United Nations ECE/CES/GE.42/2011/5

Economic and Social Council

Distr.: General
6 July 2011

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Business Registers
Twelfth session
Paris, 14-15 September 2011
Item 6 of the provisional agenda
The role of business registers in industrializing the production of statistics

Initiatives for the industrialisation of statistics and their impact on business registers

Note by the secretariat

Summary

The document presents a short summary of international initiatives to industrialise the production of international statistics. The implications of this work include significant changes in the role of statistical business registers, with much more emphasis on their use as a tool for integrating data from different sources, and as a source for statistical production in their own right. At the same time, traditional roles such as providing sampling frames and survey populations are likely to decrease.

The views expressed in this document are fully consistent with the vision of the High-Level Group for Strategic Directions in Business Architecture in Statistics, endorsed by the Conference of European Statisticians at its 59th Plenary Session in June 2011 (ECE/CES/2011/1 and ECE/CES/2011/CRP.1).

I. Introduction

1. In different international forums, at all levels, there is increasing discussion about "streamlining" and "industrialising" statistical production. All statistical organisations are under growing pressure to produce more and faster outputs with fewer resources. The concept of "streamlining" is all about improving efficiency, reducing costs, and producing outputs more quickly and in a more flexible way. "Industrialisation" is the way that this can be achieved. It includes standardising processes, tools and methodologies, and recognising that all statistics can be produced in a similar way, moving away from the idea that each statistical domain is somehow "special".

2. In most organisations, statistical production has been organised for many years according to statistical domains, creating silos (or "stovepipes") of knowledge. This approach can be compared to the "cottage industry" system in pre-industrialised economies, where individual producers choose their inputs and tools separately, and operate as individual enterprises, creating hand-crafted products with a large labour input. Industrial revolutions in manufacturing industries have typically followed the pattern of introducing large-scale automation to produce standard outputs with much less human intervention. There is a growing view that the "industry" of official statistics is now ready for this sort of industrial revolution[1].

3. The types of changes currently envisaged for statistical production systems will have a big impact on the operation and use of statistical business registers. All types of statistical registers will increasingly be seen as tools for data integration, and as sources for data outputs in their own right. At the same time, their role in providing sampling frames and survey populations is likely to decrease, as the importance of the survey as an instrument to collect statistical data declines.

4. This paper reviews recent developments in terms of the industrialisation of statistics, outlining relevant international initiatives, then returns to consider the impact on statistical business registers in more detail.

II. International initiatives on the industrialisation of official statistics

5. In 2010 the Bureau of the Conference of European Statistics created a new High-Level Group for Strategic Developments in Business Architecture in Statistics (HLG-BAS). This group comprises the heads of five national and three international statistical organisations, and has as its mission to "oversee and guide discussions on developments in the business architecture of the statistical production process, including methodological and information technology aspects".

6. The first action of the HLG-BAS was to compile a global inventory of the various groups it is supposed to oversee and coordinate with. The result was the identification of over 25 different groups working on related topics, as illustrated in Figure 1. Each of these groups was asked to provide a short description of their work, including priorities and outputs[2]. This exercise showed that there is a considerable amount of work going on, mostly in expert groups, on different aspects of the industrialisation of statistical

