

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

Женева, 23–25 июня 2021 года

Пункт 2 b) предварительной повестки дня

**После COVID: поддержка организационных
инноваций и инновационных продуктов
в национальных статистических управлениях****Управление оперативными данными о Covid-19 —
Производственная модель для статистики нового
поколения*****Подготовлен Литвой***Резюме*

В документе рассматривается опыт Статистического управления Литвы (СУЛ), касающийся управления оперативными данными о пандемии Covid, которые имеют крайне важное значение для борьбы с пандемией, мониторинга и принятия решений с целью сохранения здоровья и жизни людей. Чрезвычайная ситуация высветила важность наличия пространства в государстве, в котором можно было бы управлять всей экосистемой данных для удовлетворения информационных потребностей, начиная со сбора данных и заканчивая аутсорсингом, а также гибкого управления правами пользователей для принятия решений «здесь и сейчас».

Решение этой важной задачи укрепило СУЛ как учреждение и повысило эффективность его деятельности, которой управлению не хватало до возникновения этого вызова. Однако, несмотря на то, что это оказало огромное влияние на укрепление репутации и бренда СУЛ, этот шаг представляет собой долгосрочное стратегическое изменение в основной деятельности организации — формировании официальной статистики. В настоящее время внедряется разработанная в сжатые сроки эффективная модель управления данными для формирования официальной и экспериментальной статистики. Это позволяет ожидать значительных изменений в уровне детализации и актуальности статистики, а также новых качественных услуг для государства, науки, бизнеса и общества. Своевременные решения СУЛ и быстрые действия по оперативному управлению статистикой о Covid создали предпосылки для реализации гораздо более амбициозной цели — кардинального изменения роли СУЛ и превращения его в куратора государственных данных.

* Настоящий документ был представлен с опозданием ввиду ограниченности ресурсов.



Настоящий документ представляется для обсуждения на заседании Конференции европейских статистиков, посвященном теме «После COVID: поддержка организационных инноваций и инновационных продуктов в национальных статистических управлениях. — Инновации в организации и механизмах работы национальных статистических управлений».

I. Введение: обзор ситуации до Covid в области управления государственными данными

1. Данные государственного сектора в Литве носят полностью децентрализованный характер. Это ведет к большому технологическому и функциональному разнообразию административных источников. Высокая фрагментарность источников данных и ее негативное воздействие на возможность использования данных стали очевидными с началом пандемии Covid. В случае Литвы сбор и обработка оперативных данных стали ключевой проблемой в борьбе с пандемией. И именно пандемия Covid вдохновила нас на принятие радикальных и быстрых решений в области управления государственными данными. СУЛ стало центральным действующим лицом, демонстрируя свое стремление взять на себя фундаментально новую роль куратора данных.

2. Как складывается нынешняя ситуация с управлением государственными данными в Литве? Официально существует 96 регистров и 285 государственных информационных систем. Этими регистрами и информационными системами управляют 139 различных администраторов и 168 главных администраторов. Большинство систем и регистров ориентированы главным образом на внутренние оперативные цели, зависят от финансовых возможностей владельцев или администраторов, сферы компетенции и т. д. и не обязательно учитывают необходимость систематического подхода к государственным данным, учета более широкого контекста, потребностей пользователей и т. д.

3. Таким образом, проблемы, для решения которых мы начали реализовывать свою роль в качестве куратора государственных данных, понятны:

