

Distr.
GENERAL

CES/AC.71/2003/18
22 January 2003

ENGLISH ONLY

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
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
(Geneva, 17-19 February 2003)

Topic III: Efficient management of increasing technical complexity

A NEW DAWN FOR INFORMATION TECHNOLOGISTS

Supporting paper

Submitted by the Office for National Statistics, United Kingdom¹

I. INTRODUCTION

1. The demands made by stakeholders within government and other key national institutions for: more timely and accurate data; more openness, transparency, and accountability; delivery of more statistical information using new 'internet' based technologies. These are the drivers for IS/IT professionals to create and sustain complex technological environments.
2. The mass of information being created on a daily basis, the historical records, data from previous periods and the new technologies that will provide new media sources, provide a challenge for information managers and information users alike.
3. The use of powerful tools that enable the professional (and the amateur) to examine masses of data demands that National Statistical Institutes (NSIs) provide a significant amount of supporting data (or metadata) to ensure correct and consistent understanding of the information gathered. These services are no longer a production output managed and controlled by publishers. Instead much more statistical information is to be provided using new technologies, embracing the 'Internet' but looking beyond even Digital TV. Each new technology platform adds to, rather than replaces, a previous service so requiring the IS/IT professional to create and sustain complex integrated technological environments.
4. The nature of the IT services within statistical organisations will change greatly from that of a simple service provider to key partner in the statistical process.

¹ Prepared by Dayantha Joshua (dayantha.joshua@ons.gov.uk).

II. THE LEGACY OF STATISTICAL DATA PROCESSING

5. The statistical processes of the last 20-30 years were largely product-focused business processes, where the data was isolated from other areas of study when in its primary source state. Often systems would be developed according to departmental structures, which over time would merge and split thereby leaving a statistical organisation to nurture a range of technology products, each harbouring diverse statistical methodologies and system approaches.

6. The hardware and software that prevailed in the gestation of these systems would limit the design, number of variables and sometimes dictate the processing by having a limited range of statistical capability. Further to this, NSIs have a positive tradition of collaboration rather than competing and this has resulted in shared practice for both methodology and technology development. However, too often this has led to introverted solutions and missed 'technology-led' opportunities emerging from other industries.

7. Where there have been successes (resulting in commercial or near-commercial products that include SPSS, Superstar and Blaise) the solution has created additional constraints that has made data linking and even simple exchange between systems an unlikely proposition, certainly not one without quality risks.

A. Data Processing Practice

8. The IT professional provided discrete services and these service chains were dominated by the processing of paper based information. Even when using administrative data, all incoming data was subjected to extensive data validation and cleaning that led to significant time delays between acquisition and result publication. The processing computer modules were based upon business system design, whereas the analytical work, completed by statisticians, would be conducted using independent data tools working with data extractions.

9. The approach to systems development had tended towards art rather than classical control centres and ironically favoured the data analyst who was supported by emerging software be it the humble (but free form) spreadsheet or the more powerful desktop tools such as SPSS, Quicktab and the like.

10. Each generation of production tool (including those generated by collaborative working) created new barriers to more effective use of data, data linking and essential data management practice such as common naming and data preservation procedures.

B. Organisation Silos

11. Other Government Departments (OGDs) play an important role as data providers, traditionally in areas such as health statistics, VAT and PAYE data. Providing this data will not be a core function of the OGD and there are issues of standards, quality and timeliness which needs to be managed as part of the statistical process. The complexity of business processes led to the creation of cradle-to-grave management silos. Typically these silos would compete against one another for finance or IT resources. The need to deliver specific outcomes, at specific dates within the limited corporate service resource re-enforced localised management structures and localised solutions.

C. The loss of control through technology

12. A notable impact upon information management quality was the introduction of the PC in the mid-to-late 1980's. The mainframe computing environment which had provided some consistency of service and data standards (e.g. employing data dictionaries) formed a common bond, if not one that was capable of providing useful cross service analysis. The PC gave freedom to analysts to work creatively but destroyed the central cohesion. The PC work was managed by data analysts rather than information professionals and the products and overall environment did little to encourage documentation, data archiving practice or common data frameworks.

III. THE PRESENT DAY ISSUES & DRIVERS

13. The work of government statisticians is to monitor long-term trends in key policy areas and so specific data series are preserved for this purpose. But in the last five years much has been made of the information society and the contribution that information can make to the knowledge economy, creating an even more demanding agenda. More is expected of the Office for National Statistics (ONS) from central government by way of providing measures of and evidence for economic activity and regional social analysis. The Neighbourhood Statistics Service, accessible through the ONS web site, is a prime example of the changing deliverable of NSIs.

14. The new technology capability and the associated service expectation are forcing data owners to reappraise the use made of data, the controls needed and wider policy and privacy implications. Notable in this respect is the current draft EU Directive for the re-use and commercial exploitation of public sector documents/information; it seeks to harmonise EU data access rules and to encourage a potentially lucrative commercial information market.

A. e-Government in the UK

15. Like most leading countries the UK has embarked upon a modernisation programme that has centred upon a more effective public service through greater use of information. The goal of more efficient public services, more citizen responsive services and improved public sector knowledge has required improved Information Management policy and a realisation that IT services will need to share a common exchange language.

Electronic Government Interoperability Framework (E-GIF)

16. The Electronic Government Interoperability Framework has had an important impact on the next generation of government services and influenced the IT industry offering. E-GIF enables seamless data exchange (through mandatory use of XML though other standards will be included when mature) and has created awareness for common usage of data names and constructs through the development of a Government Standard Data Catalogue. Departments are required to build all IT services so that data exchange complies with e-GIF policy (www.govtalk.gov.uk)

17. This is refreshing for statistical and research organisations as it encourages a wider conformity to standards within administrative processes and overall there should be a higher quality data store within the public sector.

Role of Statistics in e-Government

18. The policy driven outputs associated with government statistics will remain but the information is becoming of greater social significance to a wider audience. The statistical services are therefore providing a more complete and flexible service to schools, businesses and the general public at large. These new interest groups will demand different levels of service (e.g. data streaming, analytical tools at point of collection) and a range of delivery channels, be it through paper or electronic publication.

B. Data Management

19. Data demand has been discussed and the provision of open and transparent data services introduces a number of data management issues including - Privacy, Revision control, Data Archiving, Access.

20. Two aspects will be briefly considered in this paper as they introduce some of the issues that underpin the growing complexity of technology management and the emerging role of the technology department.

Data Privacy

21. Legislation in addition to professional standards and ethics is providing an overarching framework. However, the plethora of information sources and the range of dissemination channels make privacy control complex. The ONS has developed a Code of Practice and is generating supporting protocols to guide officials and provide the necessary assurances to the public.

22. The capability of new technologies to integrate data items will have benefits (e.g. commercial and hence the EU 're-use' draft directive) but the services must be managed so as to be limited to those that have the authority or powers of access. Electronic access to data stores through each channel (e.g. Web interrogation, research tabulations, web publishing) will require the satisfaction of an authentication process. Therefore, systems now will necessarily share common data management protocols and will share security procedures. For this to be achievable there will need to be extensive investment in improved metadatabases to support not only current data sets but also historical sets. Specifically, this investment is needed to protect against unintentional disclosure through data linking of sensitive private or personal information.

Data Access

23. Data accessibility will be a critical design factor for all new data systems. Access will require:

- ◆ Physical access
- ◆ Data navigation
- ◆ Data aggregation

24. Physical access options will vary depending upon the class of user but overall the systems must be capable of servicing an increasing number of clients. Physical integration will be complex if data resides in multifarious system solutions or where, because of its age, a system may be too cumbersome to be of use in an online service.

25. Data must be properly defined and stored to ensure that data recovery tools (e.g. search engines, middleware, or purpose built routines) can locate the appropriate data and relay its underpinning context.

26. Bill Gates is reported as saying that, web technology has less to do with changing IT architectures than filling gaps 'where there was no architecture'. This is perhaps accurately reporting what is happening in large organisations - finding a path between information sources.

IV. FUTURE MODEL

27. The role of the IS/IT provider in the future must be for an even closer involvement in the design of the business process and providing expertise in all aspects of data management. The technology knowledge of the IS/IT provider will also extend beyond the current networking technologies. Accepting the issues of security and authentication - statistical organisations will be introducing web-based, Interactive Digital TV and mobile telephony platforms.

28. The traditional model of data processing compared with the new demand is analogous with the train and motor car as transport services. The old model, like the train, is very limited and primarily participated in fixed paths of action. The designs remained static over long periods.

29. The once lengthy process of survey initiation, collection of paper responses, data validation and preparation is now likely to shorten, similar to those used in face-to-face social survey work. In social surveys there is a greater use made of tools such as Blaise, which provide at point of collection - data capture; validation; and metadata substructures.

30. The new technologies, like the motor car, are more dynamic and have greater flexibility. However, this flexibility introduces new design issues and will include:

- ◆ The flexibility leads to devolved powers for end users of the data that must be actively tracked, managed, and reported.
- ◆ Explicit specification of procedures to emulate the opportunities that the longer processing chain provided to the statistics manager to use formal and informal inspection and control points.

31. In essence the IS/IT community will need to be conversant with both the technology and the purposeful activity for which the data is being used.

A. Key transformation actions for IS/IT providers

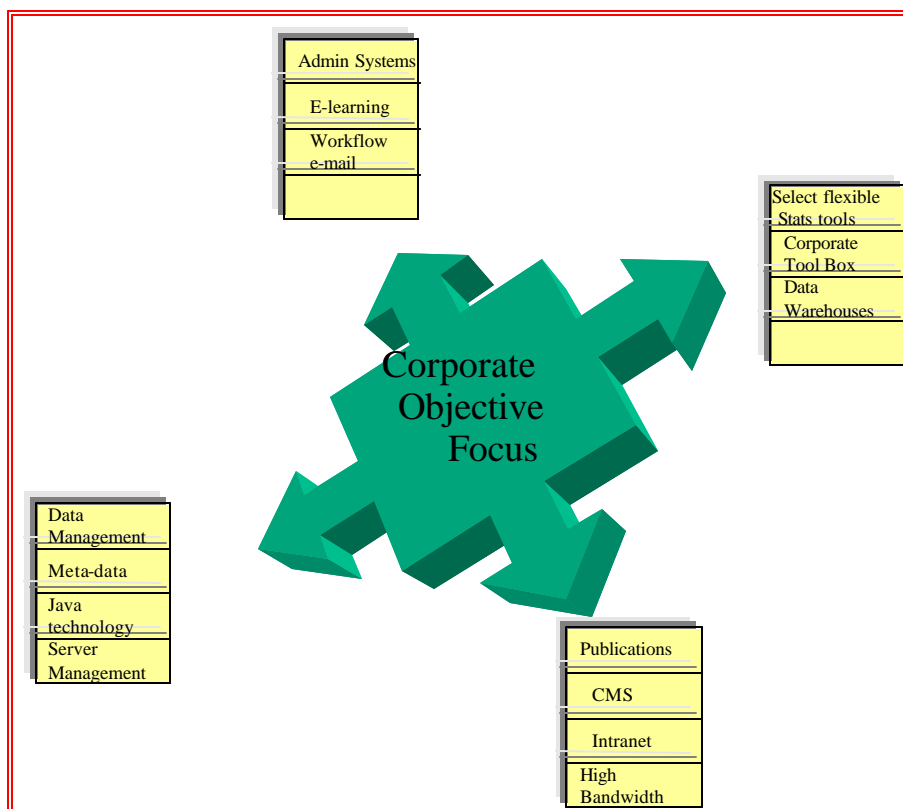
- ◆ Standardising & reducing the complexity of the technical architecture that underpins the business of the organisation.
- ◆ Developing an integrated design for all data holdings
- ◆ Revising the individual skill capability
- ◆ Redressing the service offering by the internal organisation with external opportunities
- ◆ Operating a management architecture focused upon outcomes
- ◆ Taking leadership for data standards and compliance with broader government regulation
- ◆ Balancing the open culture of the internet with the closed ethical code that NSIs are obliged to follow.

V. THE ONS IS/IT DIRECTION

32. The ONS has a rigorous transformation programme and the role of information technology is seen as central to this process. The technology side (IM Group) will therefore need to provide leadership in certain areas and instil best practice from both the commercial and public sectors.

33. The impact of a changed management structure in a newly formed IM Group (IMG) has enabled an improved service focus that has helped to identify the critical priority and to reduce attention given to low-value changes.

34. The diagram below provides a simple model that highlights the strategic directions and the issues to be addressed. This paper breaks out some of the specific actions that have arisen as a result of this focus.



A. Better Technology Management

35. The range of technology has been identified as an issue for emerging organisations. Devolved budgets have enabled silo operations to purchase without consideration for the corporate benefit. ONS has now determined a simpler and product harmonised Goal Technical Architecture that benefits from a centralised IT goods and service procurement function.

36. The coming three years will see the introduction of a single database product that has established and proven add-on modules providing a range of corporate services (financial services, HRM, property management). Through the development of a corporate data-spine it will be much easier to establish a corporate brand and overall confidence in the IT strategy through a broad product familiarity.

37. The reduction in number of products deployed for all ONS operations will lead to a tighter integration and flexibility in what can be delivered. Overall, the use of common software will ensure that the infrastructure is not burdened by inefficient activity and to the contrary the application can be tuned for better utilisation of the hardware platform. The design of back-up, archiving and server deployment can now be based upon optimum performance.

B. Better Statistical Process

38. The government targets for the UK being online are steadily being achieved. The EU has also set targets for technology take-up by the general public, including access to broadband technology. These targets present an opportunity for the redesign of statistical processes that go beyond 'filling the gaps'.

39. Data collection will be made possible through different technological and intermediary channels (assuming the appropriate authentication is in place). Supporting this aspect the Cabinet Office has led the

development of a central authentication gateway known as the Government Gateway. This capability means that the ONS can use a developing common identification service, currently employed for online tax returns or passport applications, to securely transact (because some surveys require the issue of past information) within data collection processes. The tools used for collection can, like CATI, provide point-of-collection functionality and hence deliver results cleaned and validated direct to an output source.

40. ONS is close to completing an extensive review of statistical methods (Statistical Infrastructure) and has identified a target service and the associated IT tools needed. The re-design of business processes will be based around the exploitation of the methodology/tool relationship.

41. ONS has recognised the changing relationship that will take place between IS/IT provider and business activity. Training is crucial to the success of the transition (including training in Project and Programme management) to enable the assembly of multidiscipline design and development teams.

C. Better Data Handling

42. Metadata is being harmonised across the organisation and is a key area for the ONS IM Group. It provides the basis for access and security design at one level and ensures that content is correctly applied. Metadata is core to researchers and others in this field so ONS has a multidiscipline team preparing the full range of policy and practice required in all areas of the statistical process chain.

43. The importance of secure data storage (highlighted by September 11th) and the need for improved archiving (particularly for registration data following high profile legal cases) has led an IMG review of network server and storage configuration. ONS will have low redundancy and high flexibility for hosting and switching infrastructure resources.

44. Built within the emerging network configuration will be security levels dealing with:

- Virus and hacker attack
- Physical Security
- Account control features

45. The first two will be familiar to most and will not be developed further in this paper. The third level of security reflects the developing complexity of data management and the need for flexibility within configurations to control:

- Controlled access to data within the physical premises
- Controlled access by authorised personnel working outside the physical premises
- Controlled access to authorised organisations outside the physical and organisational structure of ONS.

46. Collectively the above control environments are a necessary challenge for the new environment in which an NSI will be working. Growing anticipation of extensive data access for statistical research will mean that an NSI will be custodians of more data with an increased capability to cross-tabulate and analyse.

47. At each access point it must be possible to determine the authorisation status of the enquirer and match this to the status assigned to unique data items and the collective data holding in a session.

48. The solution will lie in a combination of devices found within firewall technology, metadata sampling and network account structures. The ONS IMG would be happy to share the developments and outcomes with this forum.

D. Communication and Publication Tools

49. The ONS is moving towards a new organisational structure that focuses expertise into functional clusters. Previously each analytical area controlled the full life cycle of collection through publication. While

there will be small pockets of such activity greater stress is being placed on a single point of collection, analysis and reporting/outcome production.

50. The production cycle will have a primary focus of Internet centred information packages that can be used in various Web delivered incarnations. Furthermore, ONS is migrating its internal information so as to be browser accessible from an Intranet.

51. To facilitate the new publishing and collaborative product development the ONS has invested in a Content Management System (CMS). It is recognised that CMS software had a roller-coaster ride but is now more stable and usable as a corporate product. The CMS provides:

- Responsiveness to changing customer demands
- A management environment for handling diverse data and content streams (ONS is the hub for GSS/NSS production)
- A structures filing environment that can house data, video, audio and other media used in survey and analytical work
- A centre of excellence built around a common corporate 'publishing' tool
- A less polarised and so more integrated business/publication workflow
- More emphasis on content tagging, storage and archiving

52. There will be increased demand for communication links and storage to service the new production model. Providing communication contingency is expensive and the overhead of the supporting an infrastructure (firewall technology, communication lines, backup services) to maintain separate public and working domains is also expensive if it handled by the host department. Recognising the importance of continuity ONS has taken the decision to use a managed external web service provider. The CMS provides a workflow discipline that is necessary in delivering public services via a hosted service.

E. Formal Service Management

53. A modern IS/IT service will need to integrate a number of services and service providers. It is important to state clearly the service structure and how the relationships between the main provider and third parties are managed.

54. IMG is developing two levels of Service Level Agreement (SLA) with ONS. The top level provides a service aegis and is an agreement between the ONS Executive Board and the IM Group. At the next level agreements can be developed specific to a unique project or service area. At both layers there will be defined service metrics to allow both sides of the agreement to monitor performance.

55. The SLA, with the metrics, are a key component to capacity planning and ensure that a review of services by the IM Group, focus upon business capability and associated risk. To this end the SLA contains a chapter that provides volumetrics and operational process details concerning the business area.

F. IM People and Group Development

56. It is essential that there is a trust between IMG and the rest of ONS so that changes within the IM group are supported. The trust is built on two supports: leadership by example; and communicating a clear vision.

57. IMG has reformed itself as a business operation. The Director is supported by a management team that has devolved responsibility for areas with clear lines of demarcation. Peer review and accountability in terms of business justification dominate performance management in each area. Regular frequent reporting via service metrics forces internal inspection and realignment of activities. Tighter management of expenditure, project approval and ONS communication is allowing IMG to set clear agendas on behalf of ONS.

58. As a result of this reformation IMG has been able to issue strong messages regarding the technological direction and how the business areas will benefit. These messages are communicated through pathfinder projects (when the solution needs to be demonstrated) or through simple presentations showing technology road-maps with associated technology and business milestones. The use of tools to create 'demonstrators' has enabled IMG to communicate, visibly, some of the capabilities that can be delivered via technology.

59. It has been necessary to target the IMG resource with changes as vigorous as the technological changes. The skills and cultural tone of an IM organisation will change radically. IMG has conducted a skills audit and are preparing transitional training that will see a significant skills change that embrace new design techniques, new mobile technologies and rapid application development methods.

60. It is understood that an investment in training requires a positive retention policy and this is underdevelopment. It is worth noting that this will not only depend upon pay and reward schemes. Training and personal aspirations will be part of the overall package. One key aspect will be to increase secondment opportunities across the range, that is, within IMG, within ONS and outside ONS, both the public and private sector.

VI. CONCLUSION

61. The IM Group of ONS is clear about the complexity of the new environment and the increasing scope of its work. It is preparing for a different relationship with the process owners that will deliver secure but flexible services to whomsoever is authorised to access ONS data services.

62. The first stage of preparation has been to improve the knowledge that the IM Group has about itself: its people, the services it delivers and the tools it has to complete its work. Continued good management is dependent upon a commitment by both the IMG personnel and the ONS generally to a simplified and targeted technology strategy. This requires a trust in IMG, not only that it can deliver technology but that it understands the emerging business. Attention is therefore being given to closer working relationships with ONS personnel at all levels. However, there is a payback in working closely with those that have a much broader influence on public services - mainly the Office of the e-Envoy based within the UK Cabinet Office.

63. During the transition IMG will need to manage legacy and new systems in parallel and so there shall be an increased burden brought on by the systems in operation and the need to displace staff for skills development. IMG will create strong partnerships with external organisations that will compliment the work commitment and skill capability at a given time.

64. Finally, it should be noted that both formal targets and human nature dictate that marked changes must be seen within a short-timetable. Planning and funding are clearly focused on a three-year target with considerable advances expected by mid-2004.

- - - - -