



Economic and Social Council

Distr.
GENERAL

CES/AC.71/1999/13
25 January 1999

ENGLISH ONLY

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Meeting on the Management of Statistical Information Technology
(Geneva, Switzerland, 15-17 February 1999)

Topic (i): The impact of Internet on the statistical production and dissemination process

INTERNET TECHNOLOGY IN PUBLISHING OF STATISTICAL INFORMATION

Submitted by INFOSTAT, Bratislava, Slovakia¹

I. BACKGROUND

1. The Internet with its Web sites offers an interesting and immediate solution for presenting information in a comprehensive form to users. Most statistical offices in the ECE region have already recognized the importance of Internet technology for the distribution of statistical data. Projects for data dissemination in statistical offices were launched using Internet services and Web technology. They generally aim at providing statistical data to users in a comprehensive electronic form without any time delay. The electronic form has certain advantages for the user, especially in cases where there is a demand for further manipulation of the selected data after transfer to the user's computer environment.

2. Most of the projects on presentation of statistical data on Internet aim at:

- the replacement, or at least limitation, of paper-based presentation and

¹ Prepared by Dusan Prazenka.

distribution of statistical data;

- providing fully interactive services for the users of statistical data;
- allowing a maximum number of users to access and obtain statistical data in comprehensive forms;
- development of applications that support the on-line transaction processing and delivery of statistical data to users;
- marketing statistical services provided by NSO's.

3. Experiences from implementation of projects on data dissemination presented in papers or at conferences show that the design of dissemination systems is a multi-professional task. These experiences also show that many components of the statistical data processing system should be integrated to achieve effective dissemination results. The design of the system in many cases may also have an impact on the existing work style of the statistical office with regard to preparation of data for publication and/or dissemination. Displaying the data on an Internet Web site does not usually involve the same data selection and presentation procedures as used in the preparation of a paper-based publication. It is better to present a concise well-structured set of data for navigation on the Internet Web site. A large amount of detailed data posted on the Web without proper structuring may discourage Internet users; the data should be presented in a more condensed form. Even though graphs and maps, as well as pictures, increase user-friendliness, they should be used carefully since they may slow down the data transmission process and presentation of data to Internet users. The development of a dynamic Web-site with interface to statistical data bases provides the users with the opportunity to access data in real time.

4. A team at Infostat that has developed the system for data dissemination on Internet gained some experience in distributing the data in electronic form from the development and operating BBS-based system. The BBS-ELIS (BBS-Electronic Information Service) for on/line dissemination of statistical data was implemented using the Major BBS software. As for the content, the BBS ELIS database is composed of:

- textual information taken from the paper based information;
- parts of publications prepared in Word for Windows editor;
- parts of publications prepared in EXCEL;
- parts of publications prepared using other software products (pictures, graphs).

5. As the BBS-ELIS system provides data to users only by downloading the pre-prepared files and hence static information, we began the design and, later, the development of a more flexible system for providing data using Web technology tools. In the first phase of the new system development, we launched in 1996 a pilot project using Web technology for the presentation of the same scope of data as in BBS-ELIS database. In the second phase of the development, a final version of the system was developed (ASIS ELIS subsystem) as part of the larger ASIS (Automated Statistical Information System) system;

this is being developed by Infostat and implemented gradually at the Statistical Office of the Slovak Republic (SO SR).

II. WWW-ELIS PILOT SYSTEM DEVELOPMENT

6. The development of WWW-ELIS for statistical data dissemination using Web technology and INTERNET services was launched as a pilot application last year. The main goal was to verify methods and technology which are used for building WWW pages on the server dedicated to statistical data dissemination. In addition to this, the pilot project was also aimed at the transformation of routines used for the transfer of data (statistical publications in electronic form) from the formerly used BBS-ELIS to HTML format. We plan to verify telecommunications services as well as services of the INTERNET provider.

