UNITED NATIONS **E**

Economic and Social Council

Distr.
GENERAL

ECE/CES/GE.42/2007/3
5 April 2007

ENGLISH
Original: FRENCH

ECONOMIC COMMISSION FOR EUROPE **STATISTICAL COMMISSION**

CONFERENCE OF EUROPEAN STATISTICIANS

Meeting of Experts on Business Registers

Tenth meeting
Geneva, 18-19 June 2007
Item 6 of the provisional agenda

QUALITY IMPROVEMENTS IN BUSINESS REGISTERS AND IMPLICATIONS OF REVISIONS OF NACE (NOMENCLATURE GENERALE DES ACTIVITES ECONOMIQUES DANS LES COMMUNAUTES EUROPEENNES) AND INTERNATIONAL STANDARD INDUSTRIAL CLASSIFICATION (ISIC)

IMPROVING THE QUALITY OF BUSINESS REGISTERS

Submitted by France

The meeting is organized jointly with the Statistical Office of the European Communities (Eurostat) and the Organisation for Economic Co-operation and Development (OECD)

SUMMARY

The present paper was prepared by the French national institute of economic and statistical information (INSEE) on request from the CES Steering Group on Business Registers for presentation and discussion at the joint UNECE/Eurostat/OECD Meeting in Geneva, 18-19 June 2007.

On the basis of experience with the French register, INSEE shows in this paper how it is possible to improve the quality of business registers for both inter-administrative and statistical purposes.

GE.07-21642 (E) 220507 230507

I. INTRODUCTION

1. Using the French *Système national d'Identification du Répertoire des Entreprises et de leurs Etablissements* (SIRENE) business register as an example, this document aims to show how the quality of a business register used for both inter-administrative and statistical purposes can be improved.

2. After a brief presentation of the SIRENE register and its contents and purpose, we will discuss the concepts of quality of a product and quality of a register. The report then goes on to describe how a "quality approach" has been developed for managing the SIRENE register. It then describes the added features of the new SIRENE management system (SIRENE3 project) in terms of improved quality. Lastly, some examples of operations carried out are given, together with indications of volumes involved.

II. DESCRIPTION OF SIRENE

3. All businesses (*entreprises*) and their local units (*établissements*) in France are registered in a national business register called SIRENE.[1] This register is used in the context of the statutory framework in which INSEE is responsible for identifying businesses (legal units) and their local units, on the basis of documents transmitted by *centres de formalités des entreprises* - CFEs (centres for business procedures), which were set up to deal with all administrative procedures falling within their remit. An act of 11 February 1994 established a unique identification number, which is now at the core of the inter-administrative system, since use of this identification number is mandatory for relations between businesses and administrative bodies.

4. The SIRENE register lists the civil status of businesses: the family name, first name, date and place of birth of the natural person (in the case of an independent business) or the company name (in the case of a legal person). Also listed are the legal form of the business, the address of its headquarters, the date it was set up and, where appropriate, when it ceased trading. The address of local units is indicated, as well as the date and purpose of their creation. In addition to these identification data, the register contains the APE (*activité principale exercée*) code assigned by INSEE identifying the principal business activity carried out, based on the NAF (*Nomenclature d'Activités Française*) French general economic classification, for each of the units listed, and the number of employees on the payroll.

5. In addition, since the creation of SIRENE, the French national institute of economic and statistical information INSEE has ensured that the business register serves both as an identification system for, and as an ongoing census of, business activities, since it can also record events in the life of a legal unit, such as restructuring or hiring of employees.

[1] Established by amended decree No. 73-314 of 14 March 1973.

6. This wealth of resources means that in addition to the statutory uses of SIRENE related to the inter-administrative nature of the register, it is also used in the areas of statistics and trade. For this reason, it was decided to record additional data for each unit registered in SIRENE, in accordance with a list drawn up after consulting statisticians using the system. Thus, in addition to identification variables, the business register includes economic classification variables (economically important units, units having employees, etc.); variables relating to the number of employees (number of employees when set up, number of employees at each year end, equivalent full-time posts, etc.); turnover; variables showing links between local units or legal units (e.g. own-risk *location-gérance* leasing arrangements); and management variables (particularly for management of dissemination of information and data confidentiality).

