

# **TWINNING CONTRACT**

**JO/13/ENP/ST/23**

## **Strengthening the capabilities of the Department of Statistics in Jordan**



## **MISSION REPORT**

**on**

### **Activity 4.3: Plan for database structure - II**

Mission carried out by

Mr. Leonardo Tininini, Istat  
Mr. Lars Knudsen, Statistics Denmark

10<sup>th</sup> to 14<sup>th</sup> March 2014

Version: Final



Expert contact information

*Mr. Lars Knudsen  
Statistics Denmark  
Sejrøgade 11  
DK-2100 Copenhagen Ø  
Denmark  
Tel: +4539173318  
Email: lak@dst.dk*

*Mr. Leonardo Tininini  
Istat  
Via Cesare Balbo 16  
IT-00184 Roma  
Italy  
Tel:  
Email: tininini@istat.it*

## Table of contents

1. General comments.....	4
2. Assessment and results.....	4
3. Conclusions and recommendations .....	6
Annex 1. Terms of Reference .....	8
Programme for the mission .....	10
Annex 2. Persons met.....	12

## List of Abbreviations

DoS	Department of Statistics of Jordan
Istat	Italian National Institute of Statistics
Dst	Statistics Denmark
ToR	Terms of Reference
SDS	Statistical Dissemination System
CNMM	Common Nordic Meta Model

## 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 third mission to be devoted to planning database structures within Component 4 of the project.

The purposes of the mission were:

- The data model used for the case study "Information Technology in homes and Employment" reviewed and refined
- Data from the case study for dissemination purposes is prepared (reconciled)
- The case study data for loading into the Statistical Dissemination System (SDS) is aggregated and prepared
- The model for the classification repository is reviewed
- The principles of GIS in SDS is presented and discussed
- Prepare a plan for the database structure
- Transfer of the Danish and in general the European Union, experience regarding database structure and online dissemination
- A lining up of work programme for the next activity (4.6, scheduled for the 21. – 25 September)

The consultant 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 consultant.

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

## 2. Assessment and results

The list of activities and discussions that took place during this mission is presented below in further detail. The main findings are summarized here according to the mission purposes above.

### **The model for the classification repository is reviewed**

As already stressed in the previous mission, merging data from different surveys and/or different rounds of the same survey often requires a preliminary reconciliation process, which can be supported by the adoption of a classification repository. In preparation to this mission the experts exchanged several e-mails with DoS to reach an agreement on the definition of a reconciliation framework based on a classification repository, that acts as a reference for all micro data to be reconciled. The repository data model was further reviewed and partially refined during the mission to cope with some special cases, e.g. those related to multi-response questions.

### **The data model used for the case study "Information Technology in homes and Employment" reviewed and refined**

Specific mappings were defined between micro data columns of the agreed surveys ("Employment and unemployment" and "ICT usage") and classifications in the repository, to formally define their meaning and interrelationships. Although the reconciliation framework was positively tested only on these two specific surveys, its realm of applicability obviously goes far beyond the specific case studies. The proposed framework does not require the original data models to be restructured for reconciliation, but rather builds a reconciliation layer (constituted by mappings and views) on top of the original (unchanged) data models. The "new" data model, based on this reconciliation layer of mappings and views, was positively tested and refined on the proposed case studies.

**Data from the case study for dissemination purposes is prepared (reconciled)**

The variables/classifications agreed before the mission were reconciled according to the proposed framework. It was stressed again by the experts that even if a reconciliation process is technically possible, this does not necessarily imply that such reconciliation is meaningful from the statistical point of view (and hence practically feasible). Reconciling data requires a deep knowledge of the corresponding survey(s), units of analysis, questionnaires and more generally domain of interest. A knowledge that is typically owned (only) by the statistical experts that were directly involved in the design as well as in the various production phases of the surveys themselves. All situations proposed in the case studies (with increasing level of complexity) were addressed and solved.

**Case study data in the SDS**

From the reconciled data on ICT a table structure for a specific cube (maintable called “ICT11”) for dissemination was decided. Three classifications were created (on “use of computer”, “age” and “Rural/Urban”), and a dataset for this was prepared by the staff of DoS. The classifications, data and relevant metadata were loaded into the data model of the SDS. Due to problems of connecting PX-Web to the database (see further below), the data could not be displayed in PX-Web from the production server. Thus, data were exported from the DoS database and loaded into a similar data model on a laptop for presentation and reviewing in PX-Web.

Metadata were established using conventional sql-statements, but it will be necessary for future purposes to establish procedure and tools for maintaining the SDS. These tools should be developed specifically to suit the working processes of DoS.

