ENUMERATE

Survey Report on Digitisation in European Cultural Heritage Institutions 2012

by

Natasha Stroeker and René Vogels, Panteia (NL)
on behalf of the EUMERATE Thematic Network
May 2012

ICTPSP

Contents

1   Executive summary

The ENUMERATE *Survey Report on Digitisation in Cultural Heritage Institutions 2012* represents the first 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 29 European countries. About 2000 institutions answered the open call to participate between January and March 2012.

The survey asked questions about

- The state of digitisation activity in the institution responding;

- Access to digital collections;

- Its digital preservation strategy;

- Expenditure of digitisation by the institution.

Highlights of the report's findings are:

Digitisation activity

- 83% of institutions said curatorial care is part mission;

- 83% of institutions have a digital collection, or is currently involved in digitisation activities;

- c20% of all collections, that need to be, are digitised;

  - Art museums are the most digitised with 42%;

  - National libraries have only 4% digitised of a target of 62% of their collections.

- Photographs are most digitised object type;

- 89% of audio visual institutions have born digital collections, while 43% of museums of art and history have them.

- 34% of institutions have a written digitisation strategy;

- About one third of the institutions are included in a national digitisation strategy. For national libraries more than half are included.

Digital access

- 31% of the institutions have a policy on the use of the digital collections. Figures range from 60% for national libraries to 22% for archaeology museums.

- 42% of institutions monitor the use of their digital collection.

- c85% of institutions use Web statistics to measure the use of their digital collections;

- By 2014 institutions estimate to make twice as much of their collections accessible through Europeana when compared to today.

Digital preservation

- 23% of institutions have a written digital preservation strategy. Figures range from 44% for national libraries to 12-25% for museums;

- About a third of the institutions are included in a national preservation strategy;

- 40% of national libraries say there is no national digital preservation strategy, compared with 13% for all institutions;

- 30% of the institutions are included in a national digital preservation infrastructure.

Digitisation expenditure

- Audio-visual institutions stand out with a digitisation expenditure of €103,000 per FTE (full time equivalent);

- The other institutions are all in the costs range of about €20,000 to €40,000.

- On average 3.3% of paid staff in all cultural heritage institutions is working full time on digitisation;

- Separate budgets include lines for: staff (in 70% of institutions), equipment and conversion (50%), and rights clearance (8%);

- In national libraries on average 15 staff are involved in the digitisation process;

- In other types of institutions the 'digitisation team' is limited to on average to about 5.5;

- Volunteers are used most at archives and records offices;

- Funding from internal budgets is a source for 87% of the institutions;

- Public grants or subsidies are mentioned by 40% of institutions.

The report is the first in a series of three in the lifetime of the ENUMERATE project. Later in 2012 there will be an in-depth 'thematic' survey and an update on this work next year. It is intended that ENUMERATE will continue its work beyond 2014.

Funding for the project is from the European Commission funded under the *ICT Policy Support Programme* part of the *Competitiveness and Innovation Framework Programme*.

2   Introduction

In the period December 2011-April 2012 the EC-funded ENUMERATE network conducted its first *ENUMERATE Core Survey*[1]. The aim of this survey was to develop a clear picture of the progress made in digitisation in the European cultural heritage sector.

In recent years there has been a growing demand for reliable data about digitisation, access to digital heritage collections, and the preservation of digital heritage materials in the memory institutions of Europe.

The survey consisted of 32 questions and was made available in 16 languages. It was distributed via internet by Digibis in Spain.

An overall population database covering all European archives, libraries, museums and other heritage institutions is not yet available. For that reason in every ENUMERATE network country the national coordinator was asked to invite institutions to complete the survey.

A result of not having a complete data base is that we cannot verify to what extent the respondents represent the whole population. We also cannot calculate which part of the heritage population is represented in the survey results. And we also cannot extrapolate the results for all institutions in Europe[2]. The discussion about improving the survey is not part of this report but will be included in a separate document.

If possible we make a global comparison[3] with the outcomes of the NUMERIC research[4].

The analysis of the results and the preparation of the report were carried out by Panteia in the Netherlands[5].

---

[1] See: **www.enumerate.eu**

[2] All national coordinators were asked to give an estimate of the number of institutions in each area of cultural heritage in their country. At the moment of writing this report 17 countries had supplied us with this information. An analysis of these estimates in order to define the total population will be included in a review of the methodology.

[3] We have to be very careful with these comparisons, because we have little information on the population and the degree to which these are comparable. And another point is that the types of institutions defined in both studies is not completely the same and thus not directly comparable (see appendix 1). That is why we call it a global comparison.

[4] Study Report NUMERIC. Study deliverable N.8. Study findings and proposals for sustaining the framework. CIPFA, IPF (UK). May 2009.

[5] Recommendations for further improvement of the survey are described in the Review of Methodology (June 2012).

3 Response to the survey

The data collection process resulted in 1951 usable responses. This included both completely filled out questionnaires, and those that were not filled out completely (about 300 cases). Non-complete responses were included in the analyses only if the major part of the questions had been answered.

Of the 1951 institutions who responded 1623 had digitised their collections or had undertaken digitisation activities. Therefore 328 had no digitisation activities at all. How the responses are divided over countries and type of institutions can be seen in the table below. By far the highest response rate was in Spain with 255. The response rate of Bulgaria (0) and France (1) are problematic. The 'other' countries are Liechtenstein (5) and Switzerland (77). The total response of all countries is quite satisfying. The number of institutions is large enough to calculate reliable data on the level of institution type and institution size:
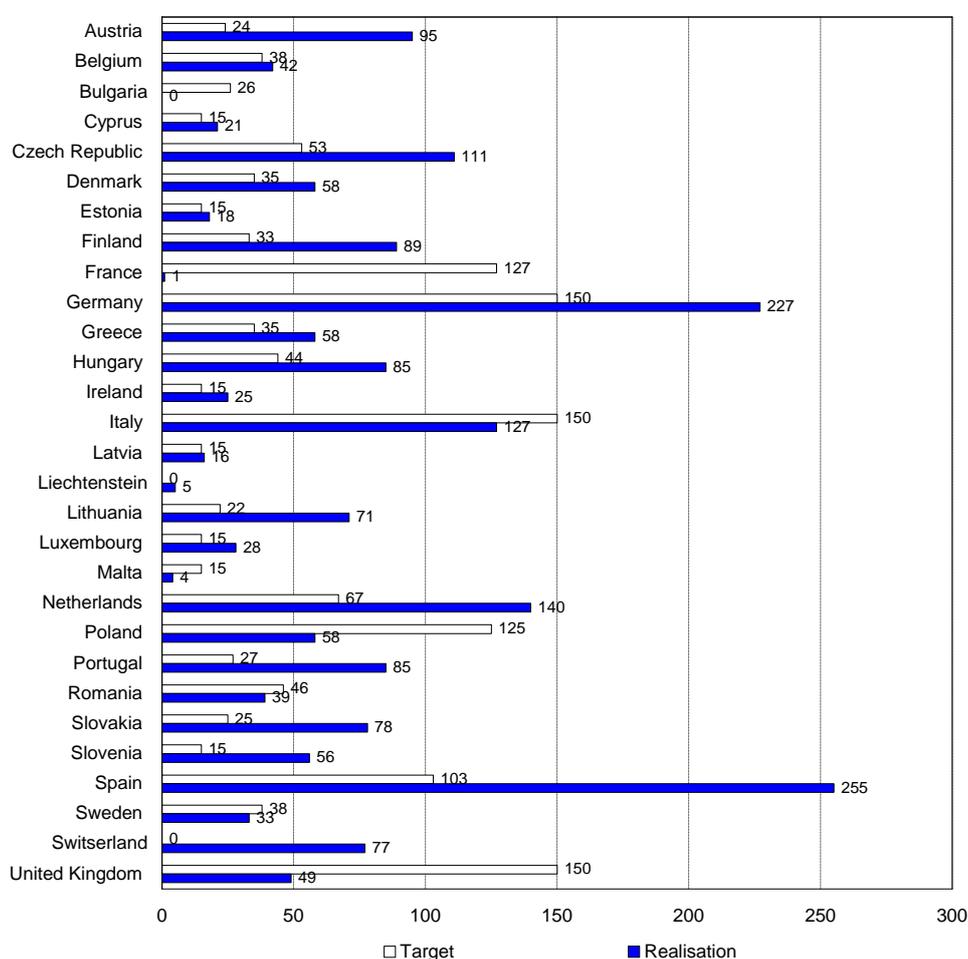

