

Интеграция данных о населении: проблемы и перспективы

**Михаил Александрович Клупт^{а)},
Олег Николаевич Никифоров^{б)}**

^{а)} Санкт-Петербургский государственный экономический университет, г. Санкт-Петербург, Россия;

^{б)} Управление Федеральной службы государственной статистики по г. Санкт-Петербургу и Ленинградской области, г. Санкт-Петербург, Россия

В статье рассматриваются методологические и организационно-технологические вопросы интеграции данных о населении, полученных государственной статистикой из административных источников и корпоративных «больших данных». Аргументируется особая актуальность взаимодействия государственной статистики с другими государственными и корпоративными информационными системами в условиях цифровизации экономики и общественной жизни, и начавшегося формирования федерального регистра данных о населении.

Предлагается система взаимосвязанных показателей, характеризующих типологические группы населения, различающиеся по следующим критериям: гражданства, месту постоянного жительства, длительности и целям пребывания на территории России. Анализируются причины трудностей практического расчета этих показателей.

Рассматриваются возможности и правовые ограничения в работе статистиков по систематизации сведений, обоснованию подбора и последующего совместного использования в анализе сведений, характеризующих физическое лицо (т.е. матчинга), для решения различных задач социальной и демографической статистики. Особое внимание уделяется различным вариантам решения вопроса о персональном коде (одном или нескольких) позволяющих увязывать находящуюся в различных базах данных информацию о данном физическом лице. Подчеркивается необходимость обеспечения прозрачности методологии, используемой различными участниками информационного взаимодействия, что, в свою очередь, создаст предпосылки для гармонизации, а там, где это возможно, и унификации такой методологии.

Показывается принципиально важная роль предварительного качественного анализа данных, полученных из различных источников, и объясняются пути дальнейшего взаимодействия органов официальной статистики с заинтересованными в данной информации организациями и общественными структурами. С использованием данных провайдеров и сотовых операторов о населении города характеризуются необходимые условия их адекватной интерпретации: прозрачность методологии; аргументированная характеристика групп населения как объектов рассматриваемого статистического наблюдения, численность которого оценивается; характеристика предположений, которые используются при такой оценке.

Ключевые слова: государственная статистика, ведомственная статистика, статистика населения, большие данные, интеграция данных, матчинг, регистр населения, персональный идентификационный номер.

JEL: C80, J10, R23.

doi: <https://doi.org/10.34023/2313-6383-2021-28-3-5-14>.

Для цитирования: Клупт М.А., Никифоров О.Н. Интеграция данных о населении: проблемы и перспективы. Вопросы статистики. 2021;28(3):5-14.

Integrating Population Data: Challenges and Prospects

**Mikhail A. Klupt^{а)},
Oleg N. Nikiforov^{б)}**

^{а)} St. Petersburg State University of Economics (UNECON), St. Petersburg, Russia;

^{б)} Rosstat Regional Office of St. Petersburg and Leningrad Region, St. Petersburg, Russia

The article deals with methodological, organizational, and technological issues of integrating population data obtained from various administrative sources and corporative big data. The article proves the particular relevance of the interaction between official statistics and other governmental and corporative information systems amidst the digitization of the economy and social life and the incipient establishment of the federal population data register.

The authors propose a system of interrelated aggregates, characterizing various categories of population, which differ according to criteria of citizenship, permanent residence, duration, and purposes of stay on the territory of Russia. Challenges associated with estimating these aggregates are analyzed.

The article considers possibilities and legal limitations in the work of statisticians on systematizing information, rationalizing the selection and subsequent joint use of information, characterizing an individual (i.e. matching) for addressing various tasks faced by social and demographic statistics. Special attention is paid to the various options for resolving the issue of a personal code (one or more) that allow linking information on the individual from different databases. The need to ensure the transparency of the methodology used by the various participants of informational interaction is emphasized, which in turn shall pave the way for the harmonization and, where possible, the unification of such methodology.

The paper demonstrates the crucial role of preliminary qualitative analysis of data from different sources and explains mechanisms for further interaction of statistical authorities with organizations, interested in this information, and social structures. Using mobile operators' and providers' data on the population of the city, necessary conditions for their adequate interpretation — transparent methodology, clear description of population aggregates to estimate, and assumptions used for such estimations — are characterized.

Keywords: state statistics, departmental statistics, population statistics, big data, data integration, matching, population register, personal identification number.

JEL: C80, J10, R23.

doi: <https://doi.org/10.34023/2313-6383-2021-28-3-5-14>.

For citation: Klupt M.A., Nikiforov O.N. Integrating Population Data: Challenges and Prospects. *Voprosy Statistiki*. 2021;28(3):5-14. (In Russ.)

