

UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on Statistical Data Collection
(Geneva, Switzerland, 25-27 September 2013)

Topic (iii): Improving the respondent experience

**SPECIFIC FEATURES OF THE ORGANIZATION OF INTERACTION
WITH RESPONDENTS DURING THE TRANSITION TO CENTRALIZED
ELECTRONIC DATA REPORTING**

Working Paper

Prepared by Vadim Bogush, National Statistical Committee of the Republic of Belarus

I. Introduction

1. The National statistical system in the Republic of Belarus is determined by the Law “On the State Statistics”, other legal acts and international agreements. This system includes National Statistical Committee of the Republic of Belarus (Belstat) as the republican body of public administration in the field of state statistics and other producers of statistics and administrative data.
2. The main tasks of the Belarusian state statistics are:
 - development of science-based methodology and its improvement in compliance with the national and international statistical standards;
 - data collection, processing, aggregation, accumulation, storage and protection on the basis of statistical methodology;
 - provision of summary statistical data to the President and the Government authorities of the Republic of Belarus;
 - dissemination of summary statistical information.
3. Today the Belstat’s system consists of central office and seven regional offices with geographically distributed statistics divisions in 133 districts and cities as well as several subordinated organizations – computing centers that support dissemination of statistical data and take part in development and implementation of modern information technologies in statistical production process. Application of modern computing technologies and data processing systems and software is one of the most important tasks of the strategy development roadmap.
4. Since the organization and management system of Belarusian state statistics is mainly centralized the technological infrastructure goes through centralized modernization as well. In this case is

particularly important to create the effective interaction between both technical and organization parts of data creators (sources) and receivers including proper level of data reliability, accuracy and authenticity, information security and technical support for all users of information system.

5. This paper presents some positive results, experience and critical issues of transition to centralized technological structure for data collection in the Republic of Belarus.

II. General overview of the collecting data technological infrastructure of the National statistical system in the Republic of Belarus

A. Organization of observations and data processing

6. Generally the procedure of formation of statistical information is based on carrying out of centralized observations where respondents are business, non-commercial organizations, householders, citizens that goes through continuous or sample surveys. Aggregated data is also generated using of other state statistical observations carried out by authorized organizations and administrative data provided according the annual Statistical program approved by the Government. The most significant problems for formation of summary information are harmonization of metadata and time synchronization of data collection process.
7. The statistical process in the past was based on paper report collecting system followed by data input at district or regional level, control and aggregate data and formation of part of general database for each observation. Then data are processed using various software based both on adapted ready-made solutions (e.g. SPSS), and on specially developed software.
8. In this case the duty distribution between the three levels of statistical system was focused on methodological issues and technology development as well as formation of summary information at national level and data input and control at lower level of districts. The main technology was based on file transfer approach.
9. Modern level of telecommunication networks and highly productive equipment with novel database management software allows transformation to centralization of data collection process and development of new applications based on common metadata including classifications and statistical unit register.
10. Belstat is now in the process of implementing new approaches to data collection from organizations, their processing and provision to users – government authorities, business community and general public.
11. One of the key projects aimed at the implementation of the new approaches is the creation of the Integrated Information System of State Statistics of the Republic of Belarus (ISS). This project was focused on modernization of telecommunication structure and development a novel technological approach and tools for metadata management, primary data collection and aggregated information processing and distribution. In the frame of the project a subsystems for metadata, micro and macro data collection and storage, publications dissemination were created.

B. Data collection process description

12. Basically, Belstat collects primary data from respondents through statistical surveys, the main form of which is the submission of statistical reports by businesses. The interaction between the state statistics bodies and businesses within the framework of ISS is aimed mainly at the submission of statistical reports in electronic form via the Internet using electronic digital signature (Figure 1). The process of electronic statistical reporting is centralized, and the interaction with respondents is carried out off-line using special software module. This software (and also corresponding guidance materials) are available on the official website of Belstat and can be installed on the respondent's computer.