7. A great many data are thus supplied by the inter-administrative system, others come from statistical sources and as at 1 January 2006, some 6.72 million businesses and 7.66 million local units were identified in the register.

III. MOVING FROM QUALITY TO A "QUALITY APPROACH"

8. The aim here is not to give an exhaustive or methodical presentation of the concept of quality but to describe aspects of the concept that are relevant to a business register, and how all this leads to the development of a "quality approach". It is based on the work done by a working group on this theme in 2000.

A. Product quality

9. Although used frequently, the word "quality" is broadly interpreted. If we take the definition used in international standards,[2] the quality of a product or service is defined as the "totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs". This definition shows that:

(a) Quality constitutes a totality of characteristics that are objective, but must be measured;

(b) The purpose of these characteristics is to satisfy the needs of users.

10. Quality is therefore a question not of *maximum* performance but of respecting *specified* - and thus expected - performance.

11. It should be noted, however, that for a given product, not all users have the same expectations, or the same ranking of priorities.

12. Moreover, quality can be measured in different ways: measurement of characteristics, enumeration of anomalies, evaluation of user satisfaction.

[2] ISO 8402.

B. A quality approach

13. Viewing quality in terms of a product has its limitations since, for the user, it is expensive to carry out systematic verification, and for the producer, rejection of a product means loss of money and time. It is therefore at the production stage that quality must be ensured. Consequently, we can differentiate between “quality control” and “quality assurance”.

14. Quality control is an internal objective. The company determines and implements the measures necessary to ensure that the product has the desired characteristics; in other words it reduces, and rectifies, errors.

15. Quality assurance concerns the relationship with the customer (who may be internal). It is in a way the systematization of confidence between customer and supplier.

16. It is also essential that this systematization should be efficient. To ensure that everybody is working towards the same goal, all stages must be specified and described in instructions that are as precise as possible. Training of staff, involving them and raising their awareness of the quality approach all play an important role. This is a standard quality control strategy known as Plan-Do-Check-Act, or PDCA, which proposes a methodological approach to quality control, as part of an ongoing process of improvement.

17. It should be emphasized that this kind of quality approach has both a spatial and a temporal dimension. A spatial dimension, since it involves all players in the production process: it is not enough for everyone to “do their best” (in their own eyes) for the overall final result to be satisfactory. For, on the one hand, there is no guarantee that one person’s priorities will be the same as the next person’s and, on the other, it is the systematization that is important. This is a paradoxical aspect of quality:

(a) Everybody is part of the quality effort, in the sense that all individuals involved in the production process do their job to the best of their ability;

(b) Nobody is part of the quality effort, in the sense that a quality approach requires stated, coordinated, standardized actions.

18. The temporal dimension is essential: a quality approach requires in-depth work, the results of which are not all immediate. It is the opposite of a “lightening raid” operation. How the relevant indicators evolve is perhaps more interesting than their actual level.

C. Inventory of requirements

19. As the definition of quality in the relevant standards indicates, quality only has meaning in terms of users’ expectations. When looking at quality in relation to SIRENE, it is essential to distinguish the different purposes of the register:

(a) Inter-administrative use;

(b) Dissemination of information contained in the register;

(c) Statistical use.

Inter-administrative use

Requirements of administrative partners involved in managing the register

20. The term “administrative partner” is used to describe the statutory partners (centres for business procedures, social security bodies or tax administration). Their requirements relate to the register itself but also to how the system operates as a whole. The three main ones are:

(a) Registration at the earliest possible opportunity;

(b) Unique identification number for each unit;

(c) Consistency between SIRENE and the legal and administrative registers for businesses (business-names register or trades register, social security and tax files), which is ensured by the mandatory unique identification number.

