TWINNING CONTRACT

JO/13/ENP/ST/23

Strengthening the capabilities of the Department of Statistics in Jordan

MISSION REPORT

on

Activity 3.8: Metadata strategy – II

Mission carried out by
Mogens Grosen Nielsen, Statistics Denmark
Karin Blix, Statistics Denmark

1-4 February 2015

Version: Final

STATISTICS DENMARK

DOS 1949 دائرة الإحصاءات العامة

Expert contact information

Mogens Grosen Nielsen
Statistics Denmark
Sejrøgade 11
DK-2100 Copenhagen Ø
Denmark
Tel: +45 3917 3894
Email: mgn@dst.dk

Karin Blix
Statistics Denmark
Sejrøgade 11
DK-2100 Copenhagen Ø
Denmark
Tel: +45 3917 3348
Email: kwb@dst.dk

Table of contents

List of Abbreviations

DoS	Department of Statistics of Jordan
ToR	Terms of Reference
BPM	Business Process Management
DDI	Data Documentation Initiative
GSBPM	Generic Statistical Business Process Model
GSIM	Generic Statistical Information Model
SDMX	Statistical Data and Metadata Exchange
SIMS	Single Integrated Metadata Structure
SMS	Statistical Metadata System

Executive Summary

The purpose of this mission has been to continue the discussions, develop and amend the Metadata strategy for DoS. The expected results are that recommendations are prepared for further developments and finalizaton of the draft metadata strategy based on international standards making use of Danish and the general European experiences. It is also expected that the work programme for the implementation of the metadata strategy is lined up.

The following recommendations have been given:

- It is recommended that the metadata strategy aligns with the overall strategy for DoS and the strategy on dissemination.
- It is recommended that DoS choose Nesstar for documenting variables.
- It is recommended that a link be given to the Nesstar documentation of variables in the quality declarations for surveys having completed documentation in Nesstar.
- It is recommended to repeating the exercise of mapping user needs before the Metadata Strategy for DoS is finally approved.
- It is recommended to go through the templates of Nesstar and the quality declarations to ensure that there is no double work.
- It is recommended to follow up on the introduction of the international standards in the organisation.
- It is recommended to use the SDMX definition of metadata. The definition can be found in missions report on mission 3.7
- It is recommended to work further on the template developed by Mr. Khalaf to ensure comprehensive guidelines for each phase that also covers the most important aspects of each phase in the GSBPM.

1. General comments

This mission report was prepared within the Twinning Project, ”Strengthening the capabilities of the Department of Statistics in Jordan”. It was the sixth mission to be devoted to Quality and Metadata within Component three of the project.

The purposes of the mission were:

- To continue the discussion on the metadata strategy
- To discuss and develop the amended draft metadata strategy prepared by DoS
- Discuss the use of the software Nesstar as the main tool for metadata

The consultants would like to express their thanks to all officials and individuals met for the kind support and valuable information which they received during the stay in Jordan and which highly facilitated the work of the consultants.

This views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of EU, DoS or Statistics Denmark.

2. Assessment and results

The purpose of this mission has been to continue the discussions, develop and amend the Metadata strategy for DoS. The expected results are that recommendations are prepared for further developments and finalizaton of the draft metadata strategy based on international standards making use of Danish and the general European experiences. It is also expected that the work programme for the implementation of the metadata strategy is lined up.

2.1 Metadata strategy and the overall strategy for DoS

The consultants had a fruitful discussion with the quality team on the metadata strategy. The first discussion was regarding the metadata strategy in connection with other strategies for DoS – like the overall strategy and the strategy for dissemination (which is actually not a strategy, but rather guidelines). It was stressed by the consultants that an alignment of the three would be important to have in mind. This might not be straight forward, as there is at the moment no overall strategy for DoS. The present overall strategy is going from 2008 till 2013 and the new one will be made in the course of 2015. It is recommended that the metadata strategy aligns with the overall strategy for DoS and the strategy on dissemination.

A presentation was given on the Danish Quality and Metadata Strategy. An overview of this can be seen in Annex 4.

