ENUMERATE

Grant Agreement 270939

ENUMERATE

Thematic Survey Schedule

Author: *Gerhard Jan Nauta (DEN)*

Document history

Date	Author	Status
30-09-2012	Gerhard Jan Nauta (DEN)	Draft
09-10-2012	Marco de Niet (DEN)	Review
10-10-2012	Natasha Stroeker (Panteia)	Final check

Contents

1 Introduction and overview

The work of ENUMERATE is centred around three surveys, which will be rolled out during the three years of the project. In the first year (early 2012), the focus was on a high level survey (Core Survey 1), as recommended by the Numeric Study Report, on which ENUMERATE builds its activities. In the second year of the project a Thematic Survey will take place, dedicated to the four main topics that ENUMERATE addresses: size of digital heritage collections; costs of digitisation of cultural heritage; use of digital heritage and digital preservation. While the Core Survey is of a quantitative nature, the Thematic Survey will focus on qualitative assessments and methodology development, especially for Core Survey 2, since in the third year, the Core Survey will be repeated, in order to create a reliable baseline of intelligence about digitisation of cultural heritage in Europe.

According to the Description of Work the Thematic Survey, as opposed to the two Core Surveys, is: "…more in-depth, (…) designed to assess digitisation, digital preservation, usage and cost in greater detail, in order to build up a more granular picture of overall activity and best practices. The thematic survey will use the original extensive Numeric questionnaire as a starting point, in order to guarantee a high degree of continuity between Numeric and ENUMERATE. However the methodologies for the thematic survey will be based on a thorough evaluation in ENUMERATE expert groups in order to remove the flaws from the Numeric questionnaire, as advised by the SIG-STATS." [from: ENUMERATE *Annex I to the Grant Agreement (Description of Work) for a Thematic Network*, Version 2.0 (2011-12-2)]

This deliverable describes the overall Thematic Survey schedule, i.e. the interrelated activities, scheduled in time, that are necessary to collect the information as indicated in the Description Of Work. It must ensure that the Thematic Survey rollout is completed effectively. The document is composed of five parts, detailing the steps necessary to do research on each of the four main themes of the ENUMERATE survey project, supplemented with what is necessary to have contextual information in each of the sub-studies. Below we will briefly explain the five parts. After that the parts will be presented in more detail, followed by a tabular representation of the various activities planned.

The main parties involved in running the Thematic Survey are: heritage institutions; National Coordinators; Panteia (the contracted partner); the ENUMERATE Team (in particular the Core Group: CT, DEN, Digibis, SPK); and the ENUMERATE Thematic Network as a whole.

The Thematic Survey will run from November 2012 to February 2013. After that a report summarising the sub-studies will be prepared, the deadline for this is the end of March 2013. Data collected in the Thematic Survey will be published on the ENUMERATE Website. This is scheduled for April 2013. The schedule for earlier activities is based on these deadlines.

1.1 Activities essential to the Thematic Survey as a whole

The Thematic Survey will not be a single questionnaire distributed among a large group of cultural heritage institutions simultaneously. For each theme and sub-study specific heritage organisations will be selected and approached to collaborate with the ENUMERATE Team and National Coordinators in the various EU member states. To make it easier to collect information, and to build up a reliable database of heritage organisations, the selection of institutions for each part of the TS will be based on a central source with contact information about European heritage institutions, preferably with the assistance of the National Coordinators. This central source will be based on the "scene setting" data that was collected in the NUMERIC and ENUMERATE Core Surveys (e.g. name of institution, contact person etc.).

In the Thematic Survey, much will depend on the commitment of the National Coordinators. The ENUMERATE Team will make efforts to create an atmosphere of involvement through consulting the National Coordinators about the themes they think can best be studied in their respective countries and about institutions in their countries that can be invited to participate in the TS.

1.2   Activities related to Theme 1: Size of digital collections

The purpose of this part of the TS is both to check and review data collected in Core Survey 1 and to further develop the methodology of quantifying the size and growth of digital collections in sufficient detail. This meeting will take on board the recommendations from the Specialist Meeting, held in The Hague on 30 January 2012. The focus will be on improving the problematic "optional table" in Core Survey 1. We will do this by applying three research methods. Approximately 7 countries will be involved in this part of the survey. In each country about 10-15 heritage institutions will be approached that will be asked to compare three ways of measuring the size of collections:

a. Measuring collection size and composition based on the records in Collection Registration Systems;
b. Detailed measuring in the three most important digital sub-collections of institutions;
c. High level, estimate-based measuring, using an improved version of the questions in the ENUMERATE Core Survey.