Цифровизация социальной жизни неизбежно влечет за собой изменения во взаимодействии государственной статистики с другими информационными системами. Цель настоящей статьи — попытаться осмыслить ход этого процесса, выявить связанные с ним императивы, возможности, проблемы и риски. Круг актуальных в научном и практическом отношении аспектов данной проблематики весьма широк: Единый федеральный информационный регистр сведений о населении (далее ЕФИР); использование в официальной статистике «больших данных»; интерпретация сведений, полученных из различных источников часто весьма разноречивых; взаимодействие с масс-медиа и т. д. Ввиду объемности тематики мы ограничились ее рассмотрением лишь одним, но крайне важным аспектом — интеграции данных о населении.

Агрегаты населения как возможность и проблема

Социально-политические изменения, произошедшие за три десятилетия после распада СССР, расширили список категорий населения, знание численности которых важно для понимания общественных процессов и принятия управленческих решений. К числу таких категорий относятся, например: иностранные граждане, находящиеся на территории России в целях трудовой деятельности и осуществляющие ее с соблюдением или нарушением законодательных норм; российские граждане, проживающие за рубежом;

граждане иностранных государств, относящиеся к категории соотечественников и т. д. Вполне естественно стремление каким-то образом упорядочить и связать между собой эти категории, не забывая, разумеется, и о той, что остается главной из них — общей численности населения.

Образцом такой увязки могут служить макроэкономические агрегаты системы национальных счетов (далее — СНС) и агрегаты денежной массы (M0, M1 и т. д.). Сходство задач очевидно — различные категории населения, подобно агрегатам СНС и денежным агрегатам, дополняя друг друга, характеризуют различные части целого, а соотношения между ними несложно формализовать с помощью системы балансовых тождеств. Один из возможных перечней таких категорий (далее мы будем именовать их *агрегатами населения*) приведен в таблице.

Все категории данного перечня за исключением граждан РФ, постоянно проживающих за границей, и соотечественников — теоретически могут быть использованы для классификации населения как России в целом, так и ее административно-территориальных образований. Научная и практическая значимость информации о численности этих категорий очевидна и, кажется, никем не оспаривается. Более того, в научной литературе и высказываниях официальных лиц периодически приводятся оценки численности проживающих за границей российских граждан, основанные на данных зарубежной статистики [2 и 3]; численности иностранных граждан, получивших российское гражданство [4]; долгосрочных (вклю-

Агрегаты населения*

№ группы	Гражданство	Место постоянного жительства	Характер пребывания на территории РФ	Наименование группы
1	Российской Федерации ¹	Российская Федерация	Постоянно	Граждане РФ, постоянно проживающие в РФ
2		За границей	Сезонный (периодический)	Граждане РФ, постоянно проживающие за границей
3	Других стран ²	Российская Федерация	Постоянно ³	Иностранцы граждане, постоянно проживающие в РФ
4		За границей	Временно более 9 месяцев	Долгосрочные мигранты, включаемые в численность населения РФ
5			Временно менее 9 месяцев	Краткосрочные мигранты, не включаемые в численность населения РФ ⁴
6			Сезонный (периодический)	Соотечественники

* Данная таблица опубликована ранее одним из авторов в [1] и приводится здесь с небольшими изменениями.

Примечания:

1. При необходимости – с выделением лиц, получивших гражданство на протяжении нескольких последних (например, трех) лет.
2. При необходимости – с выделением групп стран, например, Белоруссии, других стран ЕАЭС, прочих стран.
3. При необходимости – с выделением правового положения иностранного гражданина – наличия у него вида на жительство, статуса беженца, вынужденного переселенца и т. д.
4. При необходимости классифицируются по целям пребывания в РФ, например, в соответствии с классификацией, предусмотренной Приказом Росстата от 2 ноября 2018 г. № 652.

чаемых в численность населения РФ) и краткосрочных трудовых мигрантов, в том числе находящихся на территории РФ с соблюдением и нарушением норм российского законодательства и т. д. Однако увязать все эти показатели в единую систему, подобную агрегатам СНС или денежным агрегатам, пока не удается.

Та же проблема существует и для административно-территориальных образований. На уровне дефиниций разведение двух категорий – постоянного населения города (или его района) и численности жителей, находящихся в некоторый момент времени на его территории, не представляет сложности. Однако и здесь результаты разовых исследований транспортных потоков, а в последнее время и «больших данных» мобильных операторов связи [5] не удастся объединить во взаимозвязанную систему агрегатов населения. Непосредственная причина этого очевидна – нехватка или несопоставимость данных, необходимых для перехода от дефиниций к практическим расчетам. За ней, однако, кроются более глубокие проблемы, заслуживающие отдельного рассмотрения.

