NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Leonid Gokhberg, Tatiana Kuznetsova,
Gulnara Abdrakhmanova, Konstantin. Fursov,
Elena Nechaeva, Sergey Shashnov, Anton Suslov

PROSPECTIVE MODEL OF OFFICIAL STATISTICS FOR THE DIGITAL AGE

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: SCIENCE, TECHNOLOGY AND INNOVATION

WP BRP 111/STI/2020

Leonid Gokhberg[1], *Tatiana Kuznetsova*[2], *Gulnara Abdrakhmanova*[3], *Konstantin. Fursov*[4], *Elena Nechaeva*[5], *Sergey Shashnov*[6], *Anton Suslov*[7]

PROSPECTIVE MODEL OF OFFICIAL STATISTICS FOR THE DIGITAL AGE[8]

This paper describes key aspects of the structural and functional transformation of the Russian state statistics system as a core element of the future National Data Management System. Issues such as establishing a dialogue between the statistics service and users, integrating data from various sources, and intelligent data processing in the context of the digitalization of the economy are considered. New approaches and mechanisms should integrate and advance all of the previously achieved best results in methodology, observation areas, metrics, and other domains. Improvement areas include providing higher-quality information for policy shaping, businesses, individuals, and external partners. National statistics is expected to present an interconnected, objectively measurable model of socioeconomic processes and phenomena based on relevant theoretical concepts. In addition to using various sources of data, a necessary feature of the new system will be its reliance on a consistent conceptual framework and approaches to interpreting data, which will make it possible to integrate various data sources in the first place.

Users of intelligent statistics are becoming not only active participants in primary data collection, accumulation and application processes, but are turning into "smart" consumers who develop statistical thinking and are able to derive the greatest possible benefits from the use of statistical data. Such skills should become an inherent (and possibly mandatory) component of any.

JEL Classification: O32.

Keywords: statistics, digital age, data, Russia.

[1] National Research University Higher School of Economics. Institute for Statistical Studies and Economics of Knowledge. Director; E-mail: lgokhberg@hse.ru

[2] National Research University Higher School of Economics. Institute for Statistical Studies and Economics of Knowledge. Director of the Centre for S&T, Innovation and Information Policy; E-mail: tkuznetzova@hse.ru

[3] National Research University Higher School of Economics. Institute for Statistical Studies and Economics of Knowledge. Director of the Centre for Statistics and Monitoring of Information Society and Digital Economics; E-mail: gabdrakhmanova@hse.ru

[4] National Research University Higher School of Economics. Institute for Statistical Studies and Economics of Knowledge. Deputy Director of the Centre for Statistics and Monitoring of S&T and Innovation; E-mail: ksfursov@hse.ru

[5] National Research University Higher School of Economics. Institute for Statistical Studies and Economics of Knowledge. Director of the Centre for Socio-Economic Information Processing; E-mail: enechaeva@hse.ru

[6] National Research University Higher School of Economics. Institute for Statistical Studies and Economics of Knowledge. Head of Department for Strategic Foresight of the Centre for S&T, Innovation and Information Policy; E-mail: shashnov@hse.ru

[7] National Research University Higher School of Economics. Institute for Statistical Studies and Economics of Knowledge. Deputy Director of the Centre for Statistics and Monitoring of Information Society and Digital Economics; E-mail: asuslov@hse.ru

[8] The article is based on the study funded by the Basic Research Program of the National Research University Higher School of Economics (HSE) and by the Russian Academic Excellence Project '5-100'.

1. Statistics in the Modern Age: in Lieu of Introduction

The rapid development and wide proliferation of digital technologies have led to truly revolutionary changes in the economy and society. In particular, increased networking based on advanced platform solutions, automation, and the robotization of production processes contribute to the emergence of "unmanned", spatially distributed enterprises interacting with suppliers and consumers round the clock on a "paperless" basis.

The large-scale digitalization of various aspects of the economy and everyday life leads to radical changes in the nature and volumes of data on socioeconomic developments. Its application is changing, with demand going beyond the traditional scope. Data is generated essentially everywhere: by industrial installations, household appliances, spacecraft, organizations and households, and social networks. Its composition is becoming more varied and it is constantly updated. The range of data sources is also expanding, along with the generated data types including common quantitative and qualitative characteristics of various objects, processes or phenomena, and textual, audio-visual, and other "technological" formats. New data can supplement or refine the results of previous observations, or even radically change them.

Due to the growing rate and increasing complexity of global processes, the role of statistics has palpably changed. A report by the UK National Infrastructure Commission [NIC, 2017] recognizes the system of producing and distributing official statistical data as a new important component of national infrastructure, a public good which promotes the supply of efficient, competitive services and high-quality information products while minimizing the time needed and other production costs. Various groups of actors need objective, systemic, and up-to-date information about economic, social, demographic, environmental, and technological trends, while they are involved in the "production" of relevant data. Requirements concerning the relevance and reliability of statistical data are also becoming more stringent.

Input data requires timely and reliable verification, which implies using strict selection criteria to choose adequate methodologies and approaches for its application and to assess its suitability for decision-making purposes. National statistics systems across the world increasingly use big data and identify areas where the latter is likely to play a crucial role in the future (e.g. macroeconomic forecasting) [Eurostat, 2018a; Baldacci et al., 2016]. In 2014, the Statistics Division and UNECE conducted a survey of big data projects for the production of official statistics, in which the National Statistical Organizations indicated that the same data source could be used for different branches of statistics [Karpova et al 2016].

Statistics are only valuable when they are reliable, timely, and internationally compatible. The sooner the data becomes available, including in open access, the more the public knows about the social, economic, demographic, environmental, and other processes, the more objective this knowledge is [Statistics Denmark, 2018]. Therefore national statistical offices, the Russian one among them, have to meet a number of challenges, which implies the need for a structural and functional transformation, the active integration of statistical and administrative data and large data arrays from alternative sources, the application of cutting-edge digital technologies, and efficient strategic communications.

Responding to the challenges of the time and in line with the principles of impartiality and independence from political pressure, official statistics in developed countries has long since moved beyond the narrow scope of supporting public authorities (decision-making, public services, current administration, control, monitoring, etc.), and now successfully acts as a navigational aid in the ever-changing data environment. Businesses need it to design competitive strategies, make investment decisions, and conduct current operations. The public needs them to find their way in emerging economic and social circumstances, handle personal finances in a sensible way, obtain relevant information about education, healthcare, and the environment. Finally, civil society needs statistics to maintain social consensus and prevent tension, whereas the public authorities need to make reasonable decisions in the framework of evidence-based policies. For example, Data Strategy of Statistics Denmark for 2016-2020 [Statistics Denmark, 2016] sets two specific goals in the communications area: adapting official statistics to meet users' requirements; and playing an active role in public debates on relevant development-related issues.

