



|                                           |                            |
|-------------------------------------------|----------------------------|
| Project Number                            | IST-2006-033789            |
| Project Title                             | Planets                    |
| Title of Deliverable                      | Survey Analysis Report     |
| Deliverable Number                        | DT11-D1                    |
| Contributing Sub-project and Work-package | DT11                       |
| Deliverable Dissemination Level           | PU: Public                 |
| Deliverable Nature                        | Discussion Draft Report    |
| Contractual Delivery Date                 | 1 <sup>st</sup> April 2009 |
| Actual Delivery Date                      | 06 <sup>th</sup> May 2009  |
| Author(s)                                 | Tessella, British Library  |

**Contributors**

| Person           | Partner         |
|------------------|-----------------|
| James Duckworth  | Tessella        |
| Ann Keen         | Tessella        |
| Lewis Jardine    | Tessella        |
| Robert Sharpe    | Tessella        |
| Pauline Sinclair | Tessella        |
| Clive Billenness | British Library |
| Adam Farquhar    | British Library |
| Jane Humphries   | British Library |

**Document Approval**

| Person              | Role                    | Partner               |
|---------------------|-------------------------|-----------------------|
| Raymond van Diessen | Scientific Board Member | IBM                   |
| Manfred Thaller     | Scientific Board Member | University of Cologne |

**Distribution**

| Person                   | Partner |
|--------------------------|---------|
| PLANETS Scientific Board | All     |

**Revision History**

| Issue | Author           | Date        | Description                                            |
|-------|------------------|-------------|--------------------------------------------------------|
| 0.1   | Lewis Jardine    | 20-Mar-2009 | Initial draft version for internal review              |
| 0.2   | Pauline Sinclair | 17-Apr-2009 | Second draft version for internal review               |
| 0.3   | Pauline Sinclair | 20-Apr-2009 | Updated draft following internal review                |
| 0.4   | Pauline Sinclair | 29-Apr-2009 | Updated following external review.                     |
| 0.5   | Pauline Sinclair | 05-May-2009 | Updated following second round of reviews.             |
| 0.6   | Pauline Sinclair | 08-May-2009 | Updated following further review                       |
| 0.7   | Pauline Sinclair | 22-May-2009 | Updated following review by the scientific board.      |
| 1.0   | Pauline Sinclair | 03-Jun-2009 | Updated as part of the document approval process.      |
| 1.1   | Pauline Sinclair |             | Corrected various minor errors (mainly typographical). |

## EXECUTIVE SUMMARY

In early 2009 Planets conducted an on-line survey of people with an interest in digital preservation. In particular, national libraries and archives in Europe were targeted. The objectives of the survey were to:

- Better understand digital preservation needs within European organisations that create or hold digital content
- Establish and begin to quantify and qualify demand for Planets' tools and services
- Identify the most appropriate packaging for Planets to meet market requirements
- Inform the design of a follow-on organisation to sustain Planets' products and services after 2010.

The survey had a relatively large number of responses (over 200). While the respondents were from a wide variety of organisations, approximately two fifths were from libraries and a third from archives.

One positive finding of the survey is that organisations have become aware of the problems of digital preservation, even though they do not currently know how to solve them. It is clear from a number of the survey results that forming a digital preservation policy is one of the key steps in obtaining a digital preservation solution. Not surprisingly, organisations look to have a policy in place before embarking on significant work or allocating any budget.

Respondents need to preserve a wide range of types of digital information from a variety of sources. Libraries need to preserve more dynamic content than most organisations, and they have the widest range of different types. Interestingly, many archives can dictate which formats they receive digital material in (and thus which formats they need to preserve) but libraries do not have this control.

Respondents also have significant quantities of data to preserve. The average organisation currently stores less than 20 TB, which rises to over 500 TB in ten years' time. It is no surprise, then, that they think any solution to managing this data needs to be scalable with respect to total content and also high ingest rates.

Respondents regard the maintenance of their digital information in such a way as to demonstrate that it will not be damaged or corrupted for up to 50 years as the key capability of any digital preservation solution. Planning how to do this and adherence to standards are also very important. For libraries, the ability to handle a wide range of digital information and formats is key, reflecting the type of content they have to preserve. In general respondents deem emulation and integrating a digital archive with content producing, holding and delivery systems as less important.

Organisations are following a route of component-based development and customisation where a mix-and-match solution is used. Currently, open-source and proprietary software are used equally in solutions to digital preservation, mainly combined in the same solution. In the future only 2% of solutions are projected to be purely proprietary, a seventh of current levels. The high numbers using a mixture of open-source and proprietary software and combining components into a custom system means that digital preservation software needs to be componentised with well-defined interfaces.

As for Planets, respondents are generally aware of Planets and would like to be kept informed about Planets' work. However, the awareness of Planets is lower in academia, government departments and the public sector in general than in national archives and libraries.

There is considerable interest in the capabilities Planets technology could offer, but less interest in direct services. Many respondents hope to see Planets providing tools, especially for preservation planning, characterisation and format transformation, and also contributing to technical standards and providing information about the latest developments.

Unsurprisingly, given the tendency for people to want things for free, there is less interest in paying for Planets' output. Even so, there is significant possible interest in paying for Planets' software and support and paying more for certified software, although software vendors are not keen to pay to have their software certified.

## KEY FINDINGS

The significant findings of this report are highlighted here.

- Over 90% of organisations that responded to the survey are aware of the challenges presented by the long-term management of digital information.
- Half of all organisations that responded to the survey have a digital preservation policy in place, but only a quarter of government departments and public sector organisations that responded to the survey have one.
- Three-quarters of organisations that responded to the survey include digital preservation in their operational planning.
- Organisations with digital preservation policies are 3 times more likely to have a budget for digital preservation.
- Organisations with a digital preservation policy are 3 times more likely to have a digital preservation solution in place or planned than organisations without one.
- Organisations without a digital preservation policy are 4 times more likely to be unaware, or have no experience, of the challenges presented by digital preservation.
- 85% of respondents are working towards a digital preservation solution or have one in place.
- Only 6% of organisations with a digital preservation policy, compared with 20% of organisations without one, have no plans for digital preservation.
- Over 85% of organisations with a digital preservation policy expect to make an investment within 2 years.
- Less than a third of organisations think that they have complete control over the formats that they will accept and enter into their archives.
- Both national libraries and national archives predict large increases in the volume of material they need to store over the next ten years although, interestingly, national libraries are three times more likely than national archives to regard scalability of content as critical.
- Currently over 70% of organisations hold less than 100 TB (with a median of less than 20TB), whereas in 10 years' time over 70% of organisations expect to store over 100 TB (with a median of over 500TB).
- Over half of organisations expect to preserve digital information from file systems, document scanning programmes, the internet, electronic document management systems, e-mail systems and media digitisation programmes.
- In 10 years' time, 99% of organisations will need to preserve documents, 95% will need to preserve images and 85% will need to preserve databases.
- Nearly 40% more organisations with a digital preservation policy need to store digital information at the moment; in 10 years' time there is no difference: as many organisations without a digital preservation policy as with will need to store digital information of all types.
- The most important capability of a long-term digital information management system is that it must maintain the authenticity, reliability and integrity of records.
- 75% of respondents would like to receive electronic updates about Planets.

## TABLE OF CONTENTS

|            |                                                                           |    |
|------------|---------------------------------------------------------------------------|----|
| 1.         | Introduction .....                                                        | 8  |
| 2.         | Analysis .....                                                            | 9  |
| 2.1        | Respondents' Roles .....                                                  | 9  |
| 2.2        | Country of the Respondent .....                                           | 9  |
| 2.3        | Organisation Type .....                                                   | 10 |
| 2.4        | How Respondents Stay Informed.....                                        | 11 |
| 2.5        | Organisations with Digital Preservation Policies .....                    | 11 |
| 2.6        | Digital Preservation Policy versus Budget.....                            | 12 |
| 2.7        | Level of Awareness within Organisations.....                              | 14 |
| 2.8        | Digital Preservation's Inclusion in Organisations' General Planning ..... | 14 |
| 2.9        | Specific Plans for Digital Preservation .....                             | 17 |
| 2.10       | How the Solution will be Implemented.....                                 | 19 |
| 2.11       | Sources of Digital Content for Archiving .....                            | 21 |
| 2.12       | Digital Information Types that Need to be Managed in the Long-Term.....   | 22 |
| 2.13       | Control over the Digital Formats Archived .....                           | 25 |
| 2.14       | Important Capabilities for Digital Archives .....                         | 26 |
| 2.15       | The Importance of Scalability for Digital Archives .....                  | 29 |
| 2.16       | Quantities of Digital Content Being Stored .....                          | 31 |
| 3.         | Analysis of Planets Specific Questions.....                               | 34 |
| 3.1        | Awareness of Planets.....                                                 | 34 |
| 3.2        | Receiving Electronic Updates About Planets.....                           | 35 |
| 3.3        | Rating Planets Capabilities .....                                         | 36 |
| 3.4        | Rating Planets Services .....                                             | 38 |
| 3.5        | Funding Planets.....                                                      | 39 |
| 3.6        | Summary of Interest in Planets' Capabilities and Services.....            | 42 |
| 4.         | Conclusions.....                                                          | 44 |
| 4.1        | Implications for Planets .....                                            | 44 |
| Appendix A | Summary of Survey Results .....                                           | 46 |
| A.1        | About you and your organisation.....                                      | 46 |
| A.2        | Policies .....                                                            | 48 |
| A.3        | Current Status and Plans .....                                            | 49 |
| A.4        | Digital content.....                                                      | 51 |
| A.5        | Functionality - Capabilities .....                                        | 53 |
| A.6        | Planets.....                                                              | 54 |

## TABLE OF FIGURES

|                                                                                                                                                                                                                                  |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1: Roles of Survey Respondents (178 Responses) .....                                                                                                                                                                      | 9  |
| Figure 2: Survey Respondents by Region (196 Responses) .....                                                                                                                                                                     | 10 |
| Figure 3: Survey Respondents by Organisation Type (183 Responses) .....                                                                                                                                                          | 10 |
| Figure 4: Mechanisms Respondents Use to Stay Informed (181 Responses) .....                                                                                                                                                      | 11 |
| Figure 5: Organisations with Long-Term Digital Information Management Policies (161 Responses) .....                                                                                                                             | 11 |
| Figure 6: Digital Preservation Budgets versus Digital Preservation Policy (153 Responses) .....                                                                                                                                  | 12 |
| Figure 7: Digital Preservation Budgets by Organisation Type (153 Responses) .....                                                                                                                                                | 13 |
| Figure 8: Digital Preservation Budgets by Region (153 Responses) .....                                                                                                                                                           | 13 |
| Figure 9: Level of Awareness of Digital Preservation Within Organisations (157 Responses) .....                                                                                                                                  | 14 |
| Figure 10: Whether Digital Preservation Features in Organisations' Operational Planning (153 Responses) .....                                                                                                                    | 15 |
| Figure 11: Whether Digital Preservation Features in Organisations' Business Continuity Planning (154 Responses) .....                                                                                                            | 15 |
| Figure 12: Whether Digital Preservation Features in Organisations' Financial Planning (153 Responses) .....                                                                                                                      | 15 |
| Figure 13: Whether Digital Preservation Features in the Operational Planning of Different Types of Organisation (153 Responses) .....                                                                                            | 16 |
| Figure 14: Whether Digital Preservation Features in the Business Continuity Planning of Different Types of Organisation (154 Responses) .....                                                                                    | 16 |
| Figure 15: Whether Digital Preservation Features in the Financial Planning of Different Types of Organisation (153 Responses) .....                                                                                              | 16 |
| Figure 16: Organisations' Plans for Digital Preservation for those Organisations With and Without a Digital Preservation Policy (134 Responses) .....                                                                            | 18 |
| Figure 17: Organisations' Plans for Digital Preservation for those Organisations in Europe (103 Responses) and North America (23 Responses) .....                                                                                | 18 |
| Figure 18: Investment Timescales for Organisations with and without a Digital Preservation Policy (123 Responses) .....                                                                                                          | 19 |
| Figure 19: How Different Organisation Types Expect to Implement Their Digital Preservation Solution (132 Responses) .....                                                                                                        | 20 |
| Figure 20: Current and Expected Use of Open Source and Proprietary Software in Organisations' Digital Preservation Solutions (139 Responses) .....                                                                               | 20 |
| Figure 21: Source Systems that Organisations Currently, or Expect to in the Future, Preserve Digital Material from (141 Responses) .....                                                                                         | 21 |
| Figure 22: Types of Digital Information that Organisations Currently, or Expect to in the Future, Preserve (138 Responses) .....                                                                                                 | 22 |
| Figure 23: Types of Digital Information that Different Organisation Types Expect to Need to Preserve Within 10 Years (138 Responses) .....                                                                                       | 22 |
| Figure 24: Types of Digital Information that Organisations with and without a Digital Preservation Policy Currently Need to Preserve (136 Responses) .....                                                                       | 23 |
| Figure 25: Types of Digital Information that Organisations with and without a Digital Preservation Policy will Need to Preserve in Two Years' Time that they do not Currently Need to Preserve (136 Responses) .....             | 24 |
| Figure 26: Types of Digital Information that Organisations with and without a Digital Preservation Policy will Need to Preserve in Five Years' Time that they will not Need to Preserve in Two Years' Time (136 Responses) ..... | 24 |
| Figure 27: Types of Digital Information that Organisations with and without a Digital Preservation Policy will Need to Preserve in Ten Years' Time that they will not Need to Preserve in Five Years' Time (136 Responses) ..... | 24 |
| Figure 28: Types of Digital Information that over the next Ten Years, Organisations with and without a Digital Preservation Policy Will Need to Preserve (136 Responses) .....                                                   | 25 |
| Figure 29: Level of Control that Different Types of Organisations have over the Formats of the Content in their Digital Archives (137 Responses) .....                                                                           | 25 |
| Figure 30: Level of Control that Organisations with and without a Digital Preservation Policy have over the Formats of the Content in their Digital Archives (137 Responses) .....                                               | 26 |
| Figure 31: Importance of Different Capabilities of a Digital Archive to Respondents from Different Types of Organisations (135 Responses) .....                                                                                  | 27 |
| Figure 32: Importance of Different Capabilities of a Digital Archive to Respondents from Different Regions (135 Responses) .....                                                                                                 | 28 |

|                                                                                                                                                                                 |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 33: Importance of the Scalability of a Digital Archive to High Content Volumes (Petabytes) to Different Types of Organisation (133 Responses) .....                      | 30 |
| Figure 34: Importance of the Scalability of a Digital Archive to High Ingest Rates (Millions of Objects per Year) to Different Types of Organisation (133 Responses) .....      | 30 |
| Figure 35: Importance of the Scalability of a Digital Archive to High Access Rates (Hundreds of Objects per Second) to Different Types of Organisation (133 Responses) .....    | 30 |
| Figure 36: Growth in Volumes of Digital Content that Organisations Intend to Store over the next Ten Years (129 Responses).....                                                 | 31 |
| Figure 37: Volumes of Digital Content that Different Types of Organisation Currently Store (126 Responses).....                                                                 | 32 |
| Figure 38: Volumes of Digital Content that Different Types of Organisation Intend to Store in Ten Years' Time (126 Responses) .....                                             | 32 |
| Figure 39: Growth in Volumes of Digital Content that Organisations with and without a Digital Preservation Policy Intend to Store over the next Ten Years (126 Responses) ..... | 33 |
| Figure 40: Awareness of Planets within Different Types of Organisation (137 Responses) .....                                                                                    | 34 |
| Figure 41: Awareness of Planets in Organisations in Different Regions (137 Responses) .....                                                                                     | 35 |
| Figure 42: Awareness of Planets in Organisations with and without a Digital Preservation Policy (137 Responses) .....                                                           | 35 |
| Figure 43: Interest in Receiving Planets' Electronic Mailings from Respondents in Different Regions (196 Responses) .....                                                       | 36 |
| Figure 44: Interest in Receiving Planets' Electronic Mailings from Respondents in Different Types of Organisation (196 Responses) .....                                         | 36 |
| Figure 45: Interest in Planets' Capabilities from Respondents in Different Types of Organisation (133 Responses) .....                                                          | 37 |
| Figure 46: Interest in Planets' Capabilities from Respondents in Different Regions (133 Responses) .....                                                                        | 38 |
| Figure 47: Interest in Planets' Services from Respondents in Different Types of Organisation (131 Responses).....                                                               | 38 |
| Figure 48: Interest in Planets' Services from Respondents in Different Region (131 Responses) ..                                                                                | 39 |
| Figure 49: Interest in Paying for the Outputs from Planets (130 Responses) .....                                                                                                | 40 |
| Figure 50: Interest from Respondents from Different Types of Organisation in Paying for Planets' Outputs (130 Responses).....                                                   | 40 |
| Figure 51: Interest from Respondents from Different Regions in Paying for Planets' Outputs (130 Responses).....                                                                 | 41 |
| Figure 52: Interest in Paying for Planets' Outputs, Correlated with who will Implement the Solution (118 Responses) .....                                                       | 41 |

## TABLE OF TABLES

|                                                                                                                                     |    |
|-------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 1: Countries with Five or More Respondents .....                                                                              | 9  |
| Table 2: Cross-Correlation of Organisations' Plans for Digital Preservation (135 Responses) .....                                   | 17 |
| Table 3: Cross-Correlation of How Respondents Expect to Implement Their Digital Preservation Solution (132 Responses) .....         | 19 |
| Table 4: The Important Capabilities for a Digital Archive to have, as Rated by the Survey Respondents (135 Responses) .....         | 27 |
| Table 5: Average Volumes of Digital Content Organisations Store Now and Intend to Store in Ten Years' Time (129 Responses) .....    | 31 |
| Table 6: Interest in Paying for Planets' Offerings (130 Responses) .....                                                            | 42 |
| Table 7: Relative Rankings of Planets' Potential Capabilities and Services Correlated with Interest in Paying (133 Responses) ..... | 42 |

---

## 1. Introduction

At the start of 2009, Planets undertook a survey of national libraries, archives and other organisations in Europe to better understand their digital preservation needs and ensure Planets' technology and services are designed to meet them. This report analyses the results of the survey and draws some conclusions from them.

The survey was publicised widely and in a structured and targeted way. The survey was made available for a month in early 2009. Approximately 2000 emails were issued by the project to named individuals in libraries, archives and other organisations across Europe whose role could encompass the long-term maintenance of digital information; in particular, individuals in national archives and national libraries were targeted. Follow-up telephone calls were made to 120 of the named individuals to encourage them to take part in the survey. Additionally, the survey was publicised, by an initial announcement and two reminders, on about 30 relevant mailing lists ranging from international digital preservation mailing lists such as PADI (Preserving Access to Digital Information), to specialist mailing lists targeting sub-sections of the digital preservation community, such as those people in research institutes, government, and audiovisual (film & sound) archives. There was also a news item on Planets' website about the survey. Direct approaches were made to contacts in organisations and projects with an interest in digital preservation such as Digital Preservation Europe, the Digital Curation Centre, and the Caspar, Shaman and Protage projects, all of whom were asked to cascade notices on Planets behalf. Contact was also made with the Council of European National Libraries (CENL), International Council on Archives (ICA) and the Association of European Research Libraries (LIBER) who were asked to disseminate the message to their members.

The survey sought to target the potential user community rather than suppliers and vendors. Instead, Planets is engaging separately, and face-to-face, with almost 20 suppliers and vendors at a briefing in June 2009. The accuracy of the targeting is reflected in the high proportion of respondents (93%) derived from the user community. It was also possible to filter responses by organisation type to ensure the user community's views were appropriately analysed.

All the publicity resulted in 206 responses to the survey before it closed. The relatively large number of responses means that the results could be analysed in some depth. However, many of the respondents did not answer all the survey questions and indeed the number of responses declined throughout the survey (although 142 completed the full survey).

The appendix contains the survey results summarised question by question, while the body of the report contains comments on and further analysis of the data. Where ratings are used, respondents were offered a range from 1 to 5; the ratings have different meanings for different questions.

Note that the small sample size for respondents from outside Europe and North America means that where results have been analysed by region the answers from these respondents have not been included as a separate group.

Also note that because of the sample size for museums, academic archives, software developers or vendors, systems integrators or consultants and repository services providers is small they will not generally be examined independently by organisation type where results are broken down in this manner.



## 2. Analysis

### 2.1 Respondents' Roles

The long-term management of digital information straddles two distinct professions – IT and curation and preservation – although people have started to specialise in digital preservation. It is therefore of no surprise that respondents to the survey mainly come from these professions. About 15% of respondents work in digital preservation, 16% in preservation in general, 22% in curation (be that in an archive, a library, or looking after records in any organisation) and 16% in IT. The remainder work in a variety of professions including management and research. There are even a few representatives (8) of the producers of digital information, which then needs preserving.

It is good to see that data producers are starting to take an interest in digital preservation, since they are in a unique position to take steps, such as providing descriptions of their data, to help preserve their output. The earlier steps, such as gathering metadata about it and deciding what file format to store it in and on what physical medium, are taken towards preserving digital information the easier it is to preserve.

