

Changing Statistics Netherlands: Driving forces for changing Dutch statistics

Gosse van der Veen

Director General, Statistics Netherlands, Postbus 4000, 2270 JM, Voorburg, The Netherlands

E-mail: akin@cbs.nl

1. Introduction

Statistics Netherlands is in the middle of radical changes. The increased complexity of society and the rapid changes it is going through imply an increasing demand for reliable statistical information. The focus is shifting towards thematically presented information providing more insight into how developments and sectors are related. Statistics Netherlands wants to speed up its flexible response to the changing needs for information. This requires combining data from various sources. Data gathering is shifting from primary to secondary collection of data (i.e. Data from administrative registers). In the future, Statistics Netherlands will be able to receive, process and store much larger quantities of register data.

Statistics Netherlands laid down its vision on statistical production and information development in the document 'Statistics that count' (Statistics Netherlands, 2001). These developments have now reached the point where an outline of a Statistics Netherlands information and business architecture has been developed.

2. Driving forces behind change

Statistics Netherlands is, like all other government institutions, confronted with ongoing budget reductions due to the intention of central government to increase efficiency of government agencies in general and to reduce the number of government employees. The possibility to influence our statistical programme is diminishing due to European legislation in the field of statistical information needs of the European Commission and other European institutions. Currently, about 70% of the work programme of Statistics Netherlands is, one way or the other, subject to European legislation. Globalization makes business surveys more complex and the availability of more and better registration systems changes the statistical processes. In addition to these developments, a strong need was felt to adapt our statistical work programme to the changing information needs of Dutch society. Finally, we are heavily involved in the government policy to reduce administrative burden to Dutch society. In the period 2002–2006, a further reduction of 25% had to be realised. In order to meet all these requirements, a number of complex actions have started since 1999.

Although the legislative position of statistics Netherlands is historically strong, the need was felt to increase the independence from central government. Therefore, a new Statistical Law was drafted which guarantees independence from central government more effectively. According to this law, the Director General is solely responsible for methodology and dissemination of statistics; a financially more independent position is guaranteed by a fixed budget for a period of 5 years; Statistics Netherlands is the only national institution to produce statistics which meet European legislation; increased obligations in this field will result in additional budget funding; and free access to (government) administrative sources (registers) for the production of official statistics is guaranteed. In January 2004, Statistics Netherlands became an independent, autonomous government agency. The position before the Central Commission for Statistics is now stronger, as Statistics Netherlands is responsible for governance. However, the traditional position where the commission had to approve the statistical programme still remains. The new Statistical law came into force in January 2004.

The next response was in 2002, when it was decided that the new multi-annual programme 2004–2008 should respond to the need for adaptation of the statistical working programme to the rapidly changing information needs of Dutch society and policy-makers. The only way to find the budget for new activities was to intervene in the programme by radically ending certain activities or to reduce activities in a way which inevitably would result in less output (lower frequency, lower level of detail). In close co-operation with our advisory boards – representing users in certain fields of activities – we managed to reduce our programme in such a way that more than 175 full-time employees could be used for new programme subjects. To identify these subjects, Statistics Netherlands invited 50 key players from all fields of society to write an essay about their view on the future (statistical) information needs of Dutch society. Statistics Netherlands published the essays in a publication ‘Steering on Statistics’.¹ This publication, together with the views of our advisory boards and internal opinions within the organization, were used to depict the following spearheads for the future:

- a) Social dynamics and labour accounts;
- b) Micro data infrastructure (remote access facilities);
- c) Integrated system of health and welfare statistics;
- d) Social and economic implications of ageing;
- e) Knowledge-based economy and firm-based productivity;
- f) Statistics on justice and security;
- g) Trade and services;
- h) Spatial statistics (including mobility between spatial units).

The final driving force was the obligation to meet the government standard for the reduction of administrative burden (–25% over the period 2002–2006). Although Statistics Netherlands reduced the burden caused by statistics considerably (–60% over the period 1994–2002), new opportunities, mainly resulting from the new statistical law, 2004, occurred:

- a) Use of registers and administrations as the primary information source became possible because of free access to these registers.
- b) Close co-operation with other government register owners (e.g. tax register), created huge possibilities.