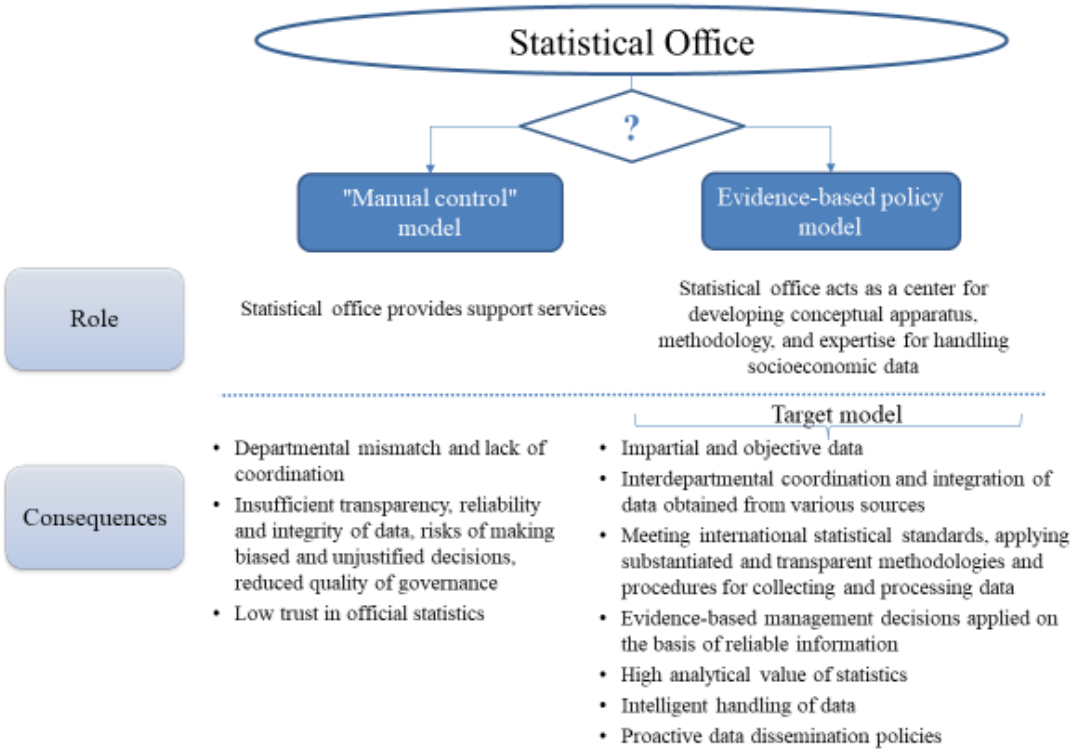
The Russian government's decision to establish the internationally competitive National Data Management System (NDMS) on the basis of the Federal State Statistics Service (Rosstat) [RF Government, 2018b] gave a major impetus for the development of a new statistics model in the country. It will serve as a framework for applying efficient tools to meet the information needs of economic agents, society, and the state as well as to promote and facilitate their active cooperation in data collection (and not simply aggregated data but primary data as well).

According to the "Digital Economy for the Russian Federation" program [Government commission, 2017; RF Government, 2018a], a digital analytical platform should serve as a basis for setting up and subsequently making use of such a system to meet the growing demand for data. It should have a flexible data access interface that is convenient for all user groups, ensure that all data are depersonified, and reduce the burden on respondents [Oksenoyt, 2018; Surinov, 2018].

While designing and implementing the NDMS, it is vitally important to understand and clearly describe the role of state statistics (and the statistical office), conceptualize the relevant model,

structure the approaches to advancing data collection and integration procedures in the context of digitalization, promote interdepartmental coordination of statistics-related activities, and apply new opportunities to analyze administrative, open-source, and big data. The choice of relevant options is largely determined by the dominant public administration model (Figure 1).

Figure 1. Statistical Office Models



As the many years of international practices show, a successful modern national statistical office operating in the scope of the evidence-based policy paradigm acts as an integrated national center for developing the conceptual apparatus, methodology, and expertise for handling socioeconomic data, regardless of its origins, be that statistical observation or alternative sources. In the model we see as the target one, statistics should ensure the following:

- Consistency, independence, and objectivity of produced data, and its relevant and timely provision;
- Integrity, compatibility, definitive interpretability of data obtained from various sources;
- Reliability of data, its high analytical potential, including for multidimensional integrated analysis and other research purposes;
- The indicator systems, classifications, methodological approaches, and toolsets applied must be open, and match international standards and best practices;
- Data collection and processing procedures must be cost-effective and transparent;

- Equal access to data, its timely provision, and adaptability to user needs.

The national statistical offices of the EU member states have interesting relevant experience: they take a proactive position in interacting with society, businesses, the state, and the global community. Their activities, coordinated by Eurostat, are focused on meeting national, interdepartmental, regional, and corporate data needs. Accomplishing its mission – the provision of high-quality statistical data – Eurostat promotes such social values as respect and trust, the dissemination of best practices, facilitating innovation, focusing on meeting user needs, and professional impartiality [Eurostat, 2018a].

Though new approaches to "producing" and distributing data are already being applied in Russian practices, a number of systemic issues persist which negatively affect the productivity of statistical work. In particular, these include lack of interdepartmental coordination of statistical activities, insufficiently transparent and inconsistent data supplied to various authorities, lack of uniform methodological supervision, lack of relevant, customized, detailed data for use by businesses (including small enterprises), R&D organizations, civil society, and various population groups. Primary focus on government agencies' "orders" as the underlying principle of planning statistical work also inhibits the development of statistical activities in Russia. Among these issues we also can emphasize limited access to primary data for analytical and management purposes and erratic work on its depersonification, inadequate research base for the statistical office, lack of competences to conduct analysis[9], insufficient use of digital technologies' potential (including platforms) and media resources for the dissemination of statistical outputs. One of more known reasons of retarding statistical development is low statistical literacy of the public, mass media, managers and lack of trust in statistical data, and dissatisfaction with it by a significant proportion of experts, officials, and the general public, among other things due to its inconsistency and the delayed provision of relevant data.

All of the above implies the need to implement a whole host of measures to create efficient mechanisms for interacting with users, improving strategic planning of statistical work, integrating data, restructuring data collection and processing practices, developing an advanced methodological basis, and designing intelligent data analysis and dissemination procedures.

[9] International experience shows that statistical services who have their own research and analytical units can meet a significant proportion of external users' demand for analysis. In Russia statistical output is limited to accounting tables built from various perspectives and a brief description of the main indicators.

2. Dialogue with Users

Goals and Principles

One of the most important features of a modern statistical office is a proactive position regarding users to better meet their data needs through an ongoing dialogue.

As was already noted, due to rapid changes in all areas of the society's life and activities, demand for all sorts of statistical data is constantly growing. Along with public authorities, it is driven by municipalities, businesses, R&D organizations, media, and last but not least, by individuals. Statistics are needed practically everywhere. This is why statistical data is positioned as a public good. In particular, this approach was reflected in the Quality Declaration of European Statistical System – a fundamental document which regulates the activities of European statistical offices. Part of them are related to interaction with users: focus user needs, requirements, and expectations; continuous improvement of work methods to take advantage of new opportunities; producing high-quality statistical information with objectivity and confidentiality; providing information on the main quality characteristics of each product; and accessibility of statistical results in a user-friendly format, enhancing user awareness of the strengths and limitations of the produced statistics, consulting users on how to use data [Eurostat, 2018a].

