



БЮРО НАЦИОНАЛЬНОЙ СТАТИСТИКИ АГЕНТСТВА ПО
СТРАТЕГИЧЕСКОМУ ПЛАНИРОВАНИЮ И РЕФОРМАМ
РЕСПУБЛИКИ КАЗАХСТАН

КОНЦЕПЦИЯ

РАЗВИТИЯ ГОСУДАРСТВЕННОЙ
СТАТИСТИКИ И НАЦИОНАЛЬНОЙ
ЭКОСИСТЕМЫ ДАННЫХ

Астана 2022

СОДЕРЖАНИЕ

1

3 стр.

ОБРАЩЕНИЕ

02

4 стр.

ОБЩИЕ ПОЛОЖЕНИЯ

03

7 стр.

АНАЛИЗ ТЕКУЩЕЙ СИТУАЦИИ

04

16 стр.

ОБЗОР МЕЖДУНАРОДНОГО ОПЫТА

05

22 стр.

МОДЕЛЬ УПРАВЛЕНИЯ ДАННЫМИ В КАЗАХСТАНЕ

06

27 стр.

ОСНОВНЫЕ ПРИНЦИПЫ

07

29 стр.

ЦЕЛИ И ЗАДАЧИ

08

36 стр.

ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ

09

39 стр.

ЭТАПЫ РЕАЛИЗАЦИИ

ОБРАЩЕНИЕ

Стремительное развитие цифровых технологий изменило многие привычные для нас процессы и открыло уникальную возможность применения накапливаемого большого объема данных как в частном, так и государственном секторе. В этих условиях данные становятся **дополнительным фактором производительности**, стимулируя инновации и качественные решения на всех уровнях.

Однако наряду с **потенциальными выгодами** данные создают и **новые вызовы** по обеспечению их качества, доступности и конфиденциальности, созданию инфраструктуры обработки и хранения, а также по внедрению прозрачных механизмов их монетизации.

Ежедневно работая с данными как со стороны производителей, так и пользователей официальной статистической информации, мы в Бюро национальной статистики Агентства по стратегическому планированию и реформам (БНС) осознаем **значительный потенциал эффективного использования данных** и их влияния на все сферы развития государства.

В этой связи перед БНС встал **непростой вопрос** – как интегрировать между собой устоявшуюся национальную статистическую систему и зарождающуюся новую экосистему данных, наряду с этим решая отмеченные вызовы.

Внимательно изучив международный опыт и лучшие практики корпоративного сектора, мы пришли к выводу, что управление данными должно быть определено как **отдельная функция государства** и выведено в качестве **приоритетной политики**.

Для реализации этой политики важно рассматривать данные как актив, имеющий экономическую ценность и требующий эффективного управления. В этой связи на примере передовых практик управления данными в основу национальной экосистемы данных предлагается заложить **5 принципов**: актуальность, конфиденциальность, правило «только один раз», инклюзивность и востребованность.

На основе этих принципов подготовлена **Концепция развития государственной статистики и национальной экосистемы данных** (Концепция). Этот документ разработан БНС и представлен Вашему вниманию для того, чтобы **поделиться видением** того, куда движется национальная статистическая система и какие вызовы ей предстоит преодолеть для формирования **эффективной экосистемы данных**.

Многие предложения **выходят за рамки** статистической деятельности и поэтому предлагаются в рамках Концепции **для обсуждения**. Приглашаем Вас принять участие в обсуждении и внести свой вклад в формирование фундамента национальной экосистемы данных.

С нетерпением ждем Вашей обратной связи

**Руководитель БНС АСПИР
Жандос Шаймарданов**



02

ОБЩИЕ ПОЛОЖЕНИЯ

Общие положения

Представленная на публичное обсуждение Концепция разработана **с целью реформирования национальной статистической системы и раскрытия потенциала экосистемы данных.**

В Казахстане национальная статистическая система развивается в соответствии с законодательством и основополагающими принципами официальной статистики. В то же время наряду с национальной статистической системой формируется **более крупная экосистема данных**, в отношении которой в настоящее время отсутствуют какие-либо регулирующие нормы.

Согласно определению McKinsey^[1] **экосистема данных** – это платформа, которая объединяет данные от многочисленных поставщиков и создает ценность за счет использования обработанных данных.

Понятие экосистемы данных только входит в оборот, в связи с чем могут применяться иные термины, которые, по сути, эквивалентны представленному выше определению. Так, в рамках отчета Всемирного Банка «Данные для лучшей жизни»^[2] введено понятие **интегрированной национальной системы данных**, которое определяется как платформа, способствующая справедливому и безопасному обмену данными между правительством, отдельными лицами, гражданским обществом, научными кругами и частным сектором.

В рамках данной Концепции национальная экосистема данных разделена на две составляющие:

1. Государственные источники, на основе которых формируются **статистические** (собираются БНС) и **административные** данные (собираются другими государственными органами).

2. Альтернативные источники – иные не государственные источники данных, к которым могут быть отнесены мобильные операторы, маркетплейсы, социальные сети, аэроснимки Земли и т.д.

Положения Концепции распространяются как на **государственные**, так и **альтернативные источники** данных при их предоставлении для включения в информационные ресурсы статистических органов и организаций государственного сектора с учетом требований законодательства Республики Казахстан в части соблюдения прав и законных интересов владельцев данных, не являющихся государственными данными, а также соблюдения принципов невмешательства в их хозяйственную деятельность.

[1] <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/data-ecosystems-made-simple>

[2] <https://wdr2021.worldbank.org/>



Для четкого понимания подходов и разграничения ролей в рамках Концепции используются следующие понятия.

Государственные данные - данные, содержащиеся в информационных ресурсах статистических органов и организаций государственного сектора, а также в информационных ресурсах, созданных в целях реализации полномочий органов и организаций государственного сектора.

Национальный регистр – эталонная база данных, удовлетворяющая критериям качества данных, содержащая систематизированный перечень субъектов и (или) объектов и их характеристик, формируемая на основе действующих информационных систем и (или) данных соответствующих отраслей.

Управление данными - совокупность процессов сбора, передачи, хранения, обработки, предоставления, распространения, уничтожения и обеспечения качества данных, а также предоставление доступа для повторного использования.

Модель управления данными – совокупность описаний данных, организационных и технологических правил и стандартов, используемых в целях управления данными, включая описание связей между видами данных, а также между определяемыми ими объектами, в том числе в целях межведомственного обмена (взаимодействия).



03

АНАЛИЗ ТЕКУЩЕЙ СИТУАЦИИ

Анализ текущей ситуации

В качестве основы для развития национальной экосистемы данных предлагается использовать систему государственной статистики, которая на сегодня уже имеет законодательный статус и накапливает большой объем структурированных данных, отвечающих международным стандартам.

Вместе с тем, **развитие национальной статистической системы и ее трансформацию в полноценную экосистему данных** в Казахстане, также как и в общей международной практике, сдерживает **ряд системных проблем.**



НЕДОСТАТОЧНОЕ ИСПОЛЬЗОВАНИЕ АДМИНИСТРАТИВНЫХ ИСТОЧНИКОВ ИНФОРМАЦИИ, В ТОМ ЧИСЛЕ ВВИДУ ИХ НИЗКОГО КАЧЕСТВА

На БНС официально возложена координация статистической деятельности, осуществляемой в стране, однако согласно действующему законодательству БНС не может существенно влиять на формирование ведомственных данных государственных органов (административных источников). Проблема усугубляется тем, что зачастую в ведомствах отсутствует единый реестр данных, не применяются национальные классификаторы и единые справочники.

Отсутствие полномочий у БНС и закрытость административных источников, а также отсутствие единого централизованного управления данными ограничивает возможность получения данных из административных источников в полном объеме, делает практически невозможным осуществление проверки их качества и соответствия международным стандартам.

В настоящее время значительная часть административных источников сосредоточена в государственных базах данных, которые имеют **разрозненную структуру и разный уровень зрелости**, что существенно затрудняет интеграцию всех этих источников. В этой связи для кардинального сокращения статистической отчетности и перехода на использование административных источников необходим **проактивный подход к интеграции** с действующими и новыми информационными системами на основе **целевой архитектуры** данных.

В то же время, государственные органы, являющиеся владельцами информационных систем, **не всегда заинтересованы в повышении качества** ведомственных данных и соответственно их дальнейшем распространении. Накапливаемая государственными органами информация практически не публикуется на их официальных сайтах, что не позволяет должным образом проконтролировать качество административных источников. В этой связи следует распространить стандарты официальной статистики на все административные источники и наделить функцией контроля качества соответствующих дата стьюардов.

**НЕВОЗМОЖНОСТЬ ИСПОЛЬЗОВАНИЯ АЛЬТЕРНАТИВНЫХ ИСТОЧНИКОВ ИНФОРМАЦИИ И ИХ ПОЛНОЦЕННОГО ВОВЛЕЧЕНИЯ В ЭКОНОМИЧЕСКИЙ ОБОРОТ**

Сегодня каждый пользователь информационных ресурсов оставляет **цифровой “след”**, который накапливается как в государственных, так и альтернативных источниках. Эти данные в сочетании с другими – общедоступными или частными – открывают большие возможности для выявления закономерностей, производства опережающих показателей, разработки новых бизнес-моделей или продуктов, а также принятия решений для общего блага казахстанцев.