1.3   Activities related to Theme 2: Cost of digitisation

In this part of the Thematic Survey a method will be developed in order to get more detailed information on the cost of digitisation than could be collected through the questions in Core Survey 1 (refer to question 27 of the CS questionnaire). In the Specialist Meeting devoted to the cost of digitisation (London, 30 March 2012) it was suggested that a workflow approach - Activity Based Costing - should be the basis for the TS. The steps consist broadly speaking of designing a workflow on paper; then testing it on a small group of heritage institutions; making improvements to the initial design; followed by a test among a larger sample of institutions.

1.4   Activities related to Theme 3: Use of digital heritage

Here too the purpose is to improve and specify the questions related to accessing digital collections in the Core Survey. After Specialist Meeting 3 in Madrid (6th June 2012) essentially two approaches to monitoring access, use and impact of digital collections stand out. In one approach monitoring access is realised using a more or less extensive questionnaire, to be distributed among a sample of institutions. A drawback of this way of monitoring is that institutions are expected to have the data at hand – which is often not the case - and are willing to fill in an extensive questionnaire. The other approach is doing some sort of action research, i.e. working together with an institution in order to improve internal evaluation procedures. The activities scheduled in the TS aim at a combination of both approaches.

In the first part, the questionnaire approach, the ENUMERATE Team will select 30 institutions from a restricted number of EU member states (3 to 5), representative of the different institution types identified in the previous ENUMERATE and NUMERIC surveys. These institutions will be asked to participate in a thorough review of the Core Survey and other access related questions (based on work done by other surveys), with special attention to the usefulness of the questions and the feasibility of answering them.

In the action research approach either the same institutions or another sample (to be determined) will be asked to monitor a particular stream of digital collection information, viz. web traffic. Institutions will be asked to install web analytics software (probably Google Analytics) and monitor during a fixed period of time a small number of standard indicators. The museum sector has seen web analytics initiatives in recent years. ENUMERATE will especially do research on cross sectoral monitoring. In this part of the TS ENUMERATE seeks collaboration with partners in existing web analytics projects, such as Museum Analytics or Culture24.

1.5   Activities related to Theme 4: Digital Preservation

Monitoring digital preservation policies and practice is an activity that is still relatively new to the cultural heritage field. Therefore the ENUMERATE Team has decided, in consultation with a group of specialists that met in The Hague on 5 October 2012 to focus on the development of a small but

balanced set of survey questions. The approach will be a combination of desk research, the consultation of a focus group on the topic, and a synthesis of outcomes. The ENUMERATE Team has experienced that the need to monitor digital preservation in the domain of cultural heritage is emerging in several places. Realising synergy with other plans and projects will be part of the activities. The main activity, the focus group session, is scheduled in the month of January 2013.

2 Role of the National Coordinators

2.1 Involvement of the National Coordinators

As stated above, the commitment of the National Coordinators is crucial to acquire data and knowledge about cultural heritage institutions across Europe. For each of the themes in the Thematic Survey the ENUMERATE Team will discuss options for participation with a selection of the National Coordinators. In this way the workload for the National Coordinators can be shared equally. National Coordinators are invited to volunteer for a specific theme. Their main activities will be to identify relevant institutions, to provide contact details and, if they desire to do so, be involved in evaluation of the results from their country.

2.2 Database of contact details

From the data collected in the ENUMERATE Core Survey an initial database of contacts will be constructed by Panteia in consultation with DEN and Digibis. The design of the database will be fairly simple. It will contain contact information about the institution, their role in the ENUMERATE network so far and their role in the TS. This database is then presented to the National Coordinators, who are expected to select institutions or submit suggestions for new institutions to participate in the TS. Institutions that did not (yet) participate in the ENUMERATE Core Survey will be added to the database. This is an on-going activity, up to the end of December 2012.

