STATISTICS DENMARK

Statistics Norway
Statistisk sentralbyrå

Statistiska centralbyrån
Statistics Sweden

MOZ:2003:24

# Exploratory mission on The creation of an Output Database

## Report from a short-term mission

## 5th – 18th December 2003


Jesper Ellemose Jensen

Annegrete Wulff


## TA for the Scandinavian Support Program to Strengthen the Institutional Capacity of the National Statistics, Mozambique

Instituto Nacional de Estatística

This report contains
restricted information
and is for official use only.


Ref. no. 104.Moz.98 December 2003

Abbreviations used:

| | |
|---|---|
| DAHR | Directorate for Administration and Human Resources |
| DICRE | Directorate for >integration, Coordination and External Relations |
| DCNIC | Directorate for National Accounting and Global Indicators |
| DESE | Directorate for Sectorial Statistics and Business Statistics |
| DCI | Directorate for Census and Surveys |
| DEMOVIS | Directorate for Demographic, Vital and Social Statistics |
| DISI | Department of Informatics and Information Systems |
| DDD | Department of Dissemination and Documentation |
| ESDEM | *Estatisticas Sociais e Démográficas de Mocambique* from UNICEF |
| INE | National Statistical Institute of Mozambique |
| LDB | *Live Database* from World Bank |

*Jesper Ellemose Jensen, jej@dst.dk*
*Annegrete Wulff, awu@dst.dk*

*Statistics Denmark*
*Sejrogade 11*
*DK-2100 Copenhagen O*
*Denmark*

# 1 EXECUTIVE SUMMARY

The mission was conducted from 5 December to 18 December. The character of the mission was exploratory.

During the mission the present status of IT-Strategy, IT-Policy, and Dissemination Policy was discussed with people inside DISI and DDD. The proposed updated IT-Strategy and IT-policy are now very mature and needs approval and implementation, so INE can move forward towards the proposed IT-Architecture.

The homepage soon to be replaced by the coming statistical portal would clearly benefit by being supplemented by a statistical output database. The proposed CMS based portal will fit nicely into the proposed IT-Strategy in terms of standards, platforms and tools.
An output database will allow users to construct their own tables, graphs and maps giving them better and more flexible information. INE will also gain, not only by increased user satisfaction but also gain reductions in the time spent answering user enquiries.

It is therefore suggested to implement a statistical output database system, with dynamic access. Demonstrations of a system build using the PX-Web system was given at a meeting with DISI staff and at a Workshop with participation from INE, and from other parties in the SEN. The reception was positive and we have therefore developed a draft implementation plan that describes how a dynamic Internet output database can be established inside 3 months.

The PX platform is used in a number of countries, and has proven to be a stable and very low risk solution. Today it works in 17 countries and is developed according to the needs expressed by the user organizations.

We had meetings with three external users and discussed their needs for access to statistics both in hard and softcopy. Reliable information on publishing schedules seems a high priority.

# 2 INTRODUCTION

The Mission was carried out 5 – 18 December 2003 in Maputo.

The consultants would like to express their thanks to all officials and individuals met for their kind support and valuable information. In particular they would thank their counterparts in INE the Head of DISI, Ms Anastásia Judas Honwana and the Head of DDD, Mr. Marcelino Silva for their preparation of an extensive program which facilitated the work of the consultants.

Several activities at INE concern how to improve the use of Internet for dissemination of statistics.

- INE is in the process of discussing and approving the updated IT-strategy and IT policy. The use of Internet and standardization plays an important role in the future information architecture based on a common data storage.
- Maintenance of the homepage will be handled in a web content management system. At the end of the year a private company is expected to carry out requirement analysis, give advice on tools, build a static page etc. INE has prepared contract and requirement specification for this activity.
- A reference group for internet is to discuss the content of the page. Subject matter areas and horizontal functions are represented in this group
- There is a group, inside the IT-department, investigating tools.
- DISI has staff starting training-courses on methodologies and tools about the Internet.

- As part of installation of ESDEM and LDB there is ongoing discussions about how to use (or not to use) these systems for disseminating data on the Internet. ESDEM and LBD are dissemination applications for statistics. They both have facilities for data input and facilities for presentation of data. ESDEM is from UNICEF. LDB is from the World Bank.

These activities are brought in mind when recommendations are given.