¹Publication ‘Sturen op statistieken’ (Steering on Statistics.) Published in Dutch, on Statistics Netherlands website www.cbs.nl. ISBN 90-357-2628-6.

- c) The introduction of a “Statistical holiday” for small and medium-sized enterprises (SMEs). Once SMEs participated in a survey for a certain period of time, they were exempt from participation for a number of years.

With these instruments, an additional reduction of 28% in the period 2002–2006 was realised. Statistics Netherlands is especially proud that our model to calculate the administrative burden is adopted by Eurostat as the EU-Model.

Finally, there is one remark to be made on this issue. The perception of the public on administrative burden caused by Statistics Netherlands does not equal the real burden (0.15% of the national total). Communication is very important and will have special attention in the years to come.

3. European Statistical System developments

Since 1957, Europe has put its demands for statistical information on the agenda. Due to the fact the European requirements have a serious effect on our work programme – more or less 70% of our work is determined by European obligations – this chapter gives some background information on the European developments.

Since the start of the European Community, the need for reliable and comparable statistics was evident. The European Statistical System (ESS) was built up gradually with the objective of providing comparable statistics for the European Union. Eurostat, the statistical office of the European Union, is one of the directorate-generals of the European Commission; it was initially set up in 1959.²

In the early days, the European demands concentrated on data for the Coal and Steel Industry, but expanded more and more with the extension of the policies under the treaties. By the end of the 1980s and in the early 1990s, the need for European statistics soared for several reasons. First the European Monetary Union (EMU) required far more data to monitor the union. This led to the Harmonised Indices for Consumer Prices (HICP), a broader coverage and increased level of detail and quality of financial and non-financial accounts for government and improvement of short-term economic indicators. On the other hand, the Lisbon agenda of the EU set the need for a large set of structural indicators covering the main areas of the EU policies, which also required development and implementation of a large range of statistics.

EU statistical requirements are laid down in regulations, which are legally binding for the EU member states and need to be implemented by EU candidate countries. The directors general of the EU statistical offices, represented in the Statistical Programme Committee (SPC), were often pointing to the fact that the new European requirement could hardly be met, given the fact that resources were not on an equal footing, or were even declining. However, it lasted until 2006 before serious attempts were made for a reprioritisation of the EU’s statistical needs and modalities were found to meet the concerns of the National Statistical Institutes and the respondents.

Statistics Netherlands integrates the European demands in its statistical work programme. In the Statistics Acts it is recognised that Statistics Netherlands is the primary producer of Community Statistics in the Netherlands and compensation is paid for meeting new EU requirements.

²From ‘Statistical Portrait of the European Union 2007, 50 years of the Treaty of Rome establishing the European Economic Community. Luxembourg 2006.

4. Redesign of the statistical production process

As summarised in Section 3, there were several ‘driving forces’ that led to the conclusion that the production system of statistics needed thorough rethinking and redesign.

The first sketches for redesign of the statistical process were already developed in the mid-1990s. In the publication ‘From assembly line to electronic highway junction; a two-track transformation of the statistical process’,³ the architecture for the new production process was outlined. It was explained that the innovation process is based on two widely recognised assets: data collection technology and integrated data frameworks. The change process from traditional data collection methods through surveys amongst respondents to a three-fold approach was outlined. The three-fold approach encompasses electronic data capture directly from external administrations, data extraction from non-statistical, administrative registers and the set-up of the flexible, multi-purpose survey synthetically harmonised with the available source registers. The optimal outcome of this restructuring is reflected in integrated statistical outputs available in an online databank. A first version of the StatLine database became available for testing by external users in 1996. Now, early 2007, StatLine is the main source of statistical information for all users, providing a treasure of statistical output – free of charge – to the users and the public at large in the Netherlands.