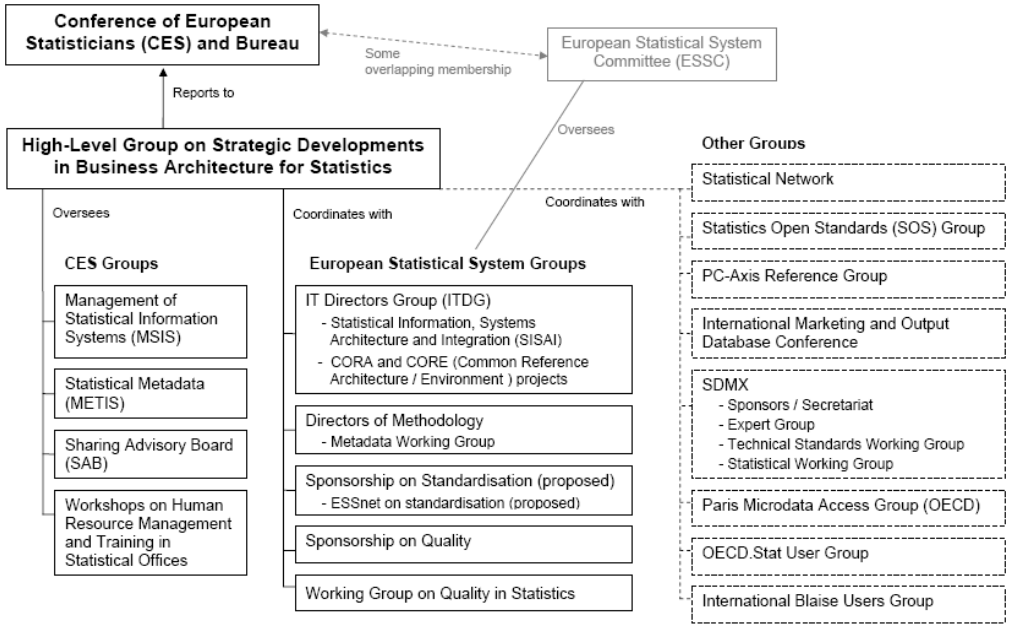
[1] See also the article "Are we becoming dinosaurs?" by Jean-Pierre Kent, in issue 4 of the Software Sharing Newsletter:
http://www1.unece.org/stat/platform/download/attachments/22478904/issue+4.pdf?version=1

[2] http://www1.unece.org/stat/platform/display/msis/Inventory+of+International+Groups

production. It also showed that the international coordination of such work is mostly informal or accidental.

Figure 1

The HLG-BAS and its relationships with other international groups



7. To improve the coordination of these groups, and to ensure that they can work towards a common goal, the HLG-BAS has prepared a strategic vision document, which was endorsed by the Conference of European Statisticians in June 2011[3]. The basic theme of this vision is that the world is changing. The amount of digital information in the world is growing exponentially, creating the potential for many new types of statistics. At the same time, statistical outputs are little changed in over 50 years, even if the production methods have sometimes been updated (e.g. from paper to web collection of data). The third factor is that data are increasingly being seen as a commodity. Most of the vast and rapidly growing pool of data is owned, controlled, or at least hosted by the private sector. Businesses are increasingly seeing the value of these data, particularly if they can be integrated with other data sets, and turned into useful information. Google has recently started producing real-time price indices, and other companies are also starting to produce outputs that are starting to bring into question the traditional monopoly enjoyed by producers of official statistics.

8. To respond to these challenges and to try to retain the relevance of national and international statistical organisations in the data industry of the future, the HLG-BAS vision focuses on the changes needed to both statistical products and processes. It concludes that industrialisation of official statistics is an essential strategic goal. This involves standardisation and the reduction of costs in the production process. Costs are defined as human labour, materials and duplication of efforts. It also involves using new sources of data, combining data sources in more flexible ways, and producing new types of outputs.

