

Содержание

Предисловие	7
Введение	10
Часть 1	
ДАнные: СМЕНА ПАРАДИГМ	
От первой научной революции до четвертой промышленной	17
Глава 1. Цифровое многообразие	18
Глава 2. От данных к мудрости	52
Глава 3. От побочного продукта к стратегическому ресурсу	76
Глава 4. Данные как ресурс: особенности и подходы к управлению	102
Глава 5. От стратегического ресурса к ценнейшему активу	126
Глава 6. Данные как актив: барьеры и ошибки на пути извлечения ценности	158
Глава 7. Построение цепочек	190
Часть 2	
ДАнные: ИЗВЛЕЧЕНИЕ ЦЕННОСТИ	
От планирования к расширению возможностей применения	227
Глава 8. Данные как объект управления	228
Глава 9. Управление данными: принципы и структуры	252
Глава 10. Руководство данными	274
Глава 11. Планирование и проектирование данных	296
Глава 12. Обеспечение доступности и обслуживание данных: основы	326

Когда шла работа над книгой «Ценность ваших данных», одному из авторов случилось оказаться в экстремальной ситуации. И та выявила все действительно важные для человека качества и ценности, но в особенности продемонстрировала ЦЕННОСТЬ ВАШИХ ДРУЗЕЙ и СЕМЬИ. Именно ДРУЗЬЯМ и СЕМЬЯМ мы и хотели бы посвятить эту книгу.

*Сергей Кузнецов,
Александр Константинов,
Николай Скворцов*

Предисловие

В XXI веке много говорят о данных: эксперты спорят, стоит ли уделять правильной организации работы с ними столь много внимания или все это хайп и надувано. Авторы этой книги уже более 20 лет наблюдают за развитием различных технологий, за тем, как рождаются методологии работы с данными, в компаниях появляются специалисты по этому направлению.

Мы постарались собрать воедино исторические, организационно-методические и технологические аспекты данных, раскрыть для читателя многообразие терминов и аббревиатур. Также вас ждет рассказ о существующем инструментарии работы с данными и о трендах, которые мы наблюдаем сегодня и увидим завтра. Все это — на языке, доступном людям без специального ИТ-образования.

Мы долго шли к решению о написании этой книги, и, как бы ни откладывали это решение, наша жизнь и работа все время подводили к необходимости описания данных, которые буквально пронизывают нашу жизнь. Для нас же ситуация получилась и вовсе уникальной, потому что мы, как профессионалы в этой области, находимся, если можно так выразиться, по обе стороны проникновения данных в жизнь общества. Именно поэтому мы все-таки решились погрузиться с головой в эту работу, которая стала, неожиданно для нас, возможностью всесторонне систематизировать не только опыт и знания, но и отраслевые компетенции заказчиков и партнеров.

Еще одним важным фактором, побудившим написать книгу, стала полуторагодичная работа нашего коллектива над переводом и научной редактурой второго издания легендарной книги «DAMA-DMBOK: Свод знаний по управлению данными». Результаты этой работы по достоинству оценены российскими читателями, которым книга позволила получить новые и систематизировать уже имеющиеся знания, а также стать частью международного сообщества и даже целой индустрии работы с данными, говорить с ним на языке единых терминов и подходов к процессам работы с данными.

Важно отметить, что создание книги — это процесс, напрямую связанный с импортозамещением программных продуктов, когда на смену зарубежному софту приходят российские разработки. И этот процесс стал гораздо важнее в условиях санкций против России. Период с конца февраля 2022 года стал крайне тяжелым для отечественных компаний, закупивших зарубежный софт. Особенно болезненным было «расставание» с иностранными организациями, которые в течение долгих лет поставляли нам продукты в области управления данными.

Такие компании, как IBM, SAP, Oracle, объявили об уходе из России, т. е. продаж, а главное, поддержки, выполнения проектов и много другого больше не будет. Конечно, в такой ситуации все большее количество организаций начинает принимать активные меры по импортозамещению. Это движение существовало и раньше, просто сейчас, в силу понятных причин, оно ускорилось. Во многом предвидя такую ситуацию, мы постарались учесть опыт применения российского программного обеспечения в области управления данными. Уверены, наша книга будет способствовать расширению круга отечественных литературных источников по тематике управления данными и тем самым сыграет значимую роль в ускорении всех процессов, связанных с заменой иностранных программных продуктов на отечественные.