The basic philosophy for these developments is the notion that the main asset of Statistics Netherlands is the production of integrated and coherent data on economic and social phenomena.

In the summer of 1999, it was decided to reorganise Statistics Netherlands and the new structure was implemented in 2000.⁴ The new structure encompassed both the organisation of Statistics Netherlands and the statistical production process as such. Until 1999 Statistics Netherlands was organised according to statistical branches. There were a large number of “stovepipes”: one for each individual statistic. Standardisation of the statistical processes made a more efficient structure possible. Within the production sphere, similarity of processes is the leading criterion for distinguishing intended production divisions. A distinction is made for input and throughput processes for data on persons and data on businesses. This led to the set-up of separate divisions for Business Statistics (BES) and persons-oriented Social and Spatial Statistics (SRS). The spatial statistics were included in the latter division mainly for reasons of balance with respect to size. A third production division was installed for macro-integration, publication and communication. Thus, the division of Macro-economic statistics and Publication (MSC) emerged. In the year 2000, a special division of Technology and Facilities was created. It encompassed both the IT and most of the methodology staff. In 2006, it was decided to separate methodology from IT functions and a separate division of Methodology and Quality was created. The role of IT then also changed as was explained in Section 2.

In the following paragraphs, developments in the field of business statistics and subsequently social statistics are described in more detail.

4.1. *The current production process for business statistics*⁵

³Based on the Van cijferfabriek naar knooppunt op the elektronische snelweg, the Strategic Development Plan ‘CBS 2000’ of Statistics Netherlands, written by C.A. van Bochove.

⁴Towards a New Statistics Netherlands; blueprint for a process oriented organization structure. Ad Willeboordse, 24 May 2000.

⁵This part of the paper is mainly derived from the Paper ‘The embedding of a uniform statistical process’, contributed paper for the UN/ECE Work Session on Statistical Data Editing, Ottawa, Canada 16–18 May 2005. Prepared by Jeroen van Velzen.

- In 1999, Statistics Netherlands started a project called IMPECT (IMPlimentation of EConomic Transformation process). This project aimed to redesign and standardise the questionnaires and the logistical and editing processes. In the first phase, the project focused on the annual structural business statistics and in the second phase short-term statistics were included.
- The first phase included almost all branches (e.g. manufacturing, construction, trade, transport, services). The degree of detail of the output was described beforehand (in terms of NACE codes) in order to be able to optimise statistical methods. In the second phase, short-term statistics were added to the system, this is the so-called IMPECT 2 system.⁶
- For the redesign of the statistical process, the layout and structure of all questionnaires were harmonised. A central database has been built to describe which specific questions are asked to the different kinds of enterprises (ordered by size and branch). From this database, both questionnaires and interfaces for data entry and editing are generated. For the definitions of the variables, the system refers to a central database containing Meta information.
- In the following stages of the process, the questionnaires (now called records) are corrected for obvious mistakes. Firstly, these mistakes and corresponding corrections are described by subject matter experts. Secondly, data that are edited interactively are used to confirm these corrections. In order to be able to discriminate between important and less important records, a plausibility indicator is calculated. The plausibility indicator consists of seven partial plausibility indicators, each representing a certain aspect of a record.
- Records which score as an implausible record and records of large enterprises are edited interactively. A Blaise application is being used for guiding this process. The other records – mainly the smaller enterprises – are corrected automatically (using SLICE components).
- The non-response of large enterprises is corrected by imputation. The records are estimated using the record of the previous year combined with a factor for development, which is estimated based on short-term statistics. The BASCULA[®] component of Blaise is used to assign weights.
- In the final stage, a central database is filled with all records on both annual structural business statistics and short-term statistics. This is the Economic Statistics Database (EDS), thus consisting of a large number of related registrations with economic business data from various sources, linked and integrated data at meso and micro level, supported by clearly defined data sets.
- The EDS is used to generate tables for StatLine (the output database of Statistics Netherlands, which can be accessed through the Internet www.cbs.nl).
- In 2004, the results of the redesign process of business statistics were evaluated. The conclusions of the evaluation were mixed: for the core of business statistics, the results were plausible, but for other areas the standardisation turned out not to be feasible or applicable. In general, the manageability has been improved and the maintenance costs reduced. However, the time for implementation of the new process turned out to be overrun. It was impossible for the old and the new system to run parallel. This made it also impossible to cross-check with the data from the old system. Initially, this led to difficulties in testing the robustness of the results.
- At the moment of writing, early 2007, the new standardised process is applied to one third of the total of business and short-term statistics.⁷