Requirements of businesses

21. The requirements of businesses dovetail to a large extent with those of the administrative partners: rapid registration; unique identification number; consistency with other administrative and legal files. Businesses also have other specific concerns, such as:

(a) The code describing their principal activity. Businesses wish to be assigned the activity code they consider most appropriate (this is the most frequent source of disputes with INSEE);

(b) The reference to economic termination in the register. If erroneous, this can penalize businesses, given the public nature of the register.

Requirements concerning dissemination of information contained in the register

22. One of the articles of the decree establishing SIRENE provides that INSEE shall give the public freedom of access to data contained in the register (other than the date and place of birth of natural persons). The staff responsible are particularly concerned to ensure that the information contained in the “address” variable is correct, and that it complies with postal standards.

23. However, while they attach importance to the reliability of addresses (no doubt owing to the immediate an obvious “penalization” effect caused by an incorrect address), the services responsible also want to ensure that the register provides a correct activity code and a reliable indication of the size of a business.

24. Another concern frequently raised by information dissemination services relates to the ability to respond quickly to users’ criticisms of the data, for example where incorrect or invalid procedures or missing variables are concerned.

Requirements concerning the use of SIRENE as a statistical register

25. SIRENE is, of course, widely used for statistical purposes.

(a) Business demography

By definition, SIRENE is used to monitor business demographics: analysis of the economic climate (statistics for business start-ups are published monthly, and are closely followed by economists and politicians); structural statistics relating to the demography of businesses and local units; description and monitoring of the characteristics of entrepreneurs and of new businesses.

(b) Basis for sampling

Its exhaustiveness and content enable the register to be used as a reference database for sampling purposes. Identification data (name, address) and stratification data (activity, size, region of location, turnover) are therefore essential and must be of good quality. In addition, to meet the requirements of statisticians, they must be supplemented by metadata (date and source of latest update, quality indicators, etc.).

SIRENE also aims at a standard error of zero for the largest units (for all variables, including address, number of employees, etc.).

(c) Survey managers

When utilizing data from administrative (tax or social-security) sources or from surveys, statisticians may sometimes query data in the register. They often have access to more recent data or are able to observe discrepancies, without being able to account for them (local unit address, for example, or economic data).

IV. RE-ENGINEERING SIRENE, TAKING QUALITY INTO ACCOUNT

26. Managers of the SIRENE system have consistently sought to provide a high-quality business register. Quality, as we have seen, is the ability to meet the requirements of users. The data managed, the functions provided and the files made available to users all contribute to the register's quality.

A. Background

27. For a long time the requirements of users, particularly statisticians, have been taken into account in the register through an approach actively based on communication. A "national quality mission" was set up, which led, for example, to the annual updating of the activity code, on the basis of annual business surveys (surveys carried out on businesses, collecting data on their structure, with the aim, inter alia, of complying with the Community regulation on structural statistics). Furthermore, thought was given to how to characterize those categories of businesses most likely to have ceased trading, given that listing businesses as still operational when they are in fact defunct was rightly considered to be one of the register's shortcomings.

28. Notwithstanding, the scant means available were not commensurate with the scale of the task. With hindsight, and after a more thorough analysis of a quality approach, it became clear that this intuitive, low-cost approach did not provide the high level of visibility needed for the introduction of a quality approach in which all would be involved (see section III above).

29. Other shortcomings had also been revealed: lack of responsiveness when an error was identified; impossibility of correcting certain errors in the information processing systems then in place. Regional experience had also shown that systematically checking units beyond a certain threshold was very time-consuming, while bringing relatively few errors to light. Conversely, analysis of tables enabled the register to be used for statistical purposes but rarely led to the identification of errors.

30. Lastly, there were - and still are - considerable limitations arising from the twofold role of SIRENE: as a national identification system used in the context of a system of inter-administrative regulations, and as a central tool for business statistics (sampling base, reference database).