В то же время доступ к этим данным ограничен по причинам, связанным с правилами конфиденциальности и конкуренции, а также последствиями, связанными с затратами на получение и обмен информацией.

Государство **упускает возможность оперативного реагирования** на те или иные острые проблемы, **не используя альтернативные источники** – показатели транзакционной активности (фискальных органов и платежных систем), данные государственных монополистов (объемы грузоперевозок, потребления электроэнергии), данные мобильных операторов и другие показатели в режиме реального времени.

В отличие от официальной статистики альтернативные данные формируются быстрее и практически в любой момент времени. В перспективе после подтверждения качества и стабильного доступа к источникам они могут заменить значительную часть статистической отчетности.

Альтернативные источники уже сегодня активно используются в целях **монетизации данных**, таким образом, извлекая **экономическую выгоду** из данных. В то же время заключаемые сделки по монетизации данных носят закрытый характер, что не гарантирует соблюдение правил конфиденциальности и добросовестной конкуренции на данном рынке.

Национальная экосистема данных в полной мере **раскроет свой потенциал** в том случае, если сделки на рынке данных будут иметь **прозрачный характер без вмешательства государственных институтов**.

**ОТСУТСТВИЕ ПОЛИТИКИ УПРАВЛЕНИЯ ДАННЫМИ**

Отсутствие политики управления данными становится все более актуальной задачей, усугубленной множеством новых законодательных требований, а также увеличением объемов данных, поступающих из многочисленных структурированных и неструктурированных источников.

Для эффективного управления национальной экосистемой данных необходимо учитывать области знаний по управлению данными согласно определениям **рамочной структуры управления данными (Data Management Association)**. Ниже на рисунке 1 представлена **экспертная оценка зрелости** казахстанской экосистемы данных по каждой из 10 функций рамочной структуры управления данными (**колесо DAMA**).

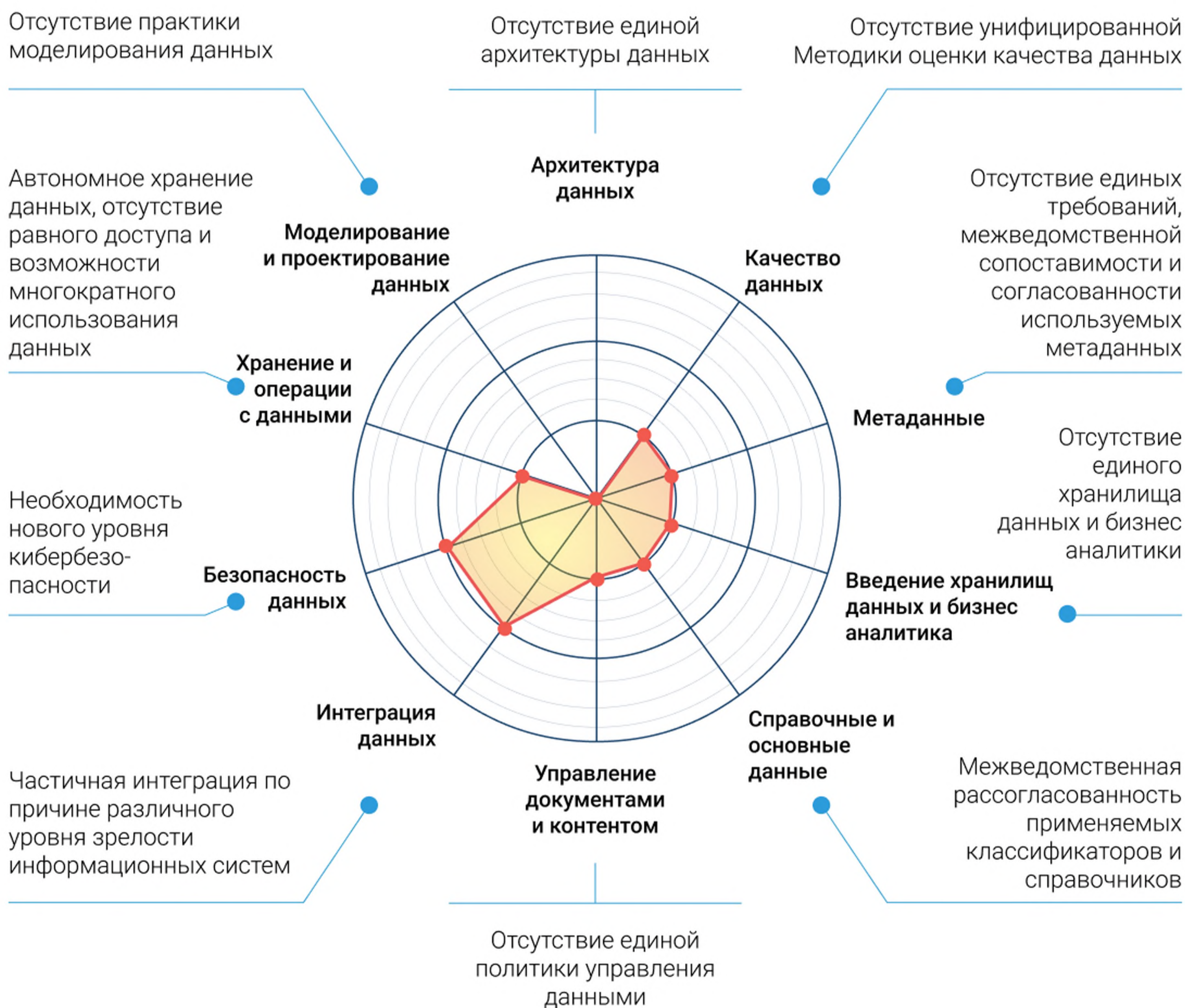


Рисунок 1. Рамочная структура управления данными (DAMA International)

Сложно реализовать общегосударственные инициативы или обеспечить функциональную совместимость, согласованность и обмен данными из-за отсутствия межведомственных стандартов данных, требований к политике управления данными. В этой связи вопросы управления данными и информацией, доступа к данным, защиты конфиденциальности и безопасности данных, качества данных, архитектуры данных являются **ключевыми компонентами политики управления данными**.



УСТАРЕВШИЕ МЕТОДЫ СБОРА И МНОГОСТУПЕНЧАТАЯ СИСТЕМА ОБРАБОТКИ ДАННЫХ

Серьезной проблемой для развития системы управления национальными данными являются **устаревшие методы** сбора и обработки данных.

Несмотря на то, что **в 2021 году более 80%** статистической отчетности представлено респондентами в **электронном виде**, сохраняются традиционные бумажные каналы предоставления отчетности. Сохранение одновременно электронного и бумажного форматов отчетности **существенно снижает качество и достоверность, увеличивает сроки** сбора и обработки данных, а также приводит к **существенным дополнительным расходам** бюджетных средств.

Используемый в настоящее время традиционный способ сбора статистической отчетности предполагает заполнение форм отчетности респондентами в ручном формате, что приводит к **появлению ошибок** и не исключает возможность **внесения недостоверной информации**.

Кроме того, многоуровневая система сбора и обработки данных приводит к увеличению сроков обработки статистической отчетности и увеличению количества ошибок, и как следствие снижения качества и актуальности формируемой статистической информации.



Рисунок 2. Многоступенчатая схема сбора и обработки данных

Аналогичная ситуация наблюдается и в других государственных органах и учреждениях, формирующих ведомственную отчетность для своих целей. Наибольшая доля бумажной отчетности приходится на МИО. Это приводит к **избытку хранимых данных, использованию различных форматов и технологий** в отношении данных. Как следствие, формируется **высокая стоимость** (или невозможность) увязки данных, а также отсутствует обзор того, какая информация доступна.



ДУБЛИРОВАНИЕ ОТЧЕТНОСТИ

Наряду со статистическими наблюдениями **государственные органы вводят дополнительные ведомственные формы отчетности**, реализуя полномочия по сбору информации в целях исполнения возложенных на них полномочий.

Основной причиной введения дополнительных форм ведомственной, региональной отчетности является законодательно установленный запрет субъектам официального статистического учета предоставлять первичные статистические данные для использования в иных целях, не связанных с формированием официальной статистической информации

Казахстан остается одной из немногих стран, в которых практически **не скоординирован сбор** статистической и налоговой отчетности. К примеру, статистические формы о финансово-хозяйственной деятельности предприятий своим содержанием, практически полностью идентичны налоговым декларациям, которые представляются в Комитет государственных доходов Министерства финансов РК. При этом, **отсутствие интеграции** БНС к информационным системам налоговых органов не позволяет в настоящее время отменить статистическую отчетность.

В данной ситуации не работает принцип эффективного управления данными «**только один раз**». Использование существующих данных должно управляться эффективно, сбор дублирующих данных не должен иметь место, если есть государственный орган, который уже собрал необходимые данные от респондента.