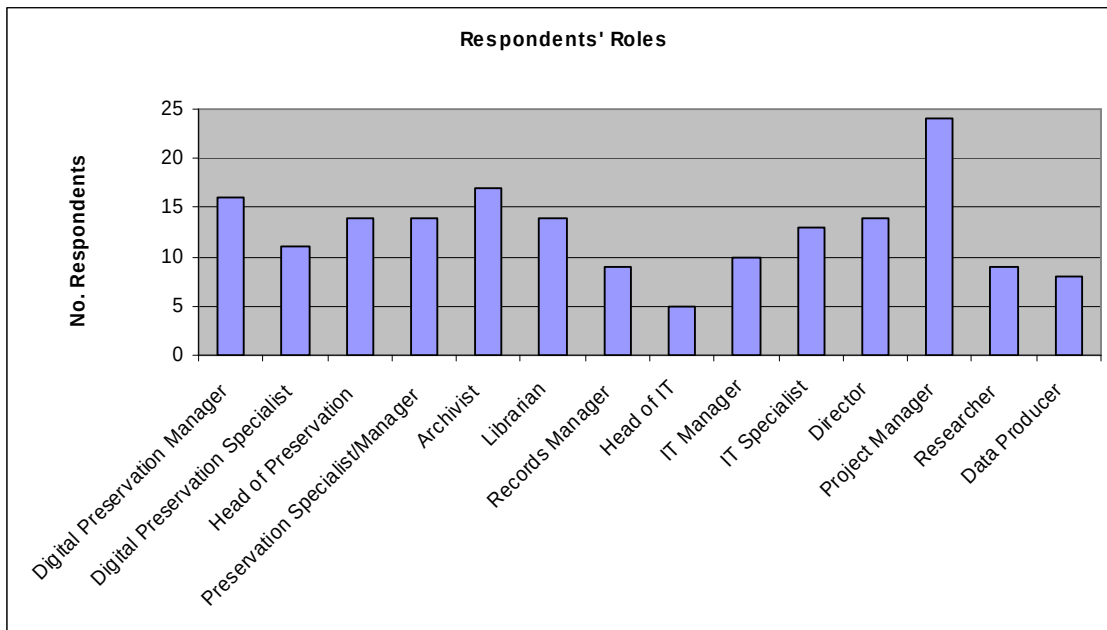


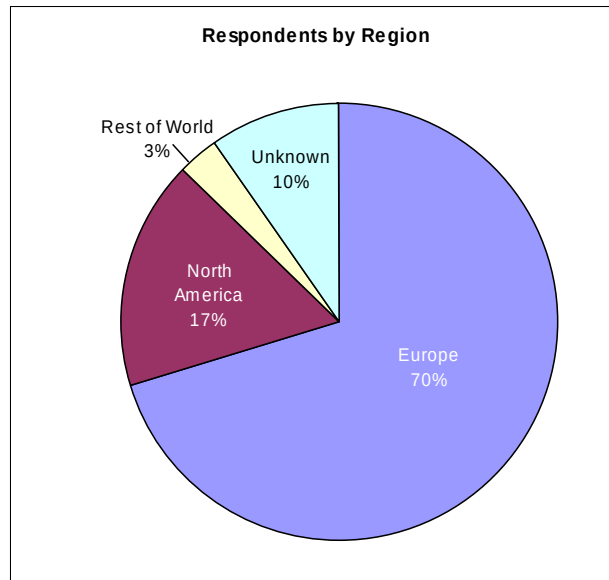
Figure 1: Roles of Survey Respondents (178 Responses)

### 2.2 Country of the Respondent

As expected, given the target audience of the survey, a majority of the respondents (65%) were from European Union countries. However there were significant numbers of respondents from elsewhere in Europe and the rest of the world. Indeed, Americans were the second largest group of respondents, after the British.

| Country     | Number of Respondents |
|-------------|-----------------------|
| UK          | 54                    |
| USA         | 26                    |
| Germany     | 16                    |
| Switzerland | 15                    |
| Netherlands | 10                    |
| Canada      | 7                     |
| Belgium     | 6                     |

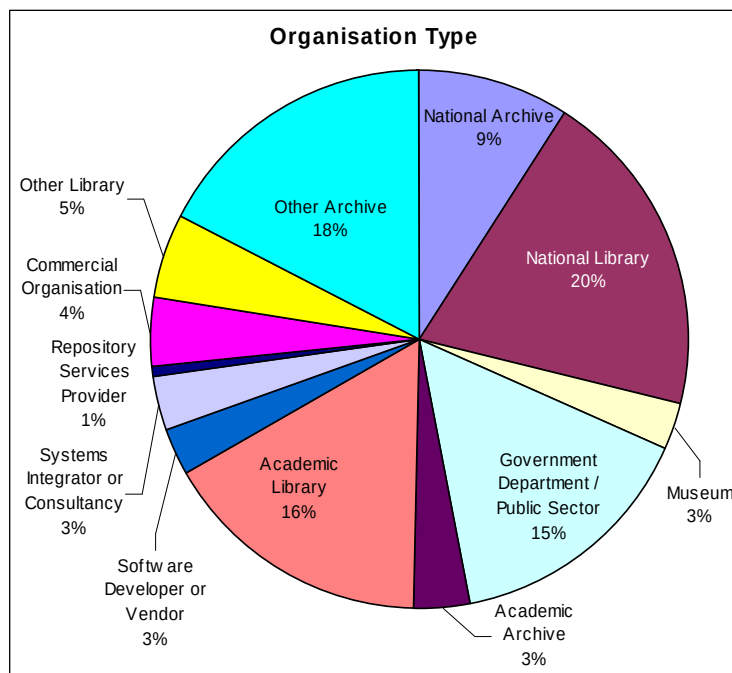
Table 1: Countries with Five or More Respondents



**Figure 2: Survey Respondents by Region (196 Responses)**

### 2.3 Organisation Type

Given the widespread potential impact of digital preservation (or more specifically the consequences of not preserving digital information), the subject is likely to be of interest to a diverse community. This survey sought to recruit participants from organisations across Europe that create or hold digital content. In particular, the 96 national archives and libraries in Europe were targeted through direct emails to individuals within these organisations, and in some cases follow-up telephone calls were made to encourage participation. The recruitment process also used specific mailing lists to target these and other communities such as academia, government and the public sector. The success of this approach in reaching its target audience is reflected in the proportion of respondents from libraries (75, or 41%), archives (55, or 30%) and government and the public sector (28, or 15%). In addition, there were 12 responses (7%) from suppliers and vendors, 8 responses (4%) from commercial organisations and 5 responses (3%) from museums.



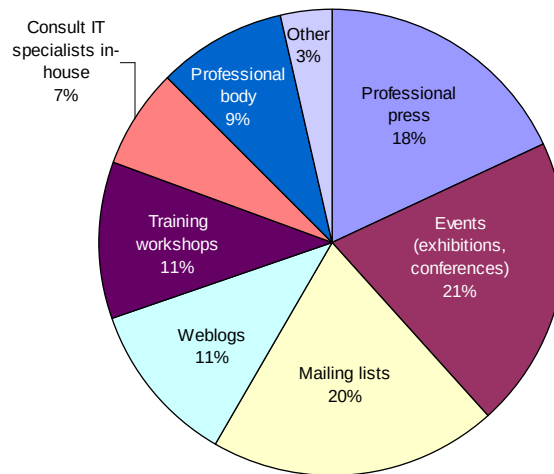
**Figure 3: Survey Respondents by Organisation Type (183 Responses)**

### 2.3.1 Implications for Planets

The diversity of respondents to the survey illustrates the difficulty of reaching all the potential “customers” for Planets.

## 2.4 How Respondents Stay Informed

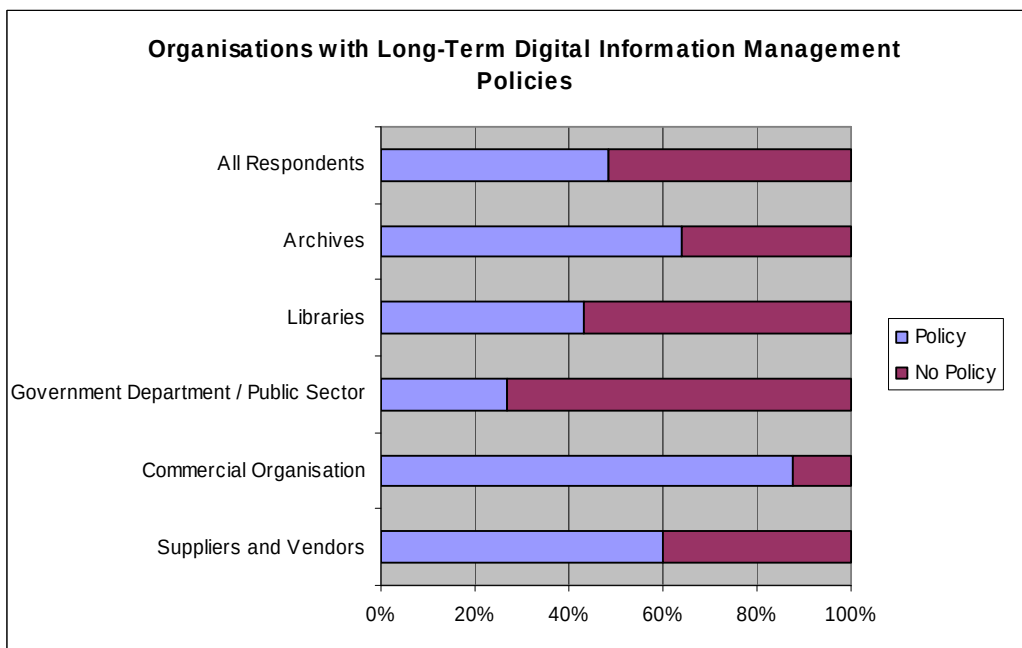
The major mechanisms respondents use for keeping themselves informed of developments in the long term management of digital information are professional press, events and mailing lists. However, other mechanisms also play a significant role in disseminating information.



**Figure 4: Mechanisms Respondents Use to Stay Informed (181 Responses)**

## 2.5 Organisations with Digital Preservation Policies

Overall, respondents are evenly split as to whether their organisation has a policy (48%) for the long-term (greater than 5 years) management of digital information in place. However, when respondents are grouped by the type of their organisation, some interesting trends emerge.



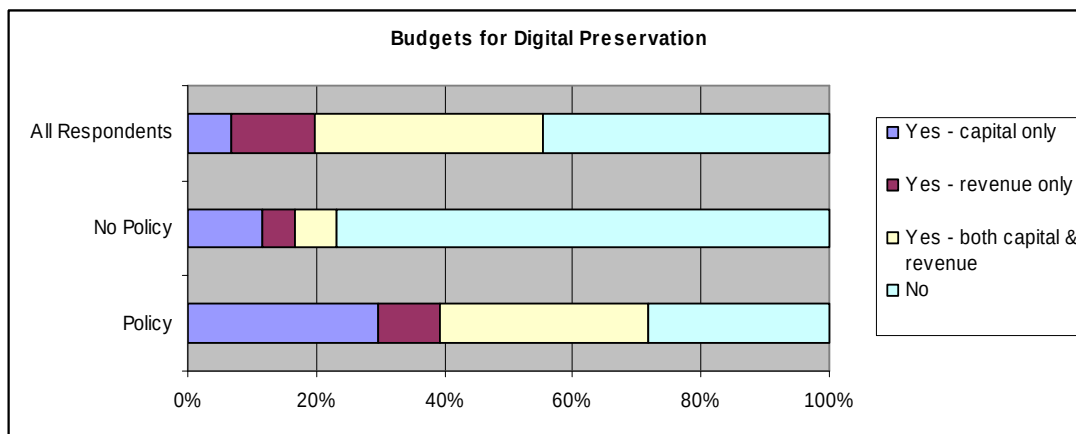
**Figure 5: Organisations with Long-Term Digital Information Management Policies (161 Responses)**

The majority (64%) of archives plan to manage digital information in the long-term, whereas fewer than half (43%) of the libraries do. It is surprising that commercial organisations and supplier and vendors combined appear more likely to have a policy in place to manage digital information in the long-term than libraries and archives combined. Of course this may not be significant since there were only 18 respondents in the first two categories compared to 105 in the latter two, and since the survey was not tailored to this community, those that responded may not be representative. Nonetheless it is an interesting finding that could be followed up in future research. If it is significant it is presumably because of accountability legislation, such as data protection and the Sarbanes-Oxley Act, and commercial pressures, such as protecting their Intellectual Property.

Government departments and the public sector are the least likely to have a digital preservation policy, with only a quarter (27%) of such organisations having one. This may be because government departments regard digital preservation as a function of the national archives, not the individual departments. However, they do still need a digital preservation policy for two reasons. Firstly, to prescribe the process required to transfer digital material to the national archive in an orderly manner and secondly because some digital preservation activities will need to occur before the material is transferred (see section 2.13 below).

## 2.6 Digital Preservation Policy versus Budget

As with digital preservation policies, only half of the organisations surveyed had specific budgets, whether capital or revenue, for dealing with the long-term management of digital preservation. Comparing those organisations with a digital preservation policy with those organisations with a budget for digital preservation, it can be seen that organisations with digital preservation policies are 3 times more likely to have a budget for digital preservation.



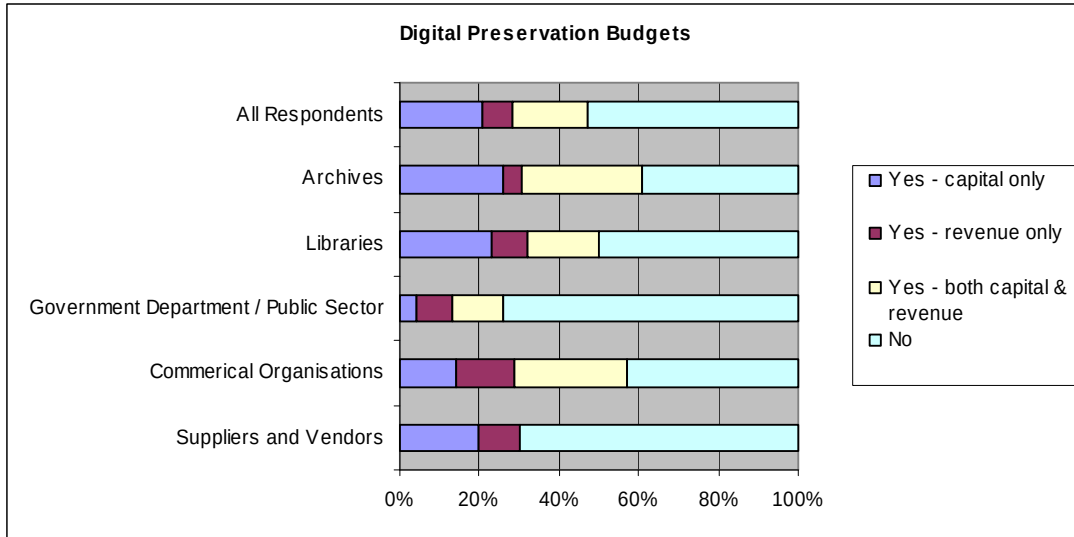
**Figure 6: Digital Preservation Budgets versus Digital Preservation Policy (153 Responses)**

Of those organisations who have a budget in place, it is mainly for capital expenditure; only a third of them have a revenue budget. This may reflect the fact that most organisations are just starting on the road to digital preservation and therefore need a high capital expenditure to put a solution in place. However, they do need to be aware of the on-going cost of digital preservation, since studies of the cost of digital preservation show significant on-going maintenance costs both in ingesting new material and in managing material already ingested. Admittedly, it is hard to set a budget for such on-going expenditure without some experience of what the organisation needs to spend; so the lack of revenue budgets may be a reflection of this lack of experience with the on-going costs of digital preservation. Comparing the type of budget to the stage that the organisations digital preservation plans are at, does not show any real difference between the three groups (capital only, revenue only, and both capital and revenue) and so does not prove or disprove any of the possible theories for the lack of revenue budgets.

When the results are broken down by sector, it can be seen that archives are most likely to have a budget for digital preservation, with three-fifths of them having one, whereas government departments and the public sector in general is the least likely to have a budget for digital preservation, with only a quarter of them having one.

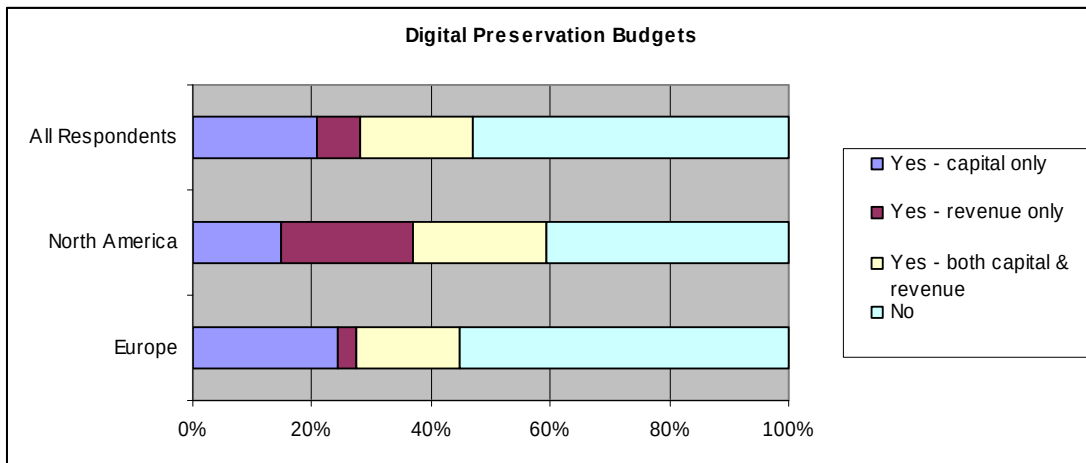
North American organisations are more likely to have a budget, than not, with nearly 60% of respondents having a digital preservation budget, and this budget is fairly evenly split between

capital only, revenue only and both capital and revenue. This is not because more North American organisations have digital preservation policies than the average, in fact the reverse is true and North American organisations are slightly less likely to have a digital preservation policy (43%) than the average organisation (48%) in the survey.



**Figure 7: Digital Preservation Budgets by Organisation Type (153 Responses)**

Where European organisations have a budget, it is 5 times as likely to be a capital only one than a revenue only one. Unsurprisingly, over half (56%) of European organisations with a capital only budget have a long-term solution in development, with most of the remainder assessing their requirements. The proportions are similar for those European organisations with both capital and revenue budgets, whereas the few European organisations with a revenue only budget are evenly spread amongst assessing requirements, tendering for a long-term solution and developing a long-term solution. This may indicate that in Europe digital preservation is seen as a one-off activity which can be funded from capital, rather than requiring an on-going revenue budget. However, it may reflect the situation that many European memory institutions 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 an on-going operation.



**Figure 8: Digital Preservation Budgets by Region (153 Responses)**

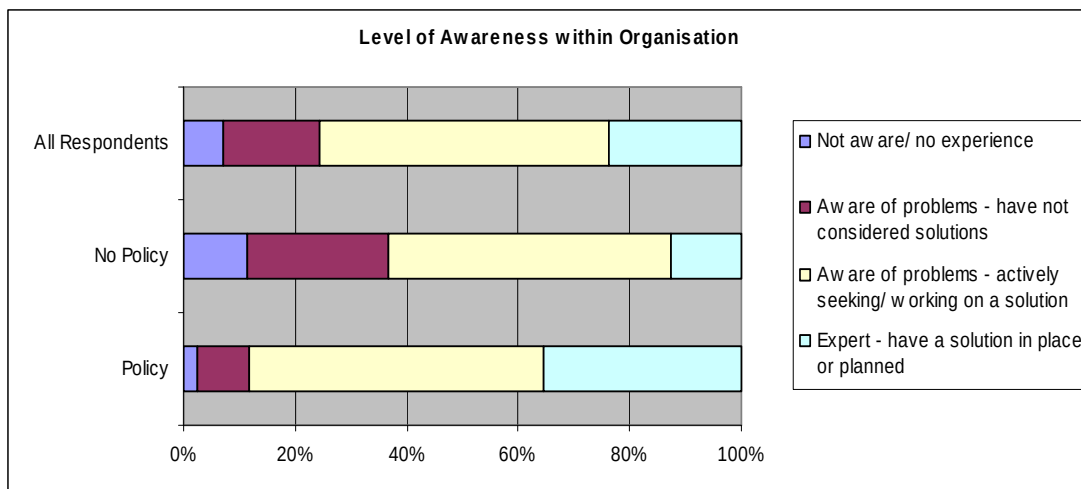
#### 2.6.1 Implications for Planets

Given that organisations with digital preservation policies are 3 times more likely to have a budget for digital preservation, a good initial starting point in promoting Planets would be to ask people whether they have a digital preservation policy and for those who do not to offer some kind of initial guide towards forming a digital preservation policy.

## 2.7 Level of Awareness within Organisations

Pleasingly, there is a high level of awareness of the problems of digital preservation within organisations, with 93% of organisations aware of the problems, although 17% of these organisations have not yet considered any solutions. Although awareness amongst respondents is high, it appears that organisations face barriers to implementing solutions, since less than a quarter of them already have 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 filled in a Planets' survey on digital preservation are more likely to be aware of the problems of digital preservation in the first place.

Having a digital preservation policy is an indication that a solution is either in place or planned. However, half of all organisations (52%) are actively seeking or working on a digital preservation solution, regardless of whether they have a policy or not. Organisations without a digital preservation policy are 4 times more likely to have no experience or be unaware of the challenges presented by digital preservation and nearly 3 times less likely to have a solution in place or planned.



**Figure 9: Level of Awareness of Digital Preservation Within Organisations (157 Responses)**

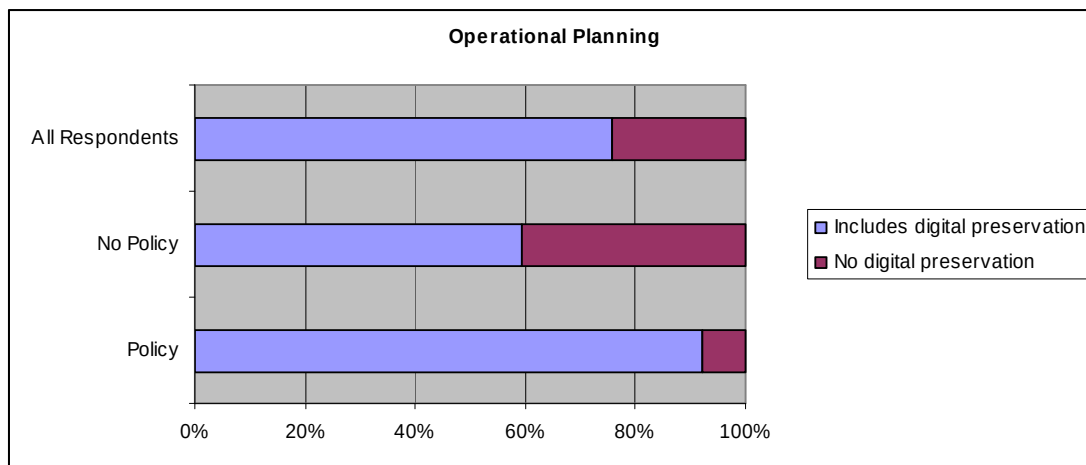
### 2.7.1 Implications for Planets

The lack of digital preservation solutions (three-quarters of organisations do not have one), despite the awareness of the problems, is an opportunity for Planets. Planets could help with information advice, and tools.

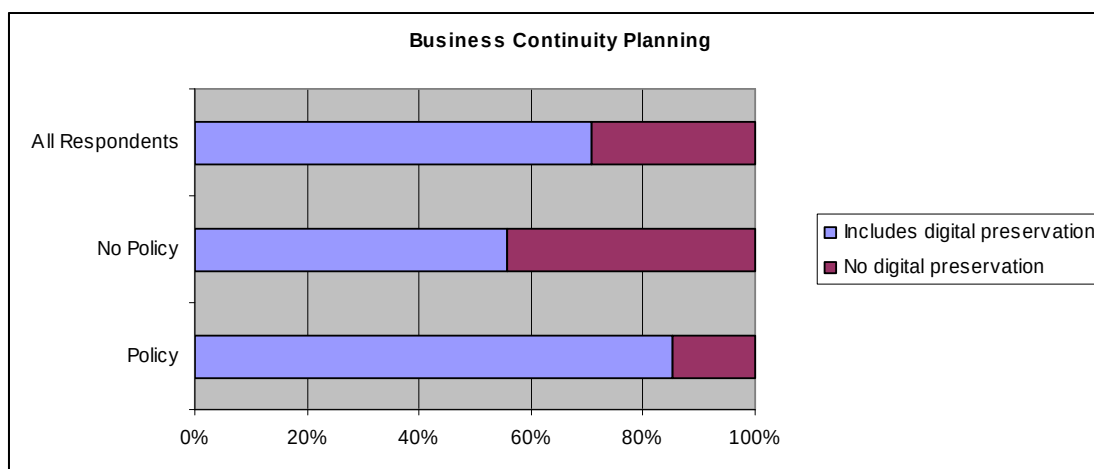
## 2.8 Digital Preservation's Inclusion in Organisations' General Planning

Following on from the high level of awareness of digital preservation, a majority of organisations include digital preservation in their operational planning (76%), business continuity planning (71%) and financial planning (62%). So organisations are not only aware of digital preservation, but they are actively starting to do something about it. Whether these plans are sufficient is not known, but it is a good start.