User needs are typically identified through specialized sample surveys. These are conducted in almost all countries and allow one to accumulate information about existing and potential consumers of statistical "products" and their requirements. All customer information is usually placed in appropriate databases, which allows for quickly responding to changing demand.

Customer Interaction Issues in Russian Statistics

Due to long-established traditions, the Russian statistical office is primarily focused on meeting the needs of the public authorities. Other user groups essentially remain unseen by providers of statistical data, so they have to "make do with what's available".

At the same time the statistical office is largely concerned with collecting and processing huge volumes of data to carry out one-off instructions, frequently within extremely short timeframes. This inevitably leads to deviating from the statistical "mainstream", which results in the reduced quality of information and duplication of results. Furthermore, a significant proportion of such data is never reused and is simply "binned".

According to businesses interactions, they remain unsatisfied the quality and "range" of publicly available aggregated statistics, among other things because the latter are not linked to specific areas,

sectors (industries), and products (services). Ordering specific detailed data (even if it can be produced at all, and disclosed, for particular small areas) involves significant financial costs. This issue is especially acute for small and medium companies and other organizations that specifically need detailed information about their home regions, which at best is available from territorial statistical offices and other agencies, but in most cases is not available at all. User requirements are not a priority for regional statistical departments, whose main objective is collecting data and maintaining centralized data flows. The population and media face a shortage of clear and meaningful figures.

Statistical data flows are organized in such a way that at the federal level they are disseminated in an aggregated form. This limits the opportunities for their secondary processing and the calculation of analytical and forecasting indicators needed by public authorities, experts, and businesses.

The current practice of providing depersonified data to users is mainly limited to the results of specific sample surveys of the population. Statistical techniques for ensuring confidentiality (such as micro-aggregation, etc.) are not applied. There is no infrastructure that would allow for the reuse of primary data for analytical and expert purposes (such as statistical research centers, networks providing remote access to depersonified microdata arrays, etc.). At the same time in international practice, possibilities of access to primary statistical information are widespread. For example, The Australian Bureau of Statistics' website has links to microdata on 40 subject areas provided by a table designer, fully depersonified microdata made available to users, and depersonified microdata which can only be used in a specially protected environment [Australian Bureau of Statistics, 2018].

Users often face problems with finding the information they need. Creating the Unified Interdepartmental Information and Statistics System (UIISS) based on organizing data by its "supplier" (as opposed to integrated subject areas) did not solve this problem. Meanwhile numerous national statistical offices have special websites where data is integrated "throughout" precisely on a thematic basis, which makes finding the right statistics and working with them much easier [Statistics Canada, 2018a; GOV.UK, 2018a].

Top-Priority Steps for the Customer Interactions Development

Launching a productive dialogue with all user groups should be the fundamental principle underlying the development of a new state statistics model in Russia. This is the only way to build an efficient system focused on user demand, extend the range of user-oriented products, and advance data culture. The statistical office needs an active marketing strategy which would not only provide for surveying statistical data users on a regular basis but also establish regular feedback channels.

A marketing strategy segmented by user groups, based on the above regular surveys of target groups to summarize their opinions about statistics as a whole, estimate demand for data, opportunities to collect it from various sources, and probable application areas should help one react to the changing demand in time, and optimize the statistical observation program to produce customized data.

The application of digital technologies for data dissemination purposes should provide rapid and convenient access and make it possible to reuse data. Setting up a unified Federal State Statistics Portal would be useful, based on a standardized conceptual apparatus, harmonized regulatory and reference information (RRI), a unified registry of indicators (with a publicly disclosed methodology for their calculation) and the integration of data, while at the same time adopting new content delivery formats (such as social media, mobile devices, etc.). Creating such a portal (which could act as a "single window" access point to the NDMS) implies a radical increase in the volume of published data, a shift to its full-fledged thematic structuring, and an extended range of data presentation formats in line with user needs to better meet their demand.

Some of the tried and true tools of this kind include the provision of data oriented towards specific user groups, market assessments, and so on [Australian Bureau of Statistics, 2018b; Statistics New Zealand, 2018]. Various interactive services and data visualization tools are being rapidly developed [Swiss Federal Statistical Office, 2018; UK Office for National Statistics, 2018; Ireland Central Statistical Office, 2018]. Infographics and videos which help users understand the information, graphically imagine the structure and dynamics of data and its interconnections command particular attention [US Census Bureau, 2010; Statistics Singapore, 2018; Statistics Canada, 2018b]. For example, Statistics Netherlands actively use iPhone applications, e-Book, YouTube, Twitter, Facebook, RSS channels among other formats in its operations and to promote its statistical products [Statistics Netherlands, 2018]. The program of digital transformation of the statistical office launched in the UK implies a radically new approach to official publications (reproducible analytical pipelines). Unlike the conventional organization of data access along the "storage – commercial software – spreadsheets – design – resulting textual document" sequence, the mass application of automated document generation tools in official information resources is envisaged, based on open programming languages and mark-up tools (such as R markdown and Python). This will reduce the time and costs of indicator development. Specialized software libraries offering researchers, analysts, and other interested users simplified access will also be created [GOV.UK, 2018b]. Rosstat also has similar experience, but it needs to be developed and stepped up.

The analytical value of the collected data is of significant importance to users. Adapting best international practices will help making progress in this area (see Section 4 for more).

To advance the statistical culture and increase trust in official information, statistical knowledge will need to be improved along with data skills, through educational programs of all levels. Setting up national information resources of the Statistics Explained type would also contribute to increasing the level of statistical literacy. Statistics Explained is an encyclopaedia/guidebook on the European statistics available on Eurostat's official website, offering information in a clear and simple form. It features a statistical glossary, numerous links to statistical data, metadata, and additional reference information. The portal is oriented not just towards professional users of statistics but also those who want to make one-off enquiries [Eurostat, 2018c].

The demand for information and an improved statistical culture form the foundation for building a brand new management system that would re-establish the national statistical office's reputation as an organization working for the public good, as opposed to being concerned solely with departmental objectives. This should be pursued by implementing digital technology-based, user-friendly data dissemination practices. The National Statistics System should increasingly interact with users. For this purpose, official statistics should become a real public good, providing impartial, reliable, timely, and transparent information to the public, state, and businesses. The statistical office have to maintain ongoing proactive dialogue with all user groups, taking into account their needs to the maximum possible extent. A strategy to improve the statistical literacy of the population, businesses, and managerial personnel in the digital economy and information society should be pursued. Efficient mechanisms for communicating social needs to the statistical office and for its interaction with businesses should be put in place, e.g. via the Federal State Statistics Portal and advanced infrastructure for the dissemination of statistical data and its provision. All this should increase trust by all user groups for official statistics [Surinov, 2018].

Thus, demand for information determines the range and content of products and services offered by the statistical office, affects all aspects of its operations including planning and standardizing observations, the production of outputs and development of tools for using them, and dissemination of results. Adhering to these standards and requirements should ultimately help produce "smart" statistics that are useful (sought after by and understandable) to all user groups. For this purpose, the data collection system should be modernized in order to provide alternative data sources meeting the requirements of state statistics and being integrated into an established system of collecting statistical data.

