Survey Report on Digitisation in European Cultural Heritage Institutions 2014

by

Natasha Stroeker and René Vogels, Panteia (NL)
on behalf of the ENUMERATE Thematic Network
January 2014

EUMERATE is funded under the
ICT Policy Support Programme part of the
Competitiveness and Innovation Framework Programme.

Contents

EXECUTIVE SUMMARY 5

1 INTRODUCTION 8

2 RESPONSE TO THE SURVEY 9

3 DIGITISATION ACTIVITY 12

3.1 Long term collection management ***12***

3.2 Digital collection ***12***

3.3 Born digital collections ***13***

4 THE COLLECTION 14

4.1 Object types ***14***

4.2 Text based resources ***14***

4.3 Visual 2D resources ***15***

4.4 Archival resources ***16***

4.5 3D man-made resources ***17***

4.6 Natural and geography based resources ***18***

4.7 Time based resources ***19***

4.8 Born digital interactive resources ***20***

4.9 Digital cataloguing ***21***

4.10 Digitisation progress ***21***

4.11 Optional questions: size of the collection ***22***

5 DIGITAL ACCESS 30

5.1 Policy on use of digital collections ***30***

5.2 Reasons for providing digital access ***30***

5.3 Measuring the use of digital collections ***31***

6 DIGITAL ACCESS 33

6.1 Digital preservation strategy ***33***

6.2 Use of international standards for digital preservation ***34***

7 DIGITISATION EXPENDITURE 36

7.1 Internal and external budgets ***36***

7.2 Internal and external budgets ***37***

7.3 Incidental and structural costs ***38***

7.4 In-house costs and outsourced costs 39
7.5 Activities.....39
7.6 Funding 41

APPENDIX – *CORE SURVEY 2* QUESTIONNAIRE..... 42

Executive summary

The *ENUMERATE Survey Report on Digitisation in Cultural Heritage Institutions 2014* represents the third major study into the current state of digitisation in Europe. It is the result of a survey carried out by the ENUMERATE Thematic Network, with the help of national coordinators, in 33 European countries (in 11 languages). Almost 1.400[1] institutions answered the open call to participate between September and December 2013.

The survey asked questions about:

- The state of digitisation activity in the institution responding;
- The collection;
- Digital Access;
- Its digital preservation strategy;
- Expenditure of digitisation by the institution.

Highlights of the report's findings are:

Digitisation activity

- 92% of the respondents are memory institutions with collections to be kept for future generations;
- 87% of institutions have a digital collection (83% in *Core Survey 1*);
- 36% of the institutions have a written digitisation strategy (34% in *Core Survey 1*);
- 53% of the institutions have born digital items (52% in *Core Survey 1*).

The collection

- Most institutions have a rich mix of heritage materials both in *Core Survey 2* as in *Core Survey 1*;
- Most mentioned object type as part of the heritage collection of the institutions is text based (84%) and visual 2D (84%) followed by archival material (64%), time based material (56%) and 3D man-made material (48%);
- Text based resources can be found most at libraries (100%);
- Visual 2D resources can be found most at museums (89%) and archives (85%);
- Archival resources can be found most at archives / record offices (85%);
- 3D man-made resources can be found most with museums (80%).

[1] Throughout this deliverable the, generally European, notation for the thousand separator, and the decimal mark is used.

Digital access

- 34% of the institutions have an explicit policy regarding the use of the digital collections (in *Core Survey 1* this was 31%);
- Academic research is perceived[2] as the most important reason to provide digital access to the collection (8,5), followed by educational use of the collection (7,9). Least important reason is sales and commercial licensing;
- 51% of all institutions measure the use of digital collections (42% in *Core Survey 1*);
- 91% of the institutions use web statistics (85% in Core Survey 1);
- 32% use social media statistics;
- Offline access (42%) and the institutional website (34%) are mentioned most concerning digital access *(Core Survey 1*: 49% offline and 31% institutional website).

Digital preservation

- 26% of the institutions have a written digital preservation strategy that is endorsed by the management (23% in Core Survey 1);
- 48% does not have a solution yet for long term preservation based on international standards for digital preservation (national libraries are 'front runners' in owning a digital archive that meets the international criteria for long term preservation).

Digitisation Expenditure

- The estimated average budget for digital collections is €295.000;
- An average of 8 people is involved on a full-time basis in the digital collection activities for an institution;
- This is made up of 5 full-time equivalent (FTE) paid staff and 3 FTE volunteers. This is about 50% more staff compared to the *Core Survey 1* results (average total: 3,5 FTE paid staff and 2,0 FTE volunteers);
- In total 33% of the paid staff is involved in digitisations activities. The lower the annual budget the higher this percentage is (for instance 92% for annual budget €10.000-50.000, 41% for annual budget €50.000-100.000);
- About 53% of the costs are qualified as being incidental cost and 47% are structural costs;
- 72% of the costs are in-house costs, 28% are out-of-pocket costs for external service providers. These results are the same as in the *Thematic Survey*;
- The analogue to digital conversion process is responsible for 37% of the incidental costs, metadata creation is 19%;
- Archiving (23%) and management (18%) are the main structural cost components;
- The digital collection activities are funded by internal budgets (88%). National public grants are available to almost 40% of the respondents. 22% receives regional or local public grants. The results are very similar to the results of *Core Survey 1*.

The report is the last in a series of three in the lifetime of the *ENUMERATE* project. The first concerned *Core Survey 1* (2012). In 2013 we published the results of an in-depth 'thematic' survey. It is intended that *ENUMERATE* will continue its work beyond 2014.

[2] Based on a scale of 1 to 10, where 1 is 'not important' and 10 is 'very important'

The project is funded by the European Commission under the *ICT Policy Support Programme part of the Competitiveness and Innovation Framework Programme*.

1 Introduction

In the period December 2011-April 2012 the EC-funded *ENUMERATE* network conducted *Core Survey 1*. The aim of this survey, the first in a limited set of surveys, was to develop a clear picture of the progress made in creation, management and preservation of digital collections in the European cultural heritage. *Core Survey 1* was a continuation of the *NUMERIC* study, which was done on behalf of the European Union in 2008.

In 2013 the work to improve the survey continued with a more qualitative test of approaches to measure digitisation efforts. The results of these thematic surveys are used to improve the survey and data collection process for *Core Survey 2*.This report presents the results of *Core Survey 2*.

An overall population database covering all European archives, libraries, museums and other heritage institutions was not available. For that reason in every *ENUMERATE* network country the national coordinator was asked to invite institutions to complete the survey. In some countries it was possible to get a complete database covering all the relevant institutions. In most countries such a database is not yet available. In those cases a sample of e-mail addresses based on a wide range of sources was used. Social media and websites were also used to invite institutions to participate in the project. An open link to the online questionnaire was available. Unforeseen respondents were asked to identify themselves, in order to un-duplicate institution contacts.

Where possible a generic comparison is made with the results of the *NUMERIC* research and *Core Survey 1*.

The data collection process, the analyses of the results and the preparation of the report is carried out by Panteia in The Netherlands.

2 Response to the survey

The first completed survey was registered on November 11th 2013. By the end of the year 5.262 individuals showed some activity on the website. Unfortunately, a large number did not take the step to start answering the questions in the survey. The data collection resulted in 1.373 usable responses including 1 response from the USA which is excluded from the analyses.

Germany, Spain, Netherlands and Sweden show the highest numbers of respondents. Iceland has an excellent response compared to the number of institutions. Some Central and East European countries have good results too, for example Lithuania and Slovenia.

Table 2.1: Response per country

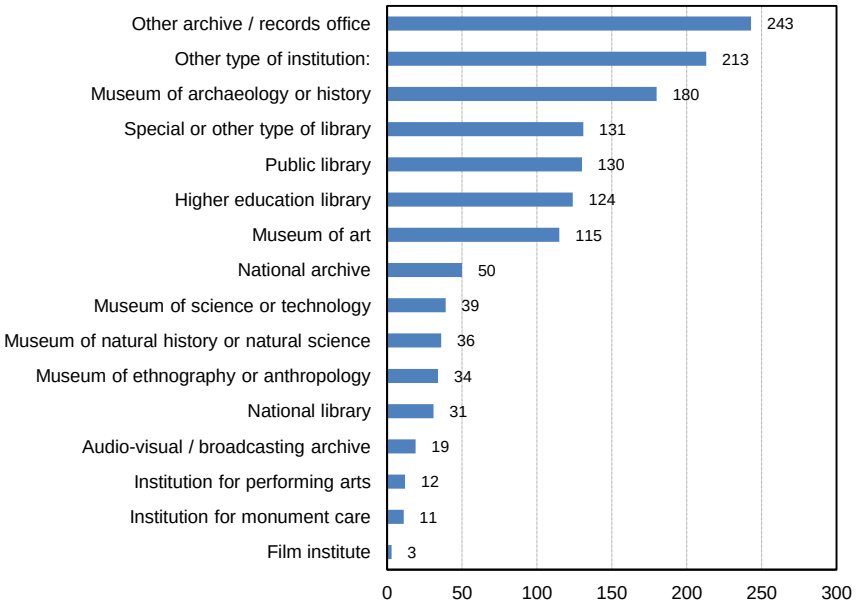
Country	Core survey 1	Core survey 2
Austria	95	36
Belgium	42	29
Bosnia and Herzegovina	0	1
Bulgaria	0	1
Cyprus	21	13
Czech Republic	111	34
Denmark	58	16
Estonia	18	16
Finland	89	59
France	0	2
Germany	227	279
Greece	58	10
Hungary	85	44
Iceland	0	38
Ireland	25	15
Italy	127	25
Latvia	16	4
Liechtenstein	5	1
Lithuania	71	61
Luxembourg	28	15
Malta	4	2
Monaco	0	1
Netherlands	140	143
Poland	58	23
Portugal	85	44
Republic of Macedonia	0	1
Republic of Moldova	0	1
Romania	39	14
Slovak Republic	78	4
Slovenia	56	57
Spain	255	180
Sweden	33	125
Switzerland	77	23
United Kingdom	49	55

Country	Core survey 1	Core survey 2
Other country: USA, not included in analyses		1
Total	1.951	1.373

Institutions by Type

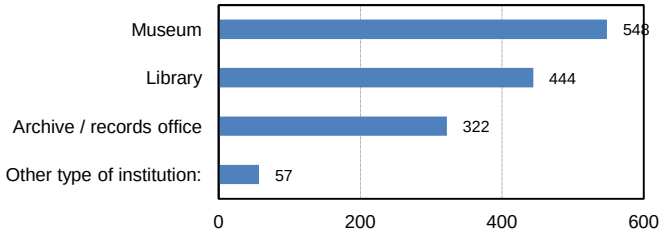
A large number of respondents find it hard to place their institution in one of the categories shown in the survey. 213 institutions are placed in the category 'Other type'. Mostly because of the mixed nature of the institutions they represent. A lot of the 'other types' have a profile matching a more general category of museum, archive or library. In order to categorise these institutions, the websites of these 213 respondents were looked at and these institutions were placed in the high level categories that contained similar institutions.

Figure 2.1: ***Response per institution type*** (n=1.372)



The results in this report will be analysed for the 4 main categories shown below.

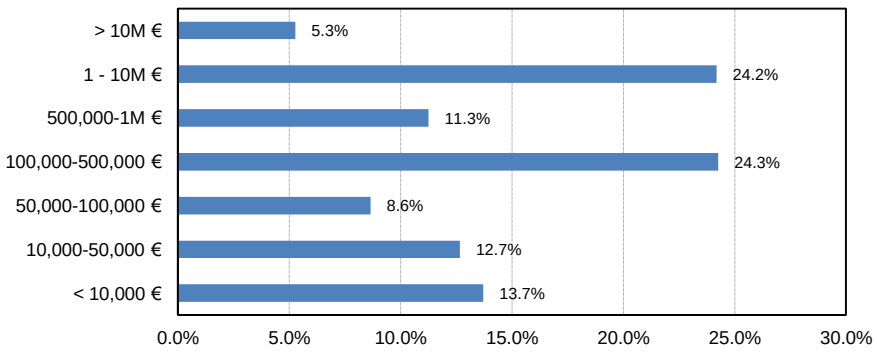
Figure 2.2: ***Response per institution type in 4 categories*** (n=1.371)



Institutions by Annual budget

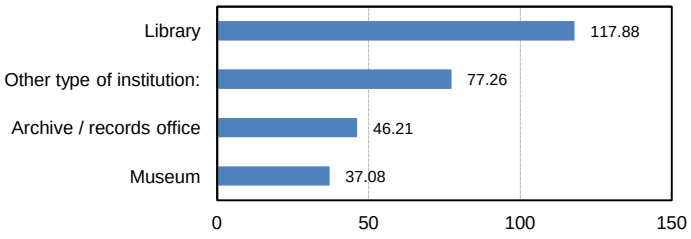
Looking at the annual budget, the responses covered all categories of institution, large as well as small. 29% have an annual budget of over €1 million. In contrast 13,7% have a rather small annual budget not exceeding €10 thousand. The median is the category €100-€500 thousand. Based on size the response is very similar to *Core Survey 1*.

Figure 2.3: ***Annual budget*** (n=1.372)



Another indicator for the size of an institution is the number of paid staff in full-time equivalents (FTE). In this survey the average institution size is almost 67 FTE. In total this survey represents a group of institutions with a total paid staff of 91.940 FTE. The overall median is 12 FTE paid staff. Core Survey 1 had a median staff size of 15FTE.