И конечно, стоит представить уважаемым читателям одного из главных героев книги. Начиная со второй ее части, мы вводим в повествование некую абстрактную корпорацию «Телеком Дубль». Это необходимо для большей связи с реальностью, с проектами внедрения, опыт в которых у коллектива авторов суммарно более семи лет. Именно с помощью «Телеком Дубль» мы будем приводить конкретные примеры, чтобы читателям было удобнее визуализировать те понятия, концепции и инструменты, о которых будет идти речь на страницах книги. Название у компании абстрактное, она нужна нам лишь для примера увеличения эффективности ее работы при внедрении различных практик в области управления данными. Соответственно, любые сходства с реальными компаниями случайны. В начале повествования компания не обладает какими-то серьезными знаниями в области управления данными. Она ведет свою деятельность на рынке телекоммуникаций, ее клиентами являются как частные лица, так и компании. Она имеет сложную филиальную организационную структуру, информационные системы, отвечающие за различные функции, включая фронт- и бэк-офисы. Особых успехов нет. Но с каждой новой главой ситуация будет меняться! В каждой главе мы будем описывать, как рассматриваемая проблематика выглядит на примере «Телеком Дубль», что для нее значит внедрение того или иного инструмента, что получается в результате как для ИТ-ландшафта, так и с организационной и эксплуатационной точки зрения.

Ценность ваших данных... Эти три слова описывают книгу целиком. Данные, информация, знания — все эти понятия окружают нас каждый день. Но ценность данных нужно еще раскрыть с помощью грамотного управления: без этого они могут быть лишь балластом, тянущим на дно.

При обсуждении процессов управления данными часто рассматривают две цепочки — цепочку ценности данных и цепочку поставок данных. Эти обобщенные концепции настолько важны, что мы решили отразить их на обложке. Что это за цепочки, как они соотносятся друг с другом и как выстроить их таким образом, чтобы извлечь из данных максимальную ценность? Об этом мы тоже поговорим на страницах предлагаемой книги.

Надеемся, что наша работа будет интересна самому широкому кругу читателей, которые интересуются данными. А учитывая тот факт, что данные в последние годы окутывают нас все больше и больше, круг будет расширяться.

Книга состоит из двух частей.

В первой части описывается многообразие окружающих нас данных и связанных с ними технологий, рассматривается последовательность преобразований «данные — информация — знания — мудрость», обсуждаются эволюция отношения к данным в организациях и задачи, стоящие перед ними в связи с переходом к представлению о данных как ценнейшем стратегическом активе.

Во второй части анализируются особенности данных как объекта управления, описываются функции управления информационными активами в контексте построения цепочек ценности данных и цепочек их поставок, обсуждаются основные тренды будущего в области технологий управления данными.

Введение

В 2006 году британский математик и бизнесмен Клайв Хамби ввел в обращение фразу «Данные — это новая нефть»¹. Он произнес ее на саммите топ-менеджеров по маркетингу, который проводился американской Ассоциацией национальных рекламодателей (Association of National Advertisers, ANA) в Школе менеджмента им. Келлога (Kellogg School of Management).

В том же году вице-президент ANA Майкл Палмер (Michael Palmer) в статье, посвященной тезису Хамби, развил его мысль: «Данные похожи на сырье. Оно полезно, но в необработанном виде непригодно для использования. Его необходимо преобразовать в газ, пластмассу, химикаты и тому подобные вещи, чтобы создать имеющие ценность объекты, которые обеспечили бы деятельность, приносящую прибыль; аналогичным образом и данные должны быть разложены на элементы и проанализированы, чтобы они обрели ценность»².

Спустя несколько лет, в 2013 году, глава IBM Вирджиния Рометти (Virginia Rometty), выступая перед представителями американского Совета по международным отношениям (Council on Foreign Relations, CFR), несколько перефразировала эту мысль, сравнив с новой нефтью большие данные (Big Data)³.

Сравнение данных с нефтью стало использоваться еще активнее после появления в 2017 году в журнале *The Economist* статьи «Самый ценный ресурс в мире больше не нефть, а данные»⁴.

¹ Клайв Хамби (Clive Humby) — создатель успешной программы лояльности «Клубная карта» для клиентов розничной сети супермаркетов Tesco. Подавляющее большинство источников приписывают авторство этой фразы именно ему.

² https://ana.blog.com/maestros/2006/11/data_is_the_new.html

³ <https://siliconangle.com/2013/03/11/ibms-ceo-says-big-data-is-like-oil-enterprises-need-help-extracting-the-value>

⁴ <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

Наконец, не менее ярко высказался авторитетный китайский ученый и эксперт Кай-Фу Ли (Kai-Fu Lee)⁵. В своей книге «Сверхдержавы искусственного интеллекта. Китай, Кремниевая долина и новый мировой порядок», вышедшей в 2019 году, он пишет: «Если данные — это новая нефть, то Китай — это новая Саудовская Аравия»⁶.

