



Grant Agreement 270939

ENUMERATE

Documentation and Guidance Materials

DRAFT version,

Authors: Gerhard Jan Nauta, Sjoerd Bakker Marco de Niet



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

Contents

- 1 Summary..... 3
- 2 Introduction 3
- 3 Approach 3
- 4 Harmonisation and Validation tools 3
- 5 Question Specific Guidance notes..... 3
- 6 Definitions 4
- 7 Invitations and reminders 4
- 8 Feedback templates 5
- Appendix 1 – Question specific guidance notes (test survey) 6
- Appendix 2 – Terminology List 12
- Appendix 3 – Invitation e-mail 16
- Appendix 4 – URL communication e-mail..... 17
- Appendix 5 – Reminder e-mail 18

1 Summary

This document describes the nature and creation of the guidance materials and the documentation that was created to aid those people working on the survey, with the survey and those completing the survey.

2 Introduction

The success of the ENUMERATE project depends on a large group of people: the consortium, the national coordinators, the respondents and many others. To make sure all the processes run smoothly a certain level of guidance is required in all aspects of the project. For this purpose a set of guidance materials was created which will be disseminated through the proper channels, as well as appended to the present document. In this document the contents of each of the guidance materials will be briefly described.

3 Approach

For many of the sets of guidance materials those that were used by the NUMERIC project were taken as a foundation, but often the differences between NUMERIC and ENUMERATE called for the creation of a new set of guidance materials. The widely differing nature of each set of guidance materials requires

4 Harmonisation and Validation tools

As a part of deliverable 2.4 (Overview of harmonisation tools) an inventory of harmonisation tools was made. This inventory brought together relevant sources regarding terminology, cost models, collection type analyses and guidelines for web statistics.

This document serves two purposes. Firstly, the cost models, collection type analysis tools and the guidelines for web statistics that are catalogued offer respondents means to answer some of the questions in the surveys (i.e. calculate digitisation expenditure). Secondly, the terminology lists that are featured in this document have helped in adapting the NUMERIC terminology to a terminology that is suitable for ENUMERATE.

A publically available version of this document can be found her:

http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE_Tools_for_validation_and_harmonisation.pdf

5 Question Specific Guidance notes

The question specific guidance was created to explain questions to the respondents. The ultimate goal of this document is to ensure that there is only one possible interpretation for each question.

When applicable, the guidance notes from the NUMERIC survey were copied and these were adapted and updated when necessary. For most questions, however, there were no relating notes in the NUMERIC survey and these had to be written specifically for the ENUMERATE survey.

The construction of the guidance notes took place in close coordination with the core survey methodology construction. In this manner the interpretation of the questions by the person creating the guidance notes could easily be checked with those of the persons responsible for the core survey methodology. This proved to be a fruitful and mutually beneficent practice.

The guidance notes are incorporated in the online version of the questionnaire in the form of small pop-up windows that can be opened for each question if desired.

The Guidance notes are appended to this document (Appendix 1).

6 Definitions

The ENUMERATE list of definitions is largely a direct copy of the list of definitions that was used in the NUMERIC survey. The definitions that are not relevant to the present surveys have been removed and the terminology of the ENUMERATE project (which is nearly identical to that of NUMERIC) was studied closely in order to identify all new terms that were in need of clarification in a definition list.

When possible, the definitions for the terms that needed clarification were taken from the terminology sources that were catalogued in D2.4 (Harmonisation and Validation Tools). Great care was put in making sure that the most accurate definition was picked. Many of these terminologies were also in use by NUMERIC.

Some terms and concepts could not be clarified by using any existing terminology. For these terms tentative definitions were drawn up which will be checked as a part of the testing phase of the core survey.

The terminology list is appended to this document (Appendix 2).

7 Invitations and reminders

Invitation e-mails are sent to all selected institutions. In each country, the selection process is carried out by the national coordinators. A list of institutions that meet the requirements is then sent to DIGIBIS, where the invitation e-mail for each institution is provided with a unique link to the questionnaire.

