Организация Объединенных Наций ECE/CES/2023/2

Экономический и Социальный Совет

Distr.: General
13 April 2023
Russian
Original: English

Европейская экономическая комиссия

Конференция европейских статистиков
Семьдесят первая пленарная сессия
Женева, 22–23 июня 2023 года
Пункт 5 b) предварительной повестки дня
Доклады, руководящие принципы и рекомендации, подготовленные под эгидой Конференции
Курирование данных

Курирование данных и роль национальных статистических управлений в новой экосистеме данных

Подготовлено Целевой группой Конференции европейских статистиков

Резюме

Настоящий документ является выдержкой из документа «Курирование данных и роль национальных статистических управлений (НСУ) в новой экосистеме данных», подготовленного специальной Целевой группой в составе Австрии, Бельгии, Венгрии, Дании, Ирландии, Казахстана, Канады, Мексики, Нидерландов (Королевство), Новой Зеландии, Польши, Соединенного Королевства, Финляндии, Швейцарии, Эстонии (Председатель), Банка международных расчетов (БМР), Евростата, Организации экономического сотрудничества и развития (ОЭСР) и ЕЭК ООН. Члены группы являются представителями НСУ и других государственных учреждений, центральных банков и геопространственных агентств.

Первый проект доклада согласовывался в электронном виде со всеми членами КЕС весной 2022 года. Полученные отзывы состояли из множества конструктивных и полезных замечаний по совершенствованию этого документа. Целевая группа обновила документ с учетом замечаний и последних изменений, а также подготовила два дополнительных документа: 1) резюме доклада, которое можно использовать как самостоятельный документ для разъяснения роли НСУ в курировании данных, и 2) документ с примерами того, как НСУ разных стран выполняют свою роль кураторов данных. Бюро КЕС рассмотрело проект доклада в феврале 2023 года и поручило секретариату направить его для проведения электронных консультаций всем членам КЕС до начала пленарной сессии КЕС 2023 года.

Настоящий документ представляет собой краткую версию вышеупомянутого доклада, подготовленную для целей перевода. Полный текст, включающий также главы о моделях руководства данными в государственном секторе, моделях оценки степени зрелости в области курирования данных, коммуникации в области курирования данных и выводы и приложения, а также вышеупомянутые дополнительные документы были разосланы для ознакомления всем членам КЕС в апреле 2023 года и доступны на веб-странице Конференции

Просьба отправить на вторичную переработку

по URL: https://unece.org/statistics/events/CES2023. Резюме отзывов участников консультаций будет представлено в документе ECE/CES/2023/3/Add.1.

При условии получения положительного отклика в рамках консультаций КЕС будет предложено одобрить данный документ.

I. Резюме

1. Быстро расширяющийся и развивающийся ландшафт производства данных и обмена ими предоставляет Национальным статистическим управлениям (НСУ) возможность задействовать свой опыт в рамках более широкого статистического и национального сообщества работы с данными. В этих условиях для полноценного выполнения своего мандата роль НСУ должна расшириться за счет включения в нее курирования данных. Поскольку эта трансформация многогранна, существуют различные интерпретации того, как может и должна измениться роль НСУ и что означает курирование данных в этом контексте.

2. В настоящем докладе Целевая группа Европейской экономической комиссии Организации Объединенных Наций (ЕЭК ООН) (созданная Конференцией европейских статистиков в 2021 году) изучила вопрос курирования данных через призму НСУ. Курирование данных рассматривается как деятельность по обслуживанию экосистемы данных, а не по контролю над ней и отражает общую смену парадигмы или эволюцию определяющих деятельность НСУ рамок, которая представляет собой переход от производства статистики к предоставлению данных и связанных с данными услуг.

3. Работа этой целевой группы была сосредоточена на определении и разъяснении основополагающих концепций и рамок, относящихся к курированию данных. Позднее, в ходе последующей работы по итогам настоящего доклада, могут быть разработаны более конкретные указания и рекомендации. Целевая группа рассмотрела дальнейшие шаги по поддержке НСУ в выполнении их кураторской роли, которые кратко изложены в разделе I B «Будущая работа».

4. Цель доклада заключается в том, чтобы служить руководством для статистических управлений, независимо от степени зрелости данных или структуры. Он поможет управлениям определиться, стоит ли им расширять свою роль в качестве кураторов правительственных данных, а также послужит источником информации и идей для тех управлений, которые уже выполняют эту функцию. Кроме того, держатели общедоступных данных получат более глубокое понимание преимуществ и услуг, предоставляемых НСУ в рамках будущей экосистемы данных. Ниже приводится краткое резюме всего доклада, а затем краткое изложение содержащихся в докладе ключевых тезисов, в частности в отношении курирования данных.

A. Основные моменты доклада

5. Развитие сложных цепочек поставок данных, новые достижения в области технологий и возрастающая важность инфраструктуры обмена данными и доступа к ним диктуют потребность мониторинга и обслуживания. Это заставляет вновь сфокусироваться на общественном доверии и, следовательно, высвечивает важность курирования данных и руководства ими (глава 2 в полном тексте доклада). НСУ могут обеспечить компетенции и возможности, способствующие совершенствованию руководства данными в государственном секторе и за его пределами, учитывая при этом правовые, организационные и технические аспекты, определяющие роль НСУ в этой области.

6. Необходимы единообразные определения руководства данными и курирования данных, которые можно использовать в контексте официальной статистики и во всем государственном секторе (глава 3 в полном тексте доклада). Курирование данных — это особенно важная концепция в современных условиях работы с данными, прежде всего в том смысле, что она способствует целостному взгляду на управление данными, благодаря четкой подотчетности в руководстве данными, когда акцент делается на курировании нескольких поколений и общественном благе.

7. Курирование данных обеспечивается посредством четкого и компетентного руководства данными. Для организации деятельности по руководству данными в государственном секторе используются три типа моделей руководства данными: централизованное (сверху вниз), распределенное (снизу вверх) и

федеративное (гибридное) руководство данными (глава 4 в полном тексте доклада). Централизованная модель характеризуется консолидацией полномочий руководства данными, а распределенная модель — независимым руководством ими. Федеративная модель представляет собой гибридную модель с центральным органом, который обеспечивает получение информации от учреждений, и является наиболее подходящей моделью для общегосударственного подхода к курированию данных. Выбор и внедрение модели руководства данными в значительной степени зависят от национальных условий.

8. Успешное обеспечение курирования данных в значительной степени обусловлено пониманием людьми того, что такое курирование, и особенно того, как оно связано с их ролями и функциями — как внутри учреждения, так и на общесистемном уровне (глава 5 в полном тексте доклада). Роли и соответствующие компетенции куратора данных можно изучить на примерах и тематических исследованиях из разных стран, в которых имеется свой уникальный контекст и свои специфические извлеченные уроки.

9. Измерение эффективности курирования данных является относительно новой сферой, в связи с чем предлагается модель оценки степени зрелости в области курирования данных. Модель степени зрелости в области курирования данных может привнести важные идеи и ценности в понимание того, что представляет собой рациональное курирование данных, и, как следствие, в организационную структуру и практику и тем самым поможет учреждениям двигаться в правильном направлении в рамках экосистемы цифровых данных (глава 6 в полном тексте доклада). Модель степени зрелости в области курирования данных может помочь прояснить значение руководства данными и высветить те способы, с помощью которых оно должно осуществляться в сочетании с курированием данных, и способствовать его реализации. Представлена одна из таких возможных моделей оценки степени зрелости на основе примера, реализованного в Новой Зеландии.

10. Хотя курирование данных может представлять собой сложную для понимания концепцию, важность курирования данных и меняющиеся функции НСУ по его осуществлению могут быть эффективно раскрыты с помощью полезных продуктов, услуг и партнерств (глава 7 в полном тексте доклада).

11. Такая приверженность курированию данных и учету меняющейся ситуации в сфере данных может привести к появлению новых партнерств и источников данных, которые снизят нагрузку, связанную с представлением ответов, и повысят своевременность, качество, совместимость и возможность повторного использования данных, а также их пригодность для конкретных целей, способствуя предоставлению гражданам более полной информации и более качественных услуг (глава 8 в полном тексте доклада).

B. Ключевые тезисы, касающиеся курирования данных

12. Концепция курирования данных является основополагающей для внедрения передовой практики работы с данными (особенно в развивающейся экосистеме) и для будущей работы НСУ. В целях обеспечения ясности и последовательности в этом вопросе Целевая группа разработала следующее комплексное определение:

> «*Курирование данных означает обеспечение этичного и ответственного создания, сбора данных, управления ими, использования и повторного использования данных таким образом, чтобы они применялись на благо общества и приносили пользу всему сообществу пользователей данных*».

13. В докладе будут приведены следующие ключевые тезисы, касающиеся курирования данных.

14. Курирование данных необходимо для того, чтобы максимально увеличить ценность активов данных.

- Курирование позволяет реализовать политику в отношении данных и обращаться с данными как со стратегическим активом.
- Курирование способствует обмену активами данных и их повторному использованию при соблюдении надлежащих этических и культурных условий, тем самым максимально повышая их ценность.
- Курирование поощряет открытые данные, совместимость и этичное и культурно приемлемое использование данных, повышая своевременность и эффективность принятия решений и обеспечивая предоставление государственных услуг на основе данных.

