COLLECTION OF RAW DATA

TASK FORCE

MEETING Nº 3

4 OCTOBER 2000

Doc. CoRD 024

A brief outline on existing data collection tools and new developments at National Institute of Statistics of Belgium

For information.
See Work Programme, Item 4 (Doc CoRD015)

National Institute of Statistics
rue de Louvain 44
B-1000 Brussels
Belgium

A brief outline on existing data collection tools and new developments at National Institute of Statistics of Belgium

Claude Delannoy
Aug. 2000

The views expressed in this paper are those of the author and do not necessarily reflect the policies of NIS Belgium.

1. Introduction

In the CoRD meeting of 27 March 2000 it was agreed that members of this Task Force will provide a short overview of available data collection tools and new developments in the data collection area that may be of interest to other NSIs (National Statistical Institutes). This paper provides information from NIS Belgium in this context.

2. Data collection tools at NIS Belgium

2.1 TourStat

TourStat is a software related to the monthly statistic of the belgian tourism (hotel business, camping and renting). This software was developed with dBase IV in the year 1989; it is a MS/DOS application for both the MS/DOS and the Windows platforms.

TourStat is a multilingual tool (dutch, french and german) being composed of different modules :

- management of the profile of the PSI (provider of the statistical information);
- collection of data by type of accommodation (hotel, camping or renting) with range checks;
- printing of data or metadata (essentially classification of countries from where the tourists are issued);
- sending of data to NIS Belgium;
- management of the local database (refresh, backup and restore of the database);
- help modules.

TourStat allows to create a file under a standard predefined format; the PSI sends this file on a floppy disk to the NIS. Those data, after consistency checks, are integrated in the MVS/DB2 database located on the mainframe.

TourStat must be rewritten to offer to the PSI a more user-friendly interface (GUI or Web form).

2.2 SESTA

Sesta is a electronic form related to the yearly Structural Business Statistics of the belgian Industry. This software was developed with MS-Access by the National Bank of Belgium with the collaboration of the NIS; it is a GUI application for the Windows platform.

Sesta is a multilingual tool (dutch and french) being composed of different modules :

- management of the general data related to the enterprise, the activity sector and the employees;
- collection of the accounting data according to the profile of the PSI with consistency checks;
- printing of data;
- sending of data to NIS Belgium;
- management of the local database (backup and restore of the database);
- helps module.

Sesta allows to create a file under a standard predefined format; the PSI sends this file on a floppy disk to the NSI. Those data, after fully consistency checks, are integrated in the MVS/DB2 database located on the mainframe.

2.3 Tools for the Collection of data developed with BLAISE

At NIS, many tools were or are being developed with the Blaise system from CBS. At this time, such applications are (1) in full operational use, in an experimental phase (2) or (3) planned, to collect data related to :

- Agricultural Census at the 1st December - yearly statistic - CAPI - (2);
- Agricultural Census at the 15th May - yearly statistic - CAPI - (3);
- Crop Production Statistic - half-yearly statistic - CADI - (1);
- Construction Cost Material Index - monthly statistic - CATI - (1);
- Labour Force Survey - quarterly statistic - CAPI - (3).

2.4 ePRODCOM

2.4.1 *Introduction*

Since Internet is accelerating change and create many new opportunities and since Internet is becoming the new communication medium with the public, NIS has decided to build Web-based Applications for the collection of statistical data, using the CASE tool COOL:Gen (Computer Associates).

A pilot project eProdcom was initiated at the beginning of 2000; the aim is to collect statistical information for the Prodcom Statistic (***monthly*** statistic in Belgium) related to the Industrial production in Belgium. The PSI corresponds to the local unit of the enterprise. The criteria to participate to the survey are number of employees >= 10 or

turnover >= 100 mio BEF; actually there are around 7000 PSIs for the Prodcom statistic in Belgium.
There exist 6 different forms depending of the industrial activity sector.

2.4.2 *Objectives*

The main objective of this project is to offer the PSI (Provider of Statistical Information) a new service: he can enter his statistical data via the web instead of filling out the paper form and sending it back to the NIS. As a result, one of the horizontal themes in the mission of the NIS, « Statistical Data Gathering », will be inline with the current evolution in society where communication via the web is gaining acceptance and is rapidly growing.

In this pilot project we will only offer this service for the PRODCOM statistics. This service will also be offered to the PSI without making any major changes to the current NIS business procedures.

The new service will have a user-friendly interface, which will lead to higher quality of the data that is entered. If the quality of the source data is higher then subsequently the NIS will have to spend less resources in correcting it. High quality of source data can be influenced by offering the PSI, wherever possible, the possibility to input only changes to previously entered data. Using radio boxes, listboxes, … wherever possible, forces the PSI to select only valid choices.