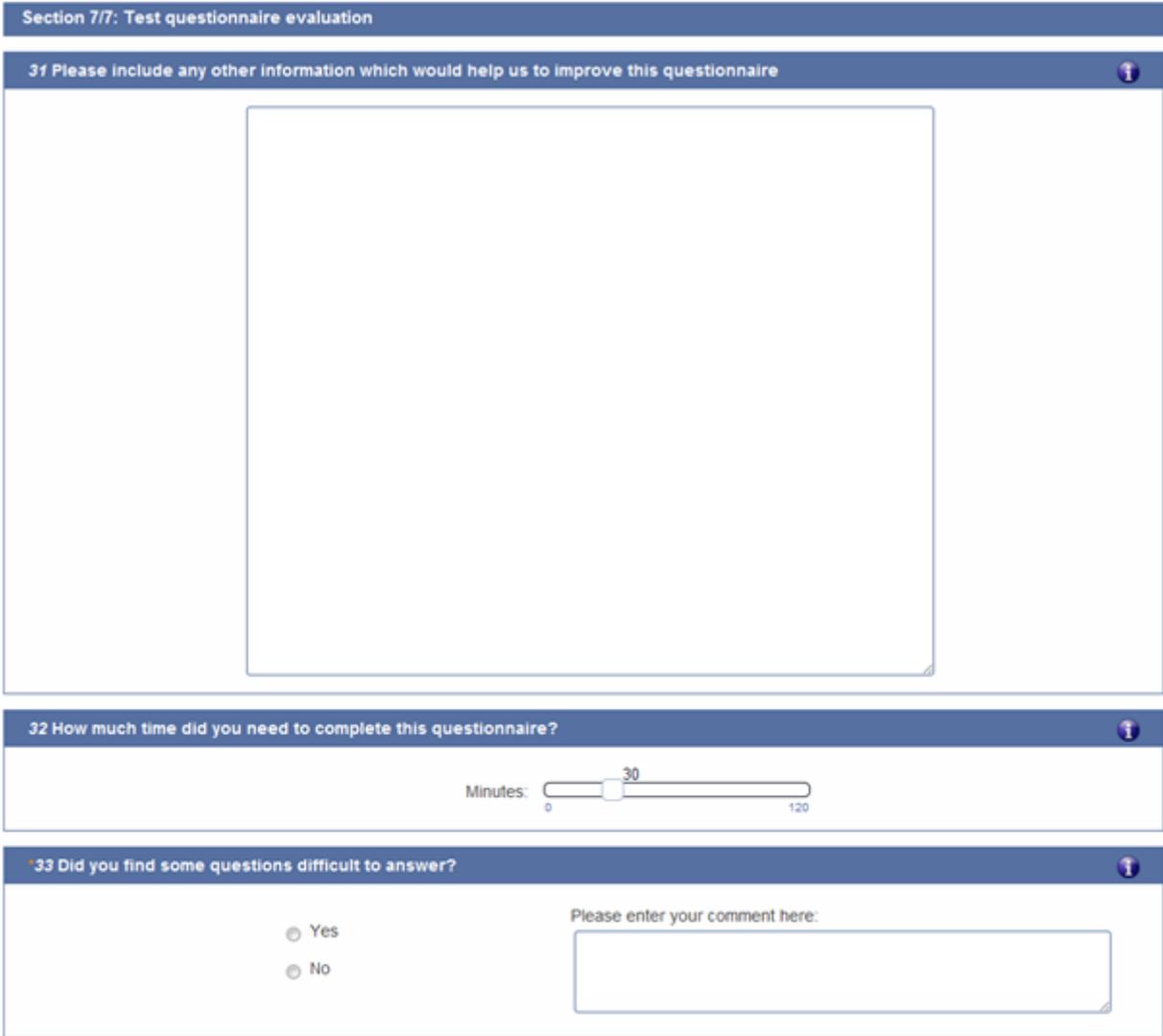
The invitation e-mails contain a brief description of the ENUMERATE project and are to be translated in as many languages as possible (or when the national coordinator indicates that this is necessary).