Матчинг и персональные идентификационные номера

Научная и практическая эффективность интеграции данных о населении, полученных из различных источников, в значительной степени определяется тем, как происходит информаци-

онное взаимодействие их распорядителей. Государственные ведомства являются важнейшими, но не единственными участниками информационного взаимодействия. Генераторами и распорядителями значительной части «больших данных» являются корпорации (операторы связи, интернет провайдеры, страховые, кредитные организации и др.) и работающие в контакте с ними научно-исследовательские коллективы, вследствие чего вопросы взаимодействия государственной статистики с ними становятся все более актуальными. Ввиду очевидной потребности в совершенствовании информационного взаимодействия, целесообразно рассматривать его не в качестве некоего статического состояния, а как динамический процесс, в ходе которого более или менее быстро, целенаправленно и эффективно меняется его характер.

Данный процесс имеет целый ряд составляющих: политическую, правовую, методологическую, организационно-технологическую, интерпретационную, медийную. Все они тесно взаимосвязаны. Так, отсутствие единой, хотя бы прозрачной и понятной всем участникам информационного взаимодействия методологии не только снижает сопоставимость данных, но и затрудняет их интерпретацию специалистами и увеличивает риск ошибочных интерпретаций в масс-медиа. В то же время достоинства даже хорошо отработанной методологии могут быть сведены на нет дефектами в технологии сбора и передачи информации от одного ведомства к другому. В случае же *матчинга*¹ –

¹ Термин «матчинг» многозначен и по-разному понимается в разных областях науки и практики, но удобен своей краткостью. Ниже в этой статье мы будем использовать его только в соответствии с данным определением.

подбора и последующего совместного использования в анализе сведений, характеризующих физическое лицо — например, сведений о возрасте достижения им того или иного уровня образования, вступления в брак, рождения детей, переезда на новое место жительства и т. д., переплетаются, как мы покажем ниже, сразу несколько из вышеперечисленных составляющих.

Матчинг, например, обеспечивает сопоставимость числителя и знаменателя при расчете коэффициентов рождаемости, смертности, брачности, разводимости и миграции, дифференцированных в разрезе образовательных групп, лиц родившихся в данной стране и за ее пределами и т. д. Наличие набора данных, относящихся к одному лицу, позволяет строить комбинационные таблицы по нескольким признакам, использовать процедуры кластер-анализа и т. д. Матчинг дает также возможность проводить анализ последовательности событий (*event history analysis*) в биографии индивида — например, анализировать влияние рождения детей на профессиональную деятельность женщин, сравнивать профессиональные карьеры выпускников различных вузов и т. д. С помощью матчинга можно также реализовывать возможности когортного метода — например, сравнивать демографические, профессиональные или «профессионально-демографические биографии» женщин в когортах, образованных по году рождения или вступления в брак, истории партнерских и супружеских отношений и т. д. Большие возможности открывает увязка данных об одном лице, содержащихся в различных источниках информации — ведомственных базах данных, регистрах, переписях населения, выборочных обследованиях и так далее [6].

Матчинг, таким образом, позволяет решать целый ряд задач, представляющих значительный интерес для социально-демографического анализа. В то же время результатом матчинга являются более или менее полные электронные профили каждого физического лица, данные о котором находятся в объединяемых системах. Ввиду этого применение матчинга и, в частности, комбинирование различных ведомственных источников данных с целью получения индивидуальных электронных профилей требует постоянного контроля со стороны общества и правовой регламентации — например — требования использования результатов только в обоб-

щенном и обезличенном виде и применения эффективных технологий защиты персональных данных.

Наиболее эффективным инструментом матчинга, если рассматривать его исключительно в информационно-технологическом ключе, является присвоение каждому физическому лицу, сведения о котором находятся в базах данных, персонального идентификационного номера (ПИН) — кодификатора, значения которого определяется либо в случайном порядке, либо с дополнением полученного случайным образом уникального номера цифровым кодом, содержащим базовые сведения об индивиде, например, его пол, дата рождения, дата приобретения гражданства и т. д. В настоящее время ПИН используют 20 европейских стран [7, p. 54].

Другим, менее распространенным среди европейских стран вариантом решения проблемы матчинга является использование одного или нескольких «секторальных» кодов — например, налогового, используемого службами социального обеспечения. Единый универсальный ПИН не используется, например, в Ирландии, Испании, Италии и Швейцарии. В этом случае приходится применять более сложные формы матчинга.