This service shifts the data entry activity to the source of the data, thereby eliminating possible data entry errors made at the NIS. Validating the data as much as possible at the source will result in fewer contacts with the WebPSI, thereby freeing up resources for more value-added activities.

The website has been designed to offer the possibility to use it in one of the three following languages (as does the official NIS Website): dutch, french and german. We have actually use the legal basis related to the paper forms (Does there exist a specific legal basis in this domain for official web forms ?)

The « look and feel » of the official NIS Website has been adopted as much as possible, besides the language choice.

Besides the functional objectives who are more related to the actual application, there are also a number of objectives targetting the way the NIS handles projects like this.

A first objective is to prove that the technology to take MVS transactions to the Web works.
A second objective is to set out a path to acquire the knowledge to implement projects like this.
A third objective is to prepare for future similar projects by setting standards and procedures.

2.4.3 *Technical architecture*

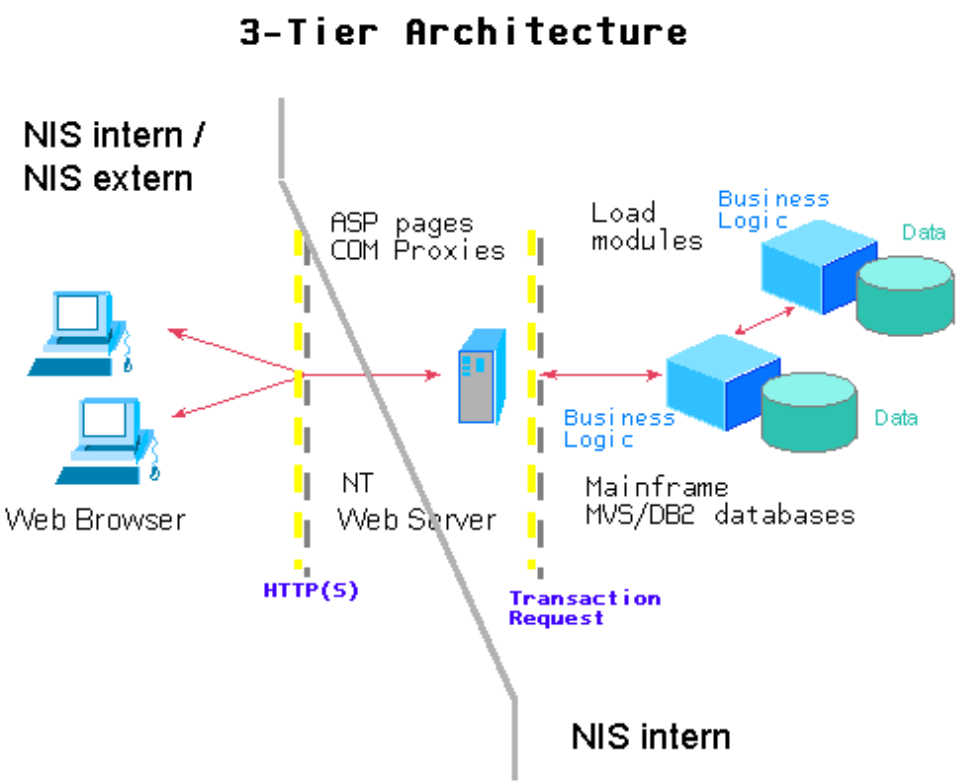
Technically the application has a 3-tier architecture:

- The front-end
    - Webbrowsers displaying the HTML-pages that represent the PRODCOM forms

    - WebApplication must function on different flavors of Webbrowsers, eg. IE, NetScape