The reminders are to be sent at a predetermined time before the end of the response period of the survey.

The invitation e-mail (Appendix 3), the e-mail in which the survey URL is communicated (Appendix 4) and the reminder (Appendix 5) are appended.

8 Feedback templates

In the test survey three meta questions are included that are aimed at eliciting feedback from the respondent. One which can be used to provide any information that may be helpful in improving the questionnaire in the future, one to indicate the amount of time the respondent spent completing the survey and one to point out which question were considered to be difficult, and why.



A screenshot of the feedback screen from the test survey.

Appendix 1 – Question specific guidance notes (test survey)

Section 1: Organisational Information

Question 1:

Name of institution / organization

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

Question 2:

Type / Domain of institution / organization

- Specify the primary heading under which your company or institution classifies itself.

Question 3:

Country in which your institution is located

- Select the country

Question 4:

Website of your institution

- Provide the address of your institution's main website that is aimed at the general public.

Question 5:

Your department

- The name of your team or department. If your team or department does not have an official name, please indicate so.

Question 6:

Your name

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

Question 7:

Your role in the institution

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

Question 8:

Your e-mail address

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

Question 9:

Your telephone number or Skype contact details

- The primary phone number and/or the Skype details of the person completing this survey.

Question 10:

What is your institution's annual revenue budget?

- Provide the annual budget for the entire institution, for instance as indicated in the last set of published accounts.

Question 11:

Number of paid staff

- Include the total number of full time equivalents for all of your institutions employees including management positions.

Section 2: Digitisation Activity

Question 12:

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

- Please, only answer this question with 'no' if your institution does not have any collections of digital materials, is not involved in any digitisation activities and does not have any established plans start developing its digital collection in the near future (in 2012).

Question 13:

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

Explanatory notes for each of the 4 columns:

- Estimated number of analogue items: Provide the estimated number of analogue items that are part of the institutions own collections.
- Estimated % catalogued in digital (textual) metadata records: Provide the estimated percentage of analogue items that are only catalogued in metadata records and are not digitally reproduced.
- Estimated % digitally reproduced: Include projects that are approaching completion. Include materials that are recorded as images, even without character recognition.
- Estimated % of the entire collection with no need to be digitally reproduced:

The part of your collection that does not form an important part of the nation's cultural heritage, or is duplicated and had been (or will be) digitised by another institution. Digitisation of reference items will obviate the need to digitise everything.

Question 14:

Does your organisation collect born digital heritage?

- 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.). The manner of collection is not of importance in answering this question.

Question 15:

If yes: what types of born digital heritage?

- Use the classifications of born digital heritage that are in use in your institution. If no established typology is in use, provide the classifications that you think best describe your collections.

Question 16:

Does your organisation have a written digitisation strategy?

- The strategy may be for any period up to 2016.

Question 17:

Is your organisation included in a national digitisation strategy?

- Answer this question with 'yes' if your institution is actively involved in, or obligated to comply with, a formal, national or federal, cross-institutional policy regarding the digitisation of cultural heritage.

Section 3: Digital Access

Question 18:

Does your organisation have an explicit (written) policy to monitor access to digital collections?

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

Question 19:

Does your organisation measure the # of times digital metadata and/or digital objects are being accessed?

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

Question 20:

If yes, how?

- Indicate all ways in which the access of digital metadata and objects is measured.

Question 21:

Please indicate estimated percentage of your digital collections that is and/or will be accessible through the access options indicated in the left column.

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

Section 4: Digital Preservation**Question 22:**

Does your organisation have a written Digital Preservation Strategy?

