

## МЕЖДУНАРОДНАЯ СТАТИСТИКА

### Будущее международного обмена статистическими данными и новые проблемы взаимодействия

Ольга Эмильевна Башина<sup>a</sup>,  
Наталья Алексеевна Комкова<sup>b</sup>,  
Лилия Валериевна Матраева<sup>c</sup>,  
Валерия Евгеньевна Косолапова<sup>c</sup>

<sup>a</sup> Московский гуманитарный университет, г. Москва, Россия,

<sup>b</sup> Федеральная служба государственной статистики (Росстат), г. Москва, Россия,

<sup>c</sup> Российский государственный социальный университет, г. Москва, Россия

*Статья посвящена проблемам и перспективам внедрения международного стандарта обмена данными и метаданными (SDMX) и осуществления международного обмена статистическими данными и метаданными с применением данного стандарта. Определены возможные области использования стандарта, описан механизм реализации обмена данными и метаданными в соответствии со стандартом SDMX. На основе анализа отечественного и зарубежного опыта внедрения и использования стандарта выделены основные проблемы, которые могут быть классифицированы на три группы: общие, статистические, информационно-технологические. При этом они могут возникнуть на национальном уровне (при внедрении стандарта внутри страны), на международном уровне (при применении стандарта внутри международных организаций) и на национально-международном уровне (при осуществлении информационного взаимодействия между страновыми поставщиками статистических данных и международными организациями).*

*Общие проблемы возникают на нормативно-правовом уровне и связаны с установлением границ ответственности организаций-контрагентов на всех трёх уровнях взаимодействия, а также в части наращивания потенциала применения стандарта SDMX. Проблемы чисто статистического характера встречаются наиболее часто по причине осуществления обмена большим количеством данных и метаданных, относящихся к различным тематическим областям статистики, структура формирования и передачи которых должна быть унифицирована.*

*С развитием информационного обмена возникают проблемы постоянного мониторинга и расширения списков кодов, используемых в стандарте SDMX; при этом отмечается отсутствие универсальной структуры данных на международном уровне и, как следствие, сложность понимания и применения на национальном уровне существующих структур данных, разработанных международными организациями. Информационно-технологические вызовы связаны с построением ИТ-инфраструктуры для обмена данными и метаданными с использованием стандарта SDMX. ИТ-инфраструктура включает следующие элементы в зависимости от статуса участника процесса: инструменты, необходимые для принимающих организаций, инструменты для предоставляющих данные организаций и инфраструктура для ИТ-специалистов.*

*По каждой из сформулированных проблем обоснованы авторские практические рекомендации на основе принципа комплексности применительно к внедрению международного стандарта обмена данными и метаданными SDMX.*

*Ключевые слова:* международный стандарт обмена информацией, обмен данными и метаданными, SDMX, информационная модель, международное информационное взаимодействие, международная статистика, описание структуры данных (DSD), описание структуры метаданных (MSD).

*JEL:* C80, D80, M15, O19.

*doi:* <https://doi.org/10.34023/2313-6383-2019-26-7-55-66>.

*Для цитирования:* Башина О.Э., Комкова Н.А., Матраева Л.В., Косолапова В.Е. Новые проблемы взаимодействия и будущее международного обмена данными. Вопросы статистики. 2019;26(7):55-66.

## The Future of International Statistical Data Sharing and New Issues of Interaction

Olga E. Bashina<sup>a</sup>,  
Natal'ya A. Komkova<sup>b</sup>,  
Liliia V. Matraeva<sup>c</sup>,  
Valeriya E. Kosolapova<sup>c</sup>

<sup>a</sup> Moscow University for the Humanities, Moscow, Russia,

<sup>b</sup> Federal State Statistics Service (Rosstat), Moscow, Russia,

<sup>c</sup> Russian State Social University, Moscow, Russia

*The article deals with challenges and prospects of implementation of the Statistical Data and Metadata eXchange (SDMX) standard and using it in the international sharing of statistical data and metadata. The authors identified potential areas where this standard can be used, described a mechanism for data and metadata sharing according to SDMX standard. Major issues classified into three groups - general, statistical, information technology - were outlined by applying both domestic and foreign experience of implementation of the standard. These issues may arise at the national level (if the standard is implemented domestically), at the international level (when the standard is applied by international organizations), and at the national-international level (if the information is exchanged between national statistical data providers and international organizations).*

*General issues arise at the regulatory level and are associated with establishing boundaries of responsibility of counterpart organizations at all three levels of interaction, as well as in terms of increasing the capacity to apply the SDMX standard. Issues of statistical nature are most often encountered due to the sharing of large amounts of data and metadata related to various thematic areas of statistics; there should be a unified structure of data and metadata generation and transmission.*

*With the development of information sharing, arise challenges and issues associated with continuous monitoring and expanding SDMX code lists. At the same time, there is a lack of a universal data structure at the international level and, as a result, it is difficult to understand and apply at the national level the existing data structures developed by international organizations. Challenges of information technology are related to creating an IT infrastructure for data and metadata sharing using the SDMX standard. The IT infrastructure (depending on the participant status) includes the following elements: tools for the receiving organizations, tools for sending organization and the infrastructure for the IT professionals.*

*For each of the outlined issues, the authors formulated some practical recommendations based on the complexity principle as applied to the implementation of the international SDMX standard for the exchange of data and metadata.*

*Keywords:* international standard for information exchange, data and metadata sharing, SDMX, information model, international information interaction, international statistics, Data Structure Definition (DSD), Metadata Structure Definitions (MSD).

*JEL:* C80, D80, M15, O19.

*doi:* <https://doi.org/10.34023/2313-6383-2019-26-7-55-66>.

*For citation:* Bashina O.E., Komkova N.A., Matraeva L.V., Kosolapova V.E. The Future of International Statistical Data Sharing and New Issues of Interaction. *Voprosy Statistiki*. 2019;26(7):55-66. (In Russ.)