Terms of reference for the mission is in annex 1

# 3 MISSION PROGRAM

The general objective of the mission was to continue the development and improvement of the webpage in order to consolidate the mechanisms of accessibility to statistics.

More specific the consultants should have clear and operational ideas

- About how to define content in order to dynamically publish basic statistical information and fulfill user needs.
- About the creation of information architecture in order to support the use of an output database and Internet database
- About tools to use for development and operation of the output- and internet database.

All these aspects were discussed according to the program, which included meetings with end-users as well as subject matter staff and the Dissemination department. Existing software and databases were demonstrated and Internet solutions were discussed with the IT-department during a presentation.

Annex 2 contains the mission program.

# 4 ACTIVITIES DURING THE MISSION

## Dissemination policy

The approved dissemination policy was discussed together with the counterpart in DISI. In particular the following goals advocate for an Internet database:

- Accessibility to data should be improved
- Timeliness of presentations
- User needs shall be taken into account
- INE shall spread awareness of statistic and the statistical products
- Information shall be coherent across different media and sources.

An Internet Reference Group has been established with participants from several departments: DEMOVIS, DESE, DCNIG, DICRE, DARH, DCI. The task is to coordinate and secure relevant data on the Internet including defining the content of the Internet database..

## Improvement of the homepage in relation to output database

INE is presently working on a modernization of the Homepage. The modernization will include the procurement of a CMS (Content Management System). The CMS shall allow INE to change its present static homepage into a modern and user friendly statistical portal.

It is important to understand that CMS and Data warehouses / Output databases are completely separate and highly different systems that play different but complimentary parts in dissemination

on the internet. CMS's stores and present textual information to the users. The textual information can be structured in varying degreases and be associated with Meta-Data. The Data Warehouse only contains (statistical) data in the form of multi dimensional tables / cubes. All data is in principle stored using the same basic Meta-Data model.

In advanced forms of Internet dissemination the CMS and the Output Database interacts. The CMS holding the textual information of the homepage and the navigation structure it self. The output database then supplies the tables, graphs and maps that are presented on the homepage. Such a dissemination form should be the ultimate goal for INE. In the meantime it is important that the future CMS system will allow "links" to external pages. This is a functionality included in most Content Management System.

In terms of tools and standards it's important to realise that both CMS' and Data Warehouses builds on core-databases (Oracle, Sybase, MS-SQL, MY-SQL) and that these again depend on operating systems software (Windows, Unix or Linux). It is highly recommended that all systems inside INE share the same basic platform. According to the present IT-Strategy INE has Microsoft as the basic software platform and ACCESS as a database system. Building the output database INE should consider moving from ACCESS to a larger database system.

### Discussions and demonstration on Internet database with DISI

A meeting was held with the staff of DISI to present the PX-Web as an example of a dynamic statistical output database and to share our experiences in building Internet databases. The presentation also introduced the Danish concept of a National Data Warehouse. It was stressed that the word Data Warehouse has different meanings to different people in different organization, and there for also can imply different tools and standards. In relation to statistical offices, the Data Warehouse concept is about storing aggregated data. The data will be stored using a common meta-data model, which shall be capable of containing data for all subject matter areas in a national statistical office. As the concept is used by most countries (Statistical Offices) the raw data from the collecting process (individual data) will not and shall not be part of a common Data Warehouse concept. All subject area departments should use the same platforms and tools in the data collecting process and then load the aggregated data into the Common Data Warehouse. This is also part of the IT-Strategy and IT-Policy which INE at the present moment works to finalize and adopt.

During the meeting it was also discussed how a Common Data Warehouse in the future could be used to produce output to different media i.e. CD-ROM, Homepage and printed publications.

The DISI staff told about how they store and process data from the subject area departments. During this it became clear that it will be straight forward to convert existing data into PX-files which can be used in a pilot project to construct a dynamic internet output database using the PX-Web software. Compared with the complexity of present systems, conversion to PX-files should involve no technical risks and will only require a small amount of additional training.

### IT-Strategy and IT-Policy discussed

The consultants, the long term IT consultant (Mogens Grosen Nielsen) and Anastásia Judas Honwana had a very detailed discussion of the proposed updated IT-Strategy and the IT-Policy.

