



**Thematic Survey Questionnaire –
SIZE of Digital Collections**

by

**Natasha Stroecker and René Vogels, Panteia (NL)
Gerhard Jan Nauta and Marco de Niet, DEN (NL)**



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

SIZE

ENUMERATE SURVEY ON THE SIZE OF DIGITAL COLLECTIONS

In this experimental thematic survey the aim is to develop better procedures for measuring the size of the digital heritage collections and the rate at which they are growing, both through the digitisation of analogue collections and the acquisition of new (born digital) materials.

We have developed three alternative approaches to measure the size of your collections. We ask each institute to fill out the questionnaire for two approaches, which we have selected at random for you.

It is important that the same person will try out the alternative approaches. At the end of the questionnaire we will ask you a few questions about these approaches: which one was most easy to handle (in terms of effort and time), etc.

Since the questionnaire aims to improve the methodology of monitoring the access and use of digital heritage collections, some questions have an accompanying field for evaluative remarks. Please help us with any comments that may be useful.

The time needed to fill in the questionnaire will depend highly on the availability of management information about your digital collections. If this information is in place, we estimate that answering the questions will take about 30 to 45 minutes of your time.

We hope you will be able to complete the questionnaire by **March 20th**. The questionnaire is designed in such a way that you can pause and return at any time to continue.

If you have any further questions, please contact your national coordinator, or:

Best Regards, the ENUMERATE Team

Natasha Stroeker: enumerate@panteia.nl

SECTION 1/5 - ORGANISATIONAL INFORMATION

[1] Name of institution / organisation *:

(Please provide the full, legal name of the company or institution about which you are answering the questions in this survey.)

[2] Type / Domain of institution / organisation *

Please specify the primary heading under which your company or institution classifies itself:

- National archive
- Other archive/Records office
- Audio-visual or broadcasting institute
- Film institute
- Museum of art
- Museum of archaeology, history
- Museum of natural history and natural science
- Museum of science and technology
- Museum of ethnography and anthropology
- Other type of museum
- National library
- Higher education library
- Special or other type of library
- Institution for Monument Care

[3] Your name:

(The full name and title of the person completing this survey.)

[4] Your telephone number and/or Skype contact details:

(The primary phone number (e.g. +44 0123456789) and/or the Skype details of the person completing this survey.)

[5] Your role in the institution:

(The position (job title) of the person completing this survey.)

[6] Your e-mail address:

(The primary e-mail address of the person completing this survey.)

[7] What is your institution's annual revenue budget?

Please choose only one of the following:

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

(Provide the annual budget for the *entire* institution, for instance as indicated in the last set of published accounts. The budget may include government funding, project funding, revenues from the organisation's primary or commercial activities, etc. If choosing between two of the options is difficult (like when the budget is exactly 50,000 €), please choose the option with the lower estimate.)

[8] Is curatorial care for the collections of your institution part of its mission?

- Yes
- No

(Answer this question with 'No' if your institution does not hold heritage collections or if you have collections (for example of books, films, music) to be lent by or sold to contemporary users without the explicit task of safeguarding the collections for future generations).

[9] Does your organisation have digital collections or is it currently involved in collection digitisation and digital born objects collecting activities?

- Yes
- No

(Only answer this question with 'no' if your institution does not have any collections of digital materials (including collections of digital born heritage materials), is not involved in any digitisation activities as to its heritage collections and does not have any established plans to start developing its digital collection in the near future/in 2013.)

SECTION 2/5 - COLLECTION REGISTRATION SYSTEM APPROACH

In this fairly straightforward approach the aim is to measure the size of your digital collections based on the quantities of descriptive metadata records in your Collections Registration System(s): How many objects are registered in the CRS? How many of these are available as digital born and digitally reproduced objects? (Everything that is not in the CRS need not be considered here.)

Before we start it is important to clearly mark the difference between (*item level*) *descriptive metadata* and *digital born and digitally reproduced objects*:

- **(item level) descriptive metadata:** the bibliographic and descriptive metadata needed to include item level records in your (online) catalogue.
- **digital born and digitally reproduced objects:** *digital surrogates*, created as a result of converting analogue materials to digital form (digitisation), and *born digital objects*, for which there has never been and is never intended to be an analogue equivalent.