2.3 Information about national policies and strategies

As was suggested during the 5th October meeting in Den Haag, we will also ask the National Coordinators about the state of the national strategy on digitisation and /or digital preservation in their country:
- Is there an established national preservation strategy in your country?
- What is the status of this strategy?
- Is there an established national digital preservation infrastructure in your country?
- What is the status of this infrastructure? etc

These questions are necessary to avoid misinterpretation of question about these issues by the individual respondts.

2.4 Schedule of activities

Activity	Start date	Due date	Actor
Contact database design and construction	01/09/2012	01/11/2012	DEN, Digibis, Panteia
involving the Nat. Coord.	15/10/2012	15/11/2012	DEN, Panteia
Selecting and proposing institutions	15/111/2012	15/12/2012	Nat. Coord
Processing input from the National Coordinators	01/11/2012	31/12/2012	Panteia

3 Theme 1: Size and Growth of digital collections

3.1 Methodology development

Acquiring in-depth information on the size and growth of digital collections is a real challenge for any monitoring initiative in the domain of cultural heritage. The varying size of memory institutions - with few or many staff involved in collection management and digitisation - and the circumstance that overall collection management information is often incomplete (the precise size of analogue and digital collections is often not known), causes ENUMERATE to choose a multifaceted approach. In a restricted sample of institutions quantifying digital collections is done in three slightly different ways. Purpose of this approach is to assess the variance that results from using different methods of counting. Below these methods are explained in more detail.

Theme 1a: Measuring based on the records in Collection Registration Systems

In this part of Theme 1 no effort will be taken to collect data other than what is registered in collections registration systems (CRS). This can relate to both physical and digital objects. For this part of the survey staff members with a responsibility for information management at the level of collections will be addressed. This approach will contribute to acquire a better understanding of and better definitions for measuring cultural heritage collections..

Theme 1b: Detailed measuring in the three most important digital sub-collections

In addition to this global way of measuring, the TS will invite institutions to measure the three most important digital collections in their holdings. As was suggested by the Specialist Meeting, it is worthwhile to try and approach the measuring from the digital perspective instead of the physcial perspective. This means we will use a similar table as the optional question #14 of the ENUMERATE Core Survey 1, but this time with a (draft version of a) classification of digital collections. Methodology development will consist of improving this classification, deciding on the final units to measure and improving definitions of digital collections. For this part of the survey staff members with a responsibility for (digital) collection curation will be addressed, who have in depth knowledge of the (composition) of the specific collections.

Theme 1c: High level, estimate-based measuring, using an improved version of the questions in the ENUMERATE Core Survey

The third part of the survey is a slightly extended version of the present Core Survey 1 questions about estimated percentages of digitised and to be digitised collections. In the questions where the respondent is asked to indicate the object types that are include in the estimates an additional qualifier - estimated percentages of the relative volume of these collections - will be asked. These questions will be sent to managers of the institutions, who should generic knowledge about the collections.

With these three approaches, partly chosen to deal with complications inherent to the different sizes of cultural heritage institutions, ENUMERATE uses a layered approach to collecting data. This approach was inspired by the recent ITHAKA S+R Survey on *Sustaining Digitized Special Collections*. ITHAKA S+R designed three separate questionnaires in order to collect data on the level of: 1. The institutional perspective; 2. Collections in the aggregate; 3. Single collections. The ITHAKA methodology provided a mechanism for passing the request to participate in the survey from a high (management) level in the organisation to more specialised levels, where there is probably a better knowledge of collection specifics.

Methodology development for Theme 1 runs from September to October 2012.

3.2 Desk research and check-up of Core Survey 1 data on the size of digital collections

In this phase of the survey the outcomes of the size and growth estimations from Core Survey 1 will be reviewed by a modest number of specialists in the various institution types. This review will be contribute to improvement of the classification of physical collections and their measuring units.

3.3 Guidance materials

As in the earlier Core Survey there will be both materials aimed at guiding the National Coordinators in working together with the institutions, and materials guiding respondents when providing information about their institutions. In principle questionnaires for Theme 1 will be made self-explanatory. For most of the Thematic Survey activities guidance materials will therefore be kept as simple as possible. The preparation of guidance materials and questions on a meta-level (questions about the actual work being done) is a task for DEN in collaboration with Panteia.