### Цифровая трансформация межгосударственного взаимодействия: предпосылки создания международного стандарта в области передачи данных и метаданных

Трансформация вектора мирового экономического развития в сторону цифровых технологий делает уровень их развития одним из главных факторов конкурентоспособности отдельных стран. Как отмечается в исследовании Глобального ин-

ститута McKinsey: «...процесс развития цифровой экономики по масштабам будет сопоставим с промышленной революцией XVIII-XIX вв., которая радикально изменила весь мир, дав многим странам толчок к стремительному росту, изменив саму парадигму развития» [1, с. 44; 2].

В последние годы развитие цифровой экономики в России стало одной из приоритетных правительственных задач, в рамках которой был реализован ряд инициатив на субнациональном и национальном уровнях. Так, в июле 2017 г. Рос-

сийская Федерация приняла Программу цифровой экономики России с ожидаемым ежегодным бюджетом в размере 1,8 млрд долларов США с целью до 2025 г. устранения текущих недостатков, препятствующих присоединению страны к мировым лидерам цифровой экономики [3].

Современная стадия цифровой трансформации характеризуется усилением процессов взаимодействия, интеграции и функционирования национальных цифровых экосистем, сопровождающихся активным созданием единого информационного пространства, под которым понимается не только связка программных продуктов, сочетающих различные методы и инструменты управления данными, но и использование единых машиночитаемых форматов обмена данными. Все это послужило предпосылками создания на международном уровне стандарта в области передачи данных и метаданных (далее по тексту SDMX). Несмотря на то, что разработка стандарта началась еще в 2001 г., в настоящий момент продолжается обсуждение международным сообществом возможностей расширения его функционала, совершенствование схем данных и глобализация формата, что свидетельствует об актуальности его изучения.

Международный стандарт обмена данными и метаданными (SDMX) - это утвержденный стандарт ISO, призванный обеспечить эффективный обмен статистическими данными и метаданными и способствующий их распространению среди национальных и международных статистических организаций. Он обеспечивает комплексный подход для упрощения обмена статистическими данными и метаданными, что позволяет реализовать взаимодействие внутри и между системами, участвующими в обмене, представлении и распространении статистических данных и связанных с ними метаданных. При этом стандартизация достигается за счет использования утвержденных списков кодов и описаний структур статистических данных и метаданных. Иными словами, это стандарт распространения данных среди статистического сообщества и пользователей статистической информации, разработанный на стыке статистической науки и информационных технологий. Комплекс документов, описывающих SDMX, можно разделить на три блока:

- технические стандарты для данных и соответствующих метаданных;
- статистические руководства, ориентированные на раскрытие содержимого;

- рекомендации ИТ-архитектуры для эффективного обмена и использования статистических данных и метаданных.

Таким образом, SDMX представляет собой совокупность технических стандартов, статистических руководств и ИТ-архитектуры, которая может содействовать улучшению бизнес-процессов для любой статистической организации, а также гармонизации и унификации статистических метаданных.

Стоит также отметить, что использование стандарта SDMX для передачи данных в международные организации - это лишь один из возможных способов применения стандарта. Он также может использоваться для автоматического обмена данными «машина к машине», для визуализации данных с помощью использования графических интерфейсов, для повышения эффективности управления базами данных, а также для сбора данных и обмена ими в рамках внутриведомственного и межведомственного информационного обмена.

Большие объемы информации, нужной для передачи, неоднородность природы передаваемых данных, необходимость унификации структуры передаваемых данных, отсутствие достаточного нормативно-правового обеспечения - все это порождает множество проблем для применения существующего стандарта.

Таким образом, целью статьи является изучение основных проблем, возникающих при внедрении международного стандарта обмена данными и метаданными SDMX и при осуществлении обмена данными с использованием указанного стандарта, а также путей их решения.

Для достижения этой цели в статье приводится описание причин и анализ природы возникающих сложностей, а также предлагаются возможные способы их решения.

### **Стандарт SDMX - новые возможности комплексной трансформации**

Международный стандарт обмена данными и метаданными SDMX представляет собой систему эффективного обмена статистической информацией с использованием структуры, основанной на общих статистических стандартах и рекомендациях, а также ИТ-архитектуры и ИТ-инструментов, что делает стандарт SDMX междисциплинарным [4].

Таким образом, эффективность применения стандарта зависит от соблюдения требований статистического характера и реализации их через ИТ-инфраструктуру. Отсюда следуют и возможные причины возникновения сложностей в применении данного комплексного стандарта.

Стандарт предусматривает передачу как данных, так и метаданных. При этом в рамках стандарта выделяются два вида метаданных: структурные и справочные.

Важно подчеркнуть, что данные и метаданные, подлежащие обмену, тесно взаимосвязаны. *Структурные метаданные* выступают идентификаторами и дескрипторами данных (например, переменные, кодовые списки, наборы данных) и необходимы для определения данных. Иными словами, значение 110226326 само по себе не будет нести информационной нагрузки, если не снабдить его надлежащим набором описаний и характеристик (метаданными). А вот если сказать, что в 2016 г. *численность населения трудоспособного возраста составила 110226326 человек*, то уже становится понятным, к чему относится данное число. Таким образом, наименование динамического ряда, временной интервал и единица измерения по отношению к данному значению являются структурными метаданными

*Справочные метаданные* включают используемые понятия и их практическую реализацию («концептуальные» метаданные), описывают методы, используемые для производства статистических данных («методологические» метаданные), различные измерения качества полученных статистических данных («качественные» метаданные). К примеру, справочные метаданные будут содержать информацию о том, что указанный показатель формируется по результатам обследования рабочей силы, которое проводится ежегодно и охватывает все население в возрасте 15-72 лет [5].

Для того чтобы обеспечить эффективный обмен данными и метаданными с применением стандарта SDMX и использовать его преимущества стандартизации, необходимо убедиться в том, что передаваемые пакеты данных и метаданных построены на основе руководящих принципов по конкретным областям и в соответствии с определениями структуры данных и метаданных с использованием заранее определенных перечней кодов для отдельных отраслей статистики [6].