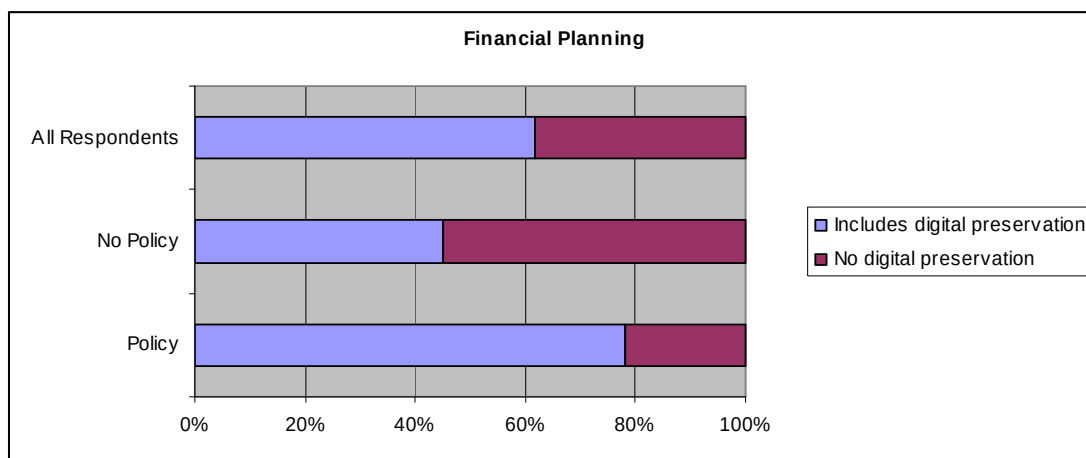
Once again it is organisations with a digital preservation policy which are farthest advanced. Over 90% include digital preservation in their operational planning, compared with just 60% of organisations without a digital preservation policy. Similarly, having a digital preservation policy increases the likelihood that digital preservation forms part of an organisation's business continuity planning (from 56% to 85%) and its financial planning (from 45% to 78%).



**Figure 10: Whether Digital Preservation Features in Organisations' Operational Planning (153 Responses)**

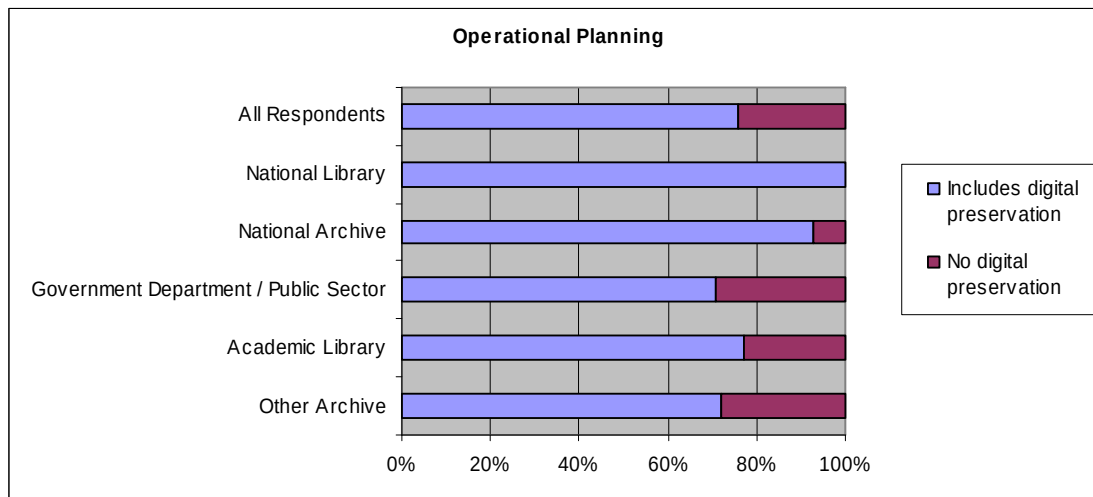


**Figure 11: Whether Digital Preservation Features in Organisations' Business Continuity Planning (154 Responses)**

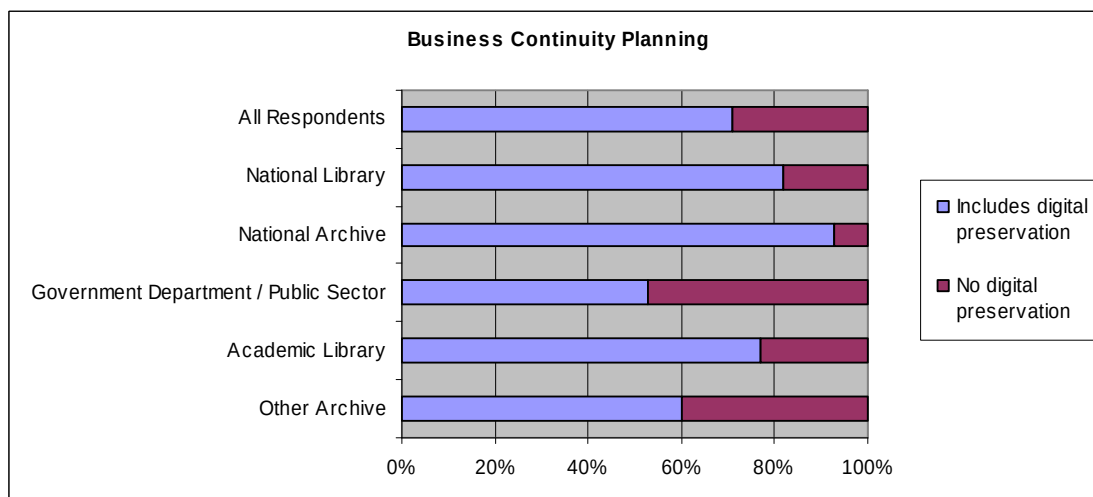


**Figure 12: Whether Digital Preservation Features in Organisations' Financial Planning (153 Responses)**

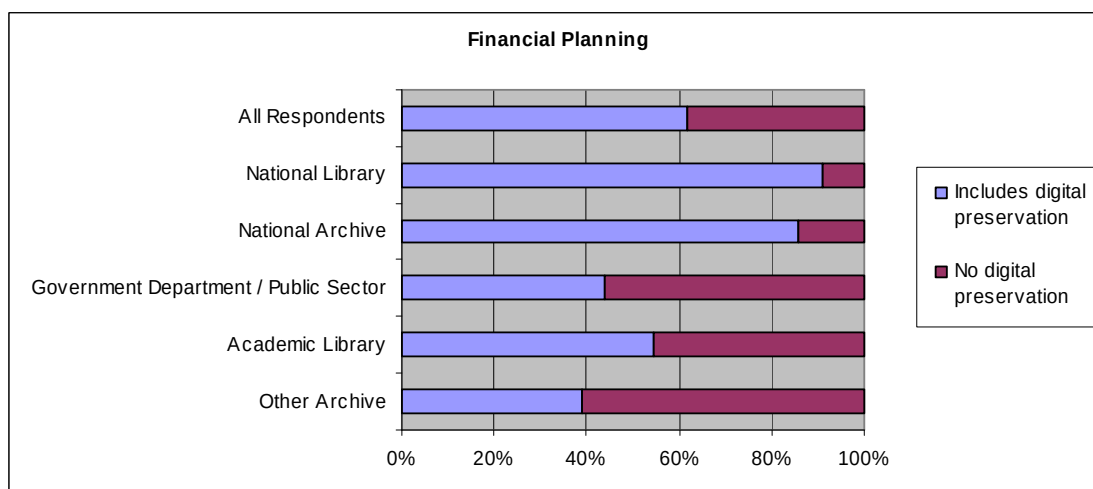
In general it is the national archives and national libraries which are most likely to include digital preservation in their operational, business continuity and financial planning and other archives and government departments and the public sector in general which is least likely to.



**Figure 13: Whether Digital Preservation Features in the Operational Planning of Different Types of Organisation (153 Responses)**



**Figure 14: Whether Digital Preservation Features in the Business Continuity Planning of Different Types of Organisation (154 Responses)**



**Figure 15: Whether Digital Preservation Features in the Financial Planning of Different Types of Organisation (153 Responses)**



## 2.9 Specific Plans for Digital Preservation

Respondents were asked to describe their plans for the long-term management of digital information and were allowed to select more than one answer (see table below for the cross-correlations). Apart from those with no plans, many respondents were at more than one stage in working towards a long-term solution. Even some of those who already have a long-term solution in place are still assessing their needs and requirements and looking to improve or extend their current solution.

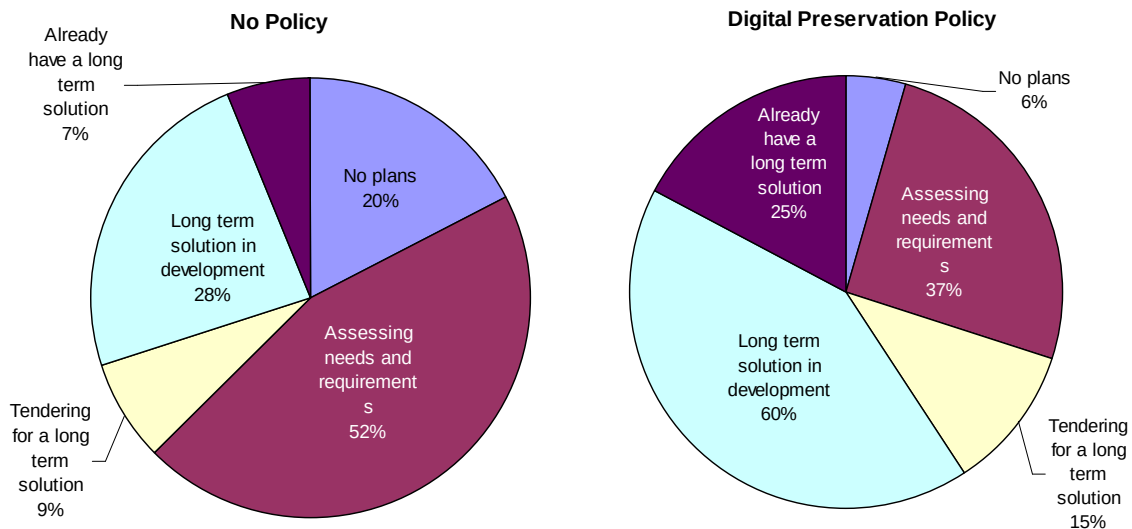
|                                                    | Response Totals | Cross-correlations                                 |                                                    |                                    |                                   |                                   |
|----------------------------------------------------|-----------------|----------------------------------------------------|----------------------------------------------------|------------------------------------|-----------------------------------|-----------------------------------|
|                                                    |                 | Assessing needs and requirements using consultancy | Assessing needs and requirements using a prototype | Tendering for a long term solution | Long term solution in development | Already have a long term solution |
| No plans                                           | 18<br>(13%)     | 0                                                  | 0                                                  | 0                                  | 1                                 | 0                                 |
| Assessing needs and requirements using consultancy | 33<br>(24%)     |                                                    | 6                                                  | 4                                  | 7                                 | 1                                 |
| Assessing needs and requirements using a prototype | 27<br>(20%)     |                                                    |                                                    | 4                                  | 11                                | 3                                 |
| Tendering for a long term solution                 | 16<br>(12%)     |                                                    |                                                    |                                    | 8                                 | 2                                 |
| Long term solution in development                  | 58<br>(43%)     |                                                    |                                                    |                                    |                                   | 5                                 |
| Already have a long term solution                  | 22<br>(16%)     |                                                    |                                                    |                                    |                                   |                                   |

**Table 2: Cross-Correlation of Organisations' Plans for Digital Preservation (135 Responses)**

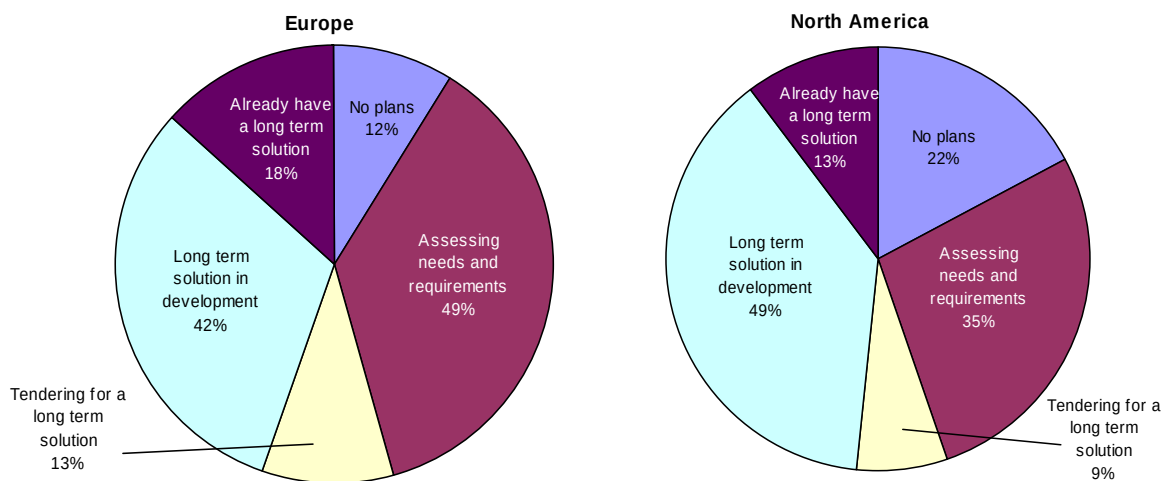
While 85% of those who answered the question are working towards a solution or already have one in place, that leaves 15% who have no plans to deal with digital preservation. For those who are working towards a solution, nearly half have a long-term solution in development, while 45% are assessing their needs and requirements.

Only 6% of organisations with a digital preservation policy have no plans; the rest are assessing needs and requirements (37%), tendering (15%), developing a solution (60%) or already have a solution (25%). Over 3 times as many organisations without a digital preservation policy, as with, have no plans for the long-term management of digital information. Conversely, over 3 times as many organisations with a digital preservation policy, as without, already have a long-term solution. This reinforces the point that having a policy is one of the key steps towards digital preservation.

The differences between North American organisations and European ones are small. The main differences are that more North American organisations (22%) have no plans for digital preservation than European organisations (12%) and more European organisations (49%) are assessing their needs and requirements than North American organisations (35%).



**Figure 16: Organisations' Plans for Digital Preservation for those Organisations With and Without a Digital Preservation Policy (134 Responses)**

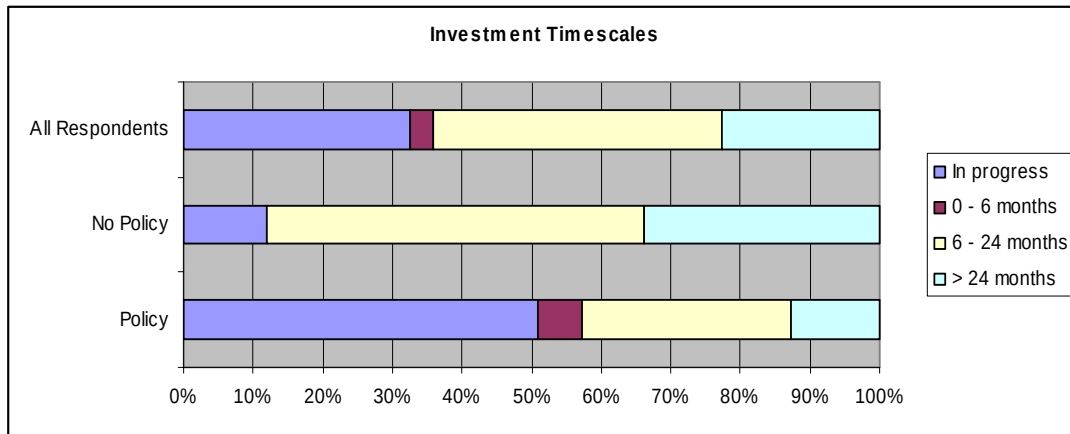


**Figure 17: Organisations' Plans for Digital Preservation for those Organisations in Europe (103 Responses) and North America (23 Responses)**

### 2.9.1 Timescales for those Assessing Needs or Looking for a Solution

A third of organisations are currently investing in a digital preservation solution, two-fifths are looking to make an investment in the next six months to two years, and only just over a fifth have no plans for the next two years.

Separating out the results for institutions with and without digital preservation policies, shows that the presence of a policy frequently indicates that a system is scheduled for the near future. Over half of organisations with a digital preservation policy are currently looking to acquire a solution and over 85% of organisations with a digital preservation policy expect to make an investment within 2 years. Contrast that with the results for organisations without a digital preservation policy where just 12%, or 4 times fewer organisations than those with a digital preservation policy, are currently investing in a solution, and for over a third of organisations without a digital preservation policy it will be more than 2 years before they are looking to make an investment.



**Figure 18: Investment Timescales for Organisations with and without a Digital Preservation Policy (123 Responses)**

## 2.10 How the Solution will be Implemented

Two-thirds of organisations are integrating components into a custom solution, with the remainder evenly split between developing a custom solution and using an off-the-shelf package. In practice, respondents are combining these approaches, as is revealed when the cross-correlations between the three answers (respondents were allowed to select more than one) are investigated (see table below). Half of those developing a custom solution are also integrating components into that solution and over two-fifths of those using an off-the-shelf package are also integrating components into a custom solution. This indicates that organisations are using a pick-and-mix approach to developing their solutions using off-the-shelf components where they are available, and developing their own custom components where they are needed.

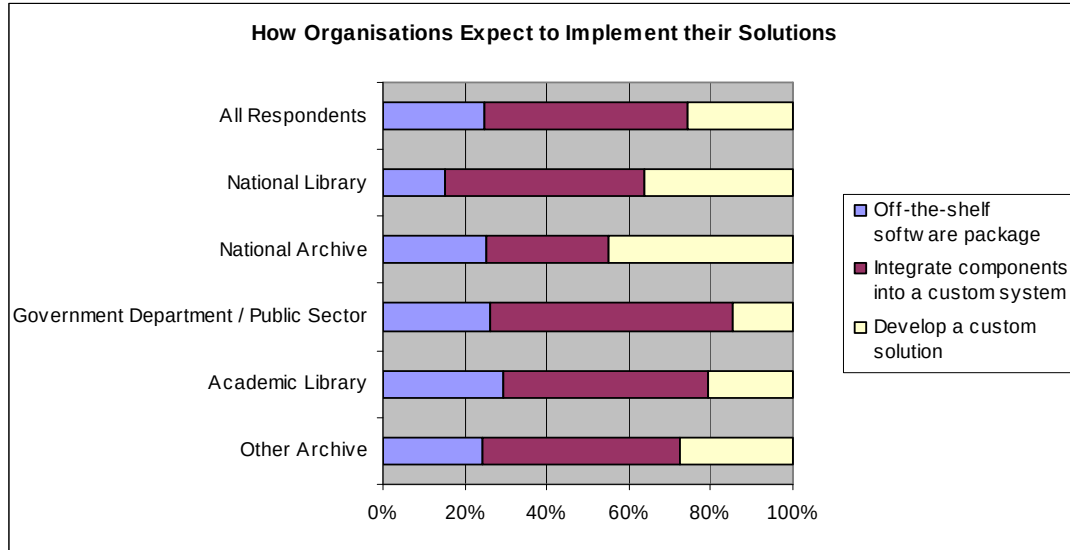
|                                           | Response Totals | Cross-correlations                        |                           |
|-------------------------------------------|-----------------|-------------------------------------------|---------------------------|
|                                           |                 | Integrate components into a custom system | Develop a custom solution |
| Off-the-shelf software package            | 42<br>(32%)     | 17                                        | 7                         |
| Integrate components into a custom system | 85<br>(64%)     |                                           | 22                        |
| Develop a custom solution                 | 44<br>(33%)     |                                           |                           |

**Table 3: Cross-Correlation of How Respondents Expect to Implement Their Digital Preservation Solution (132 Responses)**

Although most respondents expect to either use an off the shelf solution and/or integrate existing components into a custom solution, one third of people expect to implement their own solution. Of those developing a custom solution 17% are doing so from scratch and are not using existing components or off-the-shelf software packages. Given the complexity of digital preservation, developing a custom solution from scratch, rather than building upon existing solutions, is not the simplest approach and may not be the best approach.

National archives comprise the group which is most likely to develop, or have developed, their own custom solution. This may reflect the fact that many national archives have pioneered solutions to digital archiving. At the other end of the spectrum, government departments and the public sector in general are least likely to develop their own custom solution and more likely to integrate components into a custom solution. National libraries are least likely to use an off-the-shelf

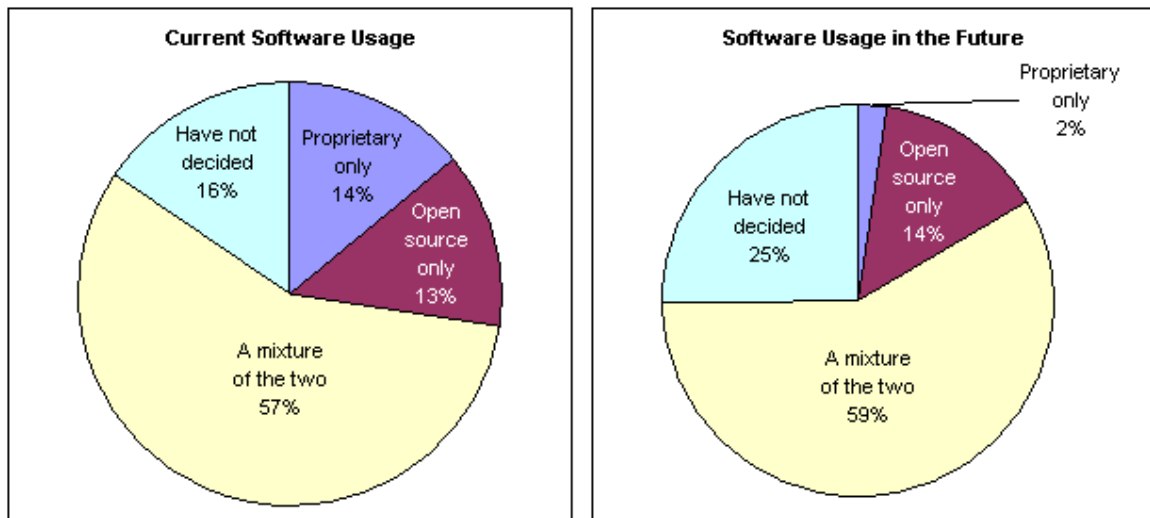
software package, preferring to either develop their own custom solution or to integrate components into a custom solution.



**Figure 19: How Different Organisation Types Expect to Implement Their Digital Preservation Solution (132 Responses)**

Slightly more than half of respondents expect their in-house software team to implement any digital preservation solution, rather than relying on a third party to do so. Interestingly, respondents are less willing to pay for support for Planets software (see section 3.5) if their in-house software team does the implementation than if a third party does so. This suggests that institutions with a tradition of relying on service providers are more willing to pay for another service provider than institutions which have traditionally relied on in-house solutions, or it may reflect the budget constraints that organisations are working under.

As for whether the software for the solution uses open-source or proprietary software: over half of respondents currently use a mixture of the two, with the rest of the responses even split between open-source only, proprietary only and undecided. In the future, a mixture is again the choice for nearly three-fifths of the respondents, but the balance between the other choices has changed. A quarter of respondents have not decided what type of software they will use, whilst the appetite for proprietary software has dwindled to 2%, and that for open-source software remains unchanged at 14%. Although respondents expect the balance to shift in favour of open-source, nearly two-thirds of them think that proprietary software will form all or part of the solution.