2.2 Choice of tools

At the last mission in August 2014 for activity 3.7 the consultants where asked to introduce Colectica to DoS. This has now been abandoned by DoS as there is no possibility for financing of the full Colectica package in the Twinning project. It was actually the free version of Colectica – Colectica Express that was in question, but as DoS is already in possession of the DDI based system, Nesstar, DoS intend to go for this which has also been on the agenda for this activity.

Nesstar is based on DDI version 2.0. This is an older version of DDI than Colectica is based on. An important feature that is missing in this version of DDI is the possibility of having one solution for all metadata elements i.e. the possibility of combining e.g. the quality declarations and the metadata for the variables. Nesstar has so far been used on 12 surveys in DoS to describe in detail the contents of the variables in the surveys. As DoS at the present have no intentions to combine all metadata in one

system their need for software to handle metadata can, for the time being, be satisfied by using Nesstar. Nesstar is based on DDI 2.0 and it will be possible to migrate to newer versions of DDI. The consultants recommend DoS to start with Nesstar. This solution should be replaced by other software or newer standards as the need arises in DoS.

Choosing Nesstar will not give a software solution for the quality declarations in DoS. As part of this component quality declarations for 35 surveys have been disseminated on the DoS web page. These quality declarations have been completed using Word in a given template and transferred as PDF-files to the web page. This is a manageable and practical solution with the limited number of quality declarations DoS is completing. The consultants recommend that a link be given to the Nesstar documentation of variables in the quality declarations for surveys having completed documentation in Nesstar.

2.3 Engaging the organisation in the work with metadata

It is a challenge to engage the organisation in the metadata work. The consultants stressed the importance of engaging the (top) management. This cannot stand alone, but it is necessary to have the backing of (top) management for the process of metadata implementation to succeed. It was decided to arranging a meeting with the directors of DoS about the metadata strategy. What role should top management play and–what should be included in the strategy to ensure that their needs are or will be fulfilled. The consultants and the quality team prepared an agenda and a presentation including a set of questions for the group of directors to consider in order to get feedback from them. The agenda was sent to the directors before the meeting.

The questions asked were as follows:

- How can your directorate benefit from good metadata?
- How can we ensure cooperation throughout DoS with common metadata?
- What can you do in your directorate to minimize the burden of preparing the metadata?

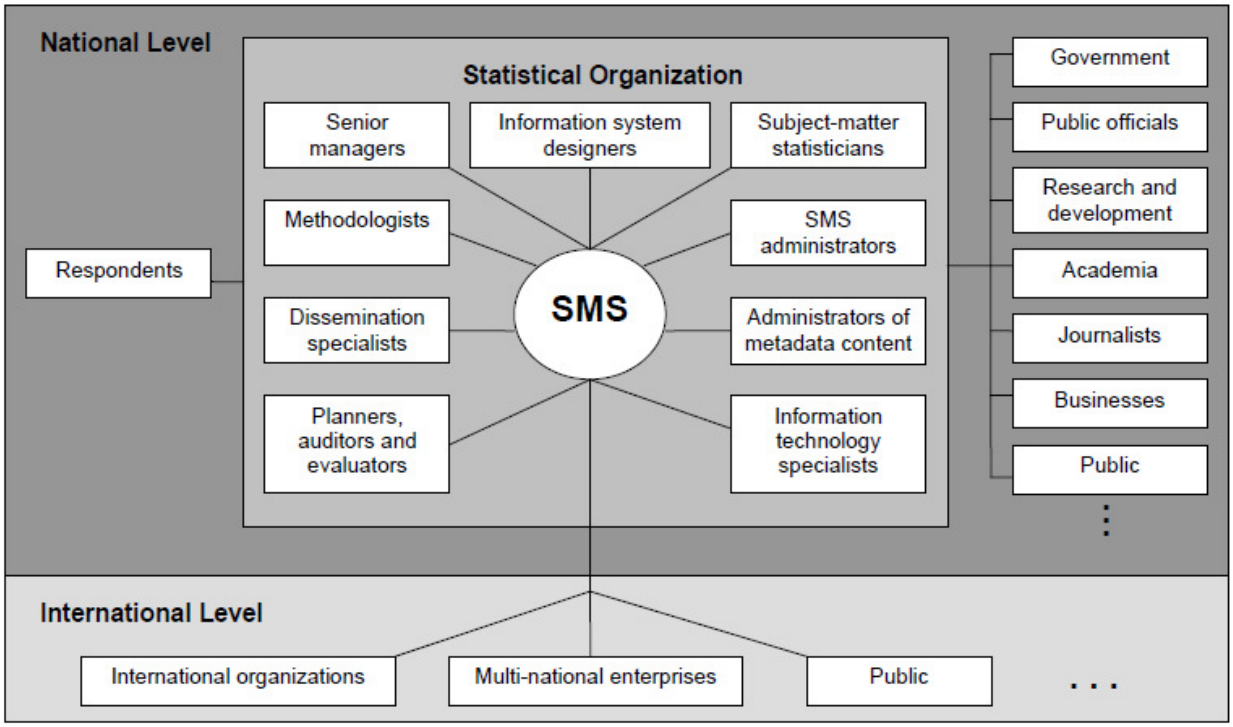
The impression from the meeting is that the directors see the importance of a metadata strategy and benefit of using international standards to ease the work on metadata. There was a common perception that it is important to make sure that the work with metadata be divided on more people so that more than one person in each directorate can handle the work on metadata.