The discussions focused on the implications of the proposed future information architecture to Internet dissemination and organization of work procedures. The general conclusion was that for the strategy and policy to really work they must be strongly supported by the top management level of INE. This corresponds to the findings from Lars Thygesen's mission.

On the strategic level INE is committed to Microsoft products as its software and operating system platform. This platform will allow the implementation of the PX- family soft-ware.
The possibility of LINUX and Open-Source platforms was discussed. It was stressed by the consultants that the Microsoft platform is the platform used by Statistical Offices in general and by the ScanStat consortium in particular. There for it is important to continue on the Microsoft platform in order to have access to specialized statistical software applications (tools) and to be able to have corporation with other statistical institutions. Also the present database (LDB and ESDM) are developed to the Microsoft platform.

Both the IT-Strategy and the IT-Policy have been developed with support from other short term missions. They now seem to have reached their final forms and the main task is now to have final approval, and to start implementation. From our meeting with the President we understand that final approval should happen shortly.

## Meetings with users

As the mission was of an explorative character it was important to get input from users of statistics concerning their use and needs of statistics and electronic access to data.
INE had arranged three meetings with some of their main users. These were UNICEF- who is responsible for the ESDEM database, TDM –Telecommunications Mozambique and SOARES DA COSTA – a big construction company.

- During our discussions they all expressed a need to know which data INE can deliver. An improved *awareness of statistics and products* was asked for. This could be provided through a public advance calendar for releases- or as subscriptions on updates of statistics as soon as they are available. INE is already working on this.

- Another requirement mentioned was a possibility to get data *aggregated at a lower level* than today and with more details. This will give room for more flexibility in the selection of statistics for the single user.

- To have access to more updated data - preferable on the Internet - was another whish.

- Finally, documentation in terms of definitions, source, reliability etc would be of utmost value for the users.

## Work shop on Output Database

A workshop on output database was arranged for directors from INE as well as stakeholders from outside INE.
The objective was to inform and discuss the future Information architecture in INE. A dissemination database with dynamic access on the Internet is part of the vision
The consultants demonstrated a stepwise solution to such an Internet database using the software PX-WEB. PX-WEB was demonstrated. It has proven to be the database solution in 17 countries and the strong sides like flexibility in selection and presentation of data was demonstrated. Showing data on maps and in graphs is a functionality appreciated by most users. Population pyramids, stables, curves are easily produced from a table. Another valued functionality is to link documentation and footnotes directly to a table.

PX-WEB had been translated to Portuguese and a few tables with Mozambique data had been created. The cartographic department had been helpful in delivering maps of Mozambique of the provinces and the districts.

The participants discussed what possibilities the new information architecture could bring INE and how to comply with new user groups.

Agenda 5 contains the invitation for the workshop, program for the workshop, presentations made, conclusions from the user discussions and a list of participants.

# 4 CONCLUSIONS AND RECOMMENDATIONS

## 4.1 Recommendations

**Recommendation 1:**
**The updated IT- strategy and IT—policy should be approved and implemented**.
The IT policy and IT strategy should be approved as soon as possible. This is important in order to secure that all IT-activities and developments take place within the same frame that is set up in the strategy. It is also recommended to present and explain the policy and strategy at the next common steering committee meeting. One item on the agenda for the subsequent meetings should report on how the policy and the strategy are followed.

**Recommendation 2:**
**A stepwise approach to a common data warehouse**.
The project of building a common data warehouse as a core for dissemination on the Internet in general, in a dynamic databank on the Internet as well as for publishing on other media like paper and CD-ROMs is recommended. The process should be approached stepwise in order to gain experience and more important to be able to present a database on the Internet within 3 months from approval.

**Recommendation 3**
**Internet Reference group to organize the database contents**
Establishing the Internet database is a task involving the DISI, DDD as well as the subject matter directorates. The work can be organized and coordinated within the Internet-reference group that holds members from these areas. The group should propose the content of the pilot database. It is recommended that the group meets monthly during the pilot project.

An outline of a project description / implementation plan for a pilot project to develop a Internet Output Database using PX-WEB software has been developed and is included in the rapport as Annex 4

**Recommendation 4:**
**User involvement.**
The dynamic database shall be driven by user needs. It is important to keep up with these needs through user satisfaction surveys, by registration of the usage and by meeting the users. 6 months after launching of the database INE it is recommended to evaluate the project involving comments from the users. INE should conduct annual user meetings to get input to the development of the web site and the database.