3. Diversification of Integration of Data Sources

A Departmental vs. Centralized Approach

State statistics in the Russian Federation is currently a hybrid system: in addition to Rosstat, more than 60 ministries, departments and organizations collect, process and provide official information in areas of their respective responsibilities. It is a fragmented structure with numerous shortcomings including the lack of uniform methodological principles and proper coordination of statistics collectors' activities, the lack of transparency, and the incompatibility and often inconsistency of data. Departmental statistical systems rely on a wide variety of technological and organizational solutions, which resulted in a kind of a "mixed structure" and disintegration of state statistics. This results in significant disparities between federal-, departmental-, and regional-level data, its reduced quality and analytical value, the duplication of statistical work, and an unnecessary increased burden upon respondents among other things due to the insufficient use of administrative data (available from the Federal Tax Service, the Pension Fund, Rosreestr, etc.). This is most painfully manifested in areas where the same "objects of accounting" act as respondents and are repeatedly surveyed by several different government departments.

Growing demand for statistics combined with the required data being unavailable from the specialized service together with the lack of access to primary data for analytical purposes, lead to numerous unauthorized initiatives by various departments to collect information. However, with very few exceptions, such initiatives do not match the principles and standards of the official statistical methodology. For example, in 2018 ten Russian national technology development programs essentially had no compatible (cross-cutting) statistical indicators for measuring R&D expenditures, commercialization of results, practical application of the technologies etc., which makes objective assessing the programs' implementation and their contribution to national development a rather challenging task. Another typical example is the "Digital Economy for the Russian Federation" program, along with its action plan. Only one of the 12 targets set in the "Information security" section can be assessed on the basis of the available statistics; for the "Information infrastructure" section it is 3 out of 17, and for the section "Developing research competences and obtaining groundwork results" - none out of eight. They mostly arise due to the need to measure progress in achieving various targets set for the departments. In the end a paradoxical situation arises: departments in effect decide the composition of the indicators themselves, develop algorithms for their calculation, collect data, and report that they have successfully achieved their targets, without using existing official statistical data.

Using various data sources, including primary and analytical accounting, administrative, financial and tax reports, and alternative resources is becoming the strategic direction of transforming data collection, processing, and analysis practices based on the principles of methodological and technological compatibility and the efficient use of the NDMS as an integrated information environment. It should be stressed that we are speaking about integrating data links, while in physical terms the system remains distributed. However, even that will help optimize statistical and administrative reporting, eliminate excessive forms thus reducing the load on respondents, and save a significant amount of public money. For example, in 2017, the Agency for Housing Mortgage Lending (AHML) initiated the development of the Unified Housing Construction Information System (UHCIS) which was expected to accumulate all data on the housing sector available in the country. The main objective was to increase market transparency and provide timely and reliable information to all interested parties – public authorities, developers, credit organizations, and individuals. At the initial stage, the UHCIS included aggregated data provided by Rosstat and the Central Bank as well as specific information about apartment buildings under construction collected by the Institute for Development of the Construction Industry. The system was being improved by integrating other sources of specific data, such as Rosreestr, Housing and Utilities Fund, and the Fund for Protection of Citizens – Co-investors in Housing Construction. This is expected to subsequently replace a significant portion of the statistical aggregates and expand the functionality of the UHCIS services, including through the flexible configuration of displayed information, giving the users information about the current situation at specific construction sites with map references, and so on.

A simple "mechanical" combination of various statistical data sources would be impossible due to the varying nature of the data they contain. A common basis is needed, which entails the unification and harmonization of definitions, classifications, algorithms for calculating indicators, and techniques for transferring and processing primary data. The very possibility of restoring a centralized statistical system in the current situation is further complicated by the prolonged and haphazard implementation of numerous different information systems by various ministries and departments. Their autonomy at the physical (due to multiple barriers preventing interagency data exchanges) and methodological level (inconsistent definitions, classifications, calculation algorithms, different data collection techniques and mismatched timing, different composition of the observed populations) not only frequently leads to the duplication of the requested information, but also its presentation in different formats and such a situation confuses respondents due to the need to calculate essentially the same parameters in different ways. Data discrepancies and errors inevitably lead to making less valid management decisions. On the contrary, the consistent implementation of the "single entry, repeated use of data" principle based on a unified methodology will significantly reduce the burden placed on respondents and improve the quality of statistical information.

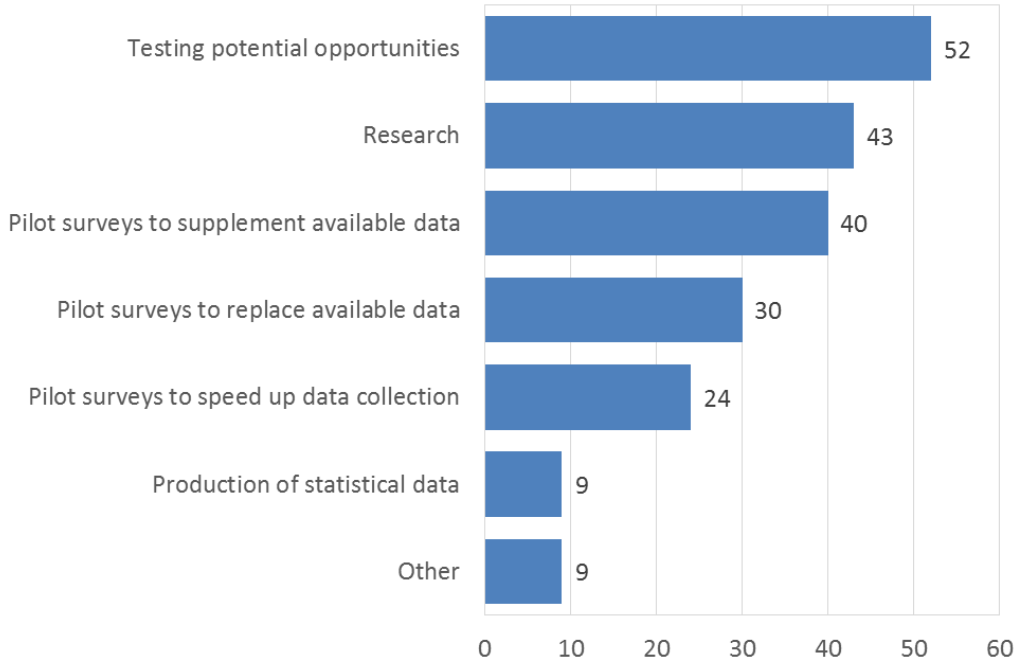
Alternative Statistical Data Sources

Recent practices, international [UN, 2014] and Russian [Oksenoyt, 2018; Plekhanov, 2017] alike, provide numerous examples of using alternative data sources to produce official statistics. Various kinds of such sources, conventional and recently emerged, can be integrated into the statistical system at different rates: information resources of executive authorities (such as state information systems (SIS), registers, cadastres, etc.) can be accessed more rapidly, while integrating data from the non-government sector (e.g. databases maintained by mobile operators, banks, commercial and non-profit organizations etc.) would take a longer amount of time. For example, in 2015, the Moscow City government launched a project to estimate the number of residents. The results of the official census were used along with administrative data available at various departments (the Federal Tax Service, Federal Mandatory Medical Insurance Fund, Pension Fund, Central Election Commission, Ministry of Education and Science, Federal State Statistical Service, Ministry of Internal Affairs), which keep records of certain groups of citizens for their own purposes. The application of departmental data using methodologies and techniques which ensured its comparability made it possible to adjust the final statistical estimates. The use of alternative data sources would help governments more accurately forecast profit tax budget revenues, expenditures on social support and infrastructure, and other scenarios.