Несогласованность статистической и налоговой отчетности, а также отсутствие доступа к административным источникам в полном объеме является **препятствием для исключения дублирующей отчетности и соответственно снижения нагрузки на бизнес.**



НИЗКИЙ УРОВЕНЬ ДОВЕРИЯ ПОЛЬЗОВАТЕЛЕЙ

Отсутствие межведомственной коммуникационной политики, полной методологической взаимосогласованности и совместимости данных из различных источников, а также существующая закрытость административных источников существенно **сокращает доверие к официальной статистике**.

Статистические обследования различного характера могут опираться на различные подходы, классификации и методологии, поэтому не могут предоставить пользователю полностью увязанную информацию. Нередки случаи, когда пользователи сталкиваются с тем, что два различных измерения одной и той же величины публикуются с различными значениями. Не всегда имеется достаточная информация о метаданных и оценке качества данных, отсутствует сопоставимость между предметными областями.

В 2020 году со стороны населения **особое недоверие** было вызвано к публикуемым данным о смертности от Covid-19. При анализе данных по показателям заболеваемости и смертности, формируемых БНС и Министерством здравоохранения РК, выявлены ряд **несоответствий и значительные расхождения данных**. Причиной расхождения данных стал различный подход при формировании статистической и ведомственной информации и отсутствие единой информационной системы между медицинскими организациями.

Так, Министерство здравоохранения РК считало обусловленными Covid-19 только те смерти, которые они диагностировали при положительном ПЦР-тесте у пациента. В случае, если анализ не проводился, то ставился другой диагноз или причина смерти была неопределенной. Кроме этого, необоснованно разделялись случаи пневмонии и ковида. Это существенно отличалось от подхода в передовых странах.

Для обеспечения **доверия к статистике** и в целом к информации, предоставляемой государственными органами, методология и политика должны быть **транспарентными, задокументированы и опубликованы для всеобщего доступа и правильной интерпретации**.



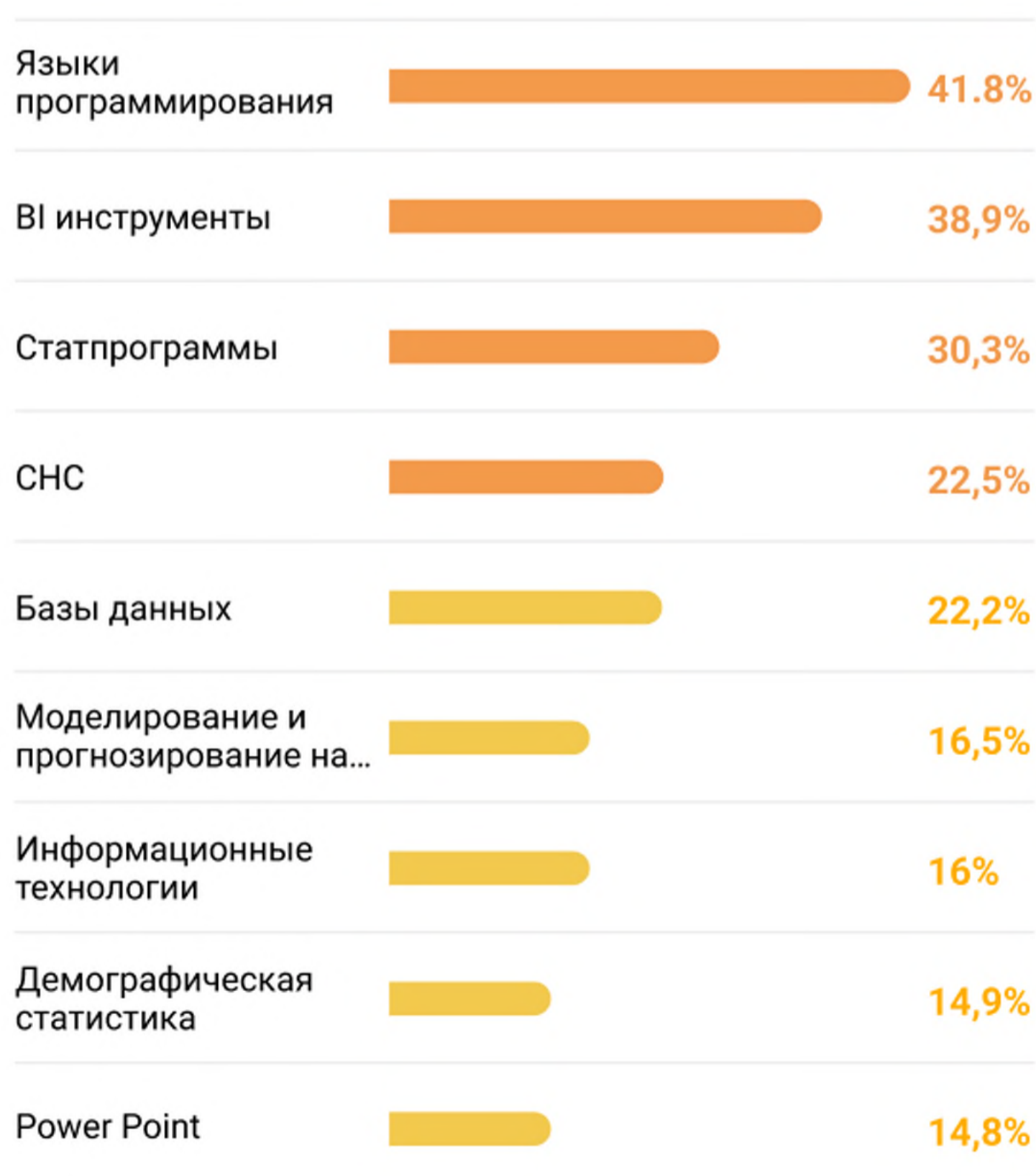
НИЗКИЙ УРОВЕНЬ КОМПЕТЕНЦИЙ ПО АНАЛИЗУ ДАННЫХ

Для постоянного улучшения качества формируемой статистики необходимо **развитие навыков по работе с данными** во всем государственном аппарате.

На сегодня сотрудники БНС являются **преимущественно статистиками**, имеющими экономическое образование и **ограниченные компетенции в бизнес-аналитики**. Практически отсутствуют специалисты с компетенциями в программировании и управлении данными. В этой связи публикуемая статистическая информация не содержит глубокого анализа, перекрестную аналитику взаимосвязанных показателей, необходимую для своевременных и упреждающих решений в режиме реального времени.

Проведенный в августе 2022 года внутренний анализ (gap-analysis) на основе самооценки самих сотрудников БНС выявил ряд проблем текущего уровня развития компетенций. В опросе приняли участие **2 221 работников БНС**. Выяснилось, что более **63%** опрошенных специалистов желают расширить свои навыки в той или иной области знаний, в частности получить навыки работы с современными **ВИ-инструментами, программировании**, а также получить необходимые **навыки бизнес-аналитики**.

