

# Дорога в мир, управляемый данными

Мир до COVID-19 уже в значительной мере управлялся данными, но с той поры темпы цифровой трансформации стремительно ускорились. Жесткая конкуренция, цифровизация, постоянно растущие ожидания клиентов и усиление контроля со стороны регулирующих органов требуют от организаций трансформироваться в современные предприятия, управляемые данными. Эта трансформация неизбежно приведет к тому, что организации станут более цифровыми и будут иметь иной взгляд на данные. Организации будут дышать данными и придерживаться философии, ставящей ИТ в центр бизнеса. Они начнут управлять данными как продуктом, принимать стратегические решения на основе анализа данных и формировать культуру, действующую на основе данных.

«Управляемые данными» (*data-driven*) — это не просто модное выражение<sup>1</sup>. Организации, управляемые данными, получают существенное конкурентное преимущество перед другими организациями. Они могут действовать проактивно и заранее предсказывать, что произойдет. Правильно используя данные, организации могут быстро реагировать на изменения. Использование данных увеличивает уверенность, потому что решения основываются на фактах, а не на интуиции. Данные позволяют быстрее выявлять новые тенденции и бизнес-возможности. Данные повышают удовлетворенность клиентов и способствуют их удержанию, потому что сообщают организациям, что думают клиенты, как они себя ведут и чего хотят. Данные помогают организациям быть более гибкими, динамичными и экономически эффективными, потому что несут информацию об измеримых результатах, лояльности сотрудников, зависимостях, приложениях и процессах. Таким образом, тенденция трансформации организаций в предприятия, управляемые данными, определенно существует.

Прежде чем перейти к самой трансформации, рассмотрим современные проблемы, требующие переоценки подходов к управлению данными. Мы выработаем общее

---

<sup>1</sup> Рынок аналитики данных стремительно растет. По сведениям IDC (<https://oreil.ly/8OZzk>), в 2022 году траты компаний на этом рынке превысили 200 млрд долларов.

определение управления данными, охватывающее все дисциплины и действия, связанные с управлением данными как активом. После этого рассмотрим несколько ключевых технологических разработок и тенденций отрасли и исследуем их влияние на управление данными. Познакомимся с передовыми практиками управления данными, появившимися за последние десять лет, которые помогут понять, почему архитектуры предыдущего поколения трудно масштабировать. Наконец, мы выясним, как может выглядеть архитектура данных следующего поколения, и затем я представлю набор практических рекомендаций, которые желательно учитывать при разработке стратегии данных.

## Последние технологические разработки и тенденции в отрасли

Трансформация организации в предприятие, управляемое данными, — непростая задача. Это долгосрочный процесс, требующий стойкости и терпения. С увеличением количества доступных данных традиционные архитектуры не могут масштабироваться из-за сложности монолитных конструкций и централизованных операционных моделей. Предприятиям нужна новая стратегия управления данными и облачная архитектура. Необходимы также смена парадигмы и изменение культуры, потому что централизованные модели данных и ИТ-операций, которые используются сегодня, перестанут работать при внедрении федеративного владения данными и моделей потребления с самообслуживанием. Это требует от организаций пересмотреть использование данных людьми, процессами и технологиями.

Последние технологические разработки и тенденции в отрасли заставляют нас пересмотреть подходы к управлению данными. Нам нужно уйти от сосредоточения всех данных в одном месте и дать возможность отделам, командам и пользователям самостоятельно распространять, потреблять и использовать данные легко и безопасно. Платформы, процессы и шаблоны должны упрощать работу. Нам нужны простые, хорошо документированные, быстрые и удобные интерфейсы. Нужна архитектура, способная масштабироваться.

В развитии организаций, управляемых данными, есть много положительных моментов, однако важно знать, какие технологические разработки и тенденции существуют в отрасли и как они влияют на ландшафты данных. В этой главе мы поговорим о них и их влиянии на управление данными. Во-первых, разнообразие аналитических подходов приводит к фрагментированию ландшафта данных. Во-вторых, новые методологии разработки программного обеспечения затрудняют управление данными. В-третьих, внедрение облачных вычислений и более быстрых сетей тоже приводит к фрагментированию ландшафта данных. Кроме того, необходимо учитывать проблемы конфиденциальности, безопасности и регулирования, а быстрый рост данных и интенсивное их потребление приводят к перегрузке операционных систем. Наконец, монетизация данных требует внедрения архитектур

«экосистема — экосистема». Влияние этих тенденций велико: вся отрасль должна переосмыслить подход к управлению данными в будущем.