**Recommendation 5:**
**Database in an Internet portal**
INE is working to change its present static homepage to a statistical portal based on a CMS system and will launch such a project in 2004. This project should make provisions for later integration with a Statistical Output Database based on the Common Data Warehouse envisioned in INE's long term IT-Strategy. In the short run the CMS must as a minimum allow links to a dynamic internet database system. CMS's systems and Output database systems are seen as complimentary but completely different systems in terms of internet dissemination.

**Recommendation 6**
**LDB and ESDEM databases**

The two existing databases LDB and ESDEM contain restricted data created with very specific goals. They can not produce multi dimensional tables and they are not accessible on the Internet. However, they fulfill some user needs, and it is recommended they will not be affected by the pilot project.

**Recommendation 7:**
**Advance publishing calendar**
The annual publication plan should be available on the Internet. It shall be used as a tool by DDD to coordinate the update of the homepage. This was a need expressed by the users interviewed during the mission, as well as on the workshop.

## 4.2 Recommendations for further support from the Scandinavian program:

Given that INE agrees on the above recommendations the following activities are recommended:

- There should be a mission to assist INE with installation of an Internet Information Server.
The costs for this will depend on decisions made in the up coming CMS project. If the project involves a dedicated server this should be used in the pilot project also. Other wise a dedicated server must be procured (expected to cost between 3,000 and 4,000 US $). It is not possible to estimate the additional costs of internet traffic due to an Internet database.
This activity is prior to the others.

- A mission to install the PX-WEB software and to conduct a training course in input and retrieval of data is recommended. This should happen as soon as the Internet reference group has started defining the content of the pilot database. The mission will also include assistance in defining content and structure. The cost of software in the PC-AXIS family – including PX-WEB- is 1,500 USD approximately which can be paid by the project.

- A follow-up mission two moth later to support the launching of the database to the Internet.

- A study visit to Scandinavia to study data warehouse technology is recommended.
This visit will benefit for phase 2 and 3 of the output database project.

Moreover INE should participate in the annual PC-AXIS Reference Group Meeting, where future development is decided. Members are statistical offices that use PC-AXIS/ PX-WEB. The first coming meeting takes place in 4$^{th}$ quarter of 2004.

# ANNEX 1 Terms of Reference for the mission

**TERMS OF REFERENCE**
within the Scandinavian Program

**for a short-term mission for a 2 week mission in December 2003**
on
**an exploratory mission about creation of an output database**

**D R A F T**

**Background**

There are many activities at INE concerned about how to improve the use the Internet.

- INE is in the process of discussing and approving the IT-strategy and IT policy. The use of Internet and standardization plays an important role especially as part of the vision for future information architecture. See annex A.
- At the end of the year a private company is expected to carry out requirement analysis, give advice on tools, build a static page etc. INE has prepared contract and requirement specification for this activity.
- There is a group of discussing the content of the page. Subject matter areas and horizontal functions are represented in this group
- There is a group, inside the IT-department, investigating tools.
- DISI has staff starting training-courses on methodologies and tools about the Internet.
- As part of installation of ESDEM and LDB there is ongoing discussion about how to use (or not to use) these systems for disseminating data on the Internet. ESDEM and LBD are dissemination applications for statistics. They both have facilities for data input and facilities for presentation of data. ESDEM is from UNICEF. LDB is from the World Bank.

The work to be carried out at mission must have the above-mentioned activities in mind. Consequently these areas must be analyzed before drawing conclusion or writing recommendation.

Definitions:

- Output database is defined more or less like datawarehouse. Database containing aggregaged data mainly for analysis and dessimintaion. The Internetdatabase contains a subset of the data in the output database.

## Purpose of the mission

Have clear and operational ideas

- on how to define content in order to publish basic statistical information / fulfill user need
- on information architecture in order to support the use of output database and internet database
- on what tools to use for development and operation of the output- and internet database
-

## Expected results

1) Course material
2) Mission report

## Activities

1. Detailed planning and prioritization
2. …
3. …
4. Workshop
5. Writing of mission report and evaluation

**Tasks to be done by INE to facilitate the mission**

- Elaborate ToR for the mission
- Prepare and supply the consultant with necessary documents and information
- Supply good working conditions for the consultant

**Consultant and Counterpart**

Consultant:
Annegrete Wulff and Jesper Ellemose Jensen from Statistics Denmark

Main counterparts:
Anastácia Judas Honwana, Head of Department, IT-department
Silva, Marcelino, Head of Department, Dissemination-department

**Timing of the mission**

Two weeks (December 8 – December 19, 2003).