15. Сфера охвата курирования данных подразделяется на две основные области: внутриведомственную и общесистемную.

- Внутреннее курирование данных может помочь последовательно распространять передовую практику работы с данными внутри учреждения, включая повышение качества данных и производство на основе метаданных.
- Общесистемное курирование данных предполагает сотрудничество и координацию между ведомствами, которые выходят за рамки различных областей данных. В силу этого достижение желаемой доступности и совместимости данных представляет собой многостороннюю работу.
- Не существует универсального подхода к курированию данных, и необходимо учитывать потребности отдельных НСУ.

16. НСУ обладают собственным уникальным опытом для руководства процессом курирования данных в Национальной статистической системе (НСС) и для выполнения обязанностей по курированию данных в рамках всей национальной экосистемы данных.

- Выполняя свою роль по производству качественной статистики, НСУ всегда участвовали в курировании данных. Этот уникальный опыт позволяет НСУ использовать свои наработки и функциональные знания для выполнения новых задач и расширения функций, связанных с курированием данных.
- В зависимости от национального контекста функции НСУ в области курирования данных могут охватывать обязанности по курированию данных в масштабах всей системы, связанные с управлением данными в государственном секторе.

17. Курирование данных повышает качество данных и релевантность НСУ и Национальной статистической системы.

- Курирование данных может способствовать предоставлению высококачественных данных во всех странах. Курирование данных приносит дополнительную пользу НСУ и НСС, учитывая национальный контекст, влияющий на его применение.

18. Курирование данных укрепляет доверие общественности к официальной статистике и к управлению данными в государственном секторе.

- Предоставление основанных на данных общественных благ и услуг требует рационального руководства данными, общественного доверия и надежности данных.
- Курирование данных способствует соблюдению Основополагающих принципов официальной статистики, которые обеспечивают конфиденциальность, высокое качество и доверие общественности.

19. Для успешного осуществления курирования данных необходимы руководство данными, возможности и ресурсы.

- Цели курирования данных должны быть встроены в систему руководства данными — систему прав принятия решений и распределения ответственности

за управление данными, а также в соответствующие законы, нормативные акты, стратегии и рамки, обеспечивающие исполнение.

- Курирование данных требует координации инвестиций и обеспечения возможностей в условиях прозрачного диалога с общественностью и принимающими бюджетные решения органами для обеспечения прогресса и необходимых инвестиций.
- Модель зрелости может быть использована для проведения самостоятельной оценки расхождений между текущим и желаемым положением дел в области курирования данных. Это позволяет определить подходящие структуры руководства данными и спланировать выделение ресурсов.

I. Введение

A. Справочная информация

1. Статистические управления работают во все более взаимосвязанном, динамичном и цифровизированном мире, который предъявляет новые требования и предоставляет новые возможности. Данные стали важнейшим стратегическим активом, и поэтому многие организации, не относящиеся к сфере официальной статистики, создают статистическую продукцию, используя новые источники данных, инструменты и методы. Это также изменило ожидания заинтересованных сторон: им требуются более открытые и подробные данные, которые можно быстро и легко получить. Пандемия COVID стала дополнительным испытанием способности официальной статистики ответственно подходить к сбору качественных и своевременных данных, ускорив модернизацию методов работы с данными в НСУ, государственных учреждениях и частных организациях.

2. Потенциал данных и информации, связанный с использованием партнерств и объединением данных из многочисленных источников, порождает вопросы, связанные с этикой, защитой конфиденциальности, безопасностью данных, доступностью, гармонизацией концепций, форматов и используемых технологий, информированностью об имеющихся данных, а также с тем, каким образом будут осуществляться управление данными и руководство ими в долгосрочной перспективе. Чтобы ответить на эти вопросы, в настоящем докладе будет дано определение термина «курирование данных» и разъяснены связанные с ним функции и компетенции, позволяющие систематическим, устойчивым и ответственным образом обеспечивать доступ к данным и их повторное использование ради общественного блага (Verhulst, 2021).

3. Для того чтобы продвинуться к единому пониманию термина «курирование данных», Конференция европейских статистиков (КЕС) создала в 2021 году Целевую группу, которая призвана помочь внести ясность в эту область и облегчить работу НСУ по приспособлению их функций к меняющейся ситуации. Членами Целевой группы являются следующие страны и организации: Австрия, Бельгия, Венгрия, Дания, Ирландия, Казахстан, Канада, Мексика, Нидерланды (Королевство), Новая Зеландия, Польша, Соединенное Королевство, Финляндия, Швейцария, Эстония (Председатель), Банк международных расчетов (БМР), Евростат и Организация экономического сотрудничества и развития (ОЭСР). Члены Целевой группы представляют статистические управления и другие государственные учреждения, а ЕЭК ООН обеспечивает секретариатское обслуживание.

4. Черновая версия документа была распространена для консультаций весной 2022 года среди стран и организаций, участвующих в работе Конференции европейских статистиков (КЕС). В настоящий проект документа вошли обширные и разнообразные отзывы, а также информация о последних изменениях. Бюро КЕС рассмотрело его в феврале 2023 года и поручило секретариату направить его для проведения широких консультаций. При условии получения положительного отклика в рамках консультаций, пленарной сессии КЕС 2023 года будет предложено одобрить данный документ.

5. Цель доклада заключается в том, чтобы служить руководством для статистических управлений, независимо от степени зрелости данных или их структуры. Он поможет управлениям определиться, стоит ли им расширять свою роль в качестве кураторов правительственных данных, а также послужит источником информации и идей для тех управлений, которые уже выполняют эту функцию. Держатели общедоступных данных также получат более глубокое понимание преимуществ и услуг, предоставляемых НСУ в рамках будущей экосистемы данных.

6. Роль национальных статистических управлений в экосистеме данных в последние годы подверглась широкому обсуждению. Однако эти дискуссии значительно различаются по направленности, и не существует единого понимания того, что такое курирование данных и какова может (или должна) быть роль НСУ в

этом меняющемся контексте. В этой связи актуальное значение имеют следующие вопросы:

- Что такое курирование данных? Какое отношение оно имеет к руководству данными?
- Какую роль могут играть НСУ в курировании данных в государственном секторе, особенно в сотрудничестве с другими компонентами НСС или в более широком смысле в экосистеме данных страны?
- Как НСУ будут реагировать на конкуренцию со стороны других организаций и какие стратегические партнерства необходимы?
- Какие критерии используются НСУ при принятии решения о том, как управлять своей позицией в качестве экспертов по вопросам курирования данных в государственном секторе? В чем выражается их руководящая роль?

B. Будущая работа

7. Во время создания Целевой группы в феврале 2021 года было отмечено, что соответствующий доклад может послужить основой для разработки дальнейших указаний и рекомендаций относительно роли НСУ в курировании данных. В ходе международных дискуссий по этой теме страны просили дать практические рекомендации относительно того, как выполнять функции курирования данных.

8. С целью обеспечения поддержки практической реализации Целевая группа рассмотрела следующие возможные последующие шаги:

- разработка Основных принципов курирования данных путем мониторинга и анализа реализации стратегий в области данных и инициатив по руководству данными в НСУ;
- предоставление более конкретных указаний и/или рекомендаций и инструментов для осуществления курирования данных;
- разработка общей «дорожной карты» по расширению ведущей роли НСУ в курировании данных (за отправную точку следует взять подготовленный Эстонией и другими сторонами документ, представленный на пленарной сессии КЕС 2020 года);
- определение базового набора обязанностей НСУ как кураторов данных в национальных экосистемах данных;
- формирование базы знаний о передовой практике и примерах успешного осуществления курирования данных, включая более подробные примеры того, как осуществляется курирование данных и какие услуги по работе с данными предоставляются на практике;
- сбор примеров моделей зрелости и других оценок курирования данных и разработка и тестирование на этой основе общей модели зрелости для оценки курирования данных в НСУ;
- подготовка методических материалов для различных аудиторий или сосредоточение внимания на курировании данных в конкретных областях — административные данные, открытые данные, данные, находящиеся в частном владении, геопространственные данные, модели руководства данными и т. д.;
- предоставление форума для обмена опытом осуществления курирования данных;
- ведение глоссария (согласованного с деятельностью глобальной Рабочей группы по вопросам курирования данных Статистического отдела Организации Объединенных Наций) и содействие переводу ключевых терминов на другие языки.

9. При проведении любых последующих мероприятий важно учитывать другую глобальную и региональную работу по этой теме, включая вышеупомянутую глобальную Рабочую группу по вопросам курирования данных (РГКД) и ее пять направлений работы, а также деятельность, проводимую в регионах Экономической комиссии для Латинской Америки и Карибского бассейна (ЭКЛАК) и Экономической и социальной комиссии для Азии и Тихого океана (ЭСКАТО). Все предложения о последующих действиях будут представлены в Бюро КЕС в октябре 2023 года.