**Figure 20: Current and Expected Use of Open Source and Proprietary Software in Organisations' Digital Preservation Solutions (139 Responses)**

### 2.10.1 Implications for Planets

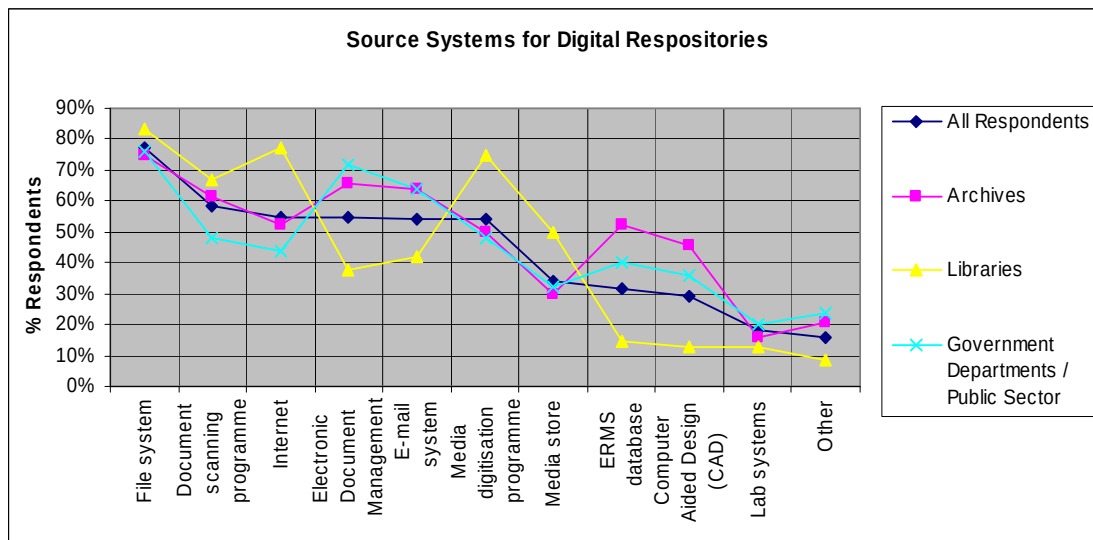
Most respondents expect to use either an off the shelf solution and/or integrate existing components and both these approaches are compatible with Planets usage. Similarly, a majority of respondents expect to use a mixture of open-source and proprietary software, which again is compatible with Planets usage.

### 2.11 Sources of Digital Content for Archiving

Unsurprisingly, the main source of digital information for archiving is file systems, as this is where most documents are stored. Those sources used by more than half of all respondents are:

- File system
- Document scanning programme
- Internet
- Electronic document management system (EDMS)
- E-mail system
- Media digitisation programme

Libraries concentrate on archiving the internet and digital surrogates for analogue resources, whereas archives have more of a focus on the systems used to manage organisations (email, EDMS, ERMS). The relative importance of the different source systems is very similar for both archives and the public sector (including government departments), which reflects the fact that these organisations are often the source of information archived by the archives. The relative unimportance of domain specific software, such as CAD and lab systems, in the survey reflects their more specialised nature and niche use.



**Figure 21: Source Systems that Organisations Currently, or Expect to in the Future, Preserve Digital Material from (141 Responses)**

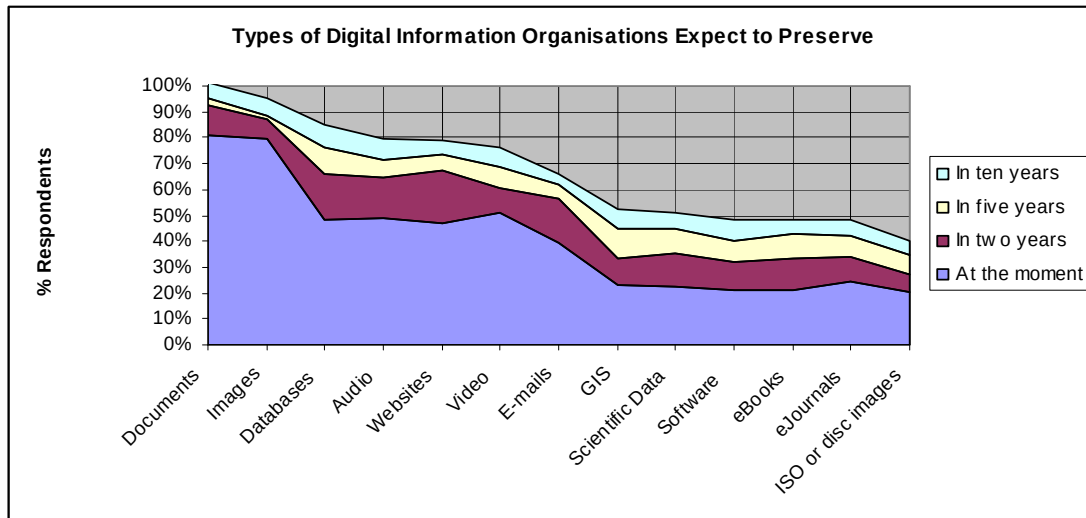
Although the listed source systems are common ones, over 15% of respondents need to take information from other systems. Amongst those mentioned are:

- Relational databases
- Geographical Information Systems (GIS)
- External media (CDs, tapes)
- Learning and teaching materials
- E-books and e-journals
- Source code

This reflects the wide range of digital information that needs to be preserved for the future. Several archives and libraries mentioned that they could not predict what would be donated or transferred to them.

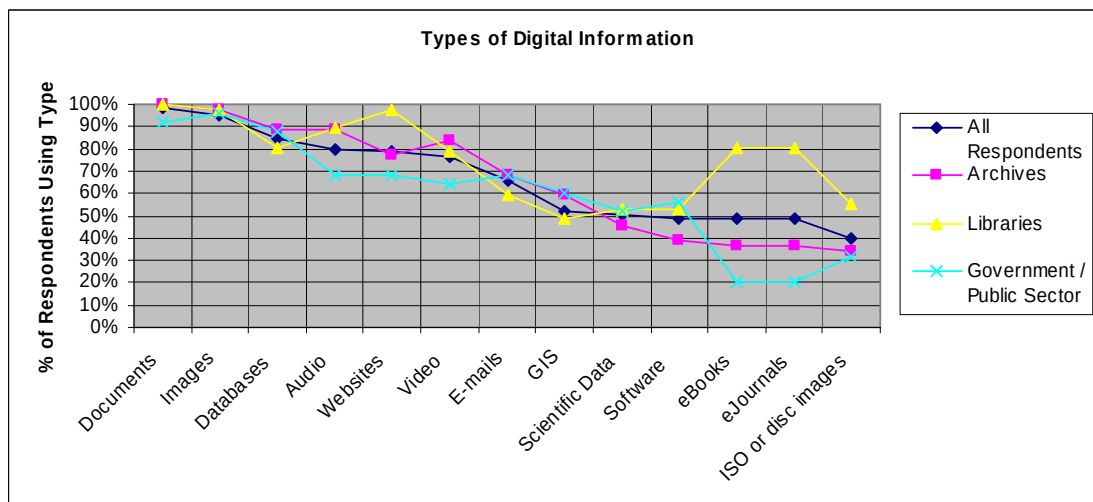
## 2.12 Digital Information Types that Need to be Managed in the Long-Term

In this question respondents were asked which types of digital information they are currently managing for the long-term and which types they think they will need to manage for the long-term in the future.



**Figure 22: Types of Digital Information that Organisations Currently, or Expect to in the Future, Preserve (138 Responses)**

Figure 22 shows the percentage of respondents who expect to need to preserve each type of digital information. Unsurprisingly, documents need to be managed by nearly all respondents, closely followed by images. More surprisingly, although only half of respondents currently need to manage databases for the long-term, 85% of respondents think they will need to manage them in ten years' time. This is a challenge as relatively little research has been carried out on preserving databases, and they are not simple things to preserve. Websites and e-mails, which are often cited as common information types needing to be preserved, are only the 5<sup>th</sup> and 7<sup>th</sup> most widespread form of digital information requiring preservation, although two thirds of respondents reckon that in 10 years' time they will need to manage these forms of digital information.



**Figure 23: Types of Digital Information that Different Organisation Types Expect to Need to Preserve Within 10 Years (138 Responses)**

When the data is broken down according to the type of organisation the respondent works for, interesting variations in the types of information being managed emerge. Libraries are more likely to preserve websites than anyone else, and for libraries nearly all respondents need to manage them (the same numbers as for documents and images). Unsurprisingly, libraries are far more likely to need to preserve eBooks and eJournals than anyone else. In general, libraries are more likely to need to preserve a wide-range of digital information types, followed by archives, then the public sector.

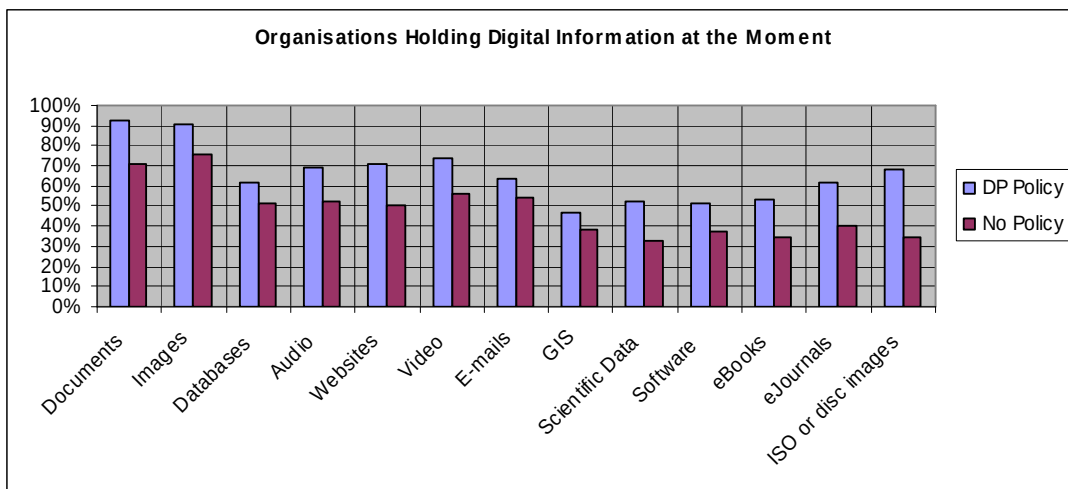
Other types of digital information that respondents need to preserve and which were not explicitly listed in the question are:

- DCPs and MAPs (digital cinema)
- Spreadsheets
- Radiographs
- CAD data
- Engineering Models
- Virtual Reality
- Games (a subset of software)
- Medical samples supporting data

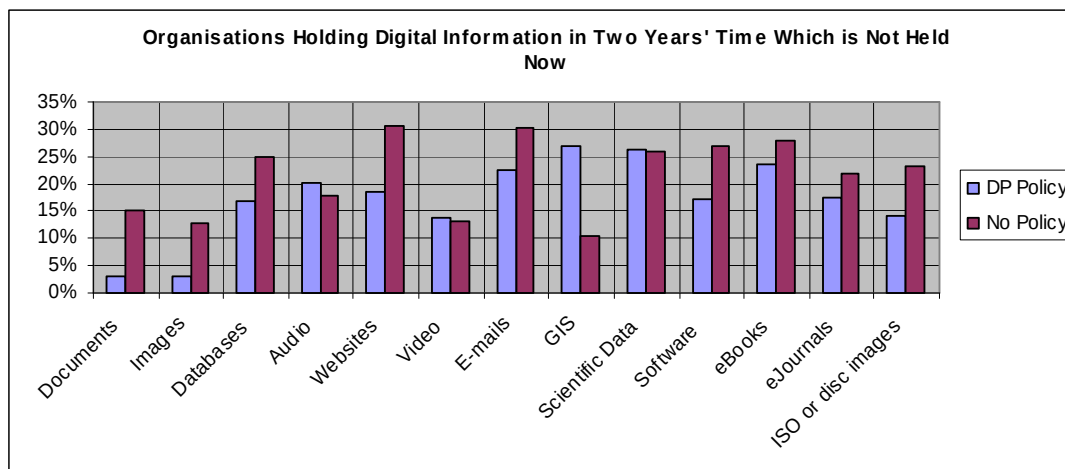
Most of these other types of digital information have dynamic content, which is a much harder challenge to preserve than static content.

#### 2.12.1 When Digital Information Needs to be Stored for Organisations with and without a Digital Preservation Policy

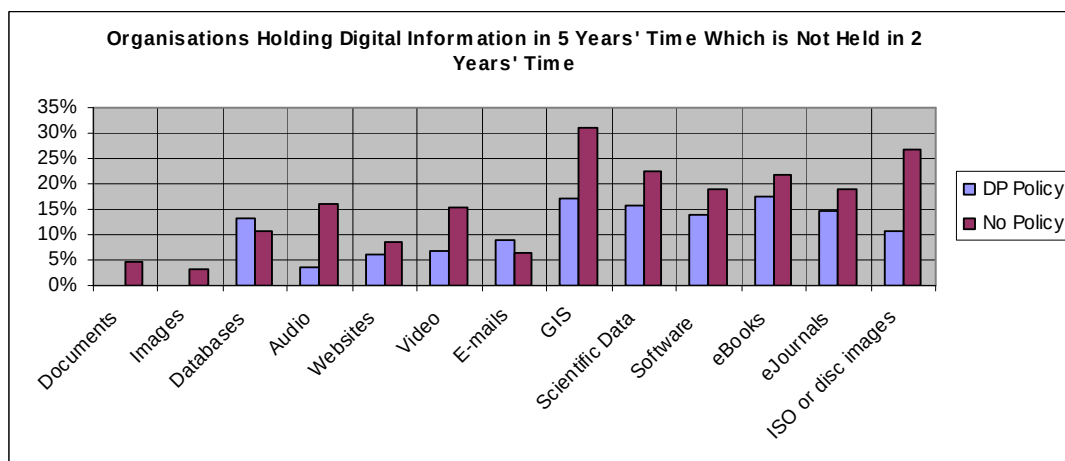
Comparing organisations which have a digital preservation policy to those without over the ten-year time period reveals that a higher percentage of organisations which currently have a digital preservation policy in place need to store each of the different types of digital information, than those organisations without a digital preservation policy. However, over time this initial difference is eroded as more organisations without a digital preservation policy foresee a need to store the different types of digital information in the future.



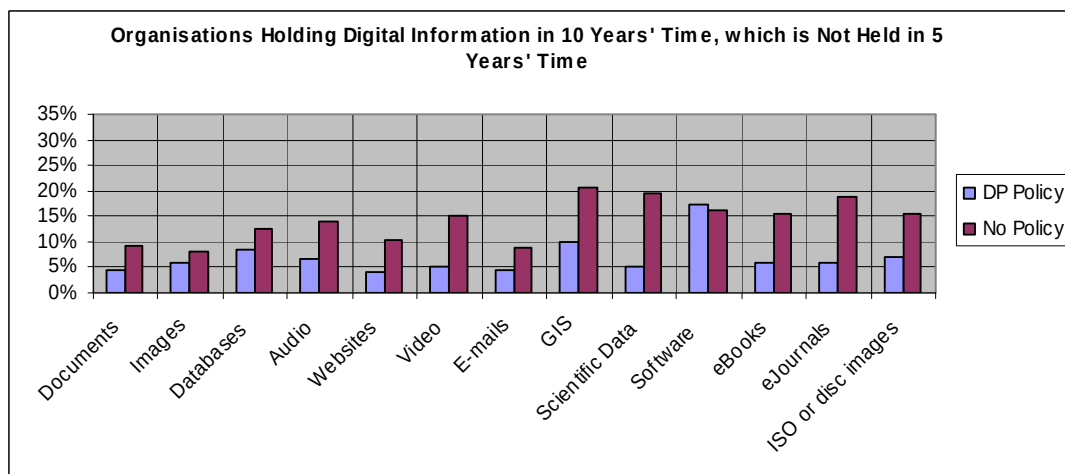
**Figure 24: Types of Digital Information that Organisations with and without a Digital Preservation Policy Currently Need to Preserve (136 Responses)**



**Figure 25: Types of Digital Information that Organisations with and without a Digital Preservation Policy will Need to Preserve in Two Years' Time that they do not Currently Need to Preserve (136 Responses)**



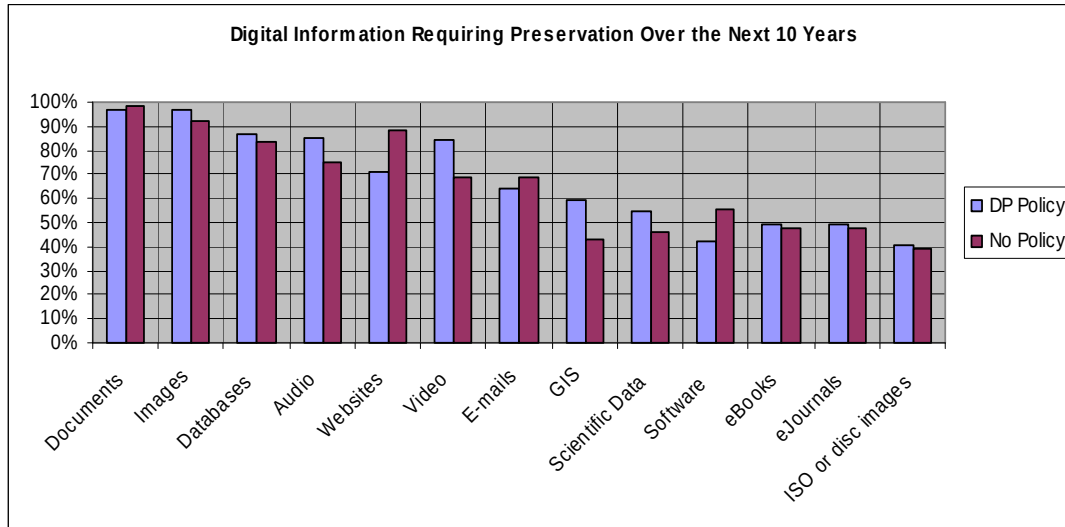
**Figure 26: Types of Digital Information that Organisations with and without a Digital Preservation Policy will Need to Preserve in Five Years' Time that they will not Need to Preserve in Two Years' Time (136 Responses)**



**Figure 27: Types of Digital Information that Organisations with and without a Digital Preservation Policy will Need to Preserve in Ten Years' Time that they will not Need to Preserve in Five Years' Time (136 Responses)**



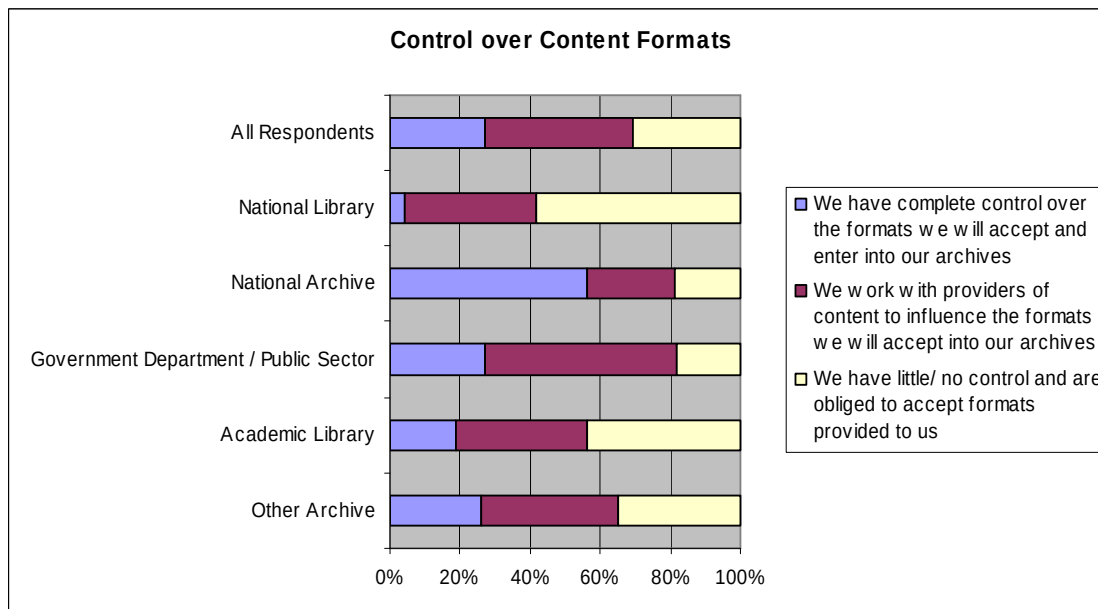
In ten years' time roughly equal numbers of organisations with and without digital preservation policies need to store each of the different types of digital information. There are some differences, though. In particular, significantly more organisations without a digital preservation policy, than those with one, think they will need to manage websites in the long-term. Whereas significantly more organisations with a digital preservation policy, than those without one, think they will need to manage videos and GIS data in the long-term.



**Figure 28: Types of Digital Information that over the next Ten Years, Organisations with and without a Digital Preservation Policy Will Need to Preserve (136 Responses)**

### 2.13 Control over the Digital Formats Archived

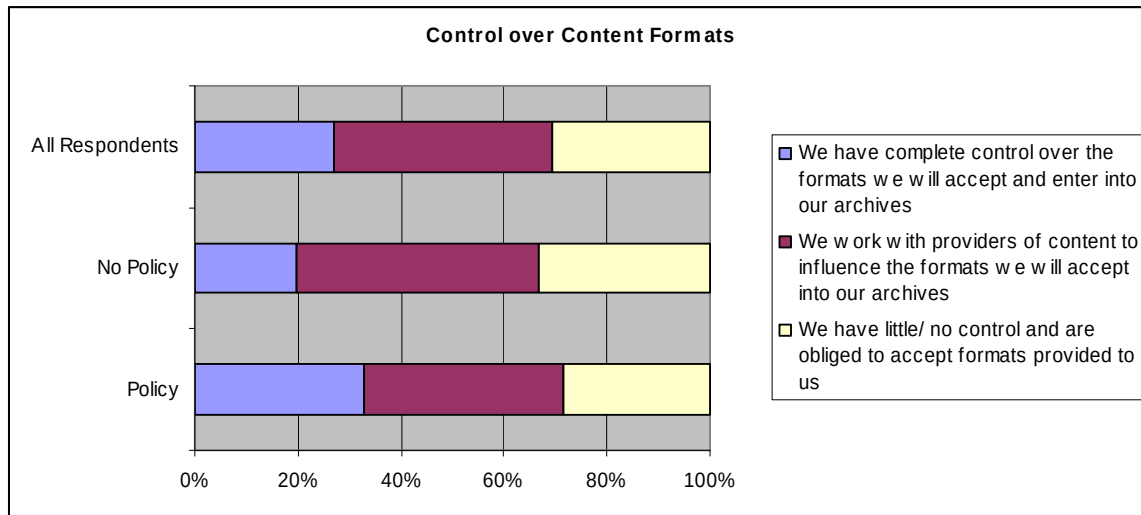
Less than a third of organisations think that they have complete control over the formats that they will accept and enter into their archives, while another third think they have little or no control over content format and the remainder hope to use their influence with content providers to limit the content formats. However, there are variations between the different types of organisations.



**Figure 29: Level of Control that Different Types of Organisations have over the Formats of the Content in their Digital Archives (137 Responses)**

In particular, national libraries are under no illusions that they can control content format but hope to have some influence. This is probably a result of experiences with legal/voluntary deposit and electronic publications. However, many national archives believe that they have complete control content over the formats that they will accept, which is surprising. Indeed, ten times more national archives than national libraries think that they can completely control the formats of the content they receive.

Some other institutions, particularly commercial organisations, also believe that they have complete control but this may be because of their particular circumstances. Having control over what software is used within an organisation, which is often the case in businesses, inevitably controls the digital formats information is stored in.



**Figure 30: Level of Control that Organisations with and without a Digital Preservation Policy have over the Formats of the Content in their Digital Archives (137 Responses)**

It is noticeable that only a fifth of organisations without a digital preservation policy think that they have complete control, compared to a third of organisations with a digital preservation policy.

## 2.14 Important Capabilities for Digital Archives

Respondents were asked to rate how important they thought various capabilities of a long-term digital information management system were. The list of capabilities, ordered by the average rating given by respondents, is shown in Table 4.