Figure 1: Response per country compared with target

The response is well spread over the different types of heritage institutions.
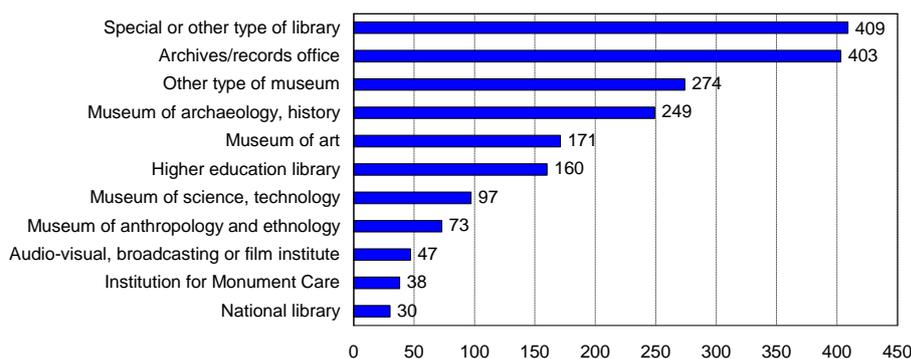
Figure 2: Response by institution type (n=1951)

In the response the number of museums is larger than the target. As mentioned before, a comprehensive database of institutions in Europe is not available. The targets for this study are based on an educated estimation of the ENUMERATE consortium and national coordinators in all countries involved in the survey.

	Target	*%*	*Response*	*%*
Archives	257	18%	403	21%
Audio, video or film	35	2%	47	2%
Libraries	380	27%	599	31%
Museums	375	26%	864	44%
Others	385	27%	38	2%
	1432	100%	1951	100%

Table 1: Target and response by institution type (n=1951)

Looking at the annual budget, the responses covered all categories of institutions, large as well as small. 27% have an annual budget of over €1 million. In contrast 16% have a rather small annual budget not exceeding €10 thousand.

	< 10	10-50	50-100	100-500	500-1M	1 - 10M	> 10M
Institution for monument care	10	7	3	21	14	34	10
Archives/records office	21	14	14	26	8	15	3
Audio-visual, broadcasting and film	7	7	10	17	15	20	24
Museum of art	16	9	7	21	13	23	11
Museum of archaeology or history	19	8	10	25	11	23	3
Museum of science or technology	16	9	8	18	12	25	12
Museum anthropology or ethnology	13	13	14	23	9	25	3
Other type of museum	19	13	11	26	13	16	2
Special or other type of library	16	14	8	26	10	22	4
National library	11	4	4	21	4	32	25
Higher education library	5	13	7	14	7	35	19
Total	**16**	**12**	**10**	**24**	**10**	**21**	**6**

Table 2: Annual budget by institution type in 1000 Euro, in % (n=1951)

Another possible indicator for the size of an institution is the number of paid staff or full time equivalents. There was a total of 165,116 FTE (full time equivalent) paid staff in organisation who responded to the survey. The size and budget in the audiovisual institutions is on average very large due to the fact that in this category national broadcasting institutions are represented.

	Average FTE	Total FTE	Response
Institution for monument care	107.1	4069	38
Archives/records office	38.5	15531	403
Audio-visual, broadcasting, film	417.0	19601	47
Museum of art	61.9	10593	171
Museum of archaeology or history	35.9	8943	249
Museum of science or technology	77.7	7538	97
Museum of anthropology or ethnology	37.7	2753	73
Other type of museum	33.1	9067	274
Special or other type of library	57.3	23417	409
National library	179.8	5394	30
Higher education library	363.8	58206	160
Total	**83.9**	**165116**	**1.951**

Table 3: Number of paid staff in FTE (n=1951)

4 Digitisation activity

4.1 Curatorial care as part of the mission

Curatorial care is part of the institution's mission for 83% of the institutions who responded:
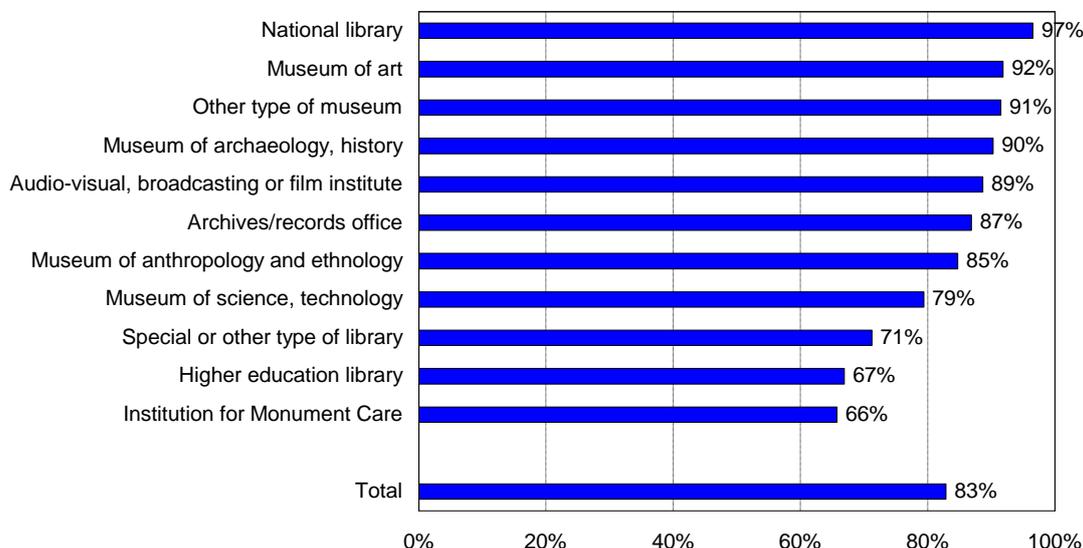

Figure 3: Is curatorial care for the collections of your institution part of its mission? (n=1951)

4.2 Digital collection?

83% of institutions have a digital collection, or is currently involved in digitisation activities. In the case of national libraries 100% have a digital collection. Of special libraries, only 75% has a digital collection:
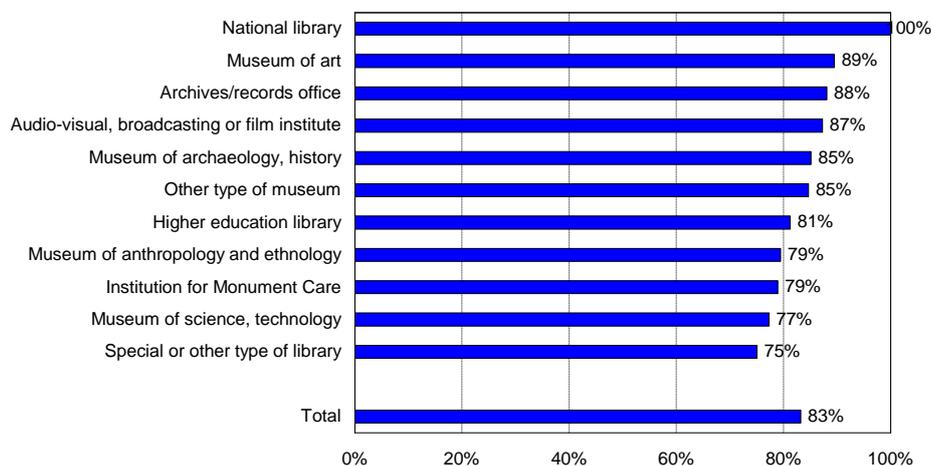

Figure 4: Does your organisation have digital collections or is it currently involved in digitisation activities? (n=1951)

When making a comparison of the results of Figure 4 with the outcomes in the NUMERIC study of the question 'Does your institution have a specific budget for digitisation activity' (Figure 3, NUMERIC report p44) the conclusion can be drawn that both National libraries and AV / film institutes have relatively high scores on these questions. So they are relatively often involved in digitisation activities and relatively often possess a specific budget for digitisation activities.

