

**Европейская экономическая комиссия****Конференция европейских статистиков****Шестьдесят седьмая пленарная сессия**

Париж, 26–28 июня 2019 года

Пункт 2 а) предварительной повестки дня

Новые источники данных – доступность и использование**Заседание 1: доступ к новым источникам данных****Доступ к большим данным для статистических целей****Записка Федерального статистического управления Германии***Резюме*

В настоящем документе представлен обзор деятельности Федерального статистического управления Германии (ФСУГ), направленной на обеспечение доступа к новым цифровым источникам данных. Для использования больших данных предстоит еще решить множество проблем. Во-первых, это вопросы качества и методологии (в частности, охват, систематическая ошибка отбора, обеспечение наличия данных на своевременной и регулярной основе). Федеральное статистическое управление уже приступило к подготовке ряда технико-экономических обоснований в этой области. Для того чтобы иметь возможность использовать новые цифровые данные для целей официальной статистики на более устойчивой основе, необходим постоянный доступ к данным, имеющимся у частных предприятий. Ввиду этого ФСУГ стремится создать юридическую базу для получения доступа к частным источникам данных, чтобы иметь возможность развивать дальнейшее сотрудничество.

Настоящий документ представлен для обсуждения на семинаре Конференции европейских статистиков 2019 года по теме «Новые источники данных – доступность и использование».



I. Введение

1. Федеральное статистическое управление Германии, а также рабочая группа по теме «Новые цифровые данные в официальной статистике» Статистического консультативного совета этого органа убеждены, что в массивах больших данных в сочетании с данными обследований и административными данными кроется колоссальный потенциал для официальной статистики. Ожидается, что использование этих новых источников данных позволит повысить своевременность выпуска официальных статистических данных и их точность и уменьшить нагрузку на респондентов, особенно в случае частных предприятий, данные которых пока собираются в рамках обследований. В результате официальная статистика отвечает ожиданиям общества в цифровую эпоху, используя все имеющиеся источники информации и стремясь, по возможности, избежать проведения трудоемких обследований. Удастся ли экономить деньги при таком подходе в официальной статистике, покажет время. Возможное сокращение затрат на проведение обследований может перекрываться – по крайней мере в краткосрочной перспективе – вложениями средств в аппаратное и программное обеспечения, а также в подготовку статистических работников. Кроме того, даже при наличии юридических гарантий доступа к частным источникам данных национальным статистическим управлениям, возможно, придется оплачивать услуги по предварительной обработке этих данных для статистических целей. К тому же потребуются специальный анализ затрат и выгод, и только после этого можно будет переходить к использованию больших данных не только для технико-экономических обоснований, но и для подготовки статистических данных.

2. Для изучения пригодности больших данных для целей официальной статистики ФСУГ приняло участие в проведении нескольких технико-экономических обоснований, например в разрезе использования данных мобильной телефонии и спутниковой связи для различных целей в сфере официальной статистики. В вопросах доступности и обработки больших данных ФСУГ сотрудничает с различными организациями, в частности с оператором мобильной связи «Ти-системз интернешнл ГмбХ» и компанией «Моушнлоджик ГмбХ» (обе компании являются дочерними предприятиями «Дойче Телеком АГ»), Германским центром авиации и космонавтики и Федеральным агентством картографии и геодезии Германии. Кроме того, ФСУГ является членом Целевой группы Европейской статистической системы (ЕСС), Руководящей группы ЕСС и Глобальной рабочей группы Организации Объединенных Наций по использованию больших данных для целей официальной статистики и сотрудничает со многими другими национальными статистическими институтами и университетами в рамках различных европейских проектов.

3. Для использования больших данных предстоит еще решить множество проблем. Во-первых, это вопросы качества и методологии (в частности, охват, систематическая ошибка отбора, обеспечение наличия данных на своевременной и регулярной основе). Кроме того, необходимо активизировать внутреннее и внешнее взаимодействие в вопросах использования больших данных. При этом следует изучить потребности в плане ИТ-инфраструктуры и интерфейсов, без которых не удастся использовать прошедшую первичную обработку продукцию частных владельцев данных, а также в плане уровня квалификации сотрудников, которые будут заниматься большими данными. Помимо этого, существуют различные правовые вопросы, которые необходимо решать с различными заинтересованными сторонами, включая министерства, службы, занимающиеся защитой данных на национальном и региональном уровнях, Федеральное сетевое агентство и, не в последнюю очередь, поставщиков данных.

4. Основное внимание в настоящем документе уделено доступу к большим данным, который крайне важен для официальной статистики как инструмент выпуска более точных статистических данных при меньшей нагрузке на респондентов и который, как показывает опыт реализации проектов использования больших данных, является одним из основных препятствий.