For Theme 1 guidance materials can in part be adapted from the Core Survey guidance materials. Questionnaire guidance will be limited to introductory remarks (covering letter) and an improved version of terms and concepts in the glossary. If feasible the tables where institutions can indicate the size of their collections and the degree to which they are (being) digitised, will be made expandable. New will be a table with a classification of (born) digital heritage collections. This will be based on work done by DEN in a previous national project, on the classification tools collected in the first year of the ENUMERATE project and on the input from the Specialist Meeting on this topic.

3.4 Translations

All survey materials will initially be developed in English. National Coordinators may choose for a translation of the questionnaire and guidance documents so that the target institutions selected can more conveniently understand it. To ease the work to be done and to reduce the workload for individual participants translations the ENUMERATE consortium partners are asked to translate the survey materials into the main European languages represented by the partners (English, French, German, Spanish). Digibis will be asked to support this translation process with their online tool used for the Core Survey 1.

3.5 Selecting institutions

Selecting institutions is an activity done in a collaboration of the ENUMERATE Consortium Partners and the National Coordinators. For Theme 1 about 7 EU member states will be involved. The National Coordinators in these countries are expected to select about 10 to 15 institutions each: n=105 (max.). In the Thematic Survey methodology document details will be given concerning the precise method of selecting the institutions.

3.6 Online survey construction

All three approaches to survey this theme will result in online questionnaires, to be developed by Panteia, in close cooperation with DEN, Digibis and the National Coordinators involved. This is done in order to stay close to later versions. For translation purposes, an offline version of the questionnaire - e.g. in the form of a spreadsheet - can be considered. In order to improve feedback some questions on a meta-level will be added to the online questionnaire. The online questionnaires will be mounted on a server hosted by Panteia, and it will be made accessible through the ENUMERATE website.

Online survey construction will be done in the month December 2012. The deadline for the publicaiton of the online questionnaire is 4 January 2013.

3.7   Online survey implementation

On behalf of the consortium, Panteia will compile a list of institutions that will be invited to participate. The actual survey will run from 4 to 31 January 2013. Coordination is a task for Panteia, in cooperation with DEN and DIGIBIS. Three tasks stand out here: maintaining the database of invited institutions, providing support to the National Coordinators, and monitoring and securing responses. Since institutions are asked to share information on quantities of materials in their analogue and digital collections there is a fair risk of faults in the data collected. Therefore in the implementation phase monitoring the incoming data will be an on-going activity and actions to correct where necessary or to ask respondents to clarify responses will be taken ad hoc. A more thorough cleaning and harmonising job is scheduled after the survey period.

3.8   Urging non-responders to respond

During the four weeks the survey will be online, the consortium contacts (National Coordinators) will be available to answer questions about the survey and if necessary urge non-responders to take action. This is a shared responsibility of the National Coordinators and the ENUMERATE Team.

3.9   Cleaning and harmonising returned questionnaires

Any corrupted or ambiguous data in the survey responses must be corrected or left out. This part of the work is scheduled between 28 January and 8 February 2013.

3.10 Data analysis

Data analysis and preparing the interim report is scheduled between 4 February and 1 March 2013.

3.11 Schedule (blueprint) for Theme 1

Activity	Start date	Due date	Actor
Theme 1 - draft methodology	01/09/2012	31/10/2012	DEN, Panteia
Review of methodology by ENUM. Partners	19/10/2012	19/10/2012	All ENUM. Partners
Theme 1 - desk research: check core survey data	22/10/2012	16/11/2012	DEN
Theme 1 - guidance materials	05/11/2012	16/11/2012	DEN
Theme 1 - translations	12/11/2012	07/12/2012	ENUM. Partners, Nat. Coord.
Theme 1 - selecting institutions	19/11/2012	14/12/2012	Nat. Coord.
Theme 1 - online survey construction	10/12/2012	04/01/2013	Panteia, Digibis
Theme 1 - survey announcing	31/12/2012	04/01/2013	Nat. Coord.
Theme 1 - survey implementation	04/01/2013	01/02/2013	Panteia
Theme 1 - urging non-responders to respond	28/01/2013	01/02/2013	Nat. Coord.
Theme 1 - cleaning/analysing questionnaires	28/01/2013	08/02/2013	Panteia
Theme 1 - data analysis	04/02/2013	01/03/2013	Panteia
Theme 1 - interim report	25/02/2013	01/03/2013	Panteia, Digibis

4   Theme 2: Cost of digitisation

4.1   Activity based costing methodology development

The schedule for Theme 2 consists of two separate activities, not overlapping in time. First a workflow model for calculating costs will be contructed based on existing examples. DEN will make the first version of this. workflow model. The model will be tested in 15 cases. After that the model is refined on the basis of an evaluation of the test results. The next step is using the improved cost model in a larger sample (n=90) of memory institutions.