B. The register's recent overhaul (SIRENE3)

31. SIRENE recently underwent a major overhaul, in the form of the SIRENE3 project, which was carried out in stages between spring 2003 and the end of 2006. The two main objectives were to improve the quality of the register and reduce the number of management tasks required.

Taking quality into account in day-to-day management

32. As already noted, the quality of a register does not depend simply on regular operations to improve quality. It is at the level of day-to-day management of the register that quality control is required.

33. Receiving standardized "administrative" information via telematics has enabled a new workstation to be developed that has powerful automated processing capabilities. Some data sent by administrative partners can thus be processed entirely automatically while other data, rejected by the system, need to be processed manually. This advanced level of automation guarantees uniform processing, which is in itself a factor in quality.

34. In addition, the introduction of checks into the data-processing flow improves quality; all information intended to be added to the register must be verified so as not to impair the quality of the information already stored in the register.

35. Finally, for the great majority of data, metadata now exist that make it possible to date information, provide an indicator of data quality and specify its source.

Taking quality into account as an "independent component"

36. Besides enhanced management, SIRENE has been refined to improve the way it meets business statistics requirements and place greater emphasis on the "quality" dimension.

37. Alongside the day-to-day management of the register (registrations, for example), INSEE performs operations unrelated to administrative requirements, either at the initiative of register managers or at the request of those in charge of statistical operations. These operations can involve, for example, updating variables for use in statistics or performing studies of "problem" data populations, and are called "quality operations" (see section V below). They are managed using a technique similar to that used for processing administrative data. Data obtained from "quality operations" are processed using the same workstation, which enables them to be processed in the same way as management operations. The workstation also allows priority to be assigned either to management operations or to quality operations. Annual work plans for staff in charge of managing the register take specific account of the quality-operations workload.

38. In addition, the intention was to foster a quality approach no longer based only on a "national quality mission" but also including a local dimension - that of each register management team. A "quality manager" was therefore designated for each INSEE regional team. The quality manager has access to specific work tools, including a management chart with which to monitor the processing of "quality operations" (volume measurement, scheduling, etc.), and its impact on the register, and with which to assess the results of the operations.

39. Depending on the progress of "quality operations" and administrative management (which of course remains the priority), the quality manager may propose new operations, the relevant decisions being made on the basis of the team's workload.

V. QUALITY OPERATIONS

40. While taking into account the constraints imposed by the inter-administrative nature of SIRENE and complying with the regulations by which it is governed, INSEE has developed a series of operations unrelated to administrative requirements.

41. Operations to improve the register, referred to as "quality operations", may take many forms (depending on their frequency and the populations concerned) but all have the same objective: to enhance the quality of SIRENE data by making up for the limitations of the inter-administrative system in terms of the requirements of user statisticians and those in charge of disseminating the register's information to the public.

A. Different ways of identifying specific units targeted

42. An operation to improve SIRENE may be carried out either at the request of one of the register's users (to clear up discrepancies or inconsistencies) or at the request of the register managers themselves (in a specific effort to improve the data available to users).

43. The way in which units are selected therefore varies from case to case:

(a) The user sends a file of units that he or she considers to contain errors, since the findings of a survey or other source differ from those of the register. The user waits for SIRENE to be checked and for an answer concerning each discrepancy to be provided.

Example: if a questionnaire is sent to a legal unit, using the address indicated in SIRENE, and the questionnaire is returned undelivered, the service carrying out the survey returns this information to the register, and asks for the address to be checked.

(b) A SIRENE manager will retrieve a list of units he or she considers to be potentially unreliable and will attempt to improve the database in accordance with certain criteria.

Example: the fact that names and dates of birth are contained in the register means that SIRENE teams can check that the same person is not registered twice.

B. Different types of operation

(a) Lists of units to be checked

Lists of units to be checked are drawn up in order to resolve a specific problem concerning one of the variables. Examples of this were given above.