Ряд стран отказались от использования ПИН, считая, что он может стать инструментом вмешательства государства в частную жизнь граждан и установления тотального контроля над ними. Важную роль при этом играют история и политическая культура страны. Вряд ли, например, можно считать случайным тот факт, что именно в ФРГ конституционный суд признал в 1983 г. использование ПИН противоречащим конституции и запретил применять данные переписей для корректировки административных регистров. В результате перепись 1983 г. была отменена и проведена только в 1987 г., причем на основе традиционного подхода, не предполагавшего использование данных регистров.

Тем не менее начиная с 2001 г. в ФРГ используется комбинированный способ проведения переписи, сочетающий три источника данных: полное обследование жилищ; получение данных об индивидах на основе регистров; выборочное обследование, позволяющее «добрать» информацию по признакам, отсутствующим в регистре. Говоря о перспективах использования регистров населения, немецкие статистики возлагают на-

дежду на то, что прогресс в технологиях защиты информации позволит совместить требования, содержащиеся в вердикте конституционного суда 1983 г., с нарастающей цифровизацией государственного управления [8, pp. 4, 14].

В российских условиях повышению информативности статистики населения существенно способствовало бы присвоение ПИН любому лицу, как проживающему в РФ, так и впервые пересекающему российскую границу. Подобное решение при соответствующей модернизации статистики пересечения границы позволило бы, среди прочего, опираться при характеристике численности и структуры миграционных потоков на данные о фактическом, а не юридическом выбытии индивида (истечении срока регистрации, фиксируемого МВД России). Другим «плюсом» такого решения является возможность получения комбинационных таблиц, фиксирующих распределения лиц, въезжающих в Россию из различных стран, не только по декларируемой цели въезда, но также по фактической длительности пребывания в России. Объединение этой крайне важной для организации туризма статистики с «большими данными», фиксирующими объемы и структуру расходов иностранных туристов в России, обеспечило бы еще больший экономический эффект.

Менее радикальным решением является сохранение нескольких «секторальных» ПИН, одним из которых предположительно будет ИНН, используемый в ЕФИР. Статьей 10 Федерального закона от 08.06.2020 № 168-ФЗ на федеральный орган, осуществляющий полномочия в сфере внутренних дел, возлагается обязанность по направлению в ЕФИР данных о постановке на миграционный учет и снятии с него ряда категорий международных мигрантов, а также о выдаче им разрешений на работу и патентов. Наличие единого идентификатора у всех физических лиц, данные о которых содержатся в ЕФИР, позволит получать данные об изменении за определенный период времени (месяц, квартал, год и т. д.) численности таких категорий международных мигрантов, как иностранные граждане в целом; иностранные граждане, проживающие в РФ на основании вида на жительство; иностранные трудовые мигранты, временно живущие в России.

Политико-правовые аспекты матчинга актуальны и для Российской Федерации. В период публичного обсуждения проекта Федерального закона от 8 июня 2020 г. № 168-ФЗ «О едином федеральном информационном регистре, содержащем сведения о населении Российской Федерации» основное внимание обращалось на вопросы, в наибольшей степени волнующие граждан, среди которых одно из центральных мест занимало обеспечение безопасности персональных данных². Нельзя исключить, что в ходе формирования ЕФИР вопросы, связанные с защитой персональных данных, вновь станут предметом обсуждения не только на экспертном, но также на медийном и политическом уровне. Это, в свою очередь, повышает значимость эффективного PR-сопровождения для успешного развития этого первого в российской истории регистра населения.

Методологические и организационно-технологические проблемы

Опыт зарубежных стран свидетельствует, что эффективность официальной статистической деятельности в значительной степени определяется законодательной базой, правоприменительной практикой и, в конечном счете, культурой межведомственного взаимодействия. Его чрезвычайно важным аспектом является наличие единой официальной и прозрачной методологии, являющейся основой статистической деятельности всех ведомств.

В российских условиях решение проблем в данной сфере требует последовательного решения двух задач. Прежде всего, необходимо обеспечить прозрачность и единство методологии, используемой всеми распорядителями источников данных, участвующих в информационном взаимодействии. Это, в свою очередь, создаст предпосылки для решения второй задачи — гармонизации, а там, где это возможно, и унификации такой методологии.

В настоящее время обе задачи далеки от решения. Статистическая методология учета населения остается единственной официально публикуемой государственной метаинформацией. Параллельно с ней существуют ведомственные

² См., например: URL: <http://duma.gov.ru/news/48646/> (дата обращения 12.03.2021).