One observation that is important to handle is that a definition of metadata has to be clear. It was agreed that the term metadata be used (and not documentation). It is recommended to use the SDMX definition of metadata. The definition can be found in missions report on mission 3.7

A new committee on metadata has been formed within DoS with Mr. Khalaf as a leader. This group will continue the work with Nesstar. The committee will have the work on quality in focus. It will be important for this group to be active in the organisation to make sure that the work on metadata has an important place in the consciousness of the people in charge of surveys.

2.4 How to fulfil users' needs for metadata

User needs in the organisation was discussed having the following figure as a starting point. For each type of user the need for metadata was prepared. This is not an easy discussion with clear answers. The consultants therefore recommend repeating this exercise before the Metadata Strategy for DoS is finally approved.



The following table shows the first mapping of user needs.

User	Needs
Subject matter	Detailed metadata on concepts, variable, understanding of the survey
Methodologist	Coverage, response rate, weights, variance for sample etc.
IT specialist	Questionnaire, variable, code-lists
Metadata administrator	All metadata for evaluation
Dissemination specialist	All metadata for user-friendly presentation
Auditors	Main metadata on quality and processes
Planners / project manager	All metadata to be used for project plan etc.
Management (second level)	Use of summary information e.g. quality declarations in order to make decisions and make new surveys
International organisations	Reference metadata. The central bank provides information
Government and government official	Information on quality (rarely). Sometimes via central bank
Research and academia	Description of content/methodology and variables
Private sector / businesses	Figures and summary descriptions
Media and public	Detailed and user-friendly information about content
Respondents	Need for goal and objectives – used in connection with interviewing

In addition the results above the following more simple categorization can be used.

Reference metadata

- Conceptual metadata. Implemented in summary form in Quality Declarations. Primarily for external users but also used internally
- Methodological and processing metadata: Implemented in process documentations including links to guidelines and methodological documents. Internal users

- Quality metadata. Implemented in Quality Declarations

Structural metadata
- Variables and code-lists for describing micro-data (entered into Nesstar). Internal users and researchers who have access to micro data.

2.5 Input from DoS to the Metadata strategy

DoS gave a presentation of the work DoS has made regarding the strategy on metadata since the last mission on activity 3.7 in august of 2014. The metadata strategy will run until 2020. Efforts will be made to describe the current situation - what is available what has been done previously. A link will be given to Jordan info – part of the DoS webpage. Here you can see the metadata for indicators only – social surveys, household survey etc. A link to NADA where there is metadata four example surveys in English and 12 in Arabic. The last example survey was created in 2012 - nothing is updated since then and only the minimum required metadata is filled in.