Where there is a digital collection then on average 20% of the whole collection is digitised (this concern estimates). Museums of art are currently the 'front runners' with 42% of their collections digitised. National libraries have a long way to go with currently only 4% digitised of a target of 62% of the collection. The data in Figure 5 do include estimates made by the respondents.
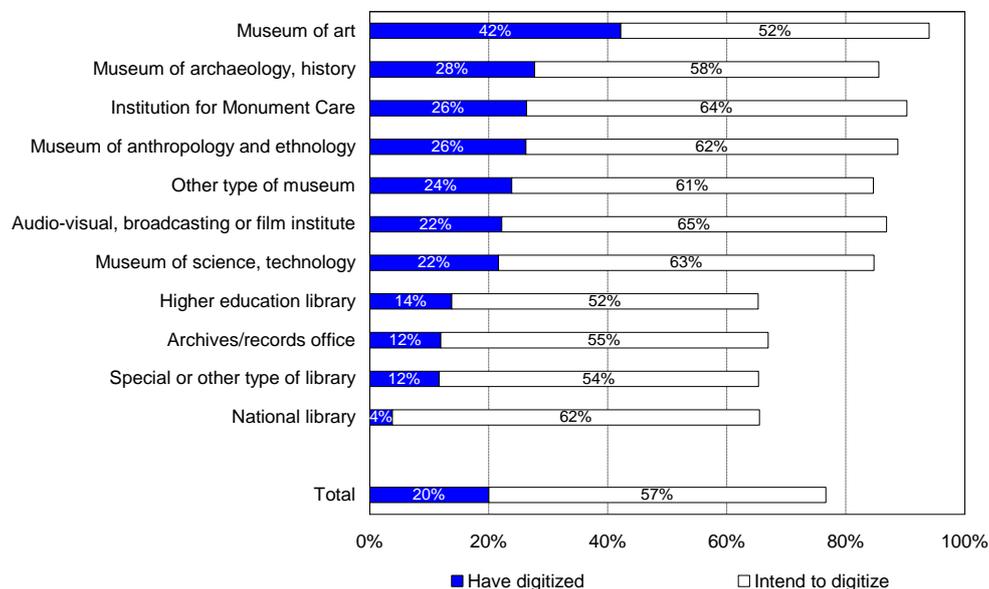

Figure 5: Estimated percentage of heritage collection already digitally reproduced and percentage still to be digitally reproduced (n=1626)

The results in Figure 5 can be compared to Figure 8 (p52) of the NUMERIC report. A relatively high percentage 'still to digitise' can be found for both national libraries and AV / film institutes in both surveys. When comparing the results of this question in both surveys on the total level, it can be concluded that the percentage of outstanding / still to digitise had increased between the NUMERIC survey and the ENUMERATE survey. The percentage 'no need' which is 23% in the ENUMERATE survey had decreased, since this was more than 30% in the NUMERIC survey. The percentage 'completed' is a few percentages higher in the ENUMERATE survey compared to the NUMERIC survey. Again one should not forget that these concern estimates made by the respondents.

4.3 Object types currently digitised

Most institutions have a wide range of object types in their collection. Institutions were asked which object types are currently in the digital collection and which types are still to be digitised. This is no indication of the total size of the collection digitised, but only an indication of objects in the collection[6].

Photographs are most digitised object type. 2 dimensional and 3 dimensional objects of art are less frequently mentioned.
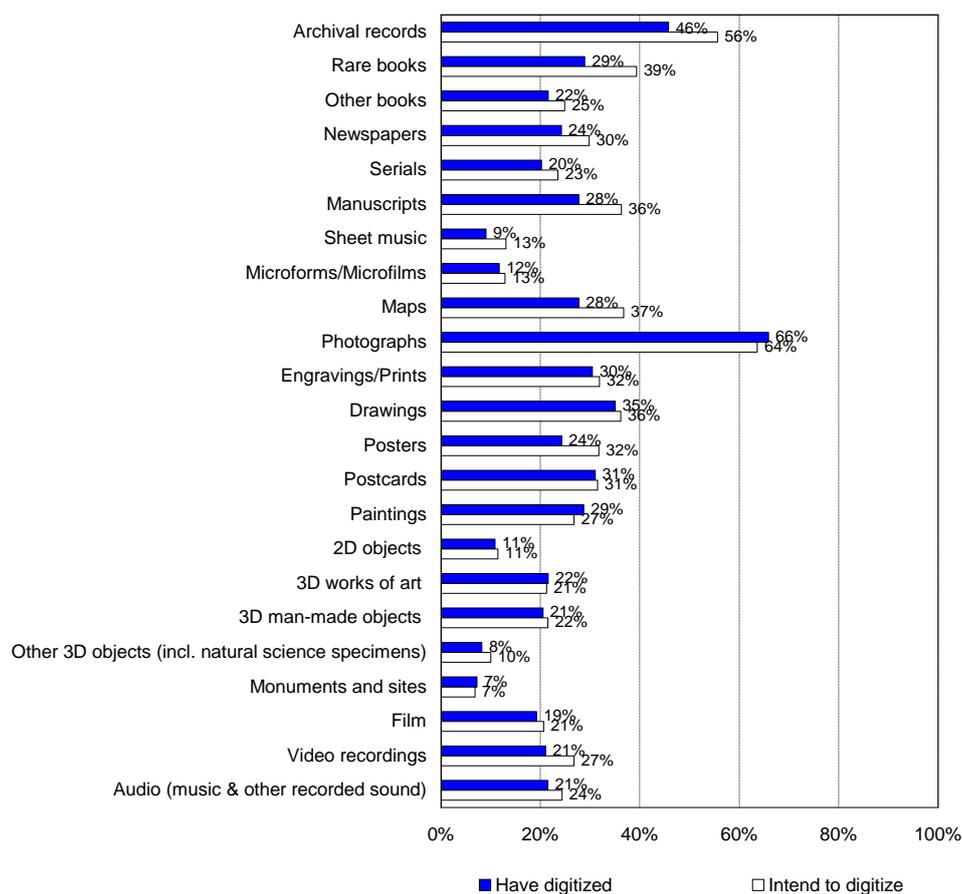

Figure 6: Object types most frequently digitised and intended to be digitised (n=1626)

In the survey an optional question was added concerning the size of the collection, the part of the collection catalogued in digital metadata, the part currently digitised, the part intended to digitize and the fraction with no intention to digitize. Unfortunately, this optional question is only answered by about 200 organisations. Only the results for these 200 organisations are included in the analysis. Of course these results are no more than an indication of the progress made by these 200 organisations. They certainly cannot be interpreted as the European average (see figures 7 and 8).