4.2   Selecting institutions

Selecting institutions is done in close collaboration between the ENUMERATE Team and the National Coordinators. To realise the first part of the survey process the ENUMERATE Team needs 15 institutions that have some expertise in cost calculations or are engaged in digitisation projects in which the monitoring of costs is put to practice. To make the selection practical, three National Coordinators will be asked to assist in finding 5 institutions in their country to participate.

For the second part of Theme 2, to test the intermediate results in a larger sample, seven EU member states will be involved. The National Coordinators in these countries are expected to select about 10 to 15 institutions each: n=105 max. In the Thematic Survey methodology document details will be given concerning the precise method of selecting the institutions.

Selecting institutions can be done fairly early in the survey period; the deadline is 16 November 2012.

4.3   Test case phase workflow survey

Workflow test cases will be done between 19 November and 14 December 2012. Prior to that the survey instruments will be developed, but since the number of participating institutions in this phase is small, the construction of the instruments can run parallel to methodology development.

4.4   Workflow test case reviewing and synthesis

The outcomes of the test case phase will be important in a review of the initial workflow design. A new improved workflow model will be developed by DEN in the month of December (3 December 2012-4 January 2013).

4.5   Workflow survey implementation

The second part of the workflow survey is planned during the month of January 2013 (7 January to 1 February 2013).

Panteia will take care of constructing the workflow questionnaire for part two of the survey on cost. Whether the questionnaire will be digital (online or offline) or on paper is still open to discussion (refer to Deliverable 2.8, the methodology document), but since the aim is to have a tool that is both focussing on the most important costs in creating digital collections and easy to handle, the questionnaire must be simple and the actual implementation mode will not have an influence on the planning cycle.

4.6   Cleaning and analysing returned questionnaires

Any corrupted or ambiguous data in the survey responses must be corrected or left out. This part of the work is scheduled between 28 January and 8 February 2013.

4.7 Data analysis

Data analysis and preparing the interim report is scheduled between 4 February and 1 March 2013.

4.8 Schedule (blueprint) for Theme 2

Activity	Start date	Due date	Actor
Theme 2 - workflow meth. development	08/10/2012	31/10/2012	DEN
Theme 2 - selecting institutions	29/10/2012	16/11/2012	Nat. Coord.
Theme 2 - workflow test cases	19/11/2012	14/12/2012	Panteia, Nat. Coord.
Theme 2 - urging non-responders	10/12/2012	14/12/2012	Nat. Coord.
Theme 2 - workflow synthesis	03/12/2012	04/01/2013	DEN
Theme 2 - workflow survey	07/01/2013	01/02/2013	Panteia, Nat. Coord.
Theme 2 - urging non-responders	28/01/2013	01/02/2013	Nat. Coord.
Theme 2 - cleaning & analysing	28/01/2013	08/02/2013	Panteia
Theme 2 - data analysis	04/02/2013	01/03/2013	Panteia
Theme 2 - interim report	25/02/2013	01/03/2013	Panteia, Digibis

5 Theme 3: Use of digital collections

5.1 Methodology development

Measuring Use of digital collections will have a dual approach: one part will be more or less a conventional survey (the 'questionnaire version'), specifically aimed at improving and deepening the questions that were posed in Core Survey 1. This will be done with a small group of contributors (ca. 30), to be able to process more qualitative feedback. The methodology for this part of the survey will be finished by 31 October 2012.

The second part (the 'data analytics approach') is will be designed in close collaboration with one or more parties already active in the field of analysing web statistics of cultural institutions and their activity/presence on social media like Facebook, Twitter, and Flickr. No fixed agreements have been made yet, but it is the intention of the ENUMERATE project to plan a *Data Analytics feasibility* study as part of an action research project, during the months December and January. Proposed partners are Culture24 (UK) and MuseumAnalytics (NL).