нормативные правовые акты силовых ведомств, определяющие порядок регистрационного учета российских граждан, миграционного учета иностранных граждан и пересечения российскими и иностранными гражданами государственной границы РФ. Данные о постановке иностранных граждан на миграционный учет и другие данные, характеризующие миграционную ситуацию, размещаются без каких-либо методологических комментариев на официальном сайте МВД России³, а данные о пересечении границы, формируемые Пограничной службой ФСБ России, также без методологических комментариев – в Единой межведомственной информационно-статистической системе⁴. О том, что при регистрационном учете иностранных граждан фиксируется их снятие с учета, а не фактический выезд из страны; а пограничная статистика фиксирует число пересечений границы, вследствие чего число таких пересечений значительно превышает число лиц, хотя бы однажды пересекавших границу в течение рассматриваемого периода, можно узнать лишь из научных публикаций и комментариев специалистов, не носящих официального характера.

Существование «параллельных» методик подсчета, вообще говоря, является следствием того, что формирование статистических данных является побочным, хотя и важным продуктом деятельности профильных ведомств. Однако при интеграции данных такой параллелизм становится серьезной проблемой. Так, ЕФИР может, вообще говоря, использоваться в разных целях, разрешенных законом. Необходимо, однако, обеспечить такое положение дел, при котором данные этого регистра выгружаются и агрегируются в строгом соответствии с действующей государственной методологией официального статистического учета ведомством, отвечающим за его организацию в Российской Федерации и соблюдение международных стандартов демографической статистики. Следует также отметить, что первостепенное значение при организации межведомственных взаимодействий имеют эталонные значения. Например, документированные факты демографических событий при множестве источников должны иметь только одно «происхождение» – органы ЗАГС.

С реализацией межведомственного информационного обмена и, в частности, межведомственного обмена базами данных о гражданах связаны также немалые организационно-технологические трудности. Первый опыт такого взаимодействия начался 1 октября 2018 года, когда оператором системы «Единый государственный реестр записей актов гражданского состояния» (ЕГР ЗАГС) постановлением правительства России № 738⁵ была определена Федеральная налоговая служба России. Сегодня ЕГР ЗАГС – ключевой элемент создаваемого ЕФИР. Наладив взаимодействие Росстата с ФНС в получении актов записей по рождению, смерти, заключению и расторжению брака, зарегистрированных органами ЗАГС в субъектах Российской Федерации, осуществлялось посредством Системы межведомственного электронного взаимодействия (СМЭВ–3). Учитывая высокую востребованность официальной статистической информации о естественном движении населения, сведения, полученные из ЕГР ЗАГС на федеральном уровне, тщательно сверялись вручную специалистами территориальных органов государственной статистики в субъектах Федерации с местными органами ЗАГС, продолжавшими некоторое время вести «двойной» учет, что потребовало дополнительных трудозатрат. В ходе опытной эксплуатации ЕГР ЗАГС была выявлена необходимость:

- совершенствования нормативных требований к заполнению реквизитов записей актов гражданского состояния и процедурам, выполняемым при их отсутствии;
- совершенствования формально-логического и предупредительного контроля при вводе данных в ЕГР ЗАГС;
- перехода к вводу адреса из Федеральной Информационной Адресной Системы (ФИАС);
- использования в Единой системе нормативно-справочной информации Минкомсвязи России (ЕСНСИ) справочника Международной классификации болезней и проблем, связанных со здоровьем 10-го пересмотра (МКБ-10), размещенного в федеральном реестре нормативно-справочной информации в сфере здравоохранения (подсистема ЕГИСЗ).

³ URL: <https://xn--b1aew.xn--p1ai/Deljatelnost/statistics/migracionnaya/item/22689548/> (дата обращения 15.03.2021).

⁴ URL: <https://www.fedstat.ru/indicator/38479> (дата обращения 15.03.2021).

⁵ Постановление Правительства РФ от 27 июня 2018 г. № 738 «Об утверждении Правил ведения Единого государственного реестра записей актов гражданского состояния». URL: <https://base.garant.ru/71978490/>.

Электронное взаимодействие органов ЗАГС, ФНС и Росстата удалось завершить до пандемии COVID-19, осложнившей регистрацию причин смерти граждан. На преодоление издержек опытной эксплуатации ЕГР ЗАГС потребовалось более года, а с января 2021 г. региональные ЗАГС прекратили ведение локальных информационных систем актовых записей гражданского состояния. Приведенный пример становления ЕГР ЗАГС и предполагаемая этапность — до 2025 г. построения ЕФИР — сохраняет надежду на решение непростых задач в определении роли эталонных данных и второстепенных сведений ведомств о населении, вносимых в ЕФИР, и регламент доступа в него.

Интерпретация данных

Интерпретация данных играет важную роль в отношениях государственной статистики и общества. В одобренных ООН Основных принципах статистической деятельности далеко не случайно говорится о том, что статистические ведомства должны способствовать правильной интерпретации данных, и «комментировать неверную интерпретацию или неправильное использование статистических данных»⁶. Неверные трактовки статистических данных препятствуют адекватному восприятию сложившейся ситуации общественным мнением, снижают качество управления и авторитет самой статистики.