НИЗКАЯ КОМПЕТЕНТНОСТЬ



ВЫСОКАЯ ПОТРЕБНОСТЬ

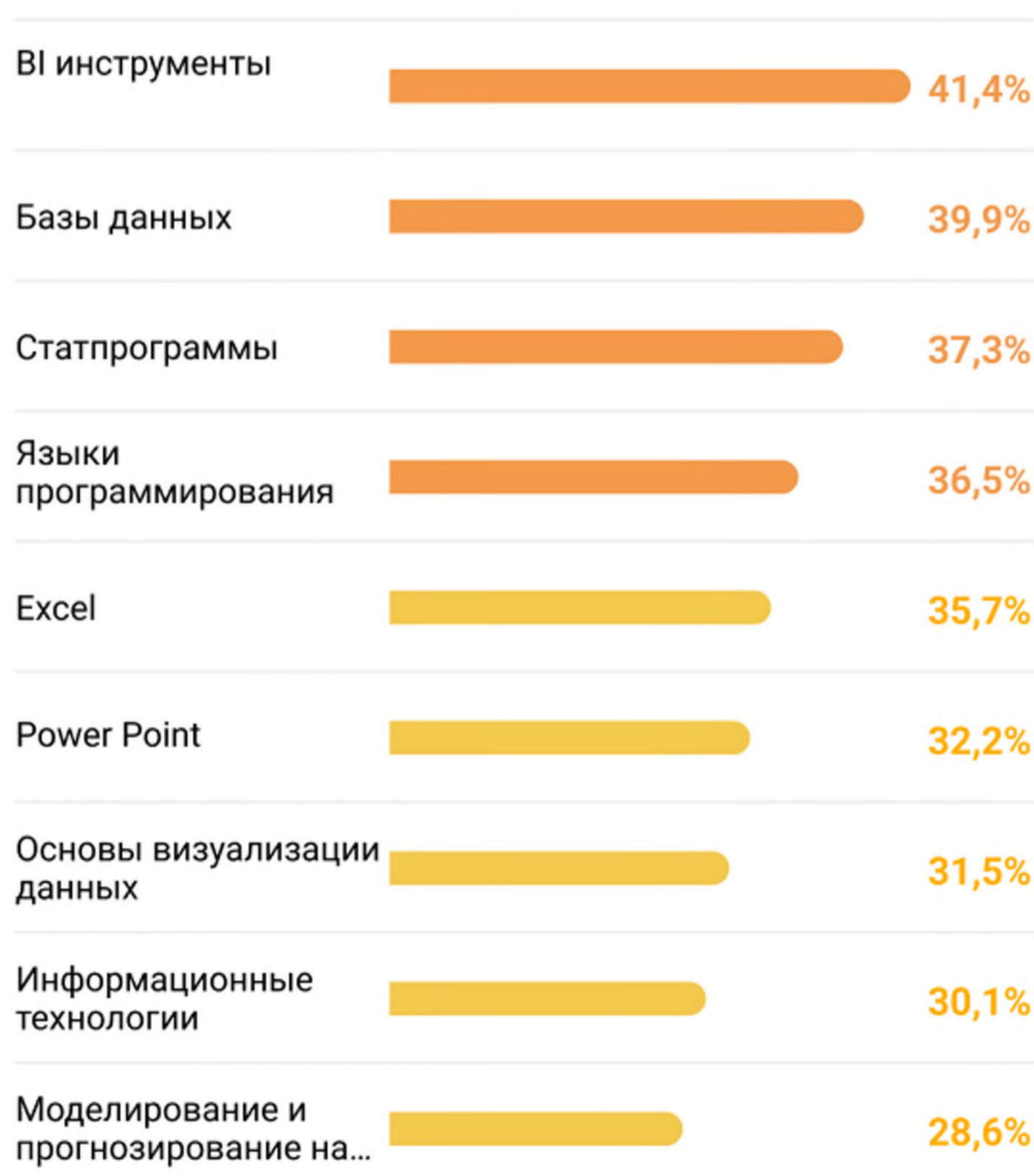
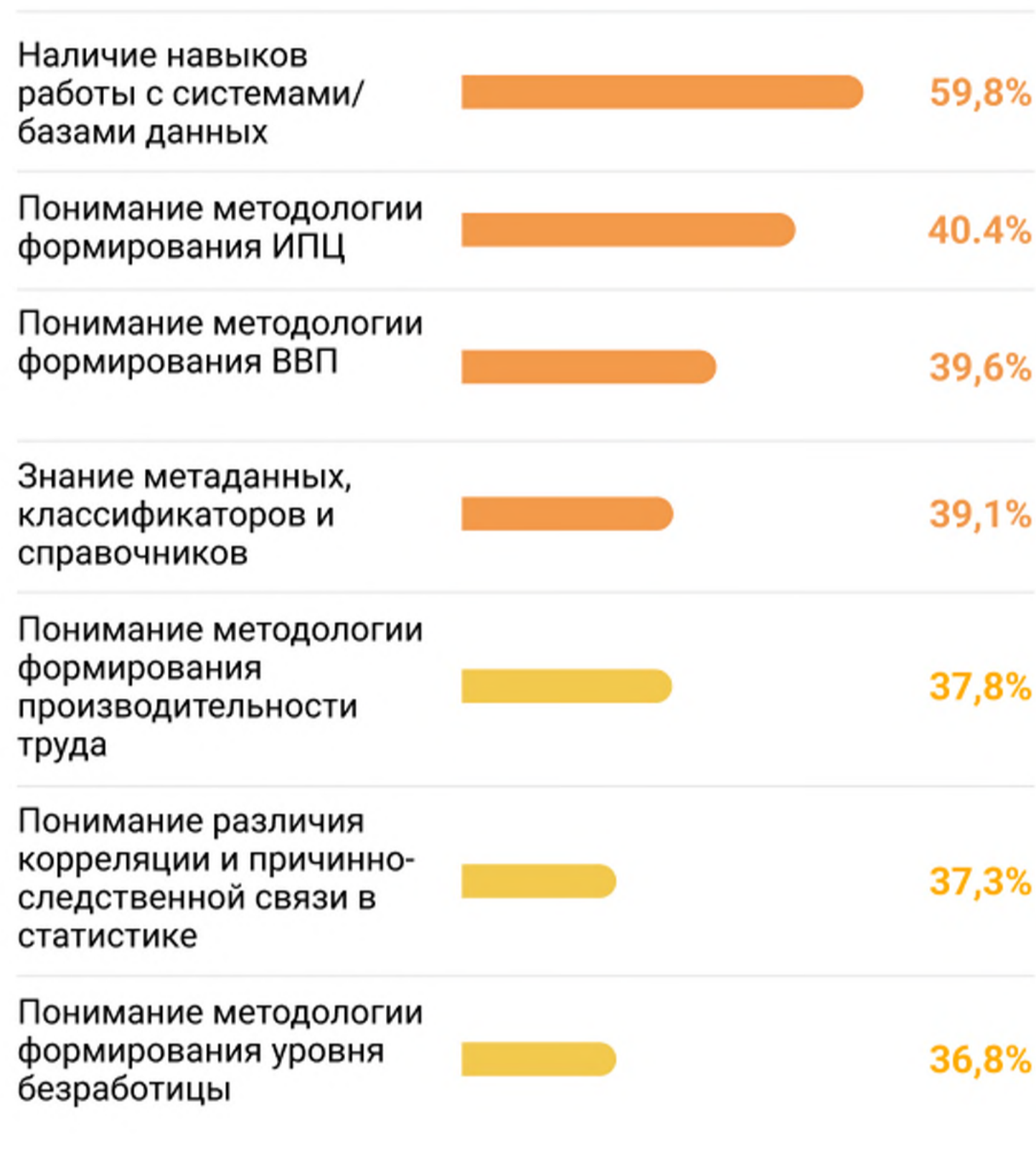


Рисунок 3. Результаты gap-analysis работников БНС

Следует обратить внимание, что в **других государственных органах** складывается **аналогичная ситуация**, наблюдается значительная нехватка специалистов, имеющих навыки работы с данными. Кроме того, у государственных служащих отсутствуют **базовые знания в области статистики**, необходимые для корректного использования официальной статистической информации при принятии решений.

В опросе приняли участие 30 ЦГО и 20 МИО. Выяснилось, что **63,8%** опрошенных государственных служащих заинтересованы в обучении по **статистической грамотности и работе с данными**.

НИЗКАЯ КОМПЕТЕНТНОСТЬ



ВЫСОКАЯ ПОТРЕБНОСТЬ



Рисунок 4. Результаты gap-analysis государственных служащих

Таким образом, все перечисленные проблемы затрудняют **развитие национальной экосистемы данных** и требуют решения в рамках представленной Концепции.



04

ОБЗОР МЕЖДУНАРОДНОГО ОПЫТА

Обзор международного опыта

По итогам 2019 года национальная статистическая система Казахстана вошла в **ТОП-50** мирового рейтинга в рамках **оценки статистической эффективности Всемирного Банка**^[1], набрав **78,6 баллов из 100**. В качестве положительных моментов отмечены правовая основа и мандат официальной статистики, достаточно высокий уровень качества и защиты формируемой информации, использование методологии, соответствующей международным стандартам и требованиям.

Существенная модернизация национальной статистической системы Казахстана была проведена в 2012-2017 годах совместно со Всемирным Банком в рамках проекта **«КАЗСТАТ»**.

Следующим важным шагом, по мнению группы экспертов Всемирного Банка, станет построение **интегрированной национальной системы данных** (экосистемы данных), в которой **одна из ключевых ролей** отводится **национальному статистическому управлению (НСУ)**.