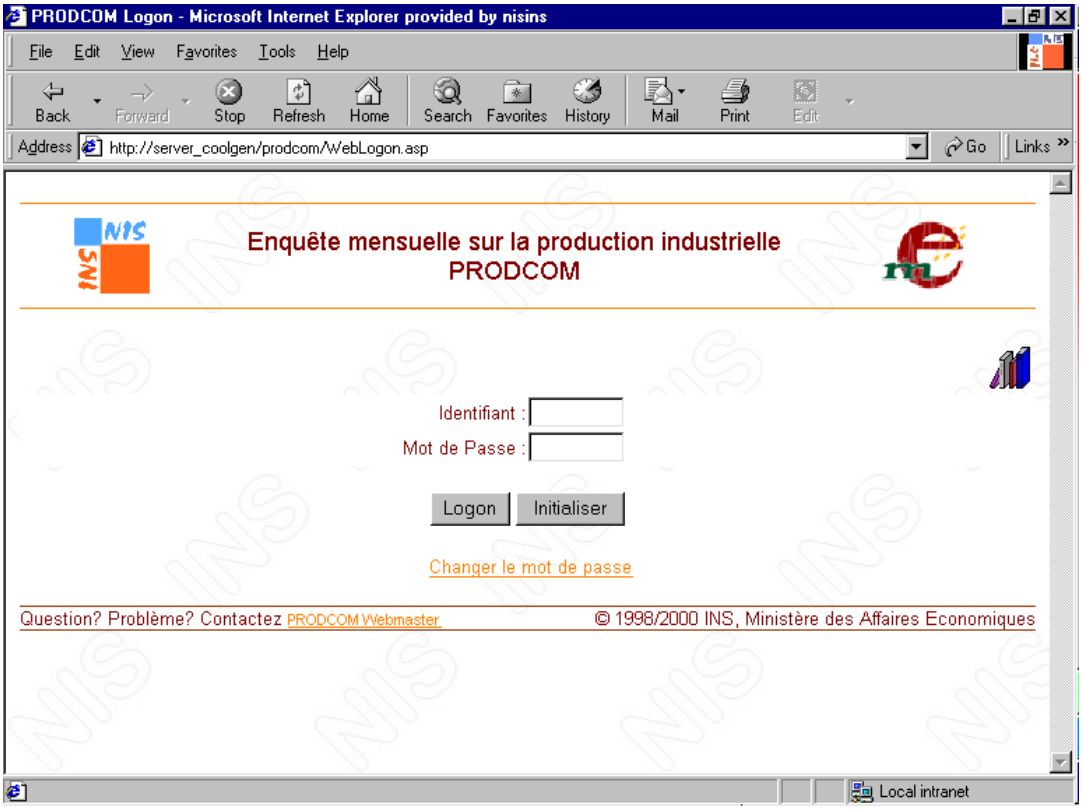
- The Middle-Tier
    - The WebServer is IIS
    - The WebApplication's navigation logic is implemented in ASP-pages
    - Using COOL:Gen's COM-proxies, the ASP-pages can access business logic located on …

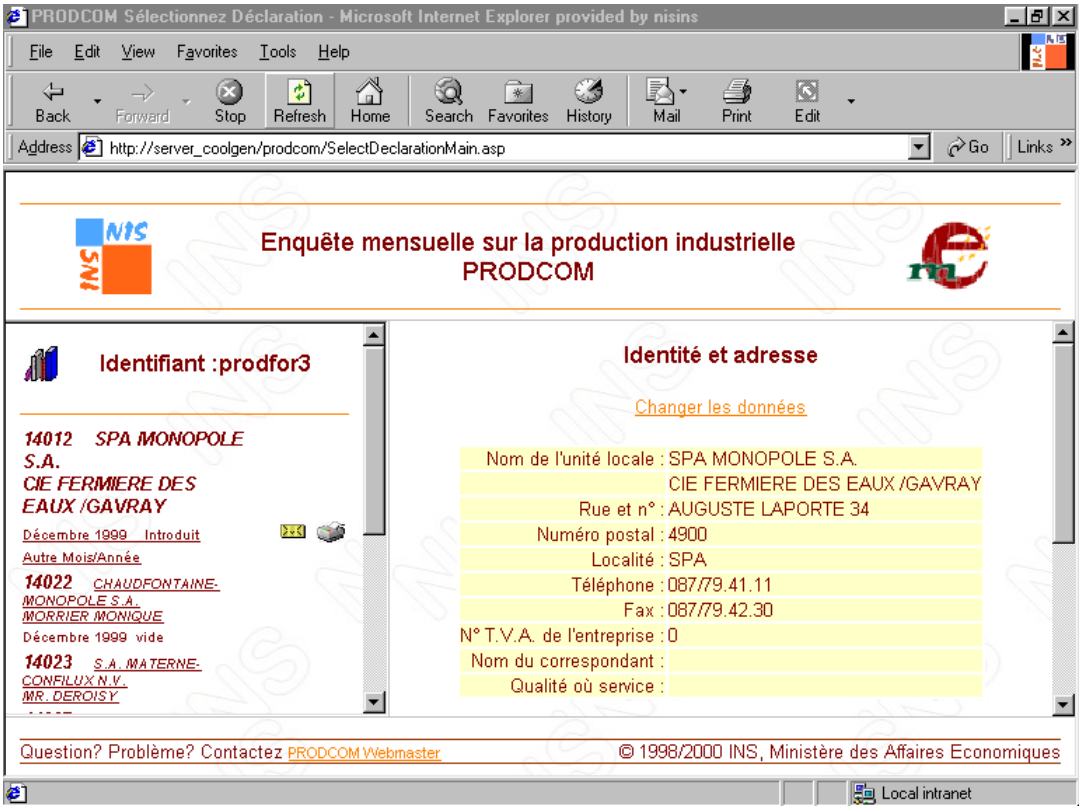
- The Back-End
    - The business logic is implemented in COOL:Gen transactions (Server Load Modules). These load modules reuse existing business logic used already in the on-line PRODCOM application.
    - The production data store is MVS/DB2