In the first part of this section we ask you to briefly describe the Collection Registration System(s) in use in your organisation.

[10] What is the name of the CRS of your institution?

(The name under which your CRS is commonly known. If several CRSs are in use, please specify the names of all CRSs.)

[11] If your CRS is online available, what is the URL of the online access point?

(Please specify the URL for the online available CRS(s) of your organisation. Leave blank if there is no such access point.)

not available online

[13] What percentage of your collections is *not* registered in your CRS?

(Please estimate the percentage of your collections that has not been registered in descriptive metadata records.)

[14] What software runs the CRS?

(Please specify the CRS software. If several software solutions are in use, please specify all.)

[15] What is the total number of descriptive metadata records in your CRS?

(If several CRSs are in use please give the sum total of all metadata records in all CRSs.)

[16] What classification system or list of object types is in use to classify the objects recorded in your CRS?

(Think of standard lists like the Art & Architecture Thesaurus (AAT) Objects Facet, etc.)

[17] What is the total # of Gigabytes/Terabytes/Petabytes of your digital heritage collections as recorded in your CRS?

(This includes both the storage size of metadata and the corresponding digitally reproduced or born digital objects, in all available qualities (master files, derivatives, etc.). Please clearly indicate the unit of measurement (Gigabyte/Terabyte/Petabyte) here. If several CRSs are in use please give the sum total for all CRSs.)

[18] Please specify the quantities of the object types that are represented in your CRS

(Please start with the most numerous object types. Leave the remaining rows empty if you have only a small number of object types recorded in your CRS).

	Object type (e.g. "Photographs")	# (=number of descriptive metadata records	# of digitally reproduced or born digital objects
1.	<e.g. manuscripts>	<e.g. 60.000>	<e.g. 10.000>
2.
3.

SECTION 3/5 - DETAILED MEASURING OF DIGITAL SUB COLLECTIONS APPROACH (2)

In this second approach the aim is to measure the size of your digital heritage collections through a detailed quantification, starting from a pre-defined table of digitally reproduced and born digital objects.

Again it is important to keep in mind the difference between item level metadata records and digital born and digitally reproduced objects. If we ask you to specify the number of records, this concerns the records with related *digital born and/or digitally reproduced objects*.

[19] Please indicate the size of your digital heritage collections per object type (columns A and B) and assess the need to digitise for objects in your entire analogue collections (columns C and D).

(Use the table below to quantify the digital heritage collections of your institution.)

	Term ([digital] objects)	Units (#=Number of) [#Records means: # Metadata records in CRS]	Estimated # of units in entire digital collection (A)	Estimated % born digital of units in entire digital collection (B)	Estimated % of <i>entire</i> analogue collection still to be digitized (C)	Estimated % of <i>entire</i> analogue collection no need to be digitized (D)
ARCHIVAL RESOURCES						
	Archives : Government documents	# Records				
	Archives : Government archives	# Archives				
	Archives : Other archival records	# Records				
	Archives : Other archives	# Archives				
TEXT BASED RESOURCES						
<i>Books</i>						
	Rare printed books	# Records				
	Other printed books	# Records				
	Electronic books (eBooks)	# Records				
<i>Serials</i>						
	Newspapers	# Issues				
	Journals	# Issues				
	Articles	# Records				
	Other serials	# Issues				

<i>Other text based resources</i>						
	Medieval Manuscripts	# Records				
	Other Manuscripts	# Records				
	Microforms / Microfilms	# Records				
	Other text based	# Records				
VISUAL (2D) RESOURCES						
	Drawings	# Records				
	Engravings / Prints	# Records				
	Maps and ground plans	# Records				
	Paintings	# Records				
	Photographs	# Records				
	Posters	# Records				
	Sheet music	# Records				
	Other visual resources	# Records				
3D MOVABLE OBJECTS						
	3 Dimensional works of art	# Records				
	Archaeological man-made objects	# Records				
	Coins and medals	# Records				
	Other 3 dimensional man-made objects	# Records				
GEOGRAPHY BASED RESOURCES						
	Landscapes	# Records				
	Archeological sites	# Sites				
	Monuments and buildings	# Records				
	Other geography based resources	# Records				