**Report**

The consultants will prepare a draft report to be discussed with INE before leaving Maputo. He will submit a final draft to INE for final comments within one week of the end of the mission. Statistics Denmark as Lead Party will print the final version within 3 weeks of the end of the mission.

**References**

- Mission report about internet (Netterstrom)
- Mission report about dissemination (Aagaard)

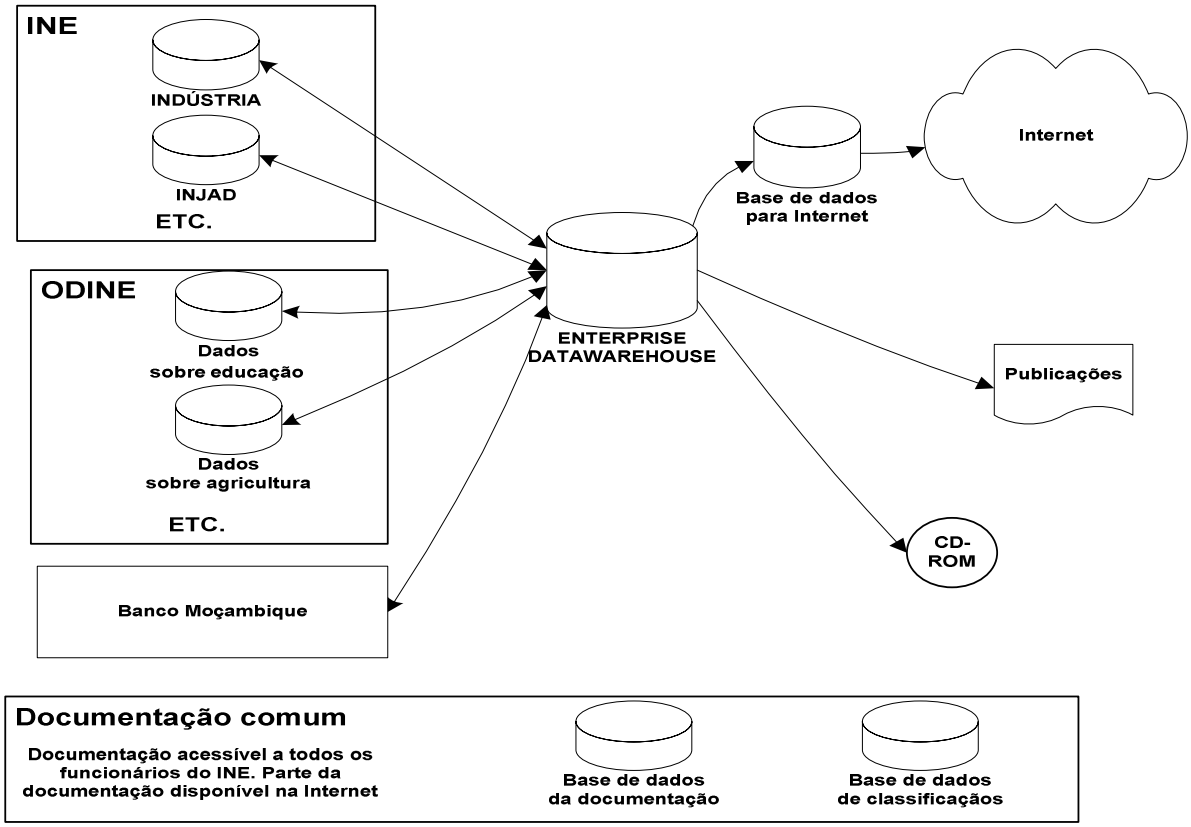
*These Terms of Reference were prepared by*

*Day / / .....*

*Approved by/in the name of the President of INE*

*Day / / .....*

Annex 1. Information architecture vision



# ANNEX 2 Agenda for the workshop etc.

## Conclusions:

Conclusions from the Work-Shop

The delegate's participation in the workshop was arranged around the following questions.