Но одновременно с ростом количества упоминаний о данных как о новой нефти возросло и число публикаций, выдвигающих обратный тезис: «Данные — это не новая нефть»⁷. В последнее время стали появляться статьи, рассматривающие вопросы токсичности данных⁸ и даже высказывающие мнение, что данные — это «новый мусор», не имеющий особой ценности⁹.

Очевидно, ключ к разрешению противоречия — приведенные выше слова вице-президента ANA Майкла Палмера о том, что данные похожи на сырье и для того, чтобы они обрели ценность, необходима их предварительная обработка. Но следует учитывать, что это особое сырье, во многом отличающееся от нефти и других природных ресурсов. Так как же правильно организовать его обработку? Существуют ли рациональные методы работы с данными? Ответы на эти вопросы — в соответствующем отношении к данным как отдельных компаний и организаций, так и целых государств.

Последние 10–15 лет в среде специалистов, занимающихся вопросами работы с данными, последовательно формируется представление об управлении данными как активом¹⁰. Созданы профессиональные организации, развивающие это направление. В частности, функционирует Международная ассоциация управления данными (Data Management Association International, DAMA), объединившая профессионалов в области управления данными по всему миру. Она призвана собирать, систематизировать и пропагандировать прогрессивный опыт. Регулярно проводятся различные тематические форумы, появилось большое

⁵ Кай-Фу Ли — экс-президент Google China, руководитель инвестиционного фонда для китайских стартапов в области искусственного интеллекта Sinovation Ventures.

⁶ <https://www.mann-ivanov-ferber.ru/books/sverhderzhavy-iskusstvennogo-intellekta/>

⁷ В качестве примера приводим следующие публикации:

<https://hbr.org/2012/11/data-humans-and-the-new-oil>

<https://towardsdatascience.com/data-is-not-the-new-oil-bdb31f61bc2d>

<https://towardsdatascience.com/data-is-not-the-new-oil-721f5109851b>

⁸ <https://enterpriseproject.com/article/2019/7/data-science-data-can-be-toxic>

⁹ <https://expert.ru/2020/08/6/dannyye-eto-novyij-musor/>

¹⁰ Под активом понимается имеющийся в собственности или контролируемый экономический ресурс, содержащий в себе или производящий ценность. При этом он может быть конвертирован в деньги (монетизирован).

количество книг и статей на эту тему, включая вышедшее в 2017 году (и переведенное на русский язык) уже второе издание руководства DAMA к своду знаний по управлению данными (DAMA-DMBOK2)¹¹.

Что же нужно делать компании или организации, чтобы начать обращаться с данными как с активом? Базовые шаги, позволяющие обеспечивать и наращивать монетизацию данных, известны. Они изложены, например, в книге ведущего аналитика Gartner Дагласа Лейни¹² «Инфономика: информация как актив: монетизация, оценка, управление»¹³. Тем не менее в этой же книге подчеркивается, что в большинстве организаций не введена практика последовательного управления информационными активами, как это сделано в отношении материальных или финансовых активов. Сотрудники компаний продолжают действовать по старинке. В частности, сохраняется несогласованность между управлением информационными технологиями и управлением данными. В своей переписке с Лейни Джон Лэдли¹⁴ отмечал: «Пока мне приходится наблюдать, как команды разработчиков поспешно производят сотни приложений и сервисов, не принимая во внимание вопросов использования связанных с ними данных, мне будет обеспечена полная занятость».

Недооценка перехода на новые подходы к работе с данными рискованна. В 2016 году Клаус Шваб (Klaus Schwab), основатель и президент Всемирного экономического форума в Давосе, в своей одноименной книге провозгласил четвертую промышленную революцию. Революцию, которая основана на цифровизации и характеризуется диджитальным (ломающим привычные представления)¹⁵ воздействием на утвердившиеся традиционные компании. Чтобы удержаться на плаву, компании вынуждены трансформировать себя, используя для этого цифровые и информационные технологии (рис. В1). Шваб подчеркивает, что особую тревогу в связи с осуществлением стоящих на повестке дня революционных

¹¹ DAMA-DMBOK: Свод знаний по управлению данными. Второе издание / Dama International. — М.: Олимп-Бизнес, 2020 (издание подготовлено при поддержке компаний «Юнидата» и BSSG).

