



Monitoring digitisation – A review of previous work

Profile of a digitisation service provider

A consensus framework

Consultation – Proving the worth of digitisation

Initial figures, but not yet facts



Monitoring digitisation – an analysis of previous studies covering the international experience

Large-scale digitisation is a priority for the development of digital libraries in Europe and for safeguarding museum and archive collections. Europe's citizens will benefit from wider access to their cultural heritage. However, significant investment is needed; and, monitoring the results of this effort is crucial.

The NUMERIC study has performed an analysis of previous attempts to measure the progress of digitising cultural materials. 32 digitisation reports have been analysed¹ to explore general trends, useful experiences and pitfalls in measuring digitisation in Europe and worldwide. Several points that will serve to improve the approach to designing the monitoring framework were concluded:

- Previous digitisation surveys have been mostly snapshots with no attempt to establish a continuous series of measures capable of describing either trends or progress from a baseline.
- These studies have attempted to trace the perspective from the top-down (policies that relate to institutions in common) and bottom-up (management information needs for the institutions). Both focus on the need for resources, but both have failed to define consistent measures for either financial or technical resources.
- Digitisation surveys have concentrated on collecting qualitative rather than quantitative information, and consequently few hard measures from which policies can be better informed have emerged. Consistent definitions are required.
- There are no uniform approaches on how to classify digitised collections. Few studies have successfully been able to relate the way in which digitised outputs can be compared with the analogue collections. Often outputs are recorded as digital files, images, pages or collections, making meaningful analysis impossible. Setting common input/output guidelines is the single most important requirement in establishing a successful statistical framework.
- Project objectives reveal different levels of aspiration in the digitisation work being pursued. The basic level would suggest that an institution ignores the social or cultural benefits of digitisation beyond its own organisation; an intermediary level would refer to digitisation for online access and preservation of the materials; and the highest level of development sets out to exploit the digitised collections by developing educational, tourism, research and other such benefits.
- Indicators of user needs, usage of materials, accessibility and usability of digitised content should be more visible in digitisation surveys and treated as important components of digitisation progress.

“There is no evidence from previous studies to show that memory institutions account for their digitisation work in a consistent way.”

1. See: “An analysis of state-of-the-art in measuring the progress of digitisation of cultural materials.” http://www.numeric.ws/download.asp?filename=NUMERIC_Desktop_Research_on_Digitisation_Studies_26102007114936.pdf

Example of services offered by a digitisation service provider

Infotechnique S.A. is one service provider that prides itself on the quality of the conversion work they undertake for their clients. The company is a division of the Getronics Group, and has invested in machines capable of handling both volume, but also complexity of content to be digitised. However, the ability to handle high volumes of complex materials is not the only determinant of good service, particularly when the contract is phased over four years – increasingly typical, if we are serious about breaking the translation backlog of Europe’s cultural heritage.

The company’s main service centre is based at La Walck in the Alsace, where each digitisation project is planned in meticulous detail – from GPS tracking of materials taken from the institution’s site, to electronic tagging and logging the precious materials at every stage in the conversion process. Their machinery is capable of reading and interpreting 800 pages of handwritten documentation per hour – and the error rate on processing 9 million written entries is no more than 4 words per thousand.

Emmanuel Amasio, head of production system design and Partnership Manager of Infotechnique, points to the improvements in digitisation standards that the new technology has brought about: the unit cost of digitising and structuring the Official Journal of the European Commission was once €15 per page – now the technology can read and file a fragile newspaper page for €1. Yet, as all those responsible for digitisation projects know, the cost can vary enormously, and interpreting handwritten pages handled in bulk can prove very much more expensive to process than, say creating simple TIFF files (simply images) that many institutions have hitherto commissioned. Emmanuel is enthusiastic in his support for gaining better statistics about the cultural heritage agenda, as he recognises the benefits that such information can provide in better planning to adopt an industrial approach with good quality control.