13. After the installation this module interacts via the Internet with Belstat's receiving centre, where the questionnaire templates and other metainformation required for filling the report are posted. There is also the possibility to interact with the Certification Centre of tax authorities, which allows for the use of a single key for the submission of both statistical reports and tax returns.

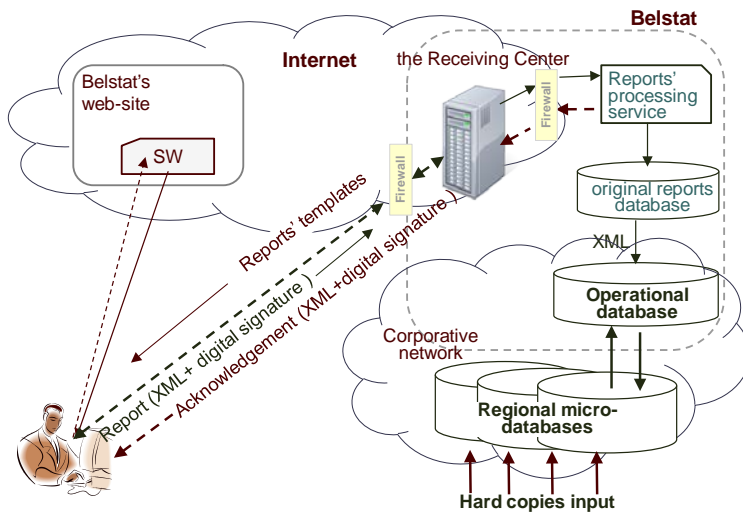


Figure 1. Collection and processing of statistical reports within ISS

14. Advantages of the respondent module are:

- user-friendly and the same interface for different reporting forms;
- the latest version of forms and required metadata (synchronization with receiving center);
- possibilities for data validation and collection data from different units within local network;
- respondent's burden decrease;
- proper level of the information security is provided by encryption and digital signature application;
- electronic confirmation of report acceptance.

15. The ISS software and hardware package allows for the geographically-distributed data entry and integration of data provided in different forms (on paper and electronically) into common databases which are used for the compilation of summary data.

C. Telecommunication and data processing infrastructure

16. Operation of applied software required a modern computing network infrastructure including datacenters which allows the workstations of statistics workers to be linked into a common telecommunication network. A general scheme of the ISS computing network infrastructure is shown in Figure 2.

17. On the basis of local area networks of territorial offices and Belstat a corporate data processing network is established which links the workstations in Belstat, seven regional offices and more than 130 district statistics divisions. Totally the network capacity is up to 2.5 thousand points (workstations, servers and other equipment). For data exchange between statistical offices the corporate e-mail system based on Microsoft Exchange, office automation and electronic document management system, and specialized data collection and processing software are used. The network is built on the basis of modern active network equipment (CISCO, Hewlett Packard, etc.), which allows for the centralized management of the key system nodes. Engineering infrastructure of data

processing centers of Belstat and regional offices is built using generally the Hewlett Packard equipment. In order to ensure failure-free operation, the technologies of clustering, data back-up and uninterrupted power supply of important system nodes are implemented. The servers and workstations use Windows OS, for databases Oracle DBMS is used.

18. In order to ensure confidentiality, consistency and integrity of data, modern security methods and systems are applied. These include firewalls, intrusion prevention system, user access control, audit and monitoring of the information security violation events. Cryptographic data protection is applied. Antivirus software is installed on all workstations and servers.