---

[6] This question does not answer which part of the archival records or other books has been digitized, but it only answers the question: if you have digitized, has that been: books (yes or no), archival records (yes of no) etc. And this holds for all types of objects that can be digitized. Thus, the percentages in Figure 6 may add up to more than 100% since a lot of institutions have digitized more than one type of material / objects.
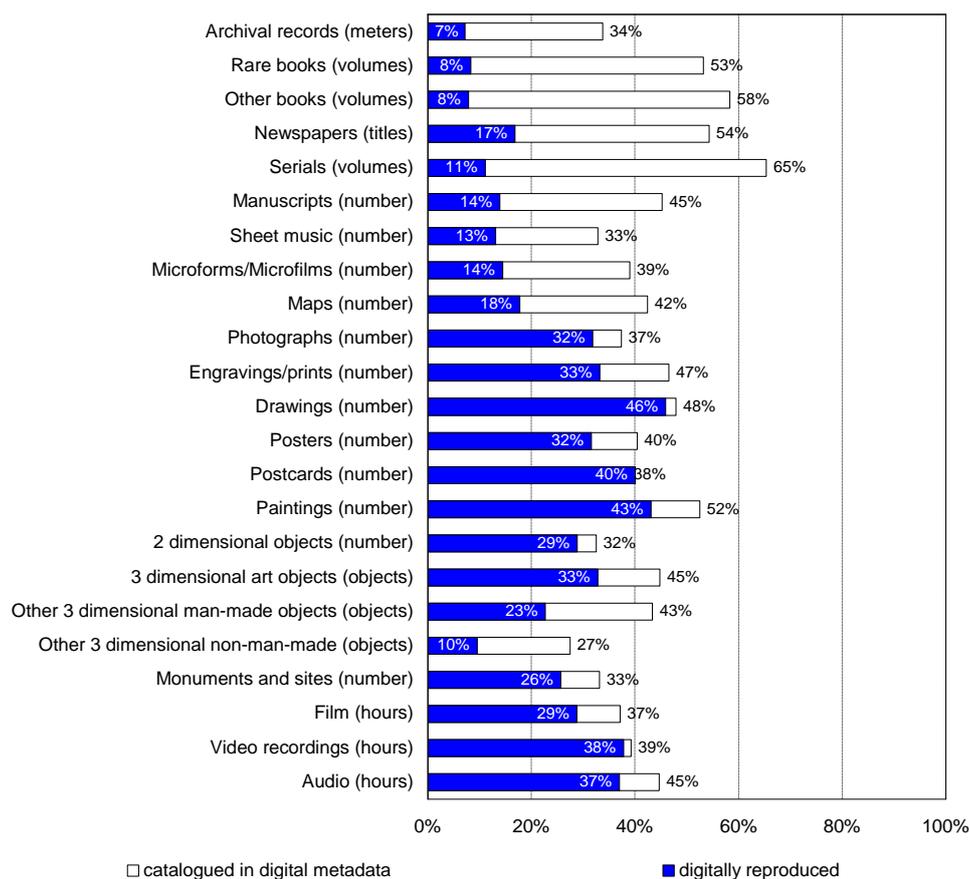
Figure 7: Part of the collection catalogued in digital metadata and percentage of the collection currently digitized (n=200)

Based on the small group completing these optional questions, we can conclude the ambition for the digital activities is quite high, given the high percentages that can be found for 'still to digitize' in Figure 8.
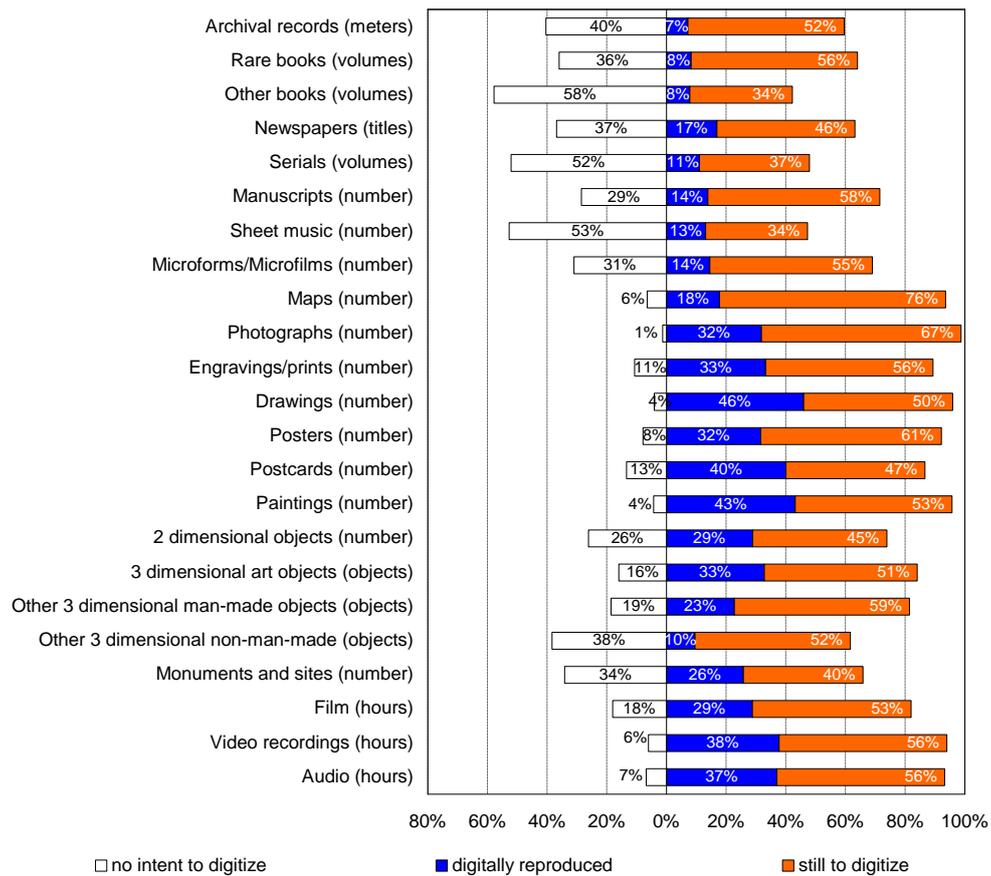
	no intent to digitize	digitally reproduced	still to digitize
Archival records (meters)	40%	7%	52%
Rare books (volumes)	36%	8%	56%
Other books (volumes)	58%	8%	34%
Newspapers (titles)	37%	17%	46%
Serials (volumes)	52%	11%	37%
Manuscripts (number)	29%	14%	58%
Sheet music (number)	53%	13%	34%
Microforms/Microfilms (number)	31%	14%	55%
Maps (number)	6%	18%	76%
Photographs (number)	1%	32%	67%
Engravings/prints (number)	11%	33%	56%
Drawings (number)	4%	46%	50%
Posters (number)	8%	32%	61%
Postcards (number)	13%	40%	47%
Paintings (number)	4%	43%	53%
2 dimensional objects (number)	26%	29%	45%
3 dimensional art objects (objects)	16%	33%	51%
Other 3 dimensional man-made objects (objects)	19%	23%	59%
Other 3 dimensional non-man-made (objects)	38%	10%	52%
Monuments and sites (number)	34%	26%	40%
Film (hours)	18%	29%	53%
Video recordings (hours)	6%	38%	56%
Audio (hours)	7%	37%	56%

Figure 8: Part of the collection currently digitized, intended to digitize and with no intention to digitize (n=200)

4.4 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, video, sound, digital art, games or websites.

Not surprisingly audio visual, broadcasting and film institutions mostly have born digital material. This is the case in 89% of these institutions. In archaeology museums born digital material is less frequently a part of the collection.

Types of born digital heritage that are mentioned most often are: photographs, video's / DVD / CD, other audiovisual objects or audio-recordings, TV and radio programmes, film, archives and archival records, E-books and E-journals, web pages and websites, computer software and games. Also mentioned, but a bit less are datasets, interviews, oral histories and PDF (or other formats). There is little diversity between the mentioned types of born digital heritage between the types of institutions except from archives that of course are more than average mentioned by archives themselves.
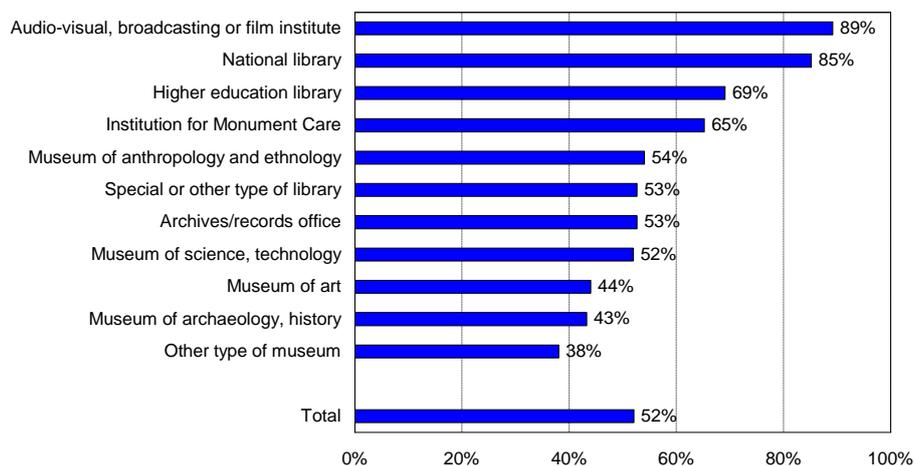
Figure 9: Does your organisation collect born digital heritage? (n=1549)