NATURAL RESOURCES						
	Natural inert specimens	# Records				
	Natural living specimens	# Records				
TIME BASED RESOURCES						
	Audio files: Music	# Records				
	Audio files: Speech & other (excl. digital audio books; incl. oral history files)	# Records				
	Digital audio books	# Records				
	Film (Cinematic)	# Records				
	Video recordings	# Records				
	Other time based resources	# Records				
DIGITAL INTERACTIVE RESOURCES [EXCLUSIVELY DIGITAL]						
	Databases (containing cultural heritage metadata)	# Records				
	Digital (3D) designs or reconstructions of objects and buildings	# Records				
	Digital art objects	# Records				
	Digital research files (incl. GIS files)	# Records				
	Games	# Records				
	Software (customised)	# Records				
	Websites (and parts of websites)	# Records				
	Other digital interactive	# Records				

	resources					
--	-----------	--	--	--	--	--

SECTION 4/5 - IMPROVED ENUMERATE CORE SURVEY QUESTIONS APPROACH

The third approach chosen here is based on estimates of the size of your digital heritage collections and the number of objects that still need to be digitised. This is supplemented by two questions about the born digital materials collected by your institution.

Again it is important to keep in mind the difference between *(item level) metadata records* and *digital born and digitally reproduced objects*. (See above.)

[20a] Estimate the percentage of your heritage collections that has already been digitally reproduced:

... %

(Digitally reproduced should be discerned from being catalogued in digital (textual) metadata records. A digital reproduction is a digital surrogate of the original analogue object. An object that has been catalogued in a digital (textual) metadata record is not -digitally reproduced- in the definition we use here. Please give a global estimate (percentage) over all (analogue) heritage collections in your institution.)

[20b] Please specify the object types and percentages that were included in the previous answer [20a].

In the table below:

- (A) Choose all object types that apply, in other words: which object types have been digitally reproduced?
- (B) Please estimate how these digitally reproduced objects are distributed in percentages amongst the different object types.

(Example: if 30% of your heritage collections has been digitally reproduced and your institution has three different types of objects in the digital collections, we ask you to divide this 30% over the three object types, e.g. 10% - 10% - 10% OR 5% - 15% - 10% etc.)

[NOTE: Table will not be presented as it is presented here. Preferably the table can be folded out, starting from the high level object classes.]

Which apply (A)	Term ([digital] objects)	Estimated % digitally reproduced (B)
ARCHIVAL RESOURCES		
<input type="checkbox"/>	Archives : Government documents	
<input type="checkbox"/>	Archives : Other archival records	
TEXT BASED RESOURCES		
<i>Books</i>		
<input type="checkbox"/>	Rare printed books	
<input type="checkbox"/>	Other printed books	
<i>Serials</i>		

☐	Newspapers	
☐	Journals	
☐	Other serials	
<i>Other text based resources</i>		
☐	Medieval Manuscripts	
☐	Other Manuscripts	
☐	Microforms / Microfilms	
☐	Other text based	
VISUAL (2D) RESOURCES		
☐	Drawings	
☐	Engravings / Prints	
☐	Maps and ground plans	
☐	Paintings	
☐	Photographs	
☐	Posters	
☐	Sheet music	
☐	Other visual (2D) resources	
3D MOVABLE OBJECTS		
☐	3 Dimensional works of art	
☐	Other (3D) man-made objects	
GEOGRAPHY BASED RESOURCES		
☐	Landscapes	
☐	Archeological sites	
☐	Monuments and buildings	

<input type="checkbox"/>	Other geography based resources	
NATURAL RESOURCES		
<input type="checkbox"/>	Natural inert specimens	
<input type="checkbox"/>	Natural living specimens	
TIME BASED RESOURCES		
<input type="checkbox"/>	Audio files: Music	
<input type="checkbox"/>	Audio files: Speech & other	
<input type="checkbox"/>	Film (cinematic)	
<input type="checkbox"/>	Video recordings	
<input type="checkbox"/>	Other time based resources	
<input type="checkbox"/>	Other (please specify below)	

Other object types (not listed above) are:

(Free text description of digitised analogue objects, not listed.)