Передача данных должна осуществляться с учетом заданной концептуальной схемы, и, в первую очередь с учетом заданного описания структуры данных (DSD), которое включает три основных компонента:

1) измерения (dimensions), описывающие статистические данные и формирующие их идентификатор (ключ) к соответствующим данным (например, время, область, пол);

2) значения (measures), которые отражают значение статистической величины за период времени;

3) атрибуты (attributes), которые отражают дополнительную информацию о какой-либо части рассматриваемого набора данных и могут характеризовать набор данных, наблюдение или группу измерений [7].

Наряду с концептуальной схемой передачи информации необходимо указание адресных данных: информацию о передающей и принимающей организации, контактные лица и иная информация, которая будет характеризовать созданный с применением стандарта поток данных или метаданных. Все это в совокупности образует файл, содержащий статистические данные, представленные с использованием утвержденных списков кодов и описаний структур данных, который может быть представлен в одном из следующих форматов: SDMX-ML или SDMX-EDI. Пример подготовленного файла в формате SDMX-ML представлен на рис. 1.

Важно отметить, что данные считаются переданными в соответствии со стандартом SDMX только тогда, когда они прошли валидацию и смогли быть должным образом прочитаны принимающей стороной. При этом отсутствует промежуточный результат, то есть при допущении ошибки, например, на стадии формирования структуры данных, набор данных будет не опознан частично или полностью.

В рамках применения стандарта предполагается осуществлять распространение данных основным тематическим областям статистики (statistical domains): демографическая и социальная статистика, экономическая статистика, экологическая и многоотраслевая статистика.

По направлениям взаимодействия процесс обмена данными с использованием стандарта можно условно разделить на *три* уровня в зависимости от уровня взаимодействия участников:

- взаимодействие на национальном уровне (при внедрении стандарта внутри страны и его

```

<?xml version="1.0" encoding="utf-8"?>
<!--NSI Web Service v6.13.1-->
<message:StructureSpecificData xmlns:message="http://www.sdmx.org/resources/sdmxml/schemas/v2_1/message"
xmlns:ss="http://www.sdmx.org/resources/sdmxml/schemas/v2_1/data/structurespecific"
xmlns:footer="http://www.sdmx.org/resources/sdmxml/schemas/v2_1/message/footer" xmlns:ns1="urn:sdmx.org.sdmx.infomodel.datastructure.
DataStructure=UNSD:SDG(0.4):ObsLevelDim:TIME_PERIOD"
xmlns:common="http://www.sdmx.org/resources/sdmxml/schemas/v2_1/common" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <message:Header>
    <message:ID>SDG_RU1_ILO</message:ID>
    <message:Test>false</message:Test>
    <message:Prepared>2018-10-02T12:33:03</message:Prepared>
    <message:Sender id="RU1">
      <common:Name xml:lang="en">Ivanov Ivan</common:Name>
      <message:Contact>
        <common:Name xml:lang="en">Natalia Komkova</common:Name>
        <message:Department xml:lang="en">SDG monitoring</message:Department>
        <message:Role xml:lang="en">Advisor</message:Role>
        <message:Email>IvanovIvan@ks.ru</message:Email>
      </message:Contact>
    </message:Sender>
    <message:Receiver id="UNSD">
      <common:Name xml:lang="en">Abdulla Gozalov</common:Name>
      <message:Contact>
        <common:Name xml:lang="en">Abdulla Gozalov</common:Name>
        <message:Email>gozalov@un.org</message:Email>
      </message:Contact>
    </message:Receiver>
    <common:Name xml:lang="en">Ivanov Ivan</common:Name>
    <message:Structure structureID="UNSD_SDG_0_4" namespace="urn:sdmx.org.sdmx.infomodel.datastructure.DataStructure=UNSD:SDG(0.4):ObsLevelDim:TIME_PERIOD"
dimensionAtObservation="TIME_PERIOD">
      <common:Structure>
        <Ref agencyID="UNSD" id="SDG" version="0.4" />
      </common:Structure>
    </message:Structure>
    <message:DataSetAction>Information</message:DataSetAction>
    <message:DataSetID>SDG_RU1_ILO_RU1_2018-10-02T123303</message:DataSetID>
    <message:Extracted>2018-10-02T12:33:03</message:Extracted>
  </message:Header>
  <message:DataSet action="Information" ss:dataScope="DataStructure" xsi:type="ns1:DataSetType" ss:structureRef="UNSD_SDG_0_4">
    <Series FREQ="A" REPORTING_TYPE="N" SERIES="SI_POV_DAY1" REF_AREA="RU" SEX="_T" AGE="_T" URBANISATION="_T" INCOME_WEALTH_QUANTILE="_T" EDUCATION_LEV="_T" OCCUPATION="_T" CUST_BREAKDOWN="_T" COMPOSITE_BREAKDOWN="_T" DISABILITY_STATUS="_T">
      <Obs TIME_PERIOD="2017" OBS_VALUE="48.5" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
    </Series>
    <Series FREQ="A" REPORTING_TYPE="N" SERIES="SI_POV_DAY1" REF_AREA="RU" SEX="_T" AGE="_T" URBANISATION="_T" INCOME_WEALTH_QUANTILE="_T" EDUCATION_LEV="_T" OCCUPATION="_T" CUST_BREAKDOWN="_T" COMPOSITE_BREAKDOWN="_T" DISABILITY_STATUS="_T">
      <Obs TIME_PERIOD="2010" OBS_VALUE="0" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
      <Obs TIME_PERIOD="2011" OBS_VALUE="0" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
      <Obs TIME_PERIOD="2012" OBS_VALUE="0" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
      <Obs TIME_PERIOD="2013" OBS_VALUE="0" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
      <Obs TIME_PERIOD="2014" OBS_VALUE="0" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
      <Obs TIME_PERIOD="2015" OBS_VALUE="0" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
      <Obs TIME_PERIOD="2016" OBS_VALUE="0" OBS_STATUS="A" UNIT_MULT="0" UNIT_MEASURE="PERCENT" />
    </Series>
  </message:DataSet>
</message:StructureSpecificData>

```

Рис. 1. Подготовленный для передачи SDXM-ML файл

Источник: пример авторов.