During the presentation it was also elaborated upon the choice of Nesstar as a metadata tool. Nesstar is a tool used internationally. At the current stage double work cannot be avoided in DoS. One aspect is that the project is awaiting a new webpage – which is part of another component in the present Twinning project.

In the case Nesstar is chosen there are five tasks to start working correctly with metadata
- Training technical staff
- Assigning one employee from each surveys responsible for using the software and make the different updates.
- Distributing software
- The arrangement
- Follow up improving adding new variables

Timetable for the metadata strategy
Stage one: Start – March 2015 – training stage of four months
Stage two: Installation and use – three months
Stage three: Dissemination – three months
Stage four: Follow up and update – continuous process hereafter

A presentation of Nesstar was given. Templates can be extracted from Nesstar. The consultants will receive a template with explanations. The consultants stressed the importance of only filling in information once. It is therefore important that any information that is reported in Nesstar is not repeated in the quality declarations or the other way around. The consultants recommend that the quality team goes through the Nesstar template and compare this with the quality declarations template to ensure that there is no double work. Guidelines exist in English – it needs to be translated to Arabic.

The consultants have looked more at the strategy. A short introduction was recommended added to the strategy giving an outline of what the strategy is about. Then the current situation will be elaborated upon. Most people in DoS actually do not know the current situation so a solid description of the current situation is in place. The consultants stressed the use of international standards. It is recommended to follow up on the introduction of the international standards in the organisation. This should cover training and information about relevant standards including the benefits of the international standards.

2.6 Analysis of processes

One important part of a metadata system is the analysis and documentation of the processes in the production of statistics. In activity 3.7 a template on agricultural statistics was developed to be piloted in the period between the two missions (this template can be seen in Annex 5). It was concluded that this template was too complex to be easy available for the subject matter statistical responsible in DoS. Mr Khalaf had developed a simpler template that would be simpler to access for the subject matter responsible (this template can be seen in Annex 6). This was discussed with the consultants. It was concluded, that the suggested template on a general level covers all the phases in the GSBPM. As the explanations or the guidelines given in the pilot were to complex DoS is recommended to work further on the template developed by Mr. Khalaf to ensure comprehensive guidelines for each phase that also covers the most important aspects of each phase in the GSBPM.

A presentation was given on the documentation of statistical processes. The documentation of processes is basically for internal use:

- For everyone involved to know the entire production process
- As a basis when changes need to me made
- For new employees – to understand the work process
- For benchmarking with other statistical units

It is recommended by the consultants to use GSBPM as a starting point for the documentation of processes. Why is GSBPM a good starting point?

- It covers all (thinkable) parts of a statistical production process
- It is used as a starting point for documentation of processes all around the world

3. Conclusions and recommendations

See executive summary.

Actions needed for moving forward as well as for preparing the next mission –add rows as needed.

Action	Deadline	Responsible person
1. Finalize the Metadata strategy for agreement with the consultant – the strategy will be translated to Arabic and hereafter Mr. Khalaf will send the Strategy to the management for approval.	1st March 2015	Mr Mohammad Khalaf

Annex 1. Terms of Reference

Terms of Reference

EU Twinning Project JO/13/ENP/ST/23

1 -4 February 2015

Component 3: Quality and metadata

Activity 3.8: Metadata strategy – II

0. Mandatory results and benchmarks for the component

- Statistics published by DoS are quality assured and documentation is improved (Apr 2015)
- Assessment report on current situation (Jan 2014)
- The consequences of applying the European Statistics Code of Practice in DoS are discussed (Apr 2014)
- Develop a standard format for a quality declaration (Jul 2014)
- A quality declaration for the national accounts is completed and published on the DoS website (Jan 2015)
- Develop a metadata strategy (Jan 2015)
- Design and test a quality audit (Apr 2015)

1. Purpose of the activity

- To continue the discussion on the metadata strategy
- To discuss and develop the amended draft metadata strategy prepared by DoS
- Discuss the use of the software Nesstar as the main tool for metadata.