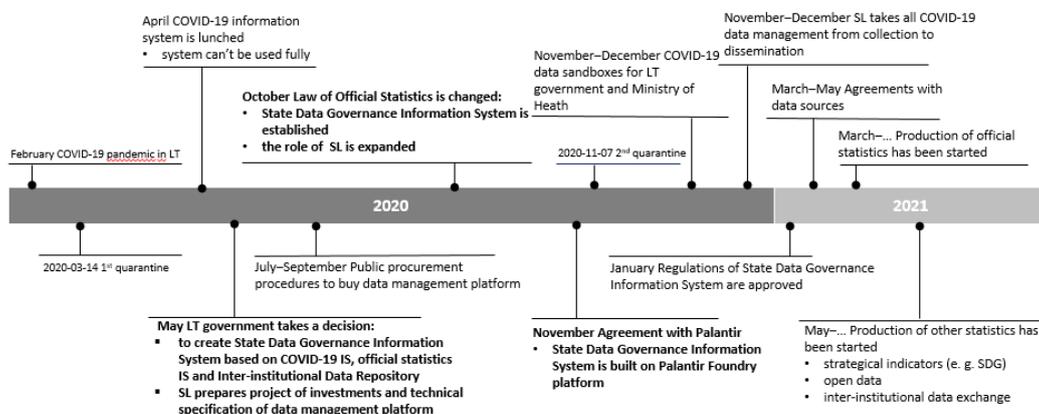
- изолированные, фрагментированные данные не имеют большой полезности. Полезная информация получается путем объединения данных из разных областей, анализа и так далее. Однако уже упомянутое большое количество источников данных, основанных на широком спектре технологических решений, означает, что данные государственного сектора имеют ограниченное использование. Это означает, что они предназначены и могут быть использованы на практике только для закрытых внутренних целей их владельцев. Большинство регистров и информационных систем являются закрытыми по своей природе и ориентированы только на потребности своих владельцев. Процессы открытия или повторного использования данных не централизованы;
- негибкая и медленная официальная статистика. Официальная статистика формируется на основе законодательства ЕС и национального законодательства в строгом соответствии с установленными методологиями и графиками. Это обеспечивает качество статистической информации. Однако в настоящее время пользователи нуждаются в своевременных и подробных данных более даже, чем в высокоточных данных. В среднем мы формируем годовую статистическую информацию за 150 дней, квартальную — за 58 дней, месячную — за 24 дня, что недопустимо в условиях пандемии или других возможных бедствий. Традиционные статистические обследования не позволяют достичь высокой эффективности, а также увеличивают нагрузку по предоставлению статистической отчетности. По этим причинам использование административных источников является одним из ключевых решений. Это позволяет формировать «более оперативную» статистику, отражающую социально-экономические изменения практически в режиме реального времени. Поэтому в нашем производстве все большее значение приобретают публичные данные. И здесь мы сталкиваемся с еще одной огромной проблемой: имея право собирать любые административные данные для целей нашей официальной статистики, мы не можем полностью раскрыть их потенциал, так как не имеем права использовать их для других целей, в отличие от владельцев этих данных (в соответствии с РОЗД, разумеется);

- неэффективный обмен данными и инициативы по повторному использованию и открытию данных. Как государственный сектор, так и бизнес владеют огромными информационными ресурсами. Однако обмен этими данными затруднен из-за уже упомянутой фрагментарности источников данных. Пользователи обязаны связываться с владельцами или администраторами индивидуальных данных для получения данных от государственного сектора в соответствии с различными процедурами. Предприятия также не имеют стимулов делиться с государством своими ценными большими данными. Существующие инициативы по открытию данных не решают проблему фрагментарности данных государственного сектора и не повышают их доступность. Кроме того, открытие данных увеличивает расходы владельцев данных на формирование и предоставление открытых данных или на их повторное использование. Кроме того, открытие данных еще больше увеличивает расходы владельцев данных на формирование и предоставление открытых данных, в то время как полезность этих открытых данных остается сомнительной из-за недостатков качества. Фрагментированный сбор и открытие данных мешают науке и бизнесу использовать эти публичные информационные ресурсы для разработки продуктов, имеющих высокую экономическую и социальную полезность.

4. Вышеупомянутые проблемы привели к разработке СУЛ инициативы по их решению путем создания Системы управления данными (или, точнее сказать, государственного «озера данных»), которая была запущена в ноябре 2020 года. СУЛ приобрело готовую платформу управления данными, которая позволяет создать устойчивую и гибкую экосистему данных с облачными инфраструктурными сервисами. Как уже упоминалось, главным стимулом к реализации идеи Системы управления данными послужила (и остается) пандемия Covid.