A rating of 3.0 means a capability is deemed important; a higher value means that respondents think the capability is more important. The actual ratings labels displayed to respondents when filling in the survey were:

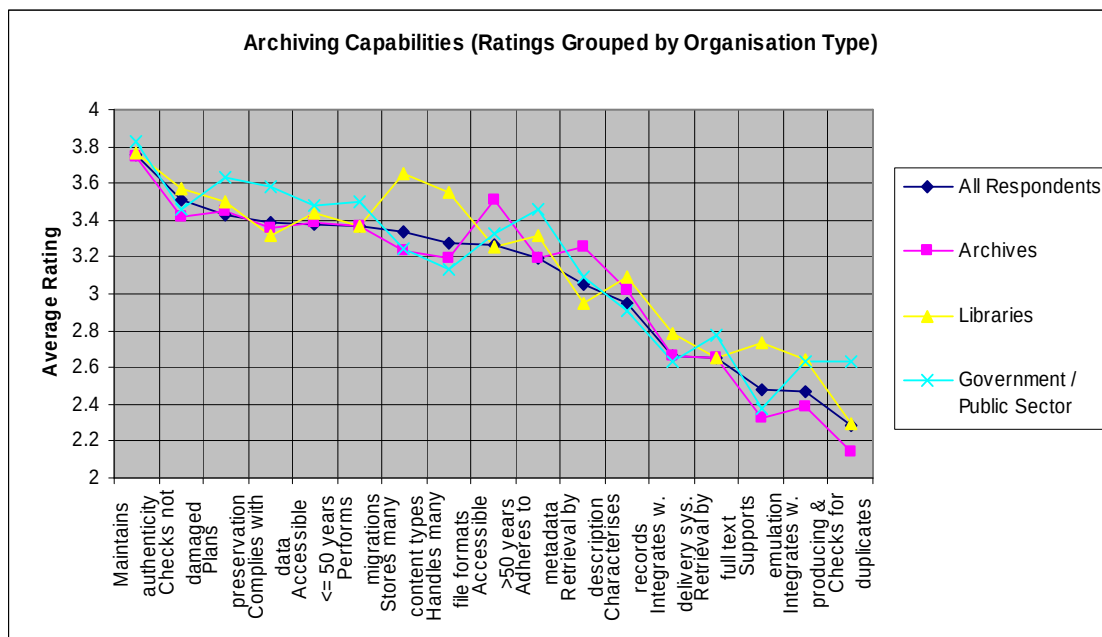
- Not applicable (1)
- Least important (2)
- Critical (5)

The last five capabilities (with a white background in Table 4) in the list were rated less important. In particular, respondents do not deem either emulation or integrating a digital archive with content producing, holding and delivery systems important. It is interesting to note that respondents are most concerned that the digital records they deem worthy of keeping are maintained in such a way as not to damage or corrupt them and demonstrably so. Standards are an important part of this, as is being able to plan how to deal with technical obsolescence. Migration is currently seen as the way to deal with technical obsolescence, rather than emulation, although this probably reflects the relative maturity of the two technologies. Preserving digital records for up to 50 years and over 50 years are both seen as important. However, archives take the longer term view, ranking ensuring records are accessible for more than 50 years as their second most important capability, after maintaining the authenticity, reliability and integrity of records.

| Capability                                                                 | Average Rating |
|----------------------------------------------------------------------------|----------------|
| 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 4: The Important Capabilities for a Digital Archive to have, as Rated by the Survey Respondents (135 Responses)**

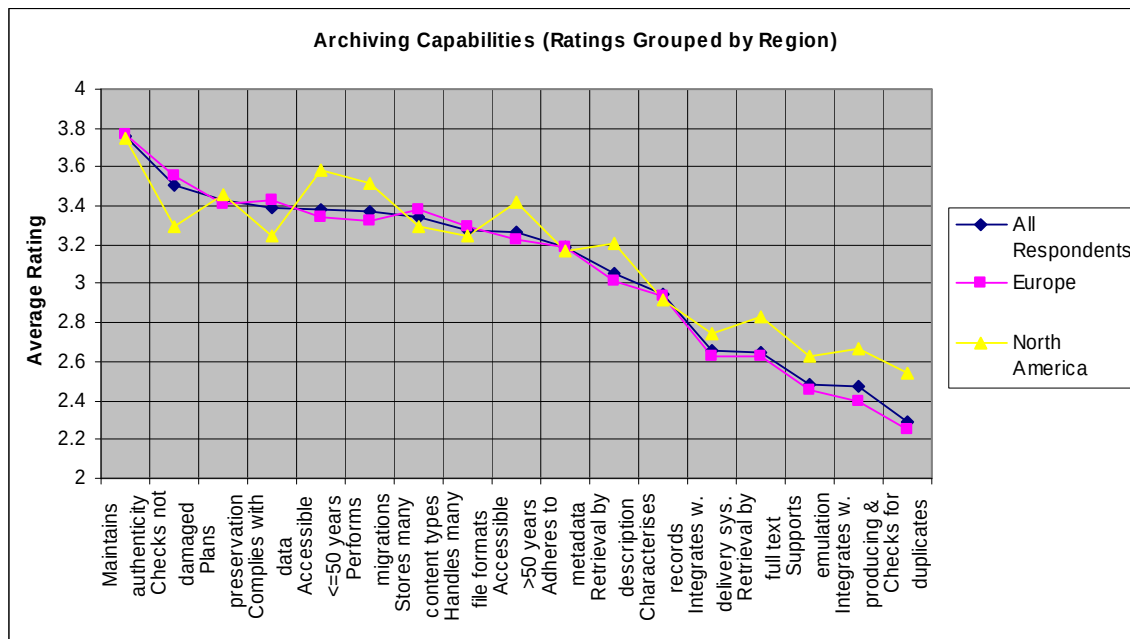
Although there is broad agreement as to the relative importance of the different capabilities, there is some disagreement between respondents. This is illustrated in the following graphs, which break down the rated importance of the different capabilities by organisation type and region. The capabilities are ordered, left to right, by their overall rating, with the most important first.



**Figure 31: Importance of Different Capabilities of a Digital Archive to Respondents from Different Types of Organisations (135 Responses)**

What is very noticeable is that the libraries rank the capabilities 'is able to store many different types of content' and 'handles a wide variety of file formats' much more highly (respectively 2<sup>nd</sup> not 7<sup>th</sup>, and 4<sup>th</sup> not 8<sup>th</sup>). This reflects the wider range of content that libraries are expected to preserve (as shown in section 2.12). It is also illustrated by the fact that libraries rate the ability to support emulation more highly than either archives or the public sector. However, the latter may have more to do with the fact that libraries need to archive more dynamic content (migration does not preserve behaviour, whereas emulation does) than archives or the public sector.

It is also noticeable that archives rank 'retrieves content by description' much more highly than the other organisation types. Presumably this is because of their experience with the ways their users search for material. The public sector (including government departments) ranks complying with standards (whether for data, digital information management or metadata) much more highly than do archives and libraries. This focus on standards reflects a culture of public accountability and regulation in the public sector.



**Figure 32: Importance of Different Capabilities of a Digital Archive to Respondents from Different Regions (135 Responses)**

There is very little difference between the rankings assigned to capabilities by respondents in Europe and North America and given the number of responses (107 from Europe and 24 from North America) these small differences are not statistically significant. However, they do give an indication of relative trends. The North Americans rank ensuring records are accessible for up to 50 years and the ability to perform migrations as the second and third most important capabilities of a digital archive, after maintaining the authenticity of records, but before checking that records have not been damaged, which they rate as much less important. Similarly, they place less importance on complying with established data or digital information management standards. Is this because North Americans place more emphasis on what a product can actually do than on whether it complies with a set of relevant standards, or do they see the current established standards as not so relevant for digital preservation? Although the North Americans agree that the last 5 capabilities are unimportant, they do think they are less unimportant than the Europeans do, the difference being most marked in checking for duplicate items.

#### 2.14.1 Comments Made by Respondents

The following comments on other capabilities seen as important for long-term digital information management systems were made by respondents, and are included verbatim.

- "We'd like the ability to repair files, not just verify or check for damage. This is fully practical for uncompressed audiovisual files, and will be pursued in EC project PrestoPRIME"

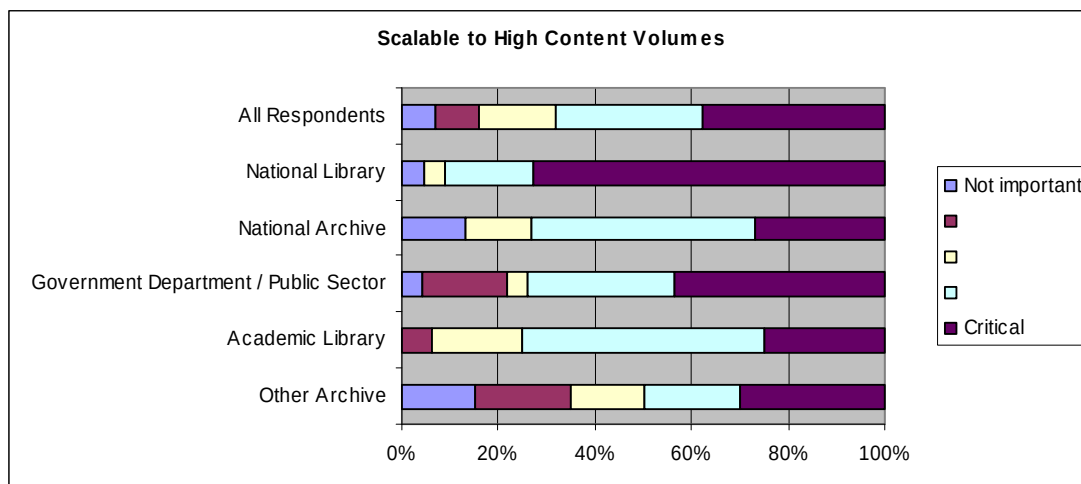
- “The main thing is to be flexible. Hard to tell whether migration or emulation will be the thing - see what FutureArchi is doing.”
- “legal hold, disposal”
- “Version control”
- “In many cases the actual choice of what is critical as opposed to what is not is influenced by other choices that you make, so in many ways this is only one view of an expectation of a solution.”
- “1. Enables content to be transferred between different DIMS systems without a loss or corruption of data occurring.  
2. Enables content to be transferred between different DIMS systems and enables assigning of diverse access/distribution rights within the different systems”  
[DIMS is Digital Information Management System]
- “Certified trustworthiness”
- “We can not see how technology will change over the next 10 years so any system will need to be able to cope with ongoing changes in the medium term.”
- “We have no aspiration to store many different types of content but we have a high aspiration to handle/migrate a wide variety of file formats without loss of information”
- “Performs migrations to deal with technical obsolescence: not necessary in the system, but system has to provide applicable content and must keep versions of objects and its information.  
Ensures records are accessible for up to 50 years: No system can ensure this.  
Responsibility of ltp-managers over the years.”  
[ltp is long-term preservation]
- “Costing functionality i.e. estimate of how much something will cost to preserve over x years.”
- “compatible with OAIS standard; has to support non-poprietary formats”
- “we don't know”
- “records provenance data (important)  
protects data in compliance with contracts with data producers  
persistent identifiers  
OAI harvesting  
open for search engines like google  
federated queries / interoperability”

---

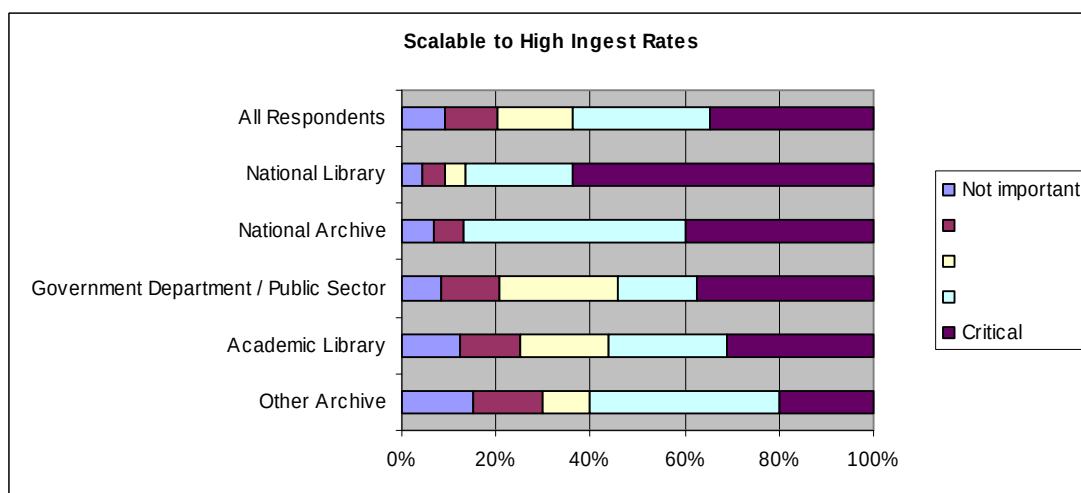
## 2.15 The Importance of Scalability for Digital Archives

It is clear from the results that scalability of total content (Petabytes of data) and high ingest rates (millions of objects per year) are regarded as extremely important with scalability of access rates (hundreds of objects per second) significantly less important.

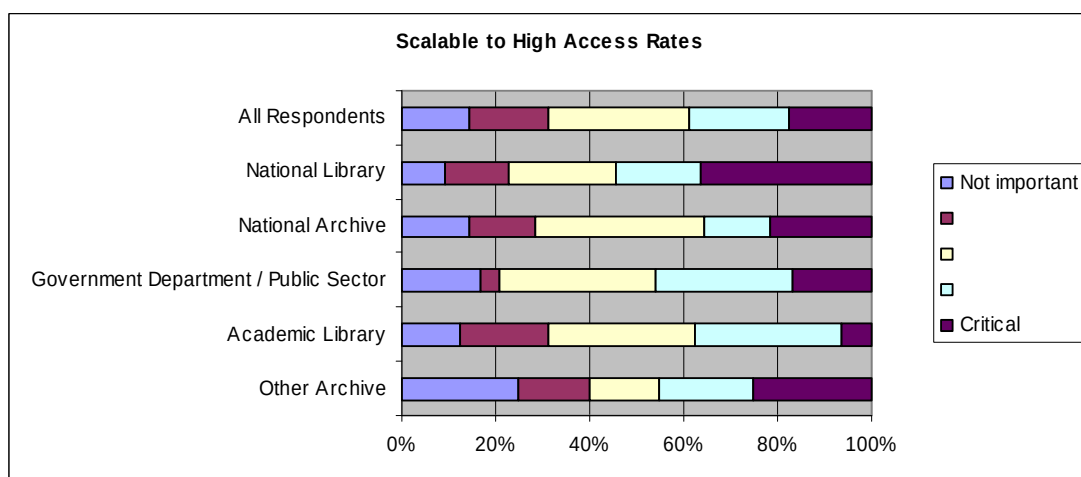
However, the importance of scalability varies significantly with the type of organisation (see graphs below). National libraries are three times more likely than national archives to regard scalability of content as critical.



**Figure 33: Importance of the Scalability of a Digital Archive to High Content Volumes (Petabytes) to Different Types of Organisation (133 Responses)**



**Figure 34: Importance of the Scalability of a Digital Archive to High Ingest Rates (Millions of Objects per Year) to Different Types of Organisation (133 Responses)**



**Figure 35: Importance of the Scalability of a Digital Archive to High Access Rates (Hundreds of Objects per Second) to Different Types of Organisation (133 Responses)**

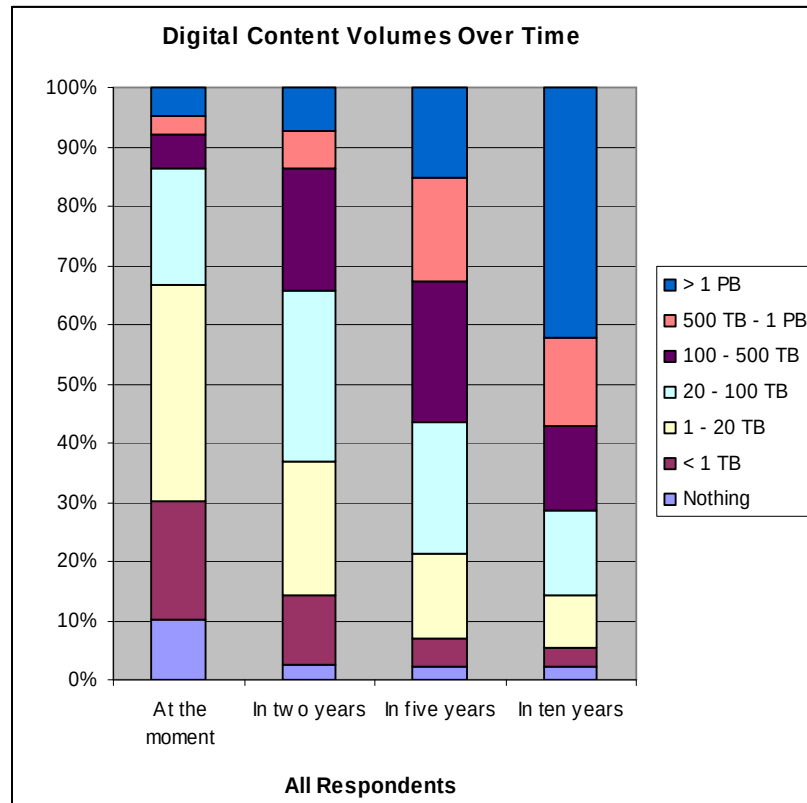
## 2.16 Quantities of Digital Content Being Stored

Unsurprisingly the amount of data that organisations expect to store increases over time. The current storage needs of most organizations are quite modest (over 80% of respondents hold less than 100TB, with a median of less than 20TB) but large increases in volume are predicted in the next decade (by which time 70% of respondents expect to hold more than 100TB, with a median of over 500TB, and many planning to hold multiple petabytes).

|                   | Data Volumes at the Moment | Data Volumes in Ten Years |
|-------------------|----------------------------|---------------------------|
| Mean <sup>1</sup> | 150 TB                     | 1.0 PB                    |
| Median            | 1-20 TB                    | 500 TB – 1 PB             |
| Mode              | 1-20 TB                    | > 1 PB                    |

**Table 5: Average Volumes of Digital Content Organisations Store Now and Intend to Store in Ten Years' Time (129 Responses)**

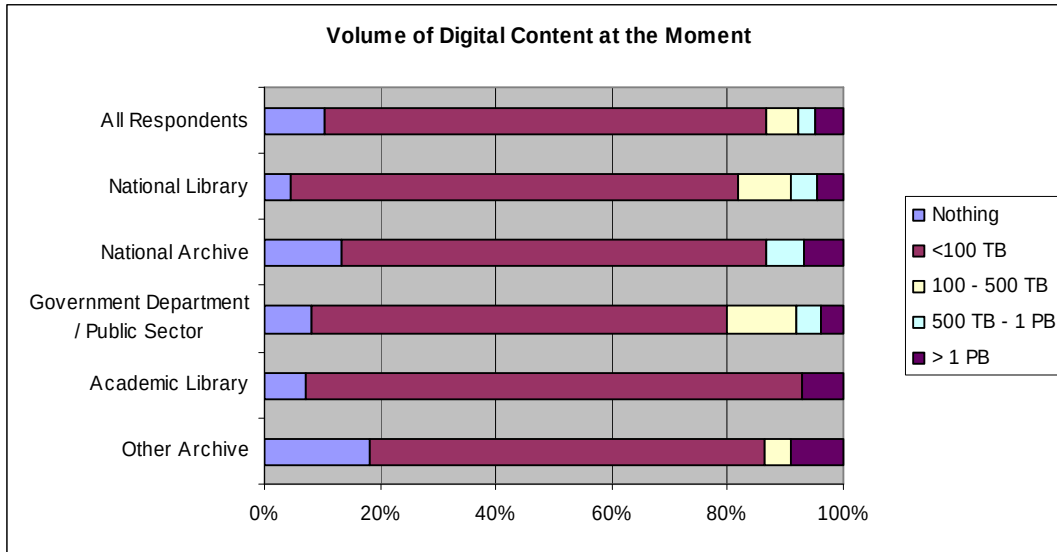
It is clear that while current average data volumes are low, typically under 20 TB, in ten years' time the average organisation expects to hold around 1 PB of digital information: a 50-fold increase.



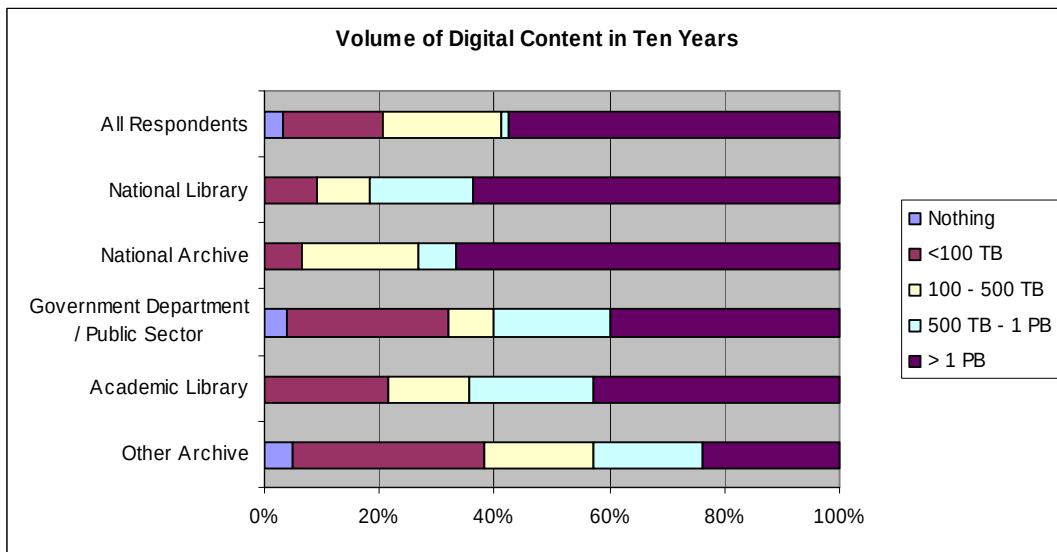
**Figure 36: Growth over the next Ten Years in Volumes of Digital Content that Organisations Intend to Store (129 Responses)**

When the results are broken down by the type of organisation, it can be seen that national archives and national libraries hold the most data, with other archives holding the least. 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 (see section 2.15).

<sup>1</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, and rounded to 2 significant figures.



