



FINAL OBSERVATIONS IT-ADVISOR

29 December 2002

TA for the 'Bridging Support Program to Strengthen the Institutional Capacity of the National Statistics, Mozambique

Bo Yttergren



Instituto Nacional de Estatística

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

FINAL OBSERVATIONS IT-ADVISOR

29.12.2002

Here follows some observations, tips and advice based on my five years of experience at INE. Due to time constraints they are written in English. As everything else I have written at INE, this is also done in good faith with the best interest of INE at heart. Many comments are well-known and may not need to be repeated but should even so make part of this personal overview of the IT-situation. All opinions expressed are my own.

INFORMATICS DEPARTMENT

Development

It is fair to say that the computing environment at INE has undergone a radical transformation over the last five years. In 1998, there were very few computers, the programmers at DISI did not have a machine of their own but had to program on other people's machines. The PCs were stand-alone and in many cases DOS-based.

Then in 1999-2000, the INE computer network, based on Windows NT, was gradually created and the PC environment became co-operative with the possibility of file sharing and e-mail for internal communication. The database environment changed to MS Access, several important new databases were created such as QUIBB and CAP. In 2001, the Intranet further improved communication while the INE Internet Web Page, created in 1998 attracts more and more visitors. Another major event was Full Internet in September 2002 whereby DISI for all practical purposes became an ISP (Internet Service Provider) for the 150+ users.

INE is now entering a third computing phase whereby the great amount of collected data needs to be consolidated and retrieved easily in tools such as the Live Data Base or ESDEM. On the medium term INE should store all databases in some form of Data Warehouse and with more web publishing, should also host its own web page. All this requires a rather complex environment and it is clear that the "amateur" days are over.

In order to achieve the goals of the *Plano Quinquenal*, it is clear that professional IT standards and IT methods will have to be implemented in practice. As a consequence the IT Plan of 2003 now being drafted, must have the highest attention of the INE management.

IT staffing

Lack of qualified staff has been, and still is a major constraint for the smooth utilization of IT technology. Over the years there has been a constant deficiency of at least three-four posts which has meant that DISI has been forced to work defensively with only the most urgent tasks and has not been able to attend to the strategic IT questions, such as planning, standards training etc as one could wish for. In my own case it has meant a lot of gap filling particularly in the areas of systems design which while regrettable was necessary.

Another important effect of DISI being short of staff is the fact that the Head of DISI is forced to do a lot of “ordinary” IT-work to the detriment of the important tasks of planning and management of staff. Further hirings should consequently have the added benefit of improved management thus giving time for better planning and follow-up of the IT resources.

The staffing situation has in fact improved in recent years with some important new recruitments. As we all know, the area in most need of new staff is the Helpdesk/Network Management Function. The reliance on only one person is extremely risky. There is in fact an Action Plan in place to remedy the situation both on the short term and the long term. It calls for temporary external support, hiring of new staff, training and more co-operation with IT-departments in other organizations.

The most important action to implement is to give the Network Administrator the Microsoft Certified Engineer Course (MSCE). This will reduce or even eliminate the dependency of external consultants.

It is urgent to hire a person that in a short time can substitute the Network Administrator. This implies that the person must have already a basic knowledge of network administration.

Ideally, a second “trainee” should be recruited to the Help Desk. This could be a recently graduated person with a keen IT-interest and a proven willingness to learn. His tasks could initially be to see to that the basic maintenance tasks of a network are being done regularly and professionally. These include:

- Keeping tabs on the back-up jobs. They should be updated regularly as new databases (surveys) are being done and there are redundant data that do not need daily back-up. As a matter of fact, the data volume of INE has increased to the point where backing-up with tapes will not be viable in the not so distant future. An automated back-up system should be in place.
- Maintaining the IT Inventory. INE has a substantial investment in IT equipment – more than 200 PCs including the DPINES, printers, scanners etc and unfortunately DISI is not at the moment able to maintain a current inventory. This makes good management more difficult and there is also the obvious risk of loss of equipment when the inventory is of bad quality.
- The Server logs should be examined every morning and every red message, indicating there is a problem, should be recorded, examined and looked into.
- As INE is constantly attacked by computer virus all occurrences should be recorded from the logs and appropriate actions taken.
- A Helpdesk Incident Database should be operated, as was the case briefly in November-December 2001. Such a database would over time be a valuable source of information, what kind of IT problems that exist, which computers and users that are affected and so forth.
- User accounts management. Persons come and go and the network user accounts should be kept up to date.
- Security Management. It is necessary to implement a good policy on which users that should have what rights, reading/writing/deleting etc on files and directories. The contents of the public Incoming network directory should be revised; there is a great mixture of files with great importance, less importance and files with no importance at all.