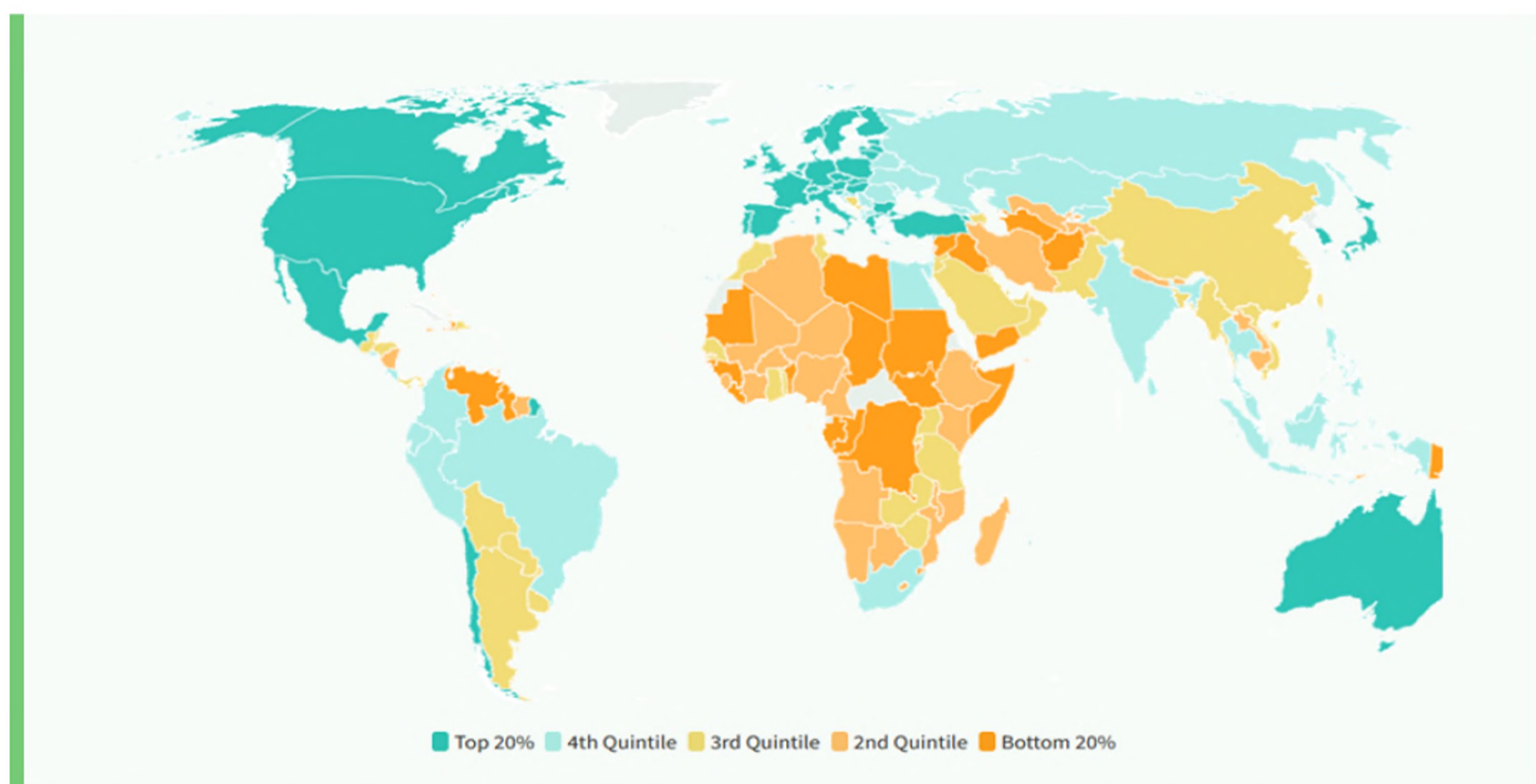


Рисунок 5. Статистические показатели эффективности (SPI) для 174 стран

Создание интегрированной национальной системы данных – актуальная задача, соответствующая современным потребностям и возможностям практического воплощения Data-Driven-подхода в государственном управлении. В свою очередь, значительно повысить качество управленческих процессов позволит **эффективная модель управления данными**.

Международная ассоциация DAMA выделяет три **типа моделей управления данными** в зависимости от того, за кем закреплена роль распорядителя данных: **централизованная, распределенная и гибридная модели**.

[1] <https://www.worldbank.org/en/programs/statistical-performance-indicators>

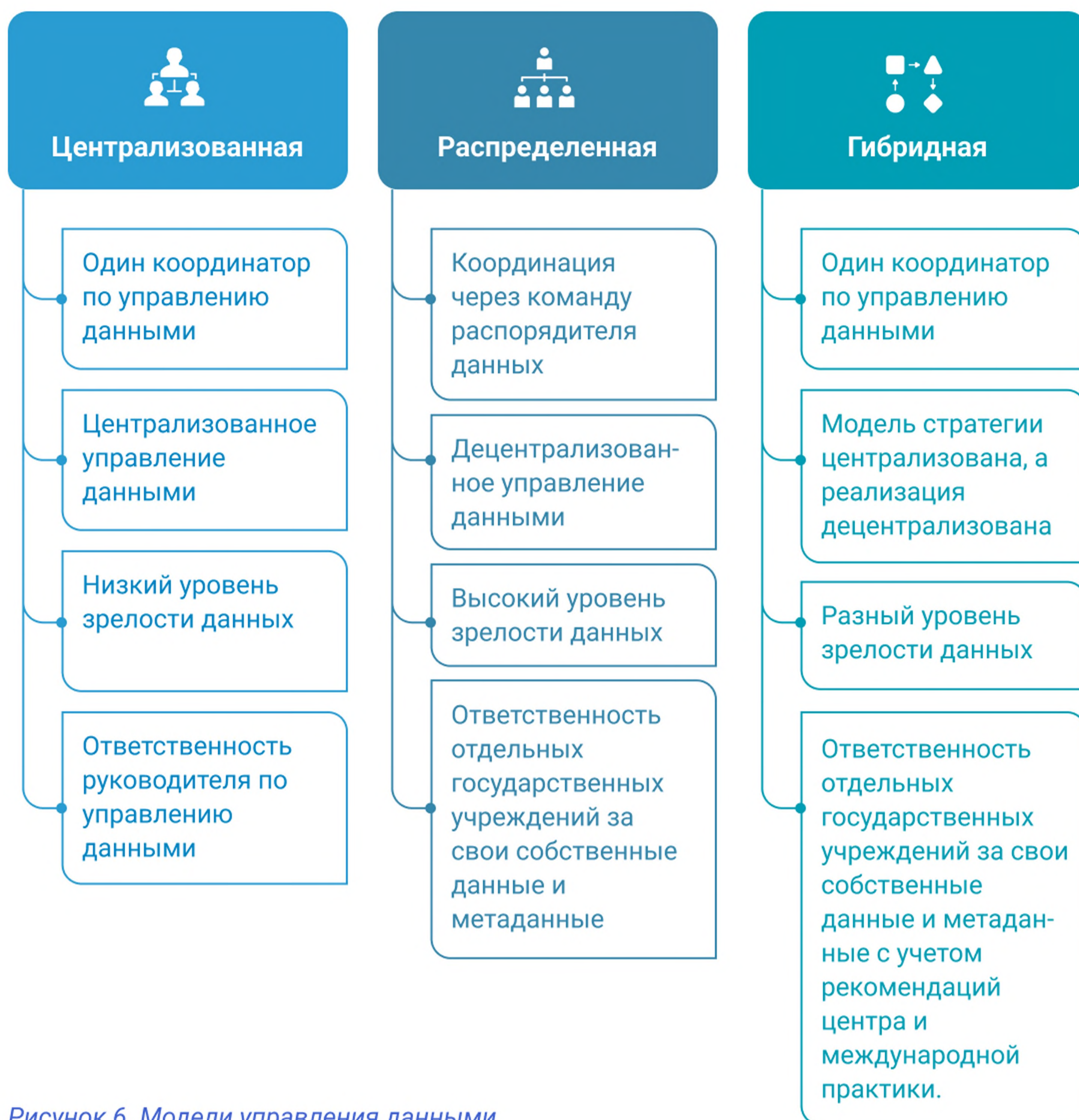


Рисунок 6. Модели управления данными

Согласно отчету **ОЭСР** по **Индексу цифрового правительства** за 2020 год [1], «**только 12%** стран имеют **единую специальную политику** (или стратегию) в отношении данных, а **82%** внедряют данные в рамках более широкой политики (например, цифрового правительства или открытых данных)».

В то же время многие страны (США, Канада, Великобритания, Германия, Новая Зеландия) уже **разработали национальные стратегии** в отношении данных, чтобы объединить в рамках единого политического инструмента различные аспекты, начиная от доступа к данным до обмена и этики использования данных.

[1] <https://www.oecd.org/gov/digital-government-index-4de9f5bb-en.htm>

В международной практике имеется несколько подходов внедрения различных моделей управления данными. Ниже представлен опыт стран, которые наиболее **активно внедряют** управление данными на **национальном уровне**.

Наименование страны	Модель управления	Владелец данных	Координатор	Роль НСУ
 Литва	Централизованная	Госорган	НСУ	Руководящая роль
 Великобритания, Новая Зеландия	Распределенная	Госорган + Информационная группа	НСУ (Chief Government Data Steward)	Руководящая роль
 Австралия	Распределенная	Госорган	Департамент обслуживания клиентов	Нет данных
 Канада	Распределенная	Госорган	НСУ Директор по данным (CDO)	Нет данных
 Россия	Распределенная	Госорган	НСУ	Руководящая роль
 Финляндия	Гибридная	Госорган + Совет по управлению информацией	Эксперт по знаниям в области управления данными	Членство в Совете по управлению данными
 Мексика	Гибридная	Госорган	Эксперт по знаниям в области управления данными	Руководящая роль
 Эстония	Гибридная	Госорган	НСУ	Руководящая роль

Рисунок 7. Сравнение моделей управления данными

Литовской Республикой выбрана **централизованная модель** управления данными, все государственные данные собираются в одних руках – Национальном агентстве данных (ранее Департамент статистики). Переход на централизованную модель позволяет существенно изменить подходы по управлению данными, создать дополнительные инструменты для сбора и использования качественных данных.

Трансформация в Национальное агентство позволяет НСУ решать новые задачи и стать эффективным государственным органом по управлению данными, который объединяет в одних руках данные из разных государственных учреждений. Такое изменение обеспечивает унифицированный подход к управлению данными, исключает дублирование собираемых данных, а используемые инструменты контроля качества существенно повышают доверие пользователей к публикуемой информации, а также создает широкие возможности для принятия правильных политических решений.