[21a] Estimate the percentage of your heritage collections that still needs to be digitally reproduced:

... %

(This is 100% - all of your collections - minus the percentage of your collections that has already been digitised, minus the percentage of your collections for which there is no need to be digitally reproduced. Digitally reproduced should be discerned from being catalogued in a digital (textual) metadata record. A digital reproduction is a digital surrogate of the original analogue object. An object that has been catalogued in a digital (textual) metadata record is not “digitally reproduced” in the definition we use here. Please give a global estimate (percentage) over all heritage collections in your institution.)

[21b] Please specify the object types and percentages that were included in the previous answer [21a].

In the table below:

- (A) Choose all object types that apply, in other words: which object types have been digitally reproduced?
- (B) Please estimate how these analogue objects that still need to be digitally reproduced are distributed in percentages amongst the different object types.

(EXAMPLE: if 70% of your heritage collections still needs to be digitally reproduced and you have three different object types in you collections, we ask you to divide this 70% over the three object types, e.g. 0% - 30% - 10% OR 20% - 15% - 35% etc.)

[NOTE: Table will not be presented as it is presented here. Preferably the table can be folded out, starting from the high level object classes.]

Which apply (A)	Term ([digital] objects)	Estimated % digitally reproduced (B)
ARCHIVAL RESOURCES		
<input type="checkbox"/>	Archives : Government documents	
<input type="checkbox"/>	Archives : Other archival records	
TEXT BASED RESOURCES		
<i>Books</i>		
<input type="checkbox"/>	Rare printed books	
<input type="checkbox"/>	Other printed books	
<i>Serials</i>		
<input type="checkbox"/>	Newspapers	
<input type="checkbox"/>	Journals	
<input type="checkbox"/>	Other serials	
<i>Other text based resources</i>		
<input type="checkbox"/>	Medieval Manuscripts	
<input type="checkbox"/>	Other Manuscripts	
<input type="checkbox"/>	Microforms / Microfilms	
<input type="checkbox"/>	Other text based	
VISUAL (2D) RESOURCES		
<input type="checkbox"/>	Drawings	
<input type="checkbox"/>	Engravings / Prints	
<input type="checkbox"/>	Maps and ground plans	
<input type="checkbox"/>	Paintings	
<input type="checkbox"/>	Photographs	

<input type="checkbox"/>	Posters	
<input type="checkbox"/>	Sheet music	
<input type="checkbox"/>	Other visual (2D) resources	
3D MOVABLE OBJECTS		
<input type="checkbox"/>	3 Dimensional works of art	
<input type="checkbox"/>	Other (3D) man-made objects	
GEOGRAPHY BASED RESOURCES		
<input type="checkbox"/>	Landscapes	
<input type="checkbox"/>	Archeological sites	
<input type="checkbox"/>	Monuments and buildings	
<input type="checkbox"/>	Other geography based resources	
NATURAL RESOURCES		
<input type="checkbox"/>	Natural inert specimens	
<input type="checkbox"/>	Natural living specimens	
TIME BASED RESOURCES		
<input type="checkbox"/>	Audio files: Music	
<input type="checkbox"/>	Audio files: Speech & other	
<input type="checkbox"/>	Film (cinematic)	
<input type="checkbox"/>	Video recordings	
<input type="checkbox"/>	Other time based resources	
<input type="checkbox"/>	Other (please specify below)	

Other object types (not listed above) are:

(Free text description of digitised analogue objects, not listed.)

[22a] Does your organisation collect born digital heritage materials?

Please choose only one of the following:

- Yes
- No
- Do not know

Answer this question with 'Yes' if your institution collects any kind of born digital heritage (i.e. software, digital documents, digital art, harvested web content, etc.).