[3] http://www1.unece.org/stat/platform/display/hlgbas/Strategic+Vision

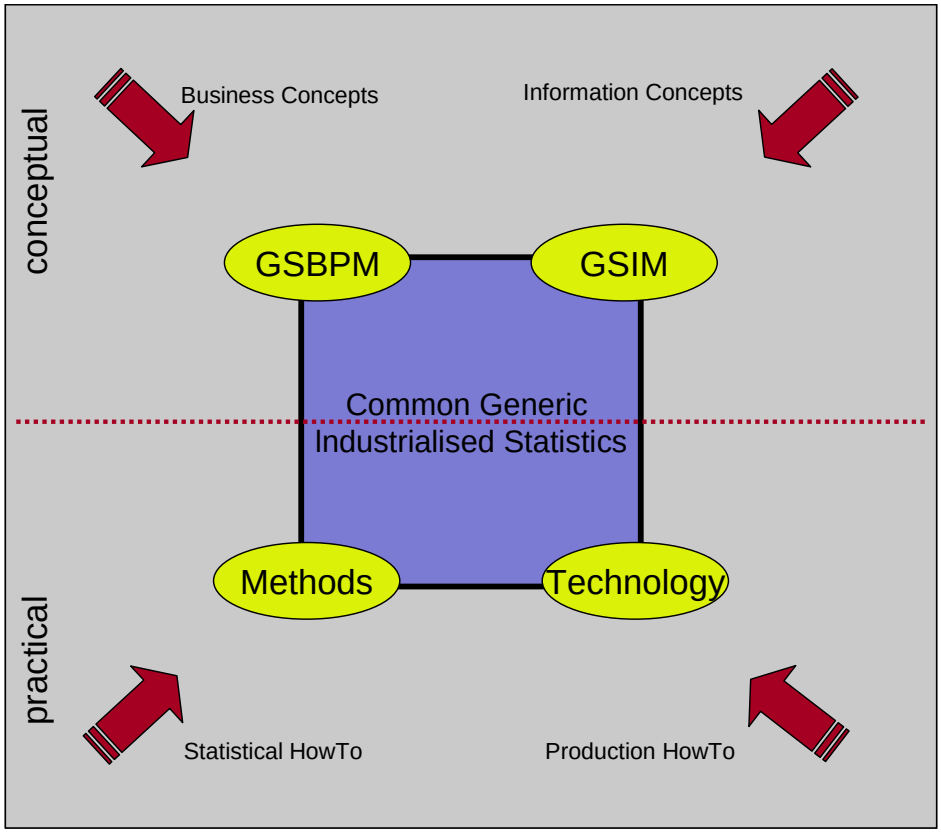
9. In Figure 2, the objective of industrialisation is symbolised by the smaller blue square. Within this square there are fewer but more standardised and cheaper ways of producing statistics. The big square symbolises the present statistical production universe with lots of different activities and a much larger cost. The basic aim is to converge to the industrialised part of the model, and reduce unnecessary diversity and duplication.

10. The blue square is divided in two parts, the conceptual part being described by architectures and models including common standards, and the implementation part comprising best practise and standardised methods and the technical implementations.

11. Industrialisation is equally applicable with regards to methodology and technology. This does not necessarily imply imposing a single solution. It implies adoption of common solutions that are optimised at the level of the organisation rather than for individual statistical outputs.

12. The new statistical process (the blue square) is to be seen as the area where the statistical production is compliant with four constraints. The Generic Statistical Business Process Model (GSBPM)[4], the Generic Statistical Information Model (GSIM)[5], standard methods and standard technology.

Figure 2
Industrialising statistics



[4] Developed by the Conference of European Statisticians Steering Group on Statistical Metadata – see www.unece.org/stats/gsbpm

[5] Currently being developed by the "Statistical Network" in cooperation with various other international groups – see http://www1.unece.org/stat/platform/display/metis/Generic+Statistical+Information+Model+(GSIM)

13. The task of industrialising statistical production is too large for most (if not all) statistical organisations to contemplate tackling it on their own. The HLG-BAS vision therefore strongly promotes the idea of international cooperation, thus adding an international dimension to the concept of industrialisation. Shared development of models, methods and tools will help to reduce costs even further. For a relatively small investment, each organisation will have access to a much larger pool of shared resources. The "Statistical Network"[6], a collaboration initiative between six national statistical organisations, aims to demonstrate how shared developments could be mutually beneficial.

14. The next step for the HLG-BAS is to develop a strategy for the implementation of this vision. To start this process, a conference is scheduled for late October 2011, to bring together representatives of the groups identified in Figure 1, identify common goals, and try to ensure that each group is contributing effectively towards implementing the vision.