использовании для внутриведомственного или межведомственного информационного статистического обмена);

- взаимодействие на международном уровне (при применении стандарта внутри международных организаций);

- взаимодействие на национально-международном уровне (например, при осуществлении информационного взаимодействия между странами поставщиками статистических данных и международными организациями) [8].

При этом национальные статистические системы с применением формата SDMX могут сочетать в себе следующие функции:

1) построение национальной статистической базы данных;

2) интеграция статистических данных;

3) передача данных и обмен данными;

4) управление статистическими данными и метаданными;

5) поддержка статистических данных для анализа и др.

Процесс передачи данных с применением стандарта SDMX можно проиллюстрировать схемой, представленной на рисунке 2.

Основная задача стандарта SDMX - это стандартизация и упрощение процесса обмена статистической информацией за счет использования унифицированных структур данных всех тематических областей (доменов) статистики и соблюдение этой стандартизации вовлеченными в процесс информационного обмена сторонами,



Рис. 2. Механизм организации передачи данных в SDMX

Источник: [8].

в том числе национальными поставщиками официальной статистики.

Процесс внедрения стандарта в статистические системы зачастую требует немало временных и трудовых затрат, а также проведения технической модернизации, а в ряде случаев и вовсе реформации внутренних статистических процессов и используемых в процессе информационного обмена технических средств. Все это порождает проблемы реализации стандарта SDMX [9].

### Вызовы на современном этапе развития и способы решения

Увеличение объемов передаваемой информации, рост вовлеченных в процесс контрагентов, изменения в структуре передаваемых данных влияют на требования к стандарту, заставляя его расширяться и отвечать вызовам со стороны статистики и информационных технологий. Эти вызовы условно можно разделить на три группы: общие, статистические и технические.

*Общие вызовы*, в первую очередь, касаются нормативно-правового закрепления ответственности по осуществлению информационного обмена за конкретными участниками процесса. Они представляют собой сферу выявления ответственности участников процесса и создания нормативно-правовой базы, что формирует условия эффективного осуществления инфор-

мационного обмена. В настоящее время имеют место случаи дублирования потоков данных на уровне взаимодействия национальных участников с международными контрагентами, когда, например, ряд показателей одновременно направляется сразу в несколько международных организаций. При этом, как отмечалось ранее, значения одинаковых показателей в различных базах данных могут не совпадать между собой, например из-за проводимых принимающей стороной корректировок для обеспечения сопоставления страновых данных, полученных из разных источников или имеющих различия в методологии расчета.

В настоящий момент осуществление обмена данными с применением стандарта SDMX возможно только в формате «push», то есть путем формирования файла в формате SDMX и направления его по электронной почте в адрес контактных лиц.

Однако более эффективным способом информационного взаимодействия является использование схемы «pull», то есть размещения данных на информационной платформе (или в ином специализированном внешнем хранилище) с возможностью доступа к ним международных организаций. Схема «pull» уменьшает трудозатраты и сокращает время, затрачиваемое на информационное взаимодействие с широким кругом различных организаций, а также позволя-

ет размещать данные по показателям на единой платформе. Получатели данных могут иметь ограниченные права доступа к внешнему хранилищу, например подписку на определенные данные, что в настоящий момент также пока не реализовано. Еще одной задачей остается также закрепление ответственности международных организаций за определенные тематические группы показателей, что может частично идти в разрез со сложившейся схемой передачи данных.

Общей проблемой для всех участников информационного обмена с применением стандарта SDMX является наращивание потенциала человеческих ресурсов. Необходимым условием для правильного использования стандарта является обучение сотрудников национальных статистических служб и других поставщиков официальной статистики использованию стандарта по осуществлению информационного обмена и его применению.

При этом для успешной реализации внедрения SDMX необходимо вовлечение и тесное сотрудничество различных субъектов информационного обмена, в которые входят: статистики (способные объяснить предметные и контентные статистические вопросы), IT-специалисты (способные поддерживать ПО, отвечать современным

техническим вызовам), ассистенты маппирования (отвечающие за соблюдение требований к созданию схем данных), координаторы (отвечающие за соблюдение нормативных требований к передаче данных в формате SDMX и осуществление информационного обмена). Наращивание потенциала в части применения стандарта SDMX предусматривает необходимость постоянной актуализации знаний и компетенций участников процесса информационного обмена, в том числе за счет проведения различных тренингов для национальных статистических служб и заинтересованных агентств, предоставления учебных материалов и информационно-технического обеспечения.

*Вызовы статистического характера* являются наиболее емкими, так как именно в этой области формируются требования к структуре данных, а, следовательно, и все атрибуты для структуры. Здесь среди направлений развития стандарта SDMX можно выделить его использование для обмена многомерными статистическими данными, микроданными и макроданными, данными на основе регистров и геопространственными данными [10].

SDMX предусматривает передачу данных в виде «срезов информации статистических кубов»

