

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

Париж, 26–28 июня 2019 года

Пункт 2 b) предварительной повестки дня

Новые источники данных – доступность и использование**Заседание 2: навыки, необходимые для использования
новых источников данных****Повышение степени интеграции данных с помощью
глобальной системы геопространственных
статистических данных****Записка Статистического управления Швеции***Резюме*

Значение вопросов геопространственной информации и ее интеграции в статистическое производство возрастает, поскольку национальные статистические управления учитывают потребности пользователей в статистических данных о доступности статистической информации в разрезе города и сельской местности или статистики на местном уровне в целом.

Статистическое управление Швеции провело на национальном уровне анализ глобальной системы геопространственных статистических данных – системы, разработанной в рамках Инициативы Организации Объединенных Наций по вопросам управления глобальной геопространственной информацией и одобренной Статистической комиссией Организации Объединенных Наций. Объектом анализа стали прежде всего пять принципов данной системы с точки зрения имеющихся возможностей, при этом опорными элементами способности интегрировать статистическую и геопространственную информацию выступают люди, системы, процессы, информация, методы, стандарты и рамки, а также институциональные механизмы.

Основное внимание в этом документе будет уделено путям расширения странами своих возможностей интеграции данных на основе использования глобальной системы в сочетании с глобальной моделью – Единой архитектурой статистического производства. Проведенный анализ помогает Статистическому управлению Швеции выявлять области для улучшений в связи с различными элементами возможностей. Он также помог выяснить, в каких областях положение является достаточно удовлетворительным.



Настоящий документ представляется для обсуждения участникам проводимого в рамках Конференции европейских статистиков 2019 года семинара на тему «Новые источники данных – доступность и использование», заседание 2 «Навыки, необходимые для использования новых источников данных».

I. Введение

1. Исходя из территориальной политики на национальном и европейском уровнях, а также Повестки дня в области устойчивого развития на период до 2030 года¹, национальные статистические управления (НСУ) европейских стран стремятся выпускать актуальную и своевременную статистику². Повестка дня на период до 2030 года призвана добиться, чтобы никто не был забыт; с географической точки зрения это означает, например, необходимость учета городской/сельской проблематики и вопросов доступности.
2. Сейчас как никогда важно соотнести официальную статистику с теми реалиями, которые привычны, например, населению бедных районов. Кроме того, статистические данные местного уровня нужны для управления рисками стихийных бедствий и принятия мер по улучшению доступности общественного транспорта, школ, медицинских учреждений или «зеленых зон». Европейская статистическая система³ признает эту необходимость, финансируя проекты ГЕОСТАТ⁴, например ГЕОСТАТ 2, тем самым ориентируя НСУ в базовых вопросах более гибкой точечной статистической инфраструктуры⁵.
3. Региональные статистические данные регулярно выпускаются всеми странами Европейского союза, как правило, с использованием классификации административных единиц во всех государствах-членах, предусмотренной в Номенклатуре территориальных единиц для статистики (НТЕС)⁶. В базе данных Евростата содержится унифицированная региональная статистика, но в тех случаях, когда пользователям нужны данные с более подробной географической детализацией, единообразия отнюдь не наблюдается. В качестве первоначального шага для улучшения положения стали использовать статистические координатные сетки, классификации по степени урбанизации и т. д., наряду с классификацией местных административных единиц в качестве новых географических элементов данных согласно пересмотренному регламенту ЕС по территориальным типологиям (Tercet)⁷.
4. Статистические координатные сетки, по всей вероятности, станут одной из областей распространения полученных результатов при проведении следующего раунда европейских переписей населения и жилищного фонда в 2021 году⁸. Выпуск статистических материалов на базе таких сеток требует модернизации методов проведения переписей в странах, прокладывая путь к созданию реестровой статистической системы.

II. Геопривязка населения и предприятий

A. Преимущества статистической геопространственной интеграции

5. Преимущества статистической геопространственной интеграции начинают получать широкое признание, и Комитет ООН по вопросам управления глобальной геопространственной информацией создал Группу экспертов по интеграции статистической и геопространственной информации (ГЭ-ИСГИ)⁹. Статистическое управление Швеции принимает активное участие в работе этой группы экспертов, а также координирует осуществление проекта ГЕОСТАТ 3 «Сеть ЕСС», направленного на внедрение в Европе глобальной системы геопространственных статистических данных (ГСГФ)¹⁰. Проект «Сеть ЕСС» представляет собой сеть нескольких организаций ЕСС, созданную с целью получения результатов, которые будут полезны всей Европейской статистической системе (ЕСС).