**Principles of GIS in SDS**

A new demo version of PX-Map for use in PX-Web has been developed, and a demo for this was prepared for the mission. However, it was decided that priority should be given to fixing the problem of connecting to PX-Web first, before focusing on the GIS part. A short demo showing the principles of GIS in PX-Web was given during the mission. Due to the importance of this topic, the brief introduction was insufficient and should be addressed in further detail in future missions on dissemination.

**Prepare a plan the database structure**

The order by which the different statistics should be prepared and disseminated is to be determined in further detail. Also it has to be determined what should be the minimum number of statistics to be included at the first launch of the SDS. When launched, the SDS will be an important representative of DoS to the public. Thus, the decisions just mentioned should be agreed upon by top management. Having this list of prioritised statistics is the first step of creating the plan.

**Transfer of experiences regarding database and online dissemination**

Involvement of statisticians in the process of defining and preparing data for dissemination is crucial to the success of a SDS. Furthermore, it is the experience of the consultants that it is important to adopt policies on standardisation of metadata as well as clear guidelines on the responsibilities sharing between stakeholders, concerning the classification repository and the SDS.

**Other issues***PX-Web and Oracle 10 Client*

Before the mission DoS staff installed the PX-Web 2013 version on a production server. The PX-Web installation was able to present PX-files on the internet. However, the preferred way of accessing data is through the data model (CNMM) in Oracle, but the connection to the Oracle database has not yet been established. During the mission the DoS staff, the experts and experts from Sweden looked into this issue. It was clarified that PX-Web requires a 32-bit oracle 10 Client. The client was re-installed on the specific machine, but it turned out not to be sufficient. Thus, the problem was not solved during this mission. The next step could be to set up a test server for installing PX-Web. This test machine could also be used to install and test other web applications before they go onto the production system.

*Arabic vs. English dissemination of data*

During the mission it was tested if the interface of PX-Web could be set up for presentation in Arabic (from Right to Left - RTL). It seems that this can be done through normal style sheet operations, but a more thorough test of this has to be carried out, to ensure that it will meet the requirements for dissemination in both Jordanian and English.

### 3. Conclusions and recommendations

#### Conclusions

Based on the assessments above, the following conclusions were drawn:

- The life cycle from data registers to aggregates have been reviewed and solutions for each step in the life cycle have been addressed
- The proposed Classification Repository has been reviewed and established (prepared by the staff before the mission)
- The proposed data model for the SDS has been established (by the staff before the mission) and case study data have been loaded into the data model
- Reconciliation of data and Dissemination is not only an IT issue, but has to be a cooperation between IT, Dissemination and Statistical divisions. Thus, steps should be taken to clarify the responsibility sharing between these parties
- Introducing the classification repository and the SDS is a big change for dissemination, and requires strong commitment from top management to be a success
- The minimum requirements, in terms of statistics to be included in the SDS, have to be determined to create a plan for reconciling and disseminating data in the new systems
- When the new classification repository and the SDS is going to be used for production purposes, a requirement for tools to maintain these systems will arise
- The installation of PX-Web on the production server has to be completed. First step should be to install it on a test server (which is not existing at the moment)
- It has to be investigated further if PX-Web has problems in relation to presentation of data in Jordanian language (Arabic, RTL)

#### Recommendations

Based on the conclusions, the recommendations are the following:

- Discussions concerning the cooperation between IT, Dissemination and Statistical divisions concerning the classification repository and the SDS should be initiated within DoS
- A training course for selected staff members in the use of PX-Web should be carried out
- A test server should be set up to test web applications before they go onto the production system
- PX-Web should be installed on the test server, before it eventually is installed on the production server
- The need for tools to maintain the classification repository and the SDS has to be considered further (e.g. in future missions on dissemination)

#### Future activities

Proposal for agenda topics to be addressed during the next missions concerning dissemination:

- Getting PX-Web to run on a test server (and preferably on a production server as well)
- Presenting best practices in organising work on classification repository and SDS
- Training course in PX-Web for selected staff (statisticians included)
- Strategy concerning the sharing of responsibility in the process of defining and creating data for dissemination

- Review and assist the process of creating and loading more data (according to a prioritised list of statistics to be included in the SDS at launch)
- Presenting best practices concerning tools for maintaining the database and view classification repository
- GIS in PX-Web

*Actions needed for moving forward as well as for preparing the next mission:*

<b>Action</b>	<b>Deadline</b>	<b>Responsible person</b>
Setup test server		DoS
Install PX-Web		DoS (e.g. remote help from PX-Web support and experts)
Investigate on RTL in PX-Web		Experts (assisted by DoS)
Send PowerPoints		Experts
Send documentation on creating SVG files for GIS presentation		Experts
Preparing a list of prioritised statistics to go into SDS		DoS