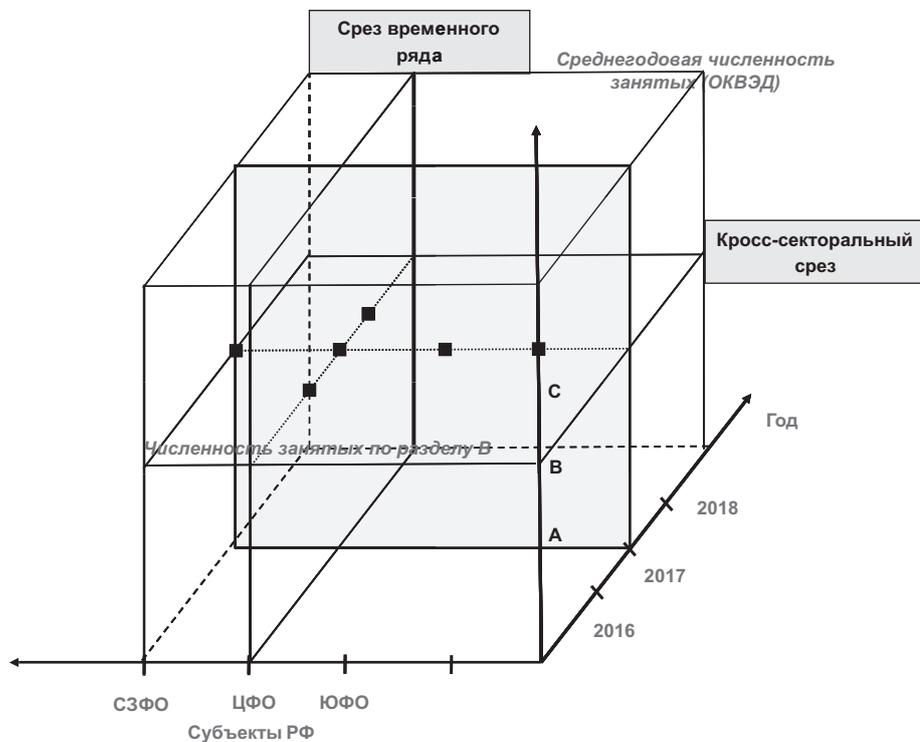


Рис. 3. Статистические данные в представлении RDF куба данных

Источник: формализовано авторами.

[11, 12], которые формируются под конкретную ситуацию или запрос. К примеру, данные могут направляться по одному виду экономической деятельности для различных периодов или, наоборот, для одного временного периода, но по нескольким видам деятельности.

На международном уровне остается нерешенным ряд вопросов передачи данных на основе временных рядов или табличных данных с использованием многомерной модели. Определение ключевых идентификаторов набора данных, таких, как различные концептуальные измерения (dimensions) и атрибуты (attributes), зачастую оказывается неочевидным (например, в ряде случаев единицы измерения сложно однозначно отнести к измерениям или к атрибутам). Помимо этого, трудности также возникают на этапе определения подхода к измерению показателей, методологии их разработки, которые тоже могут играть роль при определении объекта и характеристик показателя.

В этом вопросе при передаче данных на международный уровень национальным службам необходимо строго следовать требованиям к потокам данных, разработанным международными организациями, и согласовывать структуру передаваемых данных заранее. При передаче данных на национальном уровне решить этот вопрос несколько сложнее, так как отсутствует стандартная практика, поэтому целесообразно исходить из национальных потребностей и особенностей при определении национальной структуры данных.

При осуществлении маппирования зачастую сложно определить, к чему относится та или иная характеристика показателя (например, к измерениям или к атрибутам). Иногда необходимые для отражения в структуре данных измерения и атрибуты не очевидны и варьируют от организации к организации. Указанная проблема может иметь два варианта решения, в разной степени фиксирующих требования к структуре передачи данных (DSD):

- определить глобальную модель структуры передачи данных (DSD) (отдельные концептуальные схемы для каждого домена с поддержкой повторного использования понятий и списков кодов), которая будет способна раз и навсегда определить требования к кодированию набора данных для передачи в формате SDMX;

- использовать подход смешанных концепций, где структура передачи данных (DSD) будет осно-

вываться на нескольких концептуальных схемах - одной центральной и одной локальной. Такая структура данных позволит предъявлять более гибкие требования к кодированию набора данных для передачи в формате SDMX, но не решит проблему однозначности трактовки DSD [13].

Сложным процессом является также актуализация списков кодов, используемых в структуре данных. По сути, эти списки должны охватывать все возможные комбинации переменных. Так, например, некоторые показатели могут передаваться в региональной разрезности, однако в настоящий момент общий список внутривосточных региональных кодов отсутствует в силу ряда причин. Во-первых, такой список кодов представлял бы собой перечень огромных размеров и требовал бы постоянной поддержки и обновления (гораздо более частой, чем любые другие списки кодов). А во-вторых, он может различаться у отдельных стран и у международных организаций в силу политических причин и вызывать споры и разногласия.

По причине спорных территорий, роста числа участников обмена данными, подвижности территориальных единиц внутри стран встает вопрос о механизме обновления такого списка после его создания. В данном случае возможна ситуация, когда ответственными за ведение реестра территориальных единиц для SDMX являются сами страны-участницы, осуществляя при этом его актуализацию при наличии изменений.

Для реализации данного процесса необходимо облегчить повторное использование глобальных списков кодов, разрешить их расширение с помощью местных/дополнительных кодов, рассмотреть возможность внедрения пустого списка кодов (справочный список кодов, который, в свою очередь, ссылается на разные «родительские» списки кодов, из которых наследуются его коды).

На национальном уровне сложности возникают в понимании структуры данных и адаптации национальных данных под требования этой структуры. В первую очередь, для маппирования показателей со структурой данных необходимо приведение национальных данных под международные требования. Это очень трудоемкий процесс, осложняющийся наличием разных структур данных в различных международных организациях. Для упрощения реализации процесса необходимо четкое определение соответствия кодов для передачи данных в соответствии со стандартом SDMX и имеющихся кодов в базах данных.

Помимо этого, практически нереализованным остается обмен метаданными с применением стандарта SDMX. Для эффективного внедрения обмена справочными метаданными требуется регламентация передачи метаданных в рамках международных организаций - глобальная схема передачи метаданных, а также создание глобального определения структуры метаданных, которая, наряду со схемой передачи метаданных, позволит решить проблемы маппирования и транскодирования.

*Вызовы информационно-технического характера* связаны с построением ИТ инфраструктуры.

Информационно-техническое обеспечение внедрения стандарта SDMX заключается, в первую очередь, в создании справочной инфраструктуры SDMX (SDMX-RI), набора программных средств с открытым исходным кодом, которые позволяют организациям быстро внедрять SDMX и обмениваться статистическими данными с использованием стандарта [10].

Инструменты, поддерживающие обмены SDMX, подразделяются на три широкие группы:

1. Инструменты для «принимающих организаций», включающие «мастер-программы» для создания структуры данных (DSW), которые представляют собой программные продукты для создания артефактов SDMX, например реестр Euro SDMX (для европейского региона), который хранит созданные артефакты.