Рис. 1

График: от идеи до реализации



5. Основные этапы графика: решение Правительства о создании Системы управления государственными данными, внесение изменений в Закон об официальной статистике и покупка платформы управления данными, позволяющей реализацию такой информационной системы. В настоящее время мы находимся в процессе ассимиляции первичных данных из источников и процесса производства. Это означает, что мы подписываем соглашения с владельцами данных, организуем сбор данных из их систем и начинаем передачу данных статистических обследований и процесса производства статистической информации в Систему управления государственными данными. Параллельно мы создаем модели обеспечения конфиденциальности данных, оценки качества и предоставления услуг. Все это делается в контексте существующего процесса управления данными о Covid. Это означает, что то, чего мы стремимся достичь в большом масштабе, уже достигнуто в случае данных о Covid.

II. Правовая основа: робкий шаг вперед

6. В соответствии с новыми положениями [Закона об официальной статистике](#) СУЛ становится владельцем и администратором Системы управления государственными данными, сфера компетенции которой охватывает более широкое пространство данных и цели использования, чем официальная статистика. Таким образом, СУЛ становится актером, исполняющим две роли: СУЛ использует новую систему для формирования официальной статистики, и в то же время СУЛ получает право собирать данные и формировать информацию, отличную от официальной статистики.

7. Следует отметить, что мы видим, что поправка к Закону несколько повысила способность СУЛ удовлетворять информационные потребности государства и общества, но недостаточно, так как существуют значительные противоречия и препятствия в плане открытия и повторного использования данных. Положения Закона об официальной статистике позволяют использовать ресурсы Системы управления государственными данными только в статистических целях. Это означает, что:

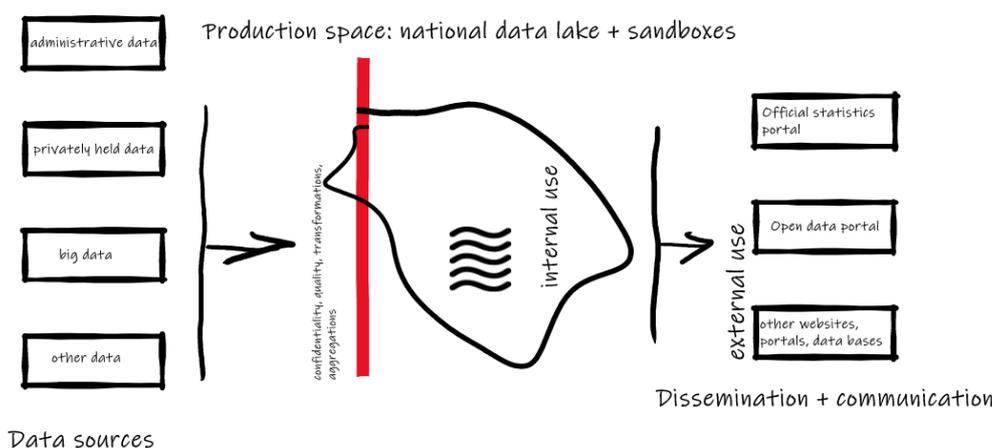
- данные не могут предоставляться в административных целях, или это требует длительных бюрократических процедур (например, подписания трехсторонних соглашений между владельцем данных, СУЛ и пользователем данных), что делает невозможной реализацию главной идеи Системы управления государственными данными — оперативного обмена и предоставления данных. Реальные примеры: эффективное аналитическое пространство для исследований теневой экономики, проводимых Государственной налоговой инспекцией, может быть реализовано в Системе управления государственными данными, но без дополнительного правового регулирования Система управления государственными данными не может использоваться в этих целях. Аналогичная ситуация складывается с предоставлением муниципалитетами подробных данных с целью оказания им помощи в борьбе с пандемией Covid и принятии соответствующих решений на местном уровне: с технологической точки зрения Система управления государственными данными может предоставлять оперативные и подробные (обезличенные или зашифрованные) данные для принятия решений, связанных с борьбой со вспышками заболевания, профилактическими мерами, материально-техническим обеспечением вакцинации и т. д., однако действующая законодательная база не позволяет это делать;
- централизованное открытие данных, хотя и очень оперативное с технологической точки зрения (например, автоматизированное открытие данных о Covid на общенациональном портале открытых данных уже реализовано). СУЛ планирует открыть, в координации с владельцами данных, все ресурсы Системы управления государственными данными, которые, вероятно, будут необходимы общественности. Для этого требуются первичные данные, поскольку это является единственным способом обеспечить максимальное использование данных. Однако существует риск того, что в отсутствие юридического обязательства владельцы данных не будут склонны делиться данными, которыми они владеют;
- высокая фрагментированность данных государственного сектора в различных системах и регистрах означает, что получение данных в Системе управления государственными данными, которая технологически быстро осуществима, может искусственно тормозиться системами или руководителями регистров из-за отсутствия интереса к предоставлению первичных данных. Это уже очевидно сейчас, когда переговоры по соглашениям о получении данных ведутся в течение относительно длительного времени, отказ в предоставлении данных приобретает наиболее изворотливые формы и т. д. Еще одним аргументом является отдельное национальное законодательство, запрещающее повторное использование данных для целей, отличных от оговоренных.