II. Подход статистических управлений

5. С точки зрения ФСУГ, данные частных предприятий, которые, как правило, накапливаются как побочный продукт их основной деятельности, должны быть доступны для общественных целей, т. е. для целей официальной статистики. Ввиду этого необходим постоянный доступ к данным частных предприятий вне зависимости от их желания или ситуации на рынке. В этой связи ФСУГ стремится создать юридическую базу для получения доступа к частным источникам данных, опираясь на которую можно строить дальнейшее сотрудничество.

6. Что касается правового регулирования, то ФСУГ и Государственное управление информации и технологий Северного Рейна-Вестфалии (IT.NRW) поручили подготовить доклад со сравнительным юридическим анализом вопроса о доступе к большим данным для общественных целей в Европе. Таким образом, можно будет гарантировать, что германские официальные статистические учреждения выпускают статистические данные на тех же условиях, что и их партнеры в рамках ЕСС. Кроме того, в докладе предполагается рассмотреть вопрос о том, как следует изменить правовые рамки германской федеральной статистики, чтобы статистические ведомства в Германии могли использовать большие данные.

7. На европейском уровне¹ Евростат в 2016 году на договорной основе поручил проанализировать юридическую ситуацию в вопросах доступа, использования и распространения статистических данных, полученных на основе больших данных, во всех государствах – членах Европейского союза. На данный момент в ЕС существуют две опорных точки в вопросах доступа к данным частных владельцев. Во-первых, это французский «цифровой закон»², обязывающий компании – при определенных обстоятельствах – открывать данные, подпадающие под определение «данных, представляющих общественный интерес», например для Национального института статистики и экономических исследований Франции (НИСЭИ). Во-вторых, это закон 2017 года о цифровой экономике Соединенного Королевства³ (см. часть 5, главу 7, раздел 80, новый раздел 45D), согласно которому Национальная статистическая служба Соединенного Королевства будет иметь доступ к данным крупных и средних предприятий и благотворительных организаций при исполнении одной или нескольких своих уставных функций, т. е. подготовки официальной статистики и статистических исследований.

8. Кроме того, в 2016 году Глобальная рабочая группа по использованию больших данных для целей официальной статистики опубликовала «Рекомендации в отношении доступа к данным частных организаций для официальной статистики»⁴, которые могут использоваться в качестве руководства для переговоров с частными предприятиями, располагающими полезными для официальной статистики источниками больших данных.

9. К тому же в апреле 2018 года Европейская комиссия опубликовала сообщение «К единому европейскому пространству данных»⁵ и методический документ по вопросу о распространении данных частного сектора в европейской экономике

¹ Относительно Целевой группы по большим данным см. неопубликованную версию документа Евростата «Legal initiatives for data access» («Правовые инициативы в области доступа к данным»), апрель 2017 года.

² См. на веб-сайте «Лежифранс» <https://www.republique-numerique.fr/pages/in-english> и <https://www.republique-numerique.fr/pages/digital-republic-bill-rationale> (обе страницы посещались 18 мая 2018 года).

³ См. на сайте национального архива Соединенного Королевства <http://www.legislation.gov.uk/ukpga/2017/30/contents/enacted> и http://www.legislation.gov.uk/ukpga/2017/30/pdfs/ukpgaen_20170030_en.pdf (обе страницы посещались 18 мая 2018 года).

⁴ См. на сайте Статистического отдела ООН <https://unstats.un.org/unsd/bigdata/conferences/2016/gwg/Item%20%20%28i%29%20a%20-%20Recommendations%20for%20access%20to%20data%20from%20private%20organizations%20for%20official%20statistics%20Draft%2014%20July%202016.pdf> (сайт посещался 18 мая 2018 года).

⁵ См. на сайте Европейской комиссии: <https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space> (сайт посещался 18 мая 2018 года).

данных⁶, которые являются частью стратегии единого цифрового рынка и касаются обмена данными между предприятиями и государственным сектором. Европейская комиссия поддерживает идею расширения доступа к частным данным для таких государственных органов, как национальные статистические институты, учитывая в то же время интересы частных предприятий, например в плане компенсации произведенных компаниями вложений в целях адаптации данных.

10. В октябре 2018 года Комитет Европейской статистической системы (КЕСС) принял Бухарестский меморандум и признал «необходимость дальнейшего развития правовой базы на европейском и национальном уровнях в интересах уменьшения препятствий для обеспечения доступности, использования и интеграции неоднородных данных в целях устойчивого производства достоверной "умной" статистики».