⁶IMPECT: Recent developments in harmonized processing and selective editing. Supporting paper UNSC/ECE-CES Work Session on Statistical Data Editing, Madrid, Spain, 20-22 October 2003. Prepared by Arjan de Jong.

⁷The functional statistics are not included in the Impect process. These are for example: statistics on agriculture, energy, environment, transport, R&D, Prodcom, and Intrastat.

- A new boost to further streamlining and standardisation will be given through the Masterplan (the work ahead), which was discussed in Section 2.

4.2. *The current production process for integrated social statistics*⁸

- For the broad range of social statistics (statistics on persons and households), Statistics Netherlands developed the Social Statistical Database (SSD) as the core system for the integration of data from both surveys and registers.
- The SSD is primarily based on register information and data from persons and household surveys. The registers frequently contain complete information on all relevant units. In the Netherlands, this is surely the case for demographic data, income tax data, participation in the labour market, dependence on social security benefits, participation in education and housing facilities.
- The files of the Population Register constitute the backbone of the database, as all the other files are linked to this register. Linking on a personal identification number appears to be successful: approximately 99 percent of the records are linked. If such an identification number is absent, the sources are linked on postal code, house number, date of birth and gender. The result is that approximately 95 percent of records are linked. The linked files form the baselines for the work process of the SSD.
- To collect survey data as efficiently as possible, the combined register data are increasingly used as a sample frame for household surveys. The data are pre-stratified for this purpose. In order to improve the linking procedure, personal identification numbers have recently been included in the samples taken from the Population Register. Then the success of the linking procedure is almost 100 percent.
- The SSD has made it possible for Statistics Netherlands to produce integrated statistical outputs at relatively low costs. An example is the output for the 2001 census round, which was almost completely based on register data.
- Because ultimately the SSD will comprise a very detailed picture of every inhabitant of the Netherlands, data security and privacy are important issues. Statistics Netherlands cannot take the risk that individual data are disclosed, because this could lead to image problems in the media. Besides, there are legal conditions which prevent the disclosure of individual data by Statistics Netherlands. Therefore, a strict security regime has been implemented.

4.3. *The role of metadata*

The move of Statistics Netherlands from a product-oriented to a process-oriented organisation structure, as well as continuing budget cuts, reinforced the need for further standardisation and harmonisation of concepts and processes. Currently, the leading idea is that metadata play a key role in achieving coherence of statistical outputs.⁹ Metadata servers will be required guiding the process in all stages, i.e. data collection, processing, analysing and dissemination. It is expected that the Metadata servers will be gradually implemented. Coherent concepts and definitions are crucial in the process. Therefore, a dictionary' containing all concepts and definitions needs to be created. The accomplishment of such a dictionary is complex and subject to intensive discussions in the office.

⁸Based on the contribution of Statistics Netherlands to OECD Newsletter, prepared by Bert Bakker.

⁹For a more elaborated article see 'On the role of Metadata in achieving better coherence of official statistics'. Ad Willeboordse, Peter Struijs and Robbert Renssen.

5. Staff development

In the past 15 years, the number of employees has decreased considerably. Partly due to ongoing budget reductions, partly also because of the need for higher educated staff. Further staff cuts are anticipated for the period 2007–2012. Staff and management skills will shift as Statistics Netherlands becomes more service-oriented. There will be new roles and new responsibilities. We expect to see new procedures and new technology. The focus will be on change management, design and statistical analysis and on process and quality management. It is imperative to train management and staff in these new aspects. A good communication plan to support this programme is required.

