



SEAGATE

# RETHINK DATA

Эффективное использование бизнес-данных  
на периферии и в облаке

ОТЧЕТ SEAGATE TECHNOLOGY

---

# Содержание

Об отчете .....	03
Кратко о результатах .....	04
Основные результаты опроса .....	05

## Раздел I. Общие результаты

Глава 1. Рост объемов и разрастание данных .....	07
Глава 2. Кладезь ценности .....	13
Глава 3. Многооблачные среды .....	19
Глава 4. Проблемы управления данными и многооблачная экосистема .....	24
Глава 5. DataOps: недостающее звено в управлении данными .....	29
Глава 6. Улучшение бизнес-результатов .....	33
Глава 7. Безопасность данных и управление ими .....	36

## Раздел II. Результаты по регионам

Глава 1. Азиатско-Тихоокеанский регион и Япония .....	40
Глава 2. Китай .....	44
Глава 3. Европа .....	48
Глава 4. Северная Америка .....	52



---

# Об отчете

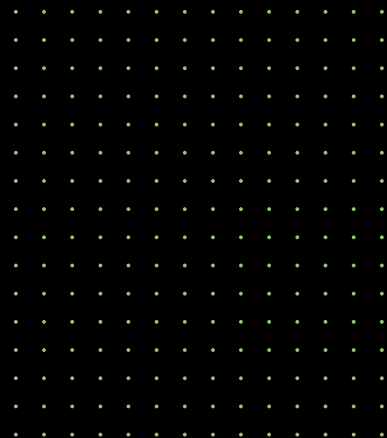
Отчет основан на результатах глобального исследования, проведенного независимой исследовательской компанией IDC в декабре 2019 — январе 2020 года и финансируемого компанией Seagate. Цель отчета — поделиться нашими прогнозами и мнениями, сформированными по результатам опроса и анализа компании IDC. Все данные, материалы, разборы и мнения, не предоставленные компанией IDC или ее аналитиками, принадлежат компании Seagate. В количественном опросе, который проводился онлайн, приняли участие 1500 респондентов со всего мира

(375 из Северной Америки, 475 из Европы, 150 из Китая и 500 из Азиатско-Тихоокеанского региона, включая Японию). Это представители средних и крупных предприятий в Канаде, США, Великобритании, Франции, Германии, России, Австралии, Японии, Индии, Южной Корее, Тайване и Китае. В число респондентов вошли главные технические директора, главные директора и вице-президенты по информационным технологиям, директора, высшие руководители, главные операционные директора, руководители направлений, архитекторы хранилищ и решений.

**Seagate Technology** уже более 40 лет является мировым лидером на рынке решений для хранения данных и управления ими. Seagate выпускает передовые высокотехнологичные решения мирового класса, которые помогают людям достигать большего. Один из наших главных приоритетов — устойчивые партнерские связи с другими компаниями.



SEAGATE



---

# Кратко о результатах

Беспрецедентный рост объемов и разрастание данных приводят к усложнению и повышению разнообразия экосистемы данных, которая сегодня включает многооблачные и периферийные среды. Владельцы компаний сталкиваются с тем, что данными сложнее управлять, поскольку они больше не хранятся централизованно. Большая часть доступных бизнес-данных так и не используется.

Исследование направлено на решение сегодняшних проблем в управлении данными. DataOps (операции с данными) — это методология, сводящая вместе создателей данных и их потребителей. Ведущие компании, внедряющие DataOps, могут рассчитывать на улучшение коммерческих показателей.

# Основные результаты опроса

- В ближайшие два года прогнозируется рост данных предприятий на 42,2% в год.
- Лишь 32% доступных бизнес-данных используются организациями, остальные 68% остаются без внимания.
- В среднем организации периодически перемещают около 36% своих данных с периферии в центр. Ожидается, что всего через два года этот показатель вырастет до 57%. Объем данных, передаваемых с периферии в центр в режиме реального времени, удвоится с 8% до 16%, то есть предприятиям придется управлять гораздо большим количеством постоянно перемещаемых данных.
- Ожидается, что в течение ближайших двух лет главной сложностью в управлении данными будет управление ими в многооблачной экосистеме. Второе место в этом рейтинге с небольшим отрывом занимает управление данными в гибридном облаке.
- Пять основных препятствий для эффективной работы с данными: 1) неспособность использовать полученные данные; 2) неспособность управлять хранимыми данными; 3) сбор не всех потенциально доступных данных; 4) недостаточная безопасность собранных данных; 5) хранение данных в разрозненных базах.
- DataOps решает множество задач управления данными. Это методология, сводящая вместе создателей данных и их потребителей. В среднем лишь 10% организаций, по словам их представителей, уже внедрили DataOps на всем предприятии. Большинство респондентов считают методологию DataOps «очень» или «чрезвычайно» важной.
- Наряду с другими решениями для управления данными DataOps обеспечивает измеримый рост бизнес-результатов: повышение лояльности клиентов, увеличение доходов и прибыли, а также множество других преимуществ.
- Важнейшим фактором, влияющим на изменение подхода к управлению центральным хранилищем в организации, является повышение безопасности данных.
- Две трети респондентов отмечают недостаточную безопасность данных, что делает этот вопрос ключевым, когда речь заходит об эффективном управлении данными.



# РАЗДЕЛ I

# Общие результаты



## ГЛАВА ПЕРВАЯ

# Рост объемов и разрастание данных

## Данные расскажут, кто мы, кем были и кем будем.

### ТОЧКА ЗРЕНИЯ SEAGATE

Численность населения планеты сегодня составляет 7,8 млрд человек и продолжает расти. Все больше людей работает из дома. Усиливается роль таких технологий, как Интернет вещей (IoT), периферийные вычисления, периферийные центры обработки данных и искусственный интеллект (ИИ). Растет и спрос на пользовательские конечные устройства. Все это приводит к быстрому увеличению количества корпоративных данных.

Чтобы эффективно управлять постоянно растущим объемом информации, владельцам компаний необходимо понимать, где и как происходит этот рост. Современную тенденцию наглядно иллюстрируют два понятия: рост и разрастание.

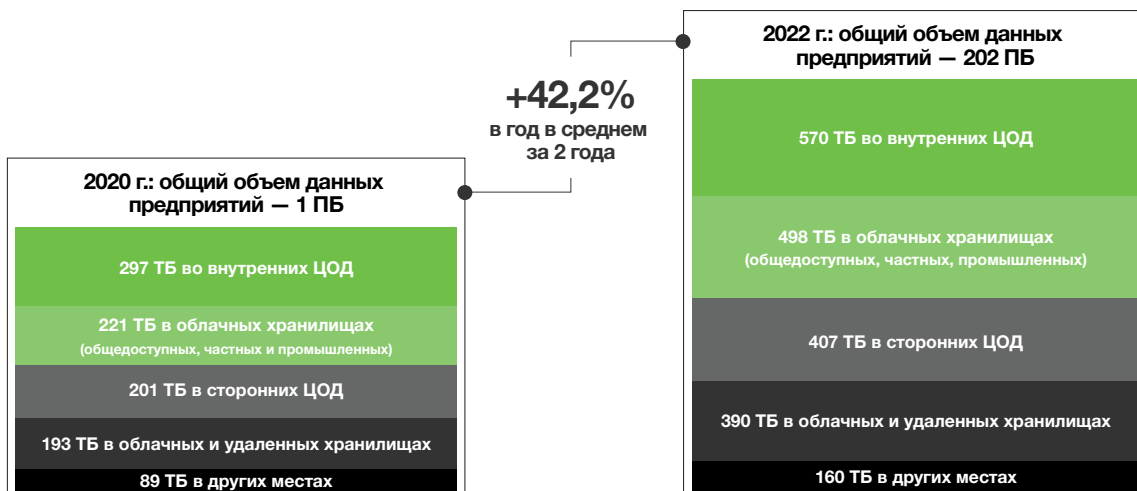
**Рост** указывает, на сколько процентов вырос объем инфосферы за определенное время. Инфосфера — это регулярно расширяющееся измерение нашей жизни, отражающее и дополняющее

ее во всех ее проявлениях и содержащее постоянно меняющиеся данные.

**Разрастание** описывает увеличение разброса и разнообразия этих данных в различных конфигурациях, от конечных точек до периферийных и облачных сред.

РИС. 1.

### Ожидаемый годовой рост объемов данных



Источник: опрос Seagate Rethink Data Survey<sup>1</sup>, IDC, 2020 г.

<sup>1</sup> Seagate Rethink Data Survey — это название опроса, используемое фирмой IDC, результаты которого представлены в настоящем отчете.



## Результаты опроса

**Рост объемов данных** остается беспрецедентным. Проведенное исследование показывает, что в течение ближайших двух лет — с начала 2020 по начало 2022 года — объем генерируемых предприятиями данных будет **увеличиваться на 42,2% в год**. Прогнозы IDC основаны на косвенном показателе — емкости корпоративных систем хранения (рис. 1).

Ответы респондентов указывают на три фактора, которыми обусловлен рост объема хранимых данных:

1. Более активное использование аналитики.
2. Увеличение количества IoT-устройств.
3. Перенос данных в облако.

**Разрастание данных** отражает степень разнесенности корпоративных данных.

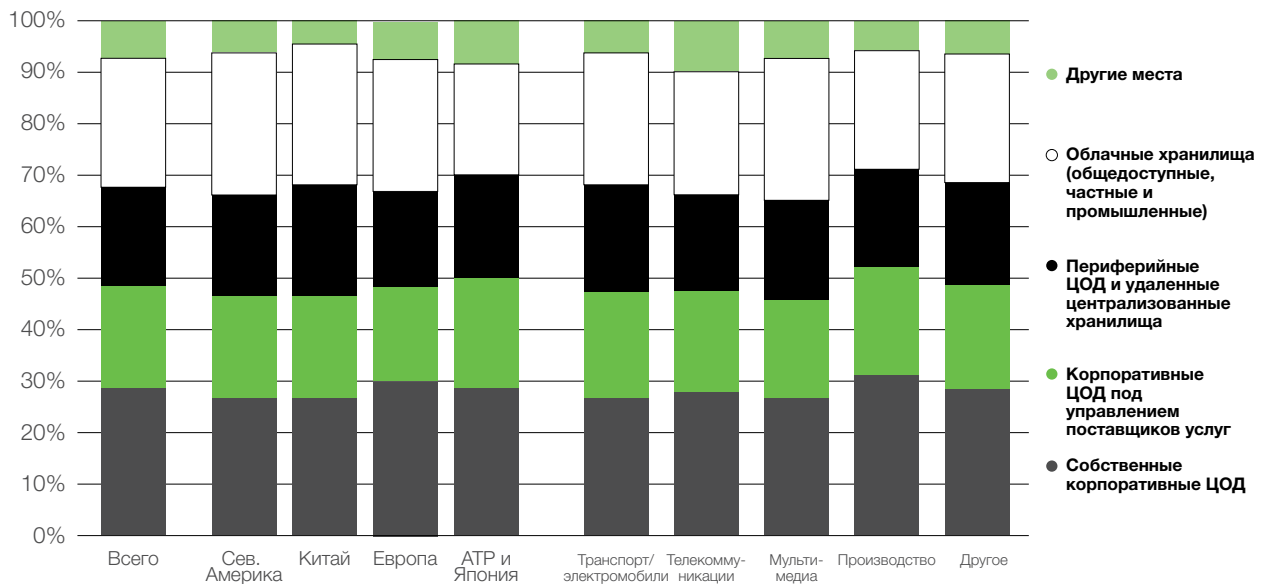
Поскольку данные предприятий хранятся распределенно, ими сложнее управлять.

Респонденты сообщили, что около 30% хранимых данных размещаются во внутренних центрах обработки, 20% — в сторонних, 19% — в периферийных или удаленных, 22% — в облачных хранилищах и еще 9% — в других местах. Ожидается, что в ближайшие два года это распределение существенно не изменится, а значит, корпоративные среды хранения в обозримом будущем останутся такими же разрозненными и непрозрачными.

Потребность предприятий в управлении распределенными данными будет ощущаться все острее.

**РИС. 2.**

### Где будут храниться данные через два года (средние показатели)



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.





Инновации порождает не мода, а необходимость создавать дополнительную ценность в условиях ограничений — а они есть везде. Одно из них — раскрытие всего потенциала данных.

Данные правят миром. Их объем растет все быстрее. Ожидается, что рост суммарного количества создаваемых данных будет экспоненциально ускоряться и к 2025 году их объем достигнет 175 зеттабайт. Сегодня мы создаем больше данных за час, чем двадцать лет назад создавали за целый год. **Данные — это человеческий потенциал.** Лучшие умы трудятся над проблемой эффективного использования их возможностей.

Ограничения, накладываемые на раскрытие всего потенциала данных, носят как системный, так и оперативный характер. Когда объемы измеряются зеттабайтами, нам требуется **простой, безопасный и недорогой способ сбора, хранения и применения данных.** Пользователи данных не хотят об этом думать. Как же решить проблему управления данными в таких масштабах и открыть путь другим инновациям, которые изменят наш образ жизни, рабочие привычки, способы передвижения и подход к защите окружающей среды?

Если нам удастся эффективно использовать возможности, обходить ограничения и при этом создавать максимальную ценность, так ли важно, следует новаторская идея модным трендам или нет? Бывает, что тренды завладевают нашими умами, но если изначальный энтузиазм не выливается в **нечто поистине ценное** для компании и для человечества в целом, инновация не приживается.

Чтобы решать проблемы на таком высоком уровне, нам необходимо **объединять разные миры** для достижения синергии.

ДЕЙВ МОСЛИ  
ГЕНЕРАЛЬНЫЙ ДИРЕКТОР SEAGATE TECHNOLOGY



## Подробные прогнозы

Дополнительные исследования компании IDC, проведенные в рамках проекта *Global DataSphere*<sup>1</sup>, проливают свет на вопрос **роста объемов данных**.

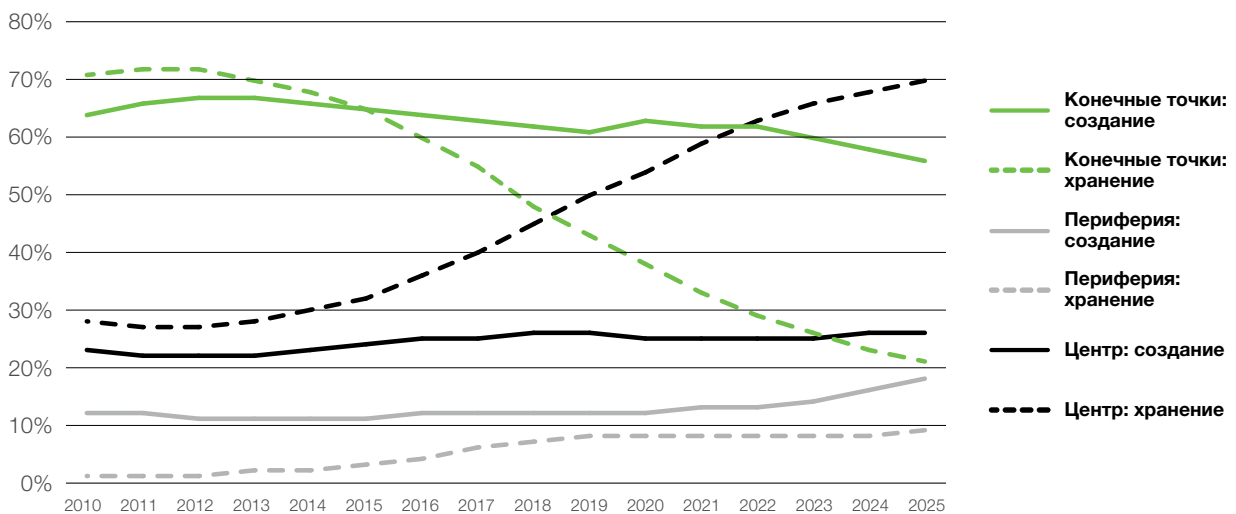
- Количество создаваемых данных ежегодно увеличивается, причем ожидается, что совокупные темпы годового прироста в 2015–2025 годах составят около 26%.
- Объем новых данных в 2025 году возрастет до 175,8 ЗБ, в то время как в 2015 году он составлял всего 18,2 ЗБ.
- В 2025 году предприятия будут хранить 9 ЗБ информации, тогда как в 2015 году этот показатель был равен 0,8 ЗБ.
- Результаты опроса *Rethink Data* подтверждают эту тенденцию. Практически все респонденты сообщали, что количество собираемых данных и емкость хранилищ в их организациях

увеличиваются. Это говорит о том, что более активное использование аналитики и IoT-устройств, а также перенос данных в облако входят в тройку главных факторов, влияющих на рост объема хранимых данных.

- Данные исследований *Global StorageSphere*<sup>2</sup> и *Global DataSphere* помогают нам лучше понять процесс разрастания данных.
- По результатам *Global DataSphere*, в 2015 году 65% данных было создано на конечных точках и 35% в центре и на периферии.
- Ожидается, что к 2025 году 44% всех данных, созданных в центре и на периферии, будет использоваться для аналитики, искусственного интеллекта и глубокого обучения, а данные с растущего числа IoT-устройств будут передаваться на периферию корпоративной сети.

- Центр тяжести данных смещается и в направлении центра, и в направлении периферии. К 2025 году почти 80% всех данных в мире будут храниться в центре и на периферии, что составит значительный рост по сравнению с 35% в 2015 году.
- По прогнозам IDC, к 2025 году емкость запоминающих устройств (жестких и оптических дисков, твердотельных и ленточных накопителей), используемых предприятиями, составит 12,6 ЗБ. Поставщики облачных услуг будут управлять 51% от этой емкости.

**РИС. 3. Где создаются и хранятся данные**



Источник: отчет *Data Age 2025*, профинансированный Seagate, с использованием данных исследований IDC *Global DataSphere* и *StorageSphere*, май 2020 г.

<sup>1</sup> Исследование *Data Age 2025*, финансируемое Seagate, с использованием данных *Global DataSphere*, IDC, май 2020 г.

<sup>2</sup> Исследование *Global StorageSphere*, IDC, 2020 г.



## Что такое периферия?

Словом «периферия» можно описать расположение тех или иных компонентов сети в самых разных отраслях. Примеры:

- производственные площадки;
- крыши зданий;
- вышки сотовой связи;
- коровники на ферме;
- беспилотные автомобили;
- нефтегазодобывающие платформы.

Периферия — это **не «что», а «где»**. Это внешняя граница сети, которая может находиться за сотни и даже тысячи километров от ближайшего предприятия или облачного центра обработки данных, как можно ближе к источнику данных. Это место, где происходит принятие решений в режиме реального времени.



## Производительность на периферии

Все большие объемы данных требуют анализа и принятия решений на периферии.

По технологическим и экономическим соображениям собирать, хранить и обрабатывать данные таким образом становится выгоднее.

Смещение центра тяжести в направлении периферии обусловлено развитием четырех технологий:

1. ИИ стал более экономичным и практичным.
2. Развернуты миллиарды IoT-устройств.
3. Операторы беспроводной связи модернизируют свои сети для поддержки пятого поколения мобильной связи (5G).

4. Инновации в периферийных центрах обработки данных решают проблемы распределенных площадок, а удельные издержки снижаются.

Помимо этих технологий, спрос на периферийные вычисления в значительной степени стимулируют такие факторы, как задержки, большие объемы данных при недостаточной пропускной способности, высокая стоимость, а также суверенитет данных и требования нормативно-правового соответствия.

Поскольку огромное количество информации создается вне традиционных центров обработки данных, облако разрастается по направлению к периферии. Но речь идет не о противостоянии, а о синергии облака и периферии.



## В центре внимания: данные на периферии

Объем данных, хранимых на периферии, растет быстрее, чем в центре.

Настоящее исследование показало, что в среднем организации периодически перемещают около 36% своих данных с периферии в центр. **Всего через два года этот показатель вырастет до 57%.**

Объем данных, передаваемых с периферии в центр в режиме реального времени, увеличится с 8% до 16%. Планы управления данными должны считаться с таким развитием событий: необходимо готовиться к масштабному перемещению данных по направлению от конечных точек через периферию к публичным, частным и промышленным облакам.