11. В целом, Общий регламент Европейского союза по защите данных (ОРЗД) призван повысить доверие к цифровым услугам и степень их защищенности, определяя общие стандарты защиты данных и обеспечивая гарантии основного права на информационное самоопределение. ОРЗД применяется во всех государствах-членах с 25 мая 2018 года с целью гармонизации законодательства о конфиденциальности данных в масштабах всей Европы.

12. Частичная обработка данных частными предприятиями может также упростить соблюдение законодательства о защите данных. В большинстве случаев статистических агрегированных промежуточных продуктов без личной информации достаточно для целей официальной статистики, и, имея дело с такими продуктами, специалистам, занимающимся вопросами защиты данных, будет легче дать согласие на использование в таких целях больших данных. Для использования статистических продуктов, предварительно обработанных частными предприятиями, нужно разработать четкие и обязательные стандарты. На европейском уровне этот процесс уже начался с появлением так называемой «достоверной умной статистики». В этом отношении статистические промежуточные продукты третьих сторон должны отвечать качественным, этическим и методологическим стандартам, чтобы их можно было сочетать с данными обследований и административными данными и использовать для выпуска официальной статистики.

III. Возможный вариант решения в Германии

13. Исходя из сегодняшних реалий, собственно большие данные должны и впредь быть бесплатными для целей официальной статистики, независимо от вида данных. Другими словами, большие данные должны поступать от респондентов без возмещения расходов.

14. Законодательство в этой области могло бы пойти по пути уже существующих законодательных актов об использовании административных данных. Для оценки приемлемости административные органы должны по запросу передавать статистическим управлениям информацию о происхождении, структуре, содержании данных и другие метаданные. Затем, если то или иное федеральное министерство придет к выводу, что следует глубже изучить вопрос, административные органы также должны направить во ФСУГ обезличенные персональные данные (т. е. без указания имен, фамилий и адресов). В случае установления приемлемости данных в законодательстве следует предусмотреть конкретные положения, предписывающие выпуск определенной статистики.

15. Кроме того, необходимо будет решить, будут ли статистические органы Германии пользоваться услугами частных предприятий для обработки данных на возмездной основе. Тот факт, что обработка данных частными предприятиями в некоторой степени будет более эффективной с точки зрения затрат, указывает на то,

⁶ См. на сайте Европейской комиссии <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy> (сайт посещался 18 мая 2018 года).

что оптимальным вариантом, вероятно, является участие статистических органов в оплате расходов или сотрудничестве между частными предприятиями и статистическими органами. В случае сотрудничества между частными предприятиями и статистическими органами следует на основе юридически обязательных прозрачных процедур обеспечить соблюдение принципа нейтральности официальной статистики.

IV. Дополнительные (правовые) вопросы

16. Помимо доступа к большим данным, нужно будет обсудить и другие правовые вопросы, например вопрос о необходимости специального законодательства о защите данных, действующего в отношении данных, передаваемых частными предприятиями для целей официальной статистики. В этой связи нужно также изучить возможность возникновения недобросовестной конкуренции вследствие использования результатов или продуктов сотрудничества с частными предприятиями, а также пути защиты конфиденциальной коммерческой информации.

17. Возникают вопросы и по поводу прав собственности на большие данные, например в секторе сельского хозяйства, причем особенно трудно отвечать на них в случае данных, генерируемых машинами. Кроме того, потребуют решения этические вопросы, например вопрос о том, влияет ли использование больших данных на профессиональные стандарты официальной статистики.

18. Весь потенциал больших данных, вероятно, проявится лишь при их сочетании с данными обследований и административными данными. Ввиду этого законодателям следует также заложить основы для необходимых законоположений в отношении сочетания различных источников данных, особенно в отношении личных данных.

19. В настоящее время данные официальной статистики предоставляются для использования в научных целях научно-исследовательскими центрами данных Федерального статистического управления и статистических управлений земель с учетом правовых требований подраздела 2 статьи 16 Закона о статистике в федеральных целях. В будущем нужно будет обеспечить возможность использования данных в научных целях и в случае больших данных. Кроме того, результаты официальной статистики должны быть открыты для научной оценки.

V. Заключение

20. Постоянный доступ к большим данным крайне важен для выпуска официальной статистики на регулярной основе. Он не должен зависеть ни от существующей ситуации на рынках, ни от готовности частных предприятий делиться своими данными.

21. ФСУГ убеждено, что в целях поддержания устойчивости официальной статистики, особенно на этапе всеобщей цифровизации, необходимо пересмотреть правовые рамки немецкой федеральной статистики в вопросах доступа к частным данным.