Figure 2.4: ***Average size of institutions*** (n=1.372)

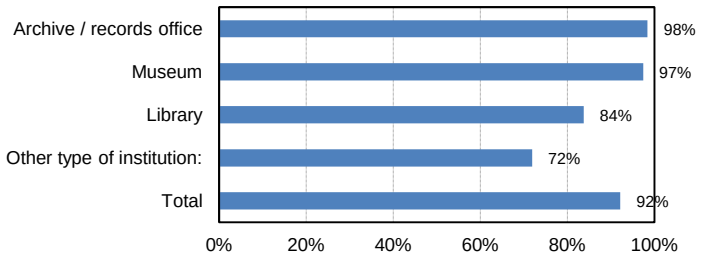


3 Digitisation activity

3.1 Long term collection management

Most institutions have collections that need to be preserved for future generations. For 8% this is not the case. These institutions can be found among libraries and monumental care institutions.

Figure 3.1: ***Does your institution have collections that need to be preserved for future generations?*** (n=1.372)

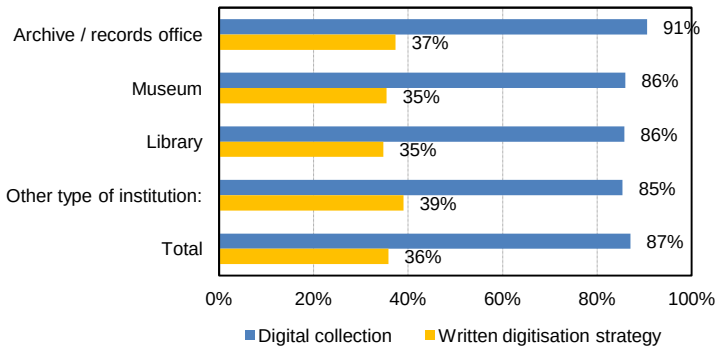


3.2 Digital collection

Most institutions, 87%, have a digital collection. This is a slight increase compared to core survey 1 with 83% having a digital collection.

Only 36% of the institutions have a written digitisation strategy. This is in fact almost the same result as *Core Survey 1* (34%):

Figure 3.2: ***Digital Collection and Written digitisation strategy*** (n=1.262)



3.3 Born digital collections

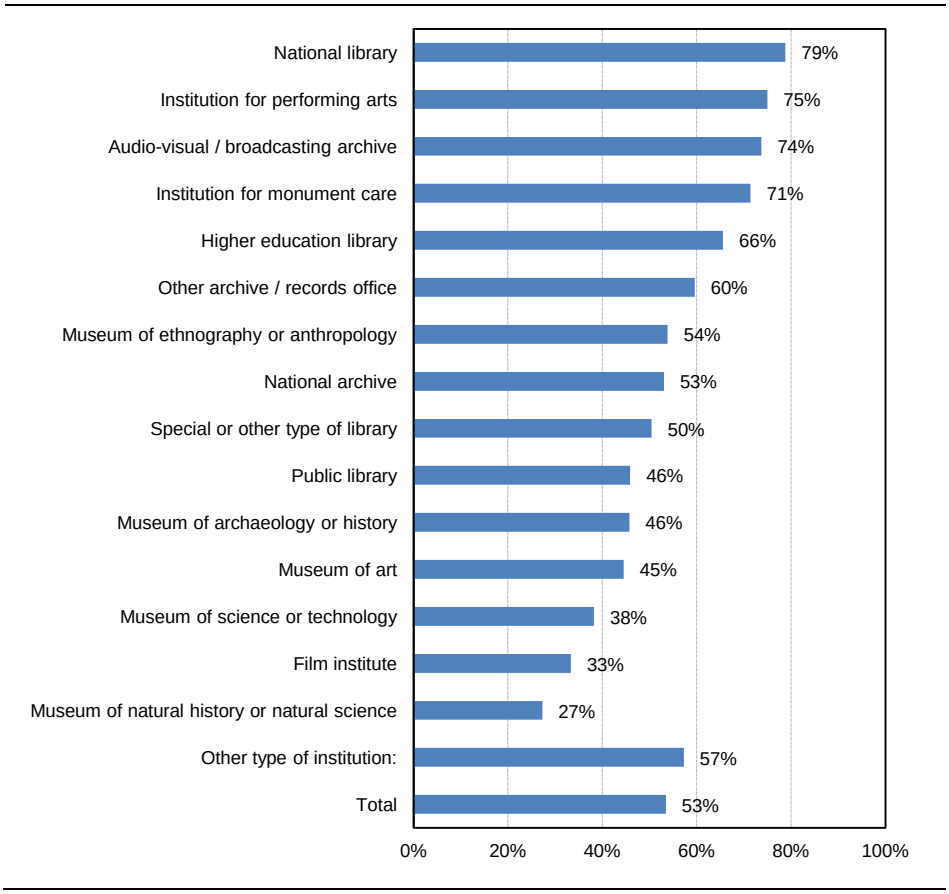
A born digital object is an object created in a digital way. The definition used in the survey is:

> "*Digital materials which are not intended to have an analogue equivalent, either as the originating source or as a result of conversion to analogue form.*"

Examples are digital images but also video, sound, digital art, games or websites. As to be expected audio visual, broadcasting and film institutions often have born digital material.

But also most national libraries and performing arts institutions have born digital material in the collections. On average 53% of the institutions have born digital items. *Core Survey 1* showed a similar result of 52%.

Figure 3.3: ***Does your organisation collect born digital heritage?*** (n=1.262)



4 The Collection

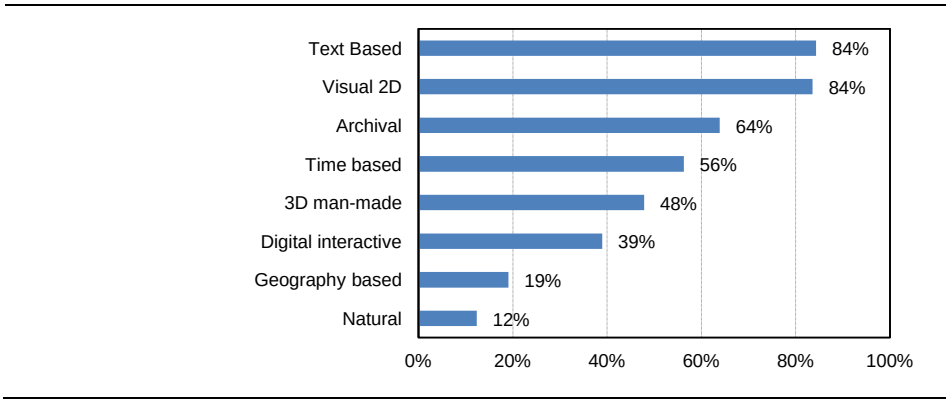
4.1 Object types

Core Survey 2 asked for indicators about physical collection management, mostly to have contextual information for the survey findings.

One of the first questions was to give information on the object categories in the collections. In this way, we were able to accommodate respondents by presenting only questions and categories that are relevant to them. Based on earlier research we classified heritage materials into seven main object types (classes). These are presented below in more detail.

Text based and Visual 2D resources are in the collections of 84% of the institutions. Archival records (64%) and time based resources (sound, film, etc.) are included in 56% of the collections. The 39% digital interactive cannot be compared with the Core Survey 1 results which asked for born digital material which is not the same: not all born digital material is interactive.

Figure 4.1: ***Object types that are part of the heritage collections of the institutions*** (n=1.252)

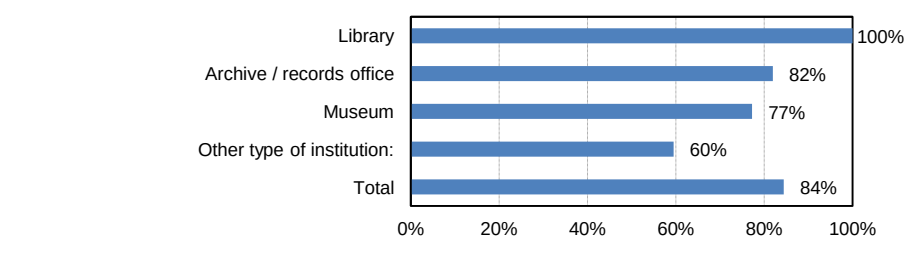


A noteworthy finding from *Core Survey 1* was that most institutions have a rich mix of heritage materials. This was confirmed in *Core Survey 2*.

4.2 Text based resources

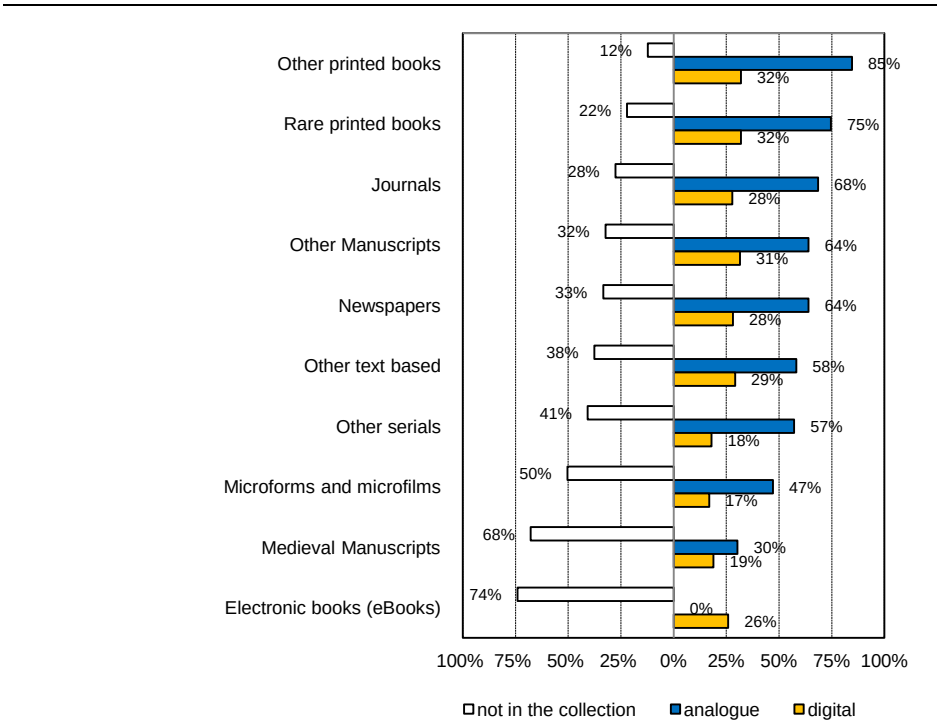
Of course text based resources can be found in every library and in most of the archives.

Figure 4.2: ***Text base resources per institution type*** (n=1.056)



10 subtypes are identified in the category text based resources. In **Figure 4.3** the bar chart shows the percentage of institutions having one or more of the 10 subtypes in their collection. The base of these percentages is of course only the group of institutions which have text based resources. This is indicated by n=1.056.

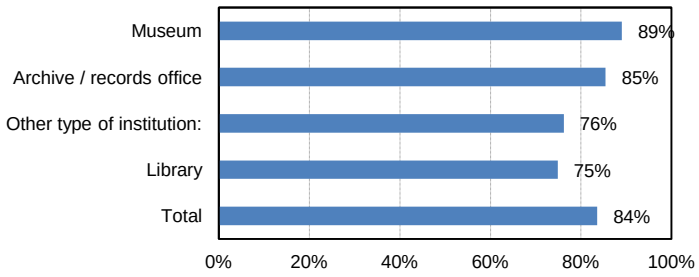
Figure 4.3: ***Text based resources (Rare printed books, other printed books, eBooks, and Newspapers)*** (n=1.056)



4.3 Visual 2D resources

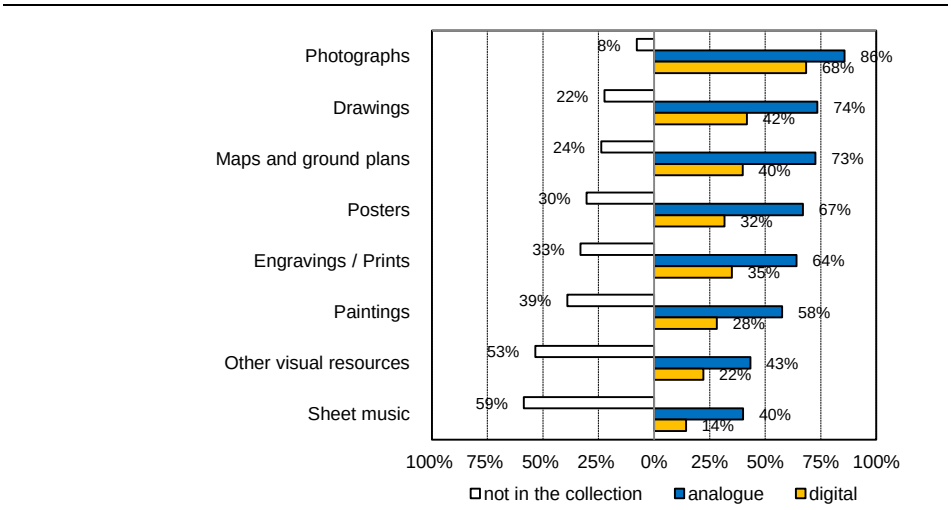
Figure 4.4 shows the degree to which visual 2D resources can be found with the types of institutions.

