Distr.
GENERAL

ECE/CES/SEM.54/21 (Summary)
10 April 2006

Original: ENGLISH
ENGLISH AND RUSSIAN ONLY

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE CONFERENCE OF EUROPEAN STATISTICIANS

EUROPEAN COMMISSION STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT (OECD) STATISTICS DIRECTORATE

Joint UNECE/Eurostat/OECD Seminar on the Management of Statistical Information Systems (MSIS)
Sofia, Bulgaria, 21-23 June 2006

Topic (iii): Dissemination and client relations

TOWARDS A SERVICE-ORIENTATED ARCHITECTURE (SOA) FOR THE STATISTICAL VALUE CHAIN (SVC)

Supporting Paper prepared by Vic Duoba, Office for National Statistics, United Kingdom

Summary

Background

1. The paper presents a technical perspective on the attempt by the UK Office for National Statistics (ONS) to modernise its data and information processing systems. The original technical objectives were very demanding, and included formulating metadata-driven processing and componentised applications accessing a single highly flexible data store for all data from unit to publication level. What transpired was a rapid learning process and a refinement of modernisation objectives to fit a more feasible and affordable plan.

Statistical Value Chain

2. The business processes of a National Statistics Institute (NSI) may be categorised as in the following table:

1 Stakeholder consultation and prioritising
2 Methodological studies
3 Administrative sources assessment
4 Sampling design (technical)
5 Sampling design (implementation)
6 Data collection (field, electronic, mail)
7 Editing, validation, coding, derivations
8 Weighting and estimation
9 Simple analyses and quality assessments
10 Index number construction
11 Time-series analyses
12 Other advanced analyses
13 Confidentiality assurance
14 Information and data dissemination
15 Data management

3. The activities tabled are in some, but not all, cases capable of being considered as fundamental services consumed by a NSI and hence the challenge arises of componentising these services. It is tempting to visualise a future scenario where NSIs (or International Statistical Institutions - ISIs) might perpetually modernise their systems by acquiring pluggable modules from other NSIs, or defray their own system development costs by selling components. Cost-sharing via joint development is also a possibility and is probably the preferred international arrangement. The pressures placed by market economies on increasing efficiency within public services should also be placing pressure on NSIs and ISIs to develop common software for common, repetitive tasks.

4. There are also purely intra-NSI drivers for componentised delivery of statistical functions and non-functional requirements. Many NSIs face the twin problems of modernising systems and maintaining their huge investments in legacy code. The legacy issue cannot be defeated simply by rushing to re-write code, as this is potentially wasteful of an investment, might not be possible in the short term because of the opaqueness of the source code and documentation or shortage of highly skilled designers, but in some cases the legacy code might be of such high quality in terms of versatility and performance that a re-write should be avoided by choice. Modernisation is also driven by the internet age fuelling a demand for information and for providing different ways of conducting business commercially, between government departments, and between citizens/businesses and government. This has caused systems scalability and security challenges for existing technology, as well as requiring novel and highly distributed applications to support new business models

5. Hence, the commonality of the SVC, pressures from governments for efficiency improvements in public services combined with the worldwide rapid adoption of electronic technologies are suggesting to and pressuring NSIs and ISIs to adapt and respond rapidly. Recent technological advances in large-scale applications development and deployment suggest that component-based development of Service Oriented Architecture (SOA) be seriously considered as a way of meeting the challenge of efficient, scalable, flexible and high-quality delivery of statistical services.

Definition of Services (as in SOA)

6. The following definitions have been obtained at random from "The Web", but they are adequate to give the flavour of the idea of a service.

A Service is an implementation of a well-defined business functionality that operates independent of the state of any other Service defined within the system. Services have a well-defined set of interfaces and operate through a pre-defined contract between the client of the Service and the Service itself.
[Samudra Gupta]

SOA is the latest name for an applications architecture for sharing and reusing code. With today's SOA, applications are built (or retrofitted) with standard interfaces most often based on XML and its derivatives: Simple Object Access Protocol (SOAP) and Web Services Description Language. SOA also defines how those services are located, executed, managed, monitored and secured.
But before the technology is relevant, companies have to choose the business processes they want to convert into services.
SOAs are loosely coupled in that services can be connected on an ad hoc basis, and that services and clients can be changed independently of one another. The services have standards-based interfaces and are coarse-grained in that they provide a distinct service or business process, ...
[John Fontana, Network World, 09/29/03]

ONS's Experience

7. ONS is currently seeking to use Oracle, Java and SAS to implement a modernised information management system that will be secure, scalable and flexible. Our intention is to use existing best-practice methodologies (e.g. GES, Tau Argus, X12ARIMA, SuperCROSS) as already incorporated within software from other NSIs, ISIs and private software companies. It is now the aim to move to an SOA in 3 stages:

(i) Establish a core of data models using a number of important surveys, basic metadata and process output types as design leaders, and implement the designs in Oracle technology
(ii) Introduce workflow and business activity monitoring
(iii) Introduce fully configurable services and use BPEL to orchestrate the services.

At the same time, the systems design and programming standards used would comply with SOA requirements (within the limitations of business budget constraints) and would be designed from the onset to be orchestrated on a common "applications stack". The tension between maintaining an SVC/SOA vision and meeting budget constraints is a difficult balancing act that has to be managed by the systems development managers in consultation with business representatives who understand the longer term cost implications of the software lifecycle.

8. The prognosis for an SOA implementation, which results in services that are shareable between NSIs, is still an open question. There is a crucial dependency on a critical mass of institutions agreeing on a common definition of a SVC and on successfully defining a practical

subset of services to implement. No single NSI can design services for the benefit of the wider community of NSIs if there is no sharing of the associated cost, as budgets are always tight and intensive modernisation is financially demanding and technically difficult. Only by designing an SOA implementation of an SVC jointly with statistical partners does it seem likely that services will be truly interchangeable between separate SOAs.