2. Инструменты для «поставщиков данных и метаданных», охватывающие программные продукты, преобразующие файлы с данными в SDMX-ML, справочную инфраструктуру SDMX (SDMX-RI) [14], которая формирует и распространяет данные в формате SDMX-ML непосредственно из баз данных, используя набор специальных инструментов. Используя SDMX-RI, поставщики данных приобретают инструменты для конвертации своих статистических данных в SDMX и передачи их организациям-получателям на основе архитектурных стандартов веб-сервисов (REST API, SOAP). Кроме того, SDMX-RI предназначен для предоставления данных и структурных метаданных на основе сопоставлений с базой данных распространения поставщика.

В специальные инструменты входят следующие:

- базы/репозитории данных и метаданных, где хранится исходная информация для конвертации;

- «конвертеры», позволяющие конвертировать многомерные таблицы Excel, содержащие статистические данные и метаданные, в соответствии с SDMX;

- программы управления данными, в которых осуществляется маппирование данных с учетом заданной структуры;

- база данных распространения - это хранилища данных (или базы данных), где хранятся статистические данные в соответствии со стандартом SDMX, готовые для передачи;

- графические интерфейсы пользователей для просмотра, загрузки, визуализации данных/метаданных, выгрузки файлов [15].

3. Инструменты для «ИТ-разработчиков», охватывающие источники исходных кодов для SDMX, веб-службы для разработки справочной инфраструктуры.

В этой области сложность на национальном уровне заключается в построении собственной ИТ-инфраструктуры. Организациям, предоставляющим данные, необходимо принимать решения, что же будет заложено в ее основу - собственные программные продукты или уже имеющиеся в свободном доступе, но требующие трансформации под уже существующую статистическую систему.

Разработки собственных программных продуктов всегда дороже использования и адаптации уже имеющихся.

На международном уровне в области информационных технологий также остается нерешенным вопрос валидации данных, представленных с применением стандарта SDMX. Язык валидации - это стандартный язык для определения правил проверки и преобразования (набора операторов, их синтаксиса и семантики) для любых статистических данных [16]. Предполагается логическая формализация правил валидации и трансформации, совместимая с другими языками программирования [17].

Частично валидация может выполняться методами, проверяющими синтаксис данных на соответствие стандарту, например процесс подтверждения элемента XML с использованием схемы XML, которая определяет допустимую структуру и содержание элементов.

Подразумевается, что язык валидации должен быть простым и понятным не только ИТ-специалистам. Массовое внедрение и распространение приложений для валидации планируется в 2020 г.

## Заключение

Распространение данных в формате SDMX становится сейчас все более предпочтительным способом получения данных для международных организаций, так как позволяет обеспечивать сопоставимость этих данных, осуществлять регулярный обмен «машина к машине», а также обеспечивать четкое структурирование данных путем процесса «маппирования». В связи с этим внедрение международного стандарта обмена данными и метаданными SDMX можно признать одним из важнейших направлений технологического развития современных статистических служб.

В условиях нарастающего интереса статистического сообщества к применению данного стандарта, его широкого распространения и применения в новых областях статистики перед ним возникают новые вызовы, которые требуют разработки новых стратегий развития стандарта.

Прежде всего, для эффективной интеграции международного формата обмена данными и метаданными SDMX в национальные статистические системы необходимы непрерывность проекта и его регулярный мониторинг. Следует повышать квалификацию и знания персонала, добиваться понимания целей и задач перехода на новый формат, устранять пробелы в коммуникации и улучшать связь между участниками проекта.

В настоящее время продолжается расширение технических возможностей применения стандарта, его глобализация, совершенствование схем данных. Однако множество выявленных проблем в существующих сценариях обмена данными требуют дальнейшего совершенствования самого стандарта SDMX, а также его адаптации к вызовам современности.

К ним относятся следующие:

- моделирование статистических данных;
- использование стандарта для осуществления информационного обмена микроданными и макроданными, геопространственной информацией;
- нормативно-правовое закрепление ответственности по осуществлению информационного обмена за конкретными участниками процесса и создание нормативно-правовой базы для эффективного осуществления информационного обмена;
- разработка глобального описания структуры данных (DSD), создание и поддержка общедоступных инструментов для конвертации и

валидации данных, внедрение механизма «Pull» и создание общедоступных платформ для организации рабочих процессов с применением стандарта SDMX.

Несмотря на то, что в рамках внедрения SDMX многие страны сталкиваются с необходимостью трансформации или реорганизации существующих национальных статистических систем и процессов обмена данными с международными организациями, необходимость применения формата SDMX на национальном уровне обусловлена преимуществами использования формата.

Основными преимуществами использования формата SDMX перед другими форматами обмена данными являются: облегчение обмена данными и метаданными, повышение доступности статистических данных и метаданных для пользователей, снижение нагрузки на поставщиков статистических данных благодаря автоматизации процесса предоставления статистических данных, направляемых в международные организации, а также обеспечение высокого качества статистической информации.

Помимо этого, стоит отметить, что SDMX является не только форматом для обмена данными и метаданными. Технические стандарты, статистические руководящие принципы, ИТ-архитектура и инструменты призваны совершенствовать производственные процессы в любой статистической организации, а также гармонизировать и стандартизировать статистические метаданные.

Осуществление информационного обмена статистическими данными и метаданными с применением международного стандарта SDMX - сложный комплексный процесс, который требует постоянного совершенствования существующих механизмов как на национальном, так и международном уровнях при их постоянном взаимодействии. В рамках данного взаимодействия национальные организации - участники процесса, предоставляющие данные и метаданные, обязаны учитывать рекомендации международных организаций, в то время как, международным организациям необходимо учитывать замечания, сформулированные на национальном уровне.

## Литература

1. **Boden M., Cagnin C., Carabias V., Haegeman K.** Facing the future: time for the EU to meet global challenges // European Commission Joint Research Centre Institute for Prospective Technological Studies European Union.

URL: [http://www.et2050.eu/docs/2010\\_IPTS\\_Facing\\_the\\_future.pdf](http://www.et2050.eu/docs/2010_IPTS_Facing_the_future.pdf) (дата обращения 28.04.2019).

2. **Берберов А.Б.** На пути к цифровизации российской экономики: проблемы и перспективы // Управление экономическими системами: электронный научный журнал. 2017. № 7(101). URL: <http://cyberleninka.ru/article/n/na-p> (дата обращения 28.04.2019).