Figure 4.4: ***Visual 2D resources per institution type*** (n=1.046)



1046 institutions have visual 2D resources and up to 8 subtypes are included in the questionnaire.

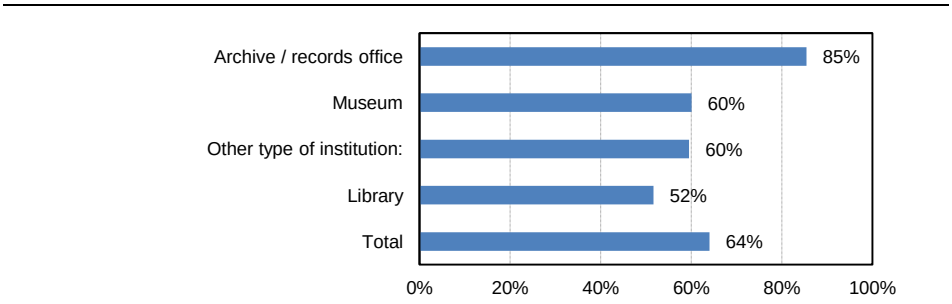
Figure 4.5: ***Visual 2D resources (Drawings, Engravings / Prints, Maps and ground plans, Paintings, and Photographs)*** (n=1.047)



4.4 Archival resources

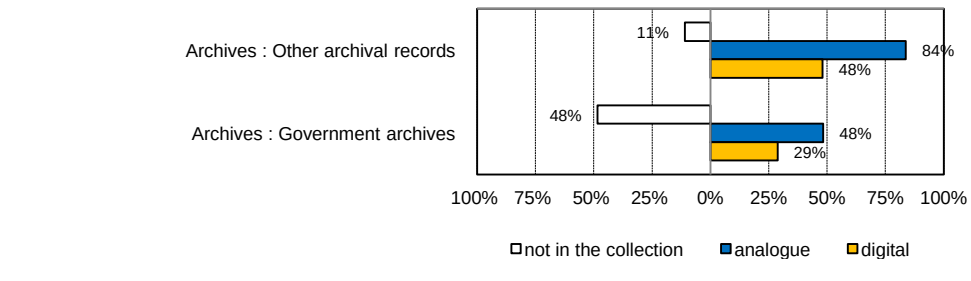
Of course archival resources are highest in archives / record offices.

Figure 4.6: ***Archival resources per institution type*** (n=801)



801 institutions collect archival records.

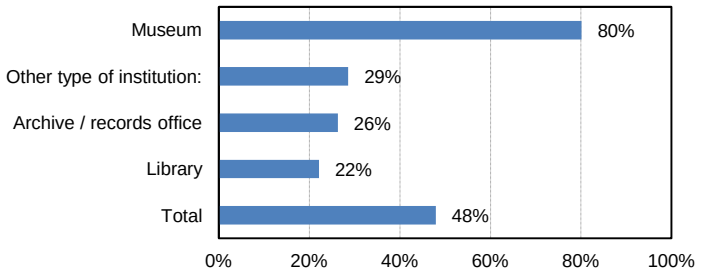
Figure 4.7: ***Archival records*** (n=801)



4.5 3D man-made resources

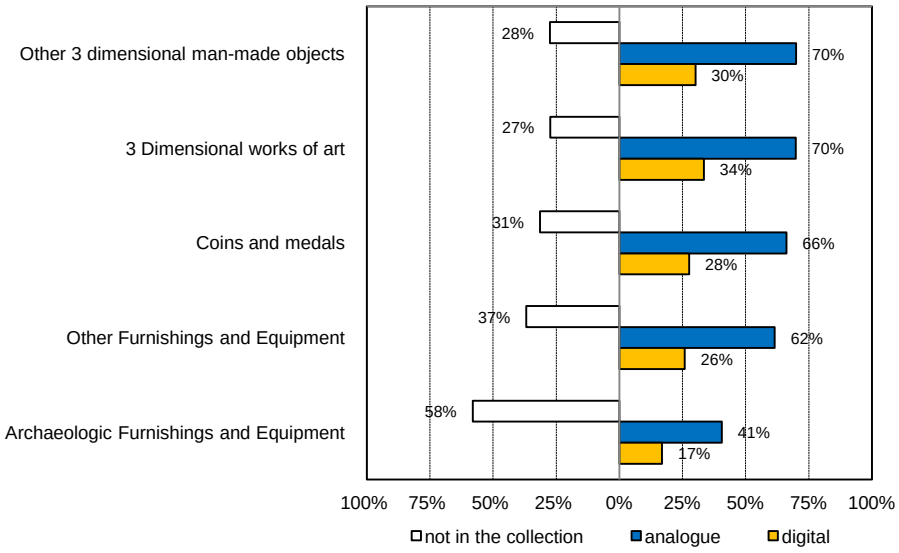
3D man-made resources can be found most with museums.

Figure 4.8: *3D man-made resources per institution type* (n=599)



599 institutions have 3D man-made objects in their current collection. These objects are classified in 5 sub types. 3 dimensional works of art and other man-made 3D objects are included in about 70% of the collections.

Figure 4.9: *3D man-made movable object (3 Dimensional works of art, Archaeological Furnishings and Equipment)* (n=599)



4.6 Natural and geography based resources

Natural and geography based resources can be found most in museums.

Figure 4.10: ***Natural and geography resources per institution type*** (n=155)

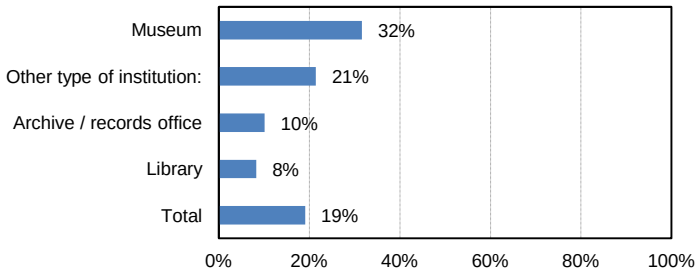
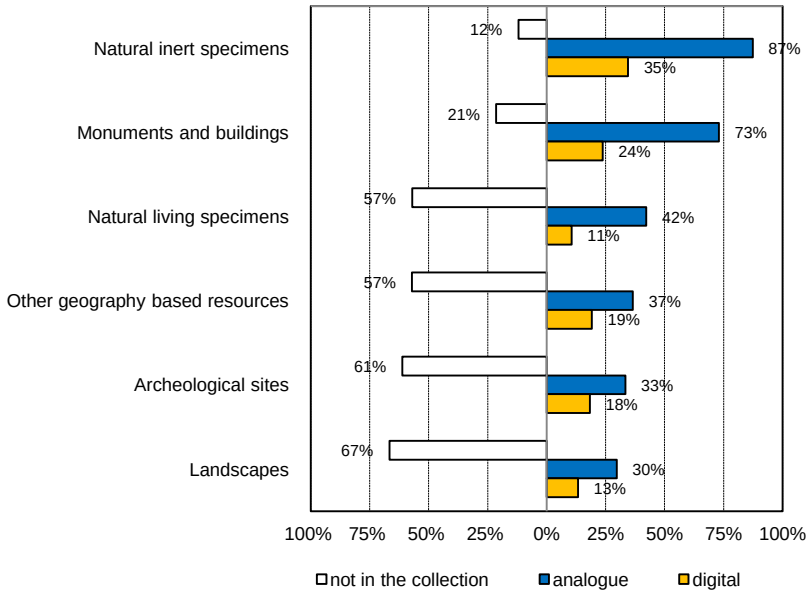


Figure 4.11: ***Natural*** (n=155) ***and geography based resources (Monuments and buildings, Landscapes, Archaeological sites)*** (n=239)



4.7 Time based resources

Time base resources are found most in the other category, which also includes audio-visual institutes and film institutes.

Figure 4.12: ***Time based resources per institution type*** (n=704)

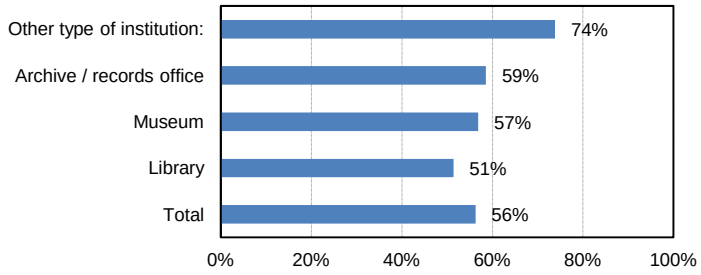
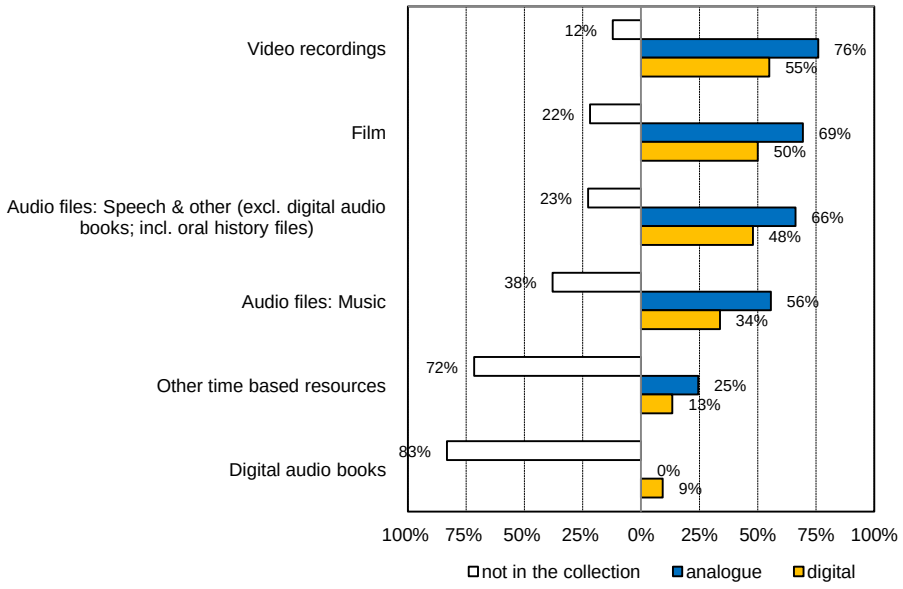


Figure 4.13: ***Time based (Audio files: Music, Audio files: Speech & other (excluding digital audio books)*** (n=704)



4.8 Born digital interactive resources

Born digital interactive resources are found with all types of institutes, most often in the other category.

Figure 4.14: ***Born digital interactive resources per institution type*** (n=488)

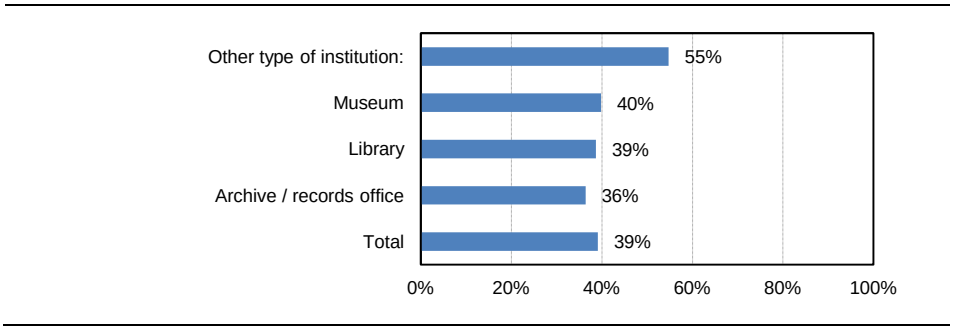
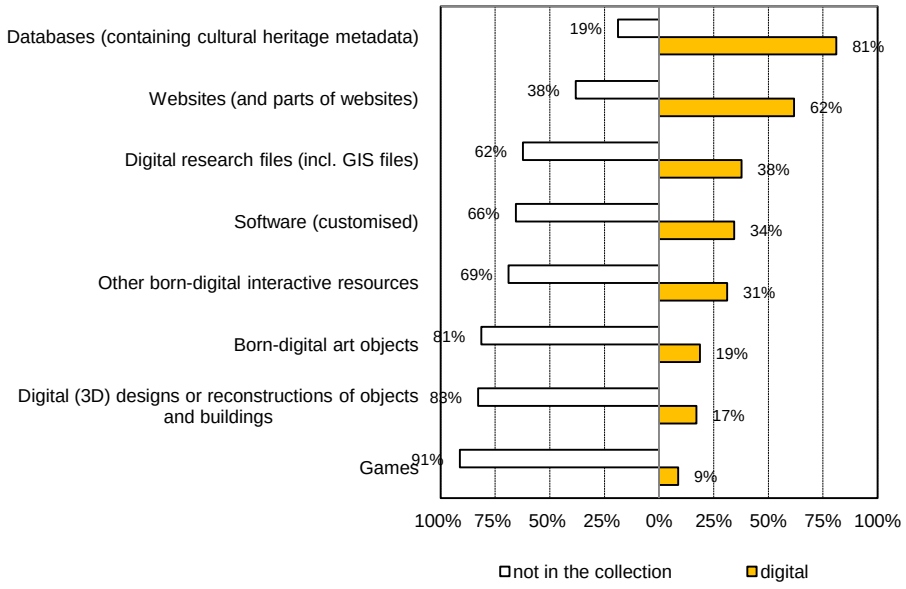