4.5 Digitisation strategy

A few questions cover the embedding of a digitisation strategy in the institutions. Only 34% have a written document covering the digitisation strategy. About half of the national libraries and museums of art have such a document. Of the archive and records offices and the higher education libraries 30% or less have a written strategy.
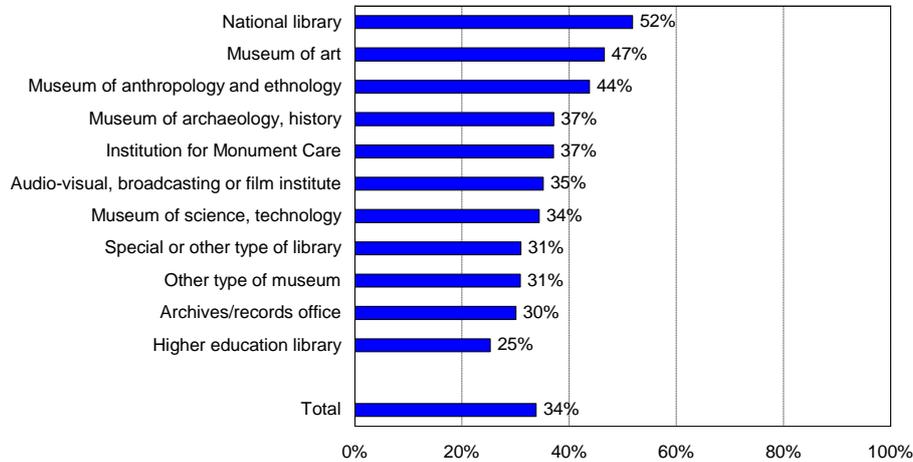

Figure 10: Does your organisation have a written digitisation strategy? (n=1557)

The results presented in Figure 8 can be compared with Figure 4 (p47) of the NUMERIC report. It seems little overall progress is made (33% in the NUMERIC study versus 34% in the ENUMERATE survey have a written digitalisation strategy). However, if you look closer you can see that especially museums of art have made a jump forward (from 33% – 34% to 47% in the ENUMERATE survey).

About one third of the institutions are included in a national digitisation strategy. For national libraries more than half of the population is included in a national strategy.
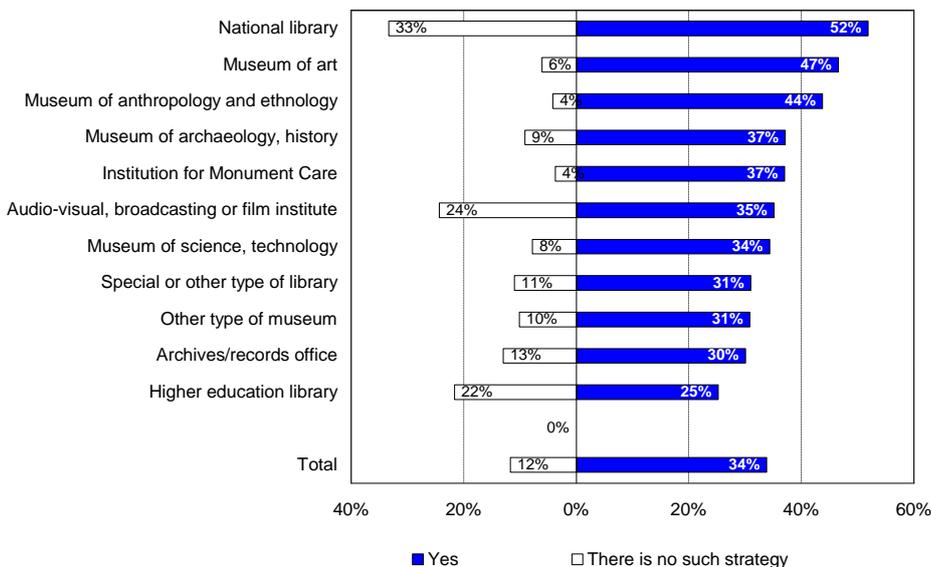
Figure 11: Is your organisation included in a national digitisation strategy? (n=1557)

5   Digital access

31% of the institutions have an explicit policy regarding the use of the digital collections. The institution type makes a large difference here. 60% of the national libraries have an explicit policy, but only 22% of the archaeology museums.
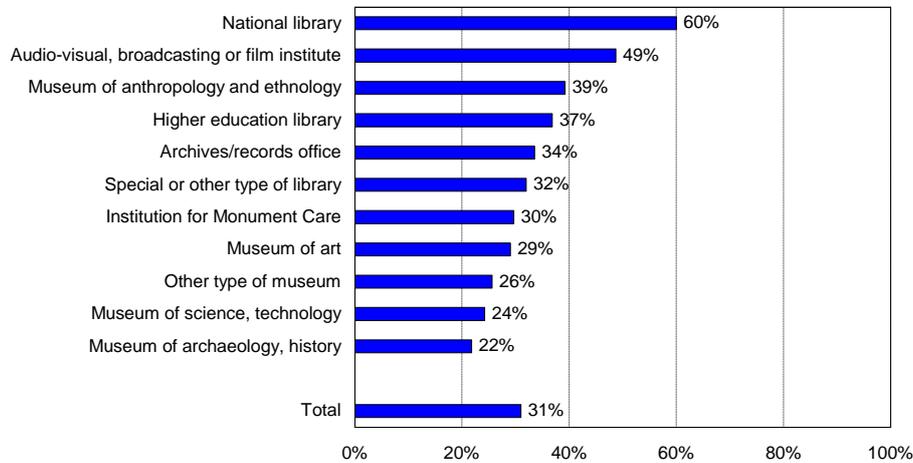

Figure 12: Does your organisation have an explicit policy regarding the use of your digital collections? (n=1488)

5.1 Measuring the use of the digital collection

Most libraries measure the number of times the digital collection is accessed. In the museum world this is far less the case. Only about a third of the museums monitor the number of people accessing the digital collection. On average 42% of all institutions monitor the use of the collection.
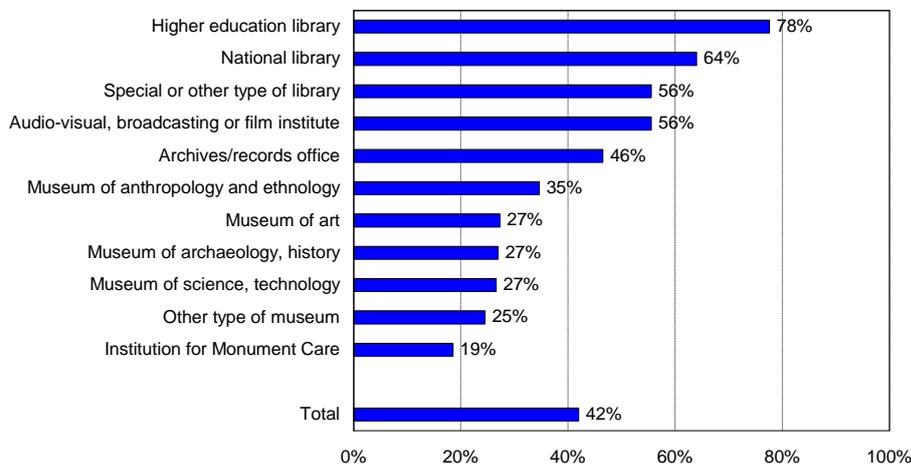
Figure 13: Does your organisation measure the number of times digital metadata and/or digital objects are being accessed by your users? (n=1495)