II. Национальные статистические управления и курирование данных

A. Контекст

10. В условиях экспоненциального роста объема данных, производимых во всем мире, и появления таких технологий, как искусственный интеллект (ИИ) и периферийные вычисления, государственные органы во всем мире все больше признают важность данных. В этом контексте становится крайне важным создание основ курирования данных и руководства ими, которые могут обеспечить доступ к данным и их совместное использование, одновременно защищая частный характер отдельных лиц и обеспечивая этичное использование данных.

11. Использование электронных устройств с помощью коммуникационных сетей позволило компаниям создавать обширные массивы данных, например, в сфере онлайн-торговли и развлекательных платформ. Эти потребительские данные в сочетании с другими (государственными или частными) данными открывают огромный потенциал для обнаружения закономерностей, формирования показателей, разработки бизнес-моделей и принятия решений для общего блага. Доступ к этим массивам данных в настоящее время ограничен из-за интенсивной рыночной конкуренции, нормативных положений, регулирующих конфиденциальность и конкуренцию, а также ввиду последствий, связанных с затратами на получение и предоставление информации. Низовое развитие совместных инициатив государственного и частного секторов позволило изучить преимущества и проблемы в области обмена информацией, которой владеют частные предприятия. Однако такой ситуативный подход ограничивает возможности получения периодической статистической информации для общественного пользования.

12. Цифровизация общества и экономики сделала доступ к данным и обмен ими ключевым фактором инноваций и общественного доверия. Использование административных реестров и экспериментальных методик позволяет получать данные из других источников и для различных целей. Такие изменения дают возможность дополнить статистику, которую традиционно производят НСУ, тем самым снижая административные расходы и нагрузку, связанную с представлением ответов.

13. Эти изменения являются источником как возможностей, так и проблем. Государственные учреждения и другие общественные институты располагают собственными информационными системами и коллекциями данных с богатой историей, что породило бесчисленные фонды данных, которые могут быть обогащены за счет альтернативных данных. Помимо снижения нагрузки, существует огромный потенциал для новых видов связанных с данными услуг, получения более своевременных и детализированных данных, формирования новых представлений путем объединения данных, связанных с различными секторами, источниками и темами. Вместе с тем существуют и огромные риски, которые необходимо минимизировать. Без надлежащего руководства данные могут быть использованы неэтично, рост цифровых возможностей влечет за собой увеличение потребности в цифровой безопасности, «цифровой разрыв» может стать «информационным разрывом», и, разумеется, в условиях таких изменений перед НСУ стоит задача оставаться бдительными и инициативными в вопросах защиты конфиденциальности. Кроме того, чтобы извлечь выгоду из этих изменений, государственные учреждения

должны будут решить проблемы, связанные с отсутствием доступности, совместимости и стандартизации данных; дублированием и избыточностью данных; и затратами на увязку данных и обмен ими.

14. Для преодоления этих рисков необходима государственная политика с более дальновидным подходом, позволяющая сосредоточиться на таких аспектах, как стратегия, культура, этика, роли и возможности людей, которые могут поддержать эффективную экосистему данных. Концепция руководства данными требует общих правил и стандартов, обеспечивающих возможность взаимодействия. Для этого на уровне всего правительства необходимы координация, признанные нормы и утвержденные стандарты. К счастью, в рамках НСС была внедрена устойчивая культура скоординированного сотрудничества, причем немаловажную роль в этом сыграло применение Основополагающих принципов.

15. Руководство данными позволяет внедрить и сделать доступным такой набор норм и стандартов, связанных с данными. Согласованные нормы управления данными обеспечат ясность и гарантии в том, что касается надзора за данными и руководства ими, а также будут способствовать доступности, качеству, безопасности, удобству использования и достоверности данных. Если руководство данными предполагает уделение особого внимания техническим компетенциям, то курирование данных в равной степени фокусируется на основополагающих концепциях, рамочных основах и передовой практике для обеспечения более зрелой политики в отношении данных.

16. В рамках курирования данных реализуются стратегии, стандарты и принципы, предусмотренные системой руководства данными. Оно включает в себя функции и компетенции, позволяющие обеспечить доступ к данным и их повторное использование в общественных интересах систематическим, устойчивым и ответственным образом. Курирование данных — это работа на благо экосистемы данных, а не контроль над ней.

17. Национальные правительства предпринимают шаги, направленные на то, чтобы ослабить политическую разобщенность и повысить согласованность стратегий, проектов и инициатив в области данных в государственном секторе (OECD, 2019). Такие страны, как Соединенные Штаты Америки, Великобритания, Нидерланды (Королевство) и Германия, разработали национальные стратегии в области данных, которые охватывают в рамках единого политического инструмента различные элементы экосистемы данных, включая доступ к данным и обмен данными, открытые данные, данные для искусственного интеллекта и этику данных в государственном секторе. На международном уровне предпринимаются усилия по содействию обмену данными и доступу к ним, такие как инициатива Группы двадцати, направленная на устранение пробелов в данных (ИПД). Ее новый этап, начавшийся в 2022 году, охватывает, среди прочих рекомендаций, доступ к частным и административным источникам информации и обмен данными (FSB Secretariat and IMF Staff, 2022).

18. В других случаях политические инструменты, связанные с данными, становятся частью более масштабных стратегий в области цифровизации. Так, данные Индекса цифрового правительства ОЭСР за 2019 год показывают, что «отдельная специальная политика (или стратегия) в области данных имеется только у 12 % стран, при этом 82 % стран рассматривают данные в рамках более масштабных смежных стратегий (например, в сфере цифрового правительства или открытых данных)» (OECD, 2020).

19. Что касается институционального руководства данными и создания фиксированных руководящих ролей, то некоторые страны предпочли назначить конкретные должности главных специалистов по данным (ГСД) для всего правительства (например, назначение главного специалиста по данным в правительстве Канады в 2022 году), а другие предпочли создать такие органы, как советы по вопросам данных, при этом руководящие работники по вопросам данных на институциональном уровне обсуждают приоритеты и договариваются об их скоординированной реализации (например, Федеральный совет ГСД США).

B. Что могут предложить Национальные статистические управления

20. Национальные статистические управления уполномочены производить и распространять в координации с другими звеньями НСС актуальные данные и информацию, необходимые для понимания эволюции общества, окружающих условий и экономики. Этот мандат исходит из Основополагающих принципов официальной статистики Организации Объединенных Наций и закреплен в статистическом законодательстве каждой страны. В течение десятилетий НСУ разрабатывали методы и технические способы получения информации, необходимой для выполнения этого мандата. При этом они последовательно и в духе сотрудничества развивают культуру стратегического использования данных и информации в процессе принятия решений на благо общества.

21. Основополагающие принципы официальной статистики Организации Объединенных Наций являются источником одобренных на международном уровне принципов в отношении оптимальных способов предоставления НСУ услуг гражданам (United Nations 2014). Обобщая, можно сказать, что в Основополагающих принципах говорится, что официальная статистика должна быть:

- практически полезной и беспристрастной;
- точной, заслуживающей доверия и этичной;
- основанной на научных принципах и стандартизированной;
- качественной, своевременной и предполагающей минимально возможную нагрузку, связанную с представлением ответов;
- конфиденциальной и надлежащим образом использующейся;
- транспарентной и общедоступной.

22. В Основополагающих принципах также указывается, что НСУ и другие статистические органы могут черпать данные из различных источников, координировать свою деятельность и сотрудничать с другими организациями (на национальном и международном уровнях), а также вносить вклад в совершенствование статистических систем (United Nations, 2014).

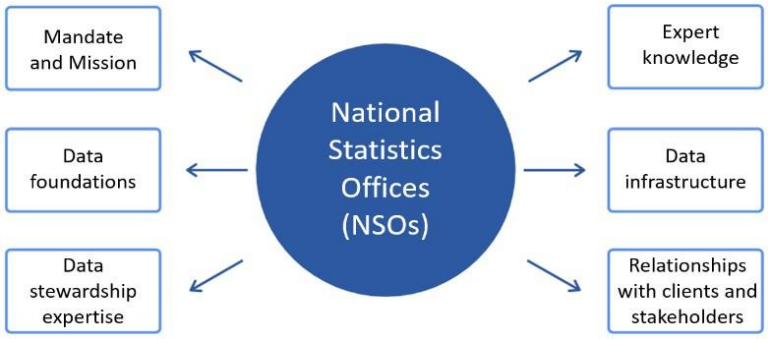
23. Разработанные и утвержденные Конференцией европейских статистиков Основополагающие принципы официальной статистики Организации Объединенных Наций (ОПОС) были приняты в 1992 году на уровне министров Европейской экономической комиссией (ЕЭК ООН) (United Nations, 2021a). Затем эти Основополагающие принципы были подтверждены и одобрены Статистической комиссией ООН на глобальном уровне (1994 год) и Генеральной Ассамблеей ООН (1994 год) (United Nations, 2021a). В преамбуле ОПОС подчеркивается важное значение официальной статистики для национального и глобального развития, выделяется крайне важная роль, которую играют НСУ в предоставлении высококачественных официальных статистических данных, и положительное влияние, которое это оказывает на принятие политических решений и на развитие стран в целом (United Nations, 2014). Связь между Основополагающими принципами и курированием данных более подробно разъясняется в приложении 2.