¹² Даглас Лейни (Douglas Laney) — известный специалист в области хранилищ данных и стратегического управления данными, автор модели зрелости процессов управления корпоративной информацией компании Gartner, а также один из авторов термина Big Data (большие данные).

¹³ Laney D. B. (2017). Infonomics: How to Monetize, Manage, and Measure Information as an Asset for Competitive Advantage; Routledge; 1st edition.

¹⁴ Джон Лэдли (John Ladley) — признанный в мире авторитет в области управления данными. Автор книг Making Enterprise Information Management (EIM) Work for Business: A Guide to Understanding Information as an Asset («Управление корпоративной информацией в интересах бизнеса: Руководство по выработке отношения к информации как к активу») и How to Design, Deploy and Sustain an Effective Data Governance Program («Как разработать, внедрить и поддерживать эффективную программу руководства данными»).

¹⁵ <https://cyberleninka.ru/article/n/gumanitarnost-v-obrazovanii-kak-otvet-na-dizruptivnye-innovatsii>

Четвертая промышленная революция основана на цифровизации и характеризуется **дизруптивным (ломающим привычные представления)** воздействием на утвердившиеся традиционные компании.

Чтобы оставаться на плаву, компании вынуждены трансформироваться, используя для этого цифровые технологии, — проводить **цифровую трансформацию**.

Компания **изобретает** заново саму себя, директор по цифровой трансформации (CDTO) выступает в роли «**внутреннего предпринимателя**».

СТРАТЕГИИ В ОБЛАСТИ ДАННЫХ И АНАЛИТИКИ — ДРАЙВЕРЫ БУДУЩЕГО

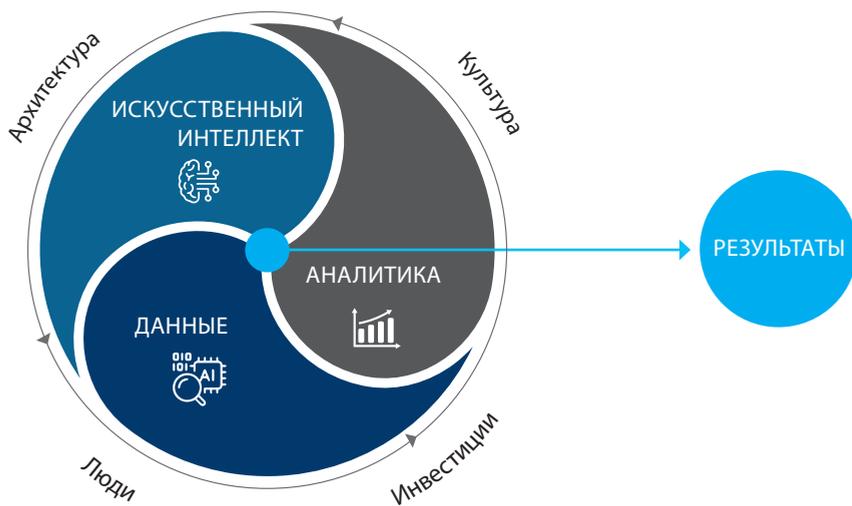


Рис. В1.

Основные драйверы цифровой трансформации (Gartner)

преобразований у него вызывают два фактора, ограничивающие их позитивный потенциал. Первый — низкий уровень управления текущими изменениями. Второй — отсутствие единой последовательной концепции их осуществления. Оба фактора распространяются и на вопросы работы с данными.