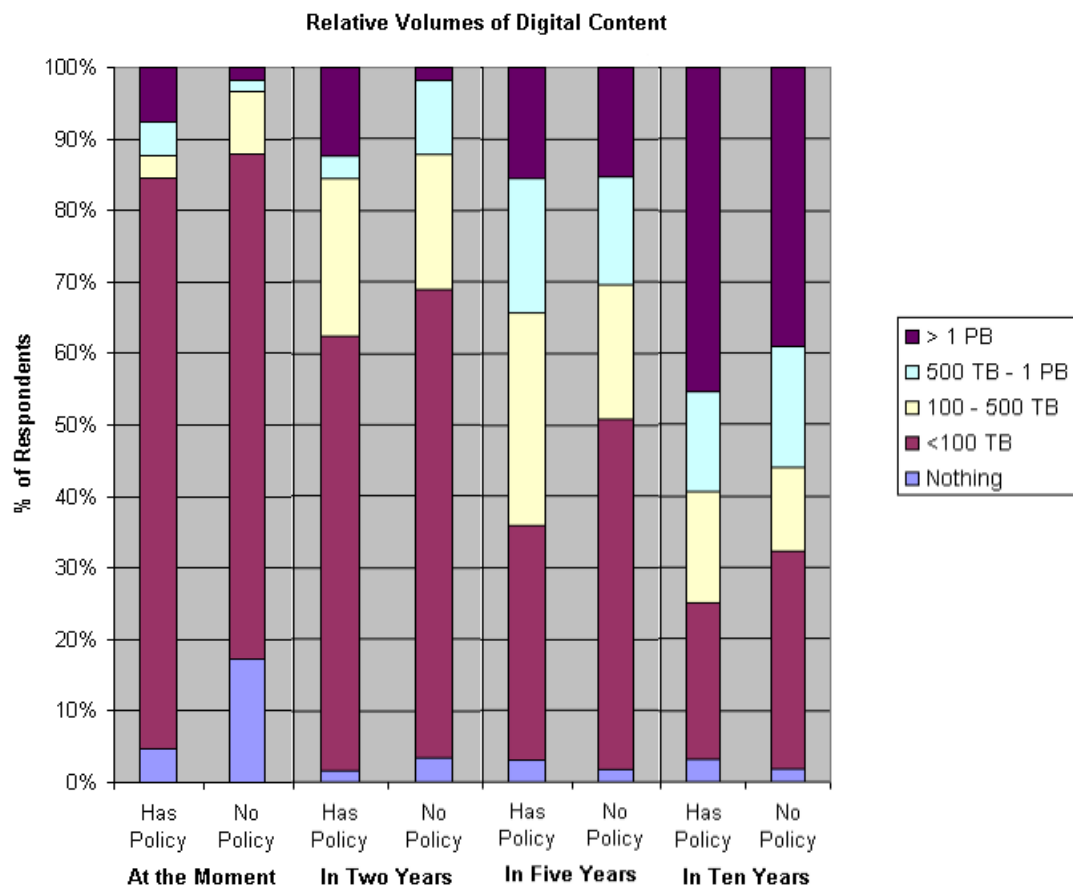
**Figure 37: Volumes of Digital Content that Different Types of Organisation Currently Store (126 Responses)**



**Figure 38: Volumes of Digital Content that Different Types of Organisation Intend to Store in Ten Years' Time (126 Responses)**

More interestingly, those organisations which do not currently have a digital preservation policy currently store and intend to store less data than those organisations with such a policy. Whether it is the case that organisations without much data do not prioritise developing a digital preservation policy, or whether it is the case that it is only when an organisation develops such a policy that it discovers how much data it needs to store is unclear.





**Figure 39: Growth over the next Ten Years in Volumes of Digital Content that Organisations with and without a Digital Preservation Policy Intend to Store (126 Responses)**

### 3. Analysis of Planets Specific Questions

This section focuses on analysing the answers to those questions in the survey that were specific to Planets, rather than to digital preservation in general.

For the questions on rating Planets' capabilities and services, the rating values used are equivalent to:

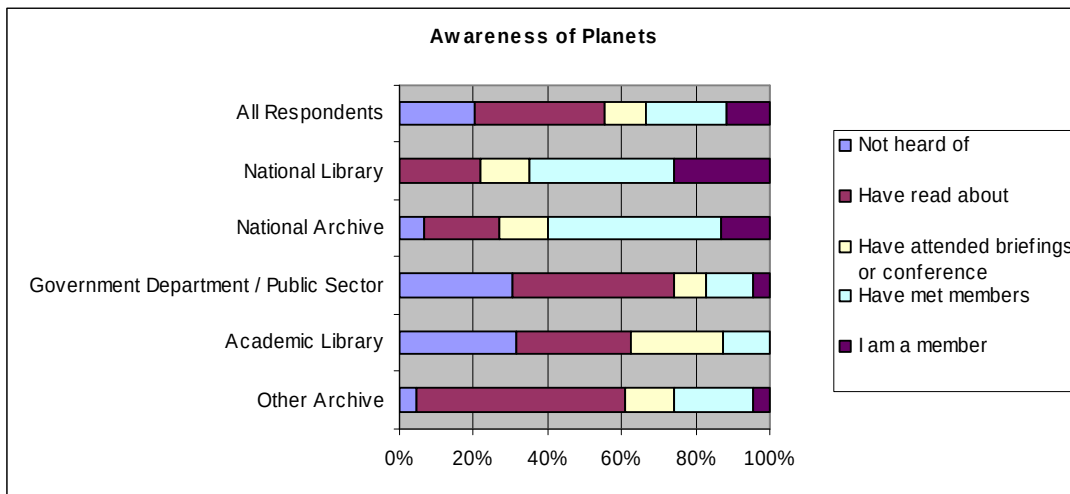
- Not interested (1)
- Probably not interested (2)
- Possible interest (3)
- Interested (4)
- Very interested (5)

In analysing Planets' capabilities, services and how they could be funded the sector breakdown includes the results for commercial organisations and for suppliers and vendors. These results have little statistical significance due to the very small sample sizes (4 for commercial organisations and 9 for suppliers and vendors) and are only indicative of trends.

#### 3.1 Awareness of Planets

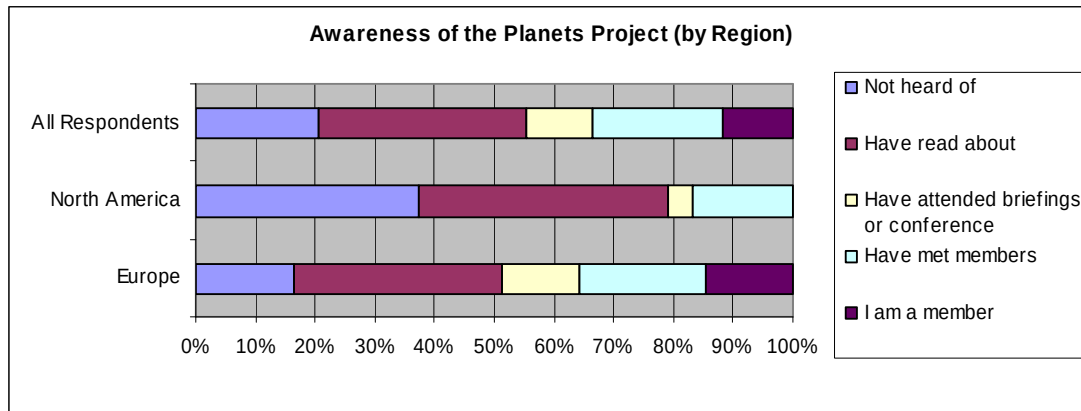
Respondents were asked whether they were aware of Planets, and if they were aware how they had heard about the project. Unsurprisingly, 80% said they were aware of Planets and around 45% of respondents had had direct contact with Planets through briefings, conference and meetings with members, or are members.

However, when the data is broken down by respondents' organisation type it becomes clearer that Planets is well known in national archives and libraries but there is still a lot of work to do in academia, government departments and the public sector in general.



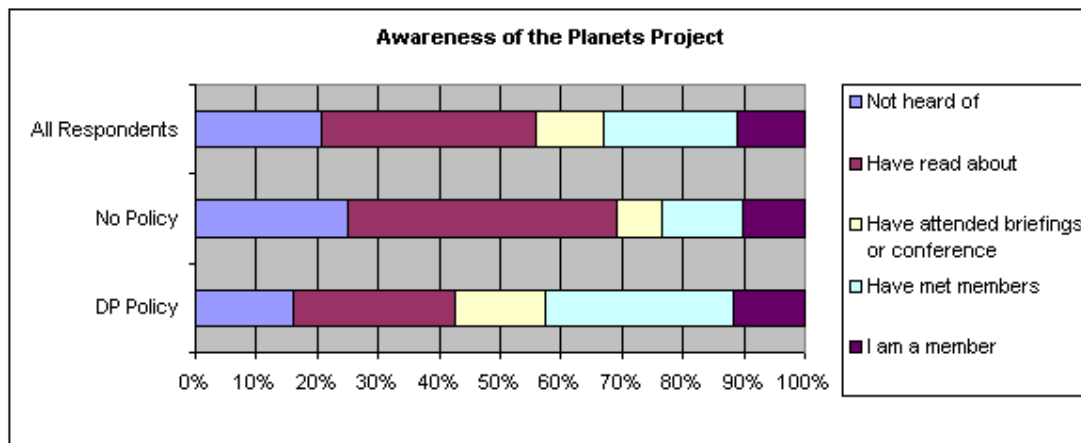
**Figure 40: Awareness of Planets within Different Types of Organisation (137 Responses)**

Planets is an European Union project so it not surprising that European respondents are more aware of Planets than those in the rest of the world, although even in North America only a third of those respondents had not heard of Planets.



**Figure 41: Awareness of Planets in Organisations in Different Regions (137 Responses)**

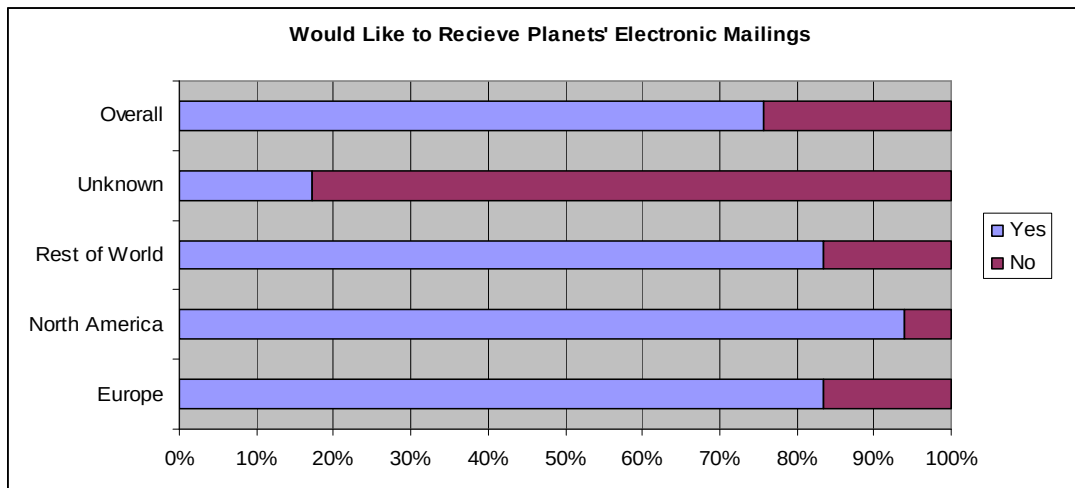
Over half of respondents whose organisation has a digital preservation policy have been actively involved (attending briefings or conferences and meeting members or being a member) with Planets. For those respondents whose organisation does not have a digital preservation this figure drops to less than a third although nearly half have read about Planets. So, assuming involvement in Planets is a reasonable indicator of involvement in the general digital preservation community, having a digital preservation policy appears to lead to greater participation in this community.



**Figure 42: Awareness of Planets in Organisations with and without a Digital Preservation Policy (137 Responses)**

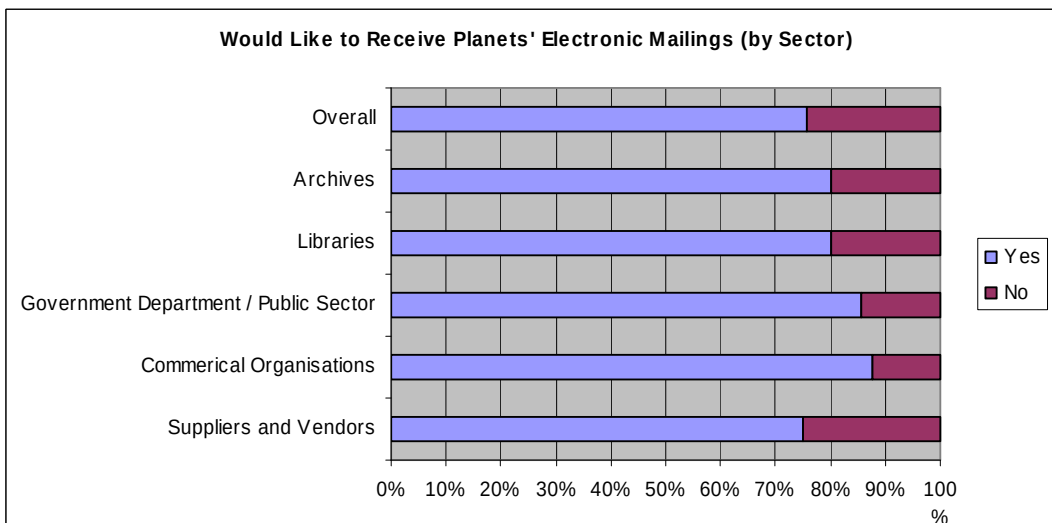
### 3.2 Receiving Electronic Updates About Planets

Over three-quarters (76%) of those people who filled in the survey indicated that they would like to receive electronic updates about Planets, which indicates a high degree of interest in Planets. It is the North American respondents who are most interested in receiving electronic updates about Planets. Presumably this is because it is harder for them to attend events in Europe, and so receiving electronic updates is a good way for them to stay informed about the Planets project's progress. Unsurprisingly, it is the people who are not prepared to identify themselves who do not wish to receive electronic mail from Planets.



**Figure 43: Interest in Receiving Planets' Electronic Mailings from Respondents in Different Regions (196 Responses)**

There is very little difference between the different sectors in interest in receiving Planets' electronic updates; the suppliers and vendors are marginally less interested (only 75% wish to receive Planets' e-mailings) and the commercial organisations, government departments and public sector in general are marginally more interested (over 85% would like to receive Planets' e-mailings).



**Figure 44: Interest in Receiving Planets' Electronic Mailings from Respondents in Different Types of Organisation (196 Responses)**

### 3.2.1 Implications for Planets

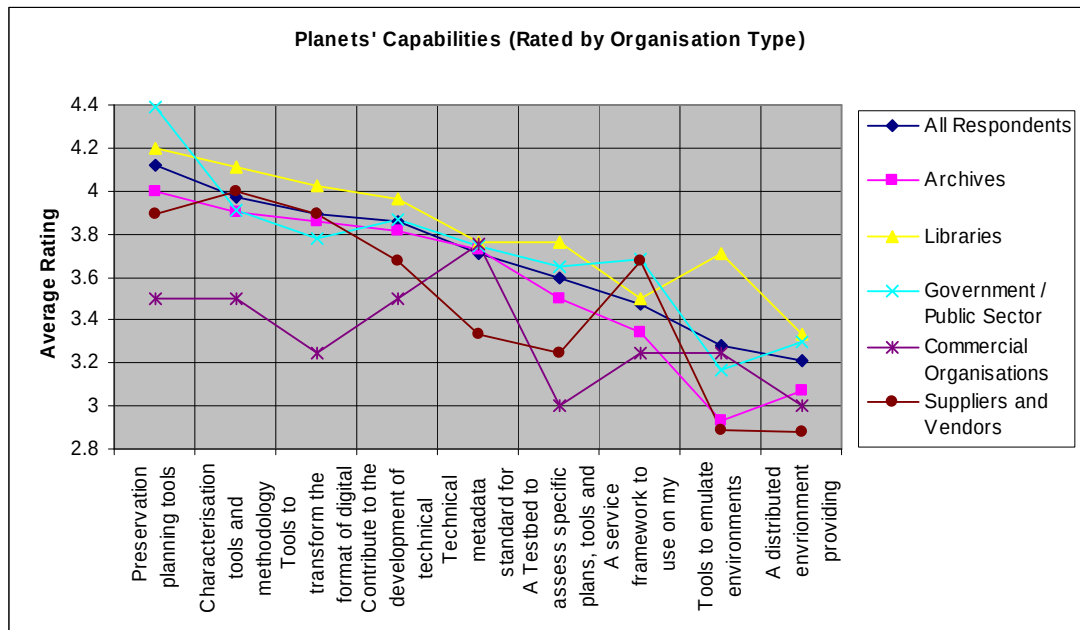
There is a high degree of interest in Planets, which can be capitalised upon with regular electronic updates about Planets' work.

## 3.3 Rating Planets Capabilities

Respondents are generally interested in the capabilities that Planets could offer, with all capabilities generating at least a possible interest and some generating a definite interest. Many respondents hope to see Planets providing tools, especially for preservation planning, characterisation and format transformation, and also contributing to technical standards.

Generally the archives, libraries and the public sector (including government departments) have similar views as to which of the various Planets capabilities they are most interested in. Although

libraries are more enthusiastic about all the capabilities than the public sector and archives are the least enthusiastic. However, the three sectors do differ in their opinions of certain capabilities.



**Figure 45: Interest in Planets' Capabilities from Respondents in Different Types of Organisation (133 Responses)**

It is particularly noticeable that libraries are much more enthusiastic about emulation tools than the other organisations. This could be because they are aware that they need to accept a wide range of digital formats and that it would be easier to manage these in emulated environments than migrating each format individually. Or, more likely, because of the larger amount of dynamic content that libraries need to store (see section 2.12), which is much better preserved using emulation since migration does not preserve behaviour.

The public sector rates preservation planning tools and a service framework for a digital archive much more highly than anyone else. Possibly this is because the public sector is only just getting started on planning for digital preservation and so has more need of these capabilities.

It is obvious that commercial organisations have a very different view of which capabilities are important to them. They rate technical metadata standards most highly, but are generally only mildly interested in any of the Planets capabilities. This may be because they have not heard of Planets and so are wary of relying on it.

It is interesting to look at the Planets capabilities from the point of view of suppliers and vendors who may want to make use of them as part of their own offerings. They share the general interest in tools (for preservation planning, characterisation and format transformation) and Planets' contribution to the development of technical standards, but they are also particularly interested in a service framework for digital archives. Presumably their interest in a service framework stems from its potential to provide them with a standard, common interface to build their own offerings around.

Given that Planets is a European project, and the lower awareness of Planets in North America, it is not so surprising that the North Americans are less interested than the Europeans in the capabilities Planets has to offer. In particular the North Americans are less interested in characterisation tools and methodology and a service framework for digital archives, possibly because they believe that such tools already exist, for example JHOVE, and further tools and frameworks are not a priority.

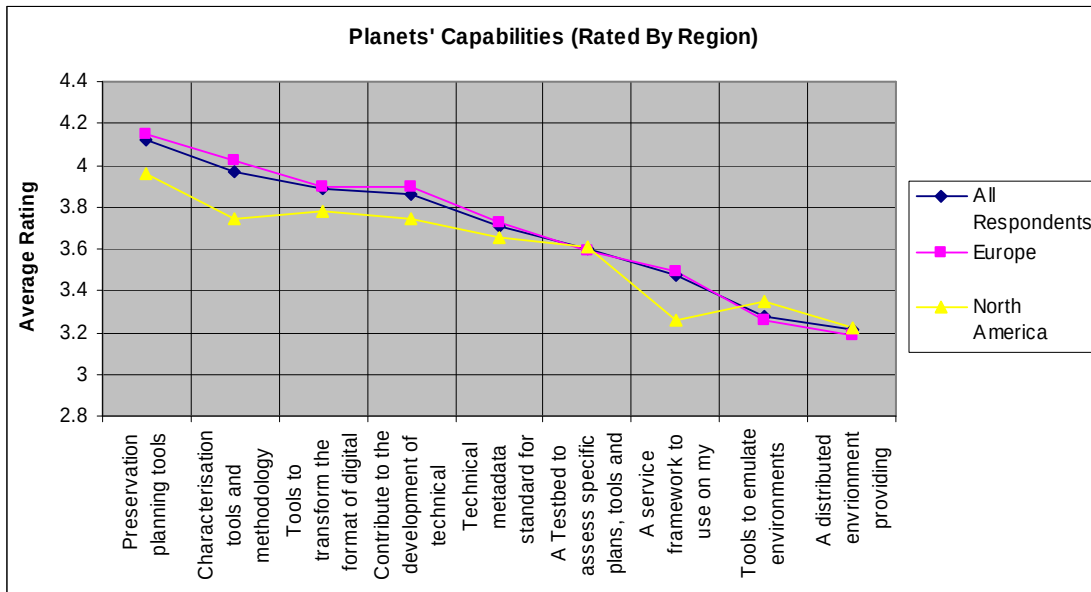


Figure 46: Interest in Planets' Capabilities from Respondents in Different Regions (133 Responses)

### 3.4 Rating Planets Services

There is generally less interest in the services that Planets could offer than in the capabilities it could offer. There are only two services in which respondents are as interested as they are in the top four capabilities, which are

- Information about the latest developments
- A portal for Planets' tools and Testbed results.

At the other end of the scale, there is little interest in Planets providing a digital rescue team and not much more interest in a register of Planets' practitioners or a service desk and technical support.

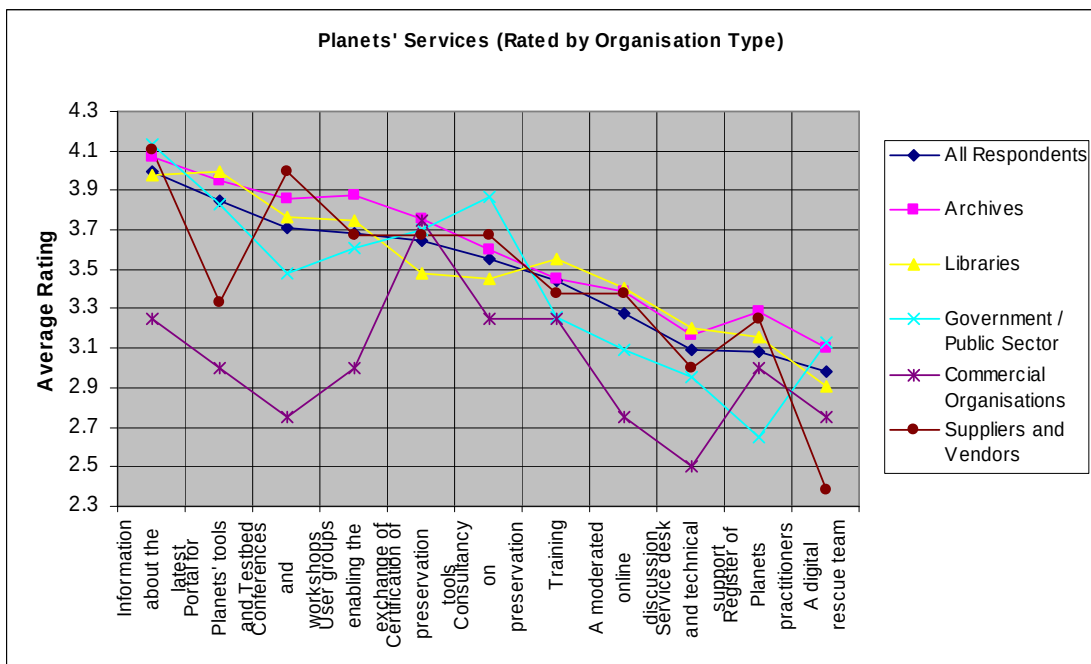


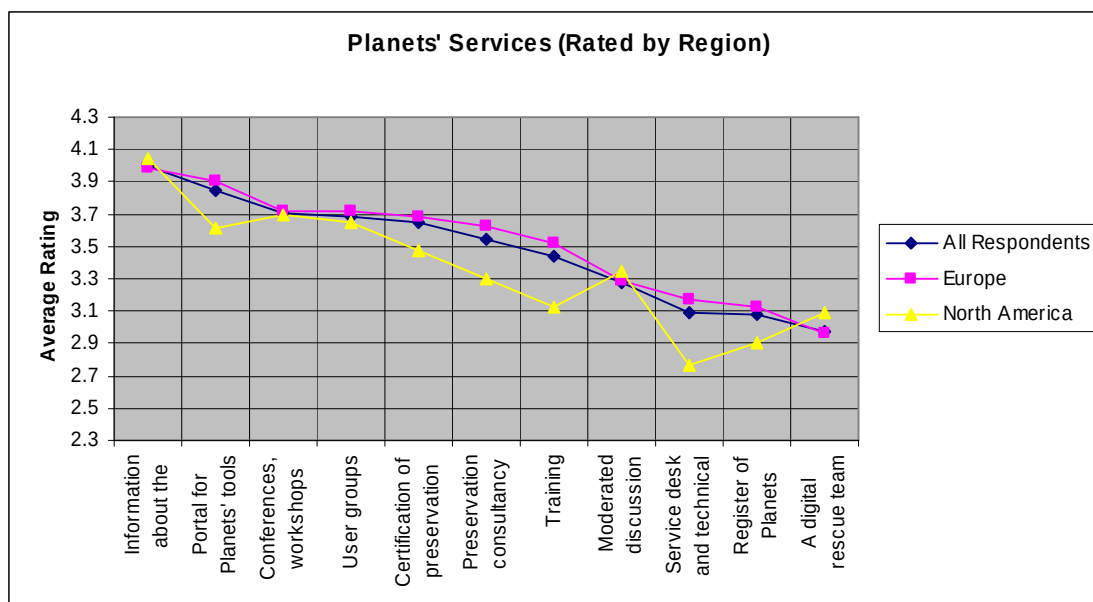
Figure 47: Interest in Planets' Services from Respondents in Different Types of Organisation (131 Responses)

One possible reason for this lack of interest in the services that Planets could offer is the general perception that projects are short-term, which raises the question of how will a short-term entity provide services in the medium to long term.