6. ГСГФ была принята в 2016 году и непосредственно предусматривает интеграцию статистической и геопространственной информации¹¹. Как показано ниже в настоящем документе, Статистическое управление Швеции приступило к внедрению ГСГФ на национальном уровне, и система зарекомендовала себя как полезный инструмент планирования и выявления улучшений в этом сегменте статистики.

7. ГСГФ помогает обеспечить дополнительную геопривязку статистических данных о физических лицах и предприятиях. Она также облегчает организацию управления геопространственными данными и процессов, необходимых для получения высококачественных статистических материалов.

Диаграмма 1

Увязка общества, экономики и окружающей среды на основе интеграции данных¹²



8. Кроме того, ГСГФ служит связующим звеном между официальным статистическим ведомством и специалистами в области геопространственных данных. Статистические единицы геокодируются с использованием геопространственных идентификаторов, таких как кадастровый участок, адрес или здание. Затем данные агрегируются и публикуются для географических границ на соответствующем граничном (или географическом) уровне. Это позволяет обеспечить интеграцию данных из различных источников и служит связующим звеном между статистическими и более традиционными наборами геопространственных данных благодаря использованию возможностей обычной географии и большей гибкости географических элементов данных.

В. Пять принципов глобальной системы геопространственных статистических данных

9. ГСГФ представляет собой высокоуровневую систему, другими словами она не предназначена для составления подробного плана или схемы осуществления; скорее, она ориентирует относительно тех возможностей, которые должны иметься в странах, обеспечивая довольно значительную свободу маневра в практических действиях.

10. Для облегчения процесса внедрения данной системы в странах в 2018 году была разработана вспомогательная документация как на глобальном уровне (ГЭ-ИСГИ), так и в Европе (в рамках проекта ГЕОСТАТ 3). Руководство по осуществлению ГЕОСТАТ 3 дает неплохое представление о том, как следует понимать эти принципы с европейской точки зрения¹³.

1. Принцип 1: использование фундаментальной геопространственной инфраструктуры и геокодирования

11. Целью этого принципа является получение высококачественных и стандартизированных атрибутов местонахождения или опорных координат. В рамках данной системы рекомендуется использовать существующую геопространственную информацию, предоставляемую специализированными геопространственными органами (например, физический адрес, кадастровый участок, здание или какие-либо другие элементы описания соответствующей точки), для привязки каждой статистической единицы (например, на уровне микроданных или единичной записи) с указанием точных координат и/или небольшого географического участка или со стандартной привязкой к координатной сетке карты.

2. Принцип 2: геокодированные данные единичных записей в среде управления данными

12. В ГСГФ рекомендуется, чтобы увязка геокода с каждой записью статистической единицы в наборе данных (лицо, домохозяйство, предприятие, здание или участок земли) производилась в среде управления данными. База данных для хранения высокоточных географических кодов с системой контроля версий позволяет работать в любом географическом контексте при подготовке данных, которые будут публиковаться в будущем (т. е. при агрегировании данных с получением целого ряда более крупных географических единиц или в целях адаптации с учетом происходящих со временем изменений географических параметров).

3. Принцип 3: единые географические параметры для распространения статистических данных

13. Для обеспечения возможности сопоставления наборов данных из различных источников в ГСГФ рекомендуется использовать единый набор географических параметров для отображения, представления и анализа социальной, экономической и экологической информации.

4. Принцип 4: Статистическая и геопространственная функциональная совместимость – данные, стандарты и процессы

14. Специалисты, занимающиеся соответственно статистическими и геопространственными данными, используют свои собственные общие модели данных и возможности обработки метаданных; однако не всегда они применяются универсально. В статистике используются типовая модель статистической информации (ТМСИ), стандарты обмена статистическими данными и метаданными (ОСДМ) и механизмы Инициативы в области документирования данных. В сфере же геопространственной информации используются другие стандарты¹⁴.