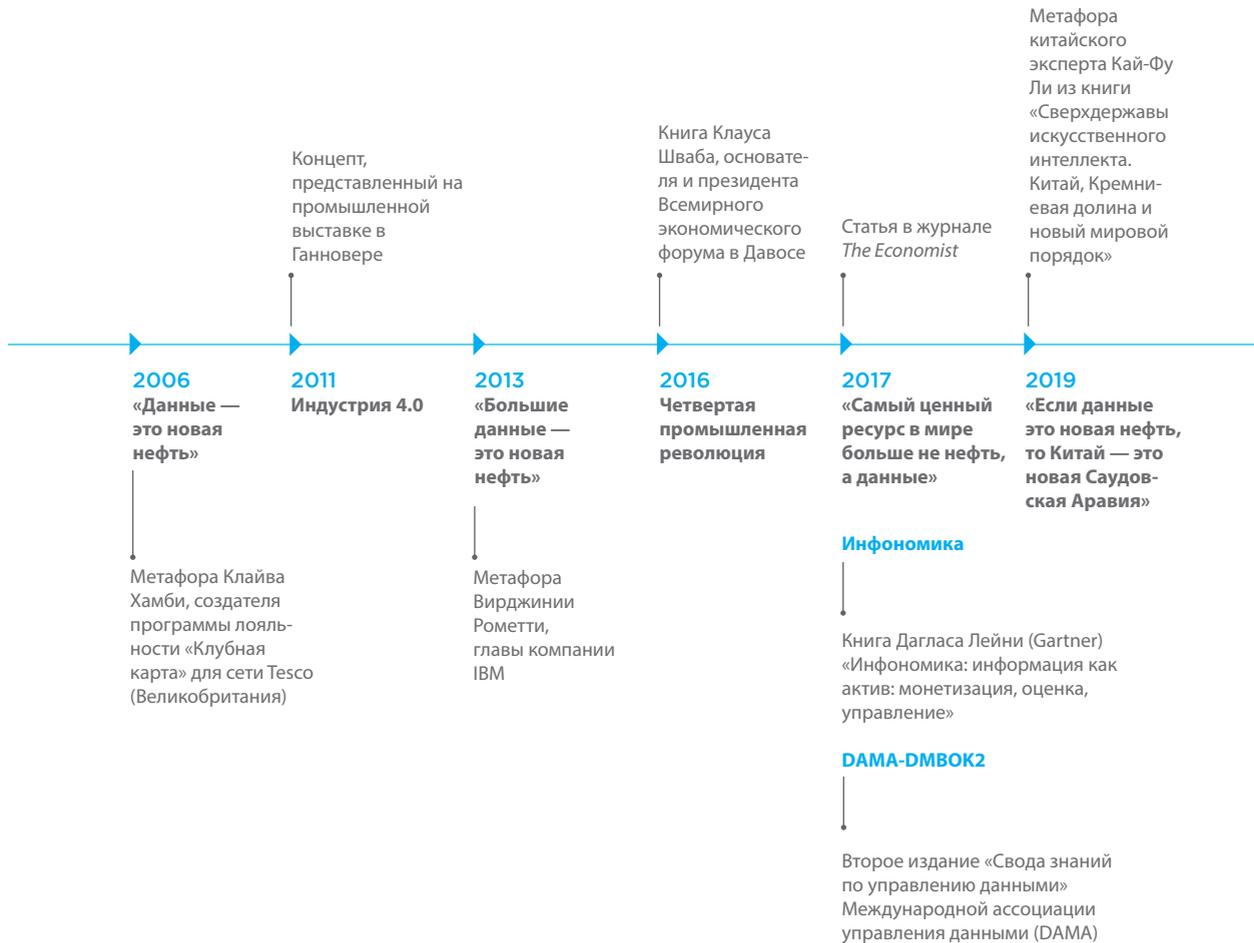


Рис. В2.

Развитие представлений о ценности данных и цифровой трансформации

Некоторые ключевые этапы развития представлений о ценности данных и цифровой трансформации начиная с начала 2000-х отражены на рисунке В2.

Появившийся во время первой президентской кампании Билла Клинтона в 1992 году популярный слоган «Это экономика, тупица» (It's the economy, stupid), подчеркивающий важность и сложность решения экономических проблем, с тех

пор трансформировался применительно ко многим контекстам. В частности, в контексте цифровизации широко распространилось клише «Это данные, тупица» (It's the data, stupid)¹⁶. А чтобы привлечь еще большее внимание к необходимости комплексного учета всех аспектов работы с данными и управления ими, стала использоваться фраза «Это не просто данные, тупица» (It's not just the data, stupid)¹⁷. Важность всестороннего взгляда на данные убедительно подтверждает практика реализации проектов, связанных с уберизацией (uberization), внедрением концепции «умного города» (smart city) и другими проявлениями того, что называется цифровой экономикой (digital economy) или экономикой, управляемой на основе данных (data-driven economy).

