

The European Library Content Strategy Workshop

The Hague, 23-02-2012

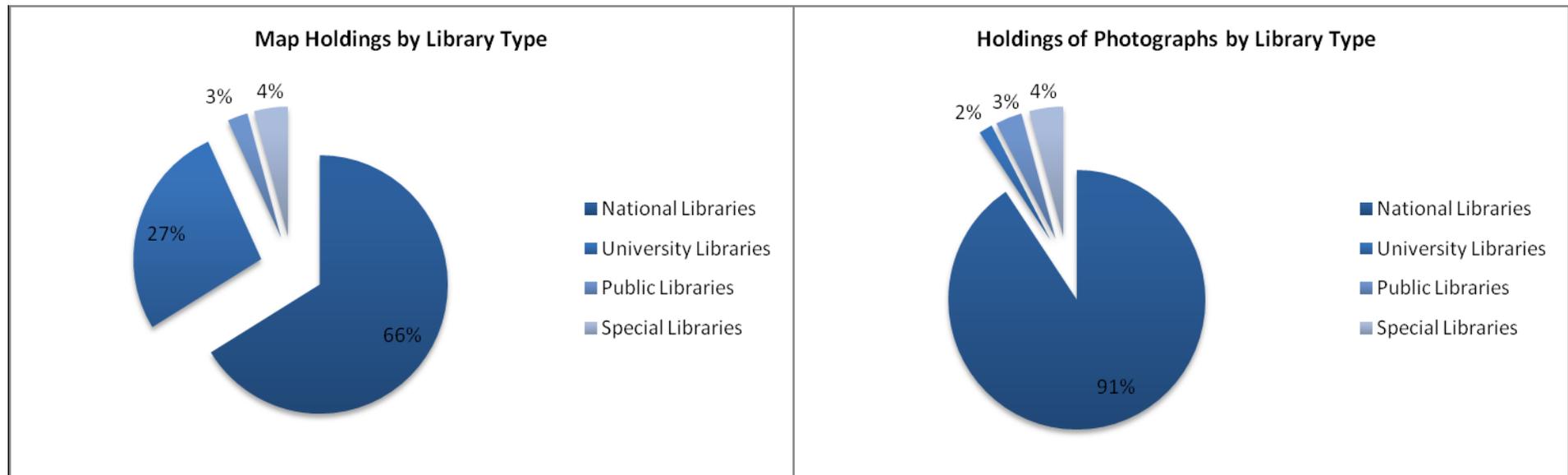
What digital content is available?

Holdings of European libraries

From: *The Cost of Digitising Europe's Cultural Heritage* (2011)