15. В сфере статистики ощущается необходимость более последовательного включения геопространственных процедур и стандартов в статистические рабочие процессы. ЕЭК ООН в 2018 году пересмотрела типовую модель производства статистической информации (ТМПСИ)¹⁵, более подробно описав в процессе подготовки статистических данных порядок использования геопространственных данных и соответствующие методы.

5. Принцип 5: доступные и пригодные для использования геопространственные статистические данные

16. В этом принципе ГСГФ подчеркивается необходимость определения или, при необходимости, разработки политики, стандартов и руководящих принципов для распространения, обеспечения доступности, анализа и визуализации геопространственной информации. Одним из важных аспектов этого принципа является обеспечение доступа к данным с использованием безопасных механизмов, которые не только защищают частную жизнь и конфиденциальность, но и обеспечивают доступ к данным для проведения различной аналитической работы в интересах принятия решений.

III. Возможности для интеграции статистической и геопространственной информации

A. Анализ возможностей Статистического управления Швеции

17. Наличие соответствующих возможностей позволяет статистической организации осуществлять один или несколько видов деятельности. Согласно Единой архитектуре статистического производства для этого, в принципе, нужны люди, процессы, методы, технологии, стандарты и соответствующая рамочная структура¹⁶. Статистическое управление Швеции, прорабатывая возможности применения ГСГФ в Европе, в рамках проекта ГЕОСТАТ 3 проанализировало на национальном уровне все принципы этой системы.

Таблица 1

Элементы потенциала

Элементы потенциала

| | |
|------------------------------------|--|
| Кадры | Наличие у сотрудников квалификации, знаний и навыков, необходимых для проведения определенного вида деятельности. |
| Системы | Приложения, аппаратные средства и платформы на базе ИКТ, требующиеся организации для проведения определенного вида деятельности. |
| Процедуры | Предпочтительный комплекс мер или задач предприятий – с указанием в необходимых случаях соответствующих методов, стандартов и систем – для эффективного и результативного осуществления определенного вида деятельности. |
| Методы | Комплекс конкретных методов или алгоритмов, необходимых для проведения определенного вида деятельности. |
| Стандарты и механизмы | Стандарты, механизмы, руководства и стратегии, необходимые для проведения определенного вида деятельности. |
| Информация | Данные, необходимые для проведения деятельности в области интеграции статистической и геопространственной информации. |
| Институциональные механизмы | Правовая и институциональная инфраструктура, необходимая для проведения и поддержки определенного вида деятельности. |

18. Анализ системы с точки зрения наличия возможностей дал полезные результаты, поскольку рамочные принципы должны подкрепляться определенными навыками, процессами, методами и т. д. Он показал, что в случае интеграции статистической и геопространственной информации, элементы, составляющие основу потенциала, можно расширить, включив дополнительно информацию и институциональные механизмы¹⁷.

Диаграмма 2

Элементы, необходимые интеграции статистической и геопространственной информации в НСУ



19. Такой анализ вполне по силам НСУ, недавно приступившим к работе по интеграции статистической и геопространственной информации, а также тем статистическим службам, которые ищут пути совершенствования такой работы. В докладе Швеции предлагаются простые меры, при разработке которых в качестве отправной точки берутся потребности пользователей.

20. Анализ каждого из элементов, лежащих в основе потенциала по каждому принципу рассматриваемой системы, выявил сильные и слабые стороны Статистического управления Швеции. Благодаря постановке целей и принятию решений о приоритетности тех или иных видов деятельности легче предусматривать мероприятия в рамках годовых планов и контролировать их осуществление. Эти мероприятия могут относиться к категории новых видов деятельности, к реализации которых следует приступить, или к тем видам деятельности, где нужно улучшить работу. Кроме того, по результатам проведенного анализа были выявлены виды деятельности, которые следует сохранить или сворачивать.

В. Необходимо выходить на новые направления деятельности

21. Статистическое управление Швеции определило девять мер, которые позволят укрепить потенциал в области интеграции статистической и геопространственной информации. Например, вынесение рекомендаций в отношении методов работы с геопространственными данными и геокодирования по типу онлайн-сервиса документации Австралийского бюро статистики (АБС)¹⁸. Это поможет упорядочить внутренние процессы в Статистическом управлении Швеции, а также станет подспорьем для других органов, отвечающих за официальную статистику. К тому же нам нужен системный подход к интеграции геопространственных данных всех реестров с повышением в результате точности таких данных. В перспективе это позволит экономить время в процессе деятельности, поскольку не нужно будет дополнительно повышать качество.