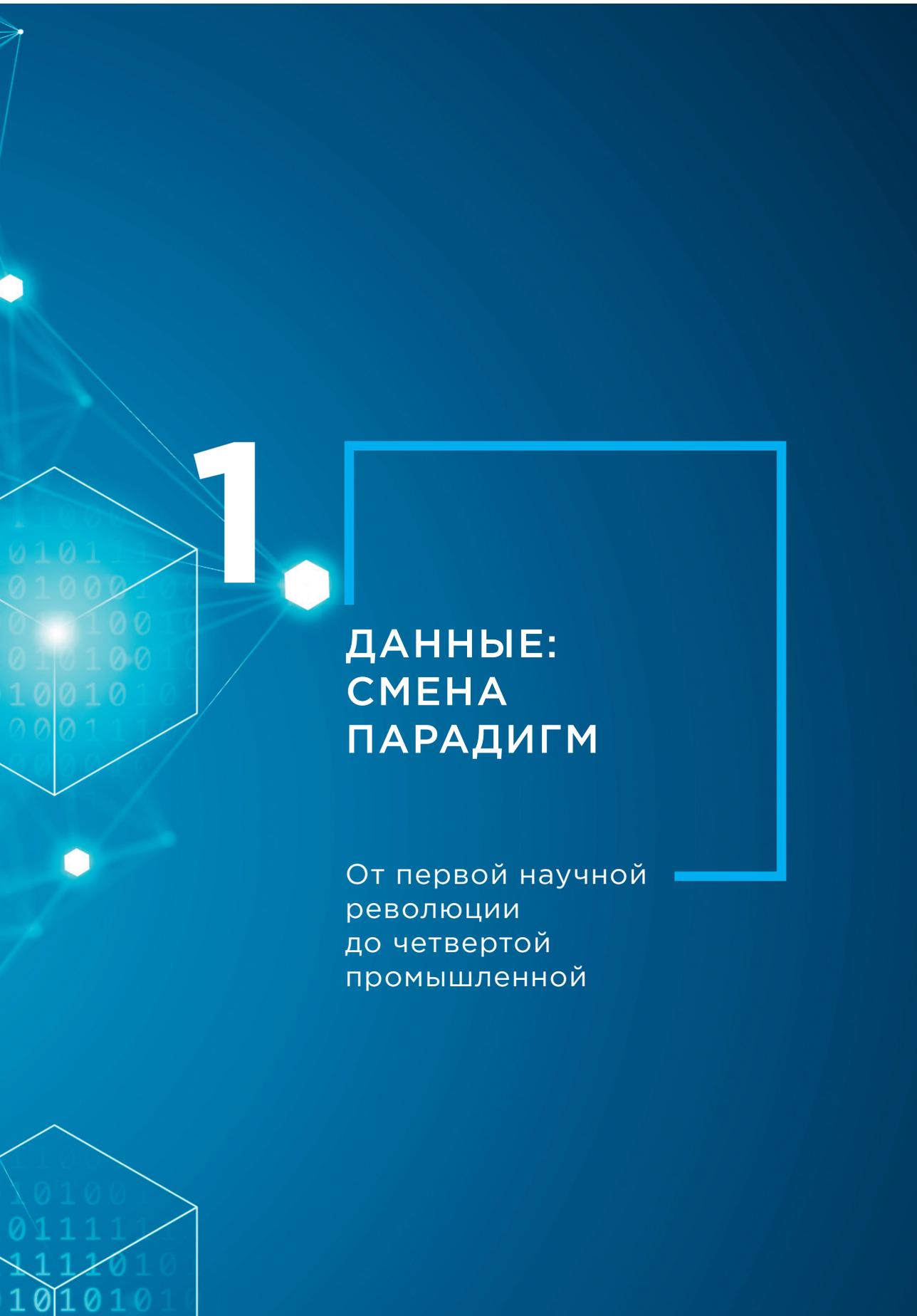
Как наладить управление данными, чтобы достойно противостоять деструктивным воздействиям, извлечь ценность из имеющихся информационных ресурсов (превратить их из «мусора» в «нефть») и неуклонно ее повышать? Разобраться в этих вопросах руководителям и рядовым сотрудникам предприятий и организаций, не имеющих прямого отношения к сфере ИТ, порой бывает непросто. Да и у специалистов в области ИТ, как показывает наш опыт, возникает немало вопросов.

К сожалению, литературы на русском языке на эту тему еще очень мало, а информация из многочисленных иностранных источников не всегда доступна, в том числе из-за языкового барьера. Хорошо осознавая эти проблемы, мы решили поделиться собственным накопленным опытом и видением ситуации. Надеемся также, что наша книга дополнит набор русскоязычных материалов, используемых преподавателями и обучающимися по такому недавно возникшему и стремительно развивающемуся образовательному направлению, как data literacy (грамотность в области данных).

¹⁶ <https://atos.net/en/blog/its-the-data-stupid>

¹⁷ <https://www.osp.ru/news/articles/2016/25/13049787>



A decorative graphic on a dark blue background. It features a network of glowing white lines connecting several hexagonal nodes. Some nodes are larger and more prominent, while others are smaller. The overall effect is a digital or data network. The number '1' is positioned to the right of the main graphic.