The UN Statistics Division's Global Working Group (GWG) on Big Data for Official Statistics collects data on relevant projects implemented by national statistical offices and international organizations. The GWG database already contains information about more than 180 projects aggregating various alternative data sources in a number of areas [UN, 2018].

A significant proportion of these projects concern research aimed at studying specific data sources' potential in the production of official statistics. At the same time there is a clear and pronounced trend in the practical application of big data to supplement and improve consumer information obtained using conventional methods. Almost every third survey was carried out to supplement the existing statistics. The same number of projects were implemented to produce data or replace official statistical information (Figure 2). Moreover, some object-specific data can be successfully combined within the same statistical areas. For example, to calculate price statistics, cash register data (barcode scanning), data obtained from medical facilities, electricity meters, unstructured information posted on the web among others can be used. Big data allow for taking a fresh look at many traditional areas of activity (monitoring economic activity of population based on bank card transaction data, analysing tourist flows based on data from mobile operators, etc.) or monitoring completely new areas previously not covered by official statistics (like consumer behaviour on the Internet) [Plekhanov 2017].

Figure 2. Objectives of Big Data Statistical Projects: International Experience (number of projects implemented by national statistical offices and international organizations)



Source: authors based on [UN, 2018].

Using administrative and other alternative data sources to the greatest possible extent will eventually contribute to substituting numerous conventional statistical reporting formats with big data analytics. Achieving this would require taking at least the following steps:

- Adjusting the legal and regulatory basis (to facilitate interdepartmental cooperation, big data application, expanding Rosstat's responsibilities as the main state statistical office and the NDMS operator);
- Developing a standardized methodological foundation (conceptual apparatus, classifications, indicator systems, etc.);
- Advancing the technological infrastructure (platform solutions, data storage systems);
- Establishing access rules and procedures (free or paid access) for data generated by businesses and non-governmental organizations; determining funding sources for data collection and storage in case a paid access model is chosen.

Conventional statistical observations would survive in such a system only in areas where alternative data sources that do not subject respondents to an additional burden. Meanwhile all available data would be used for analytical purposes.

Unified Company Register

The full integration of statistical data will not be possible without a unified register of all organizations that are the objects of statistical monitoring. This register should be mandatory to use for data collection purposes by all government bodies and based on the Unified State Register of Legal Entities (USRLE), Unified State Register of Individual Entrepreneurs (USRIE), and Unified Register of Small and Medium Business (ERSMB). To provide a fuller picture, the register should reflect companies' activities at the regional level. This will allow for *putting together an object-specific database of federal statistical observations, adopting the model based on single reporting of information by respondents and its subsequent reuse*, ensuring the transparency and comparability (spatial and temporal) of data collected by various public authorities, and increasing the efficiency of metadata management in coordination with state information systems and other administrative data sources. Such a unified register of statistical monitoring objects in effect already exists. Its official status must be legalized and its use by all government agencies must be made compulsory.

Advanced digital technologies allow one to implement distributed models for statistical data collection, processing, storage, presentation, and dissemination rather quickly.

The register of statistical monitoring objects, combined with a unified register of indicators and reporting forms (see Section 4) will help coordinate all reporting parameters and forms. This, in turn, would allow one to avoid duplication, incompatibility, and unintentional human errors, control data reliability, use data from the aforementioned alternative sources in statistical analysis, and develop an integrated data collection interface (the "single window").

Unifying techniques and technologies for the collection and processing of statistical data implies collecting all indicators and report forms from all respondents exclusively in an electronic format (e.g. via virtual "private rooms"). Ultimately the move will be made from collecting report forms to the remote provision of primary data by respondents and analytical accounting, generating all output indicators in real time.

The system for collecting, storing, updating, searching, and processing information for each statistical observation unit based on the unified register also offers another advantage: data can be entered and processed regardless of which reporting form it belongs in. Unified description tools (metadata) will allow one to build the necessary analytical aggregates from indicators related at any level of the hierarchical structure using various criteria.

4. Intelligent Data Interpretation

Improved Planning and Optimization of Statistical Work

Moving on to the new stage of development of state statistics implies altering the principles statistical work is based on and improving the quality of its planning. Key trends in and the goals of the country's socioeconomic development should be used as the main reference points here, along with coordinating statistical activities of all relevant federal executive authorities.

Meanwhile, planning practices in statistics boil down to drafting annual federal statistical work plans covering, along with the segment for which Rosstat is directly responsible, dozens of observations conducted by various federal executive agencies and a number of other authorized organizations to support their current objectives without taking into account the emerging trends and long-term goals. Planning for a longer term, with the exception of recurring (with varying frequency) observations, is essentially not done. The multiplicity of organizations conducting statistical monitoring, the absence of strict centralized expert evaluation of methodologies applied by departmental statistical units and of the quality of data produced do not contribute to improving the situation either. The consequences of occasional large-scale statistical projects (such as industry-specific censuses, etc.) which force the statistical office to focus primarily on their implementation while the work in a number of basic and new statistical areas are funded out of whatever resources may remain, complete the picture.

The strategic planning of statistical work involves using anticipatory logic (with at least a three-year horizon) based on observed trends and long-term goals, with a focus on new and emerging areas (such as digital economy, sustainable development, climate change and its consequences, social sustainability, improved quality of life, etc.) and timely implementation of advanced methodological standards. The above should be achieved in parallel with stepping up the work aimed at meeting the information needs of users other than federal executive authorities, including businesses, R&D organizations, civil society, the public, and the media (see Section 2). Areas of statistics worthy of particular attention should be identified in close cooperation with such users, along with updating survey topics and relevant toolsets. Priority should be given to activities that can produce more diverse and higher-quality statistics with fewer resources.

Overcoming interdepartmental divergence and mismatches in the state statistics domain requires coordinating the methodology of statistical work including methodological supervision, the audit of the composition and content of relevant activities, the assessment of the quality of methodology, data collection and processing procedures, and final outputs. In the shortest possible time general stock-taking of statistical observations should be carried out along with an expert evaluation of proposals

to introduce new departmental statistical reports (especially one-time requests to provide new data) to assess the feasibility and quality of monitoring techniques and programs, taking into account the requirements of a unified register of indicators and reporting forms. The collection of departmental statistics primarily intended to produce progress reports on meeting established targets must be brought to a minimum. Independent external sources should be more actively used for these purposes.

Establishing a professional dialogue between Rosstat and federal executive agencies would help to more efficiently deal with issues such as excessive data collection and duplication, increase the reliability of indicators, and reduce the load on respondents.