22. Важной мерой является предоставление всей статистико-географической информации в качестве открытых данных. Кроме того, нужно встроить геопространственные данные и метаданные в архитектуру статистического производства Статистического управления Швеции. Это позволит обеспечить более полную интеграцию производства статистической информации, управления данными, метаданных и инфраструктуры.

23. Кроме того, необходимо определить конкретные целевые показатели желательного качества географических данных в реестрах (например, доля успешно геокодированных рабочих мест в реестре предприятий). Нам также нужно

организовать сервисы интерактивных карт на сайте scb.se и выпускать в качестве открытых данных статистико-географическую информацию о малых районах.

C. Необходимость совершенствования методов работы

24. В «категории улучшений» было предусмотрено 14 мер с ориентацией на те направления, по которым уже ведется работа, но где можно добиться большего для укрепления нашего потенциала в области интеграции статистической и геопространственной информации.

25. Нам необходимо глубже прорабатывать замечания и предложения, направляемые Статистическим управлением Швеции в национальное агентство геопространственной информации «Лантметериет» по поводу качества получаемых данных. Кроме того, нам следует шире использовать согласованную терминологию и заявления о качестве данных. Полезно было бы активнее использовать SQL-сервер для пространственных операций. Необходимо повысить качество информации об адресах в реестре предприятий, чтобы можно было надлежащим образом произвести географическую привязку всех рабочих мест.

26. Необходимо лучше использовать геопространственную информацию в рамках всего производственного процесса наряду с расширением поддержки, оказываемой специалистами по ГИС остальным подразделениям организации. В порядке примера можно назвать организацию учебно-практических курсов и семинаров по методике использования геопространственных данных в статистическом анализе.

27. Кроме того, нам следует с помощью веб-службы повысить доступность всей административной и статистической географической информации, используемой любыми учреждениями в рамках национальной статистической системы, в качестве открытых данных. По мере возможности следует также предоставлять данные за предыдущие периоды, чтобы обеспечить надлежащее наложение статистических данных на точно определенные географические районы.

28. Нам также необходимо повысить эффективность методов обеспечения конфиденциальности, чтобы более гибким образом учитывать, как пользователи загружают и объединяют различные наборы данных (географическое дифференцирование). Следует также предусмотреть ввод в базу статистических данных географических карт (в дополнение к существующей в настоящее время возможности ввода данных в табличной форме).

D. Сохранение нормально работающих структур

29. Анализ также показывает, какие структуры функционируют нормально. Статистическое управление Швеции успешно сотрудничает с национальным агентством геопространственной информации «Лантметериет» и другими поставщиками геопространственных данных. Реестровая статистическая система с источниками административных данных является основой для интеграции.

30. Веб-служба РЕГИНА¹⁹ позволяет существенно повысить доступность данных об изменении административного деления за прошлые периоды. Веб-служба работает лучше с появлением возможности визуального просмотра, но предстоит еще большая работа по включению дополнительных данных об административном делении за прошлые периоды.

31. В работе по повышению эффективности наших веб-служб весьма полезен экспертный опыт по вопросам функционирования сервисов поиска, просмотра и скачивания информации в рамках Инфраструктуры пространственной информации для Европейского сообщества (ИНПРОИФЕ).

32. Действуют нормально функционирующие автоматические процессы обеспечения конфиденциальности, которые, снижая риск раскрытия информации, обуславливают эффективность производства статистики малых районов.

В Статистическом управлении Швеции создан портал для предоставления статистических данных по запросу, который является системой распространения продукции функционирующих сервисов в защищенном режиме.

Е. Борьба с неэффективностью

33. В процессе анализа элементов потенциала было очевидно, что многого можно добиться путем сокращения неэффективных процедур проверки качества и элементов дублирования работы. Следует избегать обременительной индивидуальной корректировки данных и контроля качества на завершающих этапах производственного процесса.

34. Встречаются также случаи неэффективного хранения данных и дублирования баз данных. В Статистическом управлении Швеции более не хранятся предыдущие версии геопространственной информации. Они хранятся в национальном архиве данных, что усложняет производственный процесс. Ненужная путаница существует и в вопросе различных вариантов географических параметров (масштаб, изменение границ и т. д.), которые нужно сокращать.

