

Distr.  
GENERAL

CES/AC.71/2005/20 (Summary)  
31 January 2005

Original: ENGLISH

**UNITED NATIONS STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE (ECE)  
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION  
STATISTICAL OFFICE OF THE  
EUROPEAN COMMUNITIES (EUROSTAT)**

**ORGANISATION FOR ECONOMIC  
COOPERATION AND DEVELOPMENT (OECD)  
STATISTICS DIRECTORATE**

**Joint ECE/Eurostat/OECD Meeting on the Management of Statistical Information Systems (MSIS)**  
(Bratislava, Slovakia, 18-20 April 2005)

Topic (iii): XML and web services

## **AUTOMATED ACCESS TO 100,000,000 STATISTICAL FACTS VIA STATLINE4 WEB SERVICES**

### **Invited Paper**

Submitted by Statistics Netherlands<sup>1</sup>

### **Summary**

1. The statistical output database of Statistics Netherlands, *StatLine*, evolved, in about 10 years, from a simple web-based application for browsing statistical tables, into a flexible, open and highly customizable multidimensional database system, which offers end-users the facility to interactively browse and view its contents via the Internet. Currently, StatLine contains *all* the statistical results published by Statistics Netherlands, which is a total of approximately 100 million statistical facts. Since everyday new facts are added to StatLine this number continues to grow.
2. Due to the huge amount of rather heterogeneous information in this database, finding a specific statistical fact is not always straightforward. Although an intelligent search engine assists users by pre-selecting statistical tables and although the website of Statistics Netherlands directly points to the most important key figures in StatLine, end-users often need a lot of knowledge of the thematic content structure to find more specific information. Therefore, better means for accessing StatLine are still needed.
3. The next version of StatLine, StatLine4, will offer a simpler and therefore more intuitive user interaction mechanism. But this is only one part of improving StatLine. Many institutes use a very specific part of the information in StatLine more than once. Clearly, they are most interested in those parts of the database that conform to their specific business needs. This poses the question of how a statistical database could better serve this kind of frequent and partly automated access.

---

<sup>1</sup> Prepared by Olav ten Bosch ([obos@cbs.nl](mailto:obos@cbs.nl)) and Edwin de Jonge ([ejne@cbs.nl](mailto:ejne@cbs.nl))

4. The ideal way to provide this functionality is to extend the statistical database with a set of well-defined XML web-interfaces, which offer external systems automated access to its content via the Internet. This set of web services offers third parties a way to be informed of updates with respect to a certain statistical theme, of updates of specific statistical figures and it offers a way to automatically retrieve the added or changed content. Customers may use these services to check automatically certain statistical indicators at regular intervals. In addition the web services offer third parties a way to search automatically the content of the statistical database at regular intervals and interpret the results in a way they want. One of the clients of this feature is the main website of Statistics Netherlands itself.

5. In this paper we briefly describe the design and functionality of this set of web services to be used by national statistical institutes, other institutes, or even end-users, both from a technical perspective as well as from the perspective of a consumer party. In addition, we compare this design with related approaches, such as the approach to be taken by the International SODI (SDMX Open Data Interchange) working group<sup>2</sup> and we highlight the main differences. We briefly describe the approach SODI (SDMX Open Data Initiative) Finally, we describe our ideas how statistical dissemination might be extended in future to provide even tighter integration between various statistical institutes and we relate these ideas to other base functionality such as GIS technology and semantic web technology.

-----

---

<sup>2</sup> Statistics Netherlands is one of the participating members of the EuroStat SODI task force.