III. Решение и его реализация

8. Какова формула решения СУЛ, позволяющая стать ему главным действующим лицом в управлении государственными данными? Она очень проста — создать общее пространство данных (или общенациональное озеро данных) для консолидации и обеспечения использования важных государственных данных. Это означает создание Системы управления государственными данными на основе многофункциональной и эффективной платформы управления данными, которая бы обеспечивала оперативный и эффективный доступ к различным источникам данных, с целью предоставления услуг государству, науке и бизнесу. И роль СУЛ, расширенная Законом об официальной статистике, создает правовые предпосылки для достижения этой цели.

Рис. 2

Производственное пространство: национальные озера и песочницы данных



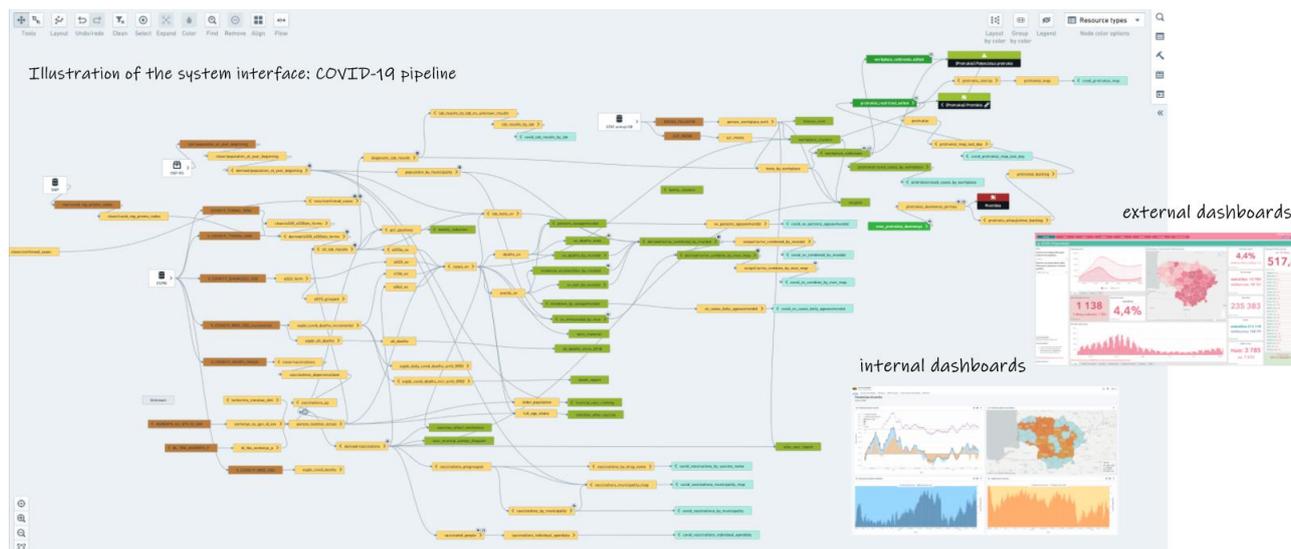
A. Управление данными о Covid

9. Основным источником оперативных данных о Covid в Литве является система *электронного здравоохранения*, управляемая Министерством здравоохранения и ведущаяся Центром регистров, которая в основном предназначена для сбора и хранения данных о пациентах учреждениями здравоохранения и врачами, а также для оказания связанных с лечением электронных услуг. Это означает, что данная система не пригодна для анализа