К счастью, за последние несколько лет появились новые подходы к управлению данными, включая *сетки данных* и *фабрики данных*.

- Сетка данных (data mesh, <https://oreil.ly/kBtUG>) — это новая методология управления данными в целом, предполагающая использование архитектуры, в которой данные сильно распределены, а масштабируемость достигается путем объединения обязанностей на федеративных началах. Основной упор в ней делается на человеческий фактор и решение проблем управления усложняющимися архитектурами данных.
- Фабрика, или матрица, данных (data fabric, <https://oreil.ly/bjl2F>) — это подход к решению современных проблем управления данными и масштабируемости за счет добавления искусственного интеллекта и упрощения доступа к данным путем внедрения методов самообслуживания. В отличие от сетки данных, этот подход больше фокусируется на технологическом уровне. Это архитектурное видение, основанное на использовании унифицированных метаданных и интегрированного сквозного уровня (матрицы) для упрощения доступа, интеграции, предоставления и использования данных.

Эти новые подходы к управлению данными являются взаимодополняющими и часто пересекаются. Несмотря на мнение, бытующее среди специалистов по работе с данными, и высказывания коммерческих поставщиков, эти подходы не следует рассматривать как жесткие или самодостаточные методы. На самом деле они будут сосуществовать и дополнять друг друга и остальные технологии хранилищ и озер данных.

При трансформации организаций в предприятия, управляемые данными, придется идти на компромиссы, чтобы сбалансировать императивы централизации и децентрализации. Одни предпочитают сделать свои бизнес-подразделения более автономными, другие, напротив, отдают приоритет качеству и контролю. Одни организации имеют относительно простую структуру, тогда как другие — чрезвычайно сложную и разветвленную. Создание идеальной структуры управления и архитектуры данных — непростая задача, поэтому при разработке стратегии я рекомендую вам рассматривать эти подходы к управлению данными лишь как фундамент, на котором вы будете строить свое здание. Нет правильного или неправильного. Например, при выборе подхода на основе сетки данных вам могут понравиться лишь некоторые из практик и принципов, и вам совершенно необязательно заставлять себя использовать то, что вам не нравится.

В этой книге я поделюсь своим взглядом на управление данными, основанным на наблюдениях с мест, когда я работал в тесном контакте со многими крупными предприятиями. Надеюсь, мой опыт поможет вам принять правильные решения. Мы выйдем за рамки концепций сетки данных и их организации, поскольку я твердо верю, что стратегия управления данными должна включать не только

операционную, но и аналитическую плоскость, и что метод декомпозиции как для предметных областей, так и для продуктов данных следует изменить и привести в соответствие с масштабом крупных предприятий. Чтобы помочь вам в этом путешествии, я поделюсь своими наблюдениями и размышлениями о стратегиях и архитектурах, разработанных различными предприятиями, и расскажу, на какие компромиссы и почему они пошли.

Прежде чем углубляться в детали, нам нужно определить, что такое управление данными и почему оно важно. После этого мы поговорим о том, как определить границы и сформировать общий ландшафт на основе различных компромиссов. Наконец, мы увидим, как проектируются и организуются современные корпоративные архитектуры и платформы данных.

Позвольте мне раскрыть маленький секрет: децентрализация — это не желаемое состояние, а неизбежное будущее данных. Поэтому я твердо убежден, что масштабируемость неизбежно повлечет за собой децентрализацию управления. Управление данными в масштабе требует объединения ключевых обязанностей, установления строгих стандартов и надлежащего согласования центральных и локальных ресурсов и видов деятельности. Это изменение затрагивает людей, процессы и технологии. Оно заставляет реорганизовать архитектуру, разделяя и группируя обязанности. Переход от централизации к децентрализации также противоречит устоявшейся передовой практике последнего десятилетия: созданию больших хранилищ данных, в которых собираются и интегрируются все данные. Архитектуры хранилищ и озер данных отлично подходят для нужд использования данных, но из-за их централизации эти модели не годятся для децентрализованной распределенной архитектуры данных.