Принятые в 2022 году поправки к закону об официальной статистике и управлении государственными данными Литовской Республики устанавливают новый подход к управлению государственными данными и регулированию их обработки, снижают административную нагрузку на учреждения, предоставляющие данные, обеспечивают более гибкое использование данных, собранных в государственном секторе, а также более быстрый обмен данными между системами государственных данных. Закон также определяет обязанности поставщиков данных – респондентов, способы предоставления статистических данных, расширил объем источников данных за счет больших данных, собираемых частными компаниями.

Великобритания, Новая Зеландия, Австралия формируют отдельную группу и осуществляют управление данными, используя распределенную модель управления, которая имеет ряд преимуществ, но также некоторые ограничения с точки зрения системных подходов к управлению данными.

Мексика, Финляндия и Эстония выбрали для своих стран гибридную модель управления данными, которая имеет прямой контроль над производственным циклом проводимых обследований для производства статистической информации и переписей, а также статистических данных, полученных на основе использования административных записей и других источников.

В отношении институциональных структур и руководящих ролей такие страны, как Новая Зеландия, Финляндия, Австралия, Мексика, Эстония предпочли определить конкретные должности главного дата стюарда по данным для всего правительства, в то время как другие предпочли создать органы, такие как Советы по данным, в которых руководители данных на институциональном уровне обсуждают приоритеты и договариваются об их скоординированной реализации.

В каждом учреждении, которое предоставляет информацию или собирает статистические данные, представляющие национальный интерес, имеется координатор, который отвечает за реализацию политики и стандартов, определенных уполномоченным в сфере управления данными. Координатор в лице НСУ разрабатывает методологию, требования и стандарты, которым должны следовать другие государственные учреждения, для получения данных при подготовке статистической информации.

Так, в **Новой Зеландии** создана должность Главного распорядителя данных правительства (GCDS), признав, что для того, чтобы помочь Новой Зеландии реализовать данные как актив, требуется сильное системное руководство. Комиссар по государственным услугам назначил на эту роль главу Статистического управления Новой Зеландии (Stats NZ).

В сочетании с ролью GCDS, Stats NZ выступает в качестве ведущего агентства по данным в рамках системы данных правительства Новой Зеландии, способствуя и обеспечивая объединенный подход к возможностям и проблемам, связанным с данными. В этом качестве Stats NZ вместе с GCDS стремится поддерживать агентства, чтобы максимально использовать потенциал их данных и обеспечивать их эффективное использование, сохраняя при этом доверие новозеландцев.

Группу лидеров цифрового правительства (DGLG) возглавляют главный правительственный распорядитель данных и главный правительственный сотрудник по цифровым технологиям.

В сентябре 2022 года вступил в силу «Закон о данных и статистике», подготовленный на основе обзора законодательства в области статистики и консультаций с общественностью по предложениям высокого уровня в отношении новых данных.

В Мексике Национальный институт статистики и географии (INEGI) является автономным органом, полностью независимым от федерального правительства. INEGI играет двойную роль в национальной информационной системе: как непосредственный производитель официальной информации и в качестве координатора, устанавливающего нормы и стандарты производства официальной статистики в разных государственных учреждениях.

Правительство Канады отдает приоритет инновациям и увеличению «горизонтальности». В стране создана должность Директора по управлению данными, который несет ответственность за обеспечение общего руководства управлением данными и информацией, разработку административных рамок и стандартов управления и качества.

Новая экосистема государственных данных и правильно выстроенная модель управления данными позволяет быстро собирать, объединять и визуализировать данные, собранные в различных регистрах и информационных системах, в соответствии с решаемой задачей.

Таким образом, практически во всех странах ключевая роль координатора политики управления данными отведена НСУ, имеющим законодательный доступ к персональным данным, соответствующие полномочия и навыки в работе с данными.

В контексте Казахстана опыт статистической деятельности и существующие полномочия уполномоченного органа в области государственной статистики могут определить БНС ключевым игроком в сфере политики управления данными.

The background features a complex, abstract data visualization. It consists of numerous small numbers in various colors (red, blue, white) scattered across the space. These numbers are connected by thin, colored lines that form a network or flow. The overall effect is that of a dynamic, interconnected data system. In the bottom right corner of this section, there is a stylized logo consisting of two parallel diagonal lines, one light blue and one white, set against a dark blue background.

05

МОДЕЛЬ УПРАВЛЕНИЯ ДАННЫМИ В КАЗАХСТАНЕ

Модель управления данными в Казахстане

На основе представленного международного опыта внедрения трех моделей управления данными, а также с учетом текущего состояния национальной статистической системы и развития государственных информационных систем в Казахстане предлагается **гибридная модель управления данными**.

Гибридная модель — это разновидность модели управления данными с **централизованным контролем и распределенным управлением**. Иными словами, в гибридной модели стратегия управления централизована, а реализация децентрализована, позволяя консолидировать преимущества централизованной и распределенной моделей, при этом существенно сократив их отдельные недостатки.

Для экосистемы с **сильно различающимся уровнем зрелости** информационных систем гибридная модель является **наиболее подходящим вариантом** для общегосударственного подхода.

Принятие решений в отношении стандартов и общих данных осуществляется централизованно. Процессы, стандарты, руководящие принципы и системы обычно устанавливаются основной командой, целевой группой или советом, а отдельные участники несут ответственность за их применение на локальном уровне.

Модель может включать принятие общей модели данных различными организациями, также может принять общий подход для всей экосистемы (межведомственный), чтобы служить ответом на проблему с данными в конкретной области.

Основные преимущества гибридной модели:

- Государственные органы несут ответственность за свои собственные данные.
- Она предоставляет автономию для разработки стандартов, политик и процедур на уровне отдельных участников, но требует, чтобы они соответствовали общегосударственной модели обмена данными для обеспечения функциональной совместимости.
- Решение проблем осуществляется по принципу «снизу вверх» (должна быть четкая ответственность и процесс для решения вопроса).

Принцип гибридной модели требует установления норм (законов, административных положений), необходимых для повторного использования данных. Принятие гибридной модели может потребовать **законодательных изменений**, чтобы избежать умножения переговоров для получения ресурсов в каждом конкретном случае. Это включает создание институциональных механизмов и уточнение ролей, необходимых для того, чтобы эта модель работала.

В случае принятия данной модели потребуется **распорядитель данных** на общегосударственном уровне.

Руководство данными определяется как деятельность по осуществлению руководящих и контрольных полномочий (планирование, мониторинг и обеспечение выполнения) в отношении управления информационными активами.

Опыт статистической деятельности и существующие полномочия уполномоченного органа в области государственной статистики могут определить **БНС ключевым игроком в сфере политики управления данными.**

Сегодня БНС – **единственный государственный орган**, имеющий значительный опыт формирования официальной статистической информации, который имеет экспертные знания в области этики данных, высокую зрелость в отношении данных, законодательный доступ к персональным данным и ресурсам. БНС обладает современными методами обеспечения качества и защиты данных.

Таким образом, учитывая большой опыт в производстве статистических данных и показателей, управления данными, доступа к ним и практики обмена ими в государственном секторе, национальная статистическая система в лице БНС должна стать распорядителем в экосистеме данных.

Роль координатора управления данными является ключевой и заключается в определении и координации политики управления данными. Государственные данные должны позиционироваться как общественное благо и, соответственно, БНС предстоит перестроить подходы к политике управления данными.

Координатор по управлению данными владеет полной информацией о том, где какие данные находятся, откуда они берутся, что означают, кто ими пользуется и каким образом они могут быть предоставлены.

Основная задача координатора по управлению данными выстроить межведомственную координацию, обеспечивающую использование унифицированной методологии, требований и стандартов для обеспечения качества государственных данных.

Мониторинг спроса на данные осуществляет координатор управления данными путем проведения анализа потребностей.

Роль координатора управления ИТ заключается в определении и координации политики управления ИТ, в том числе в части информационной безопасности.

В зоне ответственности координатора по управлению ИТ находятся процессы, связанные с цифровизацией. В фокус его внимания попадает обновление процессов за счет внедрения цифровых решений, в том числе работа с реальными данными и стратегический взгляд на области знаний по управлению данными.

Основная задача координатора по управлению ИТ выстроить процессы и внедрить программные решения таким образом, чтобы обеспечить быстрый и безопасный доступ сотрудников организации к необходимым данным.

Координатор управления IT – уполномоченный орган в сфере информатизации в Республике Казахстан.

Роль координатора обеспечения безопасности состоит в том, чтобы исключить утечку данных, доступ к которым ограничен нормативными правовыми актами, включая сведения, составляющие служебную информацию ограниченного распространения или государственную тайну.

Координатор обеспечения безопасности – уполномоченный орган в сфере обеспечения безопасности в Республике Казахстан.