24. НСУ приобрели определенную репутацию благодаря техническим знаниям, надежности и своевременности, которые они продемонстрировали в ходе тщательного управления собираемой ими информацией. Репутация НСУ свидетельствует об их опыте и компетентности в работе с данными, включая такие аспекты, как конфиденциальность и защита частной жизни, этические нормы работы с данными, обеспечение сопоставимости, использование стандартов и классификаций, распространение информации, поддержка и повышение грамотности в области данных и т. д. Хотя НСУ прежде всего рассматриваются как производители статистических данных, их описанные выше знания и опыт координации деятельности в рамках НСС могут быть использованы в качестве полезного актива в государственном секторе и за его пределами. НСУ, в свою очередь, также могут

воспользоваться опытом других государственных учреждений по решению этих вопросов в своих конкретных сферах деятельности.

25. Национальные статистические управления сталкиваются с вызовом — и с возможностью — занять место ключевых игроков в экосистеме данных, учитывая их богатый опыт в практической реализации управления данными, доступе к данным и обмене данными в государственном секторе, а также их ключевую роль в качестве производителей статистики и показателей. На рис. 1 кратко представлены возможности, которыми обладают НСУ и которые обеспечивают им хорошую основу для выполнения функции курирования данных.

Рис. 1
Возможности НСУ, позволяющие брать на себя обязательства в области курирования данных



26. Для предоставления данных, информации, аналитических материалов и визуализаций в ответ на любые возникающие потребности учреждений, граждан, коммерческих организаций и других институтов НСУ могут использовать следующие рамочные механизмы и опыт:

- законодательные рамки и юридические полномочия для сбора, обработки и распространения данных и информации (на основе законодательства в области статистики);
- хорошо отлаженные системы обеспечения качества данных, включая стандартизированные метаданные, справочные данные и основные данные;
- многолетний опыт защиты данных и передовые методы защиты частных данных и конфиденциальности;
- современные статистические методы анализа данных, интеграция и увязка данных, категоризация данных с помощью метаданных и визуализация данных;
- растущий опыт в области науки о данных, основанный на статистической и математической теории, опыт в моделировании и прогнозировании;
- экспертные знания в области этики данных и растущий опыт этичного использования искусственного интеллекта и машинного обучения для обеспечения беспристрастности методов;
- рамки доверия и разработка новых рамок, таких как пропорциональность и необходимость;
- экспертные знания в области повествовательного изложения и визуализации данных: от данных к пониманию, политике и услугам для достижения оптимальных результатов;
- устоявшаяся культура координации в рамках НСС и доступ к ресурсам и сетям совместно с другими членами НСС, а также обширный опыт партнерства и сотрудничества в области данных как внутри страны, так и на международном уровне;

- платформы для совместной работы (защищенное облако, совместное использование данных, доверительное управление данными), опыт в области услуг доступа к данным для онлайн-доступа к данным надежных и сертифицированных внешних пользователей;
- стандарты статистики и данных и системы классификации для придания данным смысла и обеспечения совместимости;
- опыт сбора данных (переписи населения, различные виды обследований, использование административных реестров, веб-скрейпинг и т. д.);
- хорошо налаженные механизмы распространения данных, включая присутствие в Интернете, центры данных, услуги хостинга данных;
- формирующийся предпринимательский дух и культура непрерывного обучения и развития инновационных экосистем.

27. Помимо продуманного управления данными и стандартов, курирование также предполагает признание и доверие со стороны общественности. Социальное признание и вовлечение общественности напрямую связаны с неприкосновенностью частной жизни, конфиденциальностью, безопасностью и транспарентностью и необходимы для того, чтобы НСУ могли выполнять свои обязанности и мандаты. Постоянная приверженность и следование НСУ принципам открытых данных, прозрачности процессов, обеспечения конфиденциальности и безопасности, а также информирование о пользе предоставления гражданами своих данных — все это является дополнительным доказательством их компетентности и оптимального положения в статистических системах.

28. Курирование данных является продолжением той роли, которую НСУ играют в деле координации национальной статистической системы, когда ключевой задачей является обеспечение эффективной координации между статистическими учреждениями внутри стран (Основополагающий принцип официальной статистики 8) (United Nations, 2014). Роль НСУ может быть расширена за счет более глубокого проникновения в статистическую систему для обеспечения совместимости, а также за счет более широкого охвата всей экосистемы данных.

29. Однако реализация этой ведущей роли в курировании данных в значительной степени обусловлена контекстом и будет варьироваться в зависимости от национальных факторов, таких как законодательная база и мандаты НСУ, степень зрелости и использование цифровых технологий, характер экосистем данных, достигнутая в рамках НСС степень сотрудничества и общественное признание. Таким образом, степень, в которой НСУ могут обеспечить ведущую роль в руководстве данными и их курировании, будет различаться от страны к стране.

30. НСУ могут по-разному позиционировать себя в экосистеме данных — от минимального изменения роли НСУ и до работы в качестве куратора данных в государственном секторе с необходимыми полномочиями по руководству данными. В международном масштабе НСУ находятся на разном уровне «степени зрелости», при этом некоторые из них уже занимают лидирующие позиции в межправительственных структурах. Какого-либо «универсального» решения, которое можно было бы применить ко всем странам, не существует. При переходе к курированию данных важно учитывать юрисдикционный и правовой контекст и понимать, что НСУ будут двигаться вперед своими собственными темпами, выбирая для себя такую роль в области курирования, которая соответствует их целям и ситуации.

C. Какова возможная выгода и риски?

31. Изменения иногда связаны с риском, и независимо от решения, которое принимает то или иное НСУ в отношении курирования данных, существуют следующие риски:

- если НСУ не будут участвовать (вообще или в руководящем качестве) в процессе преобразования экосистемы данных, то возникает не только риск утраты их востребованности, но и высокий риск создания параллельных структур для курирования данных, которые будут увеличивать, а не уменьшать нагрузку и затраты. Результатом может стать дальнейшая фрагментация хранилищ публичных данных, которые будут практически не совместимы друг с другом;
- с другой стороны, существует также риск возложения на себя слишком масштабных функций по курированию данных, что может поставить под угрозу основную деятельность НСУ и привести к репутационным потерям или снижению качества данных;
- кроме того, нельзя преуменьшать значение аспектов кибербезопасности. Хотя НСУ уже применяют самые высокие стандарты информационной безопасности, как хранители данных НСУ могут с большей вероятностью подвергнуться кибератакам, и им необходимо принимать меры для обеспечения готовности к этому.

32. Если роль куратора данных будет выполняться успешно, то НСУ смогут предоставлять важные и крайне необходимые услуги населению, а положительное воздействие, которое они могут оказать на курирование и качество данных во всем государственном секторе, возрастет. Это продемонстрирует, что НСУ могут выполнять свои основные обязанности, одновременно повышая качество и количество данных и официальной статистики. НСУ выиграют, обеспечив себе важную роль в будущем, которая предоставит им:

- доступ к более широкому кругу источников данных, позволяющий получать более своевременные, актуальные и дезагрегированные данные;
- расширение возможностей для интеграции данных из различных источников (особенно при наличии общих идентификаторов).

33. Эффективное курирование данных приносит пользу обществу и служит общественному благу, поскольку позволяет осуществлять этическую деятельность, что создает доверие, социальное признание и общественную поддержку, необходимые для статистической работы. Оно сводит к минимуму неправомерное использование данных и позволяет повторно использовать их, обеспечивая статистикам доступ к уже имеющимся в экосистеме данным. Оно также способствует обмену данными и знаниями и использованию новых, дополняющих друг друга источников данных. Все это прямо или косвенно повышает доверие общественности и увеличивает вовлеченность, экономя время и средства, снижая нагрузку, связанную с представлением ответов, повышая пользу данных и позволяя лучше информировать граждан об этой пользе.

D. Благоприятствующая среда

34. В большинстве стран уже существует несколько благоприятствующих факторов и тенденций, которые способствуют тому, чтобы НСУ взяли на себя активную роль по курированию данных. Три из этих факторов описаны ниже: 1) правовая и нормативная база; 2) технологическое развитие; и 3) организационная совместимость.

1. Правовая и нормативная база

35. Производство официальной статистики в странах и роль НСУ основаны на законодательстве по вопросам статистики. Это законодательство в разных странах имеет разный вид и может определять прямые или косвенные мандаты или роли НСУ в сфере курирования данных — от полного руководства предметной областью до выполнения определенной функции в рамках более широкой структуры НСС. Однако отсутствие законодательного мандата не должно останавливать НСУ от изучения преимуществ выполнения роли куратора в экосистеме данных.

36. В целом, правовые нормы, касающиеся данных и информации, являются довольно новым явлением по сравнению со многими другими регулируемыми вопросами. Поскольку они отражают ситуацию в сфере цифровых технологий и данных, они, естественно, были разработаны с учетом контекста отдельных стран. Унифицированное законодательство по таким вопросам, как защита данных, есть только в странах Европейского союза. Помимо того, что на многие вопросы, связанные с данными, в разных странах реагируют по-разному, они постоянно развиваются и поэтому являются постоянным предметом нормотворческого процесса как на национальном, так и на наднациональном уровнях. У НСУ имеются все возможности для того, чтобы играть здесь активную и ведущую роль за счет использования своих экспертных знаний и обеспечения последовательности и, в конечном счете, эффективности этого нормотворческого процесса.