The archives and libraries have similar views as to their interest in the range of services. The public sector however has its own opinions, being more interested in consultancy on preservation and archiving and less interested in conferences and workshops and a register of Planets' practitioners than the archives and libraries.

Again commercial organisations have a very different perspective to the other sectors; although we must not read too much into the data, due to the tiny sample size. Their main interest is in the certification of preservation tools, with some interest in receiving information about the latest developments, training and consultancy on preservation and archiving.

Suppliers and vendors are particularly uninterested in a digital rescue team, perhaps because they see this as a service they could offer themselves. They are also noticeably less interested in a Planets portal than the libraries, archives and public sector, though why this is so is open to speculation.



**Figure 48: Interest in Planets' Services from Respondents in Different Region (131 Responses)**

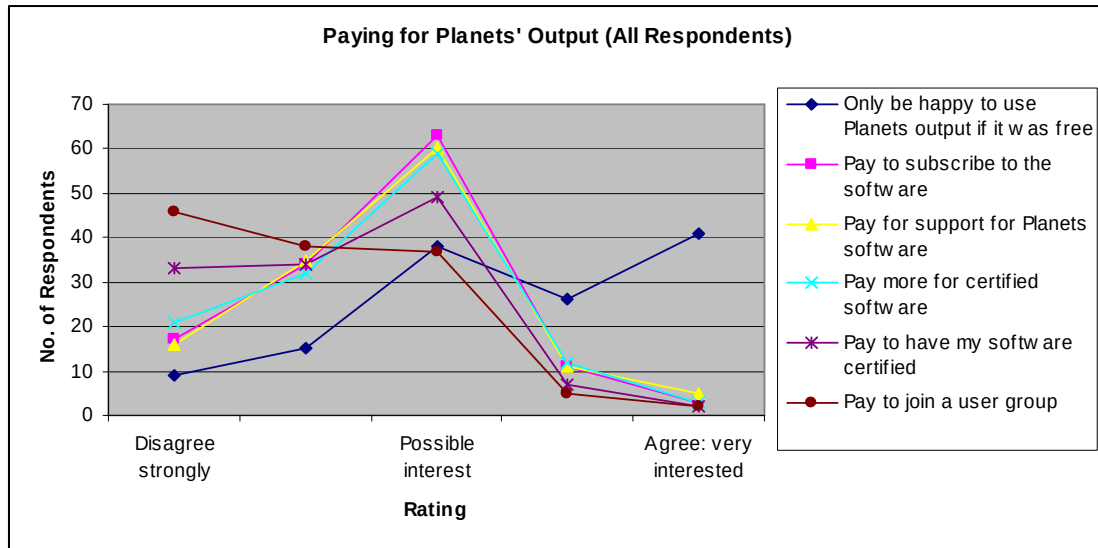
Once again the North Americans are generally less interested than the Europeans in Planets' offerings. In particular, they are noticeably less interested in a Planets portal, training and a service desk and technical support. The relative lack of interest in the last two is easy to explain as both are geographically located (with a service desk it is the different time zones and the relatively small overlap of the working day between the two regions) and it would be harder for the North Americans to make use of them.

### 3.5 Funding Planets

While people prefer not to pay for things there is significant "possible interest" in paying for Planets software, support and certified software. In general people are not prepared to pay to have their software certified or to join a user group. Since people prefer to get things for free, the results of this question were always going to be skewed towards not paying for anything. So any indication of possible interest in paying for one of the offerings is a positive statement. Therefore, it is especially important to look at the relative trends, rather than the absolute values, between the different offerings and between different groups of respondents.

The following graphs show the willingness of respondents to pay for Planets' offerings. The rating values used are equivalent to:

- Disagree strongly (1)
- Possible interest (3)
- Agree: very interested (5)



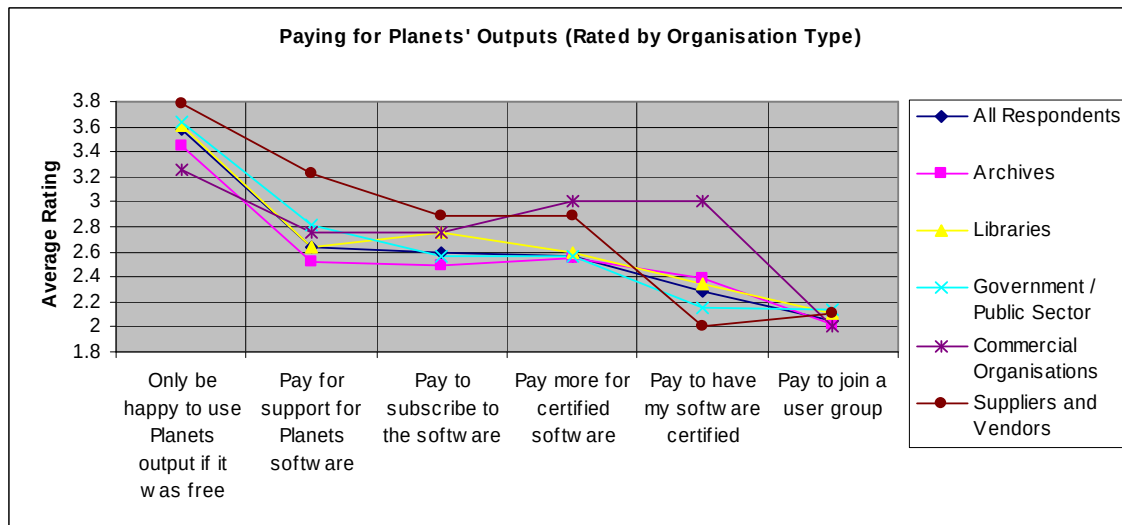
**Figure 49: Interest in Paying for the Outputs from Planets (130 Responses)**

Looking at the distribution of responses for the different offerings, it is clear that there are three different distributions of responses.

The first distribution is that for not paying for Planets' output, which is strongly skewed towards being interested in this option. The mean (3.58), mode (5) and median (4) are all in the interested or very interested rating range. This distribution is not surprising, as people prefer not to pay for things if they can avoid doing so.

The second distribution applies to joining a user group and paying to have software certified and is strongly weighted towards not being interested in paying for the offering. In both cases the median is 2 (expressing disagreement with paying for the offering). For paying to have software certified the mean is 2.29 and the mode 3 (possible interest), and for paying to join a user group the mean is 2.05 and the mode is 1 (disagree strongly).

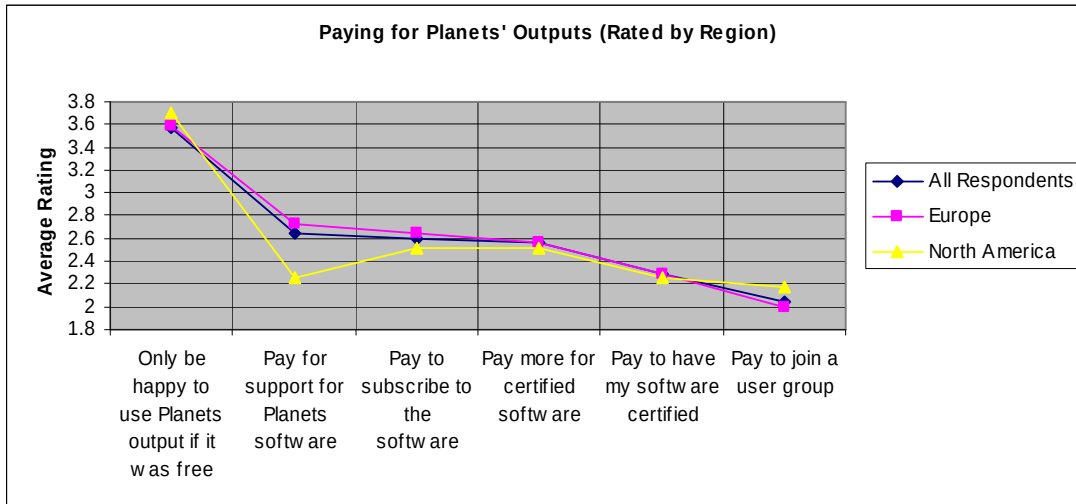
The final distribution applies to paying for Planets software, support and certified software. This distribution is the closest to a normal distribution, but is weighted towards not being interested in paying. For all three options the mode and median are 3 (i.e. "possible interest") and the means are close to 2.6 (2.6, 2.64 and 2.56 respectively). It is these three offerings, out of the five on offer, that respondents are most willing to pay for, although two fifths of respondents state that they are not willing to pay for them.



**Figure 50: Interest from Respondents from Different Types of Organisation in Paying for Planets' Outputs (130 Responses)**

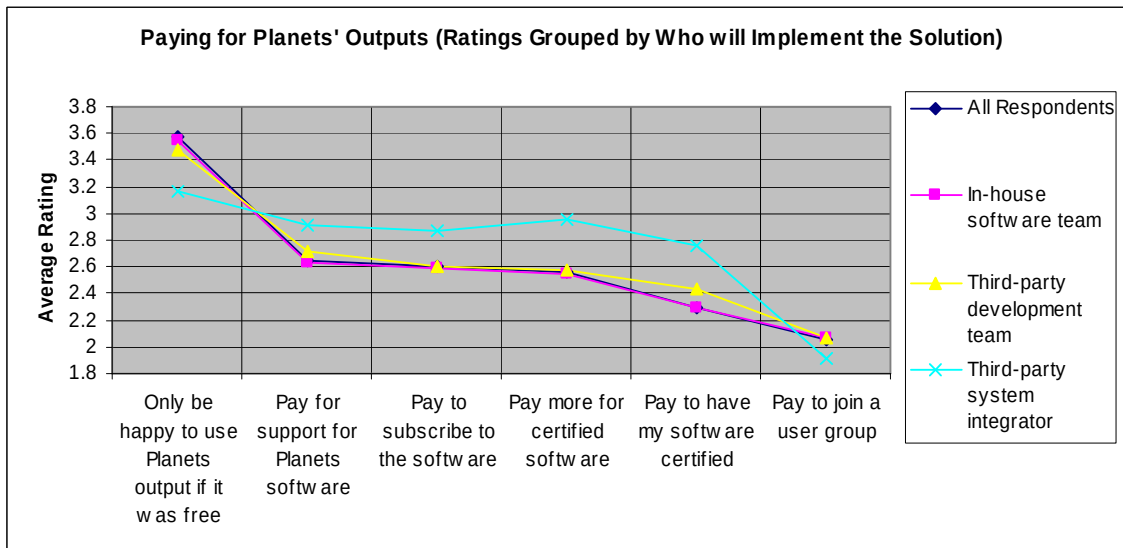


What is noticeable is that suppliers and vendors, who may be looking to build a commercial product around Planets' software, are far more willing to pay for support for Planets' software, to subscribe to the software and to pay more for certified software than the main consumers of such software: the archives, libraries and the public sector (including government departments). Interestingly, the suppliers and vendors are not prepared to pay to have their own software certified. Commercial organisations, on the other hand, are interested in having their software certified and are prepared to pay more for certified software. They are also not so definite about only using Planets' output if it is free.



**Figure 51: Interest from Respondents from Different Regions in Paying for Planets' Outputs (130 Responses)**

Analysing the data according to the region the respondent comes from, it turns out that the North Americans are far less interested in paying for support for Planets' software, and more interested in paying to join a user group. Does this indicate a North American preference for getting support from their peers, rather than a central authority?



**Figure 52: Interest in Paying for Planets' Outputs, Correlated with who will Implement the Solution (118 Responses)**

Interestingly, when the data is analysed by the type of team who will be used to implement any digital preservation solution, it becomes clear that respondents using a third-party system integrator have a different viewpoint from everyone else. They are much more willing to pay for support, to subscribe to the software, to have their software certified and to pay more for certified software but less likely to pay to join a user group.

### 3.6 Summary of Interest in Planets' Capabilities and Services

All Planets' possible outputs, both capabilities and services, were rated as being of at least "possible interest" by the survey respondents, but some outputs received more interest than others. The top three outputs are:

- Preservation planning tools
- Information about the latest developments
- Characterisation tools and methodology

As Table 6 shows, there was some "possible interest" in paying for support for Planets' software, paying to subscribe to the software and paying more for certified software.

| Planets' Offerings                                 | Interest in Paying |
|----------------------------------------------------|--------------------|
| Only be happy to use Planets output if it was free | 3.58               |
| Pay for support for Planets software               | 2.64               |
| Pay to subscribe to the software                   | 2.60               |
| Pay more for certified software                    | 2.56               |
| Pay to have my software certified                  | 2.29               |
| Pay to join a user group                           | 2.05               |

**Table 6: Interest in Paying for Planets' Offerings (130 Responses)**

When willingness to pay is combined with interest in the different capabilities and services, it is possible to get some idea of which capabilities and services might provide an income stream for Planets in the future, and this information is shown in Table 7 which ranks Planets' potential capabilities and services by their average rating. Note that it is not possible to determine whether people would be willing to pay for some of the offerings, as the relevant questions were not asked in the survey.

| Position   | Planets' Capabilities and Services                               | Average Rating | Interest in Paying |
|------------|------------------------------------------------------------------|----------------|--------------------|
| 1          | Preservation planning tools                                      | 4.12           | 2.60               |
| 2          | Information about the latest developments                        | 4.00           |                    |
| 3          | Characterisation tools and methodology                           | 3.97           | 2.60               |
| 4          | Tools to transform the format of digital objects                 | 3.89           | 2.60               |
| 5          | Contribute to the development of technical standards             | 3.86           |                    |
| 6          | Portal for Planets' tools and Testbed results                    | 3.85           |                    |
| 7          | Technical metadata standard for preservation                     | 3.71           |                    |
| 8          | Conferences and workshops                                        | 3.71           |                    |
| 9          | User groups enabling the exchange of best practice               | 3.68           | 2.05               |
| 10         | Certification of preservation tools                              | 3.65           | 2.29               |
| 11         | A Testbed to assess specific plans, tools and services           | 3.60           |                    |
| 12         | Consultancy on preservation and archiving                        | 3.55           |                    |
| 13         | A service framework to use on my digital archive                 | 3.47           | 2.60               |
| 14         | Training                                                         | 3.44           |                    |
| 15         | Tools to emulate environments                                    | 3.28           | 2.60               |
| 16         | A moderated online discussion forum                              | 3.28           |                    |
| 17         | A distributed environment providing access to different services | 3.21           |                    |
| 18         | Service desk and technical support                               | 3.09           | 2.64               |
| 19         | Register of Planets practitioners                                | 3.08           |                    |
| 20         | A digital rescue team                                            | 2.98           |                    |
| <b>Key</b> |                                                                  |                |                    |
| Orange     | Some interest in paying                                          |                |                    |
| Red        | Lack of interest in paying                                       |                |                    |
| White      | No data available about people's willingness to pay              |                |                    |

**Table 7: Relative Rankings of Planets' Potential Capabilities and Services Correlated with Interest in Paying (133 Responses)**

What the table shows is that Planets may be able to derive some income from charging customers to subscribe to the following tool sets:

- Preservation planning tools
- Characterisation tools and methodology
- Tools to transform the format of digital objects

It might also be possible to receive some income from charging for access to a portal for Planets' tools and Testbed results although there are no data to support this. Similarly, Planets could charge for attendance at conferences and workshops, although competition from other digital preservation conferences and workshops means that this is unlikely to generate much of a profit for Planets.

It is interesting that respondents are willing to pay more for certified software, but they are not willing to pay to have their software certified. However, suppliers and vendors might be willing to pay to have their software certified once they are convinced that consumers will pay more for certified software, and therefore Planets might be able to charge for this service.

Also, although people are generally interested in user groups which enable the exchange of best practice, they are not willing to pay to join them.

---

## 4. Conclusions

The message about digital preservation has got out; 90% of organisations that responded to the survey are aware of the challenges presented by the long-term management of digital information. However, only half of these organisations have taken the first step towards meeting those challenges by putting a digital preservation policy in place. Having a digital preservation policy leads onto including digital preservation in all aspects of an organisation's plans (operational, financial and business continuity), having a budget for digital preservation and starting to put, or having put, a digital preservation solution in place.

Organisations need to be encouraged to move from merely being aware of the challenges of digital preservation to tackling them. In particular work is required in government departments and the public sector in general as only a quarter of these organisations have a digital preservation policy. Further education and assistance in working towards putting a suitable solution in place, starting with developing a digital preservation policy, is likely to be required.

Currently only 48% of organisations without a digital preservation policy need to store digital information, compared with 66% of organisations with a digital preservation policy. However, in 10 years' time, organisations without a digital preservation policy will have caught up and 66% of them will need to store digital information, compared with 68% of organisations with a digital preservation policy. Therefore, over the next 10 years the increasing need to preserve their digital information is likely to provide an impetus to put a digital preservation solution in place.

Digital preservation solutions are likely to be componentised, mix-and-match solutions since over 60% of organisations expect to integrate components into a custom solution and nearly 60% of organisations expect to use a mixture of open source and proprietary software. They need to be scalable, particularly to handle large volumes of content and high ingest rates; currently over 70% of organisations hold less than 100 TB, whereas in 10 years' time over 70% of organisations expect to store over 100 TB, and 42% expect to store over 1 PB.

Digital archives will need to handle a wide range of formats since less than a third of organisations think that they have complete control over the formats that they will accept and ingest into their archives. In particular, archives will need to preserve digital information from file systems, document scanning programmes, the internet, electronic document management systems, e-mail systems and media digitisation programmes: the main types of digital information being documents, images and databases.

The most important capability of a digital archive according to the respondents is that it must maintain the authenticity, reliability and integrity of records. Migration is currently the favoured way to deal with technical obsolescence, rather than emulation.

There is a need for digital preservation solutions, and the components and tools that go to make them up, now, as three quarters of organisations expect to invest in a digital preservation solution within the next 2 years.

---

### 4.1 Implications for Planets

There is a high degree of awareness of and interest in Planets, which can be capitalised upon with regular electronic updates about Planets' work; 80% of respondents have heard of Planets and 75% of respondents would like to receive electronic updates about Planets. However, the diversity of respondents to the survey illustrates the difficulty of reaching all the potential "customers" for Planets. In particular, awareness of Planets is lower in academia, government departments and the public sector in general.

Given that organisations with digital preservation policies are 3 times more likely to have a budget for digital preservation, a good initial starting point in promoting Planets would be to ask people whether they have a digital preservation policy and for those who do not, to offer some kind of initial guide towards forming a digital preservation policy. The lack of digital preservation solutions (three-quarters of organisations do not have one), despite the awareness of the problems, is an opportunity for Planets, who could help with information advice, and tools.

Most respondents expect to either use an off the shelf solution and/or integrate existing components and both these approaches are compatible with Planets usage. Similarly, a majority

of respondents expect to use a mixture of open-source and proprietary software, which again is compatible with Planets usage.

Half of all organisations are actively seeking or working on a digital preservation solution and over 85% of organisations with a digital preservation policy expect to make an investment in a digital preservation solution within 2 years. So, now is a good time to be providing digital preservation tools and solutions and there is a large market there for Planets.

Planets needs to target its offerings at the right market. For those with a digital preservation policy, who are quite far down the road, it is Planets' tools and solutions that will be of interest. Whereas for those without a digital preservation policy, who are just starting out, there is more need for consultancy and advice before they can adopt a solution.

## Appendix A Summary of Survey Results

### A.1 About you and your organisation

#### A.1.1 What is your role?

| Answer Options           | Response Count |
|--------------------------|----------------|
|                          | 196            |
| <i>answered question</i> | <b>196</b>     |
| <i>skipped question</i>  | <b>10</b>      |

All answers were individual and have not been included in this report.

#### A.1.2 Please enter your contract information

| Answer Options           | Response Frequency | Response Count |
|--------------------------|--------------------|----------------|
| Name:                    | 100.0%             | 196            |
| Organisation:            | 100.0%             | 196            |
| Address:                 | 97.4%              | 191            |
| Address 2:               | 34.7%              | 68             |
| City/Town:               | 99.5%              | 195            |
| County:                  | 57.1%              | 112            |
| Post Code:               | 96.4%              | 189            |
| Country:                 | 100.0%             | 196            |
| Email Address:           | 100.0%             | 196            |
| Phone Number:            | 84.7%              | 166            |
| <i>answered question</i> |                    | <b>196</b>     |
| <i>skipped question</i>  |                    | <b>10</b>      |

A high percentage of respondents were prepared to answer this question, with fewer than 20 obscuring their contact details.