Web statistics are widespread as a tool to monitor the use of the digital collection. Libraries are frequent users of database statistics. User studies are mentioned by a third of the audio-visual, broadcasting and film institutions.
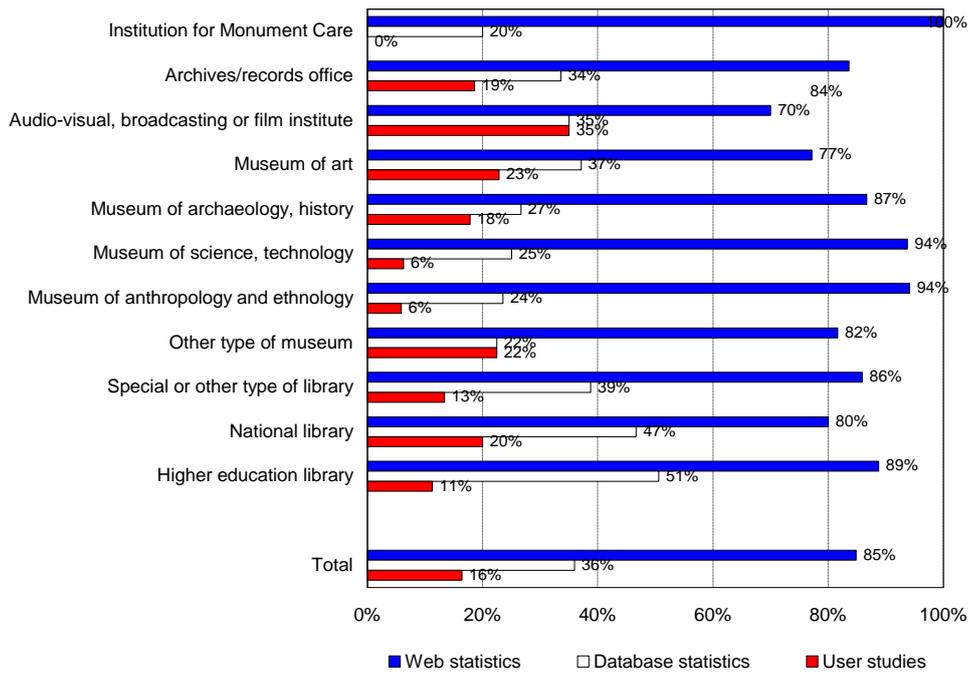
Figure 14: How is the use of the digital collection measured? (n=628)

To access the digital objects several access options are available. Institutions estimated which percentage of all the digital objects are currently accessible and will be accessible 2 years from now. Both are asked per access option.
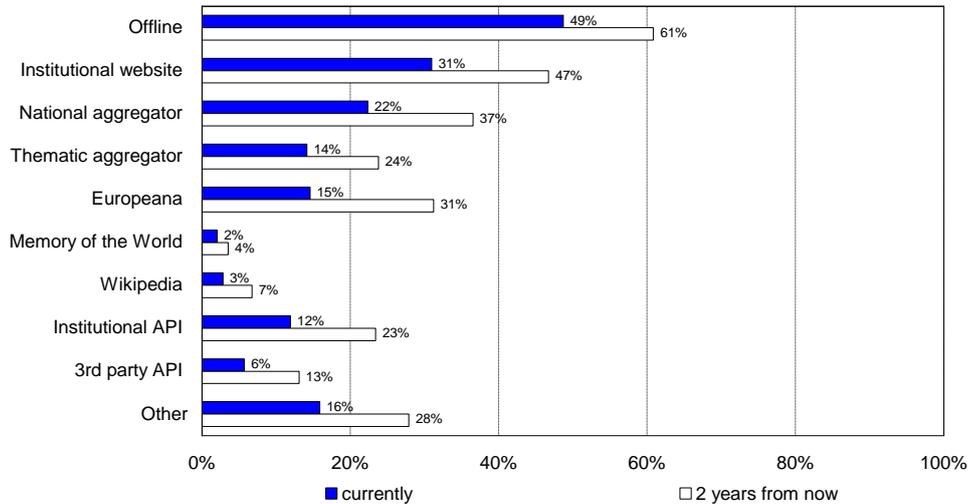

Figure 15: Which percentage of the digital collections is accessible through which options? (n=774)

6 Digital preservation

6.1 Digital preservation strategy

23% of the institutions have a written digital preservation strategy. The definition used in the survey for this document is: "*A document formally approved within an organisation, describing the way the organisation will be active in the preservation of its digitised and born digital collections.*"

On the level of national libraries about 44% have a written strategy. In the museum world a quarter to 12% can produce a written digital preservation strategy.
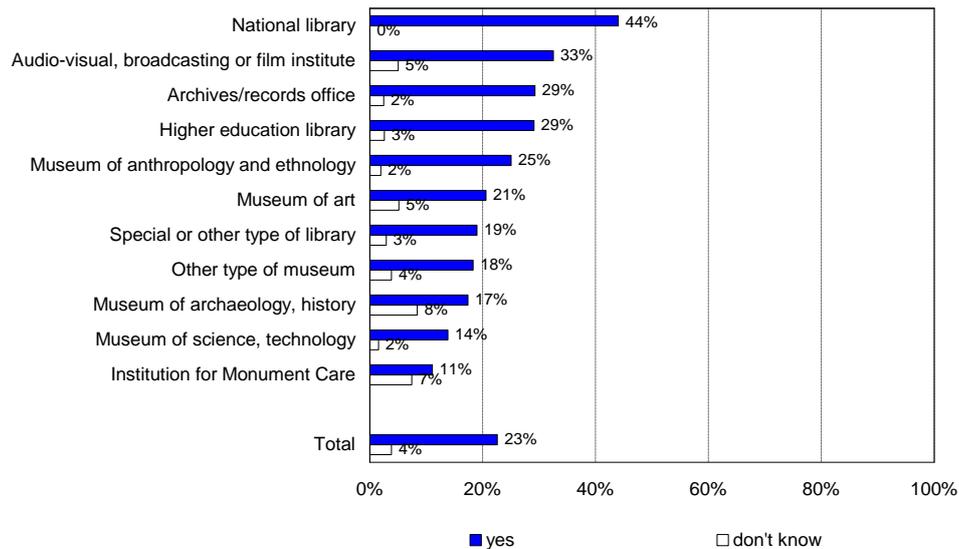
Figure 16: Does your organisation have a written digital preservation strategy? (n=1462)

6.2 National preservation strategy

About a third of the institutions in this survey are included in a national preservation strategy.

The percentage of national libraries stating there is no national digital preservation strategy is quite high, 40%, compared to 13% for the total sample of all institutions.
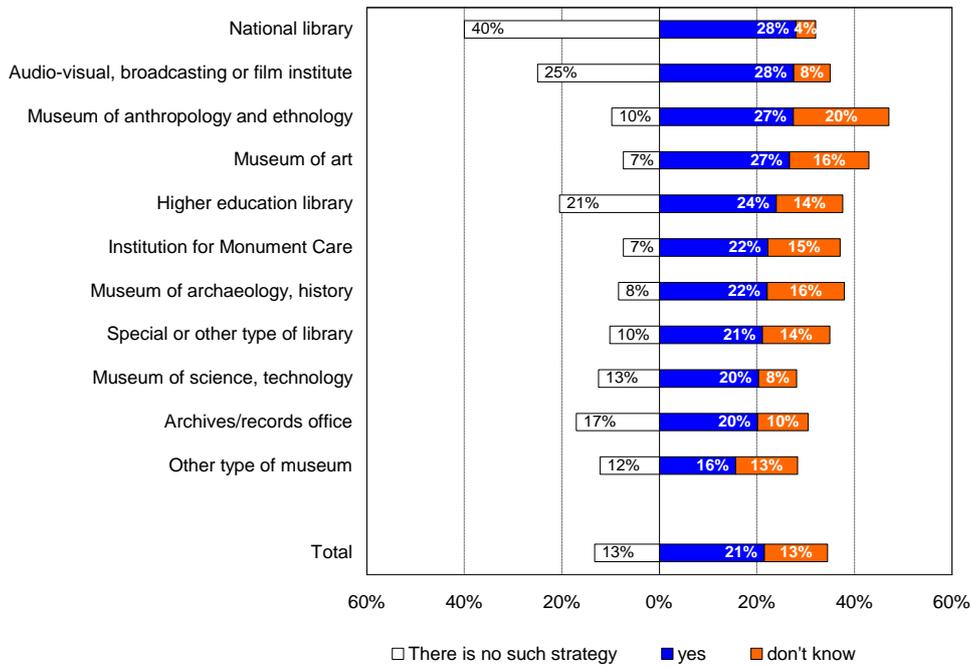