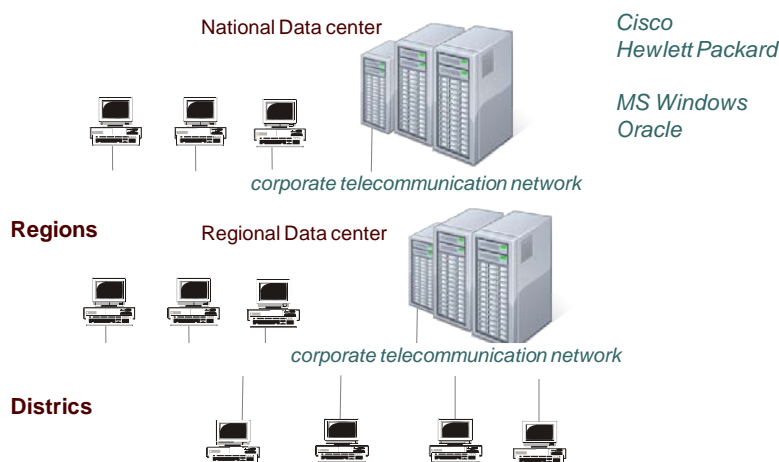


Figure 2. General scheme of the ISS computing network infrastructure

19. Local area networks of state statistics bodies are based on Ethernet 1000BASE-T technology which allows for the connection of automated user workstations at a speed of 100-1000 Mb/s. For internetworking within the corporate network Belstat and regional statistical offices are linked using the fiber optic communication lines with a bandwidth of 6-30 Mb/s, and district statistics divisions – using ADSL technology with communication channel capacity of 512 Kb/s.
20. Interaction with respondents is organized using a high-speed Internet line (10 Mb/s) for the receiving centre of Belstat with additional lines for staff to work with web-site and other global network resources. Territorial statistical offices have additional communication channels using ADSL technology to provide the staff with the Internet access.
21. Information support of both respondents and users of statistical information generally is carried out using official Belstat's website. The new Belstat's website designed within the framework of IIS is a multi-purpose information system integrating a set of information resources, such as a bank of e-data books and databases. It has seven sub-domains, each of them containing data of the relevant regional statistical office and managed as a separate site.
22. Government agencies are provided with summary data in the form of e-publications using special protected transport mail service as well.

III. Organization of centralized data collection and future plans for development

D. Technical and methodological support organization

23. Transition to centralized data collection requires redistribution of responsibilities within levels of statistical systems with increased load at republican level and reduction of duties at district level.

24. A three level support organization structure were developed where the first line of the information and methodological support is focused at district and regional level and general information and methodology development is generated at national level. All information regarding terms and organization of data collection process is published at the special section of official Belstat website.
25. Distribution of information and presentation of software were made using local media, national newspapers and TV, sending invitation letters to respondents using e-mail and ordinary mails.
26. Technical support for users of applied software was provided by technical staff at regional level, at central level that also is in charge of data collection monitoring process and support of receiving center and databases operation and by software developers.
27. Several phone hotlines for each region were organized and regional groups were determined as a main point for respondent questions collections and answers. All questions and proposals of users are analyzed at central level together with developers to software enhancement and update. Distribution of updated software goes mainly through placing the updated installation package at website and at receiving center.

E. Switching to the centralizing data collection

28. At present Belarus statistics is implementing a new technological platform for data collection and switching to centralized management structure. Development of new technological approach was started in 2007, a software testing has been launched with two pilot project for industry organization in 2010 and small business in 2011. Implementation of new software for data collection was begun in the end of 2012.
29. Organization structure of data collection process in the past was based on data processing management by the specific field of statistics using different software module for each observation. All observation was managed independently. General management of statistical production (including terms and information description) is made according to annual, quarterly and monthly plans that are generated at national level.
30. A developed approach with centralized processing and storage center allows carrying different observations at the same time but using the same metadata source. This raises additional problems with synchronization of metadata changing with terms of beginning and ending of survey.
31. Familiarize respondents with new software is also a key problem that was taken into account during preparation to electronic report collection. That's why Belstat chose a "step-by-step" strategy for expansion of observation amount and respondents involved into new data collection process started from monthly and quarterly financial and industry statistics and annual small business activity observation. A transition to electronic report submission for organization was unforced
32. At the first stage in the end of 2012 monthly statistical reports were collected through several industry and finance statistics questionnaires. The dynamics of amount of organizations switching to electronic report submission is presented in Figure 3.
33. It is characterized by growth in first several months and upper level about 90 % due to technical level of respondents and additional expenses for digital signature purchase.
34. For example, according to 2012 results, 1546 respondents (or 81.1% of all relevant reports submitted) reported on the form "Report on production costs of products (work, services)", and 5112 respondents (or 60.4% of relevant reports submitted) reported on the form "Report on financial performance" using new software.