### Что это означает для предприятий?

Прежде всего, их ожидает большее разрастание данных. Это ведет к

обособлению и разрозненности, в результате чего не все и не всегда могут получить доступ к нужной информации. Без автоматизации управление разрастанием данных требует значительных усилий со стороны сотрудников и приобретения излишних программных инструментов.

Ожидается, что все больше устройств хранения будет поддерживать некоторые вычислительные функции, а в дальнейшем, возможно, полностью обеспечивать необходимые вычисления.

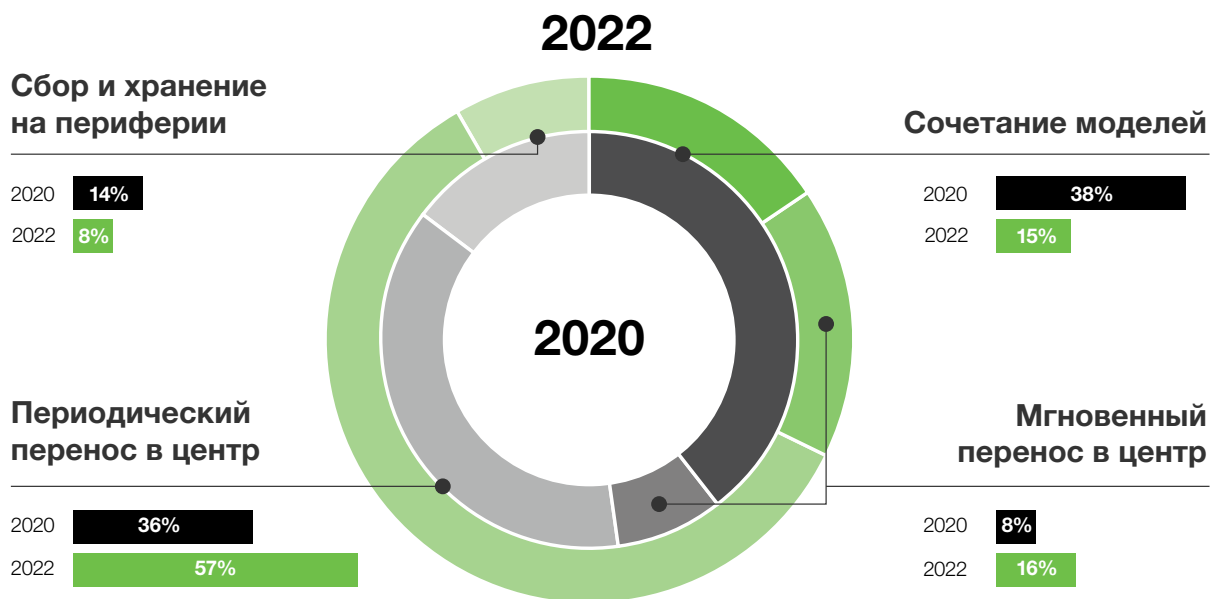
Предполагается, что на периферии будут храниться критически важные данные, необходимые для обработки запросов, чувствительных к задержкам, поступающих с конечных точек. В то же время распределенные вычисления позволят анализировать потоковые данные.

Потоковые данные, скорее всего, будут кешироваться на носителях до тех пор, пока серверы не начнут поддерживать полноценные функции аналитики.

Таким образом, **граница между хранением и кешированием на периферии может оказаться размытой**, особенно с учетом того, что данные, вероятно, будут храниться там недолго: сразу после анализа или обработки вся релевантная информация будет передаваться в центр.

РИС. 4.

### Подход к сбору данных на периферии: сегодня и через два года



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



## ГЛАВА ВТОРАЯ

# Кладезь ценности

**Инфосфера отражает все богатство и разнообразие нашей жизни. В ней можно найти и медицинские сведения, способные спасти жизнь, и хронику открытий и изобретений, и инструкции по работе с устройствами, и описание производственных процессов, и записи в дневниках. Данные — практически бездонный кладезь ценности.**

### ТОЧКА ЗРЕНИЯ SEAGATE

Компании-лидеры понимают: то, как они обращаются со своими данными, может непосредственно отражаться на росте их бизнеса.

Для предприятий, которым пользователи доверяют данные,

например поставщиков облачных услуг и потокового контента, производителей умных устройств, больниц, сетей супермаркетов и т. д., это означает необходимость инвестиций в сбор, анализ и

хранение данных в различных репозиториях.

Чем активнее эти данные будут использоваться, тем большую ценность они принесут.



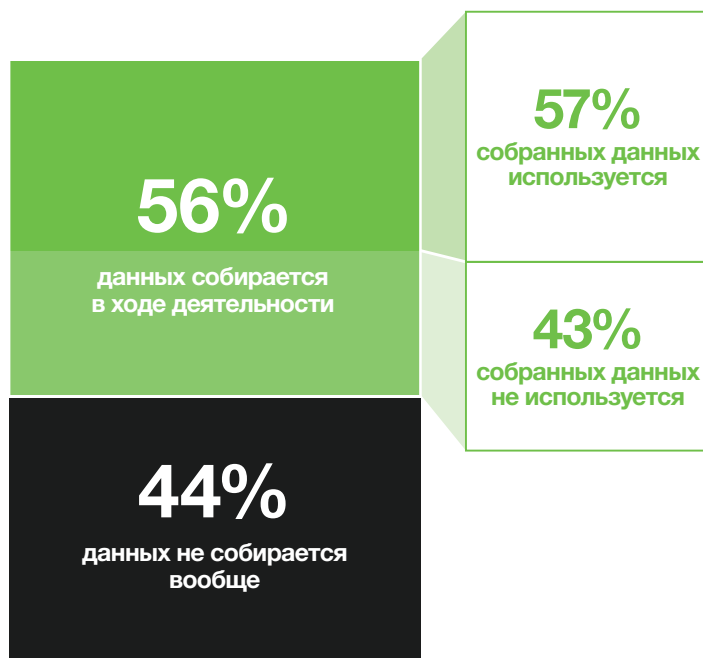
## Упущенные данные

Согласно результатам опроса *Rethink Data*, **значительная часть бизнес-данных организаций не используется**. Данные таят в себе огромные возможности извлечения ценности, но эти возможности часто упускают.

- Респонденты сообщили, что в их организациях собирается лишь 56% данных, которые потенциально могут быть получены в результате их деятельности. Таким образом, практически половина данных игнорируется.
- Из этих 56% только 57% данных были применены на практике.
- 43% собранных данных не использовались вообще.
- Это означает, что **лишь 32% доступных бизнес-данных используются организациями, а целых 68% простаивают без дела.**

РИС. 5.

Какая часть данных на самом деле используется?



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



## Ценность данных

Во многих организациях данные — это некий неосознанный актив, который не упоминается в финансовых отчетах. А ведь при этом данные все чаще используются для выявления новых источников дохода, повышения качества обслуживания и эффективности операций. Проще говоря, данные могут повысить ценность предприятия, хотя с точки зрения бухгалтерии и не являются активом.

Однако чтобы данные приносили пользу, они должны работать. Согласно результатам опроса, организации используют лишь 32% доступных бизнес-данных.

Чтобы заставить данные работать и раскрыть их потенциал, сначала нужно понять, как измерять ценность самих данных. По данным опроса IDC, проведенного для компании Seagate в 2018 году<sup>1</sup>, лишь в 25% организаций по всему миру существуют системы и процессы, позволяющие активно измерять ценность бизнес-данных — по крайней мере, в некоторых контекстах.

У предприятий есть множество возможностей более эффективного использования данных для получения конкурентных преимуществ. Тот же опрос показал, что только 11% организаций считают себя лидерами в своей отрасли в смысле способности извлекать ценность из бизнес-данных. Этот показатель оказался даже ниже для сферы здравоохранения и транспортной отрасли, где положительный ответ на этот вопрос дали менее 10% организаций.

Ясно одно: компаниям нужны «живые» озера данных, в которые будет своевременно поступать свежая информация, а старая и неиспользуемая будет передаваться в хранилища с низкой стоимостью.

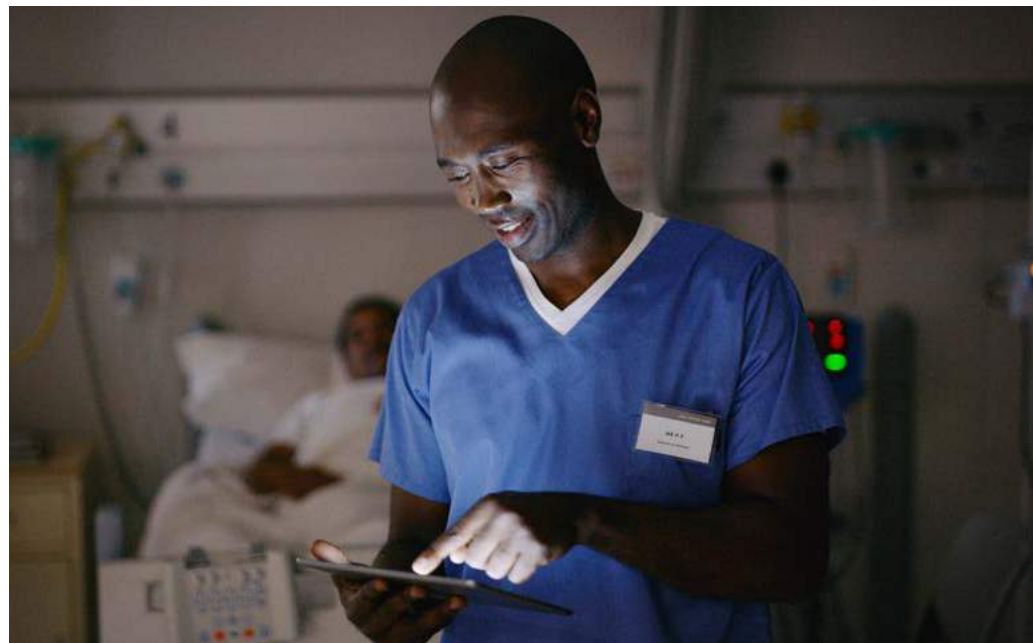
**Ни одна компания не хочет, чтобы ее озеро данных превратилось в болото, полное необработанной, но потенциально полезной информации.**

При измерении ценности данных каждая организация оперирует множеством переменных, таких как отрасль, в которой создаются данные, цель, которой они служат, и, наконец, способы их монетизации.

Рассмотрим в качестве примера типы данных, которые создаются больницей. Это информация о пациентах, расписания смен, сведения о страховке и выставленных счетах, снимки томографии, информация о методах лечения рака, оперативные и финансовые показатели, а также данные, используемые в рекламных целях. Согласно закону, больницы обязаны хранить данные в течение многих лет после смерти пациента. Скорее всего, это неактивные данные, которые, однако, можно использовать. Разные данные имеют разную ценность — это вдвойне верно для конфиденциальных данных, требования к хранению которых особо строги. В будущем видеосеансы дистанционного оказания медицинских услуг, действия хирургов во время операции или даже материалы роботизированной операции будут записываться и сохраняться по разным причинам, даже если это

будет необходимо лишь для обучения или в юридических целях. Можно ли измерить ценность таких данных?

IDC еще только предстоит дать количественную оценку ценности глобальной инфосферы. Тем не менее, если исходить из определенных предположений о ценности байта, созданного в больнице с численностью персонала от 1000 до 2500 человек и доходом свыше 1 млрд долларов США, то, по расчетам IDC, ценность всей совокупности данных, созданных в такой больнице, может достигать нескольких сотен миллионов долларов США.



<sup>1</sup> Отчеты Worldwide DATCON и VoB, IDC, сентябрь 2018 г.





**Ни одна современная компания немыслима без данных. Но какой от данных толк, если с ними ничего не делать? Чтобы они приносили пользу, в озера данных должно впадать множество бурных рек, и очень важно не захлебнуться в потоке поступающей информации. Нужно делиться данными с другими озерами, чтобы помечать информацию перекрестными ссылками и анализировать разрозненные потоки.**

Рассмотрим в качестве примера беспилотные автомобили. Анализ данных с одного автомобиля и внутри одной компании выгоден сам по себе. Но перекрестный анализ данных этого автомобиля с данными автомобилей всех остальных компаний добавляет еще один уровень понимания. Следующий шаг — объединение информации о работе одного беспилотного автомобиля с данными миллионов датчиков, установленных в умных городах. Такая комплексная картина может быть полезна региональным властям и градостроителям, которые смогут внедрить новые стандарты общественной безопасности и дорожного движения.

**Чем больше фрагментов вы составите вместе, тем более полное и точное представление о реальности получите. Обмениваясь данными и создавая перекрестные ссылки на разные потоки информации для их анализа, можно решить проблему более высокого порядка.**

**Вот почему перемещение данных так важно. Данные должны находиться в движении, чтобы было возможно выявлять связи между ними — и, как следствие, составлять точную и подробную картину происходящего.**

РАВИ НАИК  
СТАРШИЙ ВИЦЕ-ПРЕЗИДЕНТ И ГЛАВНЫЙ  
ДИРЕКТОР ПО ИНФОРМАЦИОННЫМ  
ТЕХНОЛОГИЯМ SEAGATE TECHNOLOGY



**«В наше время конкурентное преимущество определяют данные — структурированные или нет, созданные человеком или машиной, хранящиеся в центре обработки или в облаке».**

**Фил Гудвин,**  
ДИРЕКТОР  
ПО ИССЛЕДОВАНИЯМ IDC

#### АНАЛИЗ IDC

## Пусть данные работают на вас

У организаций должна быть возможность получать необходимые данные, идентифицировать их, хранить в удобном месте и эффективно использовать для принятия решений.

Активация данных — процесс их применения для целей предприятия — начинается со сбора. В связи с растущей популярностью IoT-приложений объемы данных увеличиваются в геометрической прогрессии. Неудивительно, что организации не собирают все доступные данные, ведь это чревато огромными затратами и перегруженной ИТ-инфраструктурой. Без надежного решения для управления компаниям приходится полагаться на то, что

ПО для приема данных сможет правильно идентифицировать и классифицировать их в самом начале жизненного цикла. При выполнении этого условия данные могут храниться в течение полезного периода, а затем автоматически удаляться или архивироваться согласно правилам. Отсечение данных упрощает управление ими и снижает затраты. Однако по мере развития технологий управления данными компании смогут собирать большие объемы информации и использовать ее для машинного обучения или совершенствования искусственного интеллекта.

## На периферии

Периферийные устройства представляют определенные сложности с точки зрения сбора данных. Нередко только периферийное приложение имеет инструкции о том, какие данные необходимо собирать и обрабатывать, а какие можно игнорировать. Из-за этого многие решения принимаются там же, где создаются данные. Однако в более новых приложениях для централизованного управления данными эти решения могут приниматься с помощью искусственного интеллекта (ИИ) и машинного обучения (МО). Такие программы могут идентифицировать конфиденциальные данные (например, личную или медицинскую информацию, номера кредитных карт и т. д.) и автоматически скрывать их от всех, кто не уполномочен их просматривать. Это снижает вероятность несанкционированного доступа или непреднамеренного раскрытия.

Собранные данные затем анализируются. Чаще всего это делается в озерах данных, где они оцениваются специализированным ПО согласно отраслевым стандартам и другим критериям. С помощью таких инструментов кураторы данных и исследователи извлекают полезную информацию для принятия решений. Помещение данных в озеро решает проблему обособленных баз данных и позволяет сопоставлять элементы, которые, на первый взгляд, никак не связаны между собой.

Именно это дает компании конкурентное преимущество.

Хранение является неотъемлемой частью стратегии управления данными в организации, поскольку без него невозможны сортировка и анализ. Доступ к достоверным данным в режиме реального времени необходим для совершенствования рабочих процессов, обеспечения безопасности и управления ресурсами, поэтому спрос на передовые решения в области управления данными и аналитики продолжает расти.



## Ценность данных и инновации в их хранении

От того, где и как хранятся данные, зависит, какую пользу организация может из них извлечь. Ниже перечислены области, где инновации в сфере хранения данных непосредственно отражаются на ценности последних.

1. Экономия на масштабе за счет **растущих объемов**. С ростом количества доступных данных повышается качество аналитики. Грамотное управление данными подразумевает, что к ним применяются все инновации в области ИИ или МО — насколько это физически возможно. Именно поэтому такие компании, как Seagate, усиленно работают над повышением плотности записи, чтобы на устройство можно было вместить больше информации.

Внимание к плотности записи обусловлено ситуацией в крупнейших в мире облачных хранилищах.

Данные предприятий распределяются в соотношении 90 к 10: 90% хранится на жестких дисках, а 10% — на твердотельных накопителях.

2. Еще одно направление развития в области хранения данных — **повышение пропускной способности**, обеспечивающее более быстрое и стабильное перемещение данных между хранилищем и сетью, а также увеличение скорости вычислений. Это важно для аналитики. Современные инструменты анализа используют графические процессоры, которым требуется высокая пропускная способность. Для ее увеличения в крупных корпоративных системах с ИИ, в частности, используется деагрегированная компонентная архитектура.

3. **Безопасность** также не остается без внимания. Компании продолжают инвестировать в защиту устройств: создаются специальные микропрограммы, а в вычислительных системах реализуются протоколы цифровой проверки подлинности устройств. Системные решения выигрывают от повышенной безопасности на уровне компонентов и устройств, а сеть — от более защищенных систем. И, наконец, в более безопасных сетях лучше защищены вычислительные ресурсы.

4. В долгосрочной перспективе **архитектура перемещения данных** должна будет обеспечивать аппаратное ускорение или аппаратную разгрузку через системы хранения данных. В частности, сжатие, шифрование и дедупликация наборов данных сегодня осуществляются в вычислительных системах. Из-за этого возникает потребность в масштабировании крупных архитектур, поскольку эти задачи выполняются на более высоком уровне. Но если аппаратное ускорение и аппаратная разгрузка будут выполняться на уровне хранилища или сети, это не потребует.



# Многооблачные среды

**Многие компании обязаны своим ростом общедоступным облачным хранилищам. Однако при достижении определенного размера предприятия такое решение оказывается недостаточно гибким.**

### ТОЧКА ЗРЕНИЯ SEAGATE

Компании часто сталкиваются с проблемой доступа к некоторым своим данным, хранящимся в общедоступном облаке, и испытывают трудности с их переносом — не говоря уже о затратах, которыми сопровождается извлечение данных. Это мешает использовать данные, хранящиеся в общедоступном облаке, максимально эффективно: чтобы создавать ценность, данные

должны двигаться. В результате предприятия все чаще предпочитают комбинированные варианты хранения и управления данными в многооблачной экосистеме.

#### Многооблачная среда

подразумевает использование общедоступных облачных хранилищ разных поставщиков и координированное управление хранящимися в них данными. Иногда

такие решения включают частное облако. На практике многооблачная архитектура может означать, что часть корпоративных данных хранится в общедоступном облаке, а часть — локально в **частном облаке**.

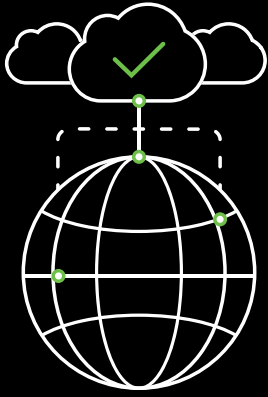
Возврат части данных в частное облако позволяет совмещать его преимущества с выгодами многооблачной экосистемы.



## Преимущества частного облака

1. **Предсказуемая экономия на масштабе** под контролем организации.
2. **Право собственности на IP-адреса хранилищ**, обеспечивающее защиту и контроль физических систем, в которых находятся данные. Это полезно, например, в ситуациях, когда доступ к данным необходим для аудита или соблюдения нормативных требований.
3. Частое **обращение к большим наборам данных**. При регулярном чтении и анализе больших объемов данных затраты на хранение возрастают, и здесь частное облако помогает сэкономить.
4. **Соблюдение нормативных требований в отношении конфиденциальных данных**. Например, закон предписывает хранить некоторые медицинские сведения в юрисдикциях пациента и поставщика услуг здравоохранения, а такие нормативные акты, как HIPAA и GDPR, накладывают строгие ограничения на раскрытие личных данных.





## Преимущества общедоступного облака, присущие многооблачной экосистеме

- Быстрый рост и масштабирование.** В многооблачной среде можно разрабатывать и развертывать приложения или службы в общей среде и получать от них прибыль без капитальных затрат.
- Доступ к обширному каталогу служб,** включая приложения, вычислительные ресурсы и инструменты для отслеживания. Многооблачная инфраструктура позволяет выполнять высокопроизводительные вычисления, которые были бы невозможны в большинстве частных облачных сред. Необходимые графические процессоры стоят от 500 тысяч до 5 миллионов долларов США. Преимущество общедоступного облака заключается в том, что организация может арендовать пространство на ограниченное время. В свою очередь, вычислительные ресурсы в общедоступном облаке обеспечивают максимальную мощность, что позволяет наращивать эффективность и не накапливать неиспользуемые данные.

**Гибридное облако** органично объединяет ресурсы частного и общедоступного облака, предоставляя интегрированную инфраструктуру как услугу. Гибридизация обеспечивает беспрепятственное взаимодействие между отдельными облаками и ресурсами. Что это означает на практике? Предположим, компания приобретает у производителя вычислительной техники корпоративное решение,

размещает его на общей площадке на условиях колокации и подключает внутреннее хранилище к протоколу поставщика облачных услуг.

Как правило, управление гибридным облаком осуществляется централизованно через портал управления — на этом уровне решаются задачи подключения и взаимодействия с приложениями.

В многооблачной среде приложение, которое обычно рассчитано

на работу в облаке, эффективно использует ресурсы и управляет ими. В случае гибридного облака подключение устаревших систем к общедоступным хранилищам может вызывать сложности.



## Многооблачность как организационная стратегия

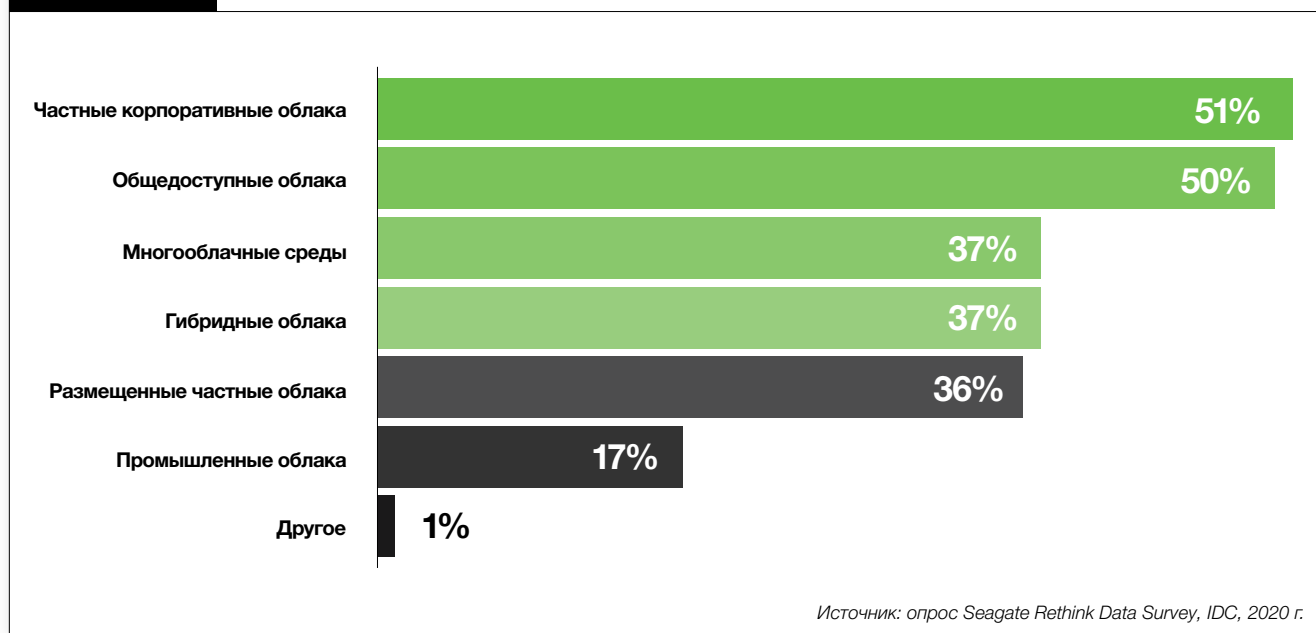
По определению аналитиков IDC, многооблачность — это «организационная стратегия или архитектурный подход к проектированию комплексных цифровых служб, которые предполагают получение облачных услуг от нескольких поставщиков». Это могут быть услуги

прямых конкурентов, например общедоступные хранилища от разных поставщиков, а также предоставление инфраструктуры как услуги (IaaS) или ПО как услуги (SaaS) одним или несколькими поставщиками. По мнению специалистов IDC, в обоих случаях многооблачная среда

охватывает гораздо больший спектр возможностей, чем гибридное облако, причем масштаб охвата определяется только стоимостью и сложностью обеспечения согласованного управления множеством различных облачных услуг.

**РИС. 6.**

**Типы облаков в корпоративных ИТ-инфраструктурах**



### ДАННЫЕ IDC

## Управление данными в нескольких облаках

Участники опроса сообщают о большом разнообразии используемых облачных услуг, включая сотрудничество с несколькими разными поставщиками.

- Исследование показало одинаковую меру использования многооблачных и гибридных облачных сред (по 37%).
- Обращение к гибриднему облаку, как правило, обусловлено используемыми приложениями.

Многооблачная среда формируется естественным образом, по мере того как отдельные подразделения начинают сотрудничать с разными поставщиками облачных услуг для решения определенных задач.

- У предприятий, где используются многооблачные среды, не обязательно есть целостный долгосрочный план интеграции.
- Многооблачные и гибридные облачные среды решают разные

задачи. Они могут облегчить доступ к данным и их анализ, повысить их безопасность, сократить расходы и обеспечить администраторам более полноценный контроль.

Однако многооблачные и гибридные облачные экосистемы также создают определенные проблемы, связанные с управлением данными.



«Многооблачную инфраструктуру сложнее развернуть, и управлять ей также сложнее, поскольку для каждой облачной платформы предусмотрены собственные инструменты, оптимизированные для работы с ней. Кроме того, смежные функции управления данными или аналитики, предлагаемые поставщиками общедоступных облаков, чаще всего разрабатываются для использования с инфраструктурой конкретной платформы и при попытке интегрировать их с другими общедоступными облачными платформами могут работать с ограниченной функциональностью или не работать вовсе».

ЭНДРЮ СМИТ,  
РУКОВОДИТЕЛЬ ИССЛЕДОВАНИЯ IDC

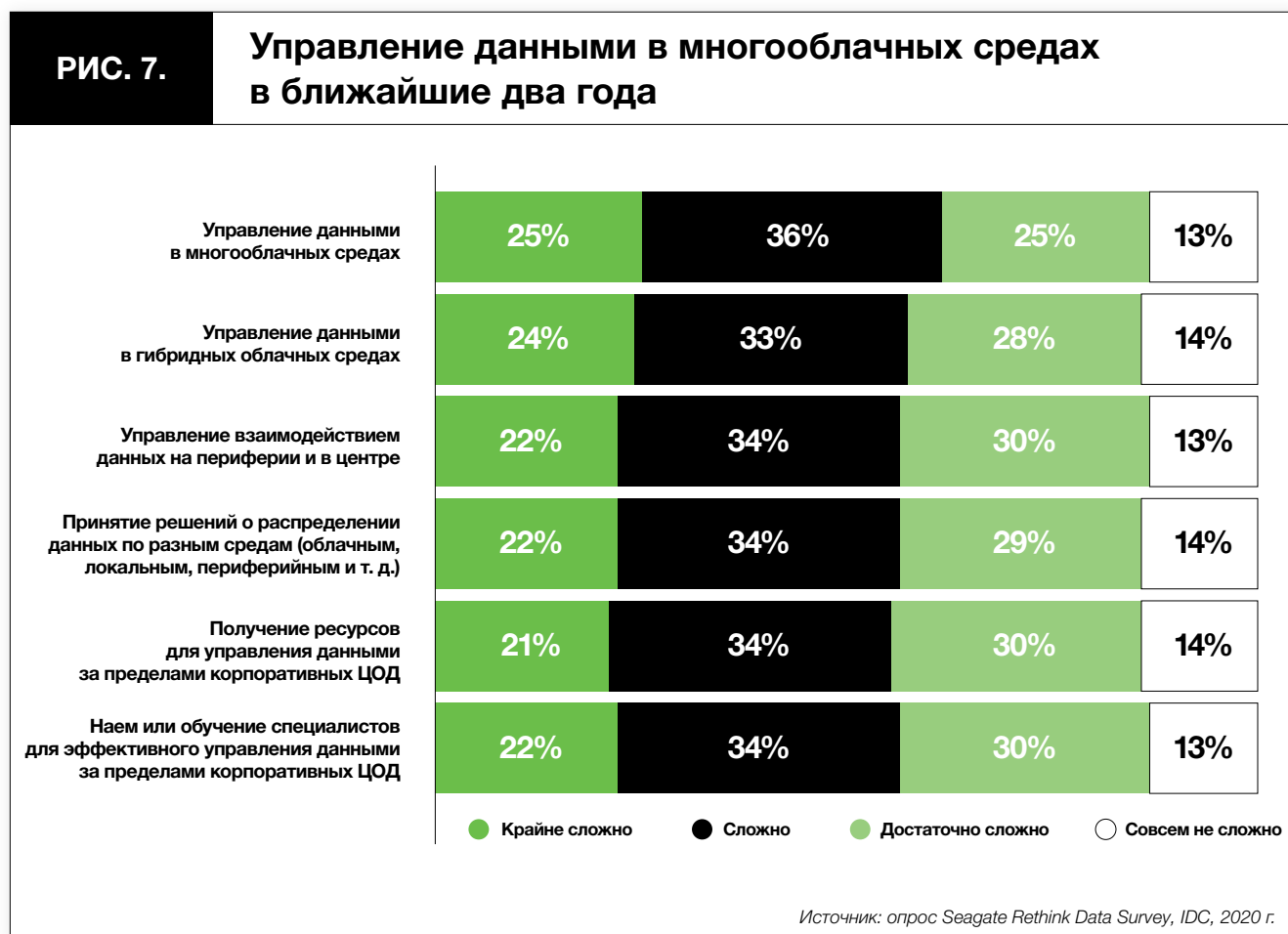
ДАННЫЕ IDC

## В двух словах

- Управление данными в многооблачных средах — самая актуальная задача в управлении данными на ближайшие два года.
- Управление данными в гибридных облаках занимает второе место в списке сложностей, связанных с управлением данными.

РИС. 7.

### Управление данными в многооблачных средах в ближайшие два года



## Подробные прогнозы

По данным исследования *IaaSView* компании IDC <sup>1</sup>, самый распространенный вариант многооблачной архитектуры — размещение приложений в разных общедоступных облаках по модели IaaS и регулярное взаимодействие этих приложений между собой. Этот вариант может работать на уровне приложений, но на уровне управления

данными остаются важные проблемы:

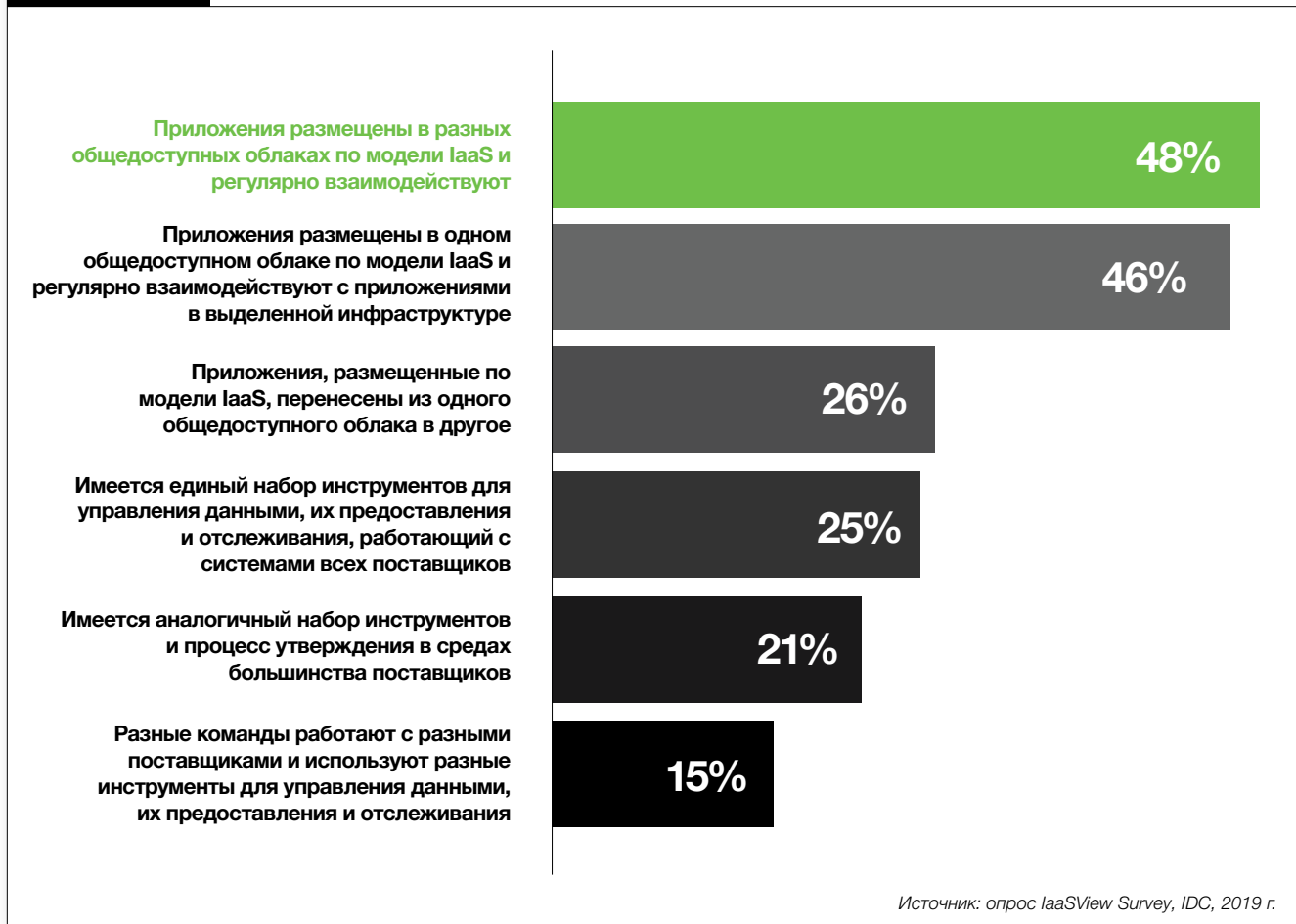
- обособленные рабочие процессы;
- разрозненные инструменты управления;
- отсутствие единой концепции управления безопасностью;
- сложности с обменом данными и перемещением больших объемов

данных между облаками разных поставщиков.

Эти проблемы усложняют и замедляют выполнение в многооблачной среде даже самых простых задач, связанных с работой с данными, таких как создание отчетов или информационных панелей.

РИС. 8.

### Интеграция общедоступных облачных служб IaaS разных поставщиков



<sup>1</sup> Опрос *IaaSView Survey*, IDC, 2019 г.



## ГЛАВА ЧЕТВЕРТАЯ

# Проблемы управления данными и многооблачная экосистема

### ТОЧКА ЗРЕНИЯ SEAGATE

Итак, мы уже установили, что:

- бизнес-данные скрывают в себе большую ценность;
- значительная доля этой ценности не извлекается,

из-за чего компании лишаются потенциального дохода;

- предприятия собирают лишь часть потенциально доступных данных.

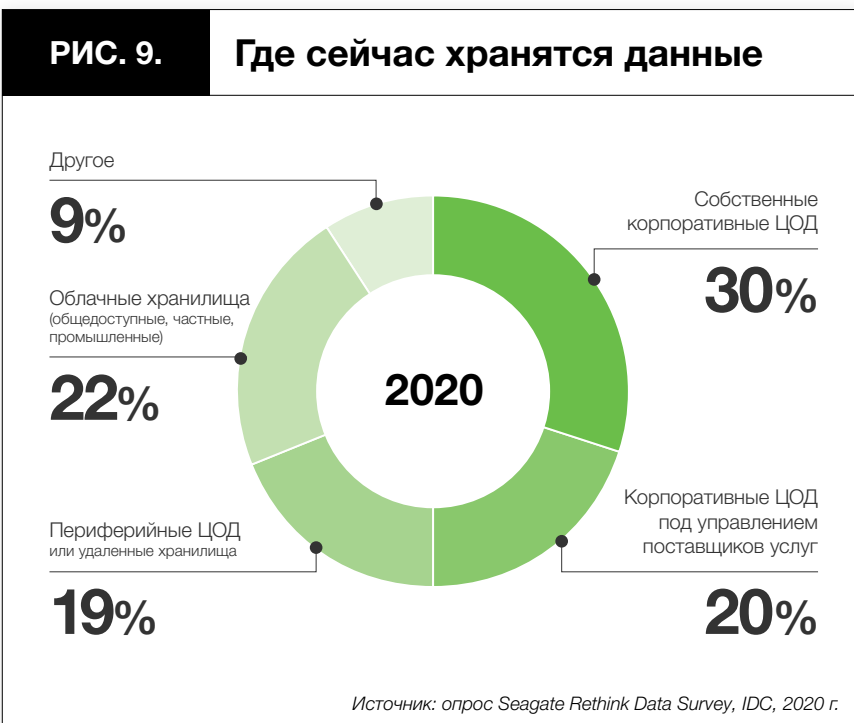
Но существуют и другие причины, по которым управлять данными так сложно. Они связаны с особенностями многооблачной экосистемы, и мы рассмотрим их в этой главе.

### ДАННЫЕ IDC

## Управление данными

Опрос показал, что когда дело доходит до управления одним из самых ценных ресурсов предприятия — данными, — компании не могут самостоятельно справиться с усложняющейся ситуацией. Респонденты сообщают, что им требуется помощь.

Главная причина — широкое и относительно равномерное распределение данных между периферийными и облачными хранилищами. Имеются и специфические отличия, характерные для отдельных отраслей. Например, в транспортной отрасли и сегменте электромобилей большинство данных сосредоточено на периферии, в то время как в производственном секторе заметна склонность к хранению информации во внутренних центрах обработки.



**РИС. 10.**

**Где хранятся данные: сегодня и через два года**

**Другое**

2020 **9%**  
2022 **8%**

**Облачные хранилища**

(общедоступные, частные, промышленные)

2020 **22%**  
2022 **25%**

2020 **19%**  
2022 **19%**

**Периферийные ЦОД**

или удаленные хранилища

**2022**

**2020**

**Собственные корпоративные ЦОД**

2020 **30%**  
2022 **28%**

2020 **20%**  
2022 **20%**

**Корпоративные ЦОД**

под управлением поставщиков услуг

Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

**Факторы, влияющие на изменения в сфере управления данными**

Ответы на вопрос, какой фактор является решающим для изменения подхода организации к управлению централизованным хранилищем, распределились следующим образом:

- 17% — повышение безопасности данных;
- 14% — расширение доступа к службам анализа и управления данными (ИИ/МО, IoT и т. д.);
- 14% — повышение прозрачности ИТ-инфраструктуры и более эффективное управление;
- 11% — снижение общей стоимости владения инфраструктурой и затрат на ее обслуживание;
- 10% — более быстрый доступ к данным для приложений и сотрудников.

**РИС. 11.**

**Факторы, влияющие на изменения в способах хранения данных**



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



## Владельцы компаний, не собирайте данные ради самого факта

Прежде чем совершенствовать процессы управления данными в компании, необходимо понять, что это за данные.

К сожалению, в большинстве организаций они просто собираются в огромные репозитории.

Такие данные сложно анализировать и еще сложнее извлекать из них какую-либо пользу. Если в компании не понимают, с какими данными имеют дело, то определить, какие сведения нужно собирать и зачем, попросту невозможно.

В первую очередь руководителям необходимо разобраться, для чего их компания получает и анализирует данные. После этого можно целенаправленно работать в одном направлении, а не пытаться вслепую решать все задачи сразу.

**Собрать данные просто. Сложно извлечь из них нужную информацию.**

На пути к эффективному сбору и сортировке данных компании сталкиваются с такими препятствиями, как дублирование инструментов, сложность данных, проблемы с их интеграцией, шум, неверное сопоставление и т. д.

Собираемые данные должны отвечать на конкретные вопросы, волнующие владельцев компаний. Без этого накопление данных само по себе ни к чему не приведет.

Рассмотрим в качестве примера производственную отрасль. Во многих сферах для управления производственными экосистемами используется огромное количество IoT-устройств и платформ, конечных устройств

с датчиками и различные функции МО и ИИ. Данные собираются в беспрецедентных масштабах, но они практически бесполезны, если владельцы не знают, чего хотят от этих данных. Например, очень сложно повысить производительность производства, если на фабриках используются десятки тысяч разрозненных IoT-устройств, развернутых на разных платформах.



## Понятное и прозрачное хранение данных

Существует и другой способ управлять данными эффективно — интеллектуальные решения для их хранения.

Проблемы обусловлены сочетанием множества факторов:

- нестандартная архитектура;
- параллельное использование все большего количества разных технологий хранения;
- ограниченное пространство для размещения хранилищ;
- непрозрачность и несогласованность при управлении хранимыми данными (объекты и хранилища файлов отображаются по-разному);
- трудности с принятием решений о том, какие данные где хранить;

- непомерно высокая стоимость некоторых решений для хранения данных;
- невозможность просматривать многооблачный пул хранилищ как единое целое, включая локальные и облачные компоненты.

Чтобы устранить эти сложности, нужно понять, как владельцы компаний **смотрят на хранимые данные**. В идеале данные должны быть как на ладони, сколько бы их ни было. Эта задача выходит за рамки демократизации данных и требует унификации хранилища. Главный директор по информационным технологиям должен иметь возможность видеть несколько облачных экосистем как единое целое.



Проблему, которую наглядно иллюстрирует перенос данных в многооблачную экосистему, можно решить за счет хранения в облаке, где предприятие сможет контролировать свои данные. Это может быть внутреннее хранилище, частное или даже общедоступное облако, хотя в последнем случае

лишь немногие поставщики дают клиентам возможность беспрепятственно экспортировать свои данные.

Еще одно возможное решение, над которым сейчас работает ряд компаний, — виртуализация хранилищ. В этом случае целостность и прозрачность

обеспечивается за счет добавления программного уровня управления хранилищем данных.

Ясно одно: чтобы руководители компаний действительно понимали свои данные, им необходимы средства наглядного и удобного просмотра хранилищ.

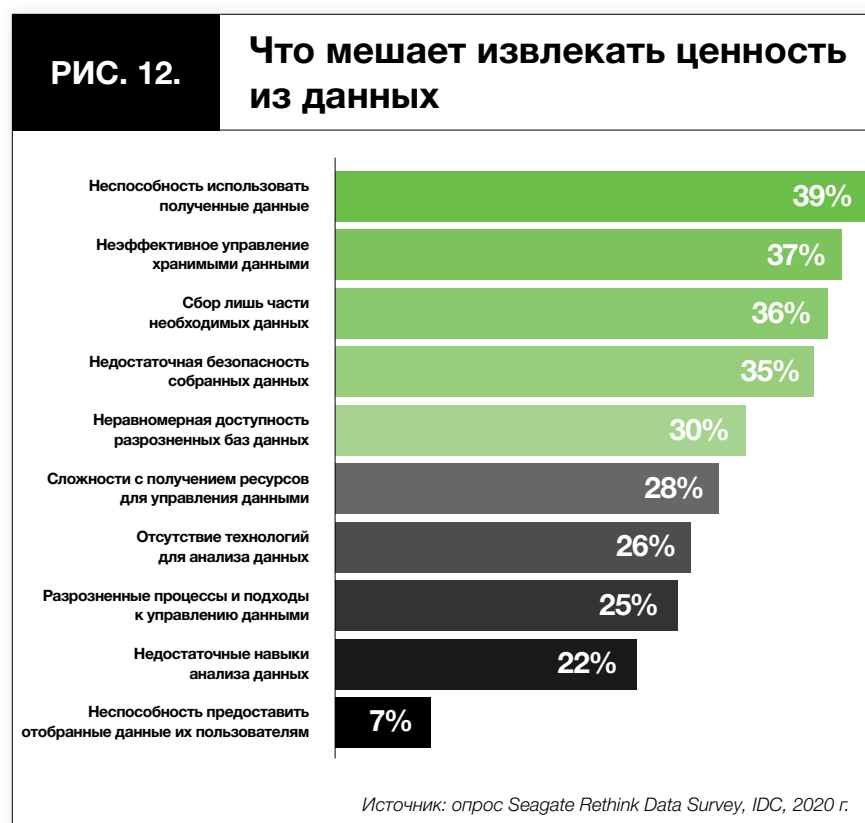
## ДАнные IDC

### Сложности управления данными

Согласно результатам опроса, организации сталкиваются с пятью основными препятствиями для использования всего потенциала собранных данных:

1. Неспособность использовать полученные данные.
2. Неэффективное управление хранимыми данными.
3. Сбор лишь части необходимых данных.
4. Недостаточная безопасность собранных данных.
5. Неравномерная доступность разрозненных баз данных.

Эти соображения сформулированы на языке ИТ, но касаются и владельцев компаний: именно они непосредственно определяют, какую ценность можно извлечь из данных, а это влияет на уровень дохода (подробнее об этом — в шестой главе). Современные решения для управления данными должны избавлять компании от этих проблем, чтобы и их владельцы, и клиенты могли успешно работать с данными, а предприятия могли задействовать



все большую долю собираемых данных и извлекать из них ценность.

В следующей главе мы рассмотрим недостающее звено для эффективного управления данными.



## Производство — исключение из правил

Согласно результатам нашего опроса, одна из самых отстающих отраслей в вопросе управления данными — это производство.

Как ни странно, в этом секторе самые низкие уровни автоматизации задач управления данными и полной интеграции соответствующих функций (на единой платформе).

- Производство отстает по уровням внедрения многооблачных и гибридных инфраструктур.
- Уровень удовлетворенности подходом к управлению данными в этой, а также в телекоммуникационной и медийной/CDN отраслях, оказался ниже среднего.
- Как и респонденты из телекоммуникационной отрасли, представители производственного сектора недовольны имеющимися инструментами для управления данными.
- Главная проблема управления данными на производстве — их хранение.

Кроме того, темпы прироста объемов данных в производственном секторе оказались едва ли не самыми низкими (37% при среднем показателе 42,2%). При этом локальные корпоративные центры обработки данных в производственном секторе, как оказалось, занимают самые большие площади по сравнению со всеми остальными охваченными опросом отраслями. Возможно, в этом и заключается причина низких темпов прироста объемов данных: такую локальную инфраструктуру довольно сложно масштабировать по емкости, особенно по сравнению с намного более гибкой облачной инфраструктурой.

Также следует отметить, что, хотя на производстве генерируются большие объемы данных с датчиков и прочих приборов, по большей части это происходит на периферии, и сами данные вскоре удаляются и не передаются в центр на долгосрочное хранение.

Производственный сектор также показал самые низкие уровни автоматизации функций управления данными и полной интеграции этих функций на единой платформе (всего 9% по сравнению с 19% и выше во всех прочих отраслях). Возможно, в обоих случаях причина заключается в огромном количестве взаимосвязанных ресурсов, с которыми приходится иметь дело современным производителям.

Как показал отчет Manufacturing Insights компании IDC за 2018 г., посвященный интеграции ИТ и эксплуатационных технологий, почти 80% производственных активов, оборудованных приборами, так или иначе связаны между собой цифровыми подключениями. Но почему же возник этот **разрыв между цифровыми активами и управлением данными** в производственном секторе? Исследование IDC показало, что, помимо уже перечисленных, в этой отрасли существуют еще две крупные проблемы:

1. Серьезная нехватка квалифицированных кадров. В экономике развитых стран главный фактор успеха заводов будущего — это квалифицированные кадры, и их нехватка является одним из главных препятствий в работе предприятий. Трудовые ресурсы, занятые на производстве, стареют,

а новые сотрудники с высокими социальными компетенциями и профессиональными навыками работы в ИТ неохотно идут на заводы.

2. Если на производственном предприятии налажен обмен данными хотя бы с половиной цехового оборудования, это уже можно считать везением. Речь идет не только о приложениях, ЦОД и сетях, но и об архитектуре центра, а также о стратегии развития инфраструктуры: безопасности, интеграции ИТ и эксплуатационных технологий и других вопросах. Во многих случаях устаревшие заводские инфраструктуры просто не выдерживают крупных поступлений новой техники с подключением к сети. В результате на многих заводах приходится придумывать ситуативные решения для подключения и администрирования этой техники, не использующие базовую инфраструктуру и не дающие возможностей комплексного управления.



## ГЛАВА ПЯТАЯ

# DataOps: недостающее звено в управлении данными

Из предыдущих разделов понятно, что перед нами возникает дилемма.

### ТОЧКА ЗРЕНИЯ SEAGATE

Как управлять корпоративными данными в эпоху их взрывного роста так, чтобы использовался весь их потенциал, а акционеры и клиенты оставались довольны?

Результаты этого опроса показывают, что решение существует, и называется оно DataOps. Возможно, эта методология является тем самым недостающим звеном в управлении корпоративными данными, которое

поможет наладить продуктивную работу с ними на любом предприятии независимо от отрасли.

IDC определяет DataOps как методологию, сводящую вместе создателей данных и их потребителей.

Но DataOps — лишь один из компонентов управления данными, наряду с оркестрацией конечных точек и центра, архитектурой данных

и обеспечением их безопасности. Задача управления данными — обеспечить комплексный взгляд на хранящуюся и передаваемую информацию, а также дать пользователям возможность работать с ней и извлекать из нее максимальную пользу.

### ДААННЫЕ IDC

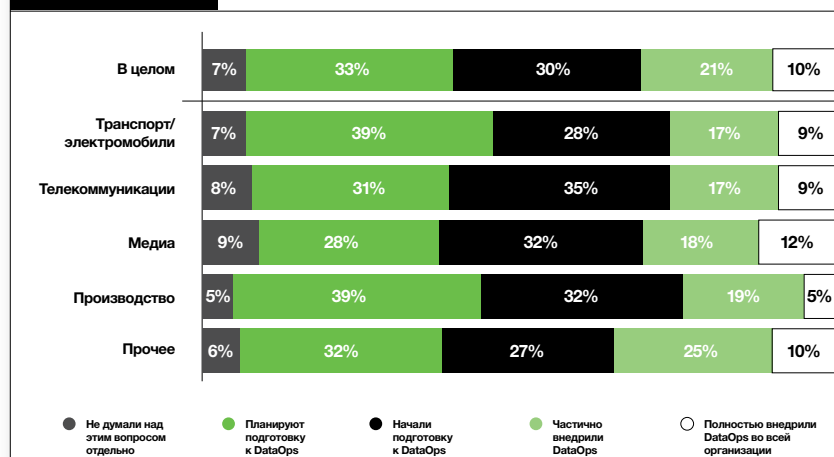
## DataOps

Согласно результатам опроса, в среднем лишь 10% организаций из разных отраслей и регионов полноценно внедрили у себя DataOps. Они уже пользуются этой возможностью.

DataOps — это не технология и не процесс, а скорее новая методология, которая сводит вместе создателей данных и их потребителей, обеспечивая тем самым эффективное сотрудничество между ними и ускоряя внедрение инноваций.

РИС. 13

Состояние DataOps



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



Потребителями данных выступают сотрудники бизнес-подразделений, отвечающие за принятие организационных решений в отношении разработки, дистрибуции и (или) маркетинга продукции, контроля затрат, производственной деятельности и т. д. Чаще всего это высшее руководство (генеральные директора, вице-президенты и т. п.) и сотрудники, помогающие ему в работе. Потребителям нужны не собственно данные, а результаты их анализа, на основании которых можно принимать взвешенные решения.

Создателями данных могут быть как машины (конечные точки, IoT-устройства), так и люди, готовящие отчеты и информацию для тех, кто принимает решения. При создании данных часто возникает вопрос, какую информацию использовать немедленно, а какую отправлять на хранение. Например, сведения о состоянии устройства (активно ли оно и т. д.) не всегда нужно обрабатывать сразу, поэтому они часто передаются на долгосрочное

архивное хранение. Однако сведения о работе того же устройства (температуре, емкости, скорости и т. д.) могут требовать незамедлительного анализа или координации, поскольку они помогают делать более точные прогнозы, устанавливать взаимосвязи между событиями и т. д.

Методология DataOps позволяет использовать такие технологии, как ИИ и МО, для поиска взаимосвязей между данными из центра, облака и с источников данных на периферии. Кроме того, в DataOps для получения данных используется процесс, построенный по принципу ELT (Extract, Load, Transform — «извлечение, загрузка, преобразование»). Он извлекает данные из нескольких разных источников и загружает в единую структуру, часто в виде озера данных. А ИИ может превращать эту массу необработанных данных в четкую полезную информацию, на основании которой можно принимать взвешенные решения.

Другие технологии не дают возможности легко устанавливать

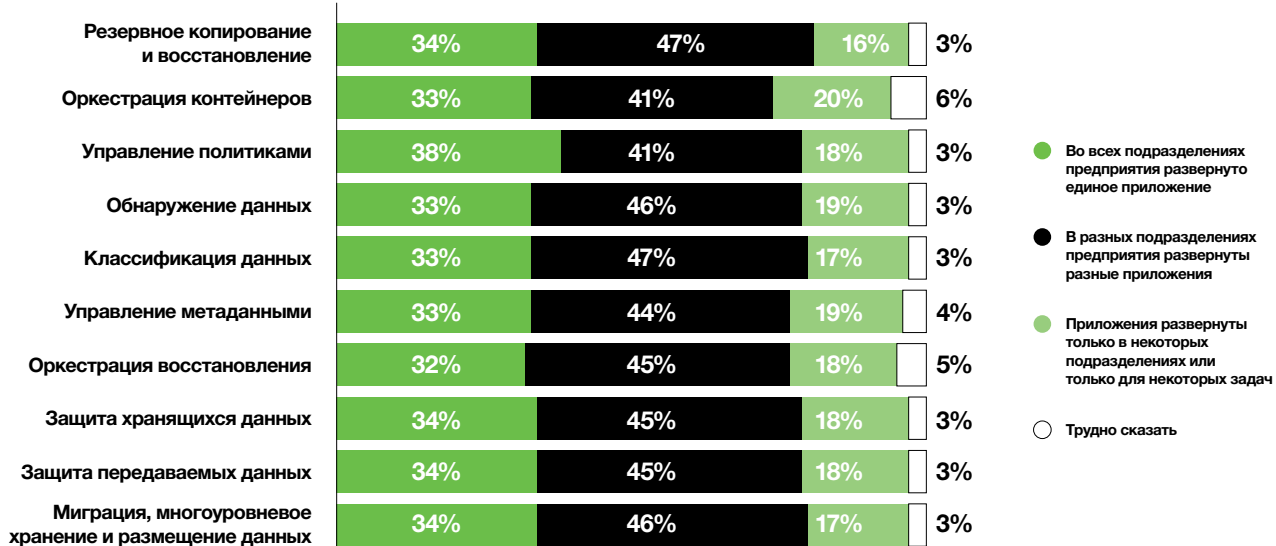
взаимосвязи между данными из разрозненных источников. Это непростая задача, и организация, которая сумеет ее решить, получит конкурентное преимущество.

Опрос также показал, что в большинстве организаций для сходных задач часто используется по несколько разных инструментов. Это заметно усложняет управление данными в масштабах предприятия.

Лишь около трети организаций используют единое решение. Это происходит по разным причинам. Иногда за закупку инструментов в пределах одной организации отвечают несколько разных человек, иногда решения несовместимы с определенными платформами, иногда старые решения используются наряду с новыми.

РИС. 14.

## Подходы к развертыванию инструментов и приложений для управления данными



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



Хотя на системном уровне действительно иногда приходится использовать специализированные продукты, централизованная система управления данными жизненно необходима для любого предприятия. И тут на помощь приходит методология DataOps, позволяющая объединить разрозненные информационные системы в структуру, поддающуюся анализу и контролю. Главные функции, необходимые для DataOps, — управление метаданными и политиками, а также классификация данных. Вот как они работают. За счет управления метаданными сопоставление данных и управление ими происходят на основании

характеристик этих данных. Функция классификации данных в сочетании с метаданными позволяет относить информацию к тому или иному типу (например, «персональные данные» или «медицинские сведения»). После этого можно разработать алгоритмы ИИ, которые будут автоматически распознавать и сопоставлять классифицированные данные.

Методология DataOps очень хорошо сочетается с подходом итеративного обучения, который используется в приложениях на основе ИИ. Этот подход прямо противоположен традиционной аналитике данных, в которой сначала берется задача, а затем для нее ищется решение. DataOps сопоставляет данные,

анализирует их и делает ценные заключения. Например, так владелец магазина может узнать, что клиенты часто покупают вместе, казалось бы, совершенно не связанные друг с другом продукты, и изменить выкладку товаров так, чтобы увеличить эти продажи. Таким же образом можно выявлять тенденции, характерные для определенных демографических групп, и успешно вести микромаркетинг в этих нишах.

#### ТОЧКА ЗРЕНИЯ SEAGATE

## DataOps: человеческий фактор

Эффективному использованию DataOps не всегда мешают *только* технические проблемы. Культурные особенности и человеческий фактор тоже играют свою роль, и каждый владелец компании должен их учитывать.

Очень часто организации страдают от разобщенности: отдельные команды соперничают и добиваются собственных целей, вместо того чтобы сотрудничать и обмениваться полезной информацией. При этом конкурирующие группы борются за контроль над данными, рассматривая их как источник власти.

В результате корпоративные данные (порой одни и те же) хранятся, контролируются и анализируются не централизованно, а несколькими разными группами в пределах компании.

Часто результаты такого анализа не совпадают, потому что для получения максимально точной и полной картины нужно использовать глобальный

репозиторий данных, а при организационной разобщенности это невозможно.

Чтобы решить эту проблему, нужно начинать со стратегии владельца предприятия. В ней должны закрепляться глобальные принципы работы: стандарты, архитектура данных, управление данными, общий доступ к общим аналитическим инструментам для всех групп.

Если передать функции формирования отчетов ИТ-отделу, можно создать централизованную инфраструктуру с инструментами, решениями и возможностями, с которыми смогут работать все группы в организации. Так можно победить разобщенность в работе с корпоративными данными.

В итоге все команды смогут принимать взвешенные решения на основании результатов анализа надежных и доступных всем в компании данных из глобальных пулов.

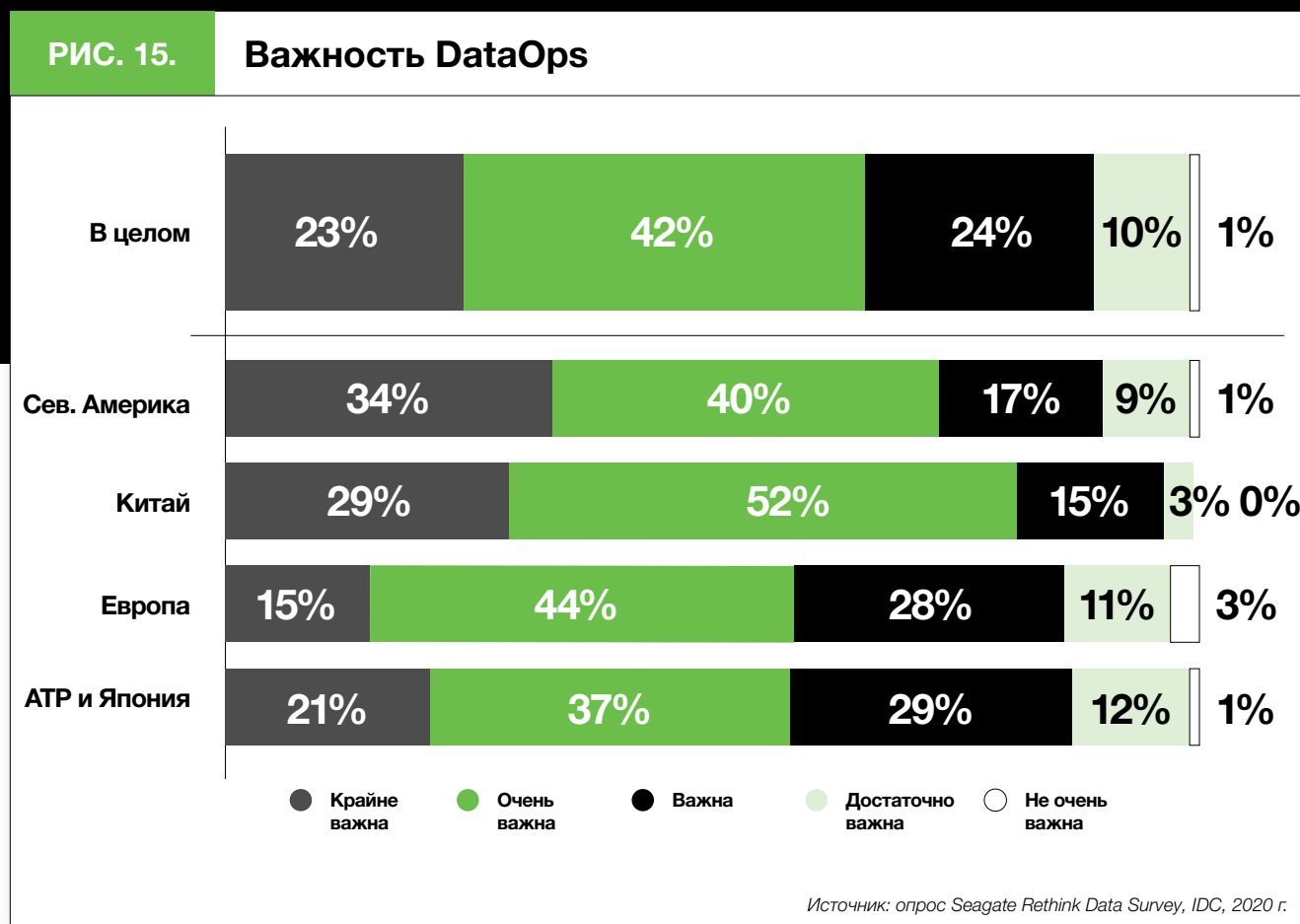


## Руководители предприятий по всему миру поддерживают идею внедрения DataOps.

- Большинство респондентов считают, что методология DataOps «очень важна» или «крайне важна» для бизнеса.
- Наибольшее значение DataOps придают в Северной Америке и Китае (анализ по регионам см. в разделе II).

- Важность DataOps признают во всех отраслях, однако в транспортной сфере потребность в этой методологии несколько выше.
- Опрос проводился до начала пандемии COVID-19. Можно предположить, что из-за нее потребность в DataOps возрастет

еще больше, поскольку после пандемии многие сотрудники продолжают работать удаленно. Из-за этой тенденции уже ускорилась миграция в облачные сервисы.



## ГЛАВА ШЕСТАЯ

# Улучшение бизнес-результатов

## Наконец-то самые приятные новости для владельцев бизнеса.

### ТОЧКА ЗРЕНИЯ SEAGATE

Результаты опроса четко показывают, что многие проблемы, связанные с управлением данными, можно решить внедрением DataOps — методологии, которая сводит вместе создателей и потребителей данных, — и соответствующих процессов. Откуда взялся вывод о пользе DataOps? На это указывают многие участники опроса.

В сочетании с прочими решениями для управления данными (оркестрация данных с использованием аналитики, хорошая функциональная архитектура данных) DataOps может заметно улучшить бизнес-результаты предприятий. Повышение лояльности и удовлетворенности клиентов, рост дохода и прибыли, снижение

текучести кадров, повышение производительности труда сотрудников — всего этого можно добиться с помощью DataOps.

Таким образом, DataOps не просто приносит пользу предприятию, но и повышает его конкурентоспособность.



## Конкурентное преимущество

Правильно используя DataOps, можно быстро создавать и обучать модели ИИ, а также внедрять аналитику данных в крупных масштабах. Благодаря результатам этой аналитики предприятия могут опережать своих конкурентов.

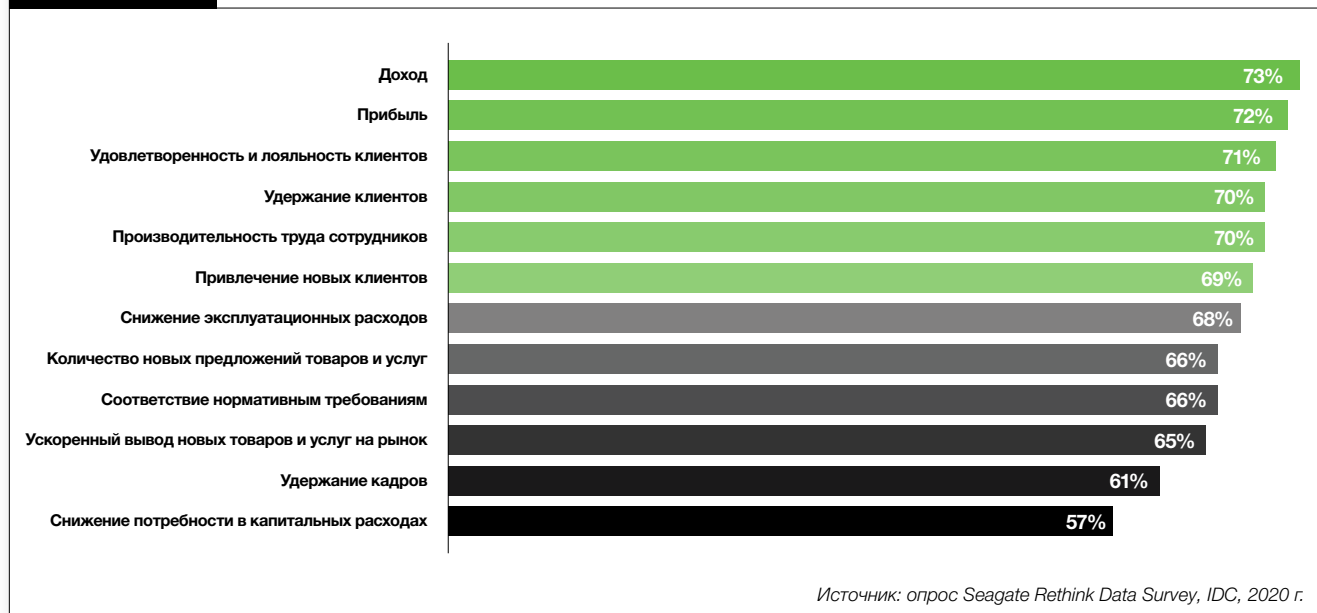
По данным опроса, чем качественнее аналитика данных на предприятии, тем лучше оно работает. И это не личные ощущения респондентов, а конкретные цифры.

Мы проанализировали показатели организаций, которые заявили о внедрении передовых средств аналитики. На диаграмме ниже показаны области, в которых улучшилась работа этих предприятий. Следует отметить, что большое значение имеет скорость улучшения показателей — ведь о постепенном улучшении может заявить большинство организаций.

Но чем быстрее компания оптимизирует свою работу, тем выше ее результаты по сравнению с конкурентами. Результаты, показанные на рис. 16, обусловлены не только внедрением DataOps, однако DataOps остается новейшим и самым передовым методом оптимизации управления данными.

РИС. 16.

### Показатели, свидетельствующие о положительном влиянии аналитики и управления данными на работу организации



Для большинства организаций «улучшение бизнес-результатов» — это три фактора:

1. Рост дохода.
2. Рост прибыли.
3. Повышение удовлетворенности и лояльности клиентов.

Но после внедрения аналитики и управления данными работа многих

предприятий улучшилась и в ряде других аспектов:

- производительности труда сотрудников;
- удержании кадров;
- снижении затрат;
- соответствии нормативным требованиям;
- привлечении новых клиентов.

А если говорить о перспективе, то более эффективное привлечение новых клиентов обещает и повышение дохода.



## DataOps: как данные помогают повысить прибыль и степень удовлетворенности клиентов

Опрос показал, что правильное использование DataOps — главное условие эффективного управления данными, которое, в свою очередь, позволяет получать от них больше выгоды и заметно улучшать бизнес-результаты предприятия — в частности, повышать прибыль и степень удовлетворенности клиентов.

Но как этого добиться?

Как мы уже говорили, нельзя недооценивать человеческий фактор. Разобщенность — это проблема, которая создается коллективом.

Поэтому для эффективного использования DataOps недостаточно просто внедрить необходимые инструменты. Конечно, они очень важны. Так, **инструменты виртуализации** особенно полезны, потому что с их помощью можно извлекать данные и манипулировать ими. Для этого можно использовать и другие средства, но слои виртуализации значительно упрощают управление данными. Можно использовать программное обеспечение (например, Kubernetes — «портативную расширяемую платформу с открытым исходным кодом для управления контейнеризованными рабочими нагрузками и сервисами, которая облегчает как декларативную настройку, так и автоматизацию») или виртуальные машины — механизмы абстракции для развертывания приложений.

Но прежде чем пользоваться благами виртуализации, нужно выбрать правильную стратегию работы с данными. Внедрить инструменты и начать их использовать довольно просто. Куда сложнее задача руководства предприятия: пользуясь средствами и процессами управления данными, принимать решения относительно этих данных.

Начинать нужно с целей владельцев предприятия. Обычно они включают повышение прибыли и степени удовлетворенности клиентов. А для этого необходимо **использовать результаты анализа данных**. При этом необходимо координировать данные из разных источников и решать проблемы, вытекающие из следующих вопросов:

- Кто имеет доступ к каким данным?
- Как классифицировать данные?
- Какие данные хранить и где?
- Что делать с данными после анализа?
- Как сделать данные доступными?
- Как установить взаимосвязи между данными?

Чтобы найти ответы на эти вопросы, владельцы компаний, администраторы данных и директора по информационным технологиям консультируются с профильными экспертами. **Совместно с этими экспертами они определяют, оценивают и проверяют данные, а также находят и исправляют записи с ошибками**. Привлекать профильных экспертов нужно обязательно, потому что в некоторых типах данных досконально разбираются только они.

Только при участии всех ключевых игроков — профильных экспертов, владельцев предприятий, администраторов данных — можно найти ответ на главный вопрос: **что вы хотите получить от тех или иных данных?** Что вы хотите узнать? Как вы хотите использовать эту информацию?

Всего лишь для одного продукта можно собирать данные приблизительно по 10 000 отправных параметров. Если вас интересуют они все и вы хотите хранить эти данные, без четкой архитектуры хранения и передачи между средами эта информация рискует

превратиться в непроходимое болото. Поэтому тем, кто принимает решения, стоит спросить у инженеров-конструкторов и инженеров ОТК, какие из этих тысяч параметров самые важные. После этого отобранные данные можно будет отслеживать и анализировать, чтобы затем оптимизировать создание компонентов и решений на основании полученных результатов.

На первых порах внедрения системы управления данными требуется довольно много работы: координация, обсуждение, анализ, согласование формулировок и классификация данных (подробнее об этом читайте в следующей главе). Большую часть последующей работы можно оптимизировать и автоматизировать с помощью инструментов виртуализации.

После этих двух этапов внедрения DataOps удовлетворенность клиентов обычно повышается: с оптимизацией администрирования и передачи данных часто растет и качество товаров (услуг), а значит, клиенты остаются довольны. Как отмечали в прошлой главе аналитики IDC, скорость доставки информации тоже имеет значение. Чем быстрее получен результат, тем больше довольны клиенты, ведь ускоренный доступ к данным означает более быстрое принятие решений и для владельцев предприятия, и для клиентов.

Вот так можно повысить прибыль и удовлетворенность клиентов, оптимизируя работу с данными посредством DataOps.

**Самое главное — целенаправленно работать над тем, чтобы извлечь из данных максимум пользы.**



## ГЛАВА СЕДЬМАЯ

# Безопасность данных и управление ими

**Помимо DataOps, отдельного внимания заслуживает такая составляющая управления данными, как обеспечение их безопасности.**

### ДАННЫЕ IDC

Две трети участников исследования отмечают недостаточную безопасность данных, что делает этот вопрос ключевым, когда речь заходит об эффективном управлении данными.

Руководители ИТ-отделов и компаний в равной степени считают безопасность данных наивысшим приоритетом. Утечка данных ведет ко множеству проблем: прямым убыткам, крупным штрафам, потере лица, репутации и клиентов. Хакерские атаки могут привести к краже корпоративных секретов, снижению производительности труда сотрудников, невозможности утере данных, а в случае атаки с целью вымогательства — к прямым убыткам и потере лица.

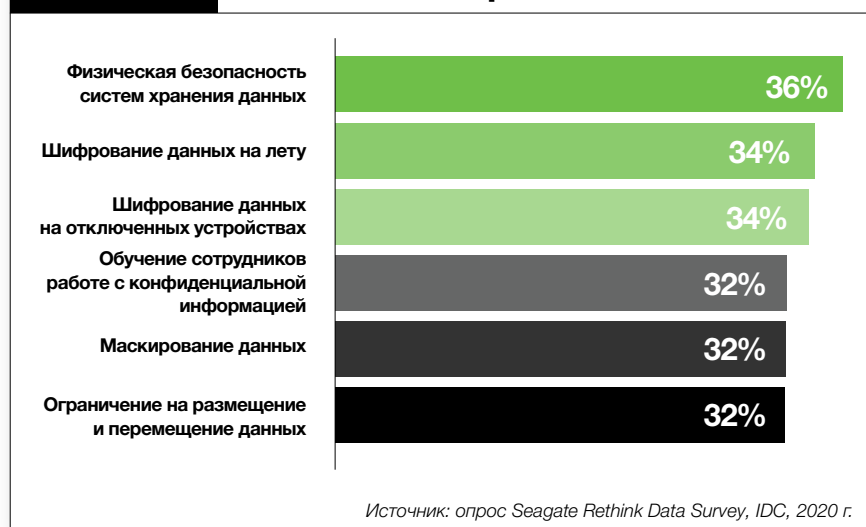
При этом во множестве организаций респондентов на корпоративном уровне не внедрены даже обычные практики защиты данных (рис. 17).

В ответах на вопросы, связанные с безопасностью предприятия (рис. 17), не применялся принцип «все или ничего». То есть, например,

респондент мог ответить «да» на вопрос о шифровании данных на отключенных устройствах и «нет» на все остальные. Таким образом,

весьма вероятно, что большая часть организаций имеет ряд существенных уязвимостей в системе безопасности.

**РИС. 17. Процент организаций, внедривших полный спектр основных мер безопасности**



## Защита данных: человеческий фактор

Глубочайшим заблуждением было бы думать, что основные трудности обеспечения безопасности данных предприятия связаны с технической стороной вопроса. Это не так.

Главная сложность — должным образом классифицировать данные по степени и видам рисков, а также по способам хранения и защиты информации. То есть, как и в случае с DataOps, — человеческий фактор.

Вопросы технологий сводятся к решениям о покупке тех или иных средств, принимаемым владельцами бизнеса. Хотите защитить данные на отключенных устройствах? Приобретите диски с технологией

самошифрования — и проблема решена.

Обучение пользователей данных (тех, кто их создает, анализирует, передает и обрабатывает) внутри компании — это другой вопрос. Руководителям следует помнить: чем выше компетенции создателей и потребителей данных внутри компании, тем лучше для бизнеса. Ключом к успеху является обучение владельцев бизнеса в плане безопасности данных. Именно так можно обеспечить поддержку с их стороны.

Обучение безопасности данных должно быть

демократичным. Защита данных (и соответствующее обучение) должны быть повсеместными. Это вовсе не исключительная прерогатива ответственных за информационную безопасность, отделов информационных технологий, администраторов, службы безопасности и юристов. Владельцы данных также должны участвовать в их защите.

Резюме: владельцам бизнеса не следует дожидаться утечки или утери данных, чтобы убедиться, что их безопасность важна для предприятия.



## Ключевые этапы обеспечения безопасности данных

**1. Классификация данных.** Это трудная, но необходимая задача, требующая активного взаимодействия внутри организации. Создатели данных, их владельцы и пользователи внутри организации должны придерживаться единых стандартов классификации информации и распределять данные в соответствии с ней. Без этого этапа определения и разработки единых стандартов никакие программы защиты данных не дадут результата. Только не усложняйте.

Например, сотрудники Seagate Technology разработали достаточно широкие стандарты. В результате получились четыре типа информации: с ограниченным доступом, конфиденциальная, внутренняя и открытая.

**2. Потоки данных.** Владельцам бизнеса необходимо понимать движение данных в компании: происходит ли оно согласно плану или нет. Решить, что правильно и обоснованно — важный шаг в определении мер контроля. Зная, как движутся потоки данных, легче выявить зоны наибольшего риска.

**3. Контроль доступа.** В качестве базовой формы контроля обычно используется управление доступом на основе ролей, но, как правило, этого недостаточно. Понятно, что доступ к данным должны иметь только те, кому он необходим, но реализовать этот принцип не так-то просто. Чем чувствительнее и уязвимее данные (в соответствии с их классификацией), тем строже должна быть реализация. Помимо управления доступом на основе ролей можно рассмотреть следующие меры:

- детальное управление правами на доступ к данным при шифровании на уровне файлов (ограничение функций печати, редактирования, копирования и вставки для конкретного объекта);
- другие механизмы доступа — например, доступ только через браузер (без скачивания) и нанесение цифровых водяных знаков для предотвращения записи видео и снимков экрана.



## Основы безопасности данных

Реализация продуманной программы защиты данных может выглядеть пугающе. Кроме того, важно не забывать, что лучшее — враг хорошего. Предположим, вы — владелец бизнеса. У вас полно дел, времени не хватает, но нужно принять хотя бы самые необходимые меры, а о менее срочных проблемах можно будет подумать позднее.

Вот список действий, которые защитят данные от наиболее серьезных угроз:

- 1. Шифруйте данные на лету.** Как минимум пользуйтесь только безопасными протоколами и сервисами: HTTPS вместо HTTP, SFTP вместо FTP, IPSec или SSL VPN для удаленного доступа. Это не так уж сложно: пользуйтесь только инструментами, которые используют защищенные соединения, правильно настройте их, выработайте стандарты и обяжите всех соблюдать их.
- 2. Шифрование данных на отключенных устройствах.** Ноутбуки и мобильные устройства представляют наибольший риск с точки зрения безопасности данных. Массивы и серверы хранения данных, как правило, располагаются в центре обработки с ограниченным доступом. Их вряд ли можно украсть из машины или выронить из кармана в такси. Сначала займитесь самыми серьезными угрозами. Введите полное шифрование дисков для ноутбуков на всем предприятии. Включите шифрование данных на мобильных устройствах на уровне файловой системы в политику управления такими устройствами.
- 3. Обучайте пользователей.** Как неоднократно подчеркивается в этом отчете, пользователи могут оказаться как самым сильным, так и самым слабым звеном. Владельцам бизнеса необходимо быть уверенными в том, что все их сотрудники прошли соответствующее обучение, осознают связанные с данными риски и следуют всем советам и рекомендациям. Остановитесь. Подумайте. Защитите.





## РАЗДЕЛ II

# Результаты по регионам



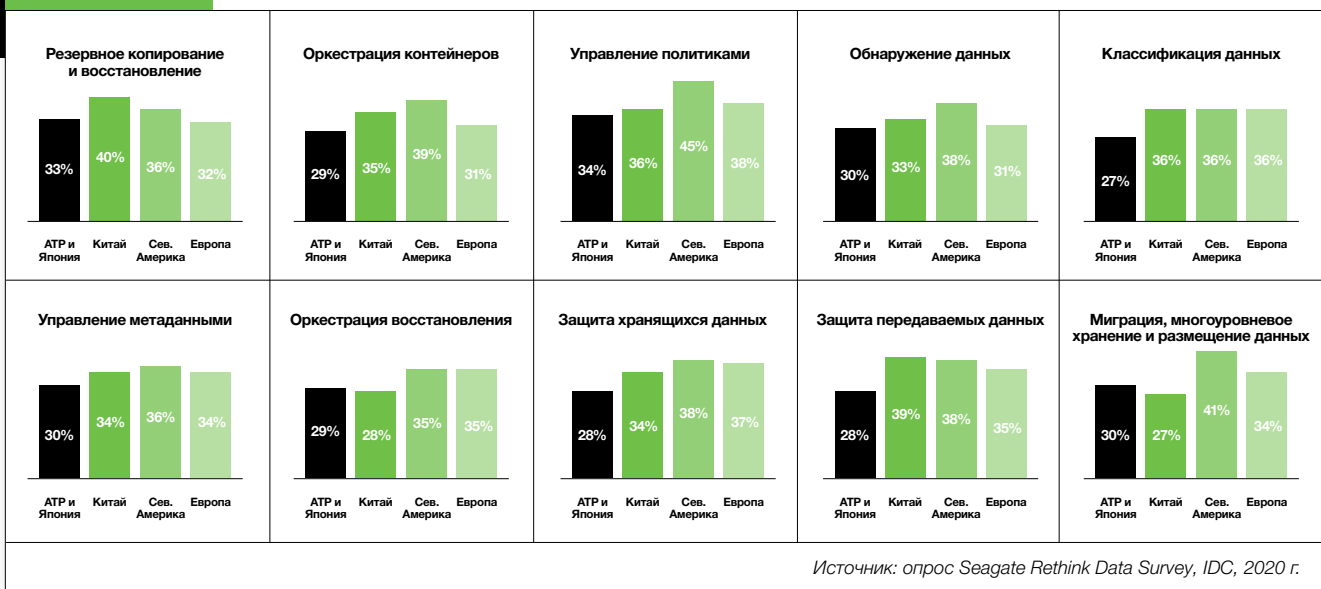
## ГЛАВА ПЕРВАЯ

# Азиатско-Тихоокеанский регион (АТР) и Япония

В число респондентов из стран АТР и Японии вошли представители Австралии, Японии, Индии, Южной Кореи и Тайваня. В рамках данного исследования Китай рассматривается отдельно и не входит в эту категорию.

РИС. 1.

Доля респондентов, отметивших наличие в своих организациях единой для всего предприятия системы управления политиками



### ДАННЫЕ IDC

Характеризовать АТР и Японию как единый регион чрезвычайно сложно. Технологический прогресс идет в этом разноплановом регионе неоднородно. С учетом этого в исследовании рассматриваются пять наиболее технологически

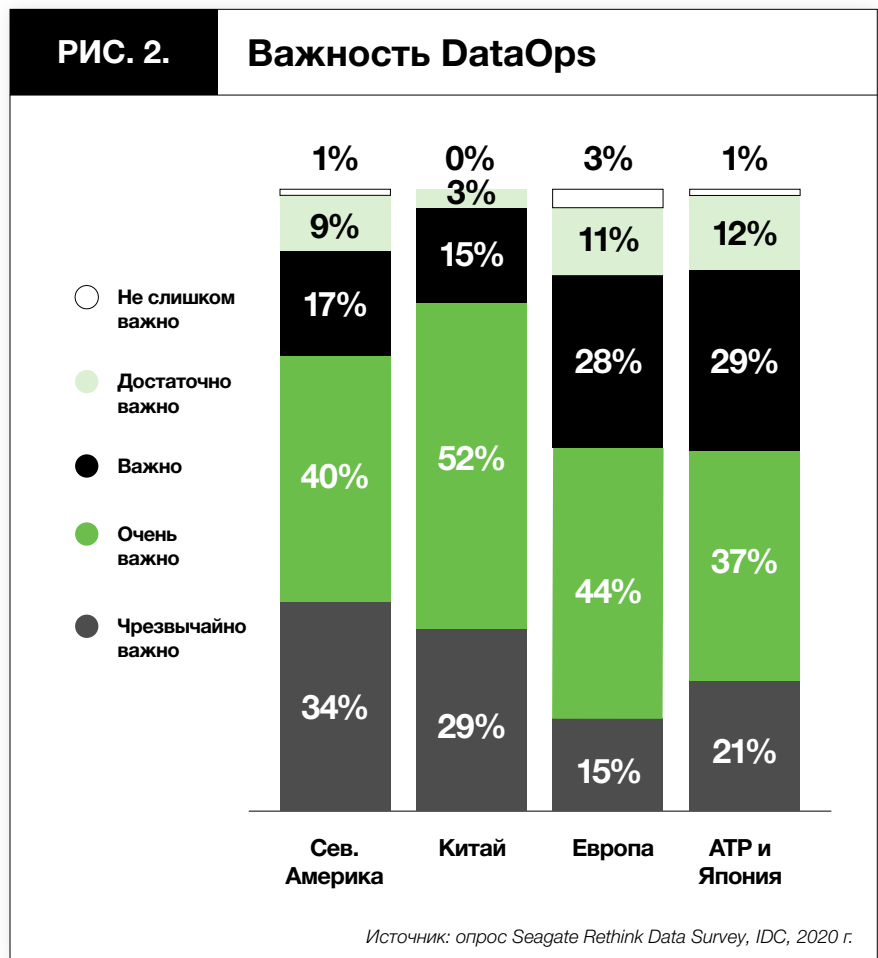
развитых стран региона. Даже с такой выборкой будет справедливо заключить, что интеграция функций управления данными в регионе все еще развивается.

Следующие результаты опроса выделяют регион из числа прочих:

- Страны АТР и Япония следуют за Северной Америкой и Европой: у 34% организаций имеется единая для всего предприятия система управления политиками (рис. 1).



- 87% респондентов отмечают DataOps «чрезвычайно важным», «очень важным» или «важным» (рис. 2).
- В ближайшие два года страны АТР и Японию ожидает рост объемов информации в результате распространения усовершенствованных средств анализа данных и автоматически собирающих данные IoT-устройств. В отличие от представителей других регионов, компании АТР и Японии не ожидают большого роста в результате миграции в облачные среды (рис. 3).
- Организации в странах АТР и Японии менее остальных опасаются проблем, связанных с управлением данными, в ближайшие пару лет (рис. 4).
- Корпоративные данные в странах АТР и Японии не показывают большой динамики. Здесь данные мигрируют не так часто, как в остальных регионах. Тем не менее ожидается, что в течение следующих двух лет 7% корпоративных данных компаний АТР и Японии переместятся в облако (рис. 5).



## Что это означает

В странах АТР и Японии отсутствует единое для региона законодательство или регулирование, которое бы ограничивало перемещение данных через границы — в отличие от, например, европейского регламента по защите данных (GDPR).

Суверенности данных там уделяется меньше внимания — по крайней мере, к ней относятся иначе. Единый механизм применения политики — самый эффективный способ автоматизированного, согласованного управления данными в масштабе предприятия.

Тот факт, что о нем говорит только четверть респондентов, означает, что большинство предприятий региона не справляются с унификацией обработки данных. Потенциально тормозить движение данных в странах АТР и Японии может языковое и законодательное разнообразие, из-за которых регион так неоднороден и которые вызывают затруднения, связанные с переводом и соблюдением соответствующих законов.

Доля респондентов, отмечающих эти вопросы управления данными как «чрезвычайно сложные» и «сложные» (рис. 4), во всех странах АТР и Японии ниже, чем среди всех участников опроса. Почему респонденты в странах АТР и Японии менее озабочены проблемами, связанными с управлением данными? Мы предполагаем, что здесь играет роль зрелость системы управления данными.

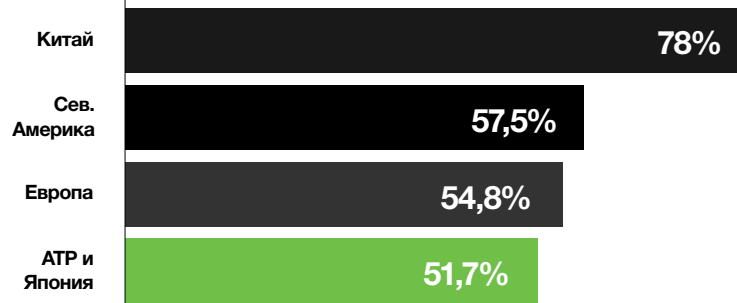
В сравнении с другими регионами страны АТР и Япония демонстрируют наиболее низкий средний уровень удовлетворенности подходом своих организаций к управлению данными. Страны АТР и Япония также показывают самые низкие средние уровни удовлетворенности инструментами управления данными. Низкая удовлетворенность процессами и инструментами управления данными может указывать на их недостаточное развитие. Если респонденты из АТР и Японии не удовлетворены процессами и инструментами, скорее всего, доля организаций в регионе, ставящих перед собой передовые задачи управления данными (например, управление данными в многооблачных и гибридных средах), будет ниже. Возможно, организации в странах АТР и Японии все еще совершенствуют процессы и технологии в своих

стратегиях управления данными и не берутся за более сложные задачи.

При этом именно в странах АТР и Японии (вместе с Китаем) основным фактором роста объемов данных являются аналитика и Интернет вещей (рис. 3). Это может объясняться ростом популярности устройств среди производителей вычислительной техники в регионе.

РИС. 4.

### Доля регионов, ожидающих проблем (в том числе серьезных) в управлении данными в ближайшие два года



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

#### ТОЧКА ЗРЕНИЯ SEAGATE

## Простор для роста и обучения

Хотя управление корпоративными данными в странах АТР и Японии не настолько интегрировано, как в других регионах, руководителям предприятий необходимо убедиться в том, что оно готово к требованиям будущего, например

мировой пандемии, в условиях которой цифровое поведение масс полностью меняется.

Управление корпоративными данными должно: 1) побуждать владельцев бизнеса правильно

классифицировать данные и понимать, каких целей им нужно достичь; 2) побуждать их принимать решения на основании аналитики; 3) подготовить работников к информационному будущему.



Как подчеркивает IDC, единая система управления политиками — наиболее эффективный автоматизированный способ согласованного управления данными в нужном масштабе. Почему же ни единая система управления политиками, ни DataOps не являются таким приоритетом в странах АТР и Японии, как в других регионах?

Во-первых, исторически сложилось, что, за исключением защиты конфиденциальности, власти многих стран АТР не вмешивались в то, как частные компании управляют своей корпоративной информацией.

Во-вторых, компаниям из большинства стран этого региона все еще не хватает сознательного отношения к оптимизации ценности данных в масштабах компании посредством соответствующей информационной политики. Руководители бизнеса должны лучше представлять преимущества единой системы управления политиками в области данных и DataOps и внедрять ее.

В-третьих, ИТ-инфраструктура компании должна поддерживать работу с искусственным интеллектом и, в частности, периферийными устройствами, которые позволяют собирать, анализировать и

использовать данные в режиме реального времени. Важно иметь надежную инфраструктуру, включающую оборудование и ПО, которая может собирать и обрабатывать большие объемы данных, поступающих от автоматических систем их сбора. Владельцы компаний в тесном сотрудничестве с ИТ-отделами должны подбирать наиболее подходящие платформы и размещать их на соответствующих участках. Для многих это означает расположение периферийных устройств по месту размещения заводского оборудования, а не за пределами производства.

В-четвертых, необходимо переучивать сотрудников, чтобы они не теряли эффективности. Все сводится к гибкой работе с данными: организациям предстоит выстроить базу кадров, специализирующихся на использовании больших данных для принятия решений. Это может быть привлечение сторонних специалистов с нужными навыками или удержание и переобучение своих сотрудников. Так ваши сотрудники будут готовы к информационному будущему.

В-пятых, организациям следует рассмотреть долгосрочное инвестирование в подготовку

к такому будущему. Как показала пандемия COVID-19, готовность действовать в долгосрочной перспективе имеет огромное значение.

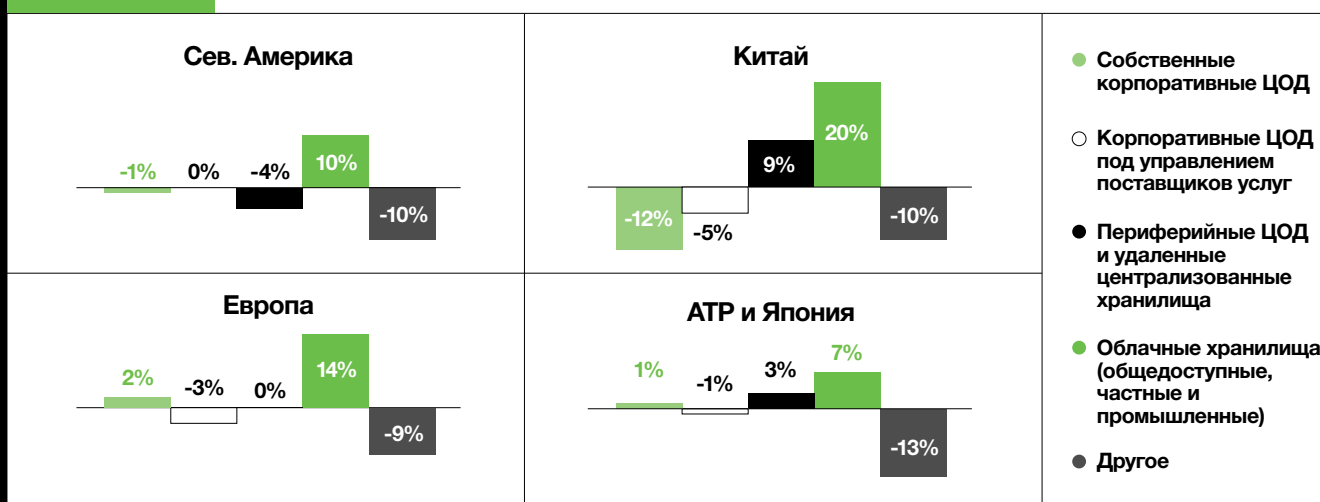
Словом, бизнес в странах АТР и Японии находится в стадии осознания того, как управление данными способно повысить их ценность. Этот регион выглядит немного более консервативным, например когда речь идет о помещении данных в облако или смене места их расположения. Повышение безопасности данных служит самым мощным толчком к развитию в регионе.

Существует определенное нежелание (связанное либо со средствами, либо с нехваткой знаний) выйти из зоны комфорта в процессах, связанных с данными: **компании в АТР и Японии могут инвестировать в DataOps**, чтобы извлекать максимальную выгоду из накапливающихся данных.

В результате глобальной пандемии, когда ряды сотрудников, работающих из дома, значительно пополнились, предприятиям необходимо относиться к управлению данными еще серьезнее.

РИС. 5.

## Ожидаемые изменения в подходах к хранению данных: прогноз на ближайшие два года



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

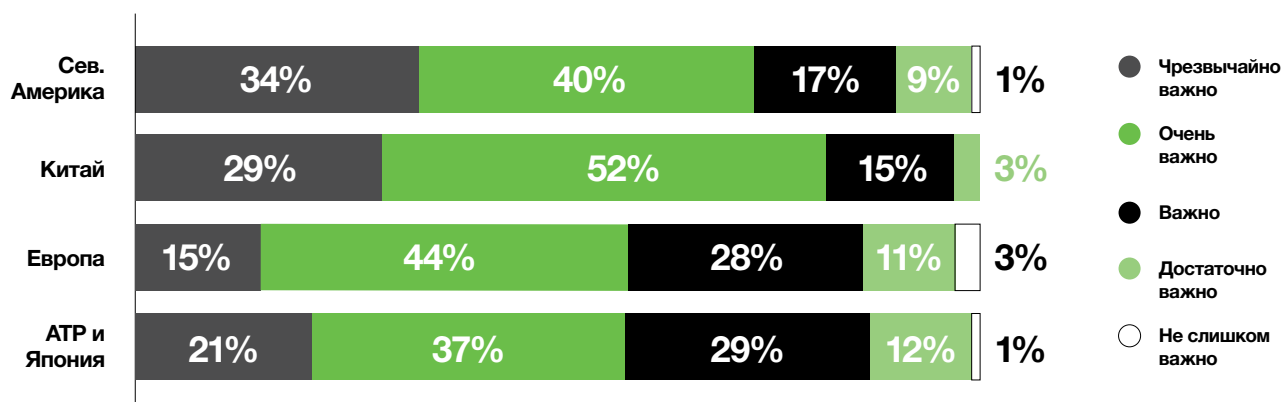


## ГЛАВА ВТОРАЯ

# Китай

В рамках настоящего исследования Китай рассматривается в качестве отдельного региона. Это имеет смысл из-за обширности, гомогенности и геополитического своеобразия данной страны.

РИС. 6. Важность DataOps



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

### ДАНЫЕ IDC

Исследование показывает, что Китай является самым прогрессивным регионом в вопросах управления данными.

- Из всех регионов именно в Китае DataOps называют «очень важным» или «чрезвычайно важным». На вопрос, насколько, по их мнению, важна концепция DataOps, 52% респондентов ответили «очень

важна» и 29% — «чрезвычайно важна», что в сумме дает 81% (рис. 6).

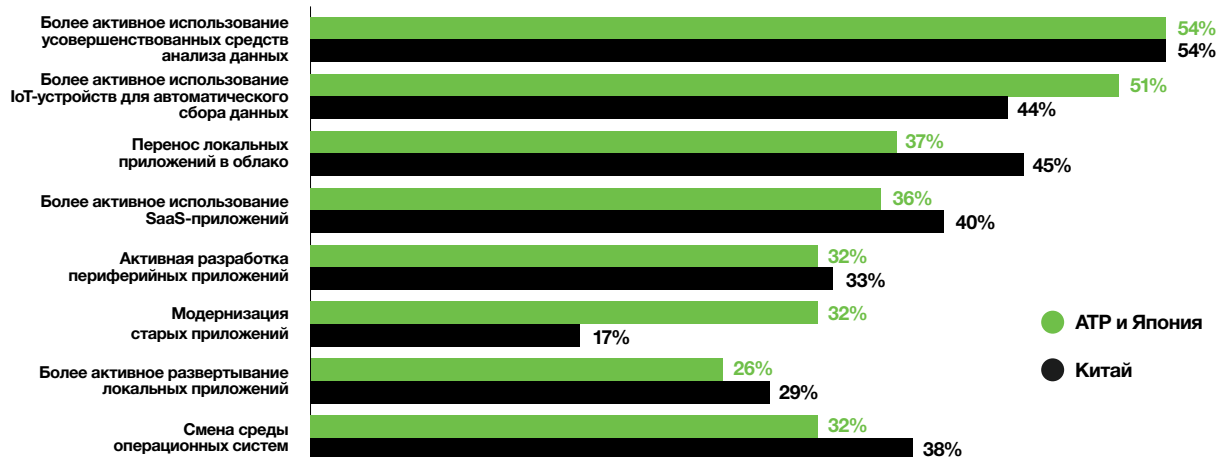
- Что касается факторов, влияющих на рост объемов данных, самым важным из них представители Китая считают аналитику данных (54% респондентов вместе с АТР и Японией), а наименее важным — модернизацию старых приложений (рис. 7).

- Респонденты ожидают, что в ближайшие два года еще 20% данных предприятий будут перемещены в облачные хранилища (самый высокий процент из всех регионов), а еще 9% переместятся в периферийные центры обработки данных (рис. 8).



**РИС. 7.**

**Факторы, влияющие на рост объемов данных**

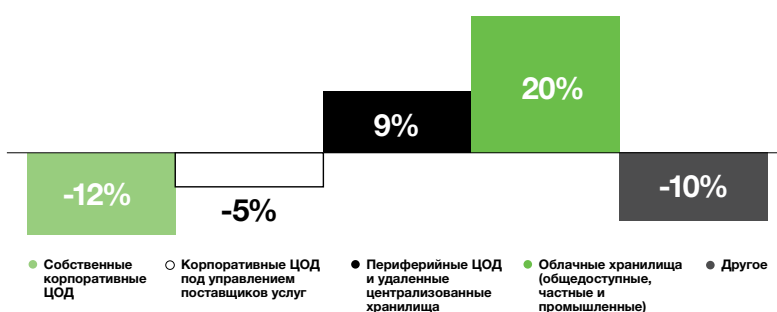


Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

- Управление корпоративными данными в Китае более централизовано по сравнению с остальными регионами. Существенная часть (67%) предприятий Китая имеют лишь одну выделенную централизованную группу, исполняющую функции управления данными (рис. 9).
- Китай отстает от других регионов в том, что касается интеграции управления данными, но он же и наиболее заинтересован в ней (рис. 10).
- Китай демонстрирует самый высокий уровень проблем в использовании полного потенциала данных, включая приведение собранной информации в пригодный к использованию вид, управление хранением собранных данных, обеспечение их безопасности и доступа к различным системам хранения (рис. 11).

**РИС. 8.**

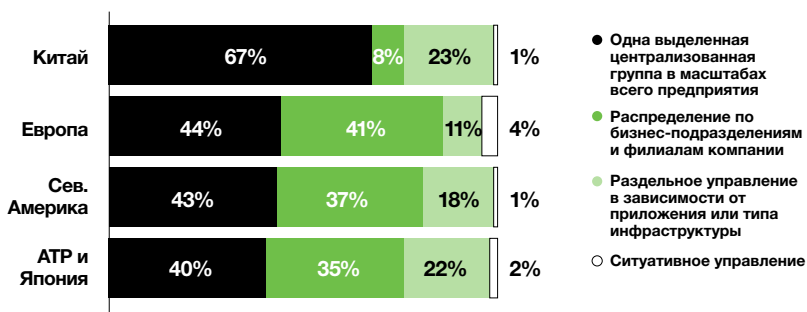
**Ожидаемые изменения в подходах к хранению данных: прогноз на ближайшие два года**



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

**РИС. 9.**

**Организация управления данными**



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



## Что это означает

Возможности организаций Китая в управлении данными быстро развиваются. Предприятия понимают важность обработки данных.

Централизованное управление данными может позволить извлекать из данных больше пользы за счет сокращения количества систем хранения. При том, что централизация управления политиками очень важна, интеграция функций управления данными также имеет большое значение как следующий уровень использования информации.

Китай не настолько преуспел в интеграции данных, но респонденты видят ее в качестве ключевой цели на ближайшие два года.

Несколько крупных поставщиков облачных услуг размещаются в Китае, так что местный рынок готов к принятию облачных сред. Возможно, поэтому регион показывает самый высокий уровень миграции в облака (общедоступные, промышленные и частные). Большая часть этих данных генерируется за счет распространения IoT-устройств, датчиков и алгоритмов ИИ.

Растущее принятие услуги IaaS в общедоступных облаках, предлагаемой региональными поставщиками, также стимулирует переход в облачные среды.

Китай больше всего стремится к повышению доступности данных и их аналитике и менее всего — к сокращению издержек. Благодаря зрелости ИТ-системы Китая мы считаем, что этот регион располагает меньшим количеством старых приложений, что является его конкурентным преимуществом.

РИС. 10.

### Интеграция управления данными: текущая ситуация и перспектива на ближайшие два года



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



## Данные в движении

В Китае данные перемещаются из локальной среды в облако и на периферию.

Китайская экономика находится в стадии активных инвестиций. Большинство китайских компаний сосредотачиваются не на текущих прибылях и убытках, а на долгосрочной прибыли. Это значит, что они инвестируют в развитие, чтобы получить существенную выгоду в будущем.

Китайским предприятиям не приходится заботиться об обновлении традиционных приложений, спроектированных и разработанных в клиент-серверную эпоху, которые обычно работают локально и с одним клиентом. (К примерам можно отнести оборудование на электростанциях, промышленные станки под управлением компьютеров с MS-DOS или устаревшие финансовые системы.) В результате они располагают свободным бюджетом для инвестиций в перемещение данных (в облако

и периферийную среду) и могут раскрывать потенциал данных благодаря оптимизированным процессам DataOps.

Китайские респонденты сообщают о больших сложностях, связанных с освоением потенциала данных: это может быть отражением высоких требований китайских компаний и их постоянного стремления к развитию, а также того факта, что в настоящее время они находятся в активной фазе внедрения DataOps.

Возможно, причина в том, что Китай отстает в интеграции функций управления данными из-за централизации, а также из-за отсутствия необходимости в интеграции традиционных приложений. Китай представляет собой развивающуюся экономику, которая строго регулируется государством и законодательством, в отличие от экономик многих западных стран.

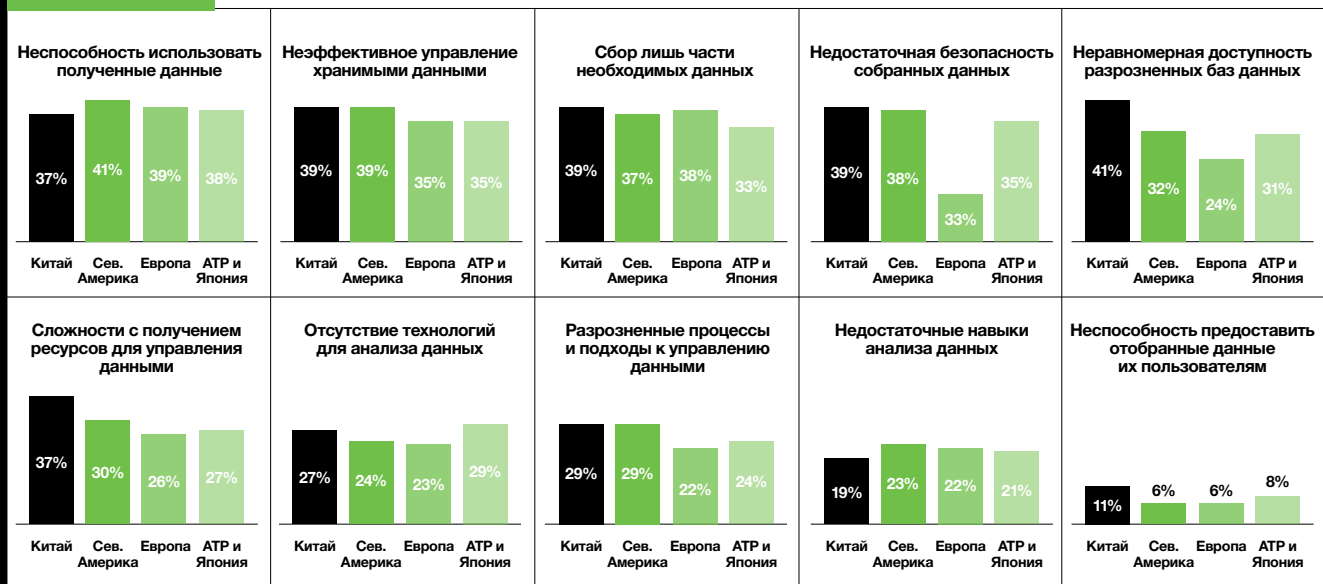
В Китае очень большое внимание уделяется искусственному интеллекту, в том числе и из-за

огромных объемов данных, создаваемых населением, численность которого составляет около 1,4 миллиарда человек. Искусственный интеллект считается одной из ведущих инициатив, которую Китай использует для успешной конкуренции в мировом масштабе. Правительство страны поддерживает разработку решений ИИ, 5G, гипермасштабируемых центров обработки данных и других средств ускорения перемещения данных.

Это значительно повышает гибкость инфраструктуры. Китай начал реализацию проекта по созданию новой инфраструктуры, чтобы стимулировать цифровую экономику. В сочетании с развитием технологий эта инициатива может претвориться в новые бизнес-возможности для управления корпоративными данными.

РИС. 11.

### Основные проблемы в раскрытии потенциала данных



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.



## ГЛАВА ТРЕТЬЯ

# Европа

В рамках исследования были опрошены представители предприятий из четырех европейских стран: Великобритании, Франции, Германии и России.

РИС. 12. Состояние DataOps в организации



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

### ДААННЫЕ IDC

Изучая данные по европейскому региону, можно заметить, что в ответах на почти все общемировые опросы прослеживается более консервативная позиция по сравнению с другими регионами. В опросе приняли участие три западноевропейские страны и Россия.

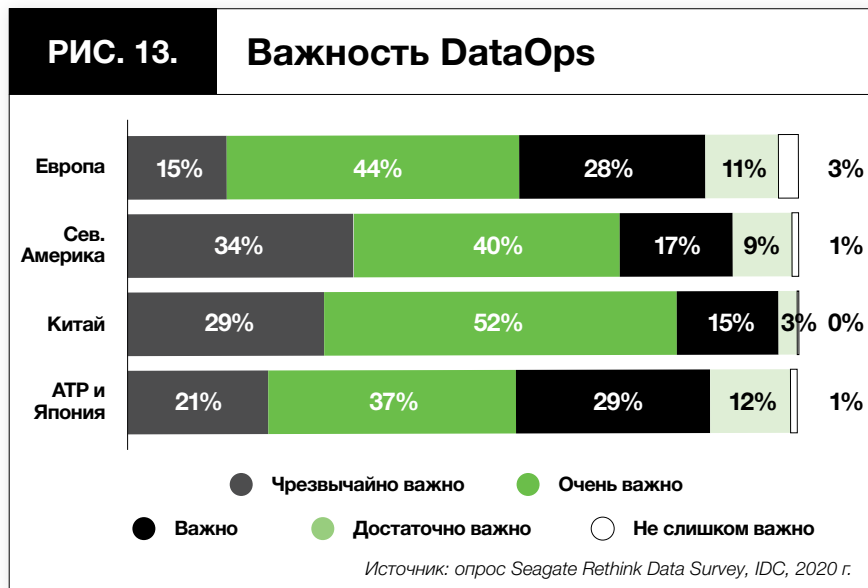
- В Европе по сравнению с другими регионами практика внедрения DataOps распространена в наименьшей степени. Только 18% респондентов в регионе заявляют о полной или частичной реализации DataOps (рис. 12).
- Однако европейские компании осознают потребность в DataOps.

86% респондентов признают внедрение DataOps «чрезвычайно важным», «очень важным» или «важным» (рис. 13).

- Скорость роста объемов данных здесь также наименьшая по сравнению с другими регионами (рис. 14).



- Европейские организации переносят данные в облако. В течение следующих двух лет в облако будет перемещено 14% корпоративных данных (рис. 15). К факторам, влияющим на рост объемов данных, относятся миграция локальных приложений в облако и увеличение количества развертываний локальных приложений (рис. 16).
- Касательно факторов, определяющих способы хранения данных, в Европе наблюдается стремление к повышению безопасности и сокращению затрат на хранение (рис. 17).

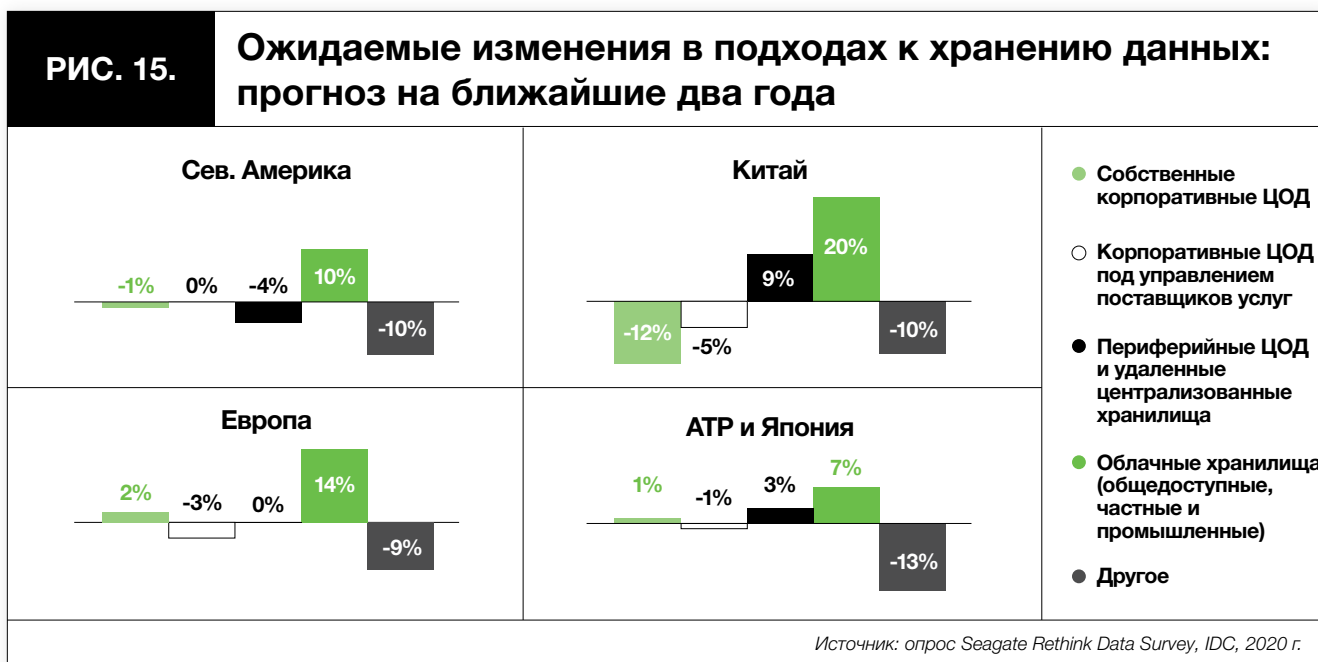
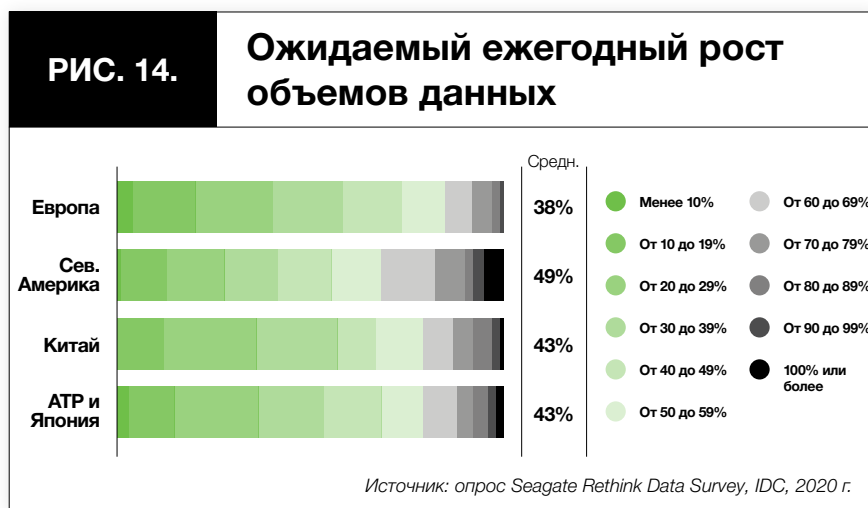


#### АНАЛИЗ IDC

### Что это означает

Чем меньше времени организации тратят на обслуживание традиционных приложений (например, на их обновление), тем больше ресурсов они могут выделить на инновации. Верно и обратное.

В Европе основное внимание уделяется обновлению существующих приложений, а также локальным приложениям. Локальные приложения из-за своей архитектуры и зависимости от локальных вычислительных ресурсов



и ресурсов хранения, как правило, менее гибкие, чем облачные. Во время пандемии COVID-19 эта проблема проявилась особенно остро, поэтому организации стали обращаться к облачным технологиям, чтобы быстро адаптироваться к меняющимся условиям. По степени гибкости инфраструктуры Европа уступает другим регионам. Это может препятствовать перемещению данных и раскрытию их ценности.

Организации, работающие на территории всей Европы, могут использовать общие приложения, но из-за GDPR и других нормативных актов они не могут передавать данные через границы. Таким образом, основным препятствием на пути к максимально эффективному

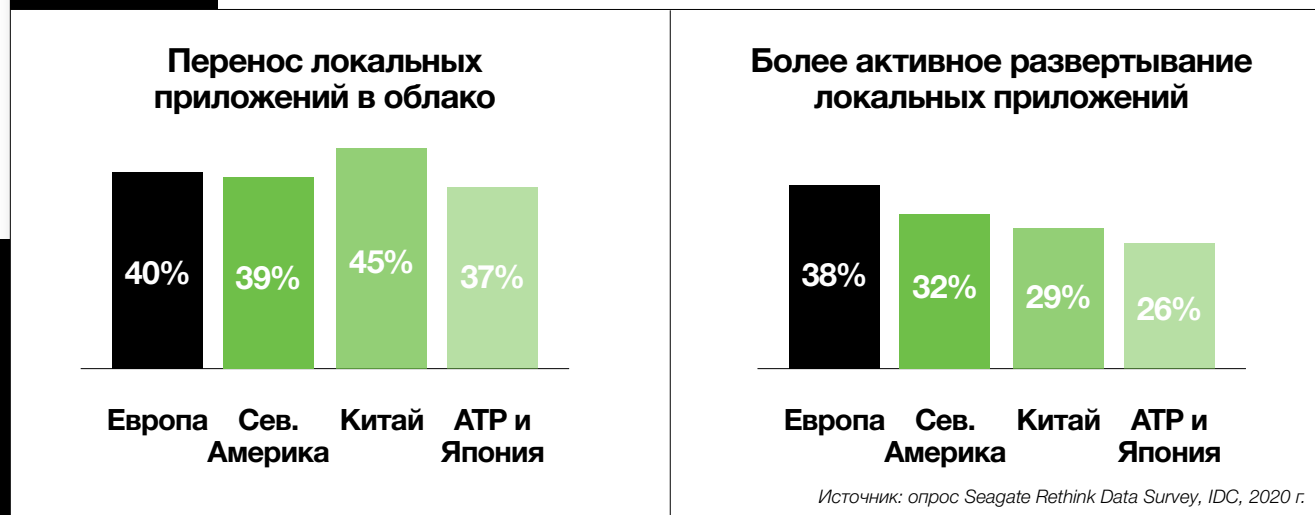
использованию данных могут оказаться именно нормативные требования, а не технологии управления данными, как, например, в Азиатско-Тихоокеанском регионе и Японии. Организациям, связанным нормативными предписаниями, приходится прилагать больше усилий и проявлять большую изобретательность, чтобы использовать доступные данные с максимальной выгодой. DataOps может помочь в этом.

Так как европейским компаниям необходимо обслуживать традиционные приложения, соблюдая при этом местные нормативные требования, совокупная стоимость владения представляет для них особую важность. Проблемы и

требования, связанные с GDPR, конфиденциальностью данных и правом на забвение, усложняют управление данными и влекут за собой дополнительные расходы. Защита конфиденциальности часто обеспечивается путем изоляции данных, что может сдерживать инновации в области управления данными.

Переход в облако может иметь разные побудительные причины. Одна из них — повышение организационной гибкости. Другая — сокращение затрат на размещение приложений или хранение данных. Европейские организации, вероятно, чаще руководствуются вторым мотивом.

**РИС. 16. Факторы, влияющие на рост объемов данных, по регионам**



**ТОЧКА ЗРЕНИЯ SEAGATE**

## Традиционные приложения и готовность к DataOps

Помимо преобладания традиционных приложений, важной особенностью европейского региона является большое значение, придаваемое суверенности и конфиденциальности данных. Связано это со строгой системой правового регулирования и геополитическим разнообразием

Европы. Зачастую требуется, чтобы инвестиции, в данном случае в существующие структуры и приложения, давали максимальную отдачу. То есть имеет место своего рода инерция развития, желание вернуть уже понесенные затраты, а не рисковать, начиная новые

дорогостоящие проекты. Значение, придаваемое конфиденциальности данных, перевешивает дополнительные затраты. Из-за этого переход в облако может замедляться, оставаясь, однако, устойчивой тенденцией на уровне 14% в следующие два года.



Темп роста объемов корпоративных данных в Европе составляет 38%. Возможно, это немного ниже, чем в других регионах, но выше, чем прогнозируемый компанией IDC общий рост объемов данных во всех секторах (32%). Учитывая степень зрелости и нормативно-правовой консерватизм в отношении использования данных в Западной Европе, а также относительно недавнее начало информатизации в некоторых регионах России, 38% — это весьма впечатляющий показатель.

Согласно профинансированному компанией Seagate отчету IDC *Инфосфера в регионе EMEA*<sup>1</sup>, общемировой рост инфосферы почти на треть обуславливается распространением систем видеонаблюдения, IoT-устройств, метаданных и развлекательных платформ.

Будь то широкое внедрение видеонаблюдения в Великобритании и Франции, совершенствование промышленного производства в Германии или развитие горнодобывающей промышленности

в России, данные сегодня играют все более важную роль. Чтобы эффективнее использовать данные, организациям в регионе следует прибегнуть к потенциалу DataOps.

В результате исследования выяснилось, что, помимо повышения безопасности данных, совокупная стоимость владения также является важнейшим фактором при выборе места для хранения данных (рис. 17). Этим может объясняться переход в облако.

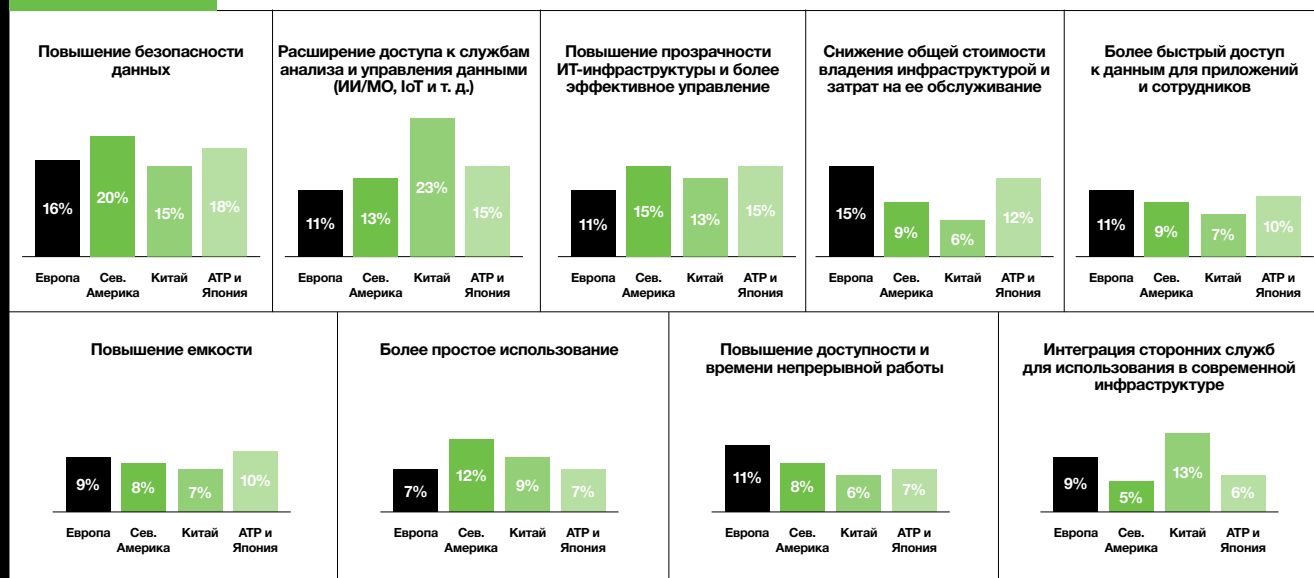
Европейские компании тратили слишком много сил на обслуживание и обновление традиционных приложений, существенно уступающих облачным в гибкости, и потому не могли выделять достаточно средств на инновации. Чтобы они могли быстро приспосабливаться к меняющимся деловым условиям, европейская инфраструктура данных должна быть более гибкой. Переход в облако также может способствовать быстрой и экономичной адаптации к новым экономическим условиям и проблемам, таким как пандемия COVID-19.

Зачастую он может стимулировать инновации и повышать динамичность бизнеса. Однако главная проблема, которая стоит сейчас перед европейскими компаниями, заключается в управлении данными. Эффективное управление данными, хранящимися в гибридных облачных и многооблачных средах, является необходимым условием для извлечения из данных максимальной ценности.

Этой потребностью может объясняться тот факт, что 40% европейских предприятий планируют внедрить методологию DataOps, а 30% уже находятся в процессе ее внедрения. **Европа осознает настоящую потребность во внедрении DataOps** для полного раскрытия потенциала данных.

РИС. 17.

## Факторы, влияющие на изменения в способах хранения данных



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

<sup>1</sup> *Инфосфера в регионе EMEA*, IDC, 2019 г.



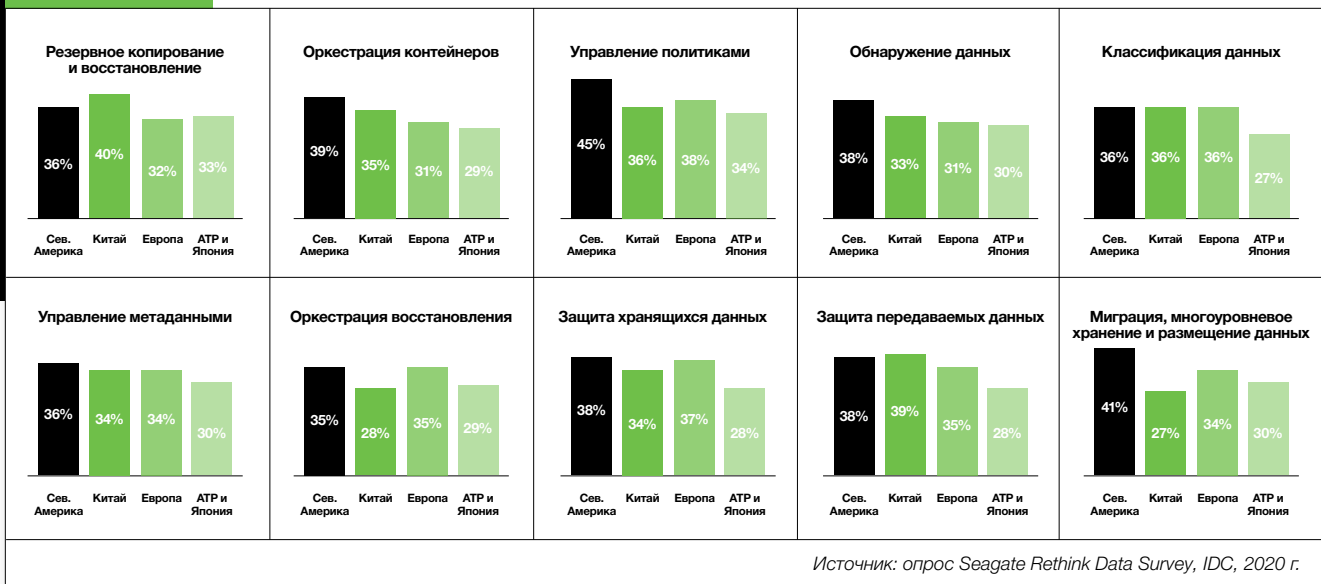
## ГЛАВА ЧЕТВЕРТАЯ

# Северная Америка

В рамках исследования были опрошены представители предприятий из США и Канады.

РИС. 18.

Доля респондентов, отметивших наличие в своих организациях единой системы управления политиками



### ДААННЫЕ IDC

Северная Америка является ведущим регионом в плане интеграции функций управления данными.

- 45% североамериканских организаций сообщают о наличии единой системы управления политиками в отношении данных.

Для сравнения: в Европе этот показатель составляет 38%, а в Азиатско-Тихоокеанском регионе — 34% (рис. 18).

- В Северной Америке прогнозируется самый высокий рост объемов данных. Ожидается,

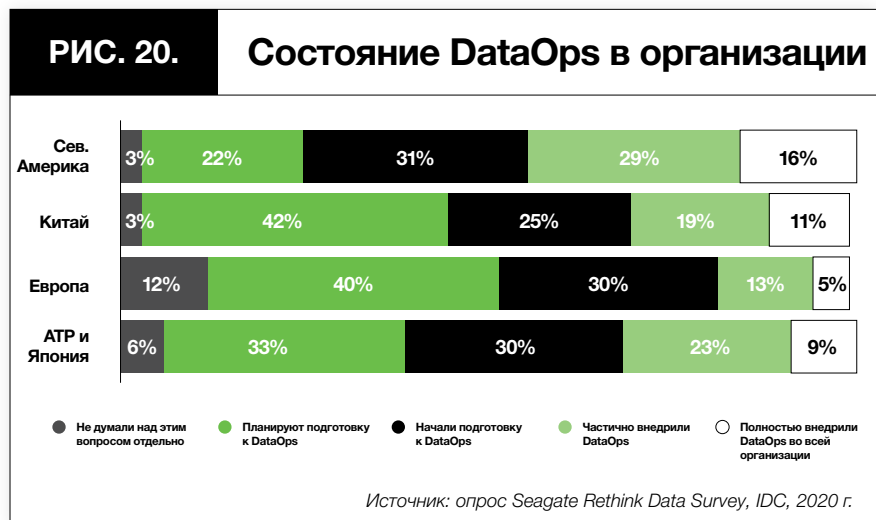
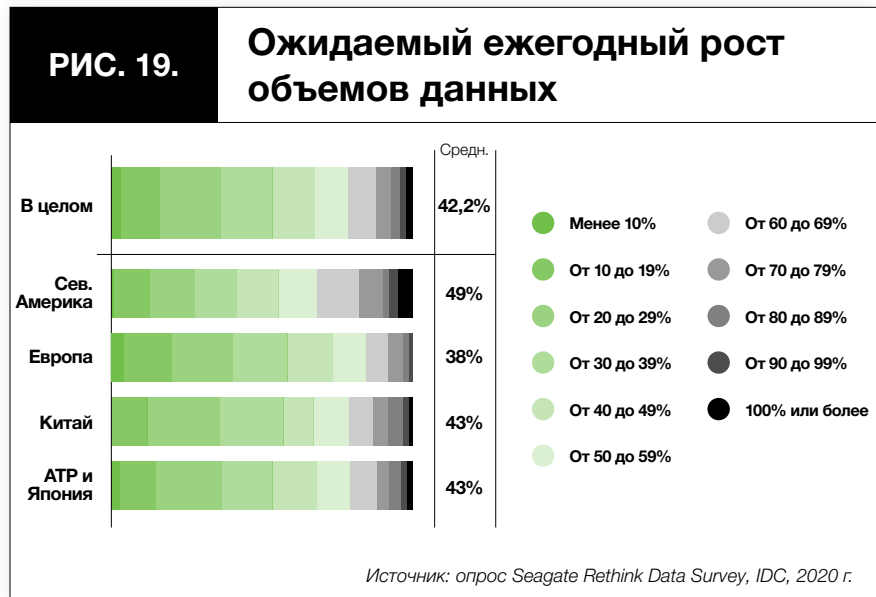
что сфера корпоративных данных в регионе увеличится почти на 50% (что больше среднего общемирового показателя, составляющего 42,2%) (рис. 19).

- Северная Америка отличается наиболее широким внедрением



DataOps. В этом регионе больше всего респондентов, заявивших о полной или частичной реализации DataOps — методологии, сводящей вместе создателей и потребителей данных. Их доля составляет 45% от общего числа (рис. 20).

- Самым важным фактором, влияющим на способ хранения данных, для североамериканских предприятий является повышение безопасности (20% респондентов) (рис. 21).
- Организации в Северной Америке добились наибольших успехов в эффективном использовании данных (рис. 22).
- Наиболее важными проблемами, препятствующими получению максимальной отдачи от данных в регионе, являются неспособность использовать полученные данные, неэффективное управление хранимыми данными, недостаточная безопасность собранных данных и сбор лишь части необходимых данных (рис. 23).



## АНАЛИЗ IDC

### Что это означает

Единый механизм применения политик — самый эффективный способ автоматизированного, согласованного управления данными в масштабе предприятия. Пожалуй, это единственный способ согласованного управления данными, так как ручные процессы вряд ли применимы в организациях среднего и большого размера. Управление политиками является фундаментом для управления данными и их использования.

Из результатов исследования можно сделать вывод, что североамериканские организации активнее всех внедряют единые политики в отношении данных и эффективнее других используют DataOps для получения максимальной отдачи.

Как правило, наличие общекорпоративной системы политик управления данными (то есть средства для управления политиками

в масштабе всего предприятия) является признаком более зрелого подхода к управлению данными. Поэтому организации, реализовавшие такую систему, способны использовать доступные данные лучше других. Безусловно, наличие общекорпоративной системы политик не является гарантией получения конкурентного преимущества, но является важной характеристикой при сравнении эффективности деятельности организаций.



Если стратегия организации направлена лишь на сокращение затрат, у нее может не оказаться средств, чтобы на равных конкурировать с организациями, уделяющими внимание оптимизации

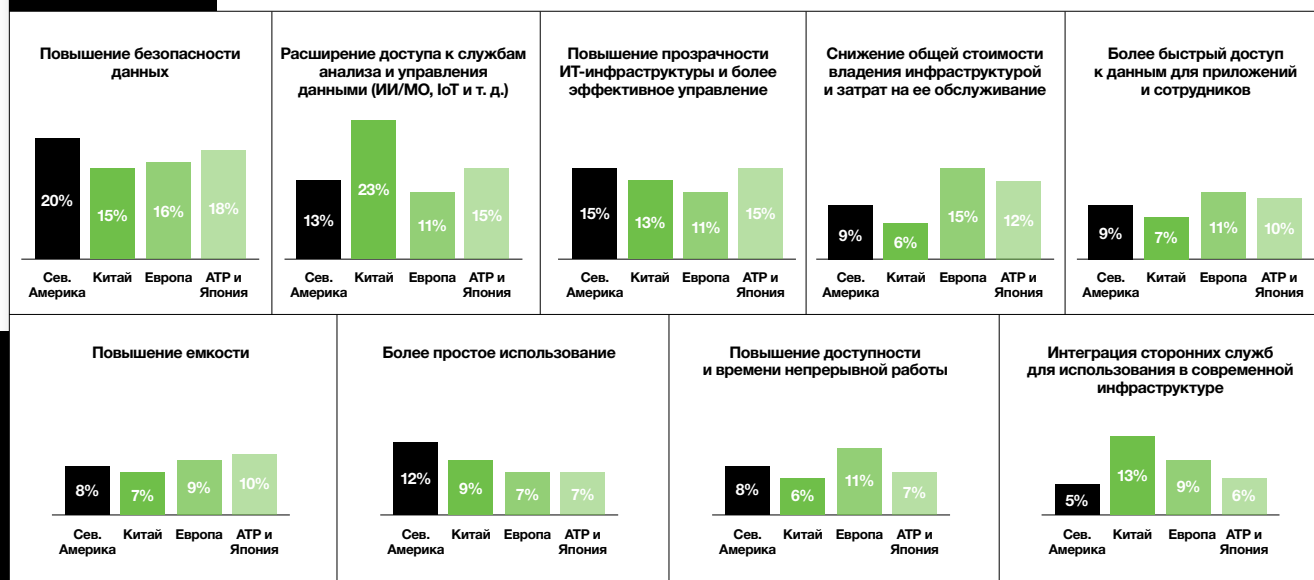
доступа к данным и раскрытию их потенциала.

Североамериканские организации, по всей видимости, уделяют больше внимания безопасности

данных, возможно, по нормативно-правовым причинам, а также из-за необходимости защищать права интеллектуальной собственности.

РИС. 21.

## Факторы, влияющие на изменения в способах хранения данных



Источник: опрос Seagate Rethink Data Survey, IDC, 2020 г.

### ТОЧКА ЗРЕНИЯ SEAGATE

## Лидер по темпу роста объемов данных

Согласно прогнозам, в ближайшие два года объем данных предприятий будет расти со средней скоростью в 42,2% в год. Это гораздо больше, чем предсказанная ранее компанией IDC средняя скорость в 32% (для всех типов данных, а не только корпоративных)<sup>1</sup>.

**Североамериканский регион выделяется на общем фоне, так как ожидаемый годовой темп увеличения объемов данных в ближайшие два года здесь превосходит среднее**

**значение в 42,2% и достигает впечатляющей цифры — 49%.**

Мы можем с уверенностью заключить, что североамериканские предприятия вносят главный вклад в увеличение объемов данных. Рост объемов данных в этом регионе обусловлен все более широким применением IoT-устройств для автоматического сбора данных, распространением средств расширенного анализа данных и приложений SaaS.

Что касается факторов, влияющих на выбор места хранения данных компаниями из Северной Америки, соображения безопасности данных перевешивают всё остальное. Контроль рисков безопасности является ключевой составляющей надежных процессов DataOps, так как руководителям североамериканских организаций постоянно приходится помнить об угрозе со стороны злоумышленников, опасности промышленного шпионажа,

<sup>1</sup> Исследование Data Age 2025, финансируемое Seagate, с использованием данных Global DataSphere, IDC, май 2020 г.

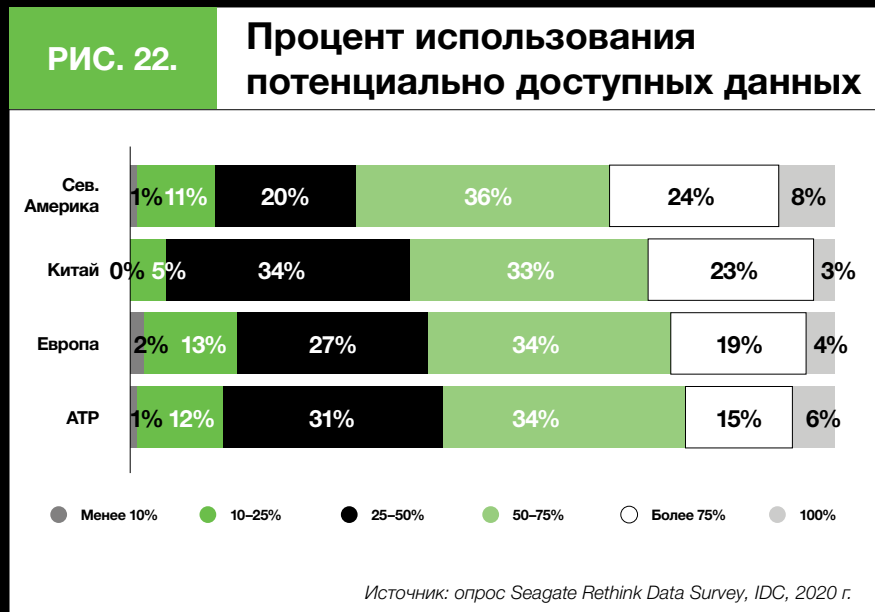


проблемах интеллектуальной собственности и защиты геополитических интересов. За стремлением защитить бизнес-данные кроется и более прозаичная, но не менее важная причина: обеспечение непрерывности бизнес-процессов. В США проблемы безопасности данных широко освещаются и обсуждаются, так как за эту безопасность здесь отвечает каждая отдельно взятая компания — национального брандмауэра нет.

Респонденты из Северной Америки, где многие компании уже внедрили или по крайней мере планируют внедрить DataOps, считают, что добились наибольших успехов в повышении эффективности использования данных (рис. 22).

В то же время в качестве наиболее важных проблем, препятствующих получению максимальной отдачи от данных в регионе, они называют неспособность использовать полученные данные, неэффективное управление хранимыми данными, недостаточную безопасность собранных данных и сбор лишь части необходимых данных (рис. 23).

Определение подходящей среды хранения для различных наборов данных, правильная классификация данных на основе взаимодействия между их создателями и



руководителями предприятий, определение назначения различных типов данных, организация сбора нужных данных, их защита, а также преобразование собранных данных в пригодную для использования форму — все это задачи DataOps.

Отличием Северной Америки от остальных регионов стало то, что респонденты из нее особо отмечали необходимость в упрощении использования данных (рис. 21). Осознание

этого факта очень полезно. Хотя в североамериканском регионе методология DataOps применяется достаточно широко и присутствует сравнительно ясное понимание необходимости ее внедрения, для более эффективного использования данных требуется дальнейшее просвещение и накопление опыта применения DataOps.





SEAGATE

[seagate.com](https://www.seagate.com)

© Seagate Technology LLC, 2020. Все права защищены. Seagate, Seagate Technology и логотип Spiral являются зарегистрированными товарными знаками компании Seagate Technology LLC в США и/или других странах. Прочие товарные знаки или зарегистрированные товарные знаки являются собственностью соответствующих владельцев. Компания Seagate оставляет за собой право изменять ассортимент или характеристики своих продуктов без предварительного уведомления. SC738.1-2006, июль 2020 г.