2. Expected output of the activity

- Recommendations prepared for further developments and finalizaton of the draft metadata strategy based on international standards
- Transfer of the Danish and in general the European Union, experience in metadata systems
- A lining up of the work programme for the implementation of the metadata strategy

3. Participants

DoS
Mr Mohammad Khalaf, Head of Quality Division (*Component Leader*)
Mr Duraid Al-Shawawreh, Quality Division

Quality team:

Mr Bassam Al-Zain, Agricultural Survey Directorate
Mr Basem Shannek, Development & Strategic Planning Unit
Mr Mohammad Damrah, Economic Survey Directorate

MS experts
Mr Mogens Grosen Nielsen, Chief Adviser, Research Services, Statistics Denmark
Mr Karin Blix, Chief Adviser, Research Services, Statistics Denmark

Annex 2. Programme for the mission

Time		Place	Event	Purpose / detail
Sunday, morning	08.30 – 10.00	Hotel /DoS	Meeting with RTA	To discuss the programme of the week
Sunday, morning	10.00 – 12.00	DoS	Meeting with BC Component Leader and BC Experts	Discussions of the week's programme Presentation by DoS on the process and content of the draft metadata strategy
	12.00 – 01.00		Break / Preparations / Report writing	Break / Preparations / Report writing
Sunday, afternoon	01.00 – 03.30	DoS	Meeting with BC Component Leader and BC Experts	Continued discussions of the metadata strategy
	03.30 – 04.00		Preparations / Report writing	Preparations / Report writing
Monday, morning	08.30 – 09.00	DoS	Preparations / Report writing	Preparations / Report writing
	09.00 – 12.00		Meeting with BC Component Leader and BC Experts	Continued discussions of the metadata strategy
	12.00 – 01.00		Break / Preparations / Report writing	Break / Preparations / Report writing
Monday, afternoon	01.00 – 03.30	DoS	Meeting with BC Component Leader and BC Experts	Continued discussions of the metadata strategy
	03.30 – 04.00		Preparations / Report writing	Preparations / Report writing
Tuesday, morning	08.30 – 10.00	DoS	Preparations / Report writing	Preparations / Report writing
	10.00 – 12.00		Meeting with BC Component Leader and BC Experts	Continued discussions of the metadata strategy
	12.00 – 01.00		Break / Preparations / Report writing	Break / Preparations / Report writing
Tuesday, afternoon	01.00 – 03.30	DoS	Meeting with BC Component Leader and BC Experts	Continued discussions of the metadata strategy.
	03.30 – 04.00		Preparations / Report writing	Preparations / Report writing

Wednesday, morning	08.30 – 09.00	DoS	Preparations / Report writing	Preparations / Report writing
	09.00 – 12.00		Meeting with BC Component Leader and BC Experts	Continued discussions of the metadata strategy
	12.00 – 01.00		Break / Preparations / Report writing	Break / Preparations / Report writing
Wednesday, afternoon	01.00 – 02.30	DoS	Meeting with BC Component Leader and BC Experts	Continued discussions of the metadata strategy
			Ad-hoc meetings	Final clarifications with BC Experts, preparation of report and presentation for BC Project Leader
Wednesday, afternoon	02.30 – 03.00	DoS	Meeting with BC Component Leader	Presentation for BC Project Leader
Wednesday, afternoon	03.00 – 04.00	DoS	Debriefing with BC Project Leader	Conclusions and decisions and their consequences for the next activity and the implied work programme for BC Experts

Annex 3. Persons met

DoS:

1. Mr Mohammad Khalaf, Head of Quality Division (*Component Leader*)
2. Mr Duraid Al-Shawasreh, Quality Division
3. Mr Bassam Al-Zain, Agricultural Survey Directorate, member of Quality Team
4. Mr Basem Shannek, Development & Strategic Planning Unit, member of Quality Team
5. Moawiah Zahlol, Director National Accounts
6. Shaher Alshawabkeh, Director Economic survey
7. Salam Quteashat, Director Finance and analysis,
8. Adel Gharaibeh, Director IT
9. Mohammad Jundi, Director Household survey
10. Mohammad Thayyan, Director Methodological statistics,

External stakeholders:

RTA Team:
Thomas Olsen, RTA

Annex 4. The Danish Quality and Metadata Strategy

A presentation was given on the Quality and Metadata Strategy for Statistics Denmark. This was approved by the management just two weeks ago. This strategy could serve as inspiration to the DoS' strategy on metadata.