(b) Register improvement surveys

"Quality operations" may also consist of actual surveys in the form of questionnaires sent to businesses in order to conduct an overall validation of their characteristics and those of their local units. This is, for example, one of the methods used to introduce into the register the new French general economic classification (NAF), Rev. 2, derived from the Statistical Classification of Economic Activities in the European Community (NACE), Rev. 2, which is in turn derived from the Standard International Trade Classification (SITC), Rev. 4.

(c) Mass updates

When the results of a survey provide information covering a wide range of the register's variables, they can be used to update the register. Naturally, not all the sources exploited by INSEE are used to update the register on an unsupervised basis. Currently, two sources are used for this type of update: one concerns the number of employees and the other the turnover. Thus, every year, all units that have taken on paid employees in the course of the year have the numbers of their employees in the register updated, using a single comprehensive source.

(d) Consistency with other registers

The SIRENE register must be consistent with the other administrative registers. An administrative body may therefore request a comparison to be carried out between those registers and SIRENE. These "consistency" operations enable the quality of the two registers compared to be enhanced.

C. Periodicity of operations and how they are carried out

44. Some operations are carried out on a periodical basis, others not.

45. The choice of whether to carry out operations on a periodical basis depends on user requirements and on the workload of SIRENE teams. *Example*: the activity code of legal units is updated on an ongoing basis, as and when the results of the annual business surveys are processed.

46. Other operations are performed on a one-off basis, such as those carried out to detect units incorrectly identified as still operational. In this operation, SIRENE identifies units that, although recorded in the register as operational, are highly likely to have ceased trading (no recent movements concerning these units).

47. Lastly, some operations start off on a “one-off” basis but are subsequently carried out periodically. In this way, a stock of units considered to be potentially unreliable is first checked. Then, once the stock has been processed, periodic checking prevents the formation of a new stock of potentially unreliable units.

D. A guaranteed response

48. One aspect of quality is that SIRENE undertakes to respond to the requests it receives. If the quality operation is carried out at the request of a user statistician, or partner involved in management of the register, this user will always receive a reply. Even if SIRENE chooses not to act on a request for a modification, the user will be informed of this and a reason provided. It should here be reiterated that quality is not about seeking maximum performance, but about respecting the specified - and thus expected - performance.

VI. CONTINUING THE QUALITY APPROACH

49. The new SIRENE workstation helps to improve the register’s quality by making the following possible: more consistent management; checks during management operations; availability of metadata; and monitoring of operations.

50. This workstation also places greater and more conspicuous emphasis on the “quality” dimension. Whereas quality was previously overlooked owing to lack of resources, or because it fell somewhat outside the “administrative” framework (although some processing operations do affect the quality of administrative data), it is now considered to be a fully fledged component in its own right. The role of the local quality managers designated two years ago was quickly adopted and appreciated by the register’s management teams.

51. Lastly, the flexibility afforded by the workstation’s design should enable further progress to be made: there are probably still data populations that need to be checked, studied and subjected to new “quality operations”.

Overview of the volume of operations carried out

52. The first “quality operations” in the SIRENE3 workstation were carried out at the end of 2004.

(a) Some 210,000 legal units had their activity code updated or validated (using data from the annual business surveys);

(b) About 15,000 units were checked at the request of statisticians (wrong address, probable termination of business);

(c) About 15,000 possible duplicate entries were examined;

(d) Some 45,000 units possibly misidentified as operational are currently being examined;

(e) At the end of 2006, 3.5 million units (legal units or local units) had had their employee statistics updated;

(f) By spring 2007, nearly 2 million businesses will have had their turnover statistics updated.

References

Re-engineering the French Business Register (2000-2006) by Marie-France Bobin and Michel Euriat (INSEE), France - 19th International Roundtable on Business Survey Frames, Cardiff, United Kingdom 17-21 October 2005 (http://forum.europa.eu.int/Public/irc/dsis/businessurvey/library?l=/2005_cardiff/developing_register/session_francedoc/_EN_1.0_&a=d).