One way to introduce new indicators is by calculating the costs and subsequently sharing them as done by the Statistical Office of Estonia, for example. A basic "price" is set for producing each new variable, which includes the development of a methodology (if necessary), data collection, processing, storage, and presentation. Interested users (i.e. government agencies) place orders for specific indicators and the costs are shared by all involved parties.

New opportunities for optimising the composition and content of statistical work also emerge due to the more active application of state information systems, specialized databases and other relevant data sources, including open ones. If currently data collection is typically preceded by the development of a methodology (identifying the object of observation, operationalizing the concepts, preparing the toolset, etc.), with the advent of integrated data based from the start on a unified (harmonized) RRI system and metadata, studying the data structure and interconnections and identifying errors and gaps would come to the forefront. This implies extracting data from various sources, matching it, verifying and processing it in order to subsequently design indicators in line with the identified user needs (requests). Reusing data will further reduce the load on respondents due to fewer one-time information requests.

When drafting statistical work plans, it is important to include steps to strengthen the research and analytical potential of the statistical office with the view to create tools and databases for more active use of all available data sources and multiple reuse of primary data or other statistical tools.

The Future-Oriented Development of Statistical Methodology

Methodology-wise, some areas of the Russian statistics significantly lag behind international standards in terms of the quality and rate of adapting new approaches and data collection, processing and presentation techniques. Despite the fact that Russian experts are represented in certain OECD and Eurostat working groups, Russian statistics on the whole is poorly involved in the development of prospective international statistical standards. Delays with adopting relevant international

standards, for example, regarding economic activity classifications (the adoption of the second edition of the All-Russian Classification of Economic Activity Types (OKVED) in 2017 and the relevant version of NACE Rev. 2 (2008) took more than seven years) result in having an outdated understanding of the present-day economy, which in turn increases the risk of making incorrect management decisions.

The issue of taking insufficient account of new developments and relevant structural shifts persists, which is fraught with a strategic failure to understand and systemically describe emerging global trends and their local effects. The lack of a unified harmonized national metadata system results in an unjustified increase of data collection costs, hinders optimizing the structure and content of statistical work, and the full-scale implementation of the forward planning model.

The previously noted inclination to adopt specific indicators in the interests of particular federal executive bodies leads to the reduced quality of the methodological processing of statistical observations and creates thematic gaps in the existing system. In particular, important statistical areas such as migration, labor and employment, jobs, investments, and the technological development of industries remain insufficiently developed in methodological terms. No harmonized approaches to the service sector, pre-school and additional education of children are applied.

In this context, the future-oriented development of statistical methodology becomes particularly relevant for anticipating major shifts in the economy and society. It requires harmonization with international standards regarding terms and definitions, local industry classifications, approaches to grouping organizations by size and activity sector, the calculation of macroindicators, and data presentation standards. The timeframes for updating major classifications and key methodologies should be synchronized to ensure timely measurements and international comparisons.

Several model examples of organizing statistical observation in the new and emerging technologies domain can be noted in Russia over the last decade. The regular measurement of science, technology, innovation, and production activities in nanotechnology has been conducted since 2009 [Gokhberg et al., 2011; Gokhberg et al., 2012]. In 2016-2017, a toolset for statistical monitoring of engineering and industrial design services was developed [HSE, 2016, 2017b], along with the one for technological and innovation activities in agriculture and related industries [HSE, 2017a]. The proposed approaches are based on the logic of technology life cycle analysis [Gokhberg et al., 2013], the use of flexible classifications and diverse data sources, and the consistent integration of key indicators into the existing national statistical observation system. This allowed for creating an information base for the economic assessment of the current state of particular high-technology sectors of the economy [Rusnano, Rosstat, HSE, 2017] and conduct international benchmarking

[OECD, 2015, 2017]. The above approaches were presented at meetings of OECD and Eurostat working groups and relevant international conferences and highly appreciated.

The need to review the understanding of objects of observation and move on from aggregating economic and statistical estimates to measuring development parameters and the quality of changes remains an important challenge.

For example, in Italy, the municipalities are one of the main actors conducting statistical observation [Morettini et al., 2013] covering a whole range of economic agents including their objectives, needs, values and motivation, which allows for producing integrated estimates of many different phenomena. Social statistics is focused not only on the integral characteristics of employment, unemployment, immigration, and so on, but also on unpaid work by members of certain social groups. This is relevant for analyzing regional development aspects which used to be in the "blind" zone [Sabbadini, 2011]. Flexible indicators adjustable to match various users' needs (such as expert community, local authorities, businesses, etc.) are coming to replace "neutral" statistics that describe general socioeconomic trends. The new indicators are also more sensitive to various characteristics of the measured objects, which allows for obtaining a more comprehensive picture of the ongoing social changes, cultural developments, and other transformations.

Data verification procedures require methodological and organizational upgrading. The techniques applied to calculate major indicators, the changes in methodologies, and the recalculation of data must be publicly presented and openly discussed. When new observations are introduced, control measurements and comparison with other sources must be made. The Swiss Federal Statistical Office's experience is worthy of attention, where the transition to using barcode scanning data from a new retail network to calculate the consumer price index for several months was accompanied by comparing it with conventional price statistics [Muller, 2010].

The development of techniques for generating representative samples, not only to improve the reliability and quality of output data but also to reduce the number of total surveys, remain dominant in Russian statistical practice, along with data collection costs, becomes particularly important for advancing statistical methodology and the organization of relevant activities. In this regard, the application of administrative data and big data plays an important role. In the latter case, despite the amount of data, its coverage can be limited or have significant systemic or random biases, which require developing complex semantic analysis algorithms to extract "useful" content. The need arises to introduce additional control diagnostic parameters, including those built upon the basis of conventional statistical indicators and accumulated information about the respondents' participation

in the surveys (paradata) and develop scientific methods for processing and maintaining time series (filling gaps, data imputation, etc.).

Strengthening the Research and Analytical Potential of the Statistical Office

Integrating statistical data obtained from various sources opens new opportunities for the integrated measurement of socioeconomic developments. For example, the training of professionals with critical skills was studied, covering a wide range of aspects from assessing young children's skills to collecting data on the number and structure of the researchers, combining information obtained via various surveys and tests [National Science Board, 2018].

The progress made in the international practice of providing access to depersonified microdata gives the leading statistical offices significant "intelligent inputs" allowing them to set up virtual research structures, among other things based on remote access using specialized software. This increases the analytical value of data and provides additional resources for developing new models and techniques for working with information. The key challenge here is opening access to statistical microdata for interested users.

40. In addition to the aspects already mentioned in Sections 2 and 3, here are several noteworthy examples of overcoming barriers hindering the use of primary data and disclosing data arrays for research purposes. Special primary data processing techniques are applied to accomplish this objective, which ensure that statistical properties remain unaffected without prejudice to confidentiality.

For example, the "Integrated Public Use Microdata Series provides a model for performing all sorts of statistical work [IPUMS, 2018]. Information about people's daily travel is used by the US Census Bureau to build generalized statistical models, which subsequently are made available to the public and are used to estimate the burden on transport infrastructure [Hayden, 2015].