Теперь, когда мы подготовили сцену, я прошу вас сделать глубокий вдох и отбросить предубеждения. Многие из нас уверены в правильности выбора централизованной архитектуры данных, поскольку она считалась лучшей в течение многих лет. Я признаю, что потребность в гармонизации и приведении больших объемов данных к определенному контексту остается и это приносит пользу организациям, но мы должны учитывать *масштаб*. Действительно ли лучшим способом управления будет применение централизации во всех измерениях в сильно распределенной экосистеме с сотнями или даже тысячами приложений? Действительно ли имеет смысл интегрировать и гармонизировать все данные?

## Управление данными

Под *управлением данными* (*data management*) подразумеваются процедуры и процессы, применяемые для организации данных. Руководство DAMA International по комплексу знаний об управлении данными (DAMA-DMBOK) (<https://oreil.ly/ВКIm6>) дает более развернутое определение: «Управление данными — это разработка, выполнение и контроль планов, политик, программ и практик, которые

доставляют, контролируют, защищают и повышают ценность данных и информационных активов на протяжении их жизненного цикла»<sup>1</sup>.

DAMA-DMBOK выделяет 11 функциональных областей управления данными, в центре которых находится руководство данными (рис. 1.1). Очень важно глубоко внедрить эти дисциплины в вашу организацию. В противном случае вы запутаетесь и станете работать неэффективно, а ваши данные выйдут из-под контроля. В результате получение максимальной отдачи от ваших данных станет невозможным. Аналитика, например, ничего не стоит, если будет основываться на некачественных данных.



**Рис. 1.1.** Одиннадцать функциональных областей управления данными

<sup>1</sup> Комплекс знаний разработан сообществом DAMA. Он редко обновляется: первый релиз был в 2008 году, второй — в 2017 году.

Виды деятельности и дисциплины управления данными разнообразны и охватывают несколько областей. Некоторые из них тесно связаны с архитектурой программного обеспечения<sup>1</sup>. В этой книге я выбрал только те аспекты управления данными, которые наиболее актуальны для современной архитектуры данных в масштабе. Ниже приводятся более подробные характеристики 11 областей, представленных на рис. 1.1, и подсказывается, в какой главе они рассматриваются.

- *Руководство данными*, представленное в центре на рис. 1.1, включает действия, связанные с реализацией и обеспечением полномочий, а также контроль над управлением данными, в том числе все соответствующие активы. Эта область подробно описана в главе 8.
- *Архитектура данных*. Это мастер-план данных. Позволяет шире взглянуть на вашу архитектуру, включая схемы<sup>2</sup>, эталонные архитектуры, видение перспективы и зависимости. Управление этими аспектами помогает организациям принимать решения. Тематика всей книги вращается вокруг архитектуры данных в целом, но сама область и ее задачи полностью раскрываются в главах 2 и 3.
- *Моделирование и проектирование данных*. Это структурирование и представление данных в определенном контексте и конкретных системах. Сюда относятся обнаружение, проектирование и анализ требований к данным. Эти темы мы обсудим в главах 4, 7 и 11.
- *Хранение данных и операции с ними* относятся к управлению проектированием баз данных, правильной реализации и поддержке с целью увеличения ценности данных. Эта область также включает управление операциями с базами данных. Некоторые темы рассматриваются в главе 11.
- *Безопасность данных* включает все дисциплины и действия, обеспечивающие безопасную аутентификацию, авторизацию и доступ к данным. К ним относятся предотвращение, аудит и меры по смягчению последствий распространения данных. Эта область более подробно описана в главе 8.
- *Интеграция и согласованность по данным* включают в себя все дисциплины и действия по транспортировке, сбору, объединению и трансформации данных для эффективного их перемещения из одного контекста в другой. Согласованность по данным — это возможность взаимодействия путем вызова функций

<sup>1</sup> Под *архитектурой* понимаются организация и высокоуровневая структура программного обеспечения, необходимого для разработки системы, соответствующей бизнес-требованиям, и обладающего такими характеристиками, как гибкость, масштабируемость, осуществимость, возможность повторного использования и безопасность. Если вы хотите узнать больше об архитектуре программного обеспечения, я рекомендую прочитать книгу *Fundamentals of Software Architecture* Марка Ричардса (Mark Richards) и Нила Форда (Neal Ford) (O'Reilly). (Книга вышла в издательстве «Питер» в 2022 году под названием «Фундаментальный подход к программной архитектуре: паттерны, компоненты, проверенные методы». — *Примеч. ред.*).