37. При построении правовой среды для обеспечения возможности курирования данных необходимо учитывать значение повторного использования и обмена на протяжении всего жизненного цикла данных — от сбора до распространения. Необходимо будет установить обязанности тех, кто хранит информацию и управляет ею (кураторы данных), а также правовую и административную основу (соглашения, другие политические инструменты и т. д.), в соответствии с которой это осуществляется, и механизмы (федеративные, централизованные, платформы, трасты и т. д.), с помощью которых это будет достигнуто.

38. Адаптация законодательной базы должна соответствовать национальным стратегиям в области данных и сопровождаться созданием гибкой системы, позволяющей вносить оперативные коррективы без необходимости частого изменения законов. То есть, она должна учитывать постоянное расширение границ экосистемы данных, обусловленное технологическим развитием и эволюцией экономических систем и моделей государственных услуг, при этом она должна установить обязанности и права всех участников этой экосистемы.

39. Правила, связанные с введением в действие механизмов повторного использования информации, должны быть административными по своей природе и в идеальном варианте разработаны на основе консенсуса вовлеченных субъектов (стандарты, модели и т. д.). Таким образом, в случае необходимости их корректировки при возникновении новых рисков или технологических изменений их модификация может быть организована без необходимости законодательного утверждения.

40. Правовая база, регулирующая деятельность НСУ в области курирования данных, должна обеспечивать возможность третьим сторонам проводить исследования и реализовывать проекты, связанные с данными. Это также позволяет выявлять новые сторонние источники, чтобы создавать и рассчитывать статистические показатели. Техническая и юридическая задача заключается в том, чтобы обеспечить регулярный доступ к этой информации, гарантируя при этом, что ее повторное использование юридически не поставит под угрозу ни одну из сторон, участвующих в генерировании данных.

2. Технологическое развитие

41. Технологическая среда постоянно развивается, и организации в разных юрисдикциях сталкиваются с задачей использования преимуществ этих вспомогательных технологий. Степень зрелости в плане использования технологий сильно различается в разных странах, но независимо от этого существует несколько связанных с курированием данных требований, которым НСУ должны уделять первоочередное внимание: 1) достижение надлежащего уровня цифровизации; 2) эффективно функционирующие реестры и использование цифровых услуг (т. е. использование источников административных данных; 3) управляемая архитектура данных, а также управление основными и справочными данными; 4) открытость и защита данных, конфиденциальность данных и кибербезопасность; 5) использование технологий искусственного интеллекта (ИИ), машинного обучения и больших данных; 6) общенациональные информационные системы и пространства данных в масштабе экономических секторов или областей.

3. Организационная совместимость

42. Степень, в которой НСУ могут выполнять роль куратора данных, и успешная реализация этой роли зависят от зрелости экосистемы данных на общенациональном уровне. На эту степень зрелости влияют правовые и технологические факторы, а также организационные факторы, такие как существующие координационные механизмы, стратегии данных департаментов государственных служб, национальные стратегии данных или другие рамки и системы, определяющие порядок действий по управлению и использованию данных в качестве стратегического актива. В процессе разработки таких стратегий НСУ следует активно привлекать к их обсуждению с использованием их знаний и опыта в области координации деятельности в рамках НСС, управления сбором данных, развития операционной совместимости, поддержки и обеспечения стандартизации, обеспечения защиты частной жизни и конфиденциальности, сохранения безопасности данных и повышения грамотности в области данных.

43. Курирование данных актуально не только для развитых стран, где государственные системы данных уже созданы. Так, например, НСУ следует стремиться влиять на процесс создания административных реестров, с тем чтобы обеспечить их пригодность для использования в процессе подготовки статистических данных. Предложение единых подходов к архитектуре данных, стандартам и классификациям до создания публичных хранилищ данных позволит сэкономить ресурсы и извлечь больше пользы из данных в долгосрочной перспективе.

III. Определение курирования данных и руководства данными

A. Курирование данных и связанные с ним термины

44. Курирование данных осуществляется в контексте экосистемы данных, в связи с чем мы начнем с определения этого термина. Данные (включая статистические данные), а также субъекты данных, широкий круг заинтересованных сторон и пользователей данных, возможности, процессы, политика и инфраструктура, используемые для сбора и анализа данных, называются экосистемой данных (European Commission, 2017b; StatCan, 2019). Экосистема включает четыре основные категории участников (European Commission, 2017b):

- производители данных — первичные источники, генерирующие данные, активно или пассивно, такие как потребители, клиенты, предприятия или граждане;
- поставщики услуг данных — субъекты, участвующие в цепочке создания стоимости данных; лица, осуществляющие сбор, организацию, хранение, обработку, извлечение данных, обмен данными, использование, повторное использование данных, ограничение сферы применения или уничтожение данных, такие как НСУ;
- коммерческие пользователи данных — те стороны, например компании и государственные органы управления, которые используют знания, полученные в результате анализа данных, для повышения показателей деятельности или качества жизни;
- конечные потребители — потребители данных, клиенты или граждане.

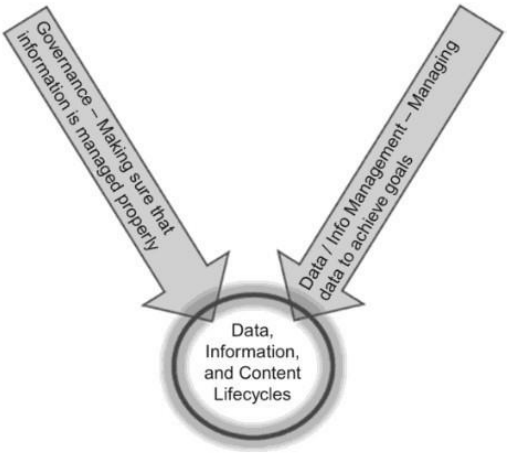
45. Информация, которой можно обмениваться в экосистеме данных, может быть самой разнообразной по своей природе (статистическая, пространственная/географическая, структурированная и неструктурированная, цифровая или в бумажном виде, качественная, количественная, в виде изображений, текста и т. д.) и, таким образом, не должна ограничиваться каким-либо конкретным типом формата. Стандарты должны быть открыты для тех, кто заинтересован в установлении форматов и типов информации, которой они хотят обмениваться.

46. В Своде знаний по управлению данными (СЗУД) руководство данными определяется как «осуществление полномочий, контроля и совместного принятия решений (планирование, мониторинг, исполнение) применительно к управлению активами данных» (DAMA, 2017). Руководство данными заключается в установлении ролей и обязанностей, что фактически представляет собой руководящую практику. Действительно, речь идет об управлении в интересах обеспечения доступности, целостности, совместимости и безопасности хранимых данных. Оно устанавливает и обеспечивает соблюдение политики в отношении доступа к данным, управления ими, их безопасности, обмена данными и их использования; определяет методы и процедуры, необходимые для процесса курирования; и устанавливает квалификацию тех, кто будет использовать данные, и условия, на которых может быть предоставлен доступ к данным (Rosenbaum, 2010). Это может относиться к стратегическому управлению активами данных на уровне частного предприятия или государственного учреждения на межведомственном или даже на национальном уровне.

47. Руководство данными можно определить как систему прав принятия решений и ответственности за управление в интересах обеспечения доступности, удобства использования, целостности и безопасности данных и информации, а также вытекающие из этого положения, политику и рамки, способствующие их исполнению. Такой целостный подход к руководству данными включает системы в рамках предприятия, организации или правительства, которые определяют, кто распоряжается активами данных и имеет контроль над ними и как эти активы данных могут быть использованы, а также круг лиц, процессы и технологии, необходимые для управления активами данных и их защиты (Data Governance Institute, n.d.); IFC, 2021; OECD, 2008, 2019; Plotkin, 2021; StatCan, 2019, 2021b). Руководство данными позволяет согласованно осуществлять и координировать деятельность по курированию данных и повышает потенциал (технический или иной) для лучшего контроля цепочки создания стоимости данных.

48. Управление данными — это дисциплина, которая направляет и поддерживает эффективное и действенное управление информацией и данными в организации, от планирования и разработки систем до утилизации или долгосрочного хранения. Управление данными включает в себя разработку, выполнение и контроль планов, практических действий, концепций, программ и сопутствующего ряда систем, которые способствуют организации и поддержанию процессов данных для удовлетворения текущих потребностей жизненного цикла информации. Управление данными обеспечивает доставку, контроль, защиту и повышение ценности данных и информационных активов с помощью интегрированных, основанных на потребностях пользователей подходов. Ключевые компоненты управления жизненным циклом данных включают реестры данных с возможностью поиска, управление справочными и основными данными и систему оценки качества (DAMA, 2017; Government of Canada, 2019; StatCan, 2019, 2020a, 2021b).