В теории статистики интерпретация данных не рассматривается, во всяком случае, явно, как самостоятельная когнитивная процедура и одна из составляющих интеграции данных. Она между тем является и тем и другим. Сталкиваясь с множеством данных из различных источников, пользователь неизбежно, хотя и не всегда осознанно, стремится к их интеграции и получению «в сухом остатке» непротиворечивых и понятных ему выводов.

В рамках рассматриваемой в данной статье тематики можно выделить несколько проблем, связанных с интерпретацией данных о населении. Одна из них связана с недостаточно точным и полным описанием исследуемой совокупности, а также цели и метода ее исследования. Новой инкарнацией этой старой проблемы стал вопрос об использовании в исследовании населения

городов данных операторов мобильной связи. На заседании Статистической Комиссии ООН в марте 2019 г. отмечалось, что такие источники данных, как «геопространственная информация... и другие виды неструктурированных данных, генерируемых в ходе деловых операций, регистрации звонков по мобильным телефонам и активности в социальных сетях» [9, р. 18] обладают значительным инновационным потенциалом. В то же время очевидно, что для эффективного использования данных мобильной связи в статистике населения необходимо выполнение ряда условий.

Во-первых, это четкое описание совокупности, на оценку численности которой направлено исследование — например, лиц, находящихся в дневное (вечернее, ночное) время в «спальных» (центральных, деловых, промышленных и т. д.) районах города.

Во-вторых, столь же четкая постановка цели исследования. В исследовании [5], построенном на основе данных операторов сотовой связи, в качестве такой цели выступает, например, анализ изменений в соотношении численности дневного и ночного населения в районах разного типа, происшедших под влиянием введения в строй Московского центрального кольца.

В-третьих, это подробное описание метода исследования и его ограничений — например, вероятных масштабов и территориальной дифференциации повторного счета, или, напротив, недоучета населения при его оценке на основе данных сотовых операторов.

В-четвертых, указание на причины различий в численности постоянного населения, полученного на основе методологии, используемой Росстатом, и данных о дневной, вечерней и ночной численности людей, находящихся, судя по данным сотовых операторов, на территории города или какого-либо из его районов.

Выполнение этих требований, способствуя получению нового знания, одновременно предотвратит бесплодные дискуссии, основанные на забвении известного принципа: «О терминах не спорят, а договариваются». Заметим также, что все вышеперечисленные требования к описанию данных и методов исследования отнюдь не новы — они, например, являются обязательными при публикации статей во многих областях социальных и естественных наук.

⁶ URL: <https://unstats.un.org/unsd/methods/statorg/FP-Russian.pdf> (дата обращения 16.03.2021).

В ряде случаев проблемы интерпретации данных тесно связаны с особенностями учета, например, его неполнотой или временными лагами в фиксации демографических событий различными ведомствами — применительно к проблемам формирования ЕФИР данный вопрос, например, подробно рассмотрен в [10]. В то же время следует иметь в виду, что сложности интерпретации могут быть обусловлены и многомерностью, и сложностью структуры, как временной, так и пространственной, самого объекта исследования. Периоды, в которых показатели, охватывающие различные стороны процесса, изменяются однонаправленно, чередуются с периодами их разнонаправленного изменения, порождают конфликтные интерпретации. Недавним примером является пришедший на 2007–2015 гг. период быстрого роста в России суммарных коэффициентов рождаемости для условного поколения, повлекший очередной виток спора о том, можно ли доверять этим показателям или, напротив, истинное положение дел можно оценить, лишь сравнивая показатели рождаемости последовательных когорт.

Завершая раздел, заметим, что неизбежная в ряде случаев конфликтность интерпретаций может, парадоксальным образом, быть не только проблемой статистики, но и ресурсом ее популяризации. Журналистика, даже в случае, когда речь идет о качественных деловых изданиях, живет по законам драматургии — конфликт и создаваемый им информационный повод относятся к числу ее краеугольных камней; конфликт интерпретаций не является в этом отношении исключением. Важно лишь, чтобы конфликтность интерпретаций становилась информационным поводом для объемного представления проблемы обществу и, сохраняя популярность изложения, не превращалась бы при этом в набор столь же броских, сколь и бессодержательных сентенций.

Заключительные замечания

Современное состояние интеграции данных о населении оставляет возможность для различного хода событий, крайние варианты которых можно свести к двум полярно противоположным сценариям. При одном из них — пессимистическом — сколько-нибудь существенного продвижения в деле интеграции данных не происходит.

При другом варианте — оптимистическом — быстрое развитие интеграции данных приносит значительный социальный и экономический эффект. Завершая статью, охарактеризуем важнейшие риски, связанные с первым сценарием, и условия для реализации второго.