[22b] If Yes, what types of born digital heritage?

Please specify the object types that were included in the previous answer:

Which apply (A)	Term ([digital] objects)
ARCHIVAL RESOURCES	
<input type="checkbox"/>	Archives : Government documents
<input type="checkbox"/>	Archives : Other archival records
TEXT BASED RESOURCES	
<input type="checkbox"/>	Electronic books (eBooks)
<input type="checkbox"/>	e-Newspapers
<input type="checkbox"/>	e-Journals
<input type="checkbox"/>	Other e-Serials
<input type="checkbox"/>	Other digital born text documents
VISUAL (2D) RESOURCES	
<input type="checkbox"/>	Digital maps and ground plans
<input type="checkbox"/>	Photographs
TIME BASED RESOURCES	
<input type="checkbox"/>	Audio files: Music
<input type="checkbox"/>	Audio files: Speech & other
<input type="checkbox"/>	Film
<input type="checkbox"/>	Video recordings
DIGITAL INTERACTIVE RESOURCES	

<input type="checkbox"/>	Databases (containing cultural heritage metadata)
<input type="checkbox"/>	Digital (3D) designs or reconstructions of objects and buildings
<input type="checkbox"/>	Digital art objects (incl. Internet art)
<input type="checkbox"/>	Digital research files (incl. GIS files)
<input type="checkbox"/>	Games
<input type="checkbox"/>	Software (customised)
<input type="checkbox"/>	Websites (incl. blogs, tweets, widgets, wikis)
<input type="checkbox"/>	Other (please specify below)

Other object types (not listed above) are:

(Free text description of born digital objects, not listed.)

SECTION 5/5 - EVALUATIVE REMARKS

These questions were different for each of the three options. Each respondent answers the evaluative remarks for two of the three approaches that are selected. Here approach 1 and 3 are selected.

In this final section we would like to know about your experiences in filling out the questionnaire. Therefore we kindly ask you how time consuming and 'easy' or 'hard' it was to fill out the two approaches tried out above and which one provides the best insight into the size of the (digital) collections of your institution.

[23] How much time / effort was needed to fill out the questions for the two different approaches?

APPROACH	# Minutes
COLLECTION REGISTRATION SYSTEM APPROACH	
IMPROVED ENUMERATE CORE SURVEY QUESTIONS APPROACH	

(Please estimate the number of minutes spent on each of the approaches. If information collected in either one of the approaches was re-used in one or both the other approaches, please try to account for that by distributing this time proportionally.)

[24] How difficult/hard was it to collect the necessary information to answer the questions in each of the two approaches?

APPROACH	1	2	3	4	5
COLLECTION REGISTRATION SYSTEM APPROACH					
IMPROVED ENUMERATE CORE SURVEY QUESTIONS APPROACH					

(Using a 5-points scale - where 1 equals "very difficult/hard" to 5 "not at all difficult/very easy" - please select only one number per row. Again, please try to account for the reuse of information, as in the previous question.)

[25a] Which one of the approaches would you prefer to fill out in future issues of the ENUMERATE Core Survey?

- COLLECTION REGISTRATION SYSTEM APPROACH
- IMPROVED ENUMERATE CORE SURVEY QUESTIONS APPROACH

(Please select only one of the alternative approaches.)

[25b] Why do you prefer the approach selected?

(Please explain you answer to question 25a.)

[26a] Which one of the alternative approaches will offer in your opinion the most faithful insight into the actual size of the (digital) collection of your institution?

- COLLECTION REGISTRATION SYSTEM APPROACH
- IMPROVED ENUMERATE CORE SURVEY QUESTIONS APPROACH

(Please select only one of the alternative approaches.)

[26b] And why?

(Please explain you answer to question 26a.)

[27] Please include any comments that would help the community of museums, libraries, archives and caretakers of monuments to improve the methodology of measuring the size and growth of digital heritage collections:

Thank you for collaborating with us in this Survey!

More information on the ENUMERATE Thematic Network is available on: www.enumerate.eu.