Рис. 2
Модель управления в виде буквы «V» (*источник*: Ladley, 2020)



49. Управление данными можно концептуально представить на основе «модели управления в виде буквы “V”» (Ladley, 2020), которая позволяет провести тонкое различие между руководством данными и управлением данными. С левой стороны находится руководство данными — полномочия и политика, о которых говорилось ранее. С правой стороны изображено «практическое» управление данными и информацией (необходимое для рационального курирования данных). В этой модели сделан упор на то, что руководство данными — это НЕ просто функция, выполняемая теми, кто управляет данными (Ladley, 2020). Она также служит для объяснения сложных взаимосвязей, существующих между курированием данных, руководством данными и управлением данными.

50. В рамках «передовой практики» и принципов руководства данными и курирования данных существует следующее понимание: данные являются стратегическим активом; данные требуют курирования и подотчетности; от качества данных зависят сохранность и ценность данных; данные должны быть надежно защищены и соответствовать правилам конфиденциальности; метаданные должны быть стандартизированы, легкодоступны и быть высокого качества. Эти понятия имеют несколько циклическую, взаимодополняющую связь, однако различие между ними важно как в концептуальном, так и практическом плане.

51. Куратор — это лицо, которое управляет чем-то или присматривает за чем-то от имени другого лица. В литературе иногда используется понятие курирования данных, с тем чтобы отличить его от владения данными (McGilvray, 2021). В настоящем докладе термин «курирование данных» используется с целью указать на то, что к данным государственного сектора не следует относиться так, как будто они являются собственностью учреждений. Вместо этого управление данными должно осуществляться от имени и на благо всего общества.

52. В этом контексте Целевая группа предлагает следующее определение курирования данных:

> «*Курирование данных означает обеспечение этичного и ответственного создания, сбора данных, управления ими, использования и повторного использования данных таким образом, чтобы они применялись на благо общества и приносили пользу всему сообществу пользователей данных*».

53. Курирование данных:

- представляет собой подход к руководству данными, который формализует ответственность за управление информационными ресурсами от имени других лиц;

- обеспечивает реализацию стратегий, стандартов и принципов, предусмотренных системой руководства данными;
- осуществляется посредством надлежащего руководства данными и управления данными, которые обеспечивают подотчетность и обслуживание или надзор за активами данных на протяжении всего их жизненного цикла для надлежащего ухода за ними, поддержания их высокого качества, безопасности и конфиденциальности;
- поощряет инициативные и ответственные методы работы с данными, помогая реализовывать стратегии в области данных, сохранять доверие к ним и способствовать подотчетности;
- выражается в долгосрочном и межпоколенческом ведении активов данных;
- стимулирует поддержку в интересах повышения зрелости политики в области данных;
- является заметным в разнообразных внутренних и внешних проявлениях, связанных с функциями курирования, включая доступ к данным, безопасность, качество данных и стандарты;
- затрагивает методы и механизмы приобретения, хранения, защиты, агрегирования, деидентификации, а также процедуры выпуска, использования и повторного использования данных, с тем чтобы обеспечить высокое качество, легкодоступность и надлежащее использование активов данных;
- применимо на всех уровнях, от национального уровня или уровня системы данных, до уровня организации или предприятия, отдельного лица или набора данных;
- обеспечивает управление взаимодействием различных участников системы и координирует его;
- имеет два основных направления: внутриведомственное и общесистемное курирование данных.

54. Несмотря на сходство сфер применения курирования данных и управления данными, они не являются синонимами. Скорее, на практике они представляются двумя сторонами одной медали. Это продемонстрировано в *Законе Европейского союза о руководстве данными*, в котором особое внимание уделяется совместимости и обмену данными, обеспечению доступности данных государственного сектора для повторного использования, а также применению Генерального регламента по защите данных (ГРЗД) для обеспечения баланса между использованием персональных данных и правами отдельных лиц (European Commission, 2016). При внедрении и применении принципов руководства данными и управления данными может наблюдаться некоторое дублирование. В то время как руководство данными связано с полномочиями и нормативно-правовой базой, которая направляет или предписывает деятельность, позволяющую рассматривать данные как стратегический актив, курирование данных представляет собой ежедневное осуществление такого руководства посредством реализации политики, стандартов и принципов, изложенных в принципах руководства данными (применительно к прикладному курированию данных см. возможные роли и обязанности в главе 5). Именно поэтому курирование данных часто присутствует в качестве важного элемента стратегий в области данных многих организаций. Управление данными представляет собой элемент или механизм курирования данных, который включает в себя такие практические действия, как поиск данных, сортировка данных, добавление метаданных, анализ и очистка данных и т. д.

55. В силу этого специалисты по курированию данных и руководству данными уделяют большое внимание этике данных. «Этика данных» — это более новая отрасль этики, которая сконцентрирована на моральных проблемах, связанных с данными. Этика данных содержит рекомендации по обращению с данными (от сбора до повторного использования), алгоритмы и соответствующие практические методы для формулирования и поддержки положительных с точки зрения морали решений

проблем этики данных. Она проявляется в полном соответствии с правами человека и направлена против рисков неправильного использования и неиспользования данных для общественного блага. Этика данных относится к знаниям, которые позволяют человеку приобретать данные, их использовать, интерпретировать и делиться ими этичным образом, включая признание правовых вопросов и вопросов этики (например, безопасность, предвзятость, частный характер, конфиденциальность и общественная поддержка или социальная приемлемость) (European Commission, 2020; Floridi and Taddeo, 2016; StatCan, 2020).

56. Работающие в этой сфере специалисты по курированию данных и руководству данными несут ответственность за данные. «Ответственность за данные» имеет форму постоянной обязанности отдельного лица, предприятия, организации или государственного органа в отношении управления данными. Этот принцип обеспечивает наличие соответствующих технических и организационных мер и записей для установления соответствия принципам и инструментам политики, связанным с данными. Примеры могут включать политику защиты данных, оценку воздействия, ведение необходимой документации, меры по обеспечению безопасности данных, роли и полномочия, связанные с защитой данных, рамки обеспечения частного характера, а также соответствующий учет и отчетность.

IV. Роли и обязанности по курированию данных

57. Обязанности куратора данных может выполнять то или иное лицо или организация. Например, Национальное статистическое управление может выступать в качестве куратора данных для НСС, для определенного аспекта данных государственного сектора или в качестве национального куратора данных, контролирующего все данные государственного сектора. В организации всегда должно быть лицо или лица (подразделение), ответственное (ответственные) за курирование данных. В каждом подразделении организации, имеющем дело с данными, могут быть кураторы данных или те, на кого возложены такие функции.

58. В рамках агентства куратор данных должен иметь четкий мандат, определяющий его обязанности на протяжении всего жизненного цикла данных, от сбора до распространения и сохранения данных. В зависимости от сложности организации и функций, связанных с производством и обработкой данных, может быть выделено специальное организационное подразделение, включающее должность главного куратора данных. Также крайне важно, чтобы такой куратор входил в сеть кураторов данных *за пределами организации* для поощрения использования данных как общественного блага в интересах общества.

59. На общесистемном уровне необходимо учитывать характер и структуру государственной системы. Речь идет о том, что должно быть назначено лицо, которое будет выполнять координирующую роль в каждой сфере управления: одно в исполнительной, одно в законодательной, одно в судебной сфере и по одному в каждом автономном органе. Эта схема будет воспроизведена на уровне штатов и муниципалитетов в соответствии с применимыми моделями управления данными (такими, как централизованная, федеративная и смешанная модели, рассмотренные в главе 4 полного текста доклада). Будет полезно также определить роль куратора данных в частных организациях и компаниях, владеющих информационными активами, которыми могут пользоваться другие субъекты.

60. В случае НСУ многие должности в их структуре де-факто являются должностями кураторов данных, лица на которых выполняют различные задачи и обязанности, необходимые для осуществления курирования данных. Однако роль распорядителя данных в масштабах всей системы должна быть четко определена, поскольку такое лицо (или организация) должно (должна) координировать работу с другими учреждениями. Особое внимание необходимо уделить отношениям с агентствами по защите частной жизни или персональных данных, чтобы избежать дублирования усилий и уточнить распределение обязанностей.

61. При использовании новых или альтернативных государственных и частных источников данных, не входящих в статистическую систему, куратору данных в НСУ необходимо будет тесно сотрудничать с другими соответствующими сотрудниками, например, ответственными за сбор данных, проектирование данных, качество данных и метаданные. Более подробно об этом говорится в разделе 5.2.3 полного текста доклада.

62. Кураторы данных — на внутриведомственном или общесистемном уровне — образуют новое и важное звено в цепочке создания стоимости данных. Такая ответственность может быть представлена в виде трех основных задач, связанных с сотрудничеством, этикой и совместным использованием (Verhulst, 2021). Ответственное сотрудничество необходимо для разблокирования данных при наличии общественного интереса, обеспечения этичного управления данными для предотвращения вреда и злоупотреблений и принятия необходимых мер к тому, чтобы информация передавалась тем, кому она необходима, что позволит придать ей преобразующее значимое воздействие.