III. What does this mean for statistical business registers?

15. The industrialisation of official statistics has significant consequences for many aspects of statistical production, including the role of statistical business registers. Traditionally, the main role of these registers has been to provide sampling frames from which statistical surveys can be conducted, and populations for which results can be calculated. Whilst industrialisation does not imply an end to statistical surveys, it does imply a much reduced role for the direct collection of data from respondents in the statistical production process. To some extent this trend has been present for many years, reflected in the increasing use of administrative sources, and the development of register-based statistical systems in some countries.

16. The HLG-BAS vision of industrialisation takes this trend several steps further, however, in that it encourages the exploration of the mostly unknown and potentially very rich domains of private sector data sources. It also places more emphasis on new types of output, looking more at what information it is possible to extract from the available data, rather than being constrained to replicating existing outputs.

17. This means that there will be a growing need to integrate data from multiple and increasingly diverse sources. Where these data relate to businesses, the natural integration tool is the statistical business register, where there is often a long tradition of integrating data from various administrative and survey sources, combined with good knowledge of matching and data linkage techniques.

18. Under this scenario, the statistical business register becomes the gateway for business data entering the statistical organisation. It is the place where different types of statistical and non-statistical units are reconciled allowing their characteristics and attributes to be combined into statistical data.

19. Of course, common identifiers and standardised unit definitions would considerably facilitate this process, but in most countries the widespread use of such identifiers and definitions, particularly by the private sector, is still more of a dream than a reality.

20. The need to optimise statistical production at the level of the organisation, and develop capacity to produce new types of statistics, also implies greater integration between statistical domains. This, in turn, implies more integration between different types of statistical registers. There will be growing demands for links between business, population,

[6] http://www1.unece.org/stat/platform/display/msis/Statistical+Network

real estate, agricultural and many other types of registers. Ultimately this could lead to a blurring of the traditional distinctions of economic, social and other statistical domains.

21. The concept of satellite (or associate) registers is likely to become more important, as new sources may often not be fully comprehensive in their coverage. Satellite registers contain a specific sub-set of units, but often hold extra information about these units than is available (or even feasible) in the main statistical business register. A traditional example is a satellite register of hotels, holding additional information about category and number of beds. Possible new examples could include e-commerce or on-line auction businesses, with information about their sales and markets.

22. Ultimately this could lead to the re-definition of statistical business registers, as their role evolves into something more like a general business statistics database, recording a much wider range of units and variables than today. These new "super registers" would then be the main source of national economic data.

23. Recent trends towards adapting statistical business registers to better reflect the global economy also support the industrialisation of statistics. From the introduction of enterprise groups in national business registers, to the creation of multi-national registers such as the European Statistical System's EuroGroups Register, these initiatives promote the ideas of collaboration between organisations, and in many cases, the use of private-sector data sources. Assuming the trend of increasing economic globalisation continues, the importance of trying to measure it will increase. The use of new types of data sources, and automated data exchange can only facilitate this task.

IV. Conclusion

24. The world we live in is changing rapidly. So far official statistics has generally failed to keep pace with these changes. This means that unless a considerable effort is made to modernise the processes and products of the official statistics community, this community risks to lose its relevance. Potential competitors from the private sector are starting to emerge. Whilst official statistics can currently rely on arguments of data quality, the threat of competition is only going to increase, particularly if competitors offer considerable savings in cost and improvements in timeliness. The "Google generation" are more concerned about finding information quickly, rather than waiting months (or even years) for "official" data. Timeliness is becoming an increasingly important dimension of quality.

25. Statistical business registers, in common with official statistics as a whole, have changed relatively little over the last 20-30 years. They need to adapt to new data sources and new statistical outputs, in order to support the industrialisation of official statistics. There are huge challenges ahead for those working with statistical registers, but also huge opportunities to be at the forefront of this industrial revolution.