35. Переход к более открытым данным снизит необходимость платных услуг, в результате чего статистика малых районов станет более доступной и будет чаще использоваться в обществе.

III. Выводы

36. По-прежнему бытует мнение, что геопространственная информация касается в основном небольшой группы экспертов Статистического управления Швеции, занимающихся вопросами ГИС. Данный анализ призван выявить те области, где нужно реализовывать новые инициативы и улучшать работу, чтобы добиться более широкого использования геопространственных данных и процессов в рамках всей организации.

37. По каждому принципу в отдельности были проанализированы элементы потенциала и на этой основе разработаны предложения в отношении того, где нужно начинать с нуля, где улучшать работу, где продолжать ее, а где и сворачивать осуществляемую деятельность. Мероприятия, круг которых был определен при проведении анализа, включаются в годовые планы работы; предприняты важные шаги в направлении повышения открытости данных. Вместе с тем предстоит проделать еще большую работу для достижения цели обеспечения интеграции геопространственных данных и процессов и их использования в рамках всей организации.

38. Мы настойчиво рекомендуем провести такую оценку всем НСУ, стремящимся расширить свои возможности в области геопространственных данных. Внедряя систему геопространственных статистических данных с акцентом на соответствующие элементы потенциала, НСУ может определять приоритетные направления инвестиций в развитие необходимых навыков, процессов и инструментальных средств.

IV. Примечания

- ¹ Генеральная Ассамблея Организации Объединенных Наций. Резолюция 70/1. «Преобразование нашего мира: Повестка дня в области устойчивого развития на период до 2030 года».
- ² Haldorson M., Zaccheddu P., Fohgrub B., Petri E. (2016). Geospatial information management in Europe – responding to the user needs, Statistical Journal of the IAOS 32.
- ³ European Statistical System, <http://ec.europa.eu/eurostat/web/european-statistical-system/overview?locale=fr>.
- ⁴ GEOSTAT Project Descriptions, <http://www.efgs.info/geostat/>.
- ⁵ Moström J. (editor) et al (2017). A Point-based Foundation for Statistics, Final report from the GEOSTAT 2 project.

-
- ⁶ NUTS Regulation (2003). Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003, on the establishment of a common classification of territorial units for statistics (NUTS), OJ L 154, 21.6.2003.
 - ⁷ Tercet – территориальные типологии Европейского союза.
<http://ec.europa.eu/eurostat/web/nuts/tercet-territorial-typologies>.
 - ⁸ European Census, <https://ec.europa.eu/eurostat/web/population-and-housing-census/overview>.
 - ⁹ UN-GGIM Expert Group on Integration of Statistical and Geospatial Information,
<http://ggim.un.org/UNGGIM-expert-group/>.
 - ¹⁰ Проект ГЕОСТАТ 3, <http://www.efgs.info/geostat/geostat-3/>.
 - ¹¹ UN EG-ISGI Background Document (2016). United Nations Expert Group on the Integration of Statistical and Geospatial Information: Global Statistical Geospatial Framework: Linking Statistics and Place. Current status and plans for development, July 2018.
 - ¹² Иллюстративный пример, подготовленный статистическими управлениями Австралии и Новой Зеландии.
 - ¹³ GSGF Europe - Implementation guide for the Global Statistical Geospatial Framework in Europe, Proposal from the GEOSTAT 3 project, Version 1.0, 28 February 2019.
 - ¹⁴ Документация рабочего совещания ЕЭК ООН по статистическим и геопространственным стандартам, Стокгольм, ноябрь 2017 года, <https://www.unece.org/index.php?id=45404>.
 - ¹⁵ Типовая модель производства статистической информации. ТМПСИ,
<http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0>.
 - ¹⁶ Common Statistical Production Architecture, CSPA,
<https://statswiki.unece.org/display/CSPA/Common+Statistical+Production+Architecture>.
 - ¹⁷ Haldorson M., Moström J. (2018): Implementing the Statistical Geospatial Framework at Statistics Sweden. National Report as part of the GEOSTAT 3 project.
 - ¹⁸ Веб-сайт АБС по вопросам статистической географии, <http://www.abs.gov.au/geography>.
 - ¹⁹ Веб-служба РЕГИНА Статистического управления Швеции, информация имеется на английском языке в документе в формате PDF <http://regina.scb.se/>.
-