Statistical microdata combined with administrative information is increasingly applied to directly assess the effectiveness of regulatory measures. A classic example of such analysis at the international level is the MicroBeRD project implemented by the OECD, focused on measuring tax breaks' impact on the amount of R&D expenditures [OECD, 2018]. Microdata allows one to employ the most advanced and precise econometric techniques to analyze factors affecting various policies' effectiveness, and in the international benchmarking context, to take into account country-specific characteristics to the greatest possible extent. Such projects have value not just as unique academic studies, but also as prospective areas for developing policy evaluation mechanisms.

The analytical function of the statistical office can also be strengthened by the timely publication of time series and data in a machine-readable format meeting international quality requirements, for subsequent processing and analysis. Specialized user applications for working with statistical data are based on this approach. For example, the US Census Bureau has developed an application for members of the general public choosing a new place of residence. The application identifies a selection of towns matching user-specified parameters. This changes the paradigm of using statistical observation results: statistics not only help justify a particular decision but offer a choice of various possible options.

5. Conclusions

This study outlines a modern model of the national statistical system. Its ultimate goals are organizing efficient management of various large-scale information flows and creating conditions for all interested actors (users in government agencies, businesses, R&D organizations, public associations, mass media, and the public) to find their way in the digital environment, understand the logic of working with data, and use it for practical purposes. An intelligent statistics system integrating big data should become the basis for the ongoing monitoring of the socioeconomic sphere, systemic research and forecasting, and determining the vector of the transformation of economy and society.

Creating this data management system is a complex and long process which is not limited to organizational and technological steps exclusively (though of course it requires applying cutting-edge technologies and the most advanced organizational and management approaches). New technologies make it possible to meet a wide range of information needs in a timely manner. The key issue is using these technologies in the right way. This is more of a methodological than "technical" issue, which take into account the following:

- The nature of changes taking place in the economy, society, and public administration, not just current transformations but also prospective ones;
- The hierarchy of goals and objectives for the socioeconomic actors;
- The particular features of and risks associated with using various data sources (which frequently compete with and occasionally contradict each other) for management (decision-making) purposes at all levels, from public administration to specific individuals.

Producing intelligent statistics has been globally considered as a way to improve conventional statistics, it has transformed into specific actions, including the reform of national statistical systems. On the basis of the available experience the following key aspects are important:

- Creating the National Data Management System, international integrated databases should not turn into a formal implementation of an "action plan". It requires a full-fledged cross-cutting project-based approach: each step (e.g. coordinating and integrating various data sources, automating primary reporting, organizing data exchanges, etc.) should yield specific, clearly defined results.
- It is important to preserve the "core" of the existing statistical system. New approaches and mechanisms should integrate and advance all of the previously achieved best results in methodology, observation areas, metrics, and other domains. As it is done in the Eurostat Development Program. The key aspect of the EU statistical office's development program is

maintaining its status as the principle supplier of relevant and reliable data to the European Commission, member states' national governments, and the public. The improvement areas include providing higher-quality information for policy shaping, businesses, individuals, and external partners. There are plans to extend data arrays and increase the number of available indicators, develop new data sources, and data flow management techniques [Eurostat, 2018a].

- The most important structural function that the new system should fully inherit should not be forgotten either. National statistics is expected to present an interconnected, objectively measurable model of socioeconomic processes and phenomena based on relevant theoretical concepts. In addition to using various sources of data, a necessary feature of the new system will be its reliance on a consistent conceptual framework and approaches to interpreting data, which will make it possible to integrate various data sources in the first place.

Increasingly complex methodologies and information bases imply that users of intelligent statistics themselves will change. They are becoming not only active participants in primary data collection, accumulation and application processes, but are turning into "smart" consumers who develop statistical thinking and are able to derive the greatest possible benefits from the use of statistical data. Such skills should become an inherent (and possibly mandatory) component of any skillset required in various spheres of activity.

References

Australian Bureau of Statistics (2018a) Available Microdata. http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Expected+and+available+Microdata (accessed on: 19.03.2018)

Australian Bureau of Statistics (2018b) Small business. http://www.abs.gov.au/websitedbs/d3310114.nsf/home/services+we+provide+to+small+business (accessed on: 19.03.2018).

Baldacci E., Buono D., Kapetanios G., Krische S., Marcellino M., Luigi Mazzi G., Papailias F. (2016) Big Data and Macroeconomic Nowcasting: from Data Access to Modelling. 2016 edition. Eurostat Statistical books. Luxembourg: Publications Office of the European Union.

Eurostat (2018a) Annual Work Programme 2018. Strategic Priorities. <http://ec.europa.eu/eurostat/documents/747709/753176/AWP2018_Strategic_priorities.pdf> (accessed on: 14.03.2018).

Eurostat (2018b) Quality Declaration of European Statistical System. http://ec.europa.eu/eurostat/documents/4187653/5752601/QUALDEC-EN.PDF/c9431f29-5fdd-4a8d-9179-bc9a054cc622 (accessed on: 14.03.2018).

Eurostat (2018c) Statistics Explained. http://ec.europa.eu/eurostat/statistics-explained/index.php/Main_Page (accessed on: 14.03.2018).

Gokhberg L., Fursov K., Karasev O. (2012) Nanotechnology development and regulatory framework: The case of Russia // Technovation. No. 32(3–4). P. 161–162.

Gokhberg L., Kuznetsova I., Dalin V., Fursov K. (2011). Nanotechnology statistics in Russia: formation of a new field. Voprosy Statistiki. 2011, 6: 3–19.

GOV.UK (2018a) Statistics. https://www.gov.uk/government/statistics (accessed on: 14.03.2018).

GOV.UK (2018b) Data in government. Blog of Civil Service. <https://dataingovernment.blog.gov.uk/2017/03/27/reproducible-analytical-pipeline/> (accessed on: 19.03.2018).

Government Commission (2017) Government Commission Protocol on the Use of Information Technologies to Improve the Quality of Life and Business Environment dated December 18, 2017 № 2.

Hayden E.C. (2015) Researchers wrestle with a privacy problem // Nature. Vol. 525. P. 440–442.

HSE (2016) Statistical Monitoring of the Engineering Services and Industrial Design Market / M. Gershman, L. Gokhberg, I. Kuznetsova et al. Moscow: National Research University Higher School of Economics.

HSE (2017b) Monitoring of Engineering and Industrial Design Market in Russia / M. Gershman, L. Gokhberg, I. Kuznetsova et al.; ed. by L. Gokhberg, V. Osmakov. Moscow: National Research University Higher School of Economics.

HSE (2017a) Guidelines for Collecting Statistical Data on Innovation in Agriculture and Related Sectors / L. Gokhberg, K. Ditkovskiy, A. Kadyrova et al. Moscow: National Research University Higher School of Economics.

IPUMS (2018) Integrated Public Use Microdata Series. http://www.ipums.org/ (accessed on: 19.03.2018).

Ireland Central Statistical Office (2018). CSO Data Visualization Tools. http://www.cso.ie/en/interactivezone/visualizationtools/ (accessed on: 14.03.2018).