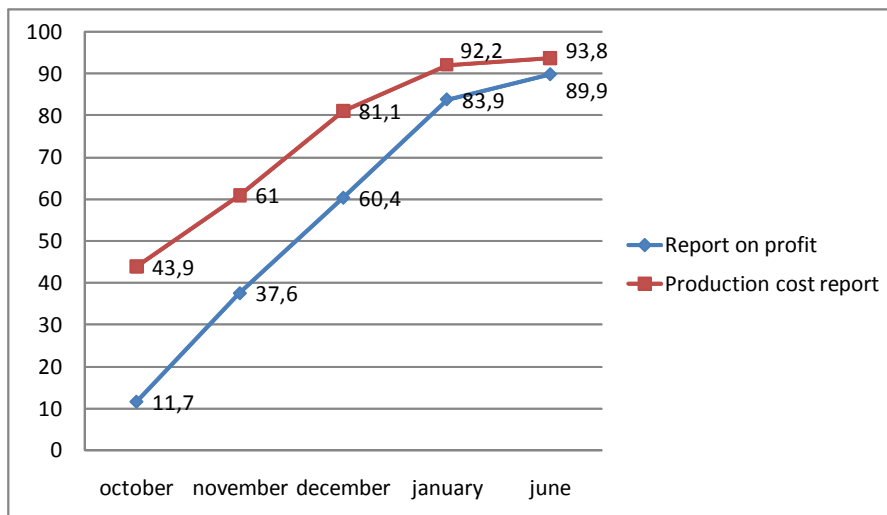


Figure 3. Dynamics of implementation of e-submission technology (in percentage of total number of respondents)

35. It should be noted that introduction of other reporting forms that can be submitted using the same software but prepared by other personnel allows realization of e-submission at level above 80 % from the first period of data collection.
36. For example, the first period of collecting data on report forms on industry production and on status of settlements is characterized by 79,4 % and 83,3 % respondent involvement respectively.
37. Interaction with small business demonstrated less positive results and depending on the region from 30% up to 52 % from total number of respondents used new software (total number of electronic submissions was about 20 thousands). This also associated with needs for additional methodological support for annual reporting and personal consulting.
38. Improving the respondent experience is reached through seminars, publication of additional information at official website and special magazines, updating answers on the frequently asked questions.
39. Practical experience has shown that maximum load on support team was at the first several months of implementing of new technology then only methodological changes support is needed. One of the major problems is interaction with respondents who participates in the survey for the first time.
40. The issues of storage of collected reports and current data solves by development of electronic archive system that also manages at the central level.

F. Future plans for development

41. The development of information technologies of state statistics bodies and interaction with respondents and data users is planned in the following main directions:

implementation of the project on the creation of Belstat's electronic archive, which would allow for the storage and work with archive e-documents, including the receipt of statistical reports in the form of e-documents from ISS, issuance of archive certificates for individuals and government agencies, modeling and certification of paper form of e-document;

development a multipurpose Belstat web portal focused on giving necessary respondent technical and methodological support.

42. The main advantages and new function of the portal would be:

ensure primary data collection in electronic format (also in the form of e-document using digital signature) both on-line and off-line with central data warehouse of respondent reports and data;

provide information support to respondent (calendar of statistical reporting, personal profile of respondent, access to regulatory and reference and other information required for primary data collection, consultative support to respondents).

43. Further development of the web portal will focus on implementation of novel technologies for information dissemination mechanisms by posting statistical databases on the web and providing access thereto to different user groups via the official website of Belstat including mobile access.

44. We also consider implementation of modern methods of data presentation, in particular, geographic information systems. Establishment of a GIS portal would allow Internet users to get statistics in the form of cartograms.