<sup>2</sup> Схемой также называют *эталонную реализацию*. Она состоит из определений инфраструктуры как кода (Infrastructure as Code, IaC), необходимых для успешного создания набора сервисов в данном конкретном случае.

или передачи данных между различными приложениями способами, не требующими знания характеристик приложения. Интеграция данных — это объединение данных из разных (нескольких) источников в одно представление. Этот процесс, который я считаю наиболее важным, часто поддерживается такими дополнительными инструментами, как инструменты репликации и ETL (extract transform and load — «извлечение, преобразование и загрузка»). Процесс подробно описан в главах 4, 5 и 6.

- *Управление документами и их содержимым* — это процесс управления данными, хранящимися в неструктурированных документах. Некоторые стороны этой деятельности будут рассматриваться в главах 5 и 6.
- *Управление справочными и основными данными* проводится для обеспечения доступности, точности, безопасности, прозрачности и надежности данных. Эта область более подробно описана в главе 10<sup>1</sup>.
- *Хранение данных и бизнес-аналитика* включают все виды деятельности, помогающие понять происходящее в бизнесе и принять обоснованное решение. Эта область описана в главе 11.
- *Управление метаданными*. Это управление всей информацией, которая классифицирует и описывает данные. Метаданные помогают сделать данные понятными, безопасными и готовыми к интеграции. Они также могут использоваться для обеспечения качества. Эта тема более подробно раскрыта в главе 9.
- *Управление качеством данных* необходимо, чтобы гарантировать пригодность данных к использованию. Некоторые дисциплины из этой области описаны в главах 2 и 3.

Есть еще одна область DAMA-DMBOK — *интеграция и взаимодействие данных* (она-то и вдохновила меня написать первое издание этой книги). Я считаю, что этой области не хватает глубины: связь с интеграцией приложений и архитектурой программного обеспечения очерчивается недостаточно ясно. В ней не обсуждаются децентрализованные архитектуры и отсутствуют современные руководства по организации взаимодействий данных, такие как передовые практики наблюдения и современное управление конвейерами данных. Кроме того, весьма слаба связь с управлением метаданными. Метаданные тоже нуждаются в интеграции и взаимодействии, будучи разбросанными среди множества инструментов, приложений, платформ и сред. Взаимодействие метаданных — способность двух и более систем или компонентов обмениваться описательной информацией о данных — недооценена, поскольку создание крупномасштабной архитектуры и управление ею во многом связано с интеграцией метаданных. Область архитектуры данных в контексте метаданных тоже мало изучена. При правильном использовании метаданных

<sup>1</sup> В книге я использую термин «управление основными данными» (Master Data Management, MDM), потому что он широко принят и на момент написания этих строк не имел общепромышленной альтернативы, но вы можете свободно использовать более инклюзивные альтернативные термины, обозначающие понятия «ведущий/ведомый».

вы четко видите, как передаются данные, как их можно интегрировать, распределять и защищать, как они связаны с приложениями, бизнес-возможностями и т. д. В DAMA-DMBOK есть ограниченные рекомендации по управлению данными в целом путем использования и подключения метаданных.

Я не согласен со взглядами DAMA и многих организаций на достижение сквозной семантической согласованности. Сегодня продолжаются попытки унифицировать всю семантику для обеспечения согласованности в масштабах всего предприятия. Это называется «единая версия истины». Однако *приложения, как и данные, всегда уникальны*. Разработка приложений часто подразумевает неявное мышление. Контекст (предметная область) бизнес-задачи влияет на архитектуру приложения и находит свое отражение в данных. Мы встречаемся с ним, когда переходим от концептуальной модели проекта приложения к его логической и физической формам<sup>1</sup>. Это важно, ведь контекст ставит в рамки будущую архитектуру. При перемещении данных между приложениями их всегда нужно преобразовывать. Выхода из этой дилеммы трансформации данных не существует! Я еще вернусь к этой мысли в следующих главах.

Есть и другая точка зрения — будто управление данными должно быть централизовано и связано со стратегическими целями предприятия. Многие организации считают, что за счет этого можно снизить эксплуатационные расходы. Также бытует мнение, будто централизованная платформа может избавить клиентов от проблем, связанных с интеграцией данных. Компании вкладывают значительные средства в корпоративные платформы данных: хранилища, озера и сервисные шины. Управление основными данными тесно связано с этими платформами, ведь объединение позволяет одновременно повышать точность наиболее важных данных.