1. How can INE improve coordination in relation to Internet publishing?
2. Will INE reach new user groups?
3. How can INE get an idea of what the users want and need?
4. Thematic organization of information? (PARPA)…
5. Specific data for the provinces?
6. Free of charge data?
7. CD-ROMs, Internet, paper? When to choose which media?
    1. Can some printed publications be substituted?
8. Etc…?

The questions where compiled by the consultants, the long term IT consultant Mogens Grosen Nielsen, Anastásia Judas Honwana and Thomas Bernado
The participants where divided in to 3 different groups who discussed the questions and then presented their conclusions to the workshop.
In short form the answers where like this:

From group 1:

- INE shall support producers
- Producers outside INE should be involved
- Identify users, areas not explored, comparable with other country's (Links)
- Questionnaires to users, about what they need (Internet / print?) both dissemination and collection
- Organized by subject area, not by instrument of collection (avoid duplication of information)
- All on homepage should be free, but list of commercial products

From group 2:

- Regular meetings among producers
- Annual publication plan (compiled / respected)
- Avoid differences in published material, clear definitions, focal points (user segmentation)
- Highly aggregated should be free
- All types of media should be used, to meet the expectations of users
- User survey (detailed)

From group 3:

- Calendar of publications, all parts of SEN
- Resume (abstract) of each publication on Internet
- Dynamic page will introduce new users
- Survey of user satisfaction on Internet
- Data by themes (Source is important), data from provinces
- Detailed data to be charged (E-business model?) Global data should be free
- Media depends of the user (Users geography) Webpage shall tell how to get more information

- Not all have access to Internet

In total, we believe that the conclusions can be summarized into:

- Calender (Annual publishing plan)
  - Good idea (The users agree), should cover entire SEN
- User survey (paper / web)
  - Satisfaction?, User needs?, Responsiveness?
    - Survey is under way?
- Media (Support multiple platforms)
  - All are relevant, depending on the users situation
- Internet reference group
  - All directorates should take responsibility

## ANNEX 3 Programme for the mission

*Program for the exploratory mission about internet and output database*

**Maputo, 2 December 2003**

**Duration of the mission: 5th to 18th of December**

**Consultants:** Annegrete Wulff e Jesper Ellemose Skou
**Counterparts:** Marcelino Silva e Anastácia J. Honwana

### December 5

13:30 –15:00 Meeting with Head of department for IT and head of department for dissemination and the IT-consultant . Introductory presentation.

### December 8

9:00 - 10:00 Meeting with dirrector for Dicre, Head of department for IT, head of department for disssemination and the IT-consultant. Presentation of the SEN-system

10:00 - 12:00 Meeting with Head of department for IT and head of department for disssemination. Discussions about term of reference and adjustments of the programme.

12:30-14:00 Lunch

14:00-15:30 Presentation of the draft versions of IT-strategy and IT-policy

### December 9

9:00 -11:00 The actual situation of the internet page and ongoing activities about restructuring the page.
11:00-12:30 Presentation of dissemination of statistics including the dissemination policy.
12:30 -14:00 Lunch
14:00 - 15:00 Meeting with staff from the demographic area
15:00 - 15:30 Presentation of ESDEM

### December 10

9:00 - 10:00 Meeting with staff from the economic area
10:00 - 10:30 Presentation of LDB
10:30 - 12:30 Meetings with main users
12:30 - 14:00 Lunch
14:30 - 15:30 Meetings with main users

### December 11

Meetings with main users

### December 12

9:00 - 11:00 Meeting with DISI
11:00 - 12:30 Preparation of the workshop
12:30 - 15:30 Preparation of the workshop

**December 15**
11:00 -12:00 Meeting with S. Excias o PINE, VPINE E
Preparation of the workshop

**December 16**
8:30 - 14:00 Workshop

**December 17**
Preparation and discussions of the mission report

**December 18**
Preparation of the mission report

## ANNEX 4 Outline of Output Database Pilot project

Phases in Step 1.

If the stepwise approach is approved by INE, the first step of the project (The pilot project) can be broken down into a number of smaller steps. The pilot project is estimated to have duration of 3 months, starting after access to a suitable Internet Server has been established. The phases in step 1 are not entirely progressive, but can to a certain degree be carried out in parallel.