Figure 17: Is your organisation included in a national or institutional digital preservation strategy? (n=1462)

30% of the institutions are included in a national digital preservation infrastructure. This is the basic physical and organisational structures and facilities (e.g. hardware, software, and system management facilities) needed for the implementation of digital preservation.
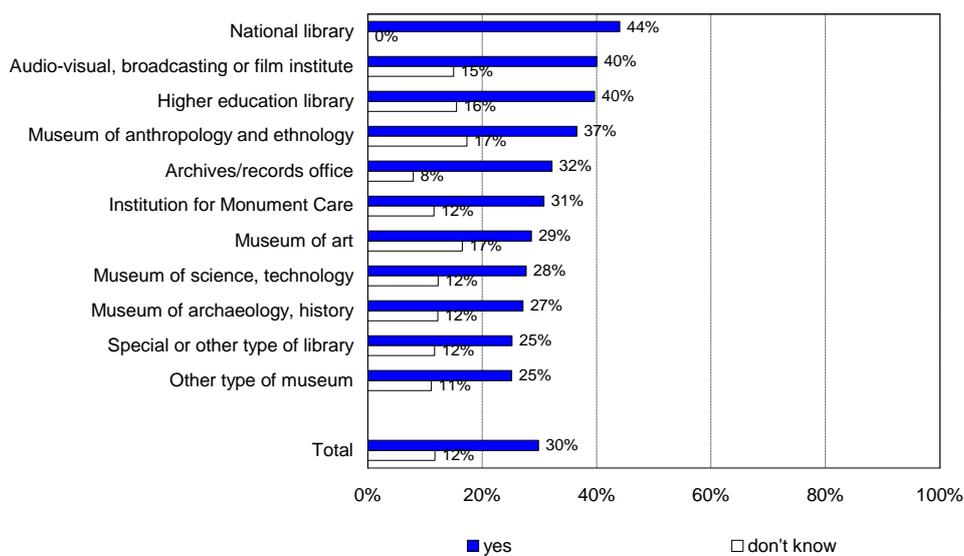

Figure 18: Is yvour organisation included in a national or institional digital preservation infrastructure? (n=1462)

7   Digitisation Expenditure

7.1 Cost components

In the closing part of the survey several cost components of the digitisation process are addressed. Institution size has of course a large influence on the costs. To make the indicators comparable we divided the costs in Euro by the number of paid staff involved in the digitisation process.

Audio-visual institutions stand out with a digitisation expenditure of €103,000 per FTE involved in the process. The other institutions are all in the costs range of about €20,000 to €40,000.
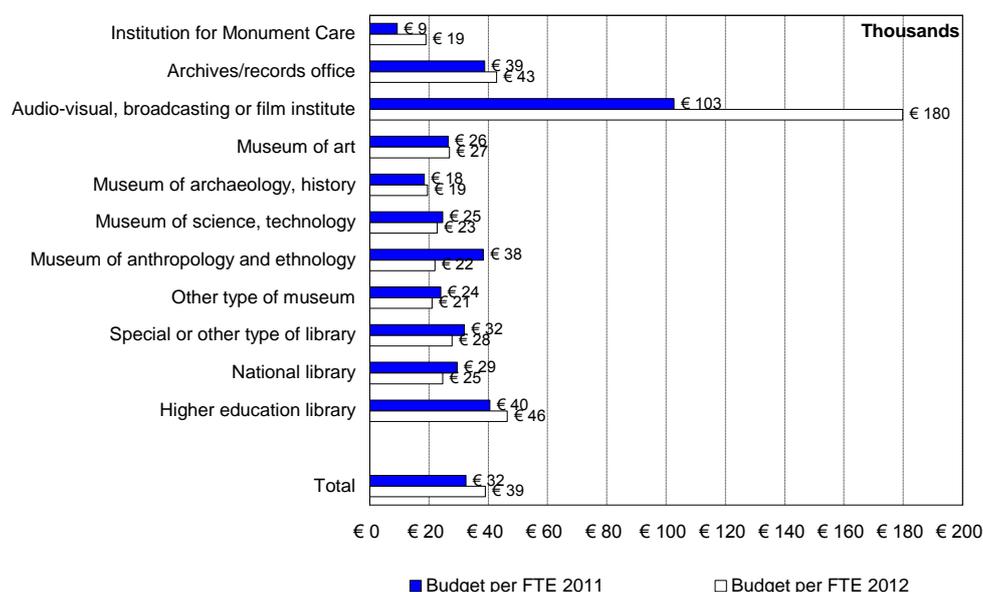

Figure 19: Annual budget per paid staff FTE engaged in digitisation activities (n=1584)

Budgets for digitisation activities can be compared to the total institutional budget. This can however only be calculated based on estimated data. The total institutional budget is registered in 7 categories in the range of <10,000 Euro as the smallest category and >10 million as the largest category. The percentages in Figure 18 are calculated only for those institutions which have a budget indication for 2011 and for 2012. Based on these calculations, the budgets for digital activities differ quite a bit between the various types of institutions. Archives have a budget of about 2.7%, while institutions for monument care can spend about 0.6% of the annual budget on digitisation activities.
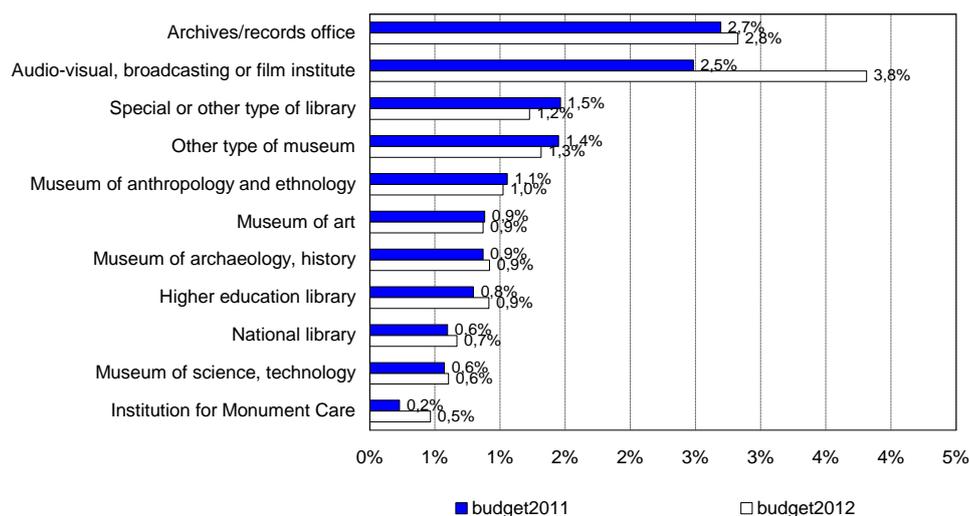
Figure 20: Annual budget for digitisation activities in % of total institutional budget (n=703)

Members of staff involved in the digitisation activities is compared to the total paid staff in FTE of the institutions. On average 3.3% of the staff is involved in the digitisation activities. This percentage is based only on those records in the Core Survey database for which both questions are answered. This percentage can be compared to the results in the NUMERIC survey (table 11, p46) which shows a total op 2.9% of the staff involved in the digitisation activities. So, an increase of staff involved in the digitisation activities is found.