данных, и поэтому в условиях пандемии Covid данные, собранные в системе *электронного здравоохранения*, невозможно использовать на практике. В настоящее время система *электронного здравоохранения* ежедневно предоставляет первичные данные, связанные с диагностикой, заболеваемостью Covid, смертностью от него и т. д. Ниже приводится краткий обзор основных процессов управления данными о Covid: заболеваемость и смертность, вакцинация и борьба со вспышками.

- *Процесс управления данными о заболеваемости и смертности.* На рисунке 3 ниже (и в этом коротком [видеоролике](#)) приведены примеры интерфейса системы: полный путь данных о Covid, касающихся тестирования, подтвержденных случаев, случаев смерти, до конечного продукта, внутренних и внешних информационных панелей и отчетов. Дополнительная потребность в новых индикаторах реализуется без разрушения общей последовательности данных. Например, если Министерству здравоохранения или другому государственному органу необходимо представить новые отчеты в международные организации, это реализуется очень быстро в Системе управления государственными данными.

Рис. 3

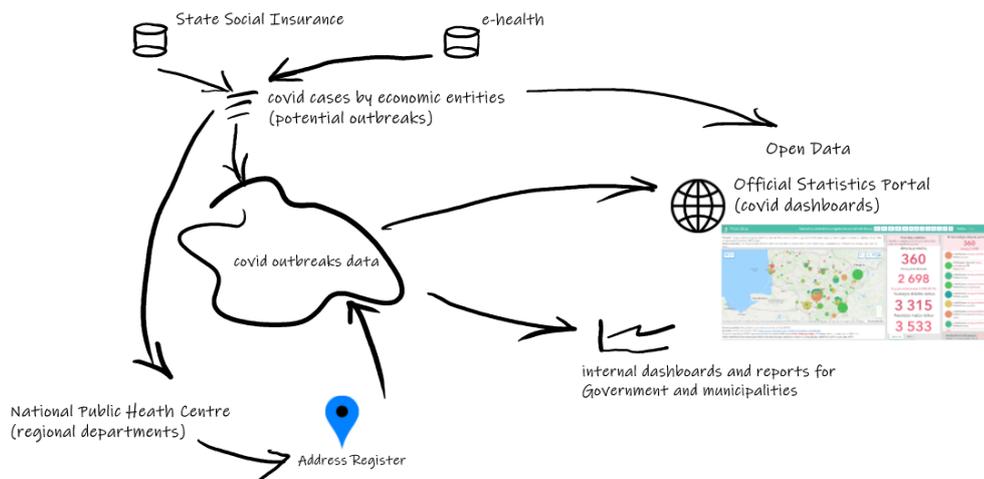
Иллюстрация интерфейса системы: процесс формирования данных о Covid-19, внутренние и внешние информационные панели