Централизованная платформа и прилагаемая к ней централизованная модель могут потерпеть фиаско из-за невозможности идти в ногу с разработками и тенденциями, лежащими в основе децентрализации, такими как аналитика, облачные вычисления, новые методологии разработки ПО, принятие решений в реальном времени и монетизация данных. Все это ни для кого не секрет, но многие компании не осознают, какое влияние они оказывают на управление данными. Давайте рассмотрим наиболее важные тенденции и определим их масштабы.

## Аналитика фрагментирует ландшафт данных

Самая заметная тенденция — это *продвинутая аналитика*. Она использует данные, чтобы сделать компании более отзывчивыми, конкурентоспособными и инновационными. Так почему же продвинутая аналитика нарушает существующий ландшафт данных? Чем больше данных, тем больше вариантов и возможностей.

---

<sup>1</sup> Больше подробностей о различных этапах моделирования данных вы найдете в моей статье *Data Integration and Data Modeling Demystified* (<https://oreil.ly/qEuCE>). Концептуальную модель иногда также называют *бизнес-моделью*, *моделью предметной области*, *объектной моделью предметной области* или *объектной моделью анализа*.

Продвинутая аналитика — это анализ возможных вариантов, прогнозирование будущих тенденций, результатов или событий, автоматизация принятия решений, обнаружение скрытых отношений и поведения. Благодаря признанной ценности и стратегическим преимуществам продвинутой аналитики появилось множество методологий, фреймворков и инструментов для ее разнообразного использования. Мы лишь коснулись способностей искусственного интеллекта, машинного обучения и обработки естественного языка.



Новая языковая модель ChatGPT от OpenAI (<https://openai.com>) — ошеломляющий пример способностей ИИ. Модели, лежащие в основе OpenAI, в том числе серия генеративных предварительно обученных трансформеров (Generative Pre-trained Transformer, GPT), способны справляться с весьма сложными заданиями, такими как анализ математических задач, написание фрагментов кода, создание эссе для книг, составление рецептов с использованием списка ингредиентов и многие другие.

Тенденции в развитии аналитики заставляют распределять данные по множеству аналитических приложений, поскольку в каждом конкретном случае необходимы разные сведения. Уникальные бизнес-задачи требуют уникального мышления, уникальных данных и оптимизированных технологий для выбора наилучшего решения. Возьмем, к примеру, маркетинговое подразделение компании, которое стремится увеличить продажи среди молодых и пожилых клиентов. Ориентация на эти две аудитории требует наличия разных характеристик — измеримых свойств — в наборах данных. Например, более молодые потенциальные клиенты будут группироваться иначе, чем пожилые. Требование к отделу маркетинга использовать один набор данных для анализа обеих целевых аудиторий вынудит согласиться с множеством компромиссов, и, скорее всего, получится хранилище характеристик, не добавляющее никакой ценности ни тому, ни другому варианту использования<sup>1</sup>. Оптимальное решение для создания наибольшей ценности — придать каждому варианту свой уникальный набор характеристик, оптимизированных для каждого алгоритма обучения. Растущая популярность продвинутой аналитики и возникающие в результате сложности, обусловленные разнообразием вариантов применения, влекут за собой две проблемы: быстрое увеличение объема данных и интенсивность их использования.

По мере разрастания одни и те же данные распределяются по множеству приложений и БД. Это происходит потому, что области, потребляющие данные, должны обрабатывать их, чтобы вписать их в свои уникальные решения. Такое распределение данных создает целый ряд проблем. Во-первых, когда данные многократно копируются и распределяются по всей организации, становится все сложнее найти их источник и оценить их качество. Это требует разработки единого логического представления одних и тех же данных, управление которыми осуществляется в разных местах.

<sup>1</sup> *Хранилище характеристик (признаков) (feature store)* — это инструмент для хранения часто используемых признаков (отдельных измеримых свойств или характеристик явления).

Кроме того, широкое распространение данных значительно затрудняет контроль, поскольку данные могут распространяться еще дальше, как только покинут какое-либо приложение. Это требует разработки структуры для эффективного повторного использования данных и управления с соблюдением внешних правил.