- The answer of this question will be 'yes' when, in your institution, a formal document exists which details the policy regarding digital preservation.

Question 23:

Is your organisation included in a National Digital Preservation Strategy?

- Answer this question with 'yes' if your institution is actively involved in, or obligated to comply with, a formal, national or federal, cross-institutional policy regarding digital preservation.

Question 24:

Does your organisation have access to a Digital Preservation Infrastructure?

- This infrastructure does not have to be owned by your institution, it may also be shared with other institutions, hired from or located at external companies.

Section 5: Digitisation Expenditure**Question 25:**

Please estimate your expenditure on Digitisation:

- This sum is the amount identified in your own institution's annual budget to provide for digitisation work. Please attempt to include the cost of the staff time devoted to such projects in this estimate, but exclude external funding that is not included in your budget. If budget year does not coincide with the calendar, please choose the calendar year that fits best (in terms of the number of months). If a full summary of costs cannot be estimated, please indicate the amount that can be estimated and specify the items considered in the next question.

Question 26:

What is included in the budgets specified in the previous question?

- Check all expenses that apply in the left column and if possible give percentages per checked item in the column to the right. Please note that these percentages need not necessarily add up to 100%.

Question 27:

What is the total number of local staff engaged in the digitisation activities of your institution?

- Include the time of your own institution's staff engaged in digitization projects, including: planning and managing in-house and contracted projects; preparing and digitising materials; enhancing digitised output to widen accessibility.

Question 28:

What is the total number of volunteers (in full-time equivalent) engaged in the digitisation activities of your institution?

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

Question 29:

From what sources are your digitisation activities funded?

- Indicate all the sources from which your digitisation activities receive funds. Grants, subsidies and investments that you include

in your answer should be directly intended for digitisation efforts, all other (general) grants, subsidies and investments become part of your institution's internal budget.

Section 6: General Notes

Question 30:

Please include any other information which would help us understand the nature of your digitisation activities

- Comments on the questionnaire itself can be given in the next question (31).

Section 7: Test Questionnaire Evaluation

Question 31:

Please include any other information which would help us to improve this questionnaire

- Please comment on the quality and clarity of the questions and the explanations above. Suggestions on how to improve the questionnaire are highly appreciated.

Question 32:

How much time did you need to complete this questionnaire?

- Indicate the estimated time you spend on answering the questions including the time spend researching the answers.

Question 33:

Did you find some questions difficult to answer?

- If any questions were difficult to answer because the situation in your institution is incompatible with the options given in this questionnaire, please indicate this here.

Appendix 2 – Terminology List

Questionnaire Definitions (source: NUMERIC Study)

- ***Access** (ISO 5127) right, opportunity or means of obtaining information from documents.
- ***Access control** (NEDLIB) verification of user rights and the terms and conditions for the access to a publication.
- ***Access policy** (ODLIS) A formal written statement issued by the person(s) or body responsible for managing archives or special collections, specifying which materials are available for access and by whom, including any conditions or restrictions on use, usually posted or distributed by some method to users.
- ***Archive** (Pierce-Moses) The division within an organization responsible for maintaining the organisation's records of enduring value. An organisation that collects the records of individuals, families, or other organisations.
- ***Archival records** (tentative def. GJN, adapted from ISO 5127) documents created or received and maintained by an agency, organization or individual in pursuance of legal obligations, in the transaction of business, or in the course of the conduct of affairs, and preserved because of their enduring value. [GJN]
- ***Audio recording** (ODLIS) any medium on which sounds are recorded for mechanical or electronic playback, including phonograph records (vinyl), audiotape, and compact disc. Synonymous with sound recording.
- ***Audio visual document** (ISO 2789) document in which sound and/or pictures are prominent, and which requires the use of special equipment to be seen and/or heard.
- ***Book** (ISO 2789) non-serial printed document in codex form.
- ***Born digital/ Born digital heritage** (Jones) 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.
- ***Collection** (ISO 5127) gathering of documents assembled on the basis of some common characteristic, without regard to their provenance.
- ***Digital materials** (Jones) A broad term encompassing digital surrogates created as a result of converting analogue materials to digital form (digitisation), and "born digital" for which there has never been and is never intended to be an analogue equivalent, and digital records.
- ***Digital preservation** (ODLIS) The process of maintaining, in a condition suitable for use, materials produced in digital formats, including preservation of the bit stream and the continued ability to render or display the content represented by the bit stream. The task is compounded by the fact that some digital storage media deteriorate quickly ("bit rot"), and the digital object is inextricably entwined with its access environment (software and hardware),