Organisation of the pilot project
Before a pilot project can begin a suitable project document must be established.
It must describe in detail the work to be undertaken by INE, and the persons who are to carry out the work. The pilot project could refer to the Internet Reference groups .

| | Activity | Duration | Responsibility |
|---|---|---|---|
| 1 + 2 | Installation of Web Server plus PX-Web software | 3 days | ScanStat. Should be combined with network support mission |
| 3 | Localisation of PX-Web | 3 days | ScanStat, INE shall provide translations and take linguistic responsibility. |
| 4 | Definition of structure and content | 5 days | DDD / DISI must be based on decisions supported by the Internet reference group<br><br>Support by short term mission from ScanStat |
| 4a. | Training in PX-WEB, PC-AXIS, and the concept of multi-dimensional cubes. Hands on training building tables. Should involve 2 people from DISI | 5 days * 2 persons = 10 days | DISI supported by short term mission from ScanStat<br><br>4 and 4a is related and should be undertaken in the same mission. |
| 5 | Identification of textual information and descriptions in relation to the data selected to be included in the pilot project | 2 days | DDD |
| 6 | Building tables | 10 to 12 days<br><br>Depending on activity 4. | DISI at INE. |
| 7 | Validation of tables | 2 days for DISI to review of the tables. About ½ day for each subject matter department . Again | Relevant subject matter departments at INE. Shall refer to Internet reference group |

| | | | |
|---|---|---|---|
| | | dependent on 4. | |
| 9 | Loading of data into production database. Follow up on 4,5,6 and 7. Opening of system on Internet. | 5 days * 1 person. | DISI supported by short term mission from ScanStat |
| 10. | Project evaluation | | INE |
| 11. | Marketing and media | ? | DDD |

1. Getting access to a suitable web server

The PX-Web solution runs on Microsoft Internet Information Sever (IIS). PX-Web is normally installed on Windows 2000 Server. End of 2003, PX-Web has not been tested on Windows 2003 server. It is expected that PX-Web will run on Windows 2003 without any problems. If INE starts a migration project to Windows 2003 Server before a project is initiated, then PX-Web should be installed on 2003 Server. PX-web does not need a dedicated IIS and can be run on the same IIS as the INE homepage. But it is important that an external PX-Web solution runs on IIS dedicated to INE. Running PX-Web will not be advisable in an environment where homepages from different organizations are hosted on the same Information Server.

The requirement of a dedicated Web Server should be taken into consideration as part of INE's procurement of a new Content Management System (CMS), in order to modernize and improve the homepage.

The PX-Web solution should have its own name on the Internet (DNS'entry) like statbank.ine.mz. The selected name must be registered with the local hosting authority and the INE's Internet Service Provider.

Installation of an Internet Information Server could be performed as part of one of the Network support missions undertaken by the Scanstat consortium.

In terms of tools and platforms the suggested PX-Web solution fully corresponds to the IT-Strategy and IT-Policy of INE.

2 Installation of software
PX-Web consists of a series of components (dll's) which will be supplied the ScanStat consortium. Depending on the future hosting of INE's homepage FTP software for file upload must be installed. This is not part of the PX-WEB software.

3. Localisation of PX-Web
PX-Web's user interface is designed for easy localisation in terms of language and so called "Look and Feel". At present the system does not support Portuguese, but this is clearly a minor project for someone with adequate language skills. Icons, logos and colours must be adapted to follow INE's corporate identity. Again this is not a major project as the product has been designed specifically to allow easy localisation. Localisation must also include preparation of on-line help texts and user guides.