The starting point for the Quality and Metadata Strategy for Statistics Denmark is the Code of Practice for European statistics (CoP) and its implementation framework, the Quality Assurance Framework for European statistics (QAF). The overall principle for quality is "fitness for use" – i.e. if our processes do not generate products that give value for our users, they have no quality. This is the measurement for all our products and processes.

The commitment to quality in principle 4 of the CoP is in Denmark fulfilled by the following:

4.1 Commitment to quality
Quality policy: published at www.dst.dk
Organisational structure: Quality Committee, Quality Coordinator, Quality and Metadata unit.
Quality guidelines: prepared and linked to GSBPM

4.2 Statistical processes
Documentation and review of processes for selected surveys

4.3 Statistical products
Quality reporting prepared, assessed and published for all surveys

The overall vision for quality and metadata in statistics Denmark

- The overall target is to fulfil both internal and external user's needs for metadata regarding our products depending of their use of statistics;
- That the quality of our products and processes is described and fulfil the quality demands given;
- To achieve internal efficiency so that employees in Statistics Denmark can use metadata (on products and processes) when solving their tasks by using metadata that are maintained continuously in one system and is available for everyone.

International standards that will be in use to achieve the quality and metadata strategy of Statistics Denmark are

GSBPM - Generic Statistical Business Process Model
: This standard was developed by the UNECE in 2009. Since 2009 this model as become the corner stone in organisations producing statistics.

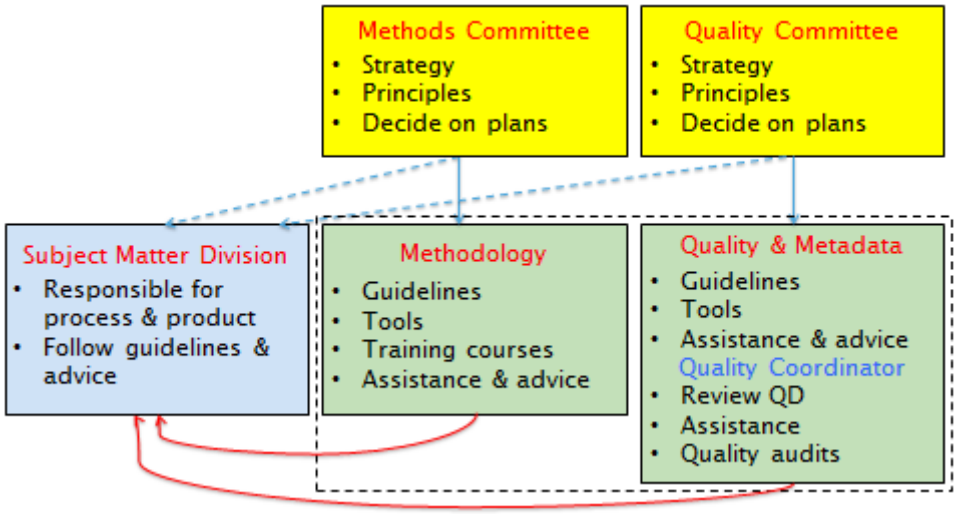
GSIM - Generic Statistical Information Model
: The purpose of this model is to describe information objects that are used and created in the single processes of GSBPM

DDI - Data Documentation Initiative
: This standard is connected to the business processes in which the statistics and its metadata are produced.

SDMX - Statistical Data and Metadata eXchange
: This standard is used by the international statistics society in connection with the exchange of macro data (tables) and the related metadata

Quality audits will be carried out based on the CoP. It will be performed by a team of experts, starting out with a self-assessment where examination of documentation will play a central role. A report will be submitted to the management followed up by an action list. The ambition is to carry through audits of 3-5 surveys per year.