5.2 Desk research and updating an overview of monitoring initiatives

In preparation of Core Survey 1 an inventory of surveys and different ways of monitoring evaluating and measuring the value, use and impact of digital collections was produced by the ENUMERATE Team. In recent years we have witnessed a lot of activity in this area - undoubtedly induced by the increasing importance attached to business models and accountability in the heritage domain. As many stakeholders are interested in a thorough overview, the ENUMERATE Team reconsiders the way the information is collected and published. It is the intention widen and deepen the overview based on the input from cultural heritage partners across Europe, using social bookmarking.

Updating the existing inventory is an explicit task in the months September and October 2012. After that the social bookmarking mechanism should be in place.

5.3 Guidance materials

As the sample in this part of the Thematic Survey will be relatively small, guidance materials will be limited. We plan for a close cooperation with 3 National Coordinators (i.e. 10 institutions per country) and guidance materials will only be needed in the phase of collecting information from the institutions.

Guidance materials can partly be adapted from the Core Survey guidance materials. Questionnaire guidance will be limited to introductory remarks (covering letter) and an improved version of terms and concepts in the glossary.

5.4 Translating the questionnaire

Depending on the choice of the National Coordinators, some translation work may be necessary. The time period for any translation work is 19 to 30 November 2012.

5.5 Selecting institutions

In the questionnaire version the ENUMERATE Team will select 30 institutions from three EU member states, representative of the different institution types identified in the Core Survey 1. These institutions will be asked to participate in a thorough review of the Core Survey and other access related questions [from other surveys], with special attention to the usefulness of the questions and the feasibility of answering them.

5.6   Online survey construction

For Theme 3 an online questionnaire will be used, where we intend to reuse questions and guidance materials from the Cor Survey. Where needed, additional other survey instruments can be developed.

Survey/questionnaire construction is a task for Panteia, in collaboration with DIGIBIS.

5.7   Survey implementation

As the survey sample is modest, the survey will be organised in a relatively short time span: from the end of December 2012 (24/12/2012) to 11 January 2013.

5.8   Urging non-responders to respond

This activity is scheduled in the week of 7 to 11 January 2013.

5.9   Cleaning and harmonising returned questionnaires

After having collected responses the data must be cleaned and harmonised. This activity is scheduled in the week of 7 to 18 January 2013.

5.10 Data analysis

Panteia will collect and collate the data as submitted by the institutions that are going to participate in this part of the TS. Analysing the outcomes of the workflow survey is a task for Panteia in close cooperation with DEN because of the domain knowledge that may be needed for the interpretation of the results. If necessary, feedback from the National Coordinators will be asked.

After having collected and analysed the survey outcomes a short report will be prepared by Panteia, summarising main outcomes. It will be reviewed by ENUMERATE Partners.

5.11 *Data Analytics* feasibility study and action research

In the action research version a small sample of institutions will be asked to monitor web traffic, using Google analytics, or some other web statistics program (to be determined). Institutions will be invited to install and monitor a small number of standard indicators over a fixed period of time. This part of the Thematic Survey is dependent on collaboration with partners in existing web analytics projects (pref. Museum Analytics (NL) and/or Culture24 (UK), who both took part in the ENUMERATE Speciliast Meeting on this theme.

The (provisional) time schedule for this part of the Survey is January-February 2013.

5.12 Processing *Data Analytics* research outcomes

Data analysis and preparing an interim report is dependent of the involvement of third parties and will be planned in dialogue with them.

5.13 Schedule (blueprint) for Theme 3

Activity	Start date	Due date	Actor
Theme 3 - (digital coll.) methodology	01/09/2012	31/10/2012	DEN, Panteia
Theme 3 - core survey data checkup	29/10/2012	16/11/2012	DEN
Theme 3 - guidance materials	05/11/2012	16/11/2012	DEN
Theme 3 - translations	19/11/2012	30/11/2012	Nat. Coord.
Theme 3 - selecting institutions	19/11/2012	30/11/2012	Nat. Coord.
Theme 3 - online survey construction	10/12/2012	21/12/2012	Panteia, Digibis
Theme 3 - survey announcing	17/12/2012	21/12/2012	Nat. Coord.