Владельцы данных - собственники государственных или альтернативных источников информации, которые принимают решения о данных в предметной области и несут ответственность за их сохранность, в то же время обеспечивая качество данных. Роли в области управления данными могут определяться как на функциональном, так и на индивидуальном уровне. При этом названия ролей, их относительная значимость и потребность могут варьироваться.

При развитии национальной экосистемы данных на государственном уровне необходимо **скоординированное взаимодействие трех ключевых ролей.**



Рисунок 8. Архитектура взаимодействия участников экосистемы государственных данных

Дата стюард – специалист по управлению данными, обеспечивающий координацию системы управления данными национального регистра^[1] и имеющий доступ к информационным системам, определяемым Правительством РК в рамках перечня национальных регистров. В зоне ответственности дата стюарда находится надзор и управление данными, а также обеспечение их качества.

Дата архитектор – сотрудник, разрабатывающий архитектуру сбора, хранения и выгрузки данных, определяет источник данных, устанавливает требования по политике управления данными.

Цифровой офицер – менеджер, отвечающий за цифровую трансформацию, определяющий стратегию цифровизации, обеспечивающий координацию с государственными цифровыми платформами и проектами.

^[1] Дата стюарды будут закреплены за национальными регистрами, а не государственными органами, которые регулярно реорганизуются. Подробнее о национальных регистрах изложено в разделе «Цели и задачи»



06

ОСНОВНЫЕ ПРИНЦИПЫ

Основные принципы

В настоящее время национальная статистическая система Казахстана опирается на **основополагающие принципы официальной статистики**, принятые **Организацией Объединенных Наций**[1]. В то же время в целях развития национальной экосистемы данных следует также учесть **принципы управления данными DMВОК** [2].

На основе этих двух не противоречащих друг другу систем сформированы следующие 5 принципов, признающие **данные как актив: актуальность, конфиденциальность, правило «только один раз», инклюзивность и востребованность**.

А

Актуальность

Актуальность (своевременность) информации – это степень соответствия информации текущему моменту времени. Актуальность информации определяется тем, насколько своевременны и важны для человека или общества данные сведения, могут ли они быть использованы в конкретной ситуации для решения проблем или принятия решения.

К

Конфиденциальность

Персональные данные, собранные государственными органами, независимо от того, касаются они физических или юридических лиц, должны носить строго конфиденциальный характер и использоваться исключительно в установленных законодательством целях.

Т

Только один раз

Использование существующих данных должно управляться эффективно, сбор дополнительных данных не должен иметь место, если имеется государственный орган, который уже собрал необходимые данные. Данный принцип обеспечивает стимул для установления общих правил и стандартов, способствующих функциональной совместимости.

И

Инклюзивность

Экосистема данных должна обеспечить равный доступ к данным на беспристрастной основе в интересах реализации права граждан на общедоступную информацию. Использование данных должно, в свою очередь, приносить справедливую долю ценности участникам экосистемы.

В

Востребованность

Востребованность данных представляет собой качественную оценку степени их нужности пользователям и, соответственно той ценности, которая извлекается из их использования. Учитывая возможность повторного использования данных, их ценность прямо пропорциональна количеству их использования. Чем больше пользователей у конкретного набора данных, тем больше его ценность.

Рисунок 9. Основные принципы управления данными

[1] <https://unstats.un.org/fpos/>

[2] <https://www.dama.org/cpages/body-of-knowledge>



07

ЦЕЛИ И ЗАДАЧИ

Цели и задачи

Главной целью данной Концепции является **раскрытие потенциала национальной экосистемы данных** на основе принципов, перечисленных в предыдущем разделе. Достижение цели будет осуществляться за счет реализации следующих задач:

ЗАДАЧА 1.

СОВЕРШЕНСТВОВАНИЕ НОРМАТИВНОЙ ПРАВОВОЙ И МЕТОДОЛОГИЧЕСКОЙ БАЗЫ

Для обеспечения доступа к данным необходимы **правовая основа** в сфере управления государственными данными, законодательная регламентация требований и доступа, порядок взаимодействия информационных систем, правила создания модели государственных данных, определение эталонных данных, приведение государственных данных в соответствие с едиными требованиями по управлению данными. Данные инициативы будут отражены в Законопроекте «О государственной статистике и данных».

Кроме того, для обеспечения качества и достоверности данных необходима **методологическая «перезагрузка»** национальной статистической системы и всех государственных органов.

В настоящее время в структуре БНС создан Экспертный совет и управление методологической координации и анализа для обеспечения качества методологического сопровождения статистической деятельности. Особое внимание уделяется вопросам унификации методологических подходов и преодолению методологической рассогласованности между ведомствами.

Важным элементом реализации этой задачи станет проведение **Глобальной оценки** казахстанской статистической системы со стороны **ЕЭК ООН** в **начале 2023 года**. Дополнительно в течение 3 лет (2023-2025 годы) также будет проведен **углубленный обзор ОЭСР** с дальнейшим получением статуса «члена Комитета ОЭСР по статистике и статистической политике».

«Перезагрузка» методологии станет первым этапом внутренней трансформации БНС и будет способствовать повышению доверия к официальной статистике.

ЗАДАЧА 2.

АКТИВНОЕ ИСПОЛЬЗОВАНИЕ АДМИНИСТРАТИВНЫХ И АЛЬТЕРНАТИВНЫХ ИСТОЧНИКОВ НА ОСНОВЕ НАЦИОНАЛЬНЫХ РЕГИСТРОВ

Эта задача потребует модернизации информационных систем государственных органов, а также расширения вычислительных мощностей БНС, максимальной автоматизации всех внутренних процессов сбора и обработки данных, кардинальной смены ИТ-архитектуры госаппарата и подходов в управлении данными в соответствии с международными стандартами управления данными (DAMA).

Для реализации эффекта синергии между различными источниками и использования административных и альтернативных данных в полном объеме потребуется **единая политика управления данными** и **новая целевая архитектура данных**. Они позволят выработать единый подход к управлению данными и унифицировать методы их обработки и анализа.

При этом, в целях повышения качества данных административных источников необходимо создание **централизованных национальных регистров** по всем сферам экономики и жизнеобеспечения страны.

Национальные регистры будут представлять собой эталонные базы, которые хранят только очищенные данные, то есть прошедшие форматно-логические и камеральные контроли, таким образом решив существующую проблему качества и достоверности данных.

Национальные регистры станут базой для государственных услуг, также интерфейсом взаимодействия между государственными органами, бизнес-структурами и частными лицами как единственный источник достоверных данных. Таким образом, решится вопрос дублирования данных в государственных базах данных разных ведомств, будут исключены множественные административные процедуры взаимодействия с госорганами, что станет основой дебюрократизации госаппарата.

Национальные регистры позволят формировать **цифровые статистические регистры** путем сбора достоверных данных по заранее согласованному формату с определенной регулярностью. Это значительно **снизит нагрузку на респондентов и повысит точность статистического производства**.

В настоящее время БНС совместно с МЦРИАП приступили к проекту **«Переход на цифровые статистические регистры»**. Он предполагает пересмотр всех действующих показателей статистики, анализ существующих баз данных государственных органов и форм отчетности, а также формирование полностью цифровых статистических регистров, которые будут наполняться на основе административных данных. Проект также позволит сопоставить существующие данные в официальной статистике и в действующих государственных базах данных, и совместно с госорганами улучшить качество данных.

ЗАДАЧА 3.

УКРЕПЛЕНИЕ АНАЛИТИЧЕСКОГО ПОТЕНЦИАЛА ГОСУДАРСТВЕННЫХ ОРГАНОВ

Для **укрепления аналитического потенциала** и проведения глубокой бизнес-аналитики необходимо формирование аналитической культуры, в которую входит формирование библиотеки показателей и стандартизация визуализации, формирование наделенных полномочиями по оптимизации бизнес-процессов, создание **центра компетенций по повышению грамотности работы с данными**.

Кроме этого, требуется государственная политика с долгосрочной направленностью на стратегию, культуру, этику, роли и условия, а также возможности людей, которые могут поддерживать эффективную экосистему данных.

Необходимо провести оценку того, насколько действующая система подбора персонала является эффективной для удовлетворения специальных потребностей национальной статистической системы и развития экосистемы данных, на основе которой в дальнейшем будут приняты меры по адаптации процесса комплектования персонала к потребностям.

В этом контексте БНС начата реализация инициативы **Data Literacy**, которая направлена на **повышение грамотности** государственных служащих по работе с данными. Недостаточно просто оснастить сотрудников аналитическими инструментами, необходимо научить их работать с данными, анализировать информацию и делать из нее выводы.

Смещение парадигмы в сторону **Data driven decision making** (от сложившего на практике Decision driven data making) требует обучить пользователей критически относиться к полученным результатам, а также обосновывать решения на основе данных.

Для массового охвата программой Data Literacy БНС формирует **серию обучающих онлайн-уроков** с контрольными заданиями, которые будут размещены на **обучающей платформе** и в перспективе станут **обязательными для всех государственных служащих**.

В дополнение к этому до конца 2022 года совместно с **Азиатским Банком развития** в рамках **Программы обмена знаниями и опытом** начнется обучение работников БНС более продвинутым инструментам бизнес-аналитики, включая BI-инструменты и погружение в Data science на основе Python.