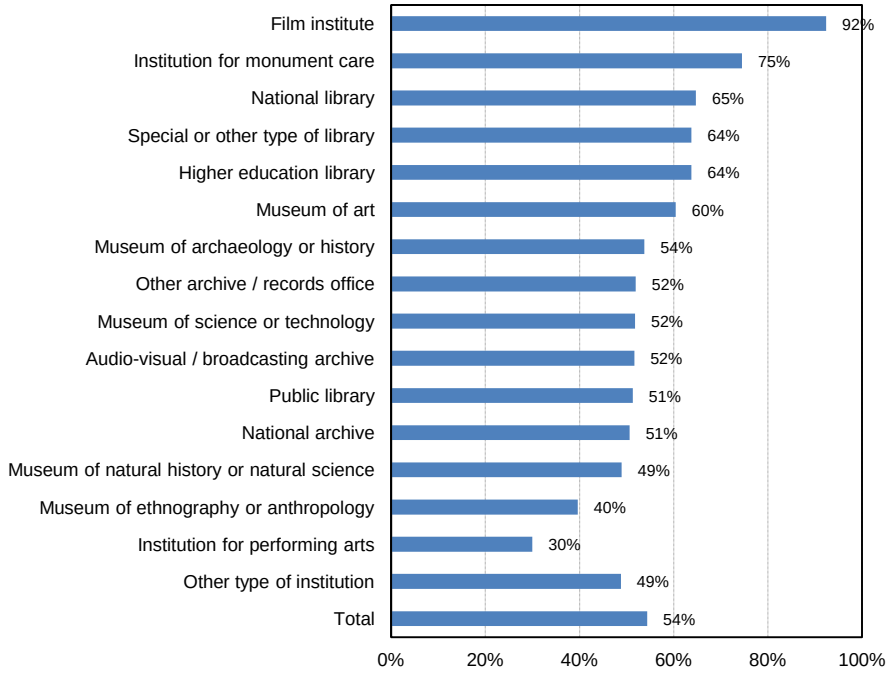
Figure 4.15: ***Born digital interactive resources*** (n=488)



4.9 Digital cataloguing

On average 54% of the collections are catalogued in a collections database.

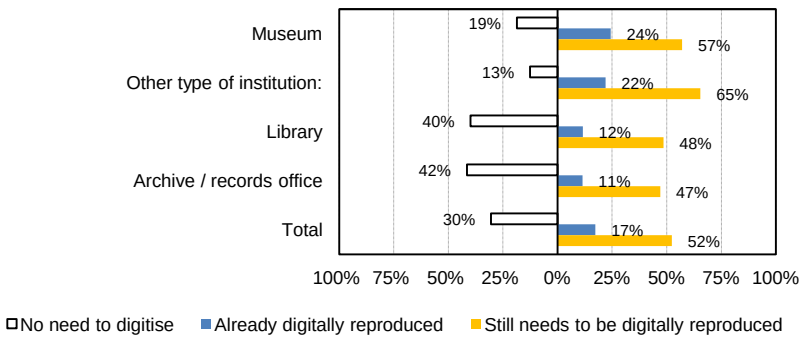
Figure 4.16: ***Estimate the percentage of your entire heritage collections that has been catalogued in a collection database*** (n=1.179)



4.10 Digitisation progress

17% of the collection is on average currently digitised and 52% still needs to be digitised. Please note that this includes not weighted averages. Institutions with small collections count for the same weight as institutions with large collections. *Core Survey 1* showed a result in the same range: 20% digitised and 57% to be digitised. The actual percentage of the digitisation level of all cultural heritage in Europe will be smaller. The exact European heritage digitisation percentage cannot be calculated based on the current information.

Figure 4.17: Estimated percentage of your analogue heritage collections that has already been digitally reproduced and still needs to be reproduced (n=1.179)



4.11 Optional questions: size of the collection

The survey included an optional section to get into more detail on the size of the collections. Answering these questions is only possible if a good insight is available in the size of the analogue collection, for example by means of a comprehensive catalogue. On top of that for many respondents it is quite a bit of work to provide the information. That is why these questions were included in an 'optional' section. 390 institutions took up the challenge to participate in this optional section.

Only the relevant object types are shown in the survey. The answers given in the survey (see: **Figure 4.1**) are used to select these categories.

Text based items

137 institutions with text based items answered the questions on size and digitisation progress. These institutions were:

- Archive/records offices: 27
- Museums: 48
- Libraries: 57
- Other institutions: 5

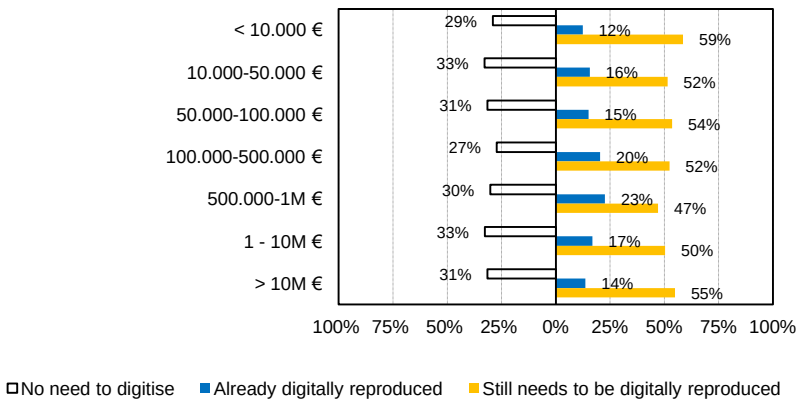
The size is shown in the table.

Table 1: ***Text based items average numbers in collections*** (n=137)

	Records in collection database(s) on item level	**Items in heritage collection**	**Born digital items**
Rare printed books	37.243	40.736 (volumes)	
Other printed books	286.371	321.657 (volumes)	
Electronic books (eBooks)	6.085	8.407 (volumes)	5.892
Newspapers	130.207	84.252 (issues)	3.172
Journals	14.846	32.302 (issues)	1.350
Other serials	7.075	7.458 (issues)	14
Medieval Manuscripts	10.371	38.930 (items)	
Other Manuscripts	31.893	286.362 (items)	1.054
Microforms and microfilms	27.409	49.994 (items)	
Other text based	25.565	547.331 (items)	55

The size of the institutions, as measured on the basis of the total annual budget, makes quite an impact on the progress made so far. For all types of institutions the group with an annual budget in the range of €100.000 – 1 million, shows the highest percentage of digitisation.

Figure 4.18: Estimated percentage of your analogue heritage collections that has already been digitally reproduced and still needs to be reproduced, compared by institution size based on total annual budget (n=1.179)



Visual 2D resources

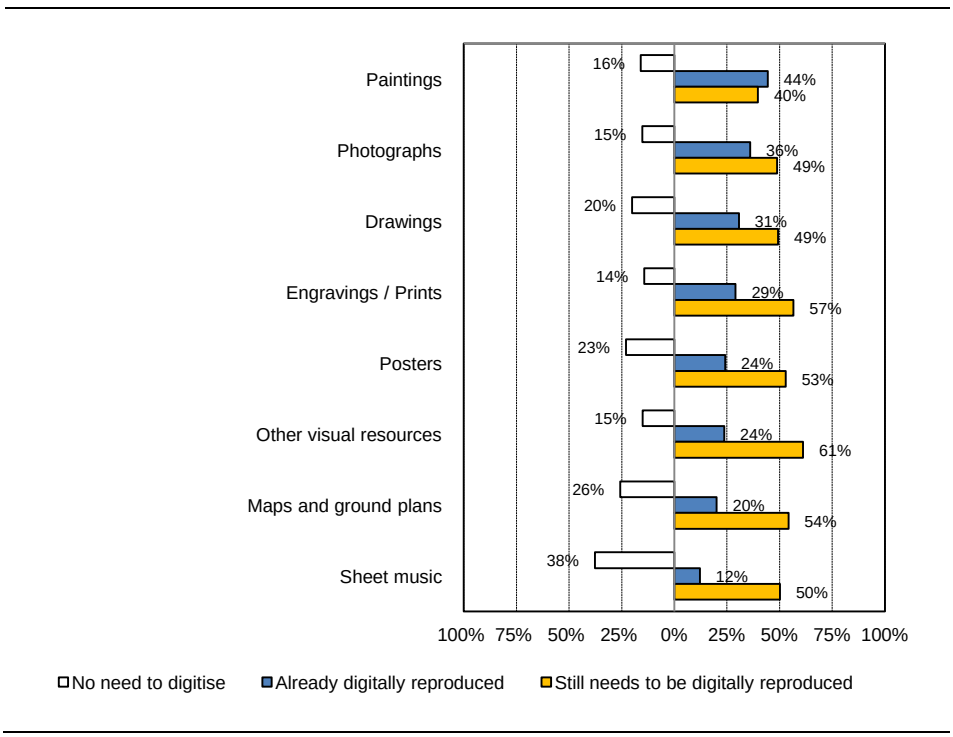
139 institutions with visual 2D based items answered the questions on size and digitisation progress. These institutions were:

- Archive/record offices: 39
- Museums: 57
- Libraries: 38
- Other institutions: 5

Table 2: ***Visual 2D resources – average numbers in collections*** (n=139)

	Records in collection database(s) on item level	**Items in heritage collection**	**Born digital items**
Drawings	3.404	10.728 (items)	87
Engravings / Prints	7.529	13.744 (items)	
Maps and ground plans	37.056	30.935 (items)	385
Paintings	757	759 (items)	3
Photographs	205.005	375.260 (items)	29.294
Posters	7.275	9.674 (items)	50
Sheet music	16.887	23.893 (items)	960
Other visual resources	2.252	6.154 (items)	50

Figure 4.19: Visual 2D resources – digitisation progress (n=137)



Archival resources

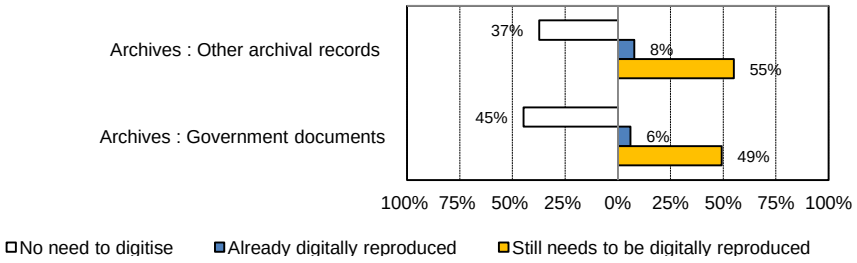
64 institutions with archival resources answered the questions on size and digitization progress. These were:

- Archives/records offices: 30
- Museums: 21
- Libraries: 11
- Other institutions: 2

Table 3: ***Archival resources – average numbers in collections*** (n=64)

	Records in collection database(s) on item level	**Items in heritage collection**	**Born digital items**
Government documents	271.123	23.165	426.166
Other archival records	16.901	6.758	21.756

Figure 4.20: ***Archival resources – digitisation progress*** (n=64)



3D man-made resources

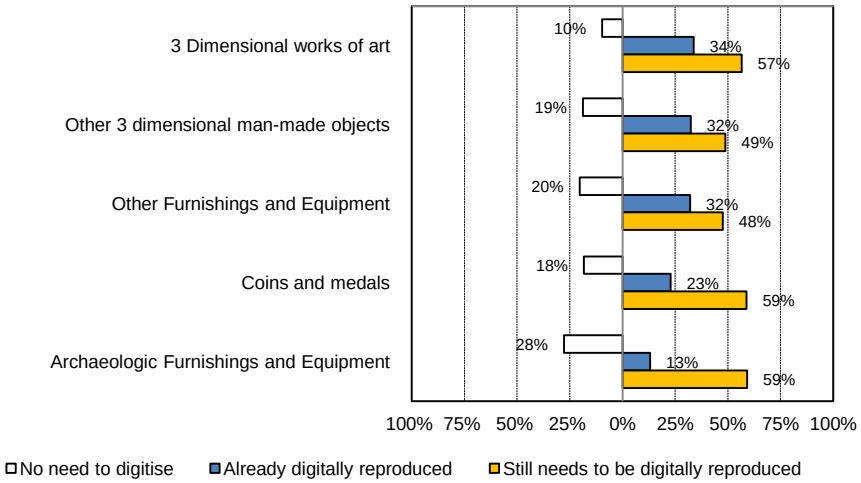
54 institutions with 3D man-made items answered the questions on size and digitization progress. These were:

- Archives/records offices: 4
- Museums: 44
- Libraries: 6
- Other institutions: 0

Table 4: *3D man-made resources – average numbers in collections* (n=54)

	Records in collection database(s) on item level	**Items in heritage collection**
3 Dimensional works of art	7.235	8.193 (items)
Archaeological Furnishings and Equipment	60.476	111.536 (items)
Other Furnishings and Equipment	4.782	5.481 (items)
Coins and medals	9.194	17.423 (items)
Other 3 dimensional man-made objects	98.216	93.719 (items)

Figure 4.21: *3D man-made resources – digitisation progress* (n=54)