Emmanuel Amasio can be contacted as follows:

T +352 40 99 44 29

Emmanuel.Amasio@infotechnique.com

www.infotechnique.com



A professional and ministerial consensus to monitor progress

The best approach to emerge from the desk research, undertaken by the NUMERIC study team, is to adopt a two tier approach to measuring the digitisation activities amongst Europe's cultural heritage institutions in 2008.

Their research shows that relatively few institutions (less than one-third) have adopted formal project plans to direct their digitisation objectives, and consequently the majority do not maintain the management information in a form that is convenient for statistical reports. This reinforces the importance of seeking information that is relatively easy to generate and that can derive from a representative sample in each country.

The professional associations active at the international level mainly represent nationally prominent institutions, and they have been largely successful in gaining survey responses from their members¹. They will be consulted over the recommended approach in order to build on their established networks, but it will be necessary to seek another tier of information covering the activities of the numerous regional and local institutions. The cooperation of the relevant cultural ministries and / or central statistical offices in each country will be required to coordinate sample surveys, and they will be consulted accordingly. The deriving information can then be combined with the data relating to the national institutions to provide a comprehensive picture across Europe.

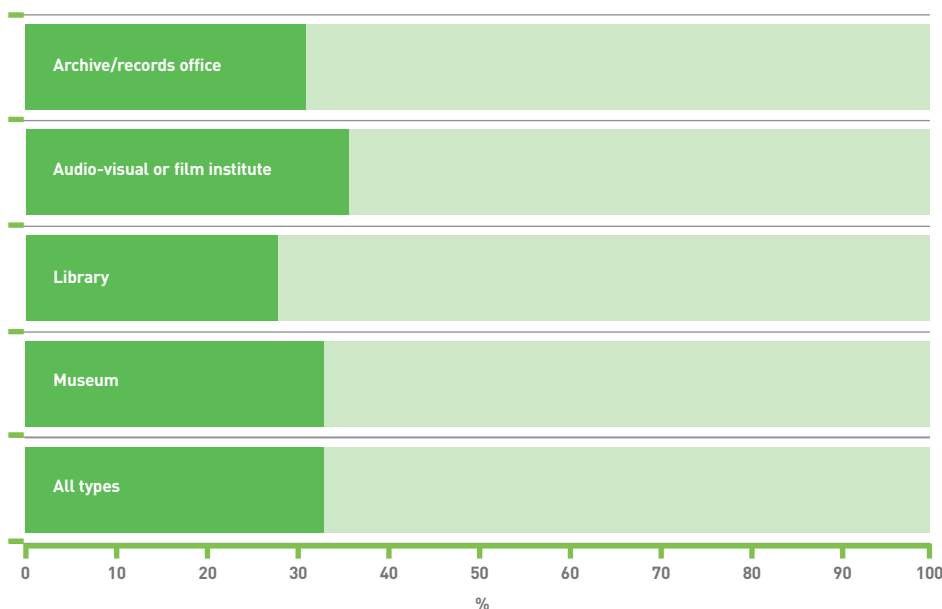
The "pathfinder" questionnaire, that the study team have used to test the approach and collect initial data, will be improved to eliminate ambiguities and clarify definitions. A lot has been learnt to more consistently measure digital output against the equivalent analogue materials. The framework questionnaire will also be simplified to reduce the perceived burden for responders.

The study team will keep in close contact with Eurostat and as part of their input, each ministry will be asked to verify the basic statistics relating to the number of institutions in each domain in their country, and to confirm the budgets and associated cultural statistics that the study team have estimated using data collated by ERICarts² and surveys conducted by CENL and EGMUS. This will allow the proper weighting of sample responses, to infer appropriate estimates for missing / non-responding institutions. UNESCO will be asked to recommend the survey to the relevant ministries outside Europe, but at least to those in India, China, Australia, Canada and USA. Regardless, the survey questionnaire will be made available on the Numeric website, just like the current framework survey; for completion directly.