7. In the pilot application we have verified:

- conversions of the BBS-ELIS formats into HTML format used in Web server;
- structure and inter-linkage of HTML documents (individual publications);
- Web client software for document browsing or for its transfer to client machine for further processing;
- technology both on server and client side and communication tools.

8. For the pilot application we selected publications from the already existing BBS-ELIS data base, converted them into HTML format and created a WWW homepage. The documents (publications) were to be sorted by forms, type and date of release.

9. Since the WWW-ELIS services, like the BBS-ELIS services, should also be provided on a commercial basis, we decided to complete the homepage with additional information on INFOSTAT, SO SR, marketing of statistical services and price list.

10. A software that was developed for Web server in the pilot application has embedded a function for registration of clients' access to particular documents. The function allows the gathering of data on access frequency to different statistical information. This data, after evaluation, may be used for the optimization of server data base and client access.

11. The defined structure of the homepage gave clients access to statistical information - a static approach. Clients can access data via a list (catalogue) of publications or via a list of indicators grouped by subject.

12. The pilot application verified the conversion routines and the ability of staff to create a homepage with structured statistical information. The contractual terms with the provider for extension of services beyond the scope of the pilot project were negotiated between INFOSTAT and the Internet service provider.

13. Experiences from operating the pilot version of the dissemination system

(WWW - ELIS) have shown that the further development of data dissemination system using Web technology should concentrate on the following:

- Providing a broader scope of statistical information for different categories of users (public, state administration). This would mean offering statistical publications and providing users with access to other statistical data resources (e.g. ORACLE databases).
- Selection of appropriate software for WEB server with a set of protocols to enable:
 - the coding (encrypting) of documents, to make an automatic conversion of documents into HTML format;
 - the use of HTML v 3.2, the latest version of HTML which deploys features as tables, text flow around images, and superscript and subscripts;
 - to provide a sufficient level of data security especially for individual data;
 - to provide statistics on client access and on use of queries.
- Selection of an operating system. The Windows NT Workstation was recommended. This platform supports an automatic conversion of source data (Word, Excel) into a form acceptable to arbitrary client viewers.

Besides the hardware dedicated for WWW server, additional workstations for data preparation and conversion are requested. An essential feature of the WWW- ELIS development should be the technological and functional compatibility with the newly developed ASIS.

III. ASIS ELIS SUBSYSTEM DEVELOPMENT

III.1 Subsystem specifications

14. As specified in the conceptual design of the ASIS, the ASIS ELIS is an integral part of the ASIS. Its functionality is oriented towards presentation of data processed either within the ASIS or off the system provided that the standard interfaces are followed. A decision was made that the ASIS-ELIS should, to a maximum extent, apply the Internet Web technology for presentation of data and should be developed in ORACLE. . The ASIS-ELIS should include developed interfaces to all kernel subsystems of the ASIS i.e. DBS (Data Base Subsystem), METIS (Meta-information subsystem), REGIS (Subsystem of Registers) and ZBER (Subsystem for Data Collection).

15. A major goal was to develop the system to provide statistical data to different categories of users (public, entrepreneurs, governmental bodies, international organizations) including internal users from the statistical office of the Slovak Republic. A by-product of the subsystem might be the preparation of data for publication in paper form.

16. The key parts of the ASIS-ELIS specifications were those of functions and of interfaces to other ASIS subsystems and, last but not least, a specification of standards for data dissemination. The recommendations of standards was a part of the next level project document - the technical specifications of the subsystem.

17. The standards that are applied on statistical data devoted to dissemination may reasonably speed up the process of data distribution, and decrease the time elapse between the reference periods of the data and time of its distribution.