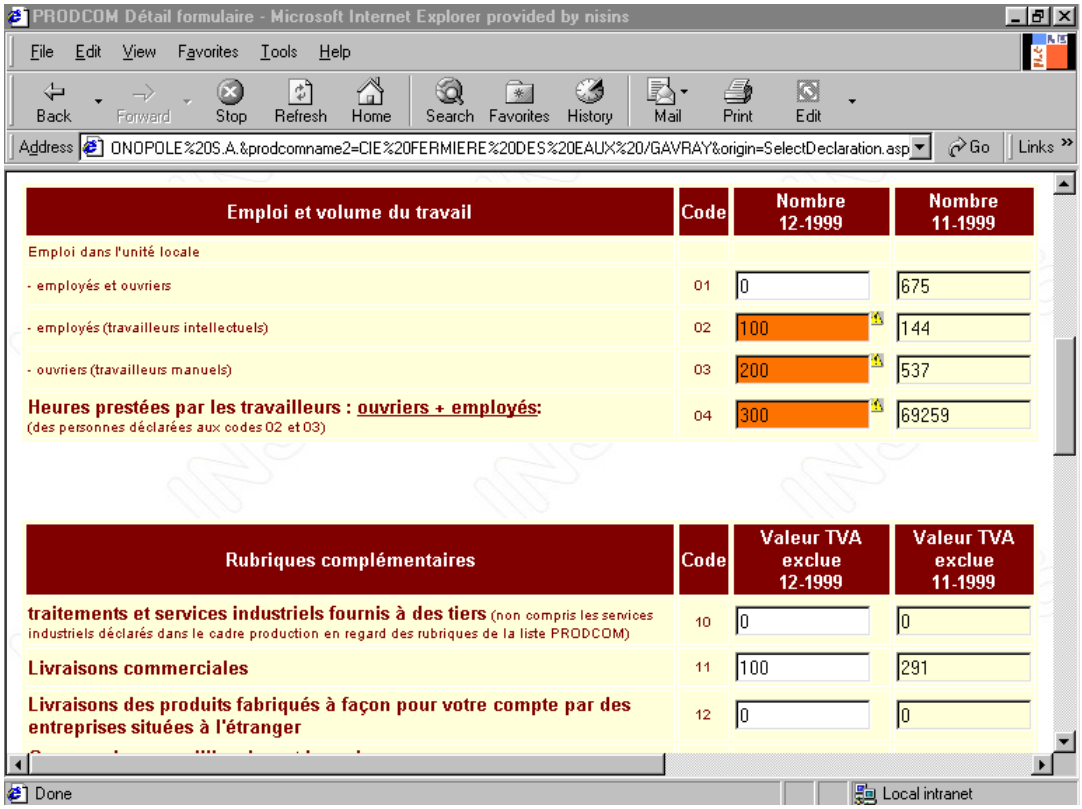
2.4.4 *Short description*

After having logged on the NSI Website (userid and password must be entered), the application shows all the PSIs (local statistical units) managed by the WebPSI and allows the WebPSI to select a declaration still to be "inputted" and/or "updated" or to update the address information for a PSI he manages.