This approach will be refined and international endorsement sought at a workshop to be held in Luxembourg in early 2008. An important outcome from this workshop will be a standard questionnaire, and the commitment of all interested parties, both professional and ministerial to support the framework.

Exhibit 1: Proportion of institutions who have formulated a digitisation plan

● Yes ● No



1. Comprehensive statistical surveys have been established by the Council of European National Libraries and the European Group of Museums, the results of which have informed the NUMERIC research. The Ligue des Bibliothèques Européennes de Recherche, the Consortium of European Research Libraries and the Association des Cinémathèques Européennes have also assisted NUMERIC. Further assistance will be sought from EBNA (Archives) to cover the professional networks and FIAT (Film Institutes).

2. ERICarts is the European Institute for Comparative Cultural Research. CENL is the Council of European National Libraries, and EGMUS is the European Group of Museums.

Consultation: Overwhelming desire to prove the worth of digitisation through exploring access to the materials

More than fifty people from fifteen countries have fed-back their views on what information they believe the statistical framework should provide. Their main message was the need to establish the link between the investment in digitisation and the benefit to those persons who would gain access to the materials.

“The worth of the framework is recognised – but unless institutions respond to future surveys, the statistics that are capable of influencing policy will never be generated. This really is an instance where actions count more than words.”

“There’s a lack of information about the access and the use of digitised content by the public; having this sort of information could really address and support the future of digitisation initiatives”, was a typical response.

Another angle on the same point stated that “There has been significant investment into digitisation. Unfortunately, I am still to be convinced that what has been digitised is being used to its capacity.” The importance of gaining hard facts to inform funding policy decisions was recognised by everyone.

Consultees said that many issues were equally important, and ranking them was therefore impossible, but preference was shown for information about the “Cost” and “Access / Usage” of digitised materials, followed by statistics about the progress “Achieved” and “Planned”. To an extent, information about the “Formats” was regarded as contextual rather than of strategic need to inform the policy debate.

“Funding” was reported as the single most problematic issue relating to digitisation projects, “Copyright” issues were also significant. The point was also made that once the materials had been digitised further investment was needed to present the information to users. However, the positive impact that digitisation programmes were making in benefiting future generations and opening up access the cultural heritage to all now was stressed.

The worth of the framework is recognised – but unless institutions respond to future surveys, the statistics that are capable of influencing policy will never be generated. This really is an instance where actions count more than words!



Books in need of preservation?



And now accessible to millions.

Emerging impressions: Initial figures, but not yet facts

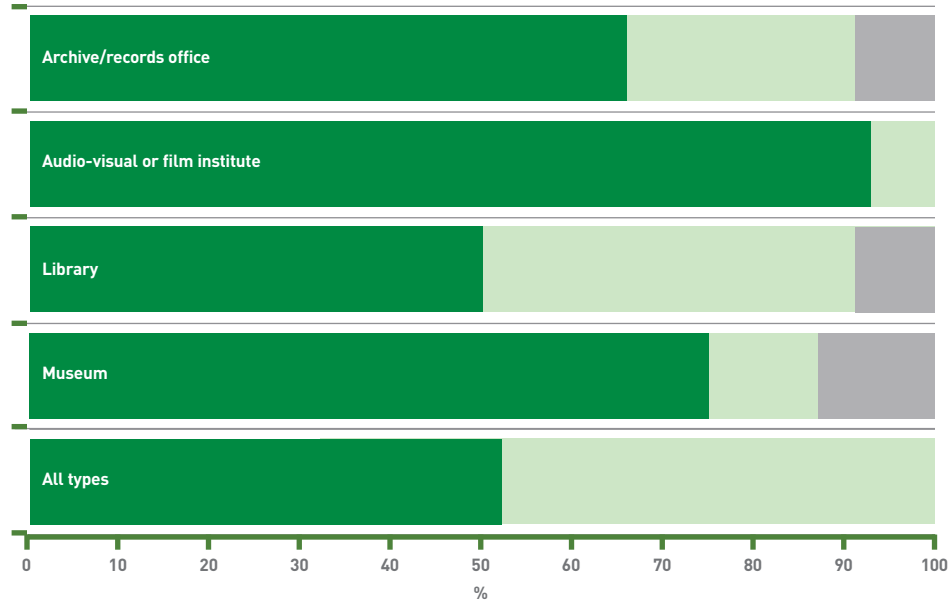
During the Autumn of 2007, a “Pathfinder” survey was conducted by the NUMERIC study, open to all institutions to report on their digitisation activities. This survey was supported by the most prominent professional associations active at the international level who encouraged the participation of their members¹.