Развитие событий в минувшей части столетия показывает, что тенденция к созданию все новых и новых ведомственных информационных систем, не объединяемых межведомственным взаимодействием, далеко не преодолена. Создание ЕФИР способно переломить эту тенденцию лишь при придании этой инновационной для России задачи статуса общегосударственной, а не очередной узковедомственной задачи, предназначенной в данном случае для решения вопросов, находящихся в ведении Федеральной налоговой службы. Потенциал ЕФИР будет в полной мере реализован лишь в случае предоставления пользователям возможности производить на основе платформенных решений, разумеется, с соблюдением требований конфиденциальности, агрегирование первичных данных.

Задачи, связанные с формированием и последующим функционированием ЕФИР, носят долговременный характер. В органах власти, ответственных за ведение информационных систем, пока не сформировались на концептуальном уровне единые подходы и регламенты, обеспечивающие «жизнь» регистров после ввода в эксплуатацию. Она, между тем, невозможна без обратной связи с пользователями. В зарубежных странах регистры населения «живут» благодаря их совместному ведению в нескольких ведомствах. Сведения о населении не только собираются из административных источников данных различных ведомств, но и сопровождаются в дальнейшем этими ведомствами, обеспечивая непрерывность процесса верификации и актуализации содержащихся в регистрах населения сведений [11]. На определенные фиксированные моменты времени (как правило, первые числа месяца, квартала, года) «счетчик» регистра населения как бы останавливается, и публикуются данные о числе происшедших демографических событий и численности населения.

Проблемы интеграции статистических данных, полученных из разных источников, при всей первостепенной важности ЕФИР, не сводятся только к его формированию и носят более широкий характер. Они выходят за пределы вза-

имодействия государственных ведомств и затрагивают вопросы их информационного взаимодействия с корпорациями и научными коллективами — примером является быстро набирающий актуальность вопрос об использовании больших данных, и в частности, данных сотовых операторов. Важнейшим условием успешного решения этих задач является преодоление узковедомственной замкнутости и следование методологическим стандартам, принятым в мировой науке и статистической практике.

Литература

1. **Клупт М.А.** Данные о международной миграции в Россию: проблемы согласованности и достоверности // Управление данными в экономике. Под науч. ред. М.В. Бочениной, И.И. Елисеевой. СПб.: Изд-во СПбГЭУ, 2020. С. 79–95.
2. **Денисенко М.Б.** Эмиграция из России по данным зарубежной статистики // Демоскоп. Weekly. 2012. № 513–514. URL: <http://demoscope.ru/weekly/2012/0513/tema01.php> (дата обращения 24.03.2021).
3. **Потапова А.А.** Эмиграция из России: текущее десятилетие // Демоскоп Weekly. 2017. № 719–720. URL: <http://demoscope.ru/weekly/2017/0719/tema01.php> (дата обращения 24.03.2021).
4. **Чудиновских О.С.** Статистика приобретения гражданства как отражение особенностей миграционной политики России // Вопросы статистики. 2018. Т. 25. № 9. С. 3–26.
5. **Махрова А.Г., Бабкин Р.А., Казаков Э.Э.** Динамика дневного и ночного населения как индикатор структурно-функциональных изменений территории города в зоне влияния Московского центрального кольца с использованием данных операторов сотовой

связи // Контуры глобальных трансформаций: политика, экономика, право. 2020. Т. 13. № 1. С. 159–179.

6. Организация Объединенных Наций. Руководящие принципы использования регистров и административных данных для переписей населения и жилищного фонда. Нью-Йорк и Женева, 2018 год. Доступно по адресу <https://unece.org/guidelines-use-registers-and-administrative-data-population-and-housing-censuses-0> (дата обращения 24.03.2021).

7. Population Registers in Different Countries. Design and developments in relation to the Netherlands. Zoetermeer. 2019. URL: <https://kennisopenbaarbestuur.nl/media/256912/population-registers-in-different-countries.pdf> (дата обращения 24.03.2021).

8. The Combined Census Model in Germany — Origins, Lessons Learned and Future Perspectives. Note by the Federal Statistical Office, Germany. Conference of European Statisticians Group of Experts on Population and Housing Censuses Nineteenth Meeting Geneva, Switzerland 4–6 October 2017. Available from: https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2017/Meeting-Geneva-Oct/WP25_ENG.pdf (дата обращения 24.03.2021).

9. ООН. Статистическая комиссия. Статистика международной миграции. Доклад Генерального секретаря. Пятидесятая сессия 5–8 марта 2019 года. URL: <https://unstats.un.org/unsd/statcom/50th-session/documents/2019-17-MigrationStats-R.pdf> (дата обращения 24.03.2021).