63. Частично роль общесистемного курирования включает деятельность, которую НСУ уже осуществляют в качестве координаторов национальных статистических систем. Однако предложенная новая роль отличается тем, что она глубже и шире охватывает данную сферу, а именно:

- координация должна осуществляться в гораздо более сложных условиях, в связи с чем требуется дополнительная детализация и конкретизация с использованием новых методов и инструментов, позволяющих соответствующее масштабирование: новые источники данных, включая частные данные; концептуальные, методологические и технические проблемы, связанные с интеграцией различных источников данных; повышенное внимание к защите конфиденциальности, этике, инклюзивности, общественному признанию и т. д. Новые методы и подходы необходимы всей статистической системе, и НСУ могут и должны стать лидерами в их разработке и продвижении;
- расширяется функция НСУ по координации за пределами статистической системы, возможно, даже за пределами государственного сектора. Это может быть сделано путем предложения консультаций и рекомендаций, путем предоставления методов и инструментов, которые могут использовать другие агентства, или путем принятия на себя ответственности в качестве кураторов данных (части) государственных хранилищ данных.

A. Общесистемное курирование данных

1. Рациональное курирование данных как одна из целей государственного сектора

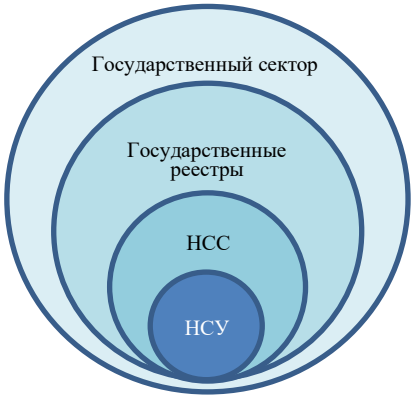
A. Стратегии в области данных и руководство данными

64. Роли и функции по курированию данных, взятые на себя НСУ, часто формализуются и реализуются в стратегиях в области данных, описывающих, как НСУ управляют данными и метаданными, обеспечивая при этом наилучший стандарт статистической информации для общественности. Наличие таких стратегий может стать важным средством коммуникации, способствующим разъяснению смысла деятельности кураторов и поясняющим заинтересованным сторонам методы обработки и защиты информация и используемые при этом стандарты. Существуют стратегии в области данных, которые определяют связанные с данными действия и процессы НСУ, а также подтверждают для заинтересованных сторон и пользователей то, что наилучшая информация, основанная на надежном анализе и стандартах, предоставляется пользователям в рамках всеобъемлющей структуры, при этом особое внимание уделяется гарантиям сохранности данных. Стратегии в области данных обычно охватывают такие аспекты, как сбор данных, использование и повторное использование данных, распространение и передача данных, а также защита и безопасность данных. Они воплощают видение и миссию на практике путем

определения стратегических целей НСУ, что крайне важно при подходе, основанном на курировании данных, поскольку эти цели часто касаются именно ролей и функций кураторов данных, т. е. интеграции данных, установления стандартов, улучшения метаданных, распространения надежных этических принципов и основ в экосистемах данных.

65. НСУ имеют большой опыт работы с данными на протяжении всего их жизненного цикла. В силу этого вопросы управления данными и информацией, доступа к данным, защиты конфиденциальности и безопасности данных занимают важное место в повестке дня правительств и ведомств государственного сектора. В ряде стран разработаны или разрабатываются стратегии в области данных в государственном секторе. Статистические управления способны внести значительный вклад в рассмотрение этих тем, и они должны быть вовлечены в такое обсуждение.

Рис. 3
Возможное освещение роли НСУ в курировании данных



66. Роль курирования данных, которую НСУ могут взять на себя в государственном секторе, может варьироваться в зависимости от контекста, в котором они действуют: от концентрации исключительно на НСС до предоставления рекомендаций и определения направленности в плане управления данными для всей НСС или даже для всего государственного сектора. В контексте широкого применения курирования данных это может происходить в рамках более широкой экосистемы данных (в том числе по отношению к частным держателям данных), применительно к открытым данным или в отношении других производителей официальной статистики, входящих в национальную статистическую систему. Кроме того, такая всеобъемлющая роль актуальна для НСУ в тех странах, в которых не существует государственных реестров, поскольку и в них НСУ в состоянии определять направленность государственных данных (например, в Новой Зеландии). При переносе акцента на внутриведомственный уровень деятельность по курированию данных предполагает «основной» набор обязанностей для НСУ, на которое возложена работа с конфиденциальными данными, и будет включать роли, соответствующие мандату НСУ.

67. На рис. 3 показаны эти различные обязанности и их охват, начиная с внешней окружности, которая определяет наибольший охват — курирование данных и управление данными для всего государственного сектора. При перемещении вглубь каждая новая окружность описывает последовательно сужающуюся роль, пока внутренняя окружность не будет отражать охват только одного учреждения — самого НСУ.

68. Если функции курирования данных для всего государственного сектора возложены на другое учреждение, при этом НСУ выполняет их только в отношении официальной статистики, последнее должно активно сотрудничать с куратором данных государственного сектора. НСУ должны стремиться играть неотъемлемую роль в разработке стандартов, рамок качества и других стимулирующих инструментов,

учитывая тот опыт и влияние, которое эти инструменты оказывают на официальную статистику.

69. При увеличении масштабов оцифровки данных все сильнее проявляется стремление к интеграции наборов данных из различных организаций для принятия решений на основе фактических данных. Необходимо эффективно управлять существующими данными. Их не следует собирать в том случае, если какая-либо организация государственного сектора уже располагает необходимыми данными (в соответствии с принципом «только один раз»). Таким образом, путь к более эффективному использованию данных лежит через повышение совместимости данных и использование общих технических интерфейсов, чему способствует такое согласование в масштабах всего государственного сектора.

70. Для достижения совместимости необходимо наличие общих рекомендаций, структур, определений метаданных, рамок качества, общих идентификаторов и т. д. Статистическое сообщество решило эту проблему десятилетия назад, внедрив использование международных концепций, классификаций и методов для содействия согласованности и эффективности статистических систем (Основополагающие принципы официальной статистики, девятое издание) (United Nations, 2014).

71. Все чаще НСУ стремятся получить доступ к данным, находящимся в частном владении, с тем чтобы выполнить свой мандат по оказанию государственных услуг, предусматривающий более оперативное предоставление в большей степени дезагрегированной информации правительству и обществу. Хотя маловероятно, что с НСУ будут консультироваться по вопросам совместимости или качества таких данных, описанные в настоящем документе концепции курирования данных еще более актуальны в вопросах повторного использования таких данных, поскольку сетевые, управленческие и юридические навыки куратора данных будут иметь важное значение для обеспечения устойчивого доступа. Такая их роль описана в разделе 5.2.2 полного текста доклада.

B. Роль НСУ в курировании данных в государственном секторе

72. Национальный контекст имеет решающее значение для той роли, которую НСУ играет в экосистеме данных. Возможно, уже существуют национальные органы, отвечающие за аспекты руководства и управления данными. НСУ должно будет сотрудничать с этими органами, с тем чтобы как минимум защитить целостность НСУ и НСС и обеспечить тесную координацию при производстве официальной статистики. НСУ может также стремиться влиять на экосистему, «продвигая» свои навыки и используя любые роли, которые ему подходят в более общей экосистеме (см. также главу 7). В силу наличия опыта, инструментов и налаженных партнерских отношений, о которых шла речь выше, НСУ может успешно взять на себя (некоторые) обязанности куратора данных для всего государственного сектора. Однако это может быть осуществимо не во всех странах. В следующем разделе изложены возможные направления или задачи в плане продвижения курирования данных и управления этим процессом на уровне государственного сектора, а также некоторые примеры различных подходов, применяемых странами в этом отношении. На рис. 5 представлены различные виды обязанностей по курированию данных, а также примеры занимающихся этим стран. Далее следуют тематические исследования и примеры инициатив и направлений действий в различных странах.

73. Следует отметить, что модель ТМРСО ЕЭК ООН (2015a) с ее акцентом на стратегию и лидерство идеально подходит для определения роли НСУ в национальном контексте. Модель ТМРСО определяет и описывает деятельность, которая имеет место в статистической организации, расширяя и дополняя Типовую модель производства статистической информации (ТМПСИ) путем включения видов деятельности, необходимых для поддержки статистического производства (2015a).

74. НСУ располагают уникальными возможностями для оценки сильных и слабых сторон национальной системы административных данных. Принцип «только один раз» — это используемая ЕС концепция, позволяющая повысить эффективность и снизить нагрузку на граждан, учреждения и компании, которые должны предоставлять

определенную информацию органам власти и управления *только один раз*. Принцип «только один раз» требует выявления пробелов в данных, а также максимизации ценности разрозненных данных путем их последующей интеграции, обмена ими и анализа таких данных. Роль НСУ в определении направления развития системы может варьироваться от создания эталонных классификаций до мониторинга использования классификаций и стандартов, влияющих на статистику (например, общих идентификаторов). Наиболее развитая или продвинутая роль курирования заключается в том, что НСУ устанавливают стандарты для «интегрированной инфраструктуры данных», которые впоследствии внедряются и контролируются во всей системе государственных данных.