Theme 3 - survey	24/12/2012	11/01/2013	Panteia, Nat. Coord.
Theme 3 - urging non-responders	07/01/2013	11/01/2013	Nat. Coord.
Theme 3 - cleaning & analysing	07/01/2013	18/01/2013	Panteia
Theme 3 - data analysis	21/01/2013	15/02/2013	Panteia
Theme 3 - data analytics feasibility/methodology	03/12/2012	25/01/2013	DEN, [partner]
Theme 3 - analysis analytics data	28/01/2013	01/03/2013	DEN, [partner]
Theme 3 - interim report	25/02/2013	01/03/2013	Panteia

6 Theme 4: the Preservation of digital collections

As indicated in the ENUMERATE Description of Work (version December 2011) there is no statistically-valid basis upon which to benchmark Digital Preservation, or through which to normalise or harmonise issues such as cost and format in the context of digital preservation. In spite of considerable progress in raising the level of awareness, especially among governments and the more advanced and often larger heritage institutions, and in spite of the development of technically advanced models, standards and practical solutions, there have been disappointingly few initiatives to improve quantification methods. Specialists have therefore welcomed the idea of developing a framework for monitoring digital preservation policies, plans and activities in memory institutions with digital collections. Because of the relative immaturity of the area - it was estimated that the bulk of the memory institutions is not able/prepared to answer detailed questions on digital preservation issues - ENUMERATE has decided to follow a tiered model of information gathering, where in the initial phase much time and effort will be spent on making explicit the inside knowledge and opinions of experts. A focus group seems to be the best way to do this.

6.1 Desk research

The usual desk research will take place especially during the first months of the survey period, but it will be setup as a continuous community effort. References will be collected using an online social citation and bookmarking tool. (Presently Delicious and CiteUlike are in use, but the ENUMERATE Team plans to switch to Diigo.)

6.2 Scoping the measurement of Digital Preservation

From the desk research an overview of topics addressed in earlier surveys will be compiled. In addition, the results of a meeting on monitoring and measuring digital preservation practices in the EU, held in The Hague on 5 October 2012, will be taken into account. This meeting recommended to start recording awareness of international standards for digital preservation (such as OAIS and PREMIS), and to organise a Focus Group as part of the Thematic Survey, to bring together specialists that have set up surveys (mostly in European projects) and work together on definitions and a set of quantitative questions on digital preservation, that both large and small memory institutions will be able to address.

6.3 Formation of Focus Group on Digital Preservation monitoring

Early in November a list of experts to be invited to the focus group will be finalised and the focus group will be announced.

6.4 Focus Group preparation

Based on desk research, the results of the October meeting and a consultation of key figures in the field of digital preservation and curation an agenda for the Focus group will be prepared for January 2013.

6.5 Focus Group event

The actual focus group will be organised taking into account the agenda of the international digital preservation community. A possible event to align with is the *8th International Digital Curation Conference*, which will take place in Amsterdam, from 14 - 17 January 2013.

6.6 Focus Group report

Following the Focus Group event a concise report, presenting the principal outcomes of the meeting will be prepared. The recommendations of this report are meant to improve ENUMERATE Core Survey 2. The contents of the report will also be incorporated into the final Thematic Survey report.

6.7   Schedule (blueprint) for Theme 4

Activity	Start date	Due date	Actor
Theme 4 desk research	01/09/2012	31/10/2012	DEN
Theme 4 meth. development	15/09/2012	31/10/2012	DEN, Panteia
Theme 4 brainstorm session topics	05/10/2012	05/10/2012	DEN, Panteia
Theme 4 selecting participants	29/10/2012	16/11/2012	DEN
Theme 4 preparation focus group	10/12/2012	04/01/2013	DEN, Panteia
Theme 4 focus group	18/01/2013	18/01/2013	DEN, Panteia
Theme 4 focus group report	21/01/2013	01/02/2013	Panteia

7 Final report Thematic Survey and publication of survey data

7.1 Thematic Survey report

The Thematic Survey report will a compilation of the four sub reports discussed above. It is the second report in a series of ENUMERATE publications: Statistical Monitoring of Digitization of Cultural Heritage in Europe, and it will be published on the ENUMERATE website. The report presents the results of the Thematic Survey, and it will contain where possible references to Numeric Final Report, in order to show developments since the Numeric project. The report will be announced through all digital channels that ENUMERATE can employ.

7.2 The ENUMERATE Data Platform

The data collected in the Thematic Survey events will be anonymised and published on the ENUMERATE Data Platform. This will be ultimately one month after the publication of the Theamtic Survey Report.