1

ДААННЫЕ: СМЕНА ПАРАДИГМ

От первой научной
революции
до четвертой
промышленной



Глава 1.

Цифровое многообразие

1.1. В чем ценность данных

Стоит только задуматься, в каком мире мы живем, и сразу обнаруживаешь, какое количество технологий нас окружает. Сейчас мы уже не представляем себе жизни без них. Мало кто задумывается, что с каждым годом динамика формирования наших потребностей во всем новом растет невероятными темпами. Даже такие крайне необходимые и уже привычные вещи, как мобильный телефон, еще 30 лет назад вызывали восхищение. Тогда пользоваться ими могли лишь избранные, сейчас же этот девайс стал постоянным спутником современного человека и уже не воспринимается как мобильное средство голосовой коммуникации. iPhone первого поколения, разработанный корпорацией Apple, поступил в продажу в 2007 году и перевернул представление о мобильном телефоне, сделав его центром коммуникаций, развлечений, устройством для получения всевозможных электронных услуг и средством платежей. Функциональные возможности современных телефонов становятся все шире. И сейчас, по прошествии 15 лет, мы себе уже не представляем, как можно жить без всего этого.

Таких примеров сотни тысяч — в самых разных областях нашей жизни. Все они стали настолько повседневными примерами нашего общения с технологиями, что мы к ним привыкли и перестали замечать.

Но несмотря на эту кажущуюся повседневность современных технологий, мировая экономика очень чутко реагирует на происходящие в нашей жизни изменения. Эта реакция проявляется повышенным интересом инвесторов в сторону технологического сегмента, который позволяет на основе человеческого интеллекта, предпринимательской смелости и современных процессов управления создавать крупнейшие мировые компании, каждая из которых может в любой момент обогнать, казалось бы, незыблемые столпы мировой экономики.

Давайте взглянем на рейтинг 100 крупнейших компаний мира (табл. 1.1)¹⁸:

¹⁸ <https://www.pwc.com/gx/en/audit-services/publications/assets/pwc-global-top-100-companies-2021.pdf>

Табл. 1.1.

Рейтинг 100 крупнейших компаний мира (по рыночной капитализации)

Позиция	Название компании	Местоположение	Сектор экономики	31 марта 2021 г.		31 марта 2020 г.	
				Изменение по сравнению с мартом 2020 г.	Рыночная капитализация, \$ млрд	Рейтинг	Рыночная капитализация, \$ млрд
1	Apple Inc	США	Технологии	2	2051	3	1113
2	Saudi Aramco	Саудовская Аравия	Энергетика	-1	1920	1	1602
3	Microsoft Corp	США	Технологии	-1	1778	2	1200
4	Amazon.Com Inc	США	Потребительские товары вторичной необходимости	0	1558	4	971
5	Alphabet Inc	США	Технологии	0	1393	5	799
6	Tencent	Китай	Технологии	1	753	8	469
7	Tesla Inc	США	Потребительские товары вторичной необходимости	75	641	83	96
8	Alibaba Gpr	Китай	Потребительские товары вторичной необходимости	-3	615	6	522
9	Berkshire Hathaway	США	Финансы	-1	588	9	443
10	Tsmc	Тайвань	Технологии	9	534	20	235
11	Visa Inc	США	Промышленность	0	468	12	316
12	Jpmorgan Chase	США	Финансы	2	465	15	277
13	Johnson & Johnson	США	Здравоохранение	-4	433	10	346
14	Samsung Electronics	Южная Корея	Технологии	6	431	21	234
15	Kweichow Mouta	Китай	Потребительские товары	12	385	28	197
16	Wallmart Inc	США	Потребительские товары вторичной необходимости	-6	383	11	322
17	Mastercard Inc	США	Промышленность	0	354	18	243
18	Unitedhealth Grp	США	Здравоохранение	0	352	19	237
19	Lvmh Moët Hennessy	Франция	Потребительские товары вторичной необходимости	12	337	32	188

Обращает на себя внимание количество технологических компаний в первой десятке. Если быть точными, их пять, а в первой сотне — уже 30. Помимо этого, в рейтинге еще 19 компаний, занимающихся здоровьем и биотехнологиями, что также относится к технологическому сегменту деятельности. По прогнозам аналитиков,

количество высокотехнологических компаний будет возрастать еще большими темпами, а одним из основных критериев успешности для компаний из других областей станет использование высоких технологий в развитии своих решений.