	Total paid staff in FTE	Digitisation paid staff in FTE	In %
Museum of anthropology and ethnology	46,7	3,9	8,3%
Other type of museum	30,3	2,3	7,6%
Archives/records office	48,3	3,6	7,4%
National library	190,2	14,0	7,4%
Special or other type of library	66,8	4,0	6,0%
Museum of archaeology, history	44,2	2,2	5,0%
Museum of art	75,3	3,2	4,3%
Museum of science, technology	88,4	3,6	4,1%
Institution for Monument Care	145,6	5,2	3,6%
Audio-visual, broadcasting or film institute	610,1	8,1	1,3%
Higher education library	513,9	5,0	1,0%
Total	**112,4**	**3,8**	**3,3%**

Table 4: Digitisation paid staff compared to total paid staff (N answered both questions = 948).

The digitisation costs have several components. In the survey only an estimation of a total budget was asked for the last year and the coming year. To get an impression of what cost components are included in this estimation, a list is included in the survey with possible cost components.

In 70% of the cases the costs cover paid staff. In about 50% equipment and conversion costs are included. Rights clearance costs are only included by 8% of the institutions.
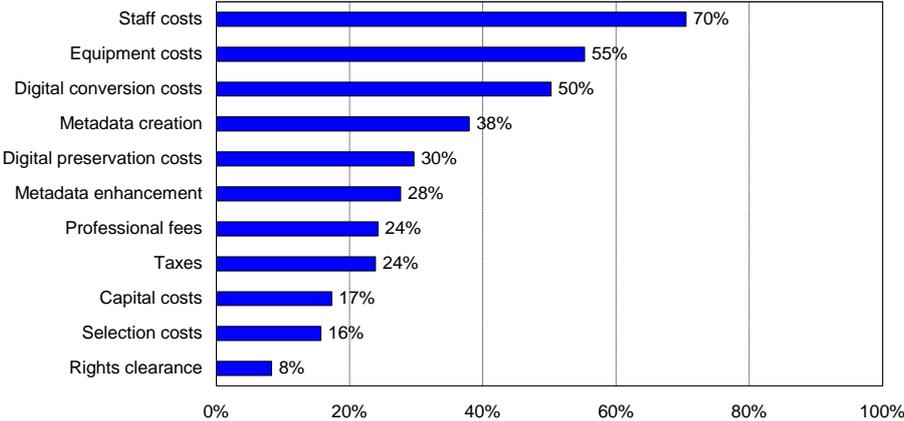
Figure 21: Included in budget estimation (n=1584)

7.2 Number of paid staff and volunteers involved

In national libraries on average 15 FTE (full time equivalent) are involved in the digitisation process. Most of these libraries are on average quite large institutions. In other types of institutions the 'digitisation team' is limited to on average about 5.5 FTE. Remarkable is the relative high number of volunteer FTE at archives and records offices.
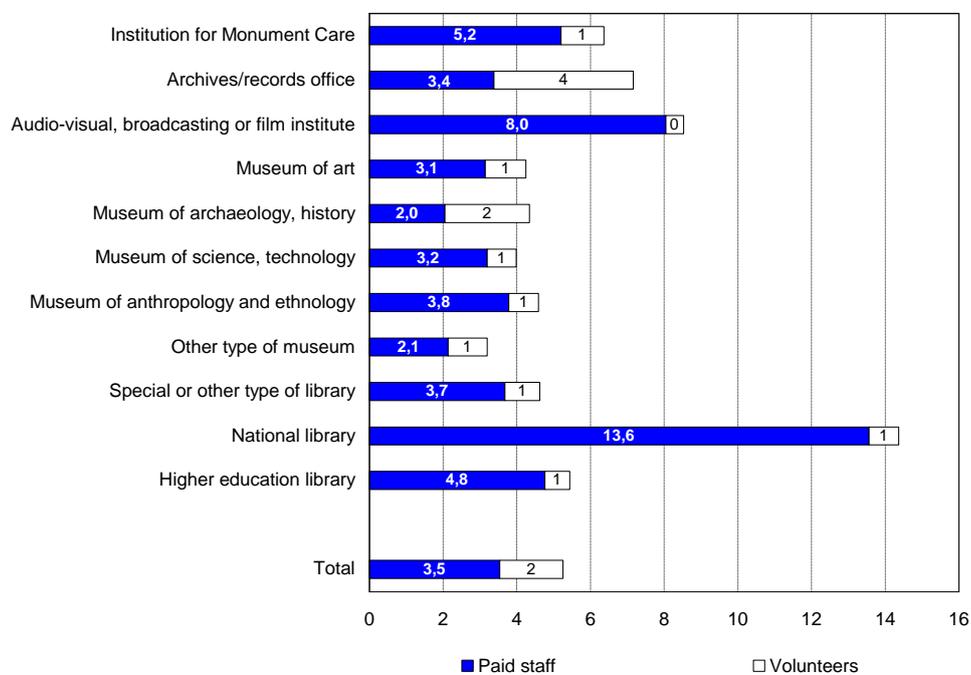

Figure 22: Staff involved in digitisation activities, in FTE (n=1584)

7.3 Funding

The digitisation activities can be funded from a variety of sources. Internal budgets are a source for 87% of the institutions. Public grants or subsidies are mentioned by 40%. The answers to this question did not indicate the amount of money in the funding. They only refer to the different sources.
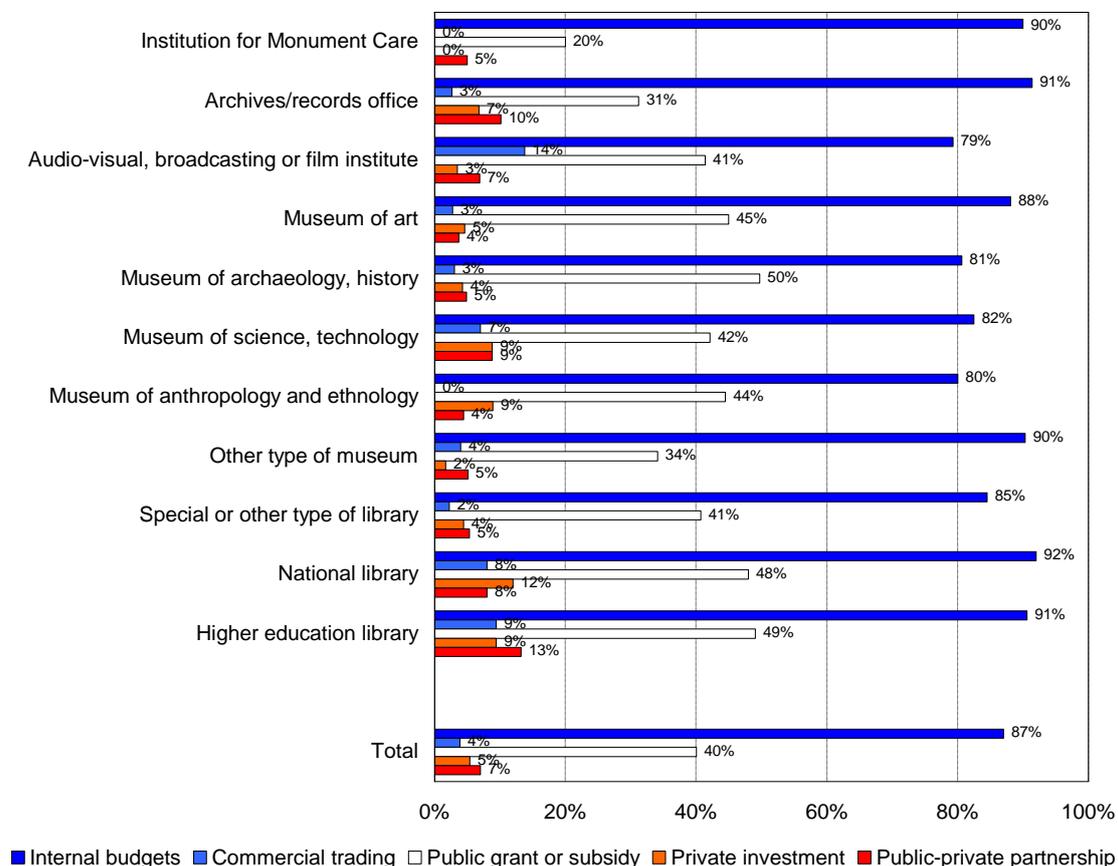

Figure 23: From what sources are the digitisation activities funded? (n=1584)