- Материально-техническое обеспечение вакцинации и процесс управления данными.* Модуль данных о вакцинации в Системе управления государственными данными включает в себя не только управление данными, но и материально-техническое обеспечение вакцинации. Это успешный пример сотрудничества с Министерством здравоохранения — когда, в некотором смысле, платформа, управляемая СУД, «сдается в аренду» для выполнения дополнительных задач. Министерство здравоохранения вводит информацию о поступлении вакцин в Систему управления государственными данными, которые автоматически распределяются между центрами вакцинации в соответствии с установленным алгоритмом. При распределении учитывается численность и состав населения по возрасту и приоритетным группам, определенным правительством.
- Карта данных о вспышках.* Мониторинг вспышек Covid является одной из самых сложных задач в борьбе с пандемией. До автоматизации управления данными о вспышках в Системе управления государственными данными региональные подразделения Национального центра общественного здравоохранения использовали собственную внутреннюю базу данных, которая, как и электронная система здравоохранения, не имеет аналитической составляющей, а любой интерфейс с другими системами сложно реализовать из-за устаревшей технологии. Кроме того, правительству и общественности не предоставлялась ясная информация о вспышках заболевания и способах борьбы с ними. Процесс управления данными о вспышках заболевания отличается от других способом сбора данных — данные загружаются сотрудниками региональных подразделений, а Система управления государственными данными позволяет проводить комплексный анализ ситуации и распространять данные.

Рис. 4
Оперативные данные о вспышках COVID-19

operational data of covid-19 outbreaks



- **Показатели госпитализации.** Сбор данных о госпитализации организован так же, как и в случае вспышек. Региональные медицинские учреждения ежедневно предоставляют данные о новых пациентах с Covid, койках для лечения Covid, а также о наличии врачей и медсестер.

10. Все ключевые данные о Covid передаются каждое утро для составления **ежедневного отчета**, который является основным источником данных для СМИ и общественности. Одновременно данные обновляются в специальных **информационных панелях о Covid**, а также на **портале открытых данных** и общенациональных **открытых данных СУЛ**.

В. Подготовка к формированию официальной статистики

11. Как можно увидеть из вышеизложенного, цели Системы управления государственными данными можно кратко сформулировать следующим образом: консолидация данных, производство информации, обмен данными и услуги аналитического пространства. В настоящее время (первая половина 2021 года) Система управления государственными данными охватывает 28 владельцев данных и 53 их регистра или информационные системы. Это обеспечивает охват практически всех текущих потребностей в официальной статистике, а также значительной части разрабатываемых государственных стратегических показателей. Следует отметить, что цель заключается в консолидации первичных данных. Это позволит централизованно формировать статистику нового поколения — гораздо более оперативную и детальную.

12. Приоритетными задачами в этот период являются: 1) ассимиляция данных, 2) разработка и применение модели анонимизации/шифрования, обеспечивающей безопасность данных, 3) разработка системы прав пользователей, обеспечивающей безопасность данных, 4) передача текущих алгоритмов расчета показателей официальной статистики.

IV. Влияние на организацию: изменения в имидже и репутации

13. После того, как СУЛ взяло на себя функции управления данными о Covid и их распространения, были отмечены серьезные изменения в бренде и репутации СУЛ. СУЛ измеряет полезность своей продукции для общества путем расчета двух показателей: удобство использования статистики и степень удовлетворенности пользователей. Удобство использования включает в себя так называемые объективные

показатели: удобство использования Портала официальной статистики, Базы данных индикаторов и социальных сетей, цитирование СУЛ и использование его материалов в СМИ, индивидуальные запросы пользователей. Цитирование бренда СУЛ и его материалов в основных общенациональных СМИ выросло с 466,8 млн до 3,8 млрд, в самой популярной социальной сети Facebook — с 420 тыс до 1,0 млн, а уровень удовлетворенности пользователей — с 72,4 до 75,5.