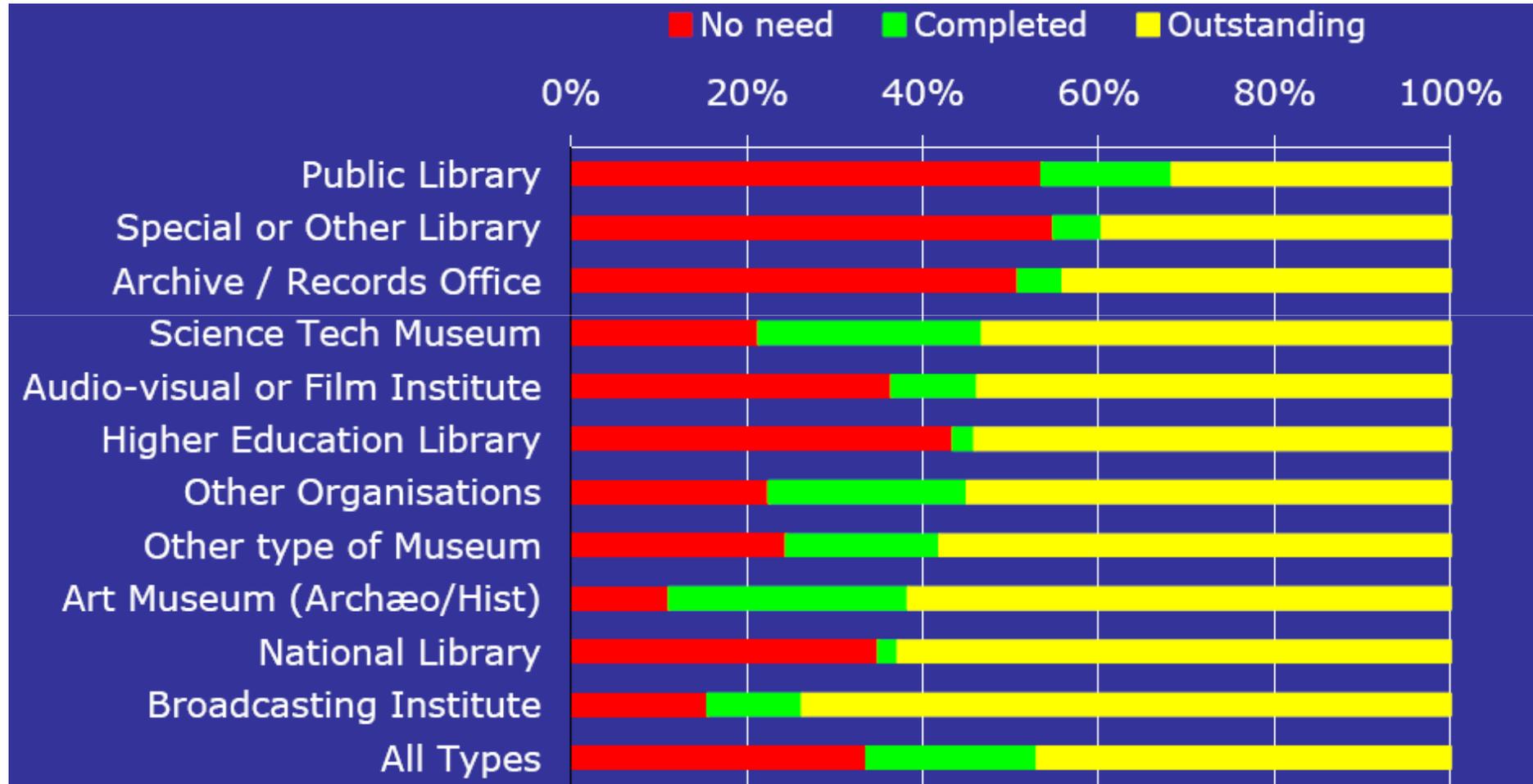
IFLA estimates in 2000⁷, for example, put the total holdings of European Libraries at approximately 16bn items (see table below), including 5.64bn books, at an estimated 1.41 trillion pages.

Media	Quantity (estimated)
Books	5.64 billion books, 1.41 trillion pages
Microform	0.02 billion units
AV	Up to 6.54 billion units
Other	Up to 3.69 billion units
TOTALS	15.89 billion units (+/- 15%)



From: Numeric Study Report (2009)

How much has been digitised?



From: Numeric Study Report

How much has been digitised?

Table 15 Progress made towards the digitisation of collections

Type of institution:	Part of collection digitised % [1]	'Order book'		Equivalent backlog [3]/[2] [4]
		Completed % [2]	Outstanding % [3]	
Archives	5.1	10.3	89.7	8.7
A-V or film institutes	9.8	15.4	84.6	5.5
Broadcasting institutes	10.8	12.8	87.2	6.8
Art/archæo museums	27.2	30.6	69.4	2.3
Science and tech museums	25.5	32.4	67.6	2.1
Other museums	17.5	23.1	76.9	3.3
National libraries	2.3	3.5	96.5	27.6
Higher education libraries	2.5	4.4	95.6	21.9
Public libraries	14.8	31.9	68.1	2.1
Special or other libraries	5.5	12.2	87.8	7.2
Other types of organisation	22.5	29.0	71.0	2.4

average 13%

From: *The Cost of Digitising Europe's Cultural Heritage* (2011)

<i>Description</i>	<i>Value</i>	<i>Units</i>
National Libraries in the Council of Europe	45	Sites
National Library service points	178	Service Points
Public Libraries	205336	Sites
HE/University Libraries	10161	Sites
School Libraries	164436	Sites
Special Libraries	29089	Sites
Google books estimated number of individual titles	130	Million titles
Lower estimate of titles currently in EU libraries	59	Million titles
Upper estimate of titles currently in EU libraries	95	Million titles
Mean estimate of titles currently in EU libraries	77	Million titles
Lower estimate of pages	1.47	Billion pages
Upper estimate of pages	2.36	Billion pages
Mean estimate of pages	1.93	Billion pages
Mean total cost for BOOK digitisation (in-house)	11.76	€bn
Mean total cost for BOOK digitisation (outsourced)	7.08	€bn
Mean total cost for BOOK digitisation (PPP)	4.79	€bn
Estimated number of RARE BOOKS	6.9	Million books
Estimated (mean) cost of RARE BOOK digitisation	8.62	€bn
Estimated (mean) cost of ARCHIVAL digitisation in libraries	2.45	€bn
Estimated pages of newspaper (based on 20m per MS)	540	Million pages
Estimated (mean) cost of NEWSPAPER digitisation in libraries	0.68	€bn
Estimated number of MAPS in library collections	7.23	Million maps
Estimated (mean) cost of MAP digitisation in libraries	61.03	€m
Estimated number of PHOTOGRAPHS in library collections	8.64	Million photos
Estimated (mean) cost of PHOTO digitisation in libraries	17.62	€m