The questionnaire is still available to download from the NUMERIC website, and further responding institutions will help to improve upon the reliability of the emerging statistical results. Until such returns are received the following summaries are based on insufficient results to be taken as anything but very tentative estimates.

The study estimates that there are about 58,000 “relevant” institutions in the European Union (EU27) where cultural materials are capable of being digitised [Archives / Records Offices, Libraries, Museums, Audio-visual or Film Institutes]. This excludes a large number of school and neighbourhood libraries where heritage materials are not kept, or only form a small proportion of the collections.

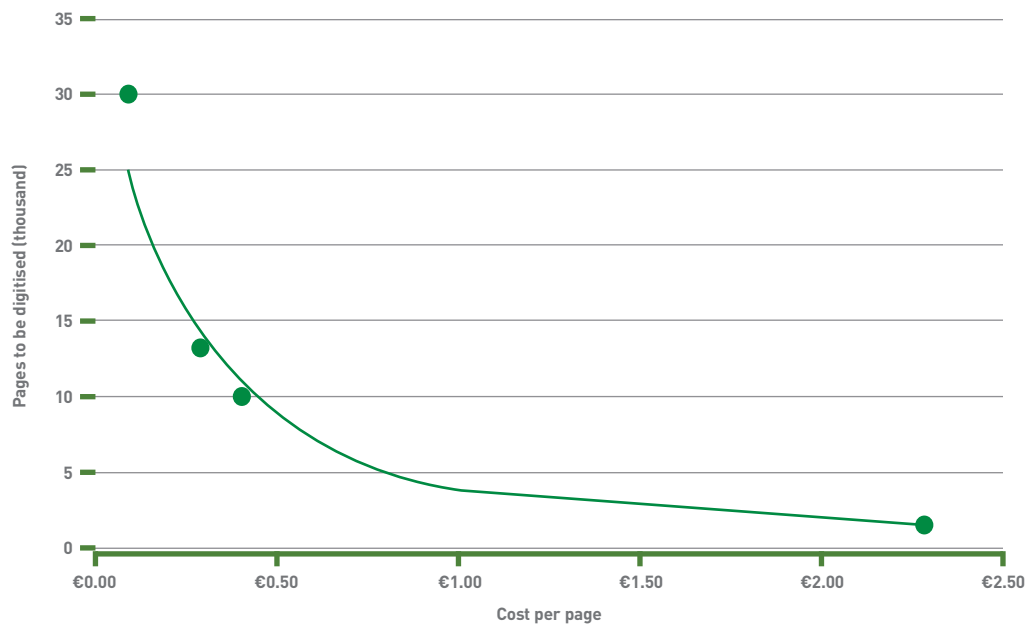
Exhibit 2: Who performs the work?

● In-house ● Contractor ● Partnership



Source: Numeric Pathfinder Survey, November 2007.

Exhibit 3: Planned costs of digitising books



1. Comprehensive statistical surveys have been established by the Council of European National Libraries and the European Group of Museums, the results of which have informed the NUMERIC research. The Ligue des Bibliothèques Européennes de Recherche, the Consortium of European Research Libraries and the Association des Cinémathèques Européennes have also assisted NUMERIC. Further assistance will be sought from EBNA (Archives) to cover the professional networks and FIAT (Film Institutes).

