simple" forms, such as documents, and images where some solutions already exist, but also in "complex" forms, such as databases, where solutions are still in development. Libraries, in particular, will need to preserve such dynamic content in the future.

Despite this need to deal with objects with behavioral properties, there was less interest expressed in emulation than migration. This may be because emulation is still



a subject for research, rather than a practical preservation strategy. However, it does point to a need for education and understanding about the role of emulation as a preservation strategy.


Over 9 in 10 respondents were aware of the issues and the challenges associated with digital preservation, reinforcing the findings of the earlier surveys. Half of the respondents' organisations had taken the vital first step of developing a digital preservation policy. Half have allocated a budget. However, where European organisations have a budget, it is five times more likely to be a capital-only one than a revenue-only one. The prevalence of capital over revenue budgets in Europe compared with North America may reflect the fact that many organisations are starting on the road to digital preservation, and therefore need a high capital expenditure to put a solution in place. In which case, we would expect the percentage of organisations with a revenue budget to increase over time, as the focus switches from the development of a digital preservation solution to its on-going maintenance, including both the ingest of new material, and the management of material already ingested. It is difficult to set a budget for on-going expenditure without experience of what the organisation needs to spend. It may also reflect the situation that many memory organisations operate under funding models where it is easier to obtain grants for individual projects than a long-term commitment from a funding body to support on-going investment.

Although awareness amongst respondents is high, it appears that organisations continue to face barriers to implementing solutions. Just one quarter currently has a solution in place or planned. Whether these barriers are due to lack of knowledge, lack of funding, or some other cause, such as low priorities, is not known. Caution should be applied in generalising this result, as those people who responded to a survey on digital preservation are more likely to be aware of the problems of digital preservation in the first place. However, findings indicate that those organisations that do plan to invest, plan to do so within the next two years.

Organisations are familiar with open-source solutions but are less familiar with commercial solutions. They plan to follow a route of component-based development and customization where a mix-and-match solution is used. Currently, open-source and proprietary software are used equally; however, findings indicate increased preference for open-source solutions in future. Such solutions need to be componentised with well-defined interfaces in order to fit in with the pick-and-mix approach used by organisations.

National archives are the most likely to develop, or have developed, a custom solution, reflecting the fact that many national archives have pioneered solutions to digital archiving. Conversely, government departments and the public sector are least likely to develop their own custom solution, and more likely to integrate components into a custom solution.

The ability of respondents' organisations to preserve digital content for the long term is limited by their ability to control the format of digital material that they need to store, mainly because such content is created externally. National archives are three times more likely to restrict the formats that they will accept than national libraries, suggesting that some digital preservation activities will have to occur before transfer to national archives. Much of the material that is transferred to national archives comes



from government departments and the public sector, but this is the group which is least likely to have a digital preservation policy. Therefore these organisations will need to develop such a policy in order to prescribe the process required to transfer digital material to the national archive in an orderly manner, as well as to cover the pre-transfer preservation activities. They may need assistance and education in order to overcome the problems they have.

Respondents are generally agreed about the key capabilities required of a digital preservation system. Such systems must maintain digital information for up to 50 years, in such a way as not to damage or corrupt it, and so that it can be accessed in future. Other important attributes in choosing a solution are the ability to plan preservation, and adherence to standards (although there is less clear agreement on which standards!).

Given the anticipated rises in volume, it is not surprising that scalability is generally regarded as one of the major criteria in assessing solutions. Given that libraries and archives predict that they will have similar levels of digital content in the future, it is surprising that archives are not as concerned about scalability, and scalability to total content in particular, as libraries. It is noticeable that scalability to high access rates is not ranked with the same importance as scalability to high volumes of content and high ingest rates. There are two possible explanations for this. One is that it reflects the fact that some organisations have restrictions on access; for archives this may be that parts of the collection are restricted for a period of time, and for libraries this may be that access is restricted to a specific group of users, such as on-site visitors to national libraries, or members of the university for academic libraries. The other explanation is that it indicates that organisations are preoccupied with ingest and storage, and have not yet reached the stage where users are requesting access to large volumes of content, which would again point to the relatively early stage of digital preservation.

The findings indicate that while archive, library, and related organisations are making progress towards long-term management of digital content, some are considerably further down the road of implementation than others. The results suggest a divide between those that have established a digital preservation policy and those that have not. The existence of a policy is a critical early step. Organisations with a policy are three times more likely to have a budget, and three times more likely to have either a solution in place or one planned for the near future, than those without a policy. This points to a need amongst those who are serious about maintaining access to digital content to start by gaining internal consensus about what must be preserved, for how long, and for whom, as a first step towards establishing an internal business case, and getting commitment to the task.

Organisations with a digital preservation policy currently store more data than organisations without a policy, although in 10 years' time the difference will have been almost completely eroded away. Similarly, more organisations with a digital preservation policy currently store each of the different types of digital information, but again, in 10 years' time there is very little difference between the two groups. It appears that organisations with little data, in relatively few formats, do not prioritise developing a digital preservation policy, whereas organisations facing the challenge of preserving large volumes of valuable content, or content in a wide variety of formats,

are taking steps to implement practical solutions. Over the next 10 years, the increasing need to preserve digital information is likely to provide an impetus for many to put a digital preservation policy and solution in place.

## Conclusions

The survey revealed that many organisations are beginning to make a transition from analysing the problem to solving it. They remain concerned that mature solutions do not yet exist. Nevertheless, 85% of organisations with a digital preservation policy expect to make an investment to create a digital preservation system within two years. Such systems are likely to be componentised, mix-and-match solutions. They will need to be scalable, particularly to handle the predicted large volumes of content, and also to handle high ingest rates. In addition, they will need to handle a wide range of formats from a variety of sources, and preserve the information contained therein for up to 50 years.

For organisations without a digital preservation policy, it is expected that the predicted increases in volume of digital information, and the range of formats needing to be preserved, will provide the impetus to focus on digital preservation and take practical steps to address its challenges.

## Acknowledgements

This work was carried as part of the Planets project (IST-033789). Planets is a four-year project co-funded by the EU under the 6<sup>th</sup> Framework Programme, and 16 European organisations, to address core digital preservation challenges. Planets is delivering a framework and set of practical tools and services that will enable institutions in Europe to manage and access digital collections for the long-term. For more information see: <http://www.planets-project.eu/>.

## References

- DigitalPreservationEurope (2007). *Market and technology trends analysis* (DPE-D6.1). Retrieved August 20, 2009, from <http://www.digitalpreservationeurope.eu/publications/dpe-market-analysis.pdf>.
- Farquhar, A., & Hockx-Yu, H. (2007, November). Planets: Integrated services for digital preservation. *International Journal of Digital Curation Vol. 2*,(2), 88-99. Retrieved November 6, 2009, from <http://www.ijdc.net/index.php/ijdc/article/viewFile/45/31>.
- Sinclair, P., & Jardine, L. (2009). *Survey Analysis Report* (DT11-D1). Planets. Manuscript in preparation.
- Waller, M., & Sharpe, R. (2006). *Mind the gap. Assessing digital preservation needs in the UK*. Digital Preservation Coalition. Retrieved November 6, 2009, from <http://www.dpconline.org/docs/reports/uknamindthegap.pdf>.