18. Specifications covered the following topics:

- functional scope of the subsystem covering the functions for different user categories;
- functions for preparation of data for dissemination;
- coverage of data for dissemination (list of publications, list of other sources directly accessible for the users via the subsystem);
- proposal for data dissemination standards;
- outline for design of homepage for WEB server for data provided in INTERNET;
- recommendations for WEB server and client software;
- recommendations for conversion routines;
- recommendations for "intranet" services;
- proposal for design of presentation and dissemination data base;
- subsystem interfaces;
- interfaces to other ASIS subsystems (METIS, REGIS, DBS);
- interfaces external to ASIS;
- the role and functions of the "public data base" (PDB) in the ASIS-ELIS.

19. Conditions to be fulfilled in order to implement the ASIS-ELIS in its full scope of designed functions have been considered in the system specifications. The specifications for hardware and software, communication lines, interconnection to WAN, provider of the INTERNET services, developing groups, system operability, conditions for ASIS-ELIS interoperability with the other ASIS subsystems were drafted and later approved as a basis for system development.

III.2 System development

20. Based on the subsystem specifications, we have developed the software using PL/SQL of Oracle 7.3 and Oracle Web Server 3.0 which provides a flexible tool for the presentation of different types of objects stored in the PDB. The ASIS-ELIS software consists of two relatively independent software packages:

- a package for administration of different objects and users as well as for the preparation of objects for presentation in the PDB;

- a package for presentation of statistical data on Web server.

21. The package for administration consists of two groups of functions.

For a Web administrator to create and maintain:

- a list of ASIS-ELIS users;
- objects in the PDB of the ASIS-ELIS;
- additional objects such as, for example, code lists, classifications etc.;
- data objects designed and pre-prepared by "internal user" from subject matter department of SO SR and their transfer into PDB of ASIS ELIS.

For an internal user from SO SR who prepares the data objects for presentation in PDB, a special menu is prepared to describe the data objects.

22. The package provides the graphical user interface (GUI) to the administrators written in the MS Visual Basic 5.0.

23. The package for presentation of statistical data uses Oracle Webserver 3.0. It allows a dynamic transfer of data objects from Oracle data base. For this reason, we have implemented an interface of the package to the ASIS Data Base Subsystem (ASIS-DBS). Based on descriptions of the data objects, it automatically creates the menus for the users connected to the system in real time. The menu provides the users with the tools for navigation and selection of data to a level of individual indicators. The menu also offers users several possibilities for selection and presentation of data on the screen.

IV. CONCLUSIONS

24. The Internet services and Web technology was a new challenge for the developing team, which has three years' experience with dissemination of data in electronic form. At the beginning of the BBS operation, we had only a few requests for statistical data in electronic form; users still preferred data in paper-based form. At present, the number of users is rapidly increasing each year and the BBS system is reaching its limits. Therefore, we have begun the design of the new system, aiming at building the Web site and using Internet services. Transfer to the Internet and Web technology has opened a new perspective for statistical data dissemination with the new facilities for users. The lessons we have learned in the application development could be summarized as follows:

- the concept of the Public Data Base in combination with the Web technology brings statistical data closer to users from the general public;
- a dynamic Web page creation gives the user the feeling of working with a real statistical data base;
- the concept of Internet/Intranet may efficiently serve both external and internal users i.e. no special effort is needed in transformation of data and its structure for either internal or external use; HTML creates a common platform for both categories;

- the style of work in the preparation of statistical data within the statistical office for publication should be changed when switching to presentation and dissemination via Internet;
- graphic design of Web pages should be modest but should still attract the Internet users; hence a professional approach is necessary;
- the correct selection of data for presentation in the Web site is crucial: well-established and long-term planning is therefore needed;
- good interfaces to the statistical data base (for data selection and transfer) are important for increasing the efficiency of the presentation.

25. The implementation of the ASIS-ELIS was completed in accordance with the specifications adopted so far. At present, the system is in the verification phase and the users can access the selected data from industrial statistics and demography via Internet. We are preparing an extension of the PDB content with additional statistics. We also plan to extend the functionality of the ASIS-ELIS upon completion of the ASIS and after evaluation of the users' requests from the verification phase.