It is estimated that these relevant institutions employ the full-time equivalent of approximately 9,500 staff on digitising cultural heritage materials. This represents fewer than 4 in every hundred staff engaged on digitisation work, even though most institutions rely on in-house resources to perform the work. This paints a picture of the majority of institutions seeking to digitise their materials as and when resources allow, and as a consequence over two-thirds have not been able to frame formal digitisation plans.

Given the need to fix more certainty around the funding of digitisation projects, the study is anxious to inform policy makers about mechanisms that appear to establish efficient and cost effective programmes. In this connection, the economies that can derive from organising larger scale projects are self-evident [Exhibit 3].

There is an enormous range in the types of material being digitised in the various types of institution, and the standards adopted and elements included in the reported costs also differ considerably. Hence the summarised results set out in exhibit 4 reflect a cross-section across all types of institution, and should not be regarded as benchmarks.

Further and more detailed results are published and regularly updated on the NUMERIC website.

Exhibit 4: Emerging unit costs of digitisation:

Objects/items/materials in collections:	Units adopted for the measure:	Pathfinder Survey – Median values			
		Digitised as % of Collection	Pages per unit	Cost per Unit €	Cost per Page €
a Books (including "Rare books")	Volumes	0.05	250	191	0.50
b Newspapers	Volumes	1.7	14	–	1.56
c Journals and other serials	Volumes	0.1	350	15	0.35
d Government publications	Volumes	16	723	2,961	3.72
e Other type printed material not classified above	Number	5.9	3	14	3.70
f Manuscripts	Number	0.3	123	184	8.74
g Maps	Number	1.2	–	11	–
h Photographs	Number	3.9	–	4.07	–
i Engravings	Number	4.4	–	20	–
j Drawings	Number	20	–	4.82	–
k Posters	Number	18	–	17	–
l Postcards	Number	33	–	3.14	–
m Sheet music	Number	1.3	23	29	1.07
n Other images not classified above	Number	18	–	5.13	–
o Archived records of government/administration	Metres	50	768	–	0.74
p Archived records of historic importance	Metres	25	500	399	0.80
q All other archived records	Metres	54	3,636	2,901	0.80
r Man-made artefacts in museums	Artefacts	53	–	3.89	–
s Natural world specimens	Objects	–	–	–	–
t Works of Art – 2 dimensions	Exhibits	72	–	11	–
u Works of Art – 3 dimensions	Exhibits	75	–	11	–
v Other objects on museum collections	Objects	10	–	–	–
w Film and video recordings	Hours	3.4	–	1,125	–
x Music and other recorded sound	Hours	9.1	–	14	–
y Other items not classified above	Number	45	–	–	–
Weighted averages		6.4	535	–	1.41

IPF

IPF is a support services company wholly owned by CIPFA. We specialise in financial advice and governance, property and asset management solutions, the supply of information, expertise and people with the skills to help you at the highest level.

info@ipf.co.uk
www.ipf.com

Registered Office: 3 Robert Street, London, WC2N 6RL

Numeric

Numeric is a study funded entirely by the Information Society and Media Directorate General of the European Commission.

If you require further information or in any other format, please contact:

Phillip Ramsdale
Research Manager
Phillip.Ramsdale@ipf.co.uk

Aidan Callaghan
Statistician
Aidan.Callaghan@ipf.co.uk

CIPFA

The Chartered Institute of Public Finance and Accountancy (CIPFA) is the leading professional body for public services, whether in the public or private sectors. It provides education and training in accountancy and financial management, and sets and monitors professional standards. Its professional qualification is high quality, relevant and practical, and is supported by a range of other products and services.

Please register your interest on the website:

<http://www.numeric.ws>
numeric@ipf.co.uk

Pascale Van Dinter, from the Scientific and Technical Information Service in Belgium, won some book tokens for her registration last year.