3. Government of the Russian Federation 2017, «Digital Economy of the Russian Federation program». Approved by the Federal Government on July 28, 2017. No. 1632. URL: <http://government.ru/docs/28653/> (дата обращения 28.04.2019).

4. SDMX // данные официального сайта МВФ. URL: <http://datahelp.imf.org/knowledgebase/articles/500102-sdmx> (дата обращения 28.04.2019).

5. SDMX Content-oriented guidelines, 2016. SDMX. URL: [https://sdmx.org/wp-content/uploads/SDMX\\_COG\\_2016\\_Introduction.pdf](https://sdmx.org/wp-content/uploads/SDMX_COG_2016_Introduction.pdf) (дата обращения 28.04.2019).

6. **Гончаров А.Н., Кочева С.Н.** Некоторые вопросы использования стандарта SDMX на международном и национальном уровнях // Вопросы статистики. 2013. № 6. С. 40-44.

7. Guidelines for SDMX Data Structure Definitions, 2013. SDMX. URL: [https://sdmx.org/wp-content/uploads/SDMX\\_Guidelines\\_for\\_DSDs\\_1.0.pdf](https://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_DSDs_1.0.pdf) (дата обращения 28.04.2019).

8. Data exchange in business statistics (part 2): SDMX, 2013. Eurostat. URL: <https://ec.europa.eu/eurostat/documents/54610/7779382/Data-exchange-in-business-statistics-SDMX.pdf/>. (дата обращения 28.04.2019).

9. SDMX Guidelines for the creation and management of SDMX code list, 2018. SDMX. URL: [https://sdmx.org/wp-content/uploads/SDMX\\_Guidelines\\_for\\_CDCL.docx](https://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_CDCL.docx). (дата обращения 28.04.2019).

10. **Karaiskos D.S., Xinidis D., Bonis V.** R&D Statistics Information System: An Interoperability Tail Between CERIF and SDMX // Procedia Computer Science. 2017. No. 106. P. 87-94.

11. **Braaksma B., Zeelenberg K.** Information management as tool for standardization in statistics. Statistics Netherlands. URL: <https://www.cbs.nl/-/media/imported/documents/2014/07/information%20management.pdf> (дата обращения 28.04.2019).

12. SDMX Glossary, version 2.0, 2018. URL: [https://sdmx.org/wp-content/uploads/SDMX\\_Glossary\\_Version\\_2\\_0\\_October\\_2018.docx](https://sdmx.org/wp-content/uploads/SDMX_Glossary_Version_2_0_October_2018.docx). (дата обращения 28.04.2019).

13. **Capadisi S., Auer S., Riedl R.** Towards linked statistical data analysis. Semantic Statistics // Proceedings of the 1st International Workshop on Semantic Statistics. 2013. Vol. 1549, urn:nbn:de:0074-1549-5. URL: <http://csarven.ca/linked-statistical-data-analysis> (дата обращения 28.04.2019).

14. **Sembiring J., Uluwiyah A.** Data Exchange Design with SDMX Format for Interoperability Statistical Data // TELKOMNIKA Indonesian Journal of Electrical Engineering. 2015. Vol. 14. No. 2. Pp. 343-352. DOI: 10.11591/telkomnika.v14i2.7505.

15. **Ward D.** 2015. SDMX started kit for national statistical agencies. SDMX. URL: [https://sdmx.org/wp-content/uploads/SDMX\\_Starter\\_Kit\\_Version\\_23-9-2015.pdf](https://sdmx.org/wp-content/uploads/SDMX_Starter_Kit_Version_23-9-2015.pdf). (дата обращения 28.04.2019)

16. **Кочева С.Н., Гончарова Е.А., Воликова Т.И.** Рекомендации по использованию Глоссария SDMX в российской системе государственной статистики. М.: 2018.

17. Validation and Transformation Language (VTL): Part 2 - Reference Manual, 2018. SDMX. URL: <https://sdmx.org/wp-content/uploads/VTL-2.0-Reference-Manual-20180416-final.pdf> (дата обращения 28.04.2019).

### Информация об авторах

*Башина Ольга Эмильевна* - д-р. экон. наук, профессор, зав. кафедрой статистики, маркетинга и бухгалтерского учета, Московский гуманитарный университет. 111395, г. Москва, ул. Юности, 5. E-mail: [bashina\\_o\\_e@mail.ru](mailto:bashina_o_e@mail.ru). ORCID: <https://orcid.org/0000-0003-1253-4029>.

*Комкова Наталья Алексеевна* - советник Управления статистики зарубежных стран и международных статистических проектов Федеральной службы государственной статистики (Росстат). 107450, г. Москва, Мясницкая, 39, стр. 1. E-mail: [KomkovaNA@gks.ru](mailto:KomkovaNA@gks.ru).

*Матраева Лилия Валериевна* - д-р. экон. наук, доцент, профессор кафедры экономической теории и мировой экономики, Российский государственный социальный университет. 129226, г. Москва, ул. Вильгельма Пика, 4 стр. 1. E-mail: [matraeva@gambler.ru](mailto:matraeva@gambler.ru). ORCID: <https://orcid.org/0000-0002-9080-7953>.

*Косолапова Валерия Евгеньевна* - аспирант кафедры экономической теории и мировой экономики Российского государственного социального университета. 129226, г. Москва, ул. Вильгельма Пика, 4 стр. 1. E-mail: [Kosolapova93@gmail.com](mailto:Kosolapova93@gmail.com).

### Финансирование

Исследование выполнено при финансовой поддержке РФФИ в рамках проекта проведения научных исследований «Исследование потенциала отечественных производителей по обеспечению импортозамещения на потребительском рынке с использованием технологий BigData», проект № 17-02-00718 и в рамках проекта проведения научных исследований «Разработка подходов к созданию системы оценки состояния и определения перспективных направлений развития научной сферы», проект № 16-02-00407.