which is evolving in a continuous cycle of innovation and obsolescence. Also refers to the practice of digitizing materials originally produced in non-digital formats (print, film, etc.) to prevent permanent loss due to deterioration of the physical medium.

***Digital preservation infrastructure** (tentative def. GJN) the basic physical and organisational structures and facilities (e.g., hardware, software, system management facilities, etc.) needed for the implementation of digital preservation.

***Digital preservation strategy (written)** (tentative def. GJN) a document, formally approved within an organisation, describing the way the organisation will be active in the preservation of her digitized and born digital collections.

***Digitisation** (IMLS) the process of converting, creating and maintaining books, art works, historical documents, photos, journals etc, in electronic representation so they can be viewed via computer and other devices.

***Digitisation strategy (written)** (tentative def. GJN) a document, formally approved within an organisation, describing the way the organisation will be active in the digitisation of her collections.

***Document** (ISO 2789) recorded information or material object, which can be treated as a unit in a documentation process.

***Drawing** (ISO 5127) picture made with a solid mineral substance or a pointed tool.

***Engraving** (ISO 5127) print made from any kind of intaglio plate, whether engraved with hand-tools or a machine, or etched with acid, so that the printing areas are lower than the non-printing areas.

***Film** (ISO 5127) series of pictures recorded on a strip of transparent material, or on an electronic data medium, which, when projected or produced rapidly one after another on a screen, give the illusion of natural and continuous movement.

***Full-time equivalent/FTE** (adapted from ISO 11620) a measurement equal to one staff person working a full-time work schedule for one year. EXAMPLE If out of three persons employed in an institution, one works quarter-time, one works half-time, and one works full-time, then the FTE of these three persons would be $0,25 + 0,5 + 1,0 = 1,75$ employees (FTE).

***Journal** (adapted from ISO 2789) serial under the same title published at regular or irregular intervals, over an indefinite period, individual issues in the series being numbered consecutively or each issue being dated. NOTE: Series of reports, transactions of institutions, series of regular conference proceedings and annuals are included, while newspapers and monographic series are excluded.

***Library** (ISO 2789) organization, or part of an organization, the main aims of which are to build and maintain a collection and to facilitate the use of such information resources and facilities as are required to meet the informational, research, educational, cultural or recreational needs of its users. NOTE

These are the basic requirements for a library and do not *exclude any additional resources and services incidental to its main purpose.

***Manuscript** (ISO 2789) original document that is handwritten or in typescript. NOTE: Bound volumes and other units (fragments, rolls, autographs, etc.) may be counted separately.

***Map** (ISO 5127) conventional representation, on a reduced scale and usually flat, of phenomena which can be localized in space and time.

***Metadata** (ODLIS) Literally, "data about data." Structured information describing information resources/objects for a variety of purposes....The term is generally used in the library community for non-traditional schemes such as the Dublin Core Metadata Element Set, the VRA Core Categories, and the Encoded Archival Description (EAD). Metadata has been categorized as descriptive, structural, and administrative. *Descriptive metadata* facilitates indexing, discovery, identification, and selection. *Structural metadata* describes the internal structure of complex information resources. *Administrative metadata* aids in the management of resources and may include rights management metadata, preservation metadata, and technical metadata describing the physical characteristics of a resource.

***Microform** (ISO 2789) photographic document requiring magnification when used. NOTE 1: Microfiche and microfilm are included. NOTE 2: Slides and similar documents are counted as audiovisual documents.

***Monument** (adapted from OECD and UNESCO) Historic monuments are fixed assets that are identifiable because of particular historic, national, regional, local, religious or symbolic significance. This includes architectural works, groups of buildings, works of monumental sculpture and painting, elements or structures of an archaeological nature, inscriptions, cave dwellings and combinations of features. The definition excludes objects in the collections of archives, libraries and museums.

***Museum (1)** (ISO 5127) organized collection of artefacts or naturalia of cultural or scientific interest, stored permanently for intended display.

***Museum (2)** (ISO 5127) organization or part of an organization responsible for collecting, preserving, and exhibiting museum documents.

***National digitization strategy** (Tentative def. GJN, MdN & SB) An official, national or federal, cross-institutional policy on the digitization of cultural heritage that is imposed upon the institutions that are included in the strategy. Institutions may take part in the formative process of such a strategy and, as such, have a role in enforcing the strategy.

***National digital preservation strategy** (Tentative def. GJN, MdN & SB) An official, national or federal, cross-institutional policy on the preservation of digitized cultural heritage and born digital cultural heritage that is imposed upon the institutions that are included in the strategy. Institutions may take part in the formative process of such a strategy and, as such, have a role in enforcing the strategy.

***Newspaper** (ISO 2789) serial, which contains news on current events of special or general interest, the individual parts of which are listed chronologically or numerically and usually appear at least once a week.

***Painting** (preliminary definition GJN) a work produced through the art of painting in oil, acrylic paint, watercolour

***Photograph** (ISO 5127) picture obtained by a process which fixes a direct and durable image on a sensitized surface by the action of electromagnetic radiation.

***Postcard** (ISO 5127) card for conveyance by post, often with a picture on one side.

***Poster** (ODLIS) A large single sheet of heavy paper or cardboard, usually printed on one side only, with or without illustration, to advertise a product/service or publicize a forthcoming event (meeting, concert, dramatic performance, etc.), intended for display on a bulletin board, kiosk, wall, or other suitable surface.

***Preservation** (ISO 5127) all measures taken including financial and strategic decisions, to maintain the integrity and extend the life of documents or collections.

***Print** (ISO 5127) copy of an image transferred to a sensitive material.

***Rare book** (ISO 2789) book published before 1800.

***Record** (ISO 5127) document created or received and maintained by an agency, organization or individual in pursuance of legal obligations or in the transaction of business.

***Serial** (ODLIS) A publication in any medium issued under the same title in a succession of discrete parts, usually numbered (or dated) and appearing at regular or irregular intervals with no predetermined conclusion.

***Sheet music** (ISO 5127) printed music issued without covers, whether actually printed on single sheets (pages) or not.

***Three-dimensional works of art** (Tentative def. SB) A work of art that has height, width and depth.

***Video recording** (ODLIS) electronic medium in which visual images, usually in motion and accompanied by sound, are recorded for playback by means of a television receiver or monitor. The

***Web analytics** (tentative def. Wikipedia) is the measurement, collection, analysis and reporting of internet data for purposes of understanding and optimizing web usage.

Appendix 3 – Invitation e-mail

Dear colleague,

Early in 2012 the EU-funded ENUMERATE network, a community of practice in the field of digital cultural heritage, led by Collections Trust in the UK, will carry out a short survey across the Europe.

It will gather data about the:

- Digitisation of collections;
- Access to digitised heritage collections;
- Preservation of digital heritage materials.

The results of the survey will be given back to the cultural heritage sector via an open data platform.

ENUMERATE is approaching a small selection of institutions willing to test the draft version of the online survey, which is intended to be distributed to several thousands of institutions.

It would be of great help to us if you were willing to test our survey, and give us any remarks that might improve it. We estimate that your research time needed to answer the questions will be about 30 minutes, possibly spread over a few days.

Your data will be integrated with the final survey, so you will not be asked to supply additional data later on.

If you are willing to participate, please send us (enumerate@digibis.com):

- Your name;
- The name of your institution;
- The type of institution (e.g. museum, archive, library);
- Your e-mail address;
- A telephone number.

Within a few days we will send you an e-mail containing the URL of the online survey and further details.

Please do not hesitate to contact us through the e-mail address below if you have any queries concerning the questions being asked or any other matter relating to this survey.

In anticipation of your kind attention and support, may we thank you in advance.

Best Regards, the ENUMERATE Team

E-mail: enumerate@digibis.com

More information about the work of ENUMERATE can be found here:

Web: <http://www.enumerate.eu/>

Twitter: <http://twitter.com/#!/EnumerateEU>

LinkedIn: <http://www.linkedin.com/groups/ENUMERATE-4042570>

Appendix 4 – URL communication e-mail

Dear colleague,

To memory institutions such as your own and policy makers it is a matter of mutual importance to measure the progress being made towards the digitisation of cultural heritage assets. Information about this can help to inform both your own institutional policies, as well as policies on a national and European level.

Early in 2012 the EU-funded ENUMERATE network, a community of practice in the field of digital cultural heritage, led by Collections Trust in the UK, will carry out a short survey across the Europe.

It will gather data about the:

- Digitisation of collections;
- Access to digitised heritage collections;
- Preservation of digital heritage materials.

The results of will be given back to the cultural heritage sector via an open data platform.

On behalf of the community of archives, libraries and museums we ask for your help.

Below is a link to an online survey we intend to distribute among several thousands of institutions. It would be of great help to us if you are willing and able to test it, and help us with any remarks that might improve the current set of questions. There are 32 questions. We estimate that your research time needed to answer the questions will be about 30 minutes, possibly spread over a few days.

[LINK TO THE QUESTIONNAIRE]

Please do not hesitate to contact us through the e-mail address listed below if you have any queries concerning the questions or any other matter relating to this survey.

More information on the ENUMERATE network is available on the website, and via Twitter and LinkedIn.

In anticipation of your kind attention and support, we thank you in advance.

Best Regards, the ENUMERATE Team

E-mail: enumerate@digibis.com

More information about the work of ENUMERATE can be found here:

Web: <http://www.enumerate.eu/>

Twitter: <http://twitter.com/#!/EnumerateEU>

LinkedIn: <http://www.linkedin.com/groups/ENUMERATE-4042570>

Appendix 5 – Reminder e-mail

Dear colleague,

May we kindly remind you of the ENUMERATE online survey test? About a week ago we have sent you an invitation to be part of the test panel for a EU funded survey in the domain of digital heritage, focusing on digitisation, (online) access, and digital preservation. We desperately hope that you are still willing and able to help us! According to our estimates doing the survey will take about 30 minutes of your time.

Any remarks that might improve the current set of questions would be much appreciated. Your responses will be saved and integrated with the results of the final survey, which we are going to do in February/March 2012 among thousands of European institutions, so don't worry: we will not approach you at that time to do the survey again!

This is a link to the online survey:

[LINK TO THE QUESTIONNAIRE]

Please do not hesitate to contact us through the e-mail address listed below if you have any queries concerning the questions or any other matter relating to this survey.

In anticipation of your kind attention and support, we thank you in advance.

Best Regards, the ENUMERATE Team

E-mail: enumerate@digibis.com

More information about the work of ENUMERATE can be found here:

Web: <http://www.enumerate.eu/>

Twitter: <http://twitter.com/#!/EnumerateEU>

LinkedIn: <http://www.linkedin.com/groups/ENUMERATE-4042570>