**A.1.3 Which title would best describe your organisation?**

| Answer Options                        | Response Frequency | Response Count |
|---------------------------------------|--------------------|----------------|
| National Archive                      | 9.3%               | 17             |
| National Library                      | 19.7%              | 36             |
| Museum                                | 2.7%               | 5              |
| Government department / public sector | 15.3%              | 28             |
| Academic Archive                      | 3.3%               | 6              |
| Academic Library                      | 16.4%              | 30             |
| Software Developer or Vendor          | 2.7%               | 5              |
| Systems Integrator or Consultancy     | 3.3%               | 6              |
| Repository services provider          | 0.5%               | 1              |
| Commercial Organisation               | 4.4%               | 8              |
| Other Library                         | 4.9%               | 9              |
| Other Archive                         | 17.5%              | 32             |
| Other (please specify)                |                    | 49             |
| <i>answered question</i>              |                    | <b>183</b>     |
| <i>skipped question</i>               |                    | <b>23</b>      |

**A.1.4 How do you stay informed about the latest developments in the long term management of digital information?**

| Answer Options                    | Response Frequency | Response Count |
|-----------------------------------|--------------------|----------------|
| Professional press                | 69.1%              | 125            |
| Events (exhibitions, conferences) | 77.3%              | 140            |
| Mailing lists                     | 76.2%              | 138            |
| Weblogs                           | 43.1%              | 78             |
| Training workshops                | 41.4%              | 75             |
| Consult IT specialists in-house   | 26.5%              | 48             |
| Professional body                 | 34.8%              | 63             |
| Other (please specify)            |                    | 24             |
| <i>answered question</i>          |                    | <b>181</b>     |
| <i>skipped question</i>           |                    | <b>25</b>      |

**A.1.5 Please indicate here if you would like to receive electronic updates about Planets.**

| Answer Options           | Response Frequency | Response Count |
|--------------------------|--------------------|----------------|
| Yes                      | 79.6%              | 156            |
| No                       | 20.4%              | 40             |
| <i>answered question</i> |                    | <b>196</b>     |
| <i>skipped question</i>  |                    | <b>10</b>      |

## A.2 Policies

### A.2.1 Does your organisation have a policy regarding the long-term (i.e. greater than five years) management of digital information?

| Answer Options           | Response Frequency | Response Count |
|--------------------------|--------------------|----------------|
| Yes                      | 48.4%              | 78             |
| No                       | 51.6%              | 83             |
| Please specify           |                    | 88             |
| <i>answered question</i> |                    | <b>161</b>     |
| <i>skipped question</i>  |                    | <b>45</b>      |

### A.2.2 How frequently do you review these policies?

| Answer Options                   | Response Frequency | Response Count |
|----------------------------------|--------------------|----------------|
| More than once per year          | 6.8%               | 9              |
| Each year                        | 12.9%              | 17             |
| Every two to five years          | 35.6%              | 47             |
| The policy has not been reviewed | 44.7%              | 59             |
| <i>answered question</i>         |                    | <b>132</b>     |
| <i>skipped question</i>          |                    | <b>74</b>      |

### A.2.3 Do you currently have specific budgets for dealing with the long-term management of digital information? If so are they capital or revenue budgets?

| Answer Options               | Response Frequency | Response Count |
|------------------------------|--------------------|----------------|
| Yes - capital only           | 20.9%              | 32             |
| Yes - revenue only           | 7.2%               | 11             |
| Yes - both capital & revenue | 19.0%              | 29             |
| No                           | 52.9%              | 81             |
| <i>answered question</i>     |                    | <b>153</b>     |
| <i>skipped question</i>      |                    | <b>53</b>      |

### A.2.4 Does digital preservation feature in your:

| Answer Options           | Yes | No | Response Count |
|--------------------------|-----|----|----------------|
| Operational planning     | 115 | 37 | 152            |
| Financial planning       | 90  | 56 | 146            |
| <i>answered question</i> |     |    | <b>153</b>     |
| <i>skipped question</i>  |     |    | <b>53</b>      |



**A.2.5 Does the long-term management of digital information form any part of your organisation's business continuity planning?**

| Answer Options           | Response Frequency | Response Count |
|--------------------------|--------------------|----------------|
| Yes                      | 70.8%              | 109            |
| No                       | 29.2%              | 45             |
| <i>answered question</i> |                    | <b>154</b>     |
| <i>skipped question</i>  |                    | <b>52</b>      |

**A.2.6 In general, what best describes the level of awareness in your organisation about the challenges presented by the long term management of digital information?**

| Answer Options                                              | Response Frequency | Response Count |
|-------------------------------------------------------------|--------------------|----------------|
| Not aware/ no experience                                    | 7.0%               | 11             |
| Aware of problems - have not considered solutions           | 17.8%              | 28             |
| Aware of problems - actively seeking/ working on a solution | 51.0%              | 80             |
| Expert - have a solution in place or planned                | 24.2%              | 38             |
| <i>answered question</i>                                    |                    | <b>157</b>     |
| <i>skipped question</i>                                     |                    | <b>49</b>      |

### A.3 Current Status and Plans

**A.3.1 Please describe your plans for the long-term management of digital information.**

| Answer Options                                     | Response Frequency | Response Count |
|----------------------------------------------------|--------------------|----------------|
| No plans                                           | 13.3%              | 18             |
| Assessing needs and requirements using consultancy | 24.4%              | 33             |
| Assessing needs and requirements using a prototype | 20.0%              | 27             |
| Tendering for a long term solution                 | 11.9%              | 16             |
| Long term solution in development                  | 43.0%              | 58             |
| Already have a long term solution                  | 16.3%              | 22             |
| Other (please specify)                             |                    | 35             |
| <i>answered question</i>                           |                    | <b>135</b>     |
| <i>skipped question</i>                            |                    | <b>71</b>      |

**A.3.2 If you are currently looking to procure or are assessing needs, when do you expect to invest in or acquire a solution?**

| Answer Options           | Response Frequency | Response Count |
|--------------------------|--------------------|----------------|
| In progress              | 32.5%              | 40             |
| 0 - 6 months             | 3.3%               | 4              |
| 6 - 24 months            | 41.5%              | 51             |
| > 24 months              | 22.8%              | 28             |
| <i>answered question</i> |                    | <b>123</b>     |
| <i>skipped question</i>  |                    | <b>83</b>      |

**A.3.3 How do you expect to implement your solution?**

| Answer Options                            | Response Frequency | Response Count |
|-------------------------------------------|--------------------|----------------|
| Off-the-shelf software package            | 31.8%              | 42             |
| Integrate components into a custom system | 64.4%              | 85             |
| Develop a custom solution                 | 33.3%              | 44             |
| <i>answered question</i>                  |                    | <b>132</b>     |
| <i>skipped question</i>                   |                    | <b>74</b>      |

**A.3.4 Who do you expect to implement your solution?**

| Answer Options                | Response Frequency | Response Count |
|-------------------------------|--------------------|----------------|
| In-house software team        | 69.2%              | 92             |
| Third-party development team  | 45.9%              | 61             |
| Third-party system integrator | 21.1%              | 28             |
| <i>answered question</i>      |                    | <b>133</b>     |
| <i>skipped question</i>       |                    | <b>73</b>      |

**A.3.5 Which of these software tools have you:**

| Answer Options                    | Not heard of | Heard of | Evaluated | Used | Short-listed | Selected | Response Count |
|-----------------------------------|--------------|----------|-----------|------|--------------|----------|----------------|
| Dspace                            | 29           | 49       | 34        | 15   | 1            | 3        | 131            |
| EMC Centera                       | 65           | 44       | 9         | 4    | 0            | 1        | 123            |
| E-prints                          | 41           | 58       | 20        | 5    | 1            | 5        | 130            |
| Ex Libris Digitool                | 40           | 60       | 15        | 3    | 2            | 4        | 124            |
| Ex Libris Rosetta                 | 49           | 62       | 8         | 0    | 4            | 2        | 125            |
| Fedora                            | 24           | 54       | 31        | 11   | 3            | 7        | 130            |
| Hitache Archivas                  | 87           | 26       | 3         | 2    | 0            | 1        | 119            |
| IBM DIAS                          | 52           | 46       | 18        | 3    | 2            | 1        | 122            |
| Livelink Archive                  | 74           | 38       | 6         | 4    | 0            | 1        | 123            |
| Tessella Safety Deposit Box (SDB) | 55           | 50       | 4         | 6    | 4            | 3        | 122            |
| VTLS Vital                        | 81           | 26       | 6         | 2    | 2            | 2        | 119            |
| Other (please specify)            |              |          |           |      |              |          | 27             |
| answered question                 |              |          |           |      |              |          | 137            |
| skipped question                  |              |          |           |      |              |          | 69             |

**A.3.6 Do you use, or plan to use proprietary or open source software?**

| Answer Options           | Now | In the future | Response Count |
|--------------------------|-----|---------------|----------------|
| Proprietary only         | 19  | 3             | 21             |
| Open source only         | 18  | 18            | 28             |
| A mixture of the two     | 76  | 73            | 96             |
| Have not decided         | 21  | 32            | 41             |
| <i>answered question</i> |     |               | <b>139</b>     |
| <i>skipped question</i>  |     |               | <b>67</b>      |

## A.4 Digital content

### A.4.1 Which source systems do you, or will you, take digital information from so that it can be managed for the long-term?

| Answer Options                               | Response Frequency | Response Count |
|----------------------------------------------|--------------------|----------------|
| File system                                  | 77.3%              | 109            |
| Electronic Document Management System (EDMS) | 54.6%              | 77             |
| E-mail system                                | 53.9%              | 76             |
| Computer Aided Design (CAD)                  | 29.1%              | 41             |
| Lab systems                                  | 18.4%              | 26             |
| Media store                                  | 34.0%              | 48             |
| Document scanning programme                  | 58.2%              | 82             |
| Media digitisation programme                 | 53.9%              | 76             |
| ERMS database                                | 31.9%              | 45             |
| Internet                                     | 54.6%              | 77             |
| Other                                        | 15.6%              | 22             |
| Other (please specify)                       |                    | 31             |
| <i>answered question</i>                     |                    | <b>141</b>     |
| <i>skipped question</i>                      |                    | <b>65</b>      |

### A.4.2 What types of digital information do you, or will you, need to manage for the long-term?

| Answer Options           | At the moment | In two years | In five years | In ten years | Response Count |
|--------------------------|---------------|--------------|---------------|--------------|----------------|
| Documents                | 110           | 12           | 3             | 9            | 134            |
| Audio                    | 68            | 21           | 10            | 11           | 110            |
| Images                   | 110           | 10           | 2             | 9            | 131            |
| Video                    | 70            | 14           | 11            | 10           | 105            |
| eBooks                   | 29            | 17           | 13            | 8            | 67             |
| eJournals                | 34            | 13           | 11            | 9            | 67             |
| Scientific data          | 31            | 18           | 13            | 8            | 70             |
| Databases                | 67            | 24           | 14            | 12           | 117            |
| GIS                      | 32            | 14           | 16            | 10           | 72             |
| ISO or disc images       | 28            | 10           | 10            | 7            | 55             |
| Software                 | 29            | 15           | 11            | 12           | 67             |
| Websites                 | 65            | 28           | 8             | 8            | 109            |
| E-mails                  | 54            | 24           | 7             | 6            | 91             |
| Documents                | 76            | 10           | 0             | 8            | 94             |
| Other (please specify)   |               |              |               |              | 12             |
| <i>answered question</i> |               |              |               |              | <b>138</b>     |
| <i>skipped question</i>  |               |              |               |              | <b>68</b>      |

#### A.4.3 How much control do you have over the format of the content in your digital archive?

| Answer Options                                                                              | Response Frequency | Response Count |
|---------------------------------------------------------------------------------------------|--------------------|----------------|
| We have complete control over the formats we will accept and enter into our archives        | 27.0%              | 37             |
| We work with providers of content to influence the formats we will accept into our archives | 42.3%              | 58             |
| We have little/ no control and are obliged to accept formats provided to us                 | 30.7%              | 42             |
| <i>answered question</i>                                                                    |                    | <b>137</b>     |
| <i>skipped question</i>                                                                     |                    | <b>69</b>      |

#### A.4.4 Which of the following do you think are important capabilities for a long-term digital information management system?

| Answer Options                                                             | Not applicable | Least important |    |    | Critical | Rating Average | Response Count |
|----------------------------------------------------------------------------|----------------|-----------------|----|----|----------|----------------|----------------|
| Complies with established data or digital information management standards | 1              | 2               | 14 | 45 | 73       | 3.39           | 135            |
| Maintains authenticity, reliability and integrity of records               | 1              | 1               | 3  | 19 | 111      | 3.76           | 135            |
| Checks records have not been damaged                                       | 2              | 2               | 12 | 28 | 90       | 3.51           | 134            |
| Checks for duplicate items                                                 | 1              | 27              | 52 | 38 | 15       | 2.29           | 133            |
| Integrates with content delivery systems                                   | 4              | 14              | 35 | 46 | 31       | 2.66           | 130            |
| Integrates with content producing and holding systems                      | 7              | 20              | 35 | 42 | 27       | 2.47           | 131            |
| Adheres to metadata standards                                              | 2              | 7               | 16 | 47 | 62       | 3.19           | 134            |
| Is able to store many different types of content                           | 2              | 4               | 17 | 34 | 77       | 3.34           | 134            |
| Handles a wide variety of file formats                                     | 1              | 5               | 21 | 36 | 71       | 3.28           | 134            |
| Retrieves content by description                                           | 3              | 3               | 25 | 54 | 47       | 3.05           | 132            |
| Retrieves content using full text                                          | 1              | 19              | 35 | 46 | 30       | 2.65           | 131            |
| Ensures records are accessible for up to 50 years                          | 4              | 3               | 13 | 29 | 80       | 3.38           | 129            |
| Ensures records are accessible for more than 50 years                      | 4              | 7               | 17 | 26 | 78       | 3.27           | 132            |
| Characterises records by extracting technical metadata                     | 3              | 5               | 26 | 58 | 39       | 2.95           | 131            |
| Plans the preservation of content to deal with technical obsolescence      | 2              | 4               | 10 | 36 | 81       | 3.43           | 133            |
| Performs migrations to deal with technical obsolescence                    | 2              | 2               | 12 | 46 | 71       | 3.37           | 133            |
| Supports emulation to deal with technical obsolescence                     | 9              | 22              | 27 | 46 | 29       | 2.48           | 133            |
| Other (please specify)                                                     |                |                 |    |    |          |                | 15             |
| <i>answered question</i>                                                   |                |                 |    |    |          |                | <b>135</b>     |
| <i>skipped question</i>                                                    |                |                 |    |    |          |                | <b>71</b>      |

**A.4.5 How important is the scalability of your digital information management solution?**

| Answer Options                                                 | Not important |    |    |    | Critical | Rating Average | Response Count |
|----------------------------------------------------------------|---------------|----|----|----|----------|----------------|----------------|
| Scalable to petabytes of content                               | 9             | 12 | 21 | 40 | 50       | 3.83           | 132            |
| Scalable to high ingest rates (millions of objects per year)   | 12            | 15 | 21 | 39 | 46       | 3.69           | 133            |
| Scalable to high access rates (hundreds of objects per second) | 19            | 22 | 40 | 28 | 23       | 3.11           | 132            |
| <i>answered question</i>                                       |               |    |    |    |          |                | <b>133</b>     |
| <i>skipped question</i>                                        |               |    |    |    |          |                | <b>73</b>      |

**A.4.6 How much digital content do you plan to store?**

| Answer Options           | Nothing | <1 TB | 1 - 20 TB | 20 - 100 TB | 100 - 500 TB | 500 TB - 1 PB | >1 PB | Response Count |
|--------------------------|---------|-------|-----------|-------------|--------------|---------------|-------|----------------|
| At the moment            | 13      | 25    | 46        | 25          | 7            | 4             | 6     | 126            |
| In two years             | 3       | 15    | 28        | 36          | 26           | 8             | 9     | 125            |
| In five years            | 3       | 6     | 18        | 28          | 30           | 22            | 19    | 126            |
| In ten years             | 3       | 4     | 11        | 18          | 18           | 19            | 53    | 126            |
| <i>answered question</i> |         |       |           |             |              |               |       | <b>129</b>     |
| <i>skipped question</i>  |         |       |           |             |              |               |       | <b>77</b>      |

**A.5 Functionality - Capabilities****A.5.1 Which of the following do you think are important capabilities for a long-term digital information management system?**

| Answer Options         | Not heard of | Heard of | Plan to use | Already use | Response Count |
|------------------------|--------------|----------|-------------|-------------|----------------|
| DDI                    | 66           | 41       | 7           | 2           | 116            |
| Dublin Core            | 6            | 35       | 23          | 67          | 131            |
| EAD                    | 27           | 53       | 18          | 26          | 124            |
| MARC                   | 22           | 53       | 6           | 41          | 122            |
| MARC-XML               | 22           | 59       | 11          | 29          | 121            |
| METS                   | 19           | 45       | 31          | 29          | 124            |
| MODS                   | 38           | 52       | 11          | 17          | 118            |
| NISO-MIX               | 64           | 35       | 10          | 12          | 121            |
| ONIX                   | 75           | 35       | 3           | 4           | 117            |
| PREMIS                 | 32           | 35       | 39          | 17          | 123            |
| TEI                    | 52           | 43       | 12          | 14          | 121            |
| ISAD (G)               | 51           | 28       | 13          | 35          | 127            |
| MAB/MAB 2              | 84           | 26       | 1           | 7           | 118            |
| Other (please specify) |              |          |             |             | 20             |
| answered question      |              |          |             |             | 134            |
| skipped question       |              |          |             |             | 72             |

## A.6 Planets

### A.6.1 How aware are you of the Planets project?

| Answer Options                        | Response Frequency | Response Count |
|---------------------------------------|--------------------|----------------|
| Not heard of                          | 20.4%              | 28             |
| Have read about                       | 35.0%              | 48             |
| Have attended briefings or conference | 10.9%              | 15             |
| Have met members                      | 21.9%              | 30             |
| I am a member                         | 11.7%              | 16             |
| <i>answered question</i>              |                    | <b>137</b>     |
| <i>skipped question</i>               |                    | <b>69</b>      |

### A.6.2 I see Planets as a potentially valuable resource for the following capabilities:

| Answer Options                                                   | Disagree: not interested |    | Possible interest |    | Agree: very interested | Rating Average | Response Count |
|------------------------------------------------------------------|--------------------------|----|-------------------|----|------------------------|----------------|----------------|
| A service framework to use on my digital archive                 | 2                        | 8  | 72                | 22 | 25                     | 3.47           | 129            |
| Preservation planning tools                                      | 1                        | 2  | 37                | 32 | 60                     | 4.12           | 132            |
| Characterisation tools and methodology                           | 0                        | 3  | 45                | 36 | 47                     | 3.97           | 131            |
| Tools to transform the format of digital objects                 | 3                        | 8  | 39                | 33 | 49                     | 3.89           | 132            |
| Tools to emulate environments                                    | 10                       | 26 | 45                | 19 | 32                     | 3.28           | 132            |
| A distributed environment providing access to different services | 9                        | 16 | 61                | 28 | 17                     | 3.21           | 131            |
| A Testbed to assess specific plans, tools and services           | 3                        | 15 | 44                | 39 | 30                     | 3.60           | 131            |
| Technical metadata standard for preservation                     | 3                        | 12 | 46                | 29 | 41                     | 3.71           | 131            |
| Contribute to the development of technical standards             | 1                        | 6  | 49                | 31 | 45                     | 3.86           | 132            |
| <i>answered question</i>                                         |                          |    |                   |    |                        |                | <b>133</b>     |
| <i>skipped question</i>                                          |                          |    |                   |    |                        |                | <b>73</b>      |

**A.6.3 I see Planets as a resource for value-added services like:**

| Answer Options                                     | Disagree: not interested |    | Possible interest |    | Agree: very interested | Rating Average | Response Count |
|----------------------------------------------------|--------------------------|----|-------------------|----|------------------------|----------------|----------------|
| Training                                           | 7                        | 16 | 52                | 23 | 32                     | 3.44           | 130            |
| Register of Planets practitioners                  | 12                       | 28 | 49                | 20 | 21                     | 3.08           | 130            |
| Service desk and technical support                 | 11                       | 23 | 55                | 22 | 17                     | 3.09           | 128            |
| Certification of preservation tools                | 4                        | 9  | 53                | 28 | 37                     | 3.65           | 131            |
| Consultancy on preservation and archiving          | 6                        | 14 | 44                | 36 | 31                     | 3.55           | 131            |
| Information about the latest developments          | 1                        | 5  | 36                | 40 | 49                     | 4.00           | 131            |
| A moderated online discussion forum                | 10                       | 18 | 54                | 20 | 27                     | 3.28           | 129            |
| User groups enabling the exchange of best practice | 5                        | 10 | 42                | 39 | 35                     | 3.68           | 131            |
| Conferences and workshops                          | 2                        | 10 | 48                | 35 | 36                     | 3.71           | 131            |
| Portal for Planets' tools and Testbed results      | 2                        | 10 | 42                | 29 | 48                     | 3.85           | 131            |
| A digital rescue team                              | 14                       | 26 | 55                | 15 | 18                     | 2.98           | 128            |
| Other (please specify)                             |                          |    |                   |    |                        |                | 3              |
| <b>answered question</b>                           |                          |    |                   |    |                        |                | <b>131</b>     |
| <b>skipped question</b>                            |                          |    |                   |    |                        |                | <b>75</b>      |

**A.6.4 Please state the extent to which you agree or disagree with the following statements. I would:**

| Answer Options                                     | Disagree strongly |    | Possible interest |    | Agree: very interested | Rating Average | Response Count |
|----------------------------------------------------|-------------------|----|-------------------|----|------------------------|----------------|----------------|
| Only be happy to use Planets output if it was free | 9                 | 15 | 38                | 26 | 41                     | 3.58           | 129            |
| Pay to subscribe to the software                   | 17                | 34 | 63                | 11 | 3                      | 2.60           | 128            |
| Pay for support for Planets software               | 16                | 35 | 60                | 11 | 5                      | 2.64           | 127            |
| Pay more for certified software                    | 21                | 32 | 59                | 12 | 3                      | 2.56           | 127            |
| Pay to have my software certified                  | 33                | 34 | 49                | 7  | 2                      | 2.29           | 125            |
| Pay to join a user group                           | 46                | 38 | 37                | 5  | 2                      | 2.05           | 128            |
| <b>answered question</b>                           |                   |    |                   |    |                        |                | <b>130</b>     |
| <b>skipped question</b>                            |                   |    |                   |    |                        |                | <b>76</b>      |

**A.6.5 What could Planets do to be more useful to your organisation?**

This question allowed respondents to write their own answers and all those that were given have been included below, grouped by the type of organisation the respondent works for.

National Archive

- Achieve the project's concrete goals
- More (technical) information about the tools and services developed as well as about possibilities to integrate those into our own digital preservation framework would be welcome
- Have IP Licensing for PLANETS products to enable them to be deployed and supported by third parties.
- As we are in the initial stages of planning for digital preservation I am not too sure just yet.

National Library

- Provide a stable environment for tried and tested tools, rather than a constantly shifting 'in test' environment. this would be very useful for benchmarking process and activities
- Set up demonstration days in Wales.
- Would be useful to integrate output of Planets into IIPC.
- Compendium handbook of tools and results
- We would be interested in the development of a community risk registry.

Museum

No comments.

Government department / public sector

- Too early in the development of digital preservation to comment. Also, UKAEA, will not be a key player in the future, it will be the Nuclear Decommissioning Authority
- I was invited to participate in this survey. Yet I think that these last questions would have to be answered by a particularly responsible unit of our organisation (such as the State Archives).
- Do not know
- Would only pay if the software was better and/or less expensive than other solutions. We use ePrints which is open source software.
- Expand participation beyond the EU.
- I was asked to complete this by a colleague (Andrew McHugh) but it's not really relevant for me to complete so please void the entry. There is a project underway at present at the University of Glasgow which involves scoping requirements for a digital preservation policy. The policy and any solution that may be put in place across the Uni is some way off and will be decided by senior management, not us, so I'm not really in a position to respond to this survey - sorry.
- I have no knowledge of your organization, so I cannot answer this question.

Academic Archive

- Give better information in electronic form. W.R.T. question 28: our financial resources are limited, which means we have to consider very carefully any engagement in paid schemes. A continuation of Planets on a (semi) commercial basis should be supported by a large number of subscribers. We are open for discussion.

Academic Library

No comments.

Software Developer or Vendor

- The current focus is on libraries, which is certainly important for the benefit of mankind. But health records, intellectual property records, and litigation support is where the money is.

Systems Integrator or Consultancy

No comments.

Repository Services Provider

No comments.

Commercial Organisation

No comments.

Other Library

- We are in a budget crisis and cutting staff but may be able to get funds and support from the IT division to purchase Planets products and/or services. An organization's relationship to its IT department is critical to the issues raised in this survey.



[Other Archive](#)

- Share more specific updates outside of members (beyond EU)
- A technical registry with structural updates
- Advise on proprietary formats that are likely to remain accessible
- Tools for long-term preservation of electronic records
- Accept more organisations as a member More insight which organisation get the control over the tools in the future.
- Organise one workshop or session of a workshop on non-textual digital content, especially if audiovisual were a focus