75. В силу своих мандатов и миссий НСУ в обязательном порядке занимаются вопросами сохранения конфиденциальности и социальной направленности данных. По мере повышения качества данных в государственном секторе и роста требований к их анализу в отдельных министерствах такое обсуждение становится все более актуальным для всего государственного сектора. Такие государственные министерства работают в многоаспектном контексте, в котором социальное воздействие их политики имеет широкий спектр последствий в других областях, что способствует росту потребности в совместимости и в увязке данных в этих областях. Хотя НСУ могут удовлетворить некоторые из этих требований, им также может потребоваться сотрудничество с другими партнерами при сохранении общей атмосферы доверия к системе.

Рис. 4
Краткий перечень областей работы по курированию данных в рамках всей системы (в скобках приведены ссылки на примеры стран)

Курирование данных как одна из целей государственного сектора

- Вклад в разработку и продвижение стратегий, политики и принципов в области данных (Австралия, Ирландия, Канада, Литва, Мексика, Новая Зеландия, США, Финляндия, Франция, Швейцария, Эстония)
    - Национальные/государственные стратегии в области данных
    - Принцип «только один раз», открытые данные, принципы FAIR, CARE
    - Укрепление доверия к системе и уверенности в ней

Координация и партнерства

- Координация процесса стандартизации и гармонизации, поддержка совместимости (Австралия, Мексика, Швейцария)
- Партнерства для расширения возможностей, разработки новых и улучшения существующих продуктов
    - Соглашения о совместном использовании данных
    - Партнерские и лицензионные соглашения
    - Предоставление рекомендаций другим учреждениям (Новая Зеландия, Финляндия)

Обеспечение качества и оказание помощи

- Оценка качества данных в национальной статистической системе и продвижение и предоставление экспертных знаний по качеству данных другим держателям государственных данных (Австралия, Германия, Норвегия, Финляндия, Швейцария, Эстония)
- Разработка и продвижение рамок обеспечения качества (например, для административных реестров, данных, находящихся в частном владении)
- Сертификация и аудит качества

Доступ к данным

- Предоставление исследовательских и информационных услуг
  - Доступ к микроданным (Австралия, Нидерланды, Эквадор)
  - Обследования для других государственных учреждений (Ирландия)
  - Информационные услуги и аналитическая поддержка пользователей (в том числе разработчиков политики) (Ирландия, Нидерланды)
- Создание платформ данных и информационных панелей (для обмена данными и их распространения) (Австралия, Германия, Ирландия)
  - Объединение данных из разных областей (например, центры хранения и обработки данных) (Ирландия)
  - Связи с геопространственными данными, геопространственная визуализация (Хорватия)
- Доступ к частным источникам данных (Австралия)

Методы, инструменты и возможности

- Стандарты и классификации (Австралия, Ирландия, Мексика, Нидерланды, Новая Зеландия, Швейцария, Эстония)
- Инфраструктура (например, для интеграции данных, доступа к индивидуальным данным/микроданным) (Австралия)
  - «Озеро данных» (Литва)
  - Каталог (мета)данных (Хорватия, Швейцария)
  - Общие идентификаторы (Дания, Ирландия, Швеция)
- Интеграция данных
- Защита данных
- Статистические методы, наука о данных и методы машинного обучения (Австралия, Нидерланды)
- Общий подход к обработке данных (Австралия)
- Помощь другим агентствам в развитии навыков, процессов, инструментов и услуг (Ирландия)
- Повышение грамотности в области данных (Австралия, Эстония)

2. Поддержка других поставщиков данных государственного сектора

A. Координация и партнерства

76. Во многих случаях поставщики данных также являются пользователями данных, поскольку учреждения государственного сектора, такие как правительственные министерства или центральные банки, обычно выполняют не только операционные, но и стратегические функции. Как минимум, информация о соответствующих результатах должна рассматриваться как часть общих отношений с такими агентствами. Более активный подход предполагает участие аналитиков этих агентств в группах по связям, регулирующих потоки данных, в рамках которых они могли бы внести свой вклад в статистическую работу, основанную на источниках данных агентств. Поддержка аналитиков в государственном секторе также может быть организована через сети, для которых НСУ выступает в качестве идеального партнера.

B. Доступ к данным

77. Многие НСУ рассматривают возможность развития услуг по предоставлению данных и информации, принимая решение о том, станут ли они управлениями «данных и статистики». Ключевым фактором для этих услуг является дифференцированная конфиденциальность. НСУ предоставляют открытые данные, при этом через стратегические партнерства они могут также содействовать другим государственным органам в выполнении их обязательств по предоставлению открытых данных. Доступ исследователей является одной из ключевых областей для развития, включая ту новую роль, которую многие НСУ взяли на себя в последнее время, обеспечив облегченный доступ к данным для исследователей COVID. Этому вопросу также уделяется внимание в законодательстве ЕС (см. приложение 2), поскольку Закон о руководстве данными предусматривает доступ исследователей к псевдонимизированным данным государственного сектора, что может стать предметом рассмотрения в деятельности НСУ по курированию данных. Наряду с услугами по предоставлению открытых и псевдонимизированных данных, многие НСУ также предоставляют услуги по предоставлению анонимных данных для развития навыков работы с данными у не обладающих обширным опытом исследователей.

C. Помощь в обеспечении качества (ОК) — методы, инструменты и возможности

78. С тем чтобы возглавлять работу в области данных и курировать качество данных общественных организаций, куратор данных должен обладать несколькими различными способностями и возможностями. При создании и поддержке национальных решений в области качества данных необходимы навыки координации, коммуникации и налаживания контактов. Это — долгосрочная работа, требующая специальных людских ресурсов, лиц, знакомых с методологиями, а также обладающих глубокими знаниями о данных, о которых идет речь.

79. Важно также отметить, что требуемый уровень качества данных зависит от того, как эти данные используются. Кроме того, могут иметь место различные виды использования одних и тех же данных, что приводит к различным предположениям о качестве и потребностям в отношении таких данных. Эти вопросы должны обсуждаться подробно с соответствующими профильными специалистами. Конкретные темы, которые НСУ могло бы рассмотреть в рамках своей роли куратора, включают:

a) Метаданные. Работа с ними крайне важна для оценки качества данных и должна вестись для административных данных в НСУ, поскольку такие данные изначально не собирались для статистических целей, и каждый источник будет располагать своими собственными показателями измерения качества в зависимости от предусмотренных мандатом/законодательных сфер деятельности, охвата и т. д. НСУ может также рассмотреть возможность предоставления метаданных в открытый доступ для исследователей или даже в более широком плане. В качестве следующего шага НСУ может определить общие стандарты метаданных или сотрудничать с организацией, разрабатывающей стандарты, для обеспечения согласованности и последовательности.

b) Разработка и продвижение стандартов данных в сообществе владельцев данных, таких как классификации входных и выходных данных. Поскольку эти стандарты необходимы для внутренних целей, они могут быть полезны для других органов, собирающих аналогичные данные, с тем чтобы помочь им в использовании и увязке данных.

c) Поощрение использования уникальных идентификаторов как особой категории стандартов. Это особенно актуально для стран, не имеющих реестров. Если охват таких идентификаторов невелик, то НСУ может продвигать стандарты для идентифицируемых данных, с тем чтобы улучшить последующий уровень вероятностного сопоставления.

d) В то время как основной проблемой НСУ в отношении данных, находящихся в частном владении, до сих пор был доступ, необходимо также рассмотреть рамки обеспечения качества этих данных. Модель качества, применяемая к данным государственного сектора, может и не подходить для данных частного сектора. Определяющие характеристики больших данных — «объем, скорость получения и разнообразие» — приведут к получению данных, отличных от продуктов из традиционных административных источников, которые будут весьма информативными при высоком и предсказуемом уровне охвата населения.

B. Связанные с внутриведомственным управлением данными роли — роль куратора данных в рамках НСУ

80. Роль куратора данных, как она изложена в настоящем документе, содержит описание *минимальных* действий и обязанностей для этой роли в том виде, в каком она существует в организации, которая получает и обрабатывает конфиденциальные данные. Контекст НСУ будет более подробно рассмотрен в последующих разделах доклада путем соотнесения роли куратора данных с другими ролями в организации на основе матриц ТМПСИ ЕЭК ООН и ОПСКИ (ответственность, подотчетность, содействие, консультирование и информированность). Содержание «максимального перечня» видов деятельности и обязанностей будет зависеть от степени внешнего участия НСУ.

81. Как минимум куратор данных в НСУ должен отвечать за следующие вопросы:

- описание данных (например, каталог данных, словарь данных);
- качество метаданных (ответственность за данные лежит на владельце данных);
- управление жизненным циклом данных;
- этика данных (куратор данных является связующим звеном между сотрудником по управлению данными/сотрудником по защите данных и комитетом по этике);
- безопасность, защита и конфиденциальность данных;
- аудит данных (мониторинг использования данных).

82. Обязанности куратора данных в НСУ можно разделить на сетевую деятельность (как внутреннюю, так и внешнюю), технические обязанности, связанные с управлением данными, и этические и юридические обязанности.