Natural and geography based resources

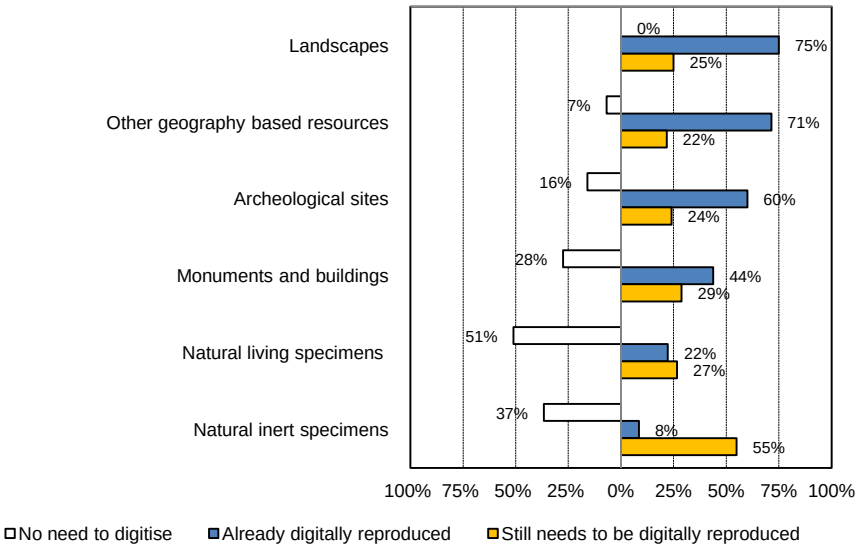
22 institutions with natural and/or geography based items answered the questions on size and digitisation progress. These were:

- Archives/records offices: 1
- Museums: 17
- Libraries: 1
- Other institutions: 3

Table 5: ***Natural and Geography based resources items – average numbers in collections*** (n=22)

	Records in collection database(s) on item level	**Items in heritage collection**
Natural inert specimens	62.263	141.9299 (items)
Natural living specimens	120.657	2.815.705 (items)
Monuments and buildings	1.636	237 (items)
Landscapes	35	2 (landscapes)
Archaeological sites	482	3 (sites)
Other geography based resources	12.587	19.887 (items)

Figure 4.22: ***Natural and Geography based resources – digitisation progress*** (n=22)



Time based resources

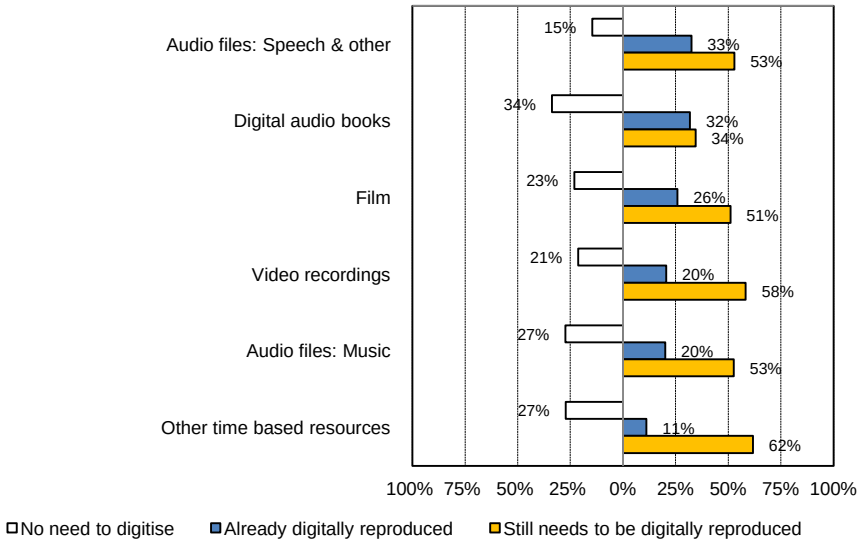
86 institutions with time-based items answered the questions on size and digitization progress. These were:

- Archives/records offices: 22
- Museums: 28
- Libraries: 29
- Other institutions: 7

Table 6: ***Time based resources – average numbers in collections*** (n=86)

	Records in collection database(s) on item level	***In heritage collection***	**Born digital items**
Audio files: Music	157.165	6.453 (hours)	2368
Audio files: Speech & other records	87.795	2.516 (hours)	853
Digital audio books	867	1.164 (volumes)	544
Film	13.287	3.986 (hours)	300
Video recordings	392.656	6.220 (hours)	4369
Other time based resources	467	33 (items)	1306

Figure 4.23: ***Time based resources – digitisation progress*** (n=86)



Born digital resources

64 institutions with born digital resources answered the questions on size and digitization progress. These were:

- Archives/records offices: 18
- Museums: 23
- Libraries: 17
- Other institutions: 6

Table 7: ***Born digital interactive resources – average numbers in collections*** (n=64)

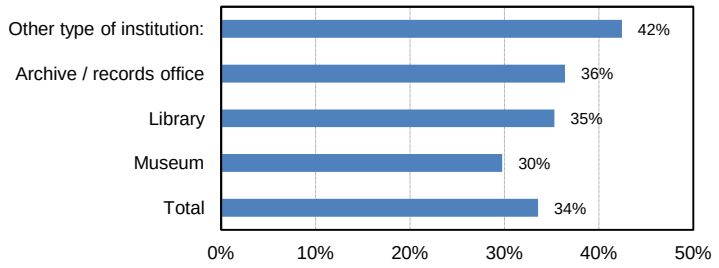
	Records in collection database(s) on item level	**Born digital items**
Digital (3D) designs or reconstructions of objects and buildings	259	418 (items)
Born-digital art objects	74	79 (items)
Digital research files (including GIS files)	8.536	11.057 (items)
Games	12	100 (items)
Software (customised)	81	1.418 (items)
Other born-digital interactive resources	17	345 (items)
Websites (and parts of websites)	42.334	40.938 (items)

5 Digital access

5.1 Policy on use of digital collections

Core Survey 1 showed that 31% of the institutions have an explicit policy regarding the use of the digital collections. This survey shows a slightly higher percentage of 34%.

Figure 5.1: ***Does your organisation have an explicit (written) policy, endorsed by the management of your organisation, that sets conditions for specific types of use of your digital heritage collections*** (n=905)

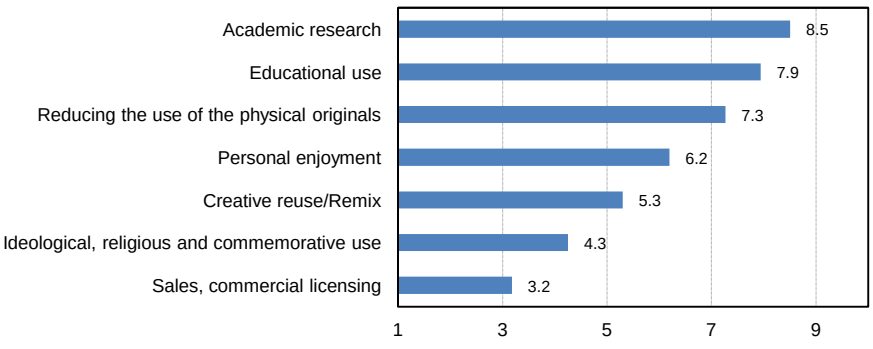


5.2 Reasons for providing digital access

Institutions were asked to indicate the importance of a number of reasons to provide digital access to the collection. Answers are given on a 10-point scale ranging from 1 = *not important* to 10 = *very important*.

Academic research is perceived as the most important reason to provide digital access to the collection. Second is the educational use of the collection. Sales and commercial licensing is the least important reason.

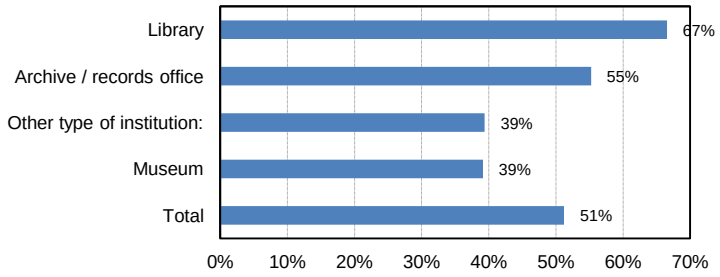
Figure 5.2: ***Collections are made accessible to the public for various reasons. How important is each of the following types of use for your institution?*** (n=906)



5.3 Measuring the use of digital collections

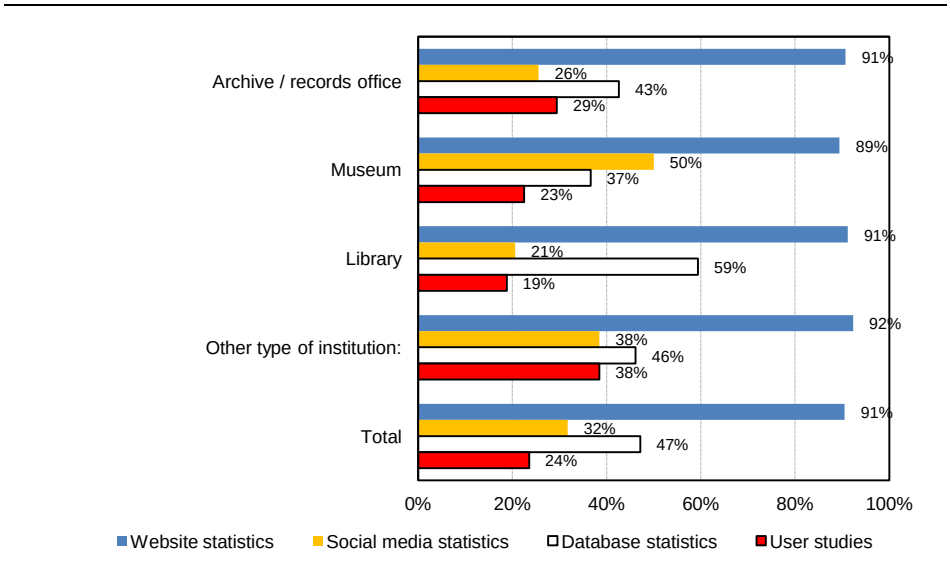
On average 42% of all institutions monitored the use of the collection in the previous core survey. This survey shows a higher percentage of 51%. The list of institution types is not the same in the last survey, so on an institution type level we cannot compare the results accurately. We can however assume that in this area progress has been made.

Figure 5.3: ***Does your organisation measure the number of times that digital metadata and/or digital objects are being accessed by your users?*** (n=902)



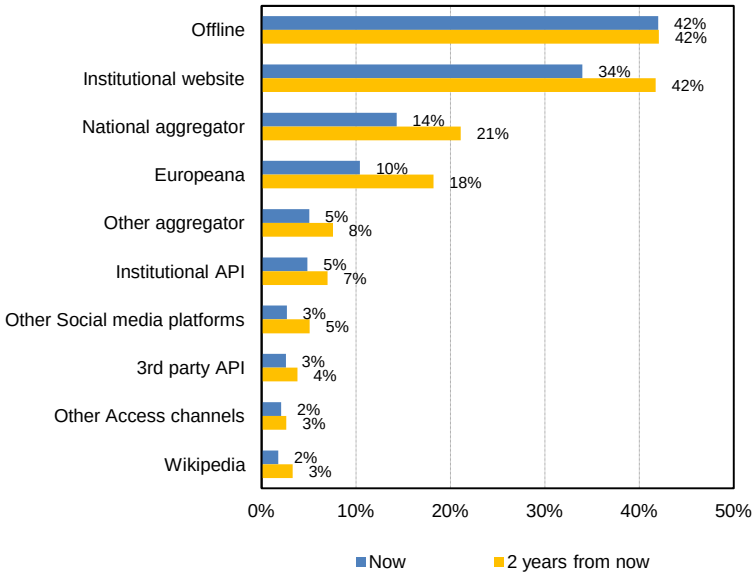
Website statistics is a widely used way to measure the use of the digital collections. 91% of the institutions that measure digital access use web statistics. Core survey 1 resulted in a use of web statistics by 85%. Database statistics are used by 47% (Core Survey 1: 36%) and user studies account for 24% (was: 16%). In the second Core Survey, we also asked for measuring the use of social media statistics: 32% of the institutions make use of these statistics. In Core Survey 1, this option was not included.

Figure 5.4: ***How is the use of the digital collection measured?*** (n=462)



A large number of digital access channels are available. Offline access and the institutional website are widely used. An increase in the future is expected from the institutional website, the national aggregator and Europeana. In *Core Survey 1* all these percentages were almost the same or a bit higher: for instance for the institutional website the access was 31% (with an expectation of growth to 47%).

Figure 5.5: Estimated percentage of all the digital objects that are and/or will be accessible by which form (n=905)



6 Digital access

6.1 Digital preservation strategy

26,1% of the institutions have a written digital preservation strategy that is endorsed by the management. In *Core Survey 1* 23% claimed to have a written strategy (the question has been made more specific in *Core Survey 2*, by adding the last part which makes the question as follows: '*Does your organisation have a written digitisation strategy, endorsed by the management of your organisation*'?)