Related to security is a Disaster Recovery Plan, which should be in place if fire or other disaster breaks out. Security concerns should be a top consideration when INE designs its new premises.

It is also necessary to reinforce the system development side of DISI. While the number of programmers probably is sufficient someone should take direct responsibility for the systems design. This includes monitoring an INE systems design model regarding development, naming conventions, application interface and not least standardized documentation. It is my opinion that the Statistics Sweden Statistical System Model would serve INE well with some minor adaptation.

Regardless of which model to be used, it is vital that the model is consistent and that it does assume close co-operation – from inception to publication – with the statistical subject matter area. If INE manages to create a strong systems design policy, INE could also more easily desist donor driven solutions, which in many cases are based on the convenience of a particular consultant rather than INE's long term goals.

INE should try and make a determined effort to absorb the QUIBB technology, from the conception of questionnaires, the scanning, the transfer of data to MS Access and finally from the publishing in MS Excel. While the QUIBB system does not suit every survey, rightly implemented it will save a lot of time and money compared to the traditional data processing.

Organization

The IT Plan of 1998 suggested dividing DISI into two sections, one for HelpDesk/Network Management and one for Systems Design/Programming.. Considering the increased number of staff and the necessity of specialization due to the more complex tasks, the proposal is even more relevant today. Whether it is possible to implement is another matter but it should surely be borne in mind when INE's future organization is considered.

Another organizational matter is the fact that the DISI is only one among 13 other departments and as such a low level in the organization does not have the power to enforce the IT Policy, as it should. It is true that there is a Directorate responsible for the IT questions at INE but it should be pointed out that presently the Director is also the Programme Director of the Scandinavian Programme which inherently must steal time and attention from the IT matters. Maybe it is still early in INE's development as an institution to make any radical changes of its structure, but it is a fact that the IT department in modern statistical bureau has a position in the organization that is reflected by the importance to its activities.

Monthly Surveys, Business Register

In the new more complex computing environment that INE will work under, co-ordination of the IT efforts will become crucial. During the past few years this has been one of the major fallacies at INE in my opinion. There are probably many reasons for this, and one of them is that I have not been able to convince the DISI clients of the advantages of having a close co-operation in the first place.

The problem is felt in particular with parts of the Economic Area, which in many cases has chosen its own solutions without proper consultations with DISI. Often a DISI

technician has been called on only at a very late stage to do a particular task without DISI management having the opportunity to discuss the overall picture.

Actually it should be the long-term goal that each department at INE can handle its computing needs by themselves to the greatest extent possible. So from that point of view the efforts of DESE to process *Comercio Externo* and *Entradas e Saídas* all by themselves were laudable. However, while successful, the data processing was far from optimal, as several server crashes were provoked by the statistician's insufficient knowledge of Access (creating databases of enormous size).

It is therefore suggested that professional data processing systems are made by DISI for *Comercio Externo* (until Eurotrace is in use) and *Entradas e Saídas*.

Another example of sub-optimization because of lacking discussions between the IT area and the statistical subject matter area is CEMPRE where DISI became involved only at a very late stage. While the new Business Register FUE certainly will serve INE's needs with some modifications it was never designed as a data entry vehicle for a business census. It is also rather amazing that an IT person never was involved in an evaluation of that application.

What is normally done in processing a census is that data entry screens are done based on the questionnaires and on the data put into a database that later can furnish data to other systems such as, for instance, the FUE.

One positive outcome of CEMPRE and FUE is that INE will have the opportunity of a fresh start in the monthly business surveys, which presently render little statistical value due to the poor sample of December 1999, the poor response rate, the exclusion of some major industrial establishments and the non-application of weights which really is fundamental in a sample survey.

The applications that were evolved during five missions from Statistics Sweden are operable but will have to be adapted to the new terminology that is being used by FUE and for that matter the new CAE. There must also be an adjustment on how the paralyzed companies are treated. Now a paralyzed company is deemed to be paralyzed permanently but the application must have more flexibility to allow the company to restart production. There is also a request of having more control of the imputations.