## Reference

1. **Boden M., Cagnin C., Carabias V., Haegeman K.** *Facing the Future: Time for the EU to Meet Global Challenges*. European Commission Joint Research Centre Institute for Prospective Technological Studies European Union; 2010. p. 44. Available from: [http://www.et2050.eu/docs/2010\\_IPTS\\_Facing\\_the\\_future.pdf](http://www.et2050.eu/docs/2010_IPTS_Facing_the_future.pdf) (accessed 28.04.2019).
2. **Berberov A.B.** On the Way to Digitalization of the Russian Economy: Problems and Prospects. *Economic Systems Management: electronic scientific journal*. 2017;7(101). (In Russ.) Available from: <http://cyberleninka.ru/article/n/na-p> (accessed 28.04.2019).
3. Government of the Russian Federation. *Digital Economy of the Russian Federation program*. Approved by the Federal Government on July 28. 2017. No. 1632. (In Russ.) Available from: <http://government.ru/docs/28653> (accessed 28.04.2019).
4. SDMX. IMF official website data. Available from: <http://datahelp.imf.org/knowledgebase/articles/500102-sdmx> (accessed 28.04.2019).
5. *SDMX Content-Oriented Guidelines, 2016*. SDMX. Available from: [https://sdmx.org/wp-content/uploads/SDMX\\_COG\\_2016\\_Introduction.pdf](https://sdmx.org/wp-content/uploads/SDMX_COG_2016_Introduction.pdf) (accessed 28.04.2019).
6. **Kocheva S., Goncharov A.** Selected Questions of Using SDMX Standard on a National and International Levels. *Voprosy statistiki*. 2013;(6):40-44. (In Russ.)
7. *Guidelines for SDMX Data Structure Definitions, 2013*. SDMX. Available from: [https://sdmx.org/wp-content/uploads/SDMX\\_Guidelines\\_for\\_DSDs\\_1.0.pdf](https://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_DSDs_1.0.pdf) (accessed 28.04.2019).
8. *Data Exchange in Business Statistics (part 2): SDMX, 2013*. Eurostat. Available from: <https://ec.europa.eu/eurostat/documents/54610/7779382/Data-exchange-in-business-statistics-SDMX.pdf> (accessed 28.04.2019).
9. *SDMX Guidelines for the Creation and Management of SDMX Code List, 2018*. SDMX. Available from: [sdmx.org/wp-content/uploads/SDMX\\_Guidelines\\_for\\_CDCL.docx](https://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_CDCL.docx) (accessed 28.04.2019).
10. **Karaiskos D.S., Xinidis D., Bonis V.** R&D Statistics Information System: An Interoperability Tail Between CERIF and SDMX. *Procedia Computer Science*. 2017;(106):87-94.
11. **Braaksma B., Zeelenberg K.** *Information Management as Tool for Standardization in Statistics*. Statistics Netherlands; 2014. Available from: <https://www.cbs.nl/-/media/imported/documents/2014/07/information%20management.pdf> (accessed 28.04.2019).
12. *SDMX Glossary, Version 2.0, 2018*. Available from: [https://sdmx.org/wp-content/uploads/SDMX\\_Glossary\\_Version\\_2\\_0\\_October\\_2018.docx](https://sdmx.org/wp-content/uploads/SDMX_Glossary_Version_2_0_October_2018.docx) (accessed 28.04.2019).
13. **Capadislis S., Auer S., Riedl R.** Towards Linked Statistical Data Analysis. *Semantic Statistics*, 2013. In: *Proceedings of the 1st International Workshop on Semantic Statistics, Vol. 1549, urn:nbn:de:0074-1549-5*. Available from: <http://csarven.ca/linked-statistical-data-analysis> (accessed 28.04.2019).
14. **Sembiring J., Uluwiyah A.** Data Exchange Design with SDMX Format for Interoperability Statistical Data. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2015;14(2):343-352. Available from: doi: 10.11591/telkomnika.v14i2.7505.
15. **Ward D.** SDMX Started Kit for National Statistical Agencies. SDMX. 2015. Available from: [https://sdmx.org/wp-content/uploads/SDMX\\_Starter\\_Kit\\_Version\\_23-9-2015.pdf](https://sdmx.org/wp-content/uploads/SDMX_Starter_Kit_Version_23-9-2015.pdf) (accessed 28.04.2019).
16. **Kocheva S.N., Goncharova E.A., Volikova T.I.** *Recommendations on the Use of the SDMX Glossary in the Russian System of State Statistics*. Moscow: 2018 (In Russ.)
17. *Validation and Transformation Language (VTL): Part 2 - Reference Manual, 2018*. SDMX. Available from: <https://sdmx.org/wp-content/uploads/VTL-2.0-Reference-Manual-20180416-final.pdf> (accessed 28.04.2019).

## About the authors

*Olga E. Bashina* - Dr. Sci. (Econ.), Professor, Head, Department of Statistics, Marketing and Accounting, Moscow University for the Humanities. 5, Yunosti St., Moscow, 111395, Russia. E-mail: bashina\_o\_e@mail.ru. ORCID: <https://orcid.org/0000-0003-1253-4029>.

*Natal'ya A. Komkova* - Adviser, International Statistics and Projects Department, Federal State Statistics Service (Rosstat). 39, Miasnitskaya St., Bldg. 1, Moscow, 107450, Russia. E-mail: KomkovaNA@gks.ru.

*Liliia V. Matraeva* - Dr. Sci. (Econ.), Docent, Professor, Department of Economic Theory and World Economy, Russian State Social University. 4, Wilhelm Pieck St., Bldg. 1, Moscow, 129226, Russia. ORCID: <https://orcid.org/0000-0002-9080-7953>.

*Valeriya E. Kosolapova* - Postgraduate Student, Department of Economic Theory and World Economy, Russian State Social University. 4, Wilhelm Pieck St., Bldg. 1, Moscow, 129226, Russia. E-mail: Kosolapova93@gmail.com.

## Funding

The study was funded by the Russian Foundation for Basic Research as part of the research project «Investigating the Potential of Domestic Producers to Ensure Import Substitution in the Consumer Market Using Big Data Technologies», grant No. 17-02-00718 and project «Developing Approaches to Creating a System for Assessing the State and Identifying Promising Areas for the Development of the Scientific Sphere», grant No. 16-02-00407.