Figure 6.1: *Does your organisation have a written Digital Preservation Strategy, which is endorsed by the management of your organisation?* (n=905)

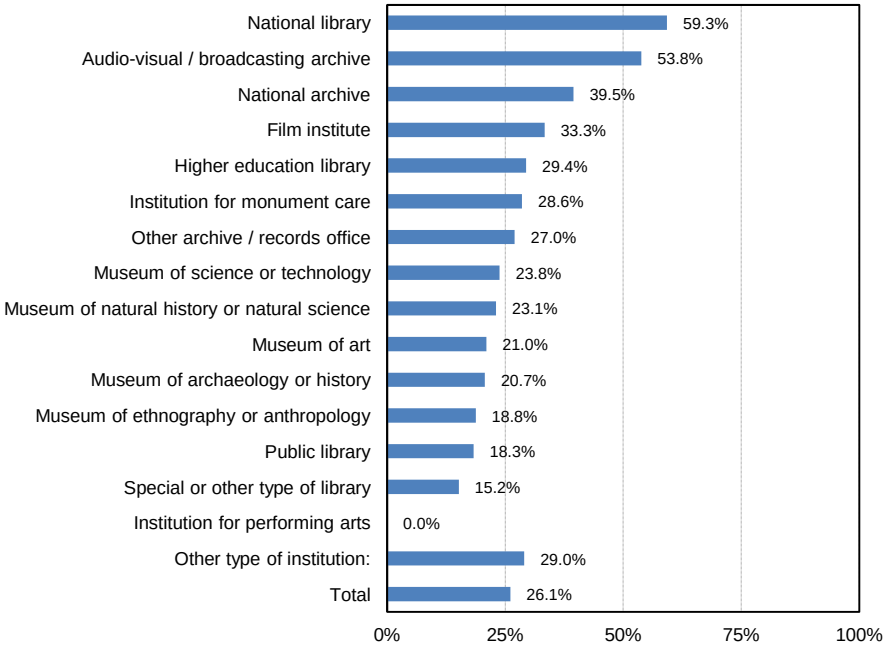
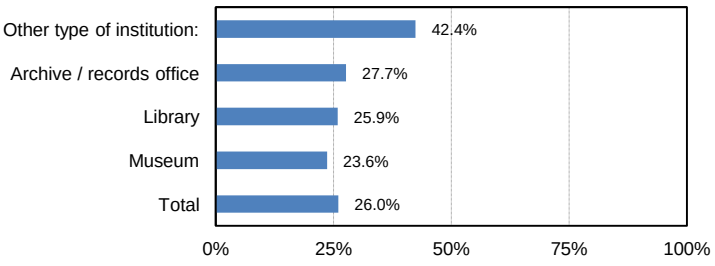


Figure 6.2: *Does your organisation have a written Digital Preservation Strategy, which is endorsed by the management of your organisation? Per main institution type* (n=905)

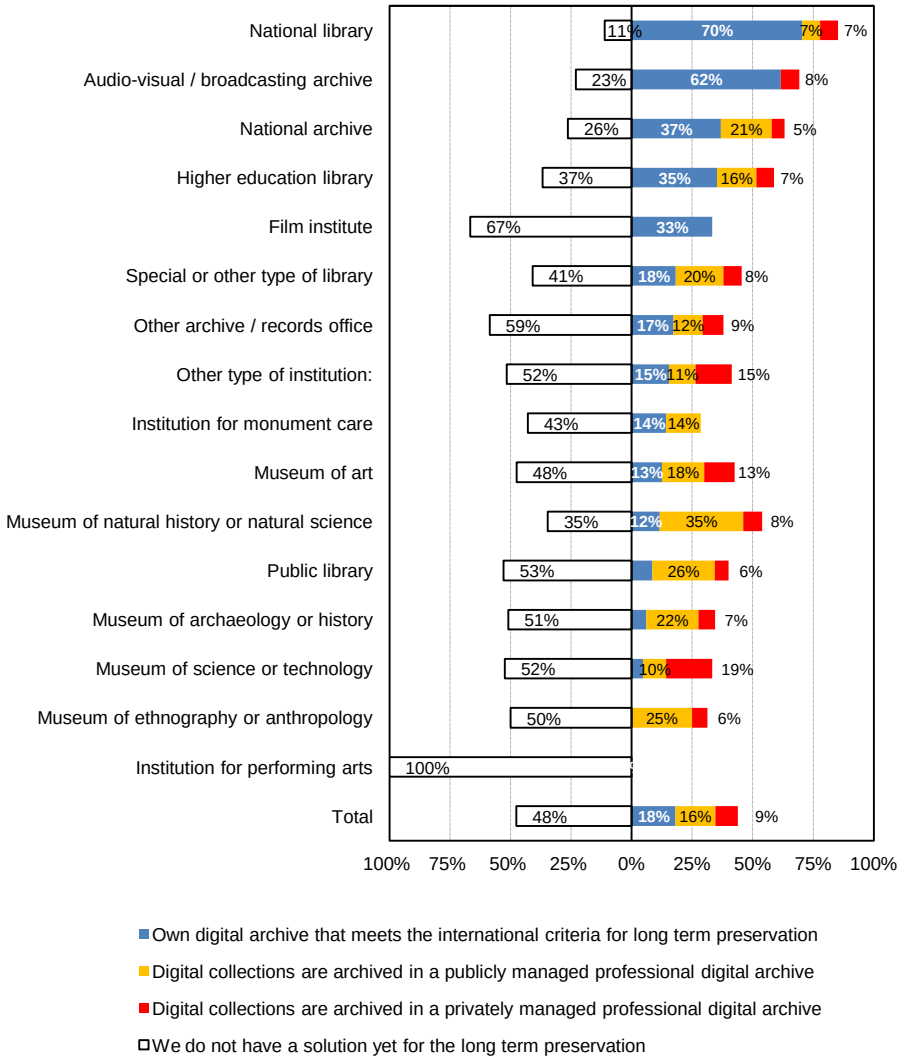


6.2 Use of international standards for digital preservation

About 48% of the respondents do not have a solution yet for long term preservation based on international standards for digital preservation. This percentage might be even a bit higher as about 5% did not know the answer to this question.

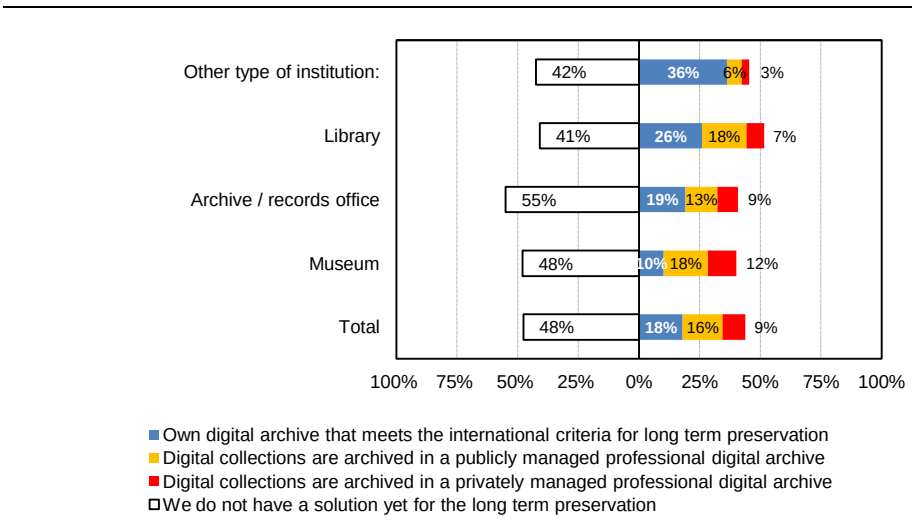
Large differences are visible if the institution types are compared. National libraries are clear 'front runners', the performing arts institutes have still a long way to go.

Figure 6.3: Are your digital collections stored in digital archives that have been set up according to international standards for digital preservation? (n=905)



The percentages in each row do not count to 100% because of the category 'do not know' which is not presented in the figure.

Figure 6.4: ***Are your digital collections stored in digital archives that have been set up according to international standards for digital preservation? Per main institution type*** (n=905)



The percentages in each row do not count to 100% because of the category 'do not know' which is not presented in the figure.

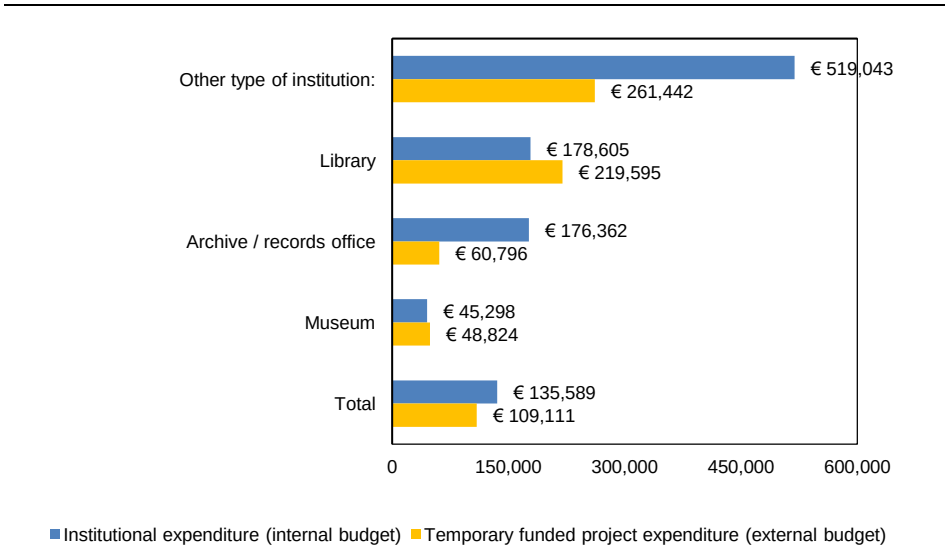
7 Digitisation Expenditure

7.1 Internal and external budgets

Costs made on an annual basis on creating, acquiring, maintaining, enhancing and preserving the digital collections can be funded by internal budgets and external budgets. On average the costs made are quite substantial summing up to €245.000. The median for the total costs is much lower: €15.600.

This sum is an estimate of all the costs related to the initial creation, ongoing maintenance, enhancement and preservation of the digital collections. The cost of the staff time devoted to such activities is included in this estimate.

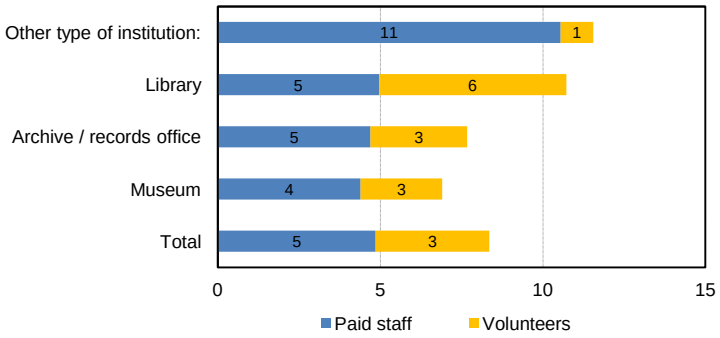
Figure 7.1: ***Estimate your annual expenditure on creating/acquiring, maintaining, enhancing and preserving your digital collections*** (n=839)



7.2 Internal and external budgets

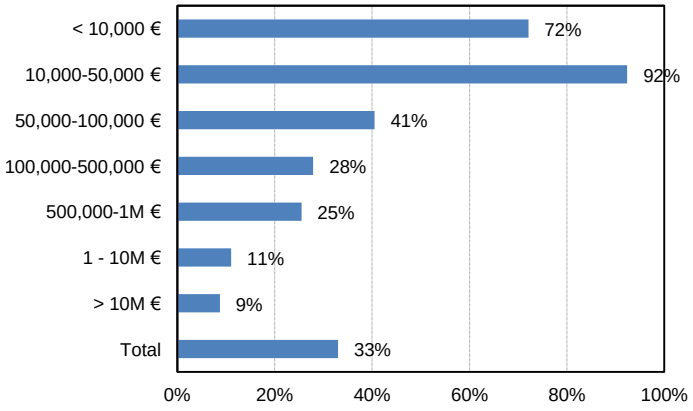
An average of 8 people are involved on a full-time basis in the digital collection activities. 5 people are paid staff and 3 are volunteers. This is about 50% more staff compared to the *Core Survey 1* results (average total: 3,5 full-time equivalent (FTE) paid staff and 2,0 FTE volunteers).

Figure 7.2: What is the total number of paid staff and volunteers (in full-time equivalent) engaged in creating/acquiring, maintaining, enhancing and preserving your digital collections on an annual basis? (n=793)



We analysed the fraction of the total paid staff involved in the digitisation activities. In total 33% of the paid staff is involved in digitisation activities. In small institutions with small budgets this percentage is much higher.

Figure 7.3: *Fraction of the total paid staff (in full-time equivalent) engaged in creating/acquiring, maintaining, enhancing and preserving your digital collections on an annual basis?* (n=793)

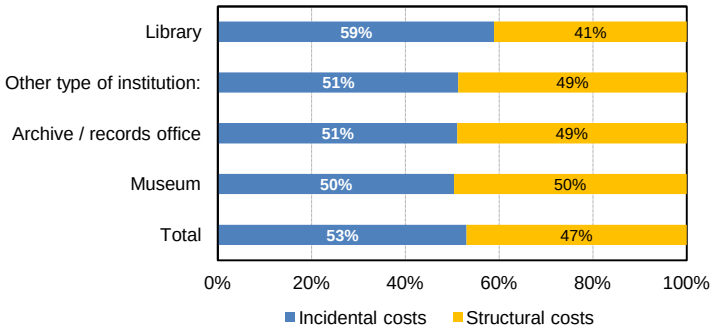


7.3 Incidental and structural costs

About 53% of the costs are qualified as being incidental cost and the remainder are structural costs.

- **Incidental costs** are defined as the costs involved with the initial creation or acquisition of a digital collection. Examples: selection of materials, acquisition of digital born materials, scanning, descriptive metadata creation, project management.
- **Structural costs** are the costs needed for the ongoing maintenance, enhancement and preservation of a digital collection. Examples: activities concerning the preservation of digital collections, licences, maintenance of web servers, user outreach and support, management.

Figure 7.4: ***Estimation of the percentage of the total annual expenditures on creating/acquiring, maintaining, enhancing and preserving digital collections. Incidental costs compared to structural costs.*** (n=814)

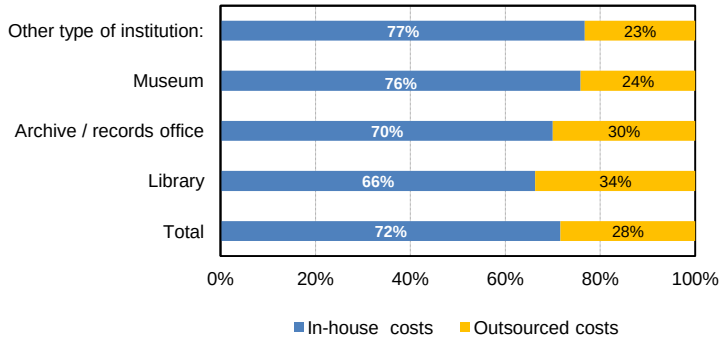


This result is different from that in the more qualitative *Thematic Survey* (in which the question was included for the first time). In the *Thematic Survey* (n=33) the figures were: 79% incidental costs and 21% structural costs (Core Survey 2: 53% incidental and 47% structural).

7.4 In-house costs and outsourced costs

72% of the costs are spent internally, 28% are out-of-pocket costs for external suppliers. These results are the same as in the *Thematic Survey*.

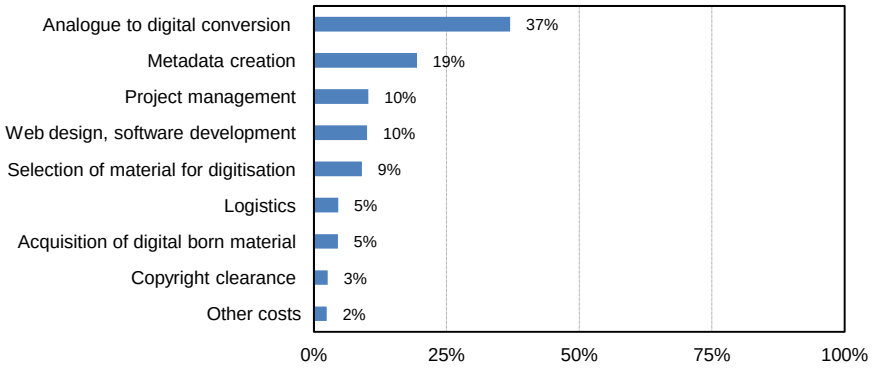
Figure 7.5: *Estimation of the percentage of the total annual expenditure on creating/acquiring, maintaining, enhancing and preserving digital collections. In-house costs compared to outsourced costs* (n=814)



7.5 Activities

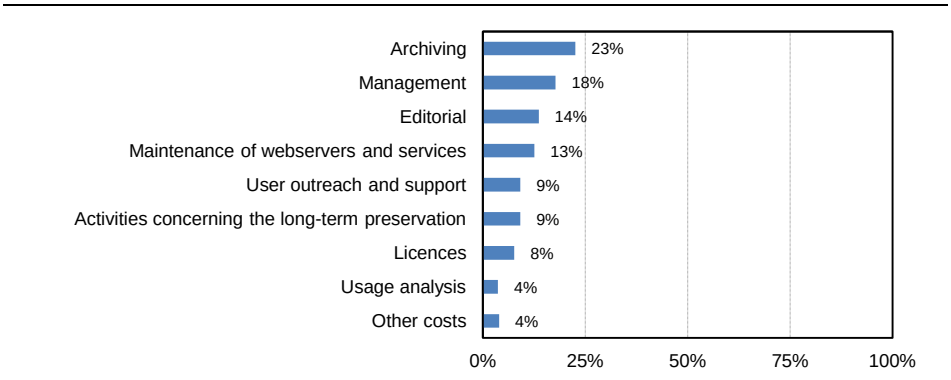
The analogue to digital conversion process is responsible for 37% of the incidental costs. Metadata creation is a substantial category too, accounting for 19% of the incidental costs.

Figure 7.6: Estimation of what percentages of the Incidental costs can be assigned to the following activities (n=185)



Archiving (23%) and management (18%) are the main structural cost components. Editorial work accounts for 14% and technology (servers) for 13%.

Figure 7.7: ***Estimation of what percentages of the Structural costs can be assigned to the following activities*** (n=185)

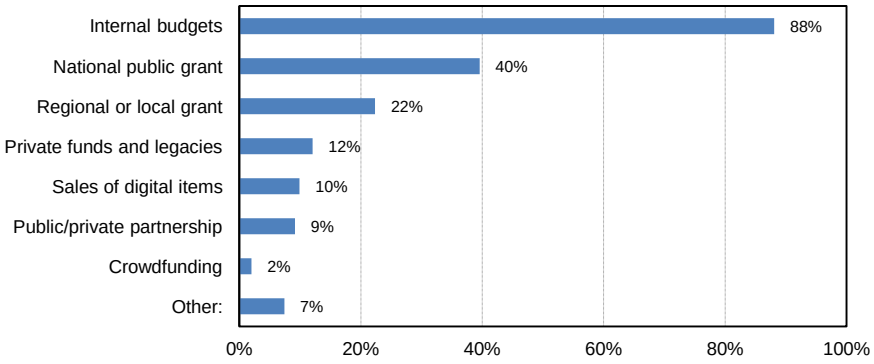


In the *Thematic Survey* structural costs were most of all assigned to archiving (storage, including backups) (n=18), and then to giving access to the digital collection (n=13), project management (n=8), user outreach and support (n=7), editorial (n=6) and activities concerning the long-term preservation (n=6).

7.6 Funding

Most institutions use internal budget for digital collection activities. National public grants are available to almost 40% of the respondents. 22% receives regional or local public grants. The answers to this question did not indicate the amount of money in the funding. They only refer to the different sources. The results are very similar to *Core Survey 1*. In the *Core Survey 2* a number of sources are added to the list. Crowd funding is a new source in this field and 2% of the institutions make use of it.

Figure 7.8: ***From what sources are your digital collection activities funded?*** (n=794)



Appendix – *Core Survey 2* Questionnaire

SECTION 1: Organisational Information

1. Name of institution / organisation

[Input box]

This information was not published in the report.

2. Type / Domain of institution / organisation

Specify the primary heading you would assign to your institution.

Please choose **only one** of the following:

- National archive
- Other archive / records office
- Audio-visual / broadcasting archive
- Film institute
- Institution for performing arts
- Museum of art
- Museum of archaeology or history
- Museum of natural history or natural science
- Museum of science or technology
- Museum of ethnography or anthropology
- National library
- Higher education library
- Public library
- Special or other type of library
- Institution for monument care
- Other (specify below)

Other type of institution: [Input box]

3. Country in which your institution is located

Please choose **only one** of the following:

- Albania
- Andorra
- Armenia
- Austria
- Azerbaijan
- Belgium
- Bosnia and Herzegovina
- Bulgaria
- Croatia
- Cyprus
- Czech Republic
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Iceland
- Ireland

- Italy
- Latvia
- Liechtenstein
- Lithuania
- Luxembourg
- Malta
- Republic of Moldova
- Monaco
- Montenegro
- Netherlands
- Norway
- Poland
- Portugal
- Romania
- Russian Federation
- San Marino
- Serbia
- Slovak Republic
- Slovenia
- Spain
- Sweden
- Switzerland
- Turkey
- Ukraine
- United Kingdom
- Former Yugoslav Republic of Macedonia
- Other country (specify below)

Other country: [Input box]

4. *Website of your institution* [Input box]
Provide the address of your institution's main website that is accessible for the general public.

This information was not published in the report.

5. *Your name*

[Input box]

This information was not published in the report.

6. *Your job title*

[Input box]
The role or position of the main person completing this survey.

7. *Your e-mail address*

[Input box]

This information was not published in the report.

8. Your telephone number and/or Skype contact details

[Input box]
The primary phone number (i.e. +44 0123456789) and/or the Skype details of the main person completing this survey.

This information was not published in the report.

9. What is your institution's total annual budget?

Please choose **only one** of the following:

- < 10,000 €
- 10,000-50,000 €
- 50,000-100,000 €
- 100,000-500,000 €
- 500,000-1M €
- 1 - 10M €
- > 10M €

Provide the annual budget for the entire cultural heritage institution as indicated in the last published annual account. If your institution is part of a larger organisation (e.g. a higher education library that is part of a higher education institution) only provide the budget of the cultural heritage related unit.

The total annual budget may include government funding, project funding, revenues from commercial activities, etc. If your budget occurs in two categories (e.g. 50,000 €), please choose the lower category.

10. Total number of paid staff (in full time equivalents, not in number of people)

[Input box] [Only 1 decimal accepted, e.g. 3.7]

The number of full time equivalents should represent the total staff employed by your institution, including permanent and temporary staff, but excluding contractors and volunteers. Part-time staff needs to be added up to represent a full working week.

Note: the number of staff engaged in digitisation activities will be asked for later in the survey (see below).

SECTION 2: Digitisation Activity

11. Does your institution have collections that need to be preserved for future generations?

[ ] Yes
[ ] No

Answer this question with 'No' if your institution does not hold heritage collections or if you only have collections (for example of books, films, music) that can be lent by or sold to users.

If the answer is No, automatically proceed to the end of the survey.

12. Does your organisation have digital collections or is it currently involved in collection digitisation activities?

[ ] Yes
[ ] No

If the answer is No, automatically proceed to questions 13, 14, 15, 16, 17, 18 and after that to the end of the survey.

13. Does your organisation have a written digitisation strategy, endorsed by the management of your organisation?

[ ] Yes
[ ] No
[ ] Do not know

The strategy may be for any period up to 2020.

14. Does your organisation collect born digital heritage?

[ ] Yes
[ ] No
[ ] Do not know

Answer this question with ´yes´ if your institution collects any kind of born digital heritage materials (i.e. software, digital documents, digital art, harvested web content, etc.) with the explicit intention of preserving these born digital materials for future generations.

15. ***Please select the collection types that are part of the heritage collections of your institution***

NOTE: Table will not be presented as it is here. The table can be folded out selectively, starting from the high level collection type classes in the left column.

Please specify the object types that are part of the heritage collections of your institution. The digital collection consists of digitally reproduced analogue objects and born digital objects. An object that has been catalogued in a database with metadata records only, is not considered to be part of the ´digital collection´.

Collection type	Object type	In *analogue* collection y/n	In *digital* collection y/n
(01) TEXT BASED RESOURCES			
	Rare printed books		
	Other printed books		
	Electronic books (eBooks)		
	Newspapers		
	Journals		
	Other serials		
	Medieval Manuscripts		
	Other Manuscripts		
	Microforms and microfilms		
	Other text based		
(02) VISUAL (2D) RESOURCES			
	Drawings		
	Engravings / Prints		
	Maps and ground plans		
	Paintings		
	Photographs		
	Posters		
	Sheet music		
	Other visual resources		
(03) ARCHIVAL RESOURCES (not included in 01 or 02)			
	Archives : Government documents		
	Archives : Government archives		
	Archives : Other archival records		
	Archives : Other archives		
(04) 3D MAN-MADE MOVABLE OBJECTS			
	3 Dimensional works of art		
	Archaeological Furnishings and Equipment		
	Other Furnishings and Equipment		
	Coins and medals		
	Other 3 dimensional man-made objects		

(05) NATURAL RESOURCES			
	Natural inert specimens		
	Natural living specimens		
(06) GEOGRAPHY BASED RESOURCES			
	Monuments and buildings		
	Landscapes		
	Archaeological sites		
	Other geography based resources		
(07) TIME BASED RESOURCES			
	Audio files: Music		
	Audio files: Speech & other (excl. digital audio books; incl. oral history files)		
	Digital audio books		
	Film		
	Video recordings		
	Other time based resources		
(08) DIGITAL INTERACTIVE RESOURCES (EXCLUSIVELY DIGITAL)			
	Databases (containing cultural heritage metadata)		
	Digital (3D) designs or reconstructions of objects and buildings		
	Born-digital art objects		
	Digital research files (incl. GIS files)		
	Games		
	Software (customised)		
	Websites (and parts of websites)		
	Other born-digital interactive resources		

Tests have indicated that answering the questions about the size of collections are rewarding, but may be challenging. Some institutions will want to be more specific than is possible here; others may find it difficult to give even the high level estimates asked for. We are convinced that all institutions will benefit from an exercise in mapping out digital collections. Please consider the OPTIONAL question below Question 18 if you want to take up the challenge!

16. ***Estimate the percentage of your entire heritage collection that has been catalogued in a collection database***

[Input box]

The estimated percentage of your entire heritage collections that has been catalogued in a collection database concerns item level descriptions (metadata records) of analogue and born-digital heritage objects.

17. ***Estimate the percentage of your analogue heritage collections that has already been digitally reproduced***

[Input box]

A digital reproduction is a digital surrogate of an original analogue object. Please note that an object that has only been catalogued in a database with metadata records is not considered to be "digitally reproduced".

18. ***Estimate the percentage of your analogue heritage collections that still needs to be digitally reproduced***

[Input box]

If it is difficult to provide global estimates of the size of your collections, please consider to map out the size of your collections per object type. The information gathered in the table below will be highly valuable both for your own institution and the international community of archives, libraries and museums.

[ ] Show table

Please indicate the size of your collections per object type and assess the need to digitise

NOTE: Table will not be presented as it is here. The table can be folded out selectively, starting from the high level collection type classes in the left column.

Collection type	(A) Number of records in your collection database(s) on item level	(B) Estimated number of analogue items in heritage collection [indicate units; see Core Survey 1]	(C) Estimated number of born digital items in heritage collection	(D) Estimated % of objects in analogue collection that is digitally reproduced	(E) Estimated % of objects in analogue collection still to be digitally reproduced	(F) Estimated % of objects in analogue collection with no need to be digitally reproduced
(01) TEXT BASED RESOURCES						
Rare printed books	… records	… volumes	not applicable	… %	… %	… %
Other printed books	… records	… volumes	not applicable	… %	… %	… %
Electronic books (eBooks)	… records	… volumes	… volumes	… %	… %	… %
Newspapers	… records	… issues	… issues	… %	… %	… %
Journals	… records	… issues	… issues	… %	… %	… %
Other serials	… records	… issues	… issues	… %	… %	… %
Medieval Manuscripts	… records	… items	not applicable	… %	… %	… %
Other Manuscripts	… records	… items	… items	… %	… %	… %
Microforms and microfilms	… records	… items	not applicable	… %	… %	… %
Other text based	… records	… items	… items	… %	… %	… %
(02) VISUAL (2D) RESOURCES						
Drawings	… records	… items	… items	… %	… %	… %
Engravings / Prints	… records	… items	not applicable	… %	… %	… %
Maps and ground plans	… records	… items	… items	… %	… %	… %
Paintings	… records	… items	… items	… %	… %	… %
Photographs	… records	… items	… items	… %	… %	… %
Posters	… records	… items	… items	… %	… %	… %

Sheet music	… records	… items	… items	… %	… %	… %
Other visual resources	… records	… items	… items	… %	… %	… %
(03) ARCHIVAL RESOURCES (not included in 01 or 02)						
Archives : Government documents	… records	… metres	… records	… %	… %	… %
Archives : Government archives	… records	… archives	… archives	… %	… %	… %
Archives : Other archival records	… records	… metres	… records	… %	… %	… %
Archives : Other archives	… records	… archives	… archives	… %	… %	… %
(04) 3D MAN-MADE MOVABLE OBJECTS						
3 Dimensional works of art	… records	… items	not applicable	… %	… %	… %
Archaeological Furnishings and Equipment	… records	… items	not applicable	… %	… %	… %
Other Furnishings and Equipment	… records	… items	not applicable	… %	… %	… %
Coins and medals	… records	… items	not applicable	… %	… %	… %
Other 3 dimensional man-made objects	… records	… items	not applicable	… %	… %	… %
(05) NATURAL RESOURCES						
Natural inert specimens	… records	… items	not applicable	… %	… %	… %
Natural living specimens	… records	… items	not applicable	… %	… %	… %
(06) GEOGRAPHY BASED RESOURCES						
Landscapes	… records	… landscapes	not applicable	… %	… %	… %
Archaeological sites	… records	… sites	not applicable	… %	… %	… %
Monuments and buildings	… records	… items	not applicable	… %	… %	… %
Other geography based resources	… records	… items	not applicable	… %	… %	… %
(07) TIME BASED RESOURCES						
Audio files: Music	… records	… hours	… hours	… %	… %	… %
Audio files: Speech & other (excl. digital audio books; incl. oral history files)	… records	… hours	… hours	… %	… %	… %

Digital audio books	… records	… volumes	… volumes	… %	… %	… %
Film	… records	… hours	… hours	… %	… %	… %
Video recordings	… records	… hours	… hours	… %	… %	… %
Other time based resources	… records	… items	… items	… %	… %	… %
(08) DIGITAL INTERACTIVE RESOURCES (EXCLUSIVELY DIGITAL)						
Databases (containing cultural heritage metadata)	… records	not applicable	… items	not applicable	not applicable	not applicable
Digital (3D) designs or reconstructions of objects and buildings	… records	not applicable	… items	not applicable	not applicable	not applicable
Born-digital art objects	… records	not applicable	… items	not applicable	not applicable	not applicable
Digital research files (incl. GIS files)	… records	not applicable	… items	not applicable	not applicable	not applicable
Games	… records	not applicable	… items	not applicable	not applicable	not applicable
Software (customised)	… records	not applicable	… items	not applicable	not applicable	not applicable
Websites (and parts of websites)	… records	not applicable	… items	not applicable	not applicable	not applicable
Other born-digital interactive resources	… records	not applicable	… items	not applicable	not applicable	not applicable

Explanatory notes for each of the 6 columns:

(A) Number of records in your collection database(s) on item level:

The (estimated) quantities of item level records in your collection database(s). (A 'record' is the container of the bibliographic and/or descriptive metadata per collection item.)

(B) Estimated number of analogue items in heritage collection

Provide the estimated number of analogue items that are part of your institutions heritage collections.

(C) Estimated number of born digital items in heritage collection

Provide the estimated number of born digital items that are part of your institutions heritage collections.

(D) Estimated % of objects in analogue collection that is digitally reproduced

You may include textual materials that are reproduced as images without character recognition.

Please include the digital output from projects that are approaching completion.

(E) Estimated % of objects in analogue collection still to be digitally reproduced

The part of your analogue collection that your institution intends to digitise.

(F) Estimated % of objects in analogue collection with no need to be digitally reproduced

The part of your collection that, for whatever reason, your institution does not intend to digitise, for instance because there is no demand for it or because it will be digitised by another institution.

SECTION 3. Digital Access

19. Collections are made accessible to the public for various reasons. How important is each of the following types of use for your institution?

Using a 10-points scale - where 1 equals "not at all important" to 10 "highly important" - please select only one number per row.

Type of use	1	2	3	4	5	6	7	8	9	10
Academic research										
Creative reuse/Remix										
Educational use										
Ideological, religious and commemorative use										
Personal enjoyment										
Reducing the use of the physical originals										
Sales, commercial licensing										
Other types of use (specify below)										

Other types of use: [Input box]

20. Does your organisation have an explicit (written) policy, endorsed by the management of your organisation, that sets conditions for specific types of use of your digital heritage collections, as specified for instance in the previous question?

[ ] Yes
[ ] No
[ ] Do not know

Answer this question with 'yes' if your organisation has a formal policy document detailing which digital materials are accessible to whom and what the terms and conditions of this accessibility are.

21. Does your organisation measure the number of times digital metadata and/or digital objects are being accessed by your users?

[ ] Yes
[ ] No
[ ] Do not know

In order to be able to answer this question with 'yes' any manner of measurement will suffice.

22. If Yes, how?

[ ] Website statistics
[ ] Social media statistics (e.g. Facebook, Flickr, Youtube, Wikipedia)
[ ] Database statistics (if not included in Website statistics and Social media statistics)
[ ] User studies
[ ] Other: [Input box]

Please indicate all ways in which the access of digital metadata and objects is measured.

23. Please indicate estimated percentage of all the digital objects you have that are and/or will be accessible through the mentioned access options

Access channel	*% of digital objects currently accessible* (estimation is OK)	*% of digital objects accessible 2 years from now (estimation is OK)*
Offline		
Institutional website		
National aggregator		
Europeana		
Other aggregator		
Wikipedia		
Other Social media platforms like Flickr, Youtube, Facebook		
Institutional API		
3rd party API		
Other Access channels (specify below)		

Other Access channels: [Input box]

Multiple access options for your individual digital collections are a possibility (i.e. Europeana and Wikipedia). Consequently, the sum total of your answers does not have to be 100%.

SECTION 4. Digital Preservation

24. Does your organisation have a written Digital Preservation Strategy, that is endorsed by the management of your organisation?

[ ] Yes
[ ] No
[ ] Do not know

The answer of this question will be 'yes' when your institution has a formal document that describes the strategy for the digital preservation and permanent access to your digital heritage collections.

25. Are your digital collections stored in digital archives that have been set up according to international standards for digital preservation?

[ ] Yes, we have our own digital archive that meets the international criteria for long term preservation
[ ] Yes, our digital collections are archived in a publicly managed professional digital archive
[ ] Yes, our digital collections are archived in a privately managed professional digital archive
[ ] No, we do not have a solution yet for the long term preservation of our digital collections based on international standards
[ ] Do not know

Answer this question with 'yes' if your institution is actively involved in safeguarding the digital heritage collections for future generations, based on international standards or best practices.

SECTION 5. Digitisation Expenditure

26. Please estimate your annual expenditure on your digital collections (total cost of ownership)

	Please estimate the budget concerned (€):	Please specify the year concerned:
Institutional expenditure (internal budget):		[drop-down list: 2010, 2011, 2012]
Temporary funded project expenditure (external budget):		[drop-down list: 2010, 2011, 2012]

These budgets should be estimates of the costs related to the initial creation, ongoing maintenance, enhancement and preservation of your digital collections. Please attempt to include the cost of the staff time devoted to digital collection related activities in these estimates. If budget year does not coincide with the calendar, please choose the calendar year that fits best (in terms of the number of months).

Costs can be divided into incidental (upfront) costs and structural (ongoing) costs:

- **Incidental costs** are defined as the costs having to do with the initial creation or acquisition of a digital collection. Examples: selection of materials, acquisition of digital born materials, scanning, descriptive metadata creation, project management.
- **Structural costs** are the costs needed for the ongoing maintenance, enhancement and preservation of a digital collection. Examples: activities concerning the preservation of digital collections, licences, maintenance of web servers, user outreach and support, management.

27. Please estimate what percentage of the total annual expenditures on creating/acquiring, maintaining, enhancing and preserving your digital collections can be assigned to incidental costs and what percentage can be assigned to structural costs

	Incidental costs:	Structural costs:	
Percentage [should add up to 100%]	… %	… %	100 %

28. Please estimate what percentage of the total annual expenditures on creating/acquiring, maintaining, enhancing and preserving your digital collections is spent In-house and what percentage is Outsourced

	In-house costs:	Outsourced costs:	
Percentage [should add up to 100%]	… %	… %	100 %

The community of libraries, archives and museums would benefit from a better understanding of the costs involved in creating and preserving digital collections. Please help us by providing a more detailed account of your costs.

[ ] Show table

I. Please estimate what percentages of the Incidental costs can be assigned to the following activities

%	**Incidental cost category**
	Project management
	Selection of material for digitisation
	Acquisition of digital born material
	Logistics (shipment of collection for digitisation, etc.)
	Analogue to digital conversion (including all technical and staff costs associated with the act of preparing materials for scanning, the scanning itself, and quality control)
	Copyright clearance
	Metadata creation
	Web design, software development
	Other costs (specify below)
100%	

The sum total adds up to 100%. Enter 0 if a cost item is not applicable. If you miss any items in the table, please help us and mention these under Other costs.

Other costs: [Input box]

II. Please estimate what percentages of the Structural costs can be assigned to the following activities

%	Structural cost category
	Management
	Archiving (storage, including backups)
	Activities concerning the long-term preservation of the digital collection (including research activities but excluding Archiving costs)
	Licences
	Maintenance of web servers and web, mobile and other services
	User outreach and support (including staff time for efforts to promote the use of the digital collections)
	Usage analysis (including user surveys, interviews, and other activities)
	Editorial (including content selection and updating)
	Other costs (specify below)
100%	

The sum total adds up to 100%. Enter 0 if a cost item is not applicable. If you miss any items in the table, please help us and mention these under Other costs.

Other costs: [Input box]

29. What is the total number of paid staff (in full-time equivalent) engaged in creating/acquiring, maintaining, enhancing and preserving your digital collections on an annual basis?

[Input box]

Include the time of your own institution's staff engaged in activities related to creating/acquiring, maintaining, enhancing and preserving your digital collections, including: planning and managing in-house and contracted projects; preparing and digitising materials; enhancing digitised output to widen accessibility.

30. What is the total number of volunteers (in full-time equivalent) engaged in creating/acquiring, maintaining, enhancing and preserving your digital collections on an annual basis?

[Input box]

Include the time of your institution's unpaid staff. Volunteers who receive compensation for their expenses (like travel costs) should also be included.

31. From what sources are your digital collection activities funded?

- Internal budgets
- Crowdfunding
- National Public grant/subsidy
- Regional/Local Public grant/subsidy
- Private funds and legacies
- Public/private partnership
- Sales of digital items
- Other: [Input box]

Indicate all the sources from which your digitisation activities are funded.

SECTION 6. General Notes

32. Please include any information that was not asked for above and that you think is relevant for understanding the nature of activities related to your digital collections.

[Free text field]

Comments on the questionnaire itself can be given in the next question (33).

SECTION 7. Questionnaire Evaluation

33. Please include any comments that would help us to improve future issues of this survey

[Free text field]