TOTALS:

Estimated total cost of digitisation in LIBRARIES (in-house)	23.67	€bn
Estimated total cost of digitisation in LIBRARIES (outsourced)	18.95	€bn
Estimated total cost of digitisation in LIBRARIES (PPP)	16.70	€bn

How to determine the collection types?

The Input-Output problem

The problem:

How to measure digital output in a way that relates to the analogue materials in a meaningful way?

Input:

Analogue materials we want to digitise

Output:

The digitised data representing the analogue materials

Two main input-output issues:

- The classification of the analogue and digitised materials
 - The units with which the analogue and digitised materials can be measured
-

From: Numeric Study Report

Table 1 Classification of materials held in collections

Framework categories	Units
Archival records	<i>Metres</i> ¹
Rare books	<i>Volumes</i>
Other books	<i>Volumes</i>
Newspapers	<i>Issues</i>
Serials	<i>Volumes</i>
Manuscripts	<i>Number</i>
Sheet music	<i>Number</i>
Microforms/microfilms ²	<i>Number</i>
Maps	<i>Number</i>
Photographs	<i>Number</i>
Engravings/prints	<i>Number</i>
Drawings	<i>Number</i>
Posters	<i>Number</i>
Postcards	<i>Number</i>
Paintings	<i>Number</i>
Any other two dimensional object	<i>Number</i>
Three dimensional works of art	<i>Objects</i>
Man-made artefacts	<i>Objects</i>
Natural world specimens	<i>Objects</i>
Other objects in collections	<i>Objects</i>
Monuments ³	<i>Number</i>
Film	<i>Hours</i>
Video recordings	<i>Hours</i>
Audio (music & other recorded sound)	<i>Hours</i>
Other items not classified above	<i>Number</i>

Term (objects)	These terms also mean:	Suggested unit of measurement
<i>Archives</i>		
Archives (macro level)	Personal archives	Quantity
Records (micro level)		Quantity
<i>Other heritage collections: museum collections, library collections, historical-architectural collections, scientific collections and archaeological collections</i>		
Audio files (excl. digital audio books)		Hours
Databases (containing cultural heritage metadata)	Possibly with user-generated content	Quantity
Digital (3D) designs or reconstructions of objects and buildings		Quantity
Digital audio books		Titles
Digital maps and ground plans		Titles
Digital art objects/installations (incl. Internet art)		Works of art
Digital research files	Such as files from archaeological excavation projects	Files
E-articles	Single-file or composite (enhanced publications)	Quantity
E-books		Titles
E-newspapers		Titles and issues
Photos	Other born-digital images	Files
Games		Games
GIS files		Titles
Online newsletters		Titles (and issues)
Software (customized)		Programmes
Video recordings/Films	Films, animated films	Hours
Websites (and parts of websites)	Blogs, tweets, widgets, wikis	Number of sites
Other digital documents	E-flyers, reports, and such	Titles

Recommendations from workshop 27-1-2012

Definitions:

1. Promote a common understanding of what cultural heritage is
2. Define criteria of what information objects can be considered as digital cultural heritage

Context:

3. For heritage wide harmonisation, only generic, high level solutions can be applied. The practice across sectors is too diverse to change.
4. Measuring the actual status of digitisation should be related to understanding the total costs of ownership.
5. Understanding the developments in information management outside the cultural heritage institutions is needed in order to create new selection and acquisition policies for digital heritage.

Survey methodology:

6. The starting point for improving the methodology of measuring digital heritage collections should be the digital realm, not the physical one.
 7. Measuring digital collections requires a better determination of digital units (which may be simple or complex)
 8. Determining digital units can best be done with a hierarchical classification that expresses dimensions and relationships.
-