6. The work ahead

Statistics Netherlands used the methodology of the Integrated Architecture Framework (IAF) to analyse the information, business and Information and Communication Technology (ICT) demands. The IAF framework is a key element in our approach to architecture. It gives a concise, complete and practical overview of all relevant aspects that are important in the discussion about information and ICT. It is useful for the development of systems, but its main function is to show the relations and interdependencies of the organisation and ICT in all its major elements. The framework is suited to all types of architectural work. For many architectural efforts only part of it will be of paramount importance. Still, the overall framework will provide the right scope and positioning of the aspects and areas involved and the way they need to be included. The framework is applicable at enterprise level of architecture, as well as at subject matter domain and project level of architecture. It describes a basic design of the statistical process at an abstract level in terms of activities, functions and objects using a uniform terminology. The key characteristics are:

- a) *Services*: concentration of similar functions in generic services. This means that the design does not focus on the difference between social and business statistics. Over time, this will have consequences for the organisation of the work.
- b) *Rules*: control of the process by business rules and chain management. The objective is no longer to maintain the software, but to maintain the rules.
- c) *Transformation*: focus on input-output transformation. The renewed statistical process will be organised as a sequence of services.
- d) *Openness*: openness to data providers and information consumers and connection with national registers

In our newly designed architecture metadata plays a central role. We use the descriptive, process, and quality metadata to design and structure our statistical process. The metadata is used to describe the processes from raw data to statistical output. The main services that support the implementation are a metadata service, a generalised data collection service and a data storage service. An important implementation guideline is that the business rules are separated from the software code.

Statistics Netherlands is now implementing (in a test environment) the following general services:

- i) a metadata server
- ii) data storage
- iii) output database (StatLine), and
- iv) data collection

Furthermore, Statistics Netherlands is implementing the necessary tools for Statistical ETL (extract transform and load), BPM (Business Process Management) and BRM (Business Rule Management). The current view on the future set-up of Statistics Netherlands shows several design principles. Statistics Netherlands opts for:

- *Product*: support of all statistical products and a fast and flexible response to changing demands for new products. ICT processes must support all statistical products, and the processes must be adjustable (relatively) fast in order to meet changes in demand. This is in line with the demands of our customers, who want faster, more specific, integrated statistics.¹⁰
- *Organisation*: a Statistics Netherlands-wide, generic set-up of processes and ICT support. ICT support must therefore be independent of the organisational structure. Research shows that primary data collection among persons and enterprises can be incorporated in one universal, common model. The ICT solution has to support this. The number of applications used must be reduced and it must be cheaper to maintain the infrastructure and the applications.
- *Process*: preventing duplication of processes (no “stove-pipes”), implementing transparency by adding process data and leaving room for exceptions. In general, the processes must be more flexible, so that client orientation will improve.
- *Methodology*: all methodology used is approved and standardised. For this purpose, we build a central register of best practices.
- *Environment*: addressing the increasing diversity of suppliers by supporting a variety of channels and formats, scalability to process large quantities of input data and to add new suppliers in a simple manner, scalability for processing and output production and preparing for direct integration with external sources (for instance, direct integration with the basic registrations).

The most important challenges are:

- a) Changing management from ‘thinking’ to ‘doing’.
- b) Keeping it simple.
- c) Management of distortion of trends in statistics.
- d) Management of expectations.
- e) Management of management.

7. Conclusion

This paper provides a brief overview of the driving forces for changing Statistics Netherlands and the measures which have been initiated. Although external pressures, such as budget cuts and the strong emphasis on reducing administrative burden have a strong impact, our organisation needs to be prepared to take initiatives of its own. An ongoing issue is the ability to adjust to the changing needs of society for statistical information and the need to improve efficiency. In the coming years, the process of modernisation and the implementation of new technologies will determine the agenda of Statistics Netherlands.

¹⁰Strategic survey, December 2006.

Acknowledgements

The author wishes to thank Ada van Krimpen, André de Boer and Dick Meuldijk for their contributions to this paper.