The organization of the quality and metadata work in Statistics Denmark can be illustrated in the following figure:



Annex 5. Analysis of existing processes on Agriculture "as-is"

Please note that many descriptions are on at outline level and should be completed
Phase 1 Needs

Elements	Description of element
Description of the work processes	In this phase subject matter statistician prepares document with user needs and objectives and a work plan. The work plan is approved by management. (just an example)
Input	Requirements from Ministry of Agriculture. Human resources centre.
Output	Decisions / project plan / Objectives of the survey (opportunities)
People involved	Management, subject matter statistics
Regulations and guidelines	Terminology from FAO
It applications	

Phase 2 Design

Elements	Description
Description of work-processes	Project manager assigns task for each employee. - Methodologist: Sampling, concept definition - Enumerator / receptor: collect - IT-staff: - Administration:
Input from phase 1 Needs	Decisions / project plan / work plan Objectives of the survey (opportunities)
Output	Methodology information Questionnaire Sampling design, Guidelines for interviewers (enumerators, supervisors) Concepts and definitions
People involved	Methodologist Data control quality divisions Public relations Dissemination division IT-staff – hard copy IT division for electronic dissemination
It application / metadata	Word Excel

Phase 3 Build

Elements	Description
Description of work-processes	It specialist receives excel and other input and based on this he builds / configures oracle applications. Survey manager output tables approves output tables
Input from phase 2 Design	Methodology information Questionnaire Sampling design, Guidelines for interviewers (enumerators, supervisors) Concepts and definitions
Output	Oracle applications (configuration) Output tables
People involved	IT-specialist
It application / metadata	

Phase 4 Data collection

Elements	Description
Description	Steps: Train employee, Test questionnaire Print questionnaires (DoS) Collect data Personal interview (according to sample) First check(audit) – field check and correction (supervisor) Second check (audit) – DoS check and correction (Validating team) Code the questionnaire Data entry (data entry team)
Input	Questionnaires, Oracle applications
Output	Raw data in Oracle Report from audit
People involved	Enumerators Supervisor Data entry team Validating team
It application	This application takes care of: enter data using “user interface”, use discoverer for data edit. Use Excel for import and export

Phase 5 Data processing (integrated with data collection)

Elements	Description
Description	Data check – by agriculture and possible improved collection (macro editing)
Input	Raw data in Oracle
Output	Oracle database with clean data
People involved	It department Subject matter agriculture
It application / metadata	Oracle Discoverer (for macro editing)

Phase 6 Analysis

Elements	Description
Description	Expenditure – detailed analysis General: Spss: Subject matter define variables and labels
Input	Clean data in Oracle
Output	Aggregated data - paper Descriptive analysis – paper Maps
People involved	IT Subject matter agriculture
It application / metadata	Oracle Word Excel Spss

Phase 7 Dissemination

Elements	
Description	Reports written Approved by DG It publish on the web-site Agenda for publishing on the web- Two divisions: public relations and it
Input	Aggregated data - paper Descriptive analysis – paper Maps
Output	Tables on internet page (IT-specialist) Bulletins (extracted hard copy) Reports Data on CD Quality Declarations
People involved	Subject matter agriculture IT-specialist
It applications	Oracle Word Excel Homepage software

Annex 6. DoS template for process analysis

GSBPM	**Survey design** Specify needs	**Instruments and tools** Design	**IT** Build	**Sample selection** Collect (4.1)	**Data collection** Collect (4.2-4.4)	**Data_analysis** Process, Analysis	**Data_dissimenation** Disseminate	**Feedback** Evaluate
Input	customer needs ngos needs	objectives	questionnaire design on pda or tablet	objectives	field survey team	data cleaning	dummy tables	getting users feedback
Processing	preparing draft proposal	preparing questionnaire	preparing templates and frames testing design s	sampling and sample methodology	filling questionnaires using pds or tablet	data analysis	publishing tables on web & hard copies	improving
Output	signing agreements	questionnaire	questionnaire on sets	sample	collected survey data	survey database & dummy tables	dessiminated results	repeating survey
guidelines								