## Annex 1. Terms of Reference

<p style="text-align: center;"><b>Terms of Reference</b></p> <p style="text-align: center;"><b>EU Twinning Project JO/13/ENP/ST/23</b></p> <p style="text-align: center;"><b>10-14 August 2014</b></p>
--

### Component 4: IT and Online Dissemination

#### Activity 4.3: Plan for database structure - II

##### 0. Mandatory results and benchmarks for the component

- New database structure defined and online dissemination improved (Apr 2015)
- Assessment report on current situation (Jan 2014)
- Develop a plan for the database structure (July 2014)
- Improve the IT-security (Jan 2015)
- Improve the online dissemination (Apr 2015))

##### 1. Purpose of the activity

- To review and refine the data model used for the case study "Employment and Unemployment and ICT"
- To prepare (reconcile) data from the case study for dissemination purposes
- To aggregate case study data for loading into the Statistical Dissemination System (SDS)
- To look into the draft model for a classification repository.
- To present the principles of GIS in SDS

##### 2. Expected output of the activity

- The data model used for the case study "Employment and Unemployment" reviewed and refined
- Data from the case study for dissemination purposes is prepared (reconciled)
- The case study data for loading into the Statistical Dissemination System (SDS) is aggregated and prepared
- The model for the classification repository is reviewed
- The principles of GIS in SDS is presented and discussed
- Prepare a plan for the database structure
- Transfer of the Danish and in general the European Union, experience regarding database structure and online dissemination
- A lining up of work programme for the next activity (4.6, scheduled for the 21. – 25 September)



### **3. Participants**

#### DoS

Mr Tayseer Deeb, Director of Information Technology (*Component Leader*)

Rana A. Swaidat, IT/ Head of Development Unit

Rania Abu Dhaim, Head of programming and analysis

Ahlam Al Rousan, Head of Programming division

Abdullah Al –sous, Web Dissemination section/IT

Manal Khuffash, Web Dissemination section/IT

Mohammad Hasan Network Engineer/ IT

Hussam abu shukor

#### MS experts

Mr Lars Knudsen, Chief Adviser, Web and Online Dissemination, Statistics Denmark

Mr. Leonardo Tininini, Head of Unit of Database Administrators, ISTAT

## 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 progress on reorganising the database structure. Presentation by DoS of the plan for the reorganisation of the database structure
	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	Discussions of the data model used for the case study "Employment and Unemployment and ICT"
	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	Discussions of the data model used for the case study "Employment and Unemployment and ICT"
	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	Discussions of the data model used for the case study "Employment and Unemployment and ICT"
				Discussions on how to aggregate case study data for loading into the Statistical Dissemination System (SDS)
	03.30 – 04.00		Preparations / Report writing	Preparations / Report writing
Tuesday, morning	08.30 – 09.00	DoS	Preparations / Report writing	Preparations / Report writing
	09.00 – 12.00		Meeting with BC Component Leader and BC Experts	Discussions of the draft model for a classification repository
	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	Discussions of the draft model for a classification repository
	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	Presentation of the principles of GIS in Statistical Dissemination System
	12.00 – 01.00		Break / Preparations / Report writing	Break / Preparations / Report writing
Wednesday, afternoon	01.00 – 03.30	DoS	Meeting with BC Component Leader and BC Experts	Presentation of the principles of GIS in Statistical Dissemination System
	03.30 – 04.00		Preparations / Report writing	Preparations / Report writing
Thursday, morning	08.30 – 09.00	DoS	Preparations / Report writing	Preparations / Report writing
			Meeting with BC Component Leader and BC Experts	Discussion of next steps in order to prepare of the work on the SDS.
	09.00 – 11.30		Ad-hoc meetings	Final clarifications with BC Experts, preparation of report and presentation for BC Project Leader
Thursday, morning	11.30 – 12.30	DoS	Meeting with BC Component Leader	Presentation for BC Project Leader
Thursday, noon	12.30 – 01.00	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 2. Persons met**

### DoS:

Mr Tayseer Deeb, Director of Information Technology (Component Leader)

Rana A. Swaidat, IT/ Head of Development Unit   Rania Abu Dhaim, Head of programming and analysis

Abdullah Al –sous, Web Dissemination section/IT

Manal Khuffash, Web Dissemination section/IT

Mohammad Hasan Network Engineer/ IT

Hussam abu shukor

### RTA Team:

Thomas Olsen, RTA

Deena N. Moghrabi, Interpreter