Modernization of Statistical Information Systems

Speaker presentation: Agency of the Republic of Kazakhstan on Statistics (ARKS): Modernization of statistical information systems /1

1 This document was contributed by the Deputy Chairman, Agency on Statistics of the Republic of Kazakhstan. It has been reproduced without formal editing. The views expressed are those of the authors and do not necessarily reflect the views of the United Nations.
Development and spreading of modern information and communication technologies today have a growing impact on all spheres of society’s activities, both within national borders and in the world as a whole.

A lot of activities aimed at development and creation of modern methods of observations, collection, storage and processing of statistical information are carried out in today's reformation of state statistics in Kazakhstan. One of the main conditions of these activities is the orientation towards the creation of an integrated system of data processing.

During 2011-2013, the Agency of the Republic of Kazakhstan on Statistics (further Agency) is planning to undertake a Project on modernization of the statistical information system. Modernization of the statistical information system involves the creation of the new information systems and development of existing ones.

At the moment, the statistics of Kazakhstan uses such information systems and components as:

- **Population Census Data Base**. The database is used for the storage of detailed information about the population of the Republic of Kazakhstan, provision of quick access to information, sampling and timely extraction of information from the database.

- **Agricultural Census Data Base**. The system is designed for the storage of agricultural census micro data with the possibility of accessing the data for sampling and reporting.

- **Information System “CLASS”**. Information system is designed for the automation of processes of the development, implementation, maintenance and dissemination of statistical manuals and classifiers and for the provision of the distributed access to the information, contained in it, for other internal and external information systems.

- **Information System “Electronic Statistical Reporting”**. The system is designed to ensure the collection and processing of primary statistical data on the basis of the use of modern information and computing technologies, in order to effectively execute the core functions of the Agency on producing the official statistical information.

It should be noted that today only less than 30% of all statistical surveys are conducted in electronic form with the use of this information system.
More than 70% of incoming raw data is stored and processed in separate local information systems based on FoxPro, Access technologies, which has the following difficulties and disadvantages:

- Plenty of isolated local databases, hard-to-integrate with each other;
- Input, processing at the territorial level and aggregation of data at the national level are complicated due to absence of a single integrated information system;
- Lack of possibilities to use data from administrative sources entails an additional burden on respondents when completing statistical forms;
- Modern methods of data entry from survey forms and data processing are not used.

These factors contribute to considerable time and labor losses of the Agency and respondents.

In 2008 the Agency launched a new portal, where all information on statistical indicators and statistical reports is published and it is free for all. Unfortunately, due to the absence of a single integrated information system, most of the statistical information is published manually, and there are no web-services on the portal for the exchange and transmission of data with information systems of other government authorities.

Thus, the main problem today is that existing information systems have different structures and formats, which hinders joint use of statistical information resources. In this regard, in order to develop and improve the current information infrastructure, the Agency decided to create an integrated information system "e-Statistics (IIS “e-Statistics”)" on a single platform.

The purpose of creation of the IIS “e-Statistics” is a comprehensive automation of the business-processes and the effective provision of public services by the Agency and its regional divisions.

IIS "e-Statistics" will combine the set of the existing and new information systems of the Agency using advanced information technologies and telecommunication systems, integrated with the infrastructure of “e-Government” and information systems of other government bodies.

The main objectives of IIS "e-Statistics" are:

- centralize collection, storage and processing of primary statistical information in one database;
- provide necessary level of protection of primary statistical information;
- integrate existing information systems in the Agency in a single system;
- improve the quality and production of statistics through the use of administrative information;
• quickly respond to and adapt internal information systems after changes in methodology;

• automate the publication of statistical information and reports on the web-portal of the Agency;

• provide access for a wide range of users to the information-analytical system (visualization tools, aggregated historical and current data bases and analysis, etc.);

• provide a mechanism for interaction and information exchange with information systems of other government authorities on the basis of web-services;

• provide to respondents an alternative way of completing and submitting statistical forms (on-line submission) using web-technologies;

• automate data entry from survey forms using a scanner processing.

IIS “e-Statistics” architecture will consist of the following technological units (see attached Structure of IIS “e-Statistics”):

• Registers;

• Collection of primary information;

• Reference information unit - RIN;

• Data warehouse;

• Single system of data output;

• Integration component.

New components of IIS “e-Statistics” will be developed in the framework of these technological units, such as:

1. The integration component that will provide a solution for the following tasks:

• interaction with the current information systems of the Agency concerning the exchange of necessary information;

• interaction with the gateway of “e-Government” in terms of obtaining a reference regulatory information on databases provided by the gateway of “e-Government”, such as:
  - Legal Persons Database;
  - Individuals Database;
  - “Real Estate Register” Database;
  - “Address Register” Database.

• interaction with the gateway of “e-Government” in terms of providing information on requests from external institutions.

2. Information system (IS) “Statistical Register of Population (SRP)” will increase the efficiency of production of statistical information in the sphere of social and
demographic statistics. Information collected by the Agency during 2009 National Population Census will serve as a basis for the formation of the SRP.

3. IS “Agricultural Statistical Register” will automate the collection and accumulation of actual information on agricultural producers, agricultural land, cultivated areas, livestock and poultry by regions of the Republic of Kazakhstan.

4. IS “On-line Data Collection” will enable the respondents to submit statistical data to the Agency without leaving their offices. The component will be implemented as a portal solution for the collection of survey forms electronically through the online service.

5. Metadata. Due to the transition to a multi-dimensional data processing, there is a need to create a single information system of statistical metadata that would enable Agencies’ employees and all users to rapidly obtain the needed statistical information.

6. “Database of Administrative Information” will form a centralized database of micro data on the basis of the primary administrative statistical information. It will ensure the availability of large volumes of historical data for processing and analysis, as well as reduce the burden on respondents by the use of administrative sources.

7. IS “Processing of the primary statistical and administrative information”. The mission of the IS is to ensure the prompt and good quality processing of primary statistical data, and to form statistical indicators for release. It is assumed that this module will replace the existing many individual software systems, developed over 10 years ago.

8. “Aggregate Statistical Indicators” Database will serve as a basis for the implementation of different information-analytical tasks, including forecasting and comprehensive analysis of statistical information.

9. Information-Analytical System is designed for dynamic data presentation, analysis of aggregated historical and current data, trend analysis, modeling and forecasting of socio-economic indicators of Kazakhstan.

10. Internal intranet portal is created for the organization of a single point of access to information systems.

The implementation of the new conceptual model of information system will:

1. Improve the quality of statistical information due to a increased capacities of automated data collection, processing and analysis;

2. Reduce the financial and time costs of processing and dissemination of statistical data;

3. Reduce the time of the preparation of information on the requests of public authorities;
4. Improve the efficiency of functioning of the system by reducing the proportion of the low-productive work and increasing of personnel’s working time for conducting of comprehensive analysis of statistical information, better work with respondents, improving the means and methods of dissemination of statistical information;

5. Provide an opportunity for the integration of statistical information system with the nationwide system of information resources;

6. Provide statistical data to a wide range of external users via Internet technologies.

In general, the introduction and modernization of information systems in the framework of IIS “e-Statistics” will help to improve the systems of collection, processing, dissemination and analysis of statistical information. Users will have the access to aggregated statistical data through user-friendly interface. We believe that this approach will not only solve the problems of integration of internal information-statistical resources and their efficient use, but also create conditions for integration with information systems of other agencies and information exchange with international organizations.