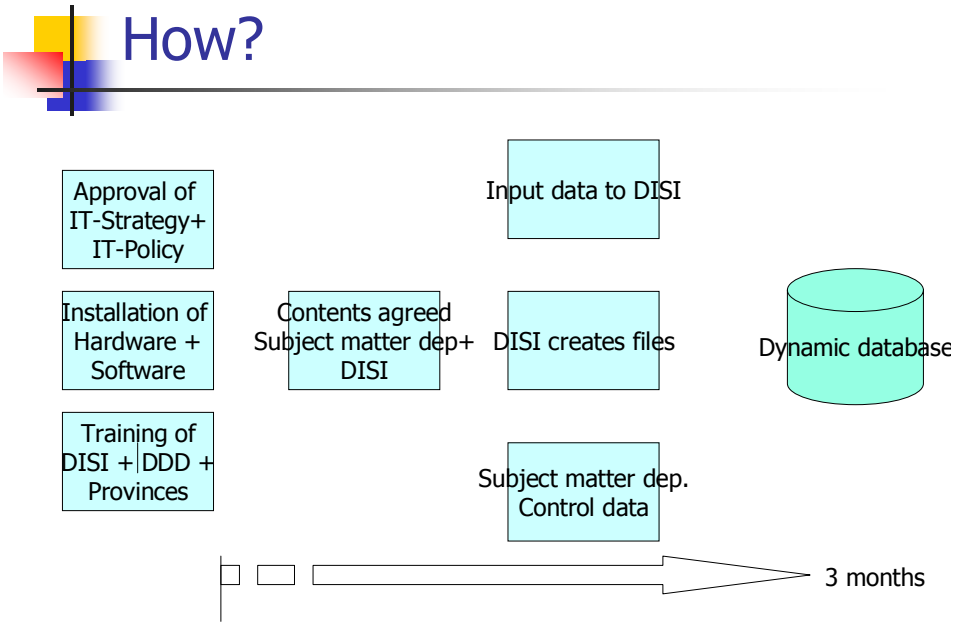
4. Definition of structure and content
The content in Dissemination Output Database is normally shown / presented in a number of subject areas. Some national statistical offices choose to reuse the subject area structure from either their yearbook or some other form of known publication structure.

When a structure has been agreed, the content (tables) inside each subject area must also be agreed. As the project is intended to be a pilot project it is important to keep the number of tables to a manageable level.
It is suggested that data is selected by the criteria of easy access and general demand. One could say that the pilot project shall focus on the low hanging / easy to pick fruits. Frequency could be criteria, but if data with a high sampling and / or publication frequency is included in the pilot project, it is important to set aside the necessary resources for continues updating.

5. Identification of metadata
The content for inclusion in the pilot project has been identified, variables and values for each table must be described in order to harmonize (codes, titles, units, etc.) meta-data for the tables

6. Building tables
The tables and meta-data has been agreed, the tables can be build. It is suggested to use the PX-Make software to do the actual table construction. Before the construction of tables can begin training in PX-MAKE and the concepts of multi-dimensional cubes must be conducted.

7. Validation of tables
It is important that the tables created bye DISI and / DDE is review by the responsible subject area statisticians before they are published on the Internet.




8. Internal / External database?
Depending on the amount of data involved in the pilot project, it should be considered to establish an internal PX-WEB solution on the Intranet. Data in the internal database should then be arranged by subject area and access to data should be limited to DISI technicians and the relevant subject matter people. The purpose of such an internal / intranet PX-WEB solution will be to allow the different subject matter divisions the possibility of applying quality assurance to the tables before they are published to general public.

9. Loading the data into the production-database
When data has been validated by the subject matter departments it can be loaded into the public access web database.

10. Evaluation of step 1 in the project.
When the data defined by the pilot project has been published on the Internet, it is important to evaluate the project. The evaluation shall address both external user's expectations to data and the organisation of work during the pilot project.

11. Marketing / dissemination
Finally potential users of the data included in the pilot project should be made aware of the new possibility's created bye the dynamic internet database.

## ANNEX 5 Persons met during the mission

**Dr. João Dias Loureiro, Presidente do INE.**
**Mr. S Excia Vice President Demography**
**Ms. Destina Eduarda Sidónio Uinge, Director DICRE.**
**Ms. Anastásia Judas Honwana, Head of DISI**
**Mr. José Inguane, DISI**
**Mr. Arao Mbalate, Director DCI**
**Ms. Fatima Zacarias, Director DEMOVIS**
**Mr. Marcelino Silva, Head of DDD**
**Mr. Azarias Marcos Nhanzimo, Director DESE**
**Ms. Suzue Saito, UNICEF**
**The Commercial DirectorTDM**
**Mr. Francisco Silva, Director Soares da Coata**
**Mr. Tomaz Bernado**
**Mr. Ralf Lanwehr, International Center for German Migration**
**Mr. Hans Erik Altvall, Consultant, Coordinator.**
**Mr. Mogens Grosen Nielsen, Consultant**