Развитие аналитических методов означает ускоренный рост объема данных: *соотношение чтения и записи* значительно меняется. Например, аналитические модели, которые постоянно переобучаются, считывают большие объемы данных. Это влияет на приложения и структуру БД, требуя оптимизировать доступность данных для чтения, и может означать следующее: чтобы избавить системы от необходимости постоянно обслуживать данные, их нужно дублировать. Это также может означать необходимость дублирования данных для их предварительной обработки из-за разнообразия и большого количества вариантов использования и шаблонов чтения, которые сопутствуют этим вариантам. Кроме того, может потребоваться создать различные представления одних и тех же данных для разных потребителей. Сложно привести в порядок такое разнообразие шаблонов чтения при одновременном дублировании данных и сохранении контроля. Решение этой проблемы описано в главе 4.

## Скорость доставки программного обеспечения меняется

В современном мире программные сервисы составляют основу бизнеса. Это означает, что нужно быстро выпускать обновления. В ответ на требования большей гибкости в таких компаниях, как Amazon, Netflix, Meta, Google и Uber, появились новые идеологии: они продвинули свою практику разработки программного обеспечения на основе двух убеждений.

Первое убеждение: разработку (development, Dev) и эксплуатацию программного обеспечения (operations, Ops) нужно объединять. Это поможет сократить жизненный цикл разработки и обеспечить непрерывную поставку качественного ПО. Такая методология называется *DevOps*. Она требует новой культуры, где есть место для большей автономии, открытого общения, доверия, прозрачности и междисциплинарной командной работы.

Второе убеждение касается размера разрабатываемого приложения. Ожидается, что гибкость и скорость разработки системы возрастут, если приложения превратятся в более мелкие сервисы, выполняющие различные задачи. Этот подход к разработке характеризуется несколькими модными словами: *микросервисы* (<https://microservices.io/>), *контейнеры*, *Kubernetes* (<https://kubernetes.io/>), *предметно-ориентированное проектирование*, *бессерверные вычисления* и т. д. Я не буду вдаваться в подробности каждой концепции, но эта эволюция разработки программного обеспечения связана с повышенными сложностью и спросом на модернизированный контроль данных.

Преобразование монолитного приложения в распределенное, например в набор микросервисов, создает множество проблем, связанных с управлением данными. При дроблении приложений данные распределяются по разным компонентам. Командам разработчиков приходится переводить свои (отдельные) уникальные хранилища данных в модель, где объекты данных распределены. Это создает несколько сложностей: интенсификацию сетевых взаимодействий, реплики чтения данных, которые необходимо синхронизировать, проблемы согласованности, ссылочной целостности и т. д.

Сдвиг в разработке ПО требует создания архитектуры, позволяющей более детализированным приложениям распределять свои данные. К тому же нужна новая культура DataOps (<https://oreil.ly/G8Bt4>) и другая философия проектирования с большим упором на совместимость данных, фиксацию неизменяемых событий, а также воспроизводимые и слабые связи. Мы подробнее обсудим это в главе 2.

## **Влияние облачных технологий на управление данными не поддается измерению**

Сети становятся быстрее, а пропускная способность из года в год увеличивается. Крупные поставщики облачных услуг доказали возможность перемещения терабайтов данных в облаке за считанные минуты, что позволяет реализовать интересный подход: вместо увеличения вычислительной мощности для обработки данных в одном месте, что было распространенной практикой из-за сетевых ограничений, мы можем распределить данные и обрабатывать в разных местах. Сеть больше не является препятствием, поэтому мы можем быстро перемещать терабайты данных между средами, давая приложениям возможность потреблять и использовать их. Такая модель становится особенно интересной с ростом популярности рынков программного обеспечения как услуги (Software as a Service, SaaS) и машинного обучения как услуги (Machine Learning as a Service, MLaaS). Вместо того чтобы выполнять сложные операции самостоятельно, мы можем использовать сети для передачи данных другим сторонам.

Этот шаблон распределения копирования (дублирования) и переноса данных на вычислительные мощности в другом месте, например в облачный центр обработки данных, еще больше фрагментирует ландшафт данных, делая четкую стратегию управления данными еще более важной, чем когда-либо. Она требует выработки руководящих принципов, поскольку фрагментация данных может негативно повлиять на производительность из-за задержки доступа к ним. Стратегия также требует иной организации и моделирования данных, поскольку поставщики облачных услуг проектируют отдельные вычисления и хранение, чтобы обеспечить возможность их масштабирования независимо друг от друга.