14. В некоторых случаях СУЛ зафиксировало рост в сотни раз. Очевидно, что это было связано с публикацией оперативных данных о Covid. Однако следует отметить, что значительный интерес вызвала не только статистика о Covid, но и статистика в других областях. Это означает, что СУЛ и его деятельность стали гораздо более заметными и известными публике. Следует отметить, что Covid также вызвал определенный прорыв в использовании открытых данных. Например, открытые данные о вакцинации с середины февраля просматривались 500 000 раз на портале открытых данных СУЛ, а в среднем открытые файлы данных просматриваются до 7 000 раз в день.

15. Что касается более тонких внутренних изменений в организации, то следует отметить, что появление совершенно новой идеи стало, в некотором смысле, «глотком свежего воздуха» для СУЛ, относительно крупной бюрократической организации Литвы. Благодаря проекту Системы управления государственными данными в СУЛ появились почти 30 новых высококвалифицированных специалистов, которые привнесли иные взгляды, культуру работы и привычки. Наряду со структурными изменениями в организации, изменением рабочей среды во время пандемии, это неизбежно приносит результаты и стимулирует изменения: эффект снежного кома, похоже, начинает работать.

V. Выводы

16. Главный вывод анализа этого чрезвычайно динамичного периода времени и экстраординарного опыта заключается в том, что выполнение роли куратора государственных данных не является технологической проблемой. Речь идет, прежде всего, о юридической проблеме. Статистические органы нуждаются в правовом мандате для расширения своей роли, с тем чтобы стать кураторами данных и обеспечить использование своих собственных компетенций и данных. В настоящее время открытие данных, обмен ими и их повторное использование затруднены законодательными положениями, ограничивающими деятельность статистических органов. И это очень важный вывод: если статистические органы, как обладающие наибольшими компетенциями в области управления данными, не расширят свою роль в сфере не только формирования официальной статистики, но и обмена своими ресурсами, то они станут ненужными, так как новые игроки выйдут на рынок данных в соответствии с Законом о данных.

VI. Планы на будущее: централизованное формирование открытых данных государственного сектора

17. Цели СУЛ на 2022–2026 годы связаны с реформой открытия государственных данных. Целью реформы является обеспечение доступности достоверных данных государственного сектора, возможности обмена ими между секторами экономики, возможности повторного использования данных и создание предпосылок для принятия учреждениями государственного сектора решений на основе данных, а также для внедрения предприятиями цифровых инноваций. Для этой цели планируется использовать средства Фонда восстановления и устойчивости.

18. В контексте реформы данных Система управления государственными данными рассматривается в качестве общенационального озера данных, объединяющего в себе важные государственные данные, которые, в свою очередь, благодаря применению стандартизированных методов анонимизации, конфиденциальности и открытия данных, используются для получения широкого спектра информации: открытых

данных, официальной, экспериментальной и оперативной статистики, мониторинга стратегического государственного развития, специальных правительственных задач, отраслевых пространств данных, исследований, инноваций в бизнесе.

19. Реформа направлена на расширение объема данных, публикуемых в Системе управления государственными данными, за счет использования 376 государственных информационных ресурсов, создания предпосылок для повторного использования данных во всех государственных информационных системах и регистрах, а также открытия высокополезных данных для общественности, бизнеса и науки. Тот факт, что вся экосистема данных реализована в Системе управления государственными данными, созданной на базе готовой платформы управления (большими) данными, обуславливает возможность реализации помимо главной цели других дополнительных сервисов: открытие данных, аналитические пространства (песочницы), извлечение больших данных из массивов предприятий или других источников и обмен различными данными друг с другом и с внешним миром.

20. Повторное использование данных, имеющихся у государственных органов, повышает их полезность для повторных пользователей, конечных пользователей и общества в целом. В то же время во многих случаях это отвечает интересам самого органа государственного сектора, поскольку усиливает прозрачность и подотчетность и обеспечивает обратную связь с пользователями и конечными пользователями и поощряет соответствующий орган государственного сектора обеспечивать качество собранной информации и выполнение им своих государственных задач.