Karpova N.S., Surinov A.E., Ulyanov I.S. (2016) Problems and possibilities for using Big Data in the Russian statistics // Voprosy statistiki. 2016; 7:3-9.

MacFeely S. (2016) The Continuing Evolution of Official Statistics: Some Challenges and Opportunities // Journal of Official Statistics, Vol. 32, No. 4, 2016, pp. 789–810, http://dx.doi.org/10.1515/JOS-2016-0041.

Morettini L. Perani G. Sirilli G. (2013) The Concentration of Knowledge Activities in Italy: An Analysis at Local Level //Foresight-Russia. V. 7. No 2. PP. 28–39.

Muller R. (2010) Scanner data in the Swiss CPI: An alternative to price collection in the field. Swiss Federal Statistical Office FSO. Room document 4/30/2010 for the UNECE/ILO meeting of experts on Consumer Price Indices, 10–12 May 2010. Geneva, Switzerland: Swiss Federal Statistical Office FSO.

National Science Board (2018) Science and Engineering Indicators 2018. Alexandria, VA: National Science Foundation.

NIC (2017) Data for the Public Good. National Infrastructure Commission report. https://www.nic.org.uk/publications/data-public-good/ (accessed on: 19.03.2018).

OECD (2015) The OECD Science, Technology and Industry Scoreboard 2015: Innovation for growth and society. Paris: OECD Publishing.

OECD (2017) The OECD Science, Technology and Industry Scoreboard 2017: The digital transformation. Paris: OECD.

OECD (2018) microBeRD. http://www.oecd.org/sti/microberd.htm (accessed on: 14.03.2018).

Oksenoyt G.K. (2018) Digital Agenda, Big Data and Official Statistics // Voprosy statistiki. 2018; 25(1): 3–16.

Plekhanov D.A. (2017) Data and Official Statistics: An Overview of International Practice in Using New Data Sources // Voprosy statistiki. 2017, 12: 49–60.

RF Government (2018a) "Digital Economy of the Russian Federation" Programme (approved by the Presidium of the Presidential Council for Strategic Development and National Projects of the Russian Federation, protocol dated December 24, 2018 № 16).

RF Government (2018b) List of Instructions Issued by the Prime Minister of the Russian Federation Following Sochi Investment Forum "Sochi-2018" dated March 1, 2018 № DM-P16-1131.

Rusnano, Rosstat, HSE (2017) Nanoindustry in the Russian Federation: statistical data book. 2011–2017. Moscow: Foundation for Infrastructure and Educational Programs of RUSNANO, Federal State Statistics Service (Rosstat), National Research University – Higher School of Economics.

Sabbadini L.L. (2011) The Development of Official Social Statistics in Italy with a Life Quality Approach // Social Indicators Research. No. 102(1). P. 39–46.

Statistics Canada (2018a) Statistics Canada. Statistics by subject. https://www.statcan.gc.ca/eng/subjects/ (accessed on: 17.03.2018).

Statistics Canada (2018b) Infographics. https://www.statcan.gc.ca/pub/11-627-m/index-eng.htm.

Statistics Denmark (2016) Communication and dissemination strategy. 2016–2020. Published by Statistics Denmark. September 2016.

Statistics Denmark (2018) Strategy and Quality of Statistics Denmark. <https://www.dst.dk/en/OmDS/strategi-og-kvalitet#> (accessed on: 15.03.2018).

Statistics Netherlands (2018). Follow us. https://www.cbs.nl/en-gb/follow-us (accessed on: 14.03.2018).

Statistics New Zealand (2018) Market Mapper. Find your target market. http://businesstoolbox.stats.govt.nz/MarketMapperMap.aspx?_ga=2.267912078.1484417099.1520070146-1924548221.1520070146/ (accessed on: 19.03.2018).

Statistics Singapore (2018) Infographics. Singapore population. https://www.singstat.gov.sg/statistics/visualising-data/infographics/population (accessed on: 14.03.2018).

Surinov A.E. (2018) Digital economy: challenges for Russian statistics: Report of Head of Federal State Statistics Service at the Extended meeting of Rosstat Board, 27.02.2018. <http://www.gks.ru/free_doc/new_site/rosstat/college/docl27022018.pdf> (accessed on: 16.03.2018).

Swiss Federal Statistical Office (2018) Our interactive and visual tools. https://www.bfs.admin.ch/bfs/en/home/services/interactive-visual-tools.html (accessed on: 14.03.2018).

UK Office for National Statistics (2018) People, Population and Community. https://visual.ons.gov.uk/category/people-population-and-community (accessed on: 14.03.2018).

UN (2014) Big data and modernisation of statistical systems. UN Statistical Commission. <http://unstats.un.org/unsd/statcom/doc14/2014-11BigData-E.pdf> (accessed on: 03.04.2018).

UN (2014) Fundamental Principles of Official Statistics. Resolution 68/261 adopted by the General Assembly on 29 January 2014 68/261. Doc. A/RES/68/261, 3 March, 2014. https:// unstats.un.org/unsd/dnss/gp/FP-New-E.pdf (accessed on: 15.06.2019).

UN (2018) Big Data Project Inventory. UN Global Working Group on Big Data for Official Statistics Inventory. <https://unstats.un.org/bigdata/ inventory> (accessed on: 19.03.2018).

US Census Bureau (2010) Census Home. https://www.census.gov/2010census/ (accessed on: 19.03.2018).

Leonid Gokhberg

National Research University Higher School of Economics (Moscow, Russia). Institute for Statistical Studies and Economics of Knowledge. Director;
E-mail: lgokhberg@hse.ru

Tatiana Kuznetsova

National Research University Higher School of Economics (Moscow, Russia). Institute for Statistical Studies and Economics of Knowledge. Centre for S&T, Innovation and Information Policy. Director;
E-mail: tkuznetzova@hse.ru

Gulnara Abdrakhmanova
National Research University Higher School of Economics (Moscow, Russia). Institute for Statistical Studies and Economics of Knowledge. Centre for Statistics and Monitoring of Information Society and Digital Economics. Director;
E-mail: gabdrakhmanova@hse.ru

Konstantin Fursov

National Research University Higher School of Economics (Moscow, Russia). Institute for Statistical Studies and Economics of Knowledge. Centre for Statistics and Monitoring of S&T and Innovation. Deputy Director;
E-mail: ksfursov@hse.ru

Elena Nechaeva

National Research University Higher School of Economics (Moscow, Russia). Institute for Statistical Studies and Economics of Knowledge. Centre for Socio-Economic Information Processing. Director;
E-mail: enechaeva@hse.ru

Sergey Shashnov

National Research University Higher School of Economics (Moscow, Russia). Institute for Statistical Studies and Economics of Knowledge. Department for Strategic Foresight of the Centre for S&T, Innovation and Information Policy. Head of;
E-mail: shashnov@hse.ru

Anton Suslov

National Research University Higher School of Economics (Moscow, Russia). Institute for Statistical Studies and Economics of Knowledge. Centre for Statistics and Monitoring of Information Society and Digital Economics. Deputy Director;
E-mail: asuslov@hse.ru

Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE.