



**Economic and Social
Council**

Distr.
GENERAL

CES/2001/5
26 March 2001

ORIGINAL: ENGLISH

**STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

Forty-ninth plenary session
(Geneva, 11-13 June 2001)

Some thoughts about the Internet and Official Statistics.

Paper prepared by the Office for National Statistics, United Kingdom¹

Introduction

1. Statistical data is very important to public policy, and to public understanding of the context of public policy. Its accessibility contributes to the effectiveness of and trust in government, and the efficiency and stability of financial, commodity, labour and other markets. The effective delivery of official statistics on the internet is not only driven by the potential efficiencies for continuing to deliver existing services to existing users, but the growing expectation of access at a low cost by citizens to government services, higher levels of numeracy and a sea change in the ways in which statistics are able to be made interesting and relevant. Our future service model may well come from the media, compared the present more academic-like approach we take.

2. The internet is a fast growing, highly integrated global facility, and the competitive position of the statistical office internet service, its comparative qualities, and the nature of relevant innovations, are highly visible through being provided alongside other public services, and alongside those of all other information providers, in each and any of the situations users are in when they draw on our services

¹ Paper prepared by Mr. Len Cook, National Statistician, ONS, UK

The special value of the internet to official statistics

3. The internet is an architecturally independent pathway that enables myriad's of linkages through public and private communications networks between a large minority of private households, many businesses, most public bodies, many community organisations and nearly all of the media, research and international organisations. The internet enables a gigantic leap in the numbers of people who can access official statistics at a very low cost, regardless of region, physical mobility or language.
4. The internet has particular value to official statistics, because of the capacity to present statistical information dynamically, in visual form through graphs and video, through real-time linkages to databases, and linked to contextual information provides for a major step change from the static nature of books, whether in printed form or acrobat file. The ability to make statistics understandable, interesting and relevant can increase greatly, at little cost.
5. Through the power of contemporary web based tools in providing content sourced from databases, the internet enables powerful searching of information with or without structure, making accessible huge repositories of information perhaps originally organised in ways that were only relevant to the uses for which the information was initially designed. Users can easily gain familiarity with what is available, and readily retrieve it, and learn of whether similar material exists, with little time or skill. The users' desktops can be customised to fit their interests, in the mapping, searching and presenting of information.
6. The internet can enable information to be efficiently provided in the great variety of forms relevant to official statistics, particularly text, statistical tables, graphs, maps. Such releases can be made available in complete studies, as tightly focused pictures of the state, organised to whatever extent is necessary to make a reasonably complete parcel of information.
7. The internet can integrate data, voice, visual and other means of interaction. The enables each and every communication to be concluded in a satisfying way, across the huge variety of circumstances through which users of statistics engage with official statistics, regardless of how it was initiated. It provides for the resulting service delivered to users to be able to be customised to their needs, and ability to absorb statistical results.
8. Monitoring of activities of the statistical office web-site, including degree and form of access, enables patterns of behaviour, levels of use, and reactions to change to be cheaply and quickly assessed, at the level of the system, the specific service, and the individual user. Changes in demand, and the value placed on new services can be quickly and accurately measured, in aggregate, and for individual or groups of users.
9. The internet enables a high level of integration with related services, including billing and promotion of new material. It also facilitates the efficient, downstream production of internet based services by added value information providers

The nature of the statistical office web-site

10. The statistical office web-site has the potential, among other things to become the prime means of delivering official statistics, managing statistical enquiries and meeting obligations to provide access to information about statistical practice. All statistical office services could be

designed to service web based access, and other forms of access will be efficient by-products of the web service and associated facilities, with few exceptions.

11. The statistical office web-site could be a wholly integrated part of internal statistical office statistical and business systems. Information will be organised so that the power of the web tools and browsers for unstructured searches can compliment any structured access paths. The web focus of all internal processes will not usually need any intervention in business processes to provide for the results to be delivered on the web, and any that is will be, of necessity, quite limited.

12. Statistical material and other information and web-site management will be through presentations that are designed to appear firstly as quality presentations on an internet screen. Where printed, or where delivered in machine readable form, such as excel spreadsheets, that will determine the design of the information.

13. A typical web-site might consist of five quite distinct elements, which will be linked where relevant.

i) Pictures of the nation: These will be a series of graphical and text based presentations with dynamic elements, which present self contained and usually short stories about particular topics. These will be dynamically updated, where they are focused on the latest available measures. They may contain videos, audio, graphics and other tools embedded in them, that provide a dynamic and interactive element. There will be a high editorial contribution from the statistical office, including design and selection not only of measures but how they are presented. Particular elements will be

- The analysis and main results associated with each first release of statistics
- The nation's yearbook in a new form
- Topic reports that reflect the evolution of existing services such economic trends, social trends, labour trends
- The nation in an international context

ii) Analytical Reports: More in depth analysis of social, economic or environmental issues including development of analytical concepts and frameworks. Will also include more policy relevant reports drawing on the pictures of the nation

iii) Information gathering: This will provide structured and unstructured access to:

- the formal publications and presentations of the statistical office, as well as available but unpublished tabulations and series.
- the structured statistical data-banks necessary for making available at low cost repeated access to indexed, machine readable data, particularly time series, regional statistical aggregates and accompanying aerial descriptions, and multi-way cross tabulations, most of which are based on a few critical cross cutting

variables such as family, industry, sector, commodity, population group, age or sex.

- Statistical summary data-files in special areas particularly foreign trade, where a well defined set of analyses of individual import or export commodities exists.
- Such data base access will cover data sets of both the authoritative estimates in a particular field of interest and access to the results of individual source material drawn from censuses, surveys and administrative sources.

iv) *Secure access to confidential unit record statistical data sources:* Retrieving information from survey and administrative data sources

- Statistical unit record database extractions, where users can define tables or other aggregates to be extracted from unit record databases using flexible self specified enquiries, and receiving results that are tested against confidentiality protection rules.
- Modelling environments, which may be made up of a mix of unit records, statistical summary information or other statistics, and to which particular models can be applied, or which are linked to generalised modelling environments such as SAS.

v) *Documentation:* This will provide structured and unstructured access to:

- Writing about statistical practice, statistical processes, methods and tools. It will also provide access to meta-data of all forms, including classifications, definitions, frameworks such as ESA,
- Information about each statistical survey
- Information about the statistical office
- Information about National Statistics
- Catalogue of services
- National statistics strategies and work programme
- Links to international and other relevant country web-sites

There will also be special services that develop for particular user communities, focus on particular products, or increase visibility of something important. Examples are;

- Education services
- Communications channels (feedback, email, voice messages, trigger for personal contact)
- Media services

Distinct faces of the statistical office web-site

14. The statistical office web-site may need to have two quite distinct faces, each of which will present a distinct menu of the available statistical office internet services.

i) A public face, which will be focused on immediately usable material, free access, supported by a fixed amount of telephone support each year. All documentation will be available, and several information gathering services will be available even where chargeable, so that users can learn of their existence. The public face will be designed to need no training for its use. The monitoring of each user's access will enable some customising of feedback to users, and we could plan to have a variety of front ends to the web-site, if distinct groups of users had specific and particularly different needs that justified that. There will be an integration of voice with data.

ii) A "subscriber managed" face, that will provide services that are either needing access to confidential records held only in statistical office custody, or which use services that are provided on a subscriber basis, or which need training and support for their effective use, such as the time series access service. The needs of each and every user of this service will be of sufficient value to the statistical office, or of sufficient importance to government, that they might justify their own access path through the web-site services, as well as their own approach to linking to parts of the web-site content either on their own internal network, through the internet or through some private network. The integration of the statistical office web-site with other services used by these users could be customised to their needs. Training will be available for services that need it. Security will be comprehensive, and user specific, using civil service wide infrastructures where available. There will be an integration of voice, data and video. A chargeable service enabling access to statistical office experts will be available.

15. Having two distinct web faces will ensure that the public face is consistent with free access to statistics, has the highest level of consistency in the look and feel of its services, has minimal security overhead, and is reproducible in many ways. The subscriber managed face will firstly be designed to deliver to policy ministries a comprehensive, easy to use, always available and well supported access to statistical information to ensure that all relevant information is discovered and accessible, for whatever question is being addressed. Other specialist users outside government will also use these specialist services, within a suitably secure environment and in their strict legal and ethical constraints. All procedures for the subscribe managed face will be transparent and open to scrutiny.

Integration with production processes

16. The statistical office web-site will be fully integrated with statistical office systems and processes to ensure that material is always up to date and properly quality assured. Document management in some sites such as Statistics New Zealand is managed through mirroring the information provided in a Lotus Notes based knowledge management system, so that the marginal costs of transferring information into the web environment includes no translation cost. The preparation of statistical pictures should involve tools that belong to the statistical office

standard tool-kit, and which by design integrate at no extra cost into the web environment. This might include Microsoft tools such as excel, or necessitate a new standard graphing tool. The dynamic access to statistical measures needs contemporary web integratable summary data management tools for time series (Fame is a good example), Regional data (Supermap is a good example), and multi-way tables (Supertable or IBM Pivot table are good examples). Statistical office tools in this area may well all need upgrading.

17. The information management capability should accept enquiries on statistical records driven by web managed tools, and should provide extractions of data in a form that readily inputs into the standard web integrated summary data management tools.

18. There should be no web content that requires operator intervention or special processes to translate the information onto the web-site. The consequence of this is that the statistical office web-site will be developed in parallel with Information management initiatives.

Web based improvements in the presentation of official statistics.

19. The web creates the potential for the presentation of official statistics to radically improve in quality in several ways. Each of these will challenge the existing practices of the statistical office, but also will generate more added value from existing major work such as the studies that make up the UK Social Trends. We will see:

- Timeliness improvements
- Quick extraction by users of indexed, machine readable statistical series, and of graphics.
- Greater emphasis on graphics, and maps
- New forms of presentation
- More relevant integration of information
- Dynamic response when changing the cross cutting variables of any analysis
- Ease of quick linkage to related sources of information, and to relevant statistical services
- Easy links to meta data associated with any series
- Easy fit of results of enquiries with other electronically provided outputs

Implications of a centralised web publishing service

20. The statistical office information technology architecture will embrace the internet, creating a wholly integrated environment. This integration covers technology, security and information, involves tools, systems, processes, management, client, suppliers, reporting. We will create high value from communicating via databases, and have efficiencies gained by structuring information that was previously unstructured. There will be a very large web site with authorship

decentralised but control centralised. There will be considerable ease of exchange of information with partners who use notes. The statistical process and data access benefits we will obtain from the fully web integrated architecture will be

- Extensive public documentation at low cost
- Hugely improved performance monitoring
- Lower cost structure for all activities
- Reduced cycle times for new activity
- Credible capacity for more innovation
- More rapid creation of new products
- Able to manage more change at any time
- Web can provide several channels for supply and delivery of services
- Much reduced communication costs
- Easy links to any relevant organisation
- All statistical office statistical services are able to be digitised
- Web compatible services penetrating down into all business processes
- Exchangeable business systems and processes

21. The new statistical office web-site will be a statistical system wide facility, relevant to all statistical outputs and services, and all statistical outputs and services will use it, as the prime means of service delivery. This will require strong corporate standards that are relevant to all the work of the statistical office, state of the art corporate systems that are used locally, everywhere, and the application of corporate practices at all times. To facilitate this, the integration of technology environments will necessitate a high degree of standardisation of systems and equipment, integrated software and data management tools, one statistical office publishing policy, and one release process. We need to determine our corporate tools, e.g. Lotus Notes, and apply universally without question. A strategy for settling on new publishing approaches for statistics needs to be prepared. We need to decide on what sort of books and other printed material we will continue with, to balance what the web can now best deliver.

22. The new statistical office web site will be tightly integrated with the new statistical office information management and knowledge management tools, and the developments will be planned and implemented in parallel.

Data collection via the web

23. The internet offers new possibilities and challenges for the *collection* of official statistics. This will be true across the spectrum of data collected by the statistical office.

24. Over the last decade, new technology has transformed the world of social surveys of the general population. Improvements in the speed, quality and cost of data collection and dissemination have flowed from the introduction of computer assisted interviewing, either face-

to-face or by telephone. In the coming years the statistical agency must examine how it can utilise the web to deliver similar step improvements over the next period.

25. Increasing numbers of citizens will become accustomed to using the internet and will expect to conduct their business with government over the web. We can expect this to be true of citizens as *providers* of official statistics, for example as respondents to censuses and surveys. It is possible that the option of completing a survey interview via the web might help arrest the falling response rates witnessed in recent years, particularly amongst certain sub-groups of the population.

26. Developing the use of the internet for use on surveys of the general population raises a number of challenges for the statistical agency. For example questions such as:

- How can the web be used in surveys that require a nationally representative sample?
- How much of the data currently collected by survey interviewers can be collected via self-administered questionnaires completed over the web?
- Is the web best used as a single mode or as part of a mixed mode strategy, for example alongside face-to-face or telephone interviewing?
- What is the optimum length of a survey questionnaire via the web and how do we discourage break-offs?
- What procedures and safeguards need developing to ensure the privacy and security of respondents' data, both real and perceived ?

27. Conclusion

28. For any statistical office, the web-site is a critical initiative. We need some clear direction on its focus, integration with statistical office infrastructure, and the potential changes that it will make possible. The technology options are considerable, and when we have succeeded in them, we will find that business change may be slower than that of the technology. New technologies that exist now give statistical offices the chance to be at the forefront of public sector modernisation, provided we rethink; and redevelop the statistical services we provide. The internet will be fully exploited by the national statistical office when there is a civil service wide secure network to work within, all agencies have a similar level of competence in IT matters, and public access to the internet is judged equitable. This paper is about what we can achieve with or without these broader successes, and about the contribution we can make to making these successes visible when they occur.

- - - - -