An understanding should be made about the possible decentralization of the monthly surveys: While typing in the information certainly can be done in the provinces (although the smaller ones like Tete and Niassa have very few questionnaires to process) it should be noted that data processing can be done **ONLY** at INE HQ. This is because the application must have the questionnaires from all provinces in order to be able calculate the weights since this is a **national** survey and **not** a provincial survey.

Even validation in the provinces may be difficult although there is the advantage of having local knowledge. In order to make a good validation there is a need to have access to all historical data. This then implies two copies of the database, one at the province and one at INE HQ for data processing. There is always the problem with synchronization when having two copies of a database and someone has to be the referee and say which copy is correct.

It is therefore suggested, once CEMPRE has finished, and FUE is up-to-date, to make a new sample and to do the statistical processing according to established statistical practices, e.g. weighting. A number of short-term missions with this object in mind should therefore be planned for 2003.

The Consumer Price Index

Here there has been a very close and fruitful co-operation over the years. Unfortunately, it was not possible to do a proper handing over of the IPC Access application as the appointed programmer had to work with the Business Census more or less full time.

IPC Access was used successfully for several months during 2001 running in parallel with the Clipper application. Unfortunately, staff was withdrawn and IPC Access could not substitute Clipper as intended.

The application needs the following interventions:

- Special treatment of missing prices (*Ausencias*) for the markets. There is a Visual Basic routine in a test database that works for the shops and a similar one for the markets that needs further testing. After positive testing the routines need to be incorporated into the application.
- Presentation of the indices via Excel. Instead of doing this work manually as is the case now Access can be programmed to produce the results in Excel
- When the application will be introduced in Nampula and Beira (and other provinces) there need to be an adjustment of certain tables (shops, markets, products) to reflect the local content.

It should be stressed that the DPC staff need training in MS Access in order to make better use of the application and permit ad hoc analysis as the need arises.

National Accounts

There has not been any work done on the systems side between DISI and National Accounts. However, it was noted that they work with more than 50,000 Excel files, which is difficult to manage. National Accounts is strongly advised to store their information in databases, a long but no doubt necessary transformation process.

GIS

The IT Plan makes few mentions of the use of GIS but the Cartography department has made great progress in recent years using this tool. (Although creating thematic maps is not strictly GIS). The Business Census is accommodating co-ordinates for the enterprises.

Since embarking on GIS is a **major** undertaking – equipment, software, human resources signify a considerable financial investment, the objectives of INE's GIS efforts should be thoroughly investigated and the practical consequences carefully analyzed.

Not least because there are many other GIS stakeholders in Moçambique, chiefly DINAGECA, Ministry of Environment, Agriculture and others. There is likely a need to co-ordinate on the national level to decide who should be in charge of the coordinates database and what the various interests are. Software should be standardized in order to facilitate data exchange and training.

Intranet, Internet Web Page

The Intranet, now in its first year of existence, has proved to be a useful tool for dissemination of internal information. However, its usefulness is entirely dependent on how often it is updated with new information. At this point the updates are erratic because there is no Web Editor who assumes the responsibility.

As long as this person does not exist, it is also difficult to decentralize to the Directorates and Departments who should be the principal content providers. Another consequence is that the web directories are not maintained, having many broken links, lots of unnecessary files. The directory structure is no longer logical as many directories are created ad hoc and then promptly forgotten.

The Web Editor should also take a close look at the INE Web Page. While it attracts a comparatively high number of visitors, the traffic could increase much more with the following actions:

- Keeping the information up to date – the economic area still publishes numbers from 1997 for instance
- Increase the content, there are many areas of statistics not represented such as tourism, traffic accidents, environment
- Introduce a local search motor to facilitate information retrieval
- Publish the pages to the various global search engines (Google, Altavista etc) so that more hits are generated
- Ask Teledata to run scripts that gives a better idea who visits the various pages, where they come from and how long they stay
- Introduce a guest-book or other tool to communicate with the visitors
- Keep a log of all e-mail questions and guarantee that all requests are replied to promptly, whether positively or negatively.
- Observe the accepted rules of web design, i.e. quick-loading, easy-to-read pages, no big PDF-files, no dead links, having a logical site structure.

Administration

Just to inform that there exists, for more than a year, a database structure for a system “Gestão de Meios de Transporte do INE”. Whenever DARH feels there is a need to get a better management of the transport costs, the creation of the system could start.