10. **Чудиновских О.С.** К вопросу о создании регистра населения и использования административных данных для нужд государственной статистики // Вопросы статистики. 2021. Т. 28. № 1. С. 5–17. URL: <https://doi.org/10.34023/2313-6383-2021-28-1-5-17>.

11. **Poulain M., Herm A.** Central Population Registers as a Source of Demographic Statistics in Europe // Population. 2013. Vol. 68. Iss. 2. P. 183–212.

Информация об авторах

Клупт Михаил Александрович — д-р экон. наук, профессор кафедры статистики и эконометрики Санкт-Петербургского государственного экономического университета. 191023, г. Санкт-Петербург, ул. Садовая, д. 21. E-mail: klupt@mail.ru. ORCID: <https://orcid.org/0000-0002-9423-4363>.

Никифоров Олег Николаевич — канд. экон. наук, руководитель Управления Федеральной службы государственной статистики по г. Санкт-Петербургу и Ленинградской области. 197376, г. Санкт-Петербург, ул. Профессора Попова, д. 39. E-mail: onik758@yahoo.com.

References

1. **Klupt M.A.** Data on International Migration to Russia: The Problems of Consistency and Reliability. In: Eliseeva I.I., Bochenina M.V. (eds.) *Data Management in Economy*. St. Petersburg: SPbGEU Publ.; 2020. P. 79–95 (In Russ.).
2. **Denissenko M.B.** Emigration from Russia According to Foreign Statistics. *Demoscope Weekly*. 2012;(513–514). (In Russ.) Available from: <http://demoscope.ru/weekly/2012/0513/tema01.php>.

3. **Potapova A.A.** Emigration from Russia: The Current Decade. *Demoscope Weekly*. 2017;(719–720). (In Russ.) Available from: <http://demoscope.ru/weekly/2017/0719/tema01.php>.

4. **Chudinovskikh O.S.** Statistics on Citizenship Acquisition as a Reflection of the Peculiarities of the Russian Migration Policy. *Voprosy Statistiki*. 2018;25(9):3–26. (In Russ.)

5. **Makhrova A.G., Babkin R.A., Kazakov E.E.** The Dynamics of the Day and Night Population as an Indica-

tor of Structural and Functional Changes in the Territory of the City in the Zone of Influence of the Moscow Central Ring Using Data from Mobile Operators. *Outlines of Global Transformations: Politics, Economics, Law*. 2020;13(1):159–179. (In Russ.)

6. United Nations. *Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses*. New York, Geneva; 2018. (In Russ.) Available from: <https://unece.org/guidelines-use-registers-and-administrative-data-population-and-housing-censuses-0> (accessed 24.03.2021).

7. *Population Registers in Different Countries. Design and Developments in Relation to the Netherlands*. Zoetermeer; 2019. Available from: <https://kennisopenbaarbestuur.nl/media/256912/population-registers-in-different-countries.pdf> (accessed 24.03.2021).

8. The Combined Census Model in Germany – Origins, Lessons Learned and Future Perspectives. Note by the Federal Statistical Office, Germany. In: *Conference*

of European Statisticians Group of Experts on Population and Housing Censuses Nineteenth Meeting Geneva, Switzerland 4–6 October 2017. Available from: https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2017/Meeting-Geneva-Oct/WP25_ENG.pdf (accessed 24.03.2021).

9. International Migration Statistics. Report of the Secretary-General. In: *Fiftieth session of the UN Statistical Commission, 5–8 March 2019*. Available from: <https://unstats.un.org/unsd/statcom/50th-session/documents/2019-17-MigrationStats-R.pdf> (accessed 24/03/2021).

10. **Chudinovskikh O.S.** On the Establishment of the Population Register and the Use of Administrative Data for the Needs of State Statistics. *Voprosy Statistiki*. 2021;28(1):5–17. (In Russ.) Available from: <https://doi.org/10.34023/2313-6383-2021-28-1-5-17>.

11. **Poulain M., Herm A.** Central Population Registers as a Source of Demographic Statistics in Europe. *Population*. 2013;68(2):183–212.

About the authors

Mikhail A. Klupt – Dr. Sci. (Econ.), Professor, Department of Statistics and Econometrics, St. Petersburg State University of Economics (UNECON). 21, Sadovaya Str., St. Petersburg, 191023, Russia. E-mail: klupt@mail.ru. ORCID: <https://orcid.org/0000-0002-9423-4363>.

Oleg N. Nikiforov – Cand. Sci. (Econ.), Head, Rosstat Regional Office of St. Petersburg and Leningrad Region. 39, Professor Popov Str., St. Petersburg, 197376, Russia. E-mail: onik758@yahoo.com.