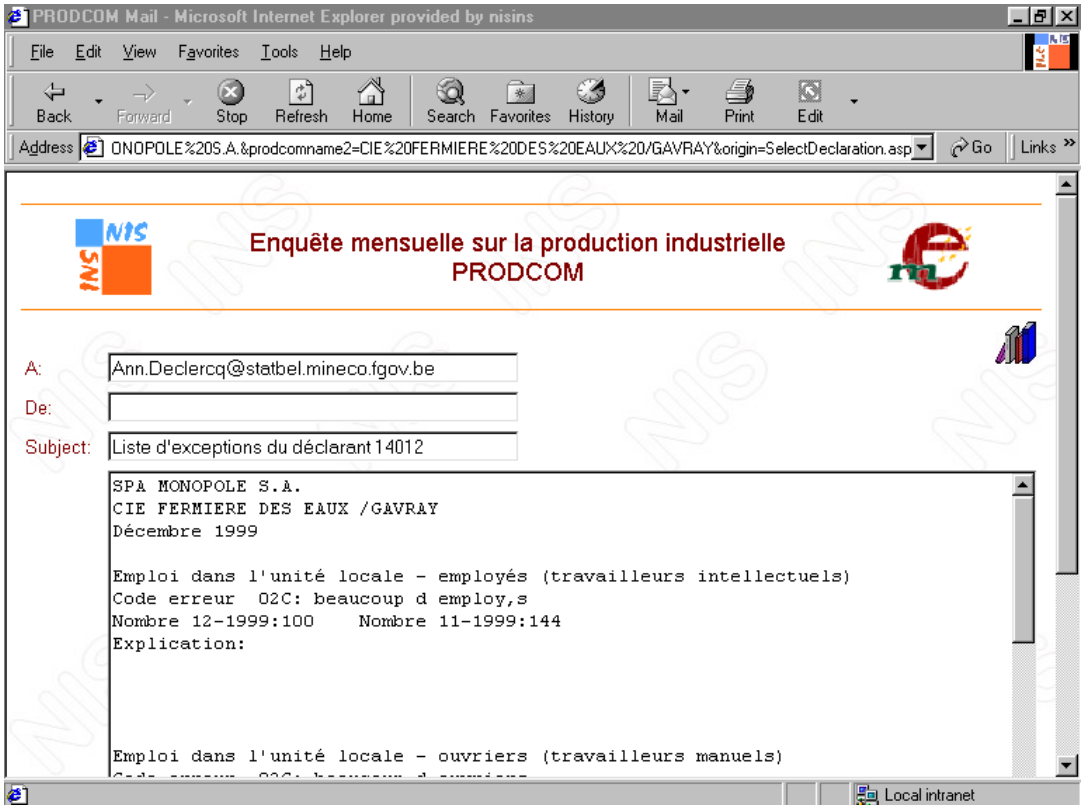




The declaration related to a specified statistical period is displayed according to the PSI profile (one of the six different forms) and with the values of the data related to the previous period.
Consistency checks are done at the browser.



After submitting a new or an updated declaration, the data are directly stored in the MVS/DB2 database on the mainframe (this allows to mix data issued from Internet with data related to paper-forms entered inside the NIS). Simultaneously an e-mail is automatically generated with the remaining “errors”; this e-mail can be edited and updated by the WebPSI to explain the “errors”. After that, the WebPSI sends this e-mail to the responsible at the NIS.



An on-line help related to the survey and to the functioning of the application is also provided to the WebPSI.

The eProdcom projet is in an experimental phase. Some WebPSIs are candidate (β-testers) to test the Web application. Those tests are forseen in the short term.

3. General consideration

Will the NIS capitalize on benefits of eg. « Statistical Data Gathering via the Web » by rethinking its internal procedures? Or in short: will it do **eGovernment Process Engineering**?

Will the NIS use the Internet as an important communication medium towards suppliers (PSI) and customers (PSI?, Government, Universities, …)? Or in short: will it have an **Technology Architecture** project?

Will the NIS offer « digital products » in the near future? Or in short: will it have an **eGovernment Strategy** project

eGovernment : definition
The priority of the Government is to reduce the administrative burden of the entreprises and the citizens.
A way to achieve this priority is the use of new IT-technologies such Internet technology.
eGovernment is thus only one of the common pre-occupations of the Government.

NIS environment is by nature a digital environment.
NIS can thus offer to his clients digital products like statistics on demand.
The use of Internet as communication medium can increase the imago of the NIS.
What's more, Eurostat and the others european NIS are orienting their strategy and developments in this direction. NIS Belgium must become integrated into the ESS.

XML can be seen as EDI over the Internet.
Offers the PSI's the possibility of automating the data entry (data source = accountancy package or ERP-package or …)

Once it is decided that eGovernment is part of the strategy (and mission) of an organization, a lot of questions will arise on how to do it? The most important question will be how eGovernment can add more value to the public services that are offered. But this is not the only question… a whole lot of them will pop up and they all have to be answered at the right level and at the right time.

Three levels are of importance:

- The business level

  How will eGovernment change the way we work inside the organization? How will it change our relations to the outside world? Will there be new business processes to implement? And how shall we implement them?

- The application level

  What applications will we need to support these new business processes? Are they existing applications that will be web-enabled? Will they be bought or build? How will we build them?

- The technical level

All these applications that we need, with what kind of technology will they be implemented? Will we use existing platforms and technologies? Must we implement new platforms and new development methods? How are we going to integrate all these applications and platforms?