На 2023 год планируется обучение ключевых сотрудников БНС стандартам DAMA и прохождение **сертификации Certified Data Management Professional CDMP**.

Таким образом, на базе БНС будет сформирован **центр компетенций**, который на возобновляемой основе продолжит дальнейшее обучение сотрудников государственных органов собственными силами.

Отмеченные программы обучения будут охватывать также **дата стюардов** для повышения их квалификации и формирования единого видения развития экосистемы данных. Обученные дата стюарды в последующем будут применять полученные навыки при работе с данными из национальных регистров.

Аналитический потенциал сотрудников БНС также будет развиваться за счет **исследовательской активности** совместно с ВУЗами. Совместные исследования будут проводиться по темам, интересующим как БНС, так и ВУЗы. БНС представляет на сайте **перечень тем научно-исследовательских работ**, направленных на решение текущих исследовательских потребностей. Также БНС будет активно вовлекаться в исследовательские работы с профессорско-преподавательским составом ВУЗов по интересующим их темам.

Наряду с предоставлением баз данных в деидентифицированном виде в рамках отмеченных направлений будет оказываться методологическая помощь с возможным соавторством публикаций.

Кроме того, будет **систематизирован механизм взаимодействия** БНС с ВУЗами в части предоставления обезличенных микроданных и запущен **информационный сервис для исследователей**, который будет отражать список получателей данных и ссылки на результаты исследовательских работ. Это позволит повысить прозрачность, координацию и результативность исследовательской деятельности на основе данных БНС.

ЗАДАЧА 4.

ОБЕСПЕЧЕНИЕ ДОСТУПА К ДАННЫМ И СОЗДАНИЕ ПРОЗРАЧНОГО КОНКУРЕНТОСПОСОБНОГО РЫНКА ДАННЫХ

Важным направлением развития экосистемы данных является обеспечение **равного доступа к данным**. Массовое использование данных позволит повысить их ценность, а также полноценно внедрить практику принятия решений на основе данных.

В этих целях БНС формируются сервисы для упрощенного доступа к данным. В частности, разрабатываются **интерактивные дэшборды** по всем основным статистическим показателям. В перспективе планируется создание **кабинета исследователя** и **виртуальной лаборатории данных** для всестороннего использования микроданных аккредитованными пользователями.

В рамках данного направления будет **законодательно расширен доступ** пользователей к микроданным. Доступ к обезличенным микроданным позволяет анализировать комплексные проблемы с учетом разнообразия населения и бизнеса, разработать инструменты **микромоделирования для оценки потенциального воздействия** изменения государственной политики и макроэкономических шоков.

Одним из ключевых партнеров по распространению данных и их массовому применению являются научные организации и высшие учебные заведения, которые имеют законодательный доступ к обезличенным микроданным. Для продвижения культуры работы с данными среди студентов создаются совместные проекты, конкурсы и исследования с использованием микроданных.

Наряду со статистическими данными важно обеспечить доступ пользователей к административным данным, имеющимся в центральных государственных и местных исполнительных органах. Зачастую административные данные не доступны пользователям и применяются только единожды, что снижает потенциальный эффект от собранных данных.

Существенный потенциал имеет **раскрытие данных из альтернативных источников**, формируемых частным сектором. Их активное использование позволит получить более **полную картину** социально-экономической ситуации.

Для обеспечения доступа к отмеченным комбинированным источникам данных важна **роль регулятора рынка данных**, который будет обеспечивать координацию и **четкие правила** игры с целью одновременного сохранения **инклюзивности и конфиденциальности данных**. В этой связи необходимо создание регулятивного поля для **сбалансированного развития экосистемы данных**.

Важным драйвером развития экосистемы данных станет возможность **монетизации данных**. В этой связи в качестве инструмента коммерциализации сервисов единой платформы данных, будут созданы условия для запуска **биржи данных**, которая позволит различным участникам рынка совершать сделки с данными и дата-продуктами. На площадке будут доступны технологии для обмена данными и их анализа, при этом площадка будет предлагать готовые сценарии монетизации данных. Биржа должна стать ядром экосистемы данных, что впоследствии позволит развить национальный рынок данных.

В долгосрочной перспективе реализация перечисленных задач позволит создать **полноценный рынок данных**. Участники рынка данных смогут **взаимно обогащаться** данными и монетизировать их, тем самым повышая прибыльность, а государство – принимать более эффективные решения на основе **комбинированных источников данных**.






08

ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ

Ожидаемые результаты

Результатом трансформации национальной статистической системы к концу **2025 года** станет создание полноценной **национальной экосистемы данных**, выгоды от которой получат **все пользователи рынка данных**, включая государственный сектор, граждан и бизнес.

ОЖИДАЕМЫЕ ЭФФЕКТЫ ДЛЯ СТЕЙКХОЛДЕРОВ

	 Государство	 Граждане	 Бизнес
Сбор отчетности и оказание услуг	Оптимизация статистических наблюдений в 2 раза (2022г – 146 форм)	Проактивное предоставление государственных услуг	Снижение нагрузки за счет исключения дублирования отчетности
Раскрытие данных	Повышение ценности данных за счет раскрытия и возможности повторного использования данных	Соблюдение требований безопасности персональных данных и конфиденциальности информации	Раскрытие ниши для ведения эффективного бизнеса за счет устранения информационных асимметрий
Использование данных	Увеличение точности планирования, скорости и качества принимаемых решений	Множество новых сервисов и продуктов на основе данных	Внедрение новых дата продуктов и законной возможности монетизации данных

09

ЭТАПЫ РЕАЛИЗАЦИИ



Этапы реализации

Предложенные в рамках Концепции задачи реализуются в **2 этапа в течение 3 лет**.
Ниже представлены мероприятия, реализуемые в данный период (2023-2025 годы).

№	Мероприятия	2023 г.	2024-2025 гг.
1	Задача 1. Совершенствование нормативной правовой и методологической базы		
1.1	Разработка Закона «О государственной статистике и данных» и соответствующих НПА	■	
1.2	Совершенствование методологических требований для проведения аудита по управлению данными	■	
1.3	Разработка Методологии обеспечения и контроля качества данных	■	
1.4	Проведение аудита информационных систем государственных органов	■	■
1.5	Глобальная оценка казахстанской статистической системы ЕЭК ООН	■	
1.6	Обзор ОЭСР с дальнейшим получением статуса «члена» Комитета ОЭСР	■	■
1.7	Исполнение рекомендации ОЭСР и ЕЭК ООН	■	■
2	Задача 2. Активное использование административных и альтернативных источников на основе национальных регистров		
2.1	Разработка и утверждение целевой архитектуры данных для страны	■	
2.2	Разработка дата-каталога и стандартизация метаданных	■	
2.3	Создание национальных регистров на основе государственных баз данных и повышение их качества	■	
2.4	«Перезагрузка» и обучение дата стюардов	■	
2.5	Использование альтернативных источников информации	■	■
2.6	Цифровизация внутренних процессов статистической деятельности	■	
2.7	Обновление ИИС «е-Статистика»	■	■
2.8	Расширение и перераспределение производственных мощностей, предназначенных для хранения и обработки данных	■	■
3	Задача 3. Укрепление аналитического потенциала государственных органов		
3.1	Обучение и развитие компетенций государственных служащих	■	
3.2	Формирование научно-исследовательских компетенций и технологических практик	■	
3.3	Разработка специализированных программ повышения квалификаций	■	■
3.4	Проведение совместных исследований с ВУЗами	■	■
3.5	Создание нового механизма предоставления обезличенных данных и запуск информационного сервиса для исследователей	■	■
4	Задача 4. Обеспечение доступа к данным и создание прозрачного конкурентоспособного рынка данных		
4.1	Обеспечение доступа к микроданным в целях проведения исследований и реализации аналитических решений	■	■
4.2	Создание «Биржи» данных		■
4.3	Подготовка к переходу с традиционной на регистровую перепись		■

Сокращения



DAMA	Ассоциация управления данными (Data Management Association)
БНС	Бюро национальной статистики Агентства стратегического планирования и реформ Республики Казахстан
МЦРИАП	Министерство цифрового развития, инноваций и аэрокосмической промышленности Республики Казахстан
ЦГО	Центральные государственные органы
МИО	Местные исполнительные органы
НСУ	Национальное статистическое управление
IT	Информационные технологии
ИПЦ	Индекс потребительских цен
ВВП	Валовый внутренний продукт
ОЭСР	Организация экономического сотрудничества и развития
GCDS	Government chief data steward (Главный распорядитель данных правительства)
Stats NZ	Статистическое управление Новой Зеландии
DGLG	Digital government leader group (Группа лидеров цифрового правительства)
INEGI	Instituto Nacional de Estadística y Geografía (Национальный институт статистики и географии Мексики)
DMBOK	Data management body of knowledge (свод знаний по управлению данными)
ЕЭК ООН	Евразийская экономическая комиссия Организации объединенных наций
НПА	Нормативный правовой акт