Экономическая важность высокотехнологических компаний уже сейчас оказывается основополагающей для развития мировой и национальных экономик. Из 20 крупнейших компаний мира 10 (Apple, Alphabet, Amazon, Microsoft, Alibaba, Walmart, Tencent, Johnson & Johnson, Procter & Gamble, Nestlé), по мнению Boston Consulting Group, являются самыми инновационными на планете (рис. 1.1)¹⁹.

Позиция: 1–10	Позиция: 11–20	Позиция: 21–30	Позиция: 31–40	Позиция: 41–50
1 Apple	11 Siemens	21 Toyota	31 Xiaomi	41 Inditex
2 ALPHABET	12 LG	22 Salesforce	32 IKEA	41 Moderna
3 AMAZON	13 Facebook	23 Walmart	33 Fast Retailing	42 Philips
4 MICROSOFT	14 Alibaba	24 Nike	34 Adidas	43 Disney
5 Tesla	15 Oracle	25 Lenovo	35 Merck & Co.	44 Mitsubishi
6 Samsung	16 Dell	26 Tencent	36 Novartis	45 Comcast
7 IBM	17 Cisco	27 Procter & Gamble	37 Ebay	46 GE
8 HUAWEI	18 Target	28 Coca-Cola	38 PepsiCo	47 Roche
9 SONY	19 HP	29 Abbott Labs	39 Hyundai	48 Astra Zeneca
10 PFIZER	20 Johnson & Johnson	30 Bosch	40 SAP	50 Bayer

Рис. 1.1.

Самые инновационные компании 2021 г.

¹⁹ The Most Innovative Companies 2021. The Serial Innovation Imperative. Boston Consulting Group 2021.

Вывод прост: сегодня, чтобы стать успешной компанией, необходимо быть инновационной и высокотехнологичной независимо от отрасли деятельности. Именно это становится фундаментом для развития.

При всем многообразии высоких технологий в различных областях нашей жизни есть несколько ключевых технологий, на которых исследователи и разработчики строят свои открытия. Одной из групп таких технологий являются решения для работы с данными, которые не только помогают в создании новых продуктов, но и позволяют повысить эффективность действующих решений.

Одним из самых ценных активов компаний в ближайшем будущем станут данные, накопленные ими за годы деятельности: информация о клиентах, поставщиках, номенклатуре закупаемой и продаваемой продукции, о сотрудниках, оборудовании, исследовании рынков, а также статистика, публикуемая различными государственными, общественными и частными организациями. Возможность анализировать и использовать огромные массивы данных позволит государствам и организациям стать более эффективными — создать огромное количество новых направлений деятельности и совсем по-иному взглянуть на давно сложившиеся стороны жизни.

В последнее время появляется все больше и больше исследований, описывающих человека как огромный массив информации, связанной с его документами, услугами, которыми он пользуется, аккаунтами в социальных сетях, транзакциями, платежными инструментами и другими цифровыми следами повседневной жизни. Еще 100 лет назад человека, его внешность, знания, устремления, жизненный опыт, чувства, таланты, эмоции и желания изображали художники и литераторы, сейчас же цифровые следы человека заметны везде.

Как мы уже говорили, мы живем в эпоху технологий и открытий, происходящих с немыслимой скоростью, когда с момента фундаментальных открытий до повседневности технологий проходят считанные годы, а объем хранимых человечеством данных прирастает ежегодно десятками процентов. Согласно результатам исследования аналитической фирмы IDC, опубликованным в докладе «Эра данных — 2025» (Data Age 2025), в ближайшие годы основной объем данных будут производить не отдельные пользователи, а организации. На промышленность и другие сферы экономики придется до 60% всех данных мира. Для сравнения, в 2015 году предприятия генерировали треть всех мировых данных. При этом, как утверждают авторы исследования, в будущем более важную роль будет играть качество данных, а не их количество. «Не все данные одинаково важны, а без контекста они и вовсе бесполезны. В этот период перемены лидерство будет принадлежать организациям, которые сумеют определить наиболее критичные подгруппы информации с максимальным влиянием на нужную сферу

деятельности и сосредоточатся именно на них» — говорится в отчете. Пятая часть всех данных к 2025 году станет критически важной, считают исследователи. Речь идет о сведениях, от которых будет зависеть жизнь и безопасность людей, международная обстановка и мир на планете.

При этом в ближайшие годы разрыв между объемом данных, нуждающихся в защите, и реально защищенной информацией будет только расти. К 2025 году до 90% всей информации должно быть так или иначе защищено. Авторы доклада также отмечают, что значительный объем данных будет исходить от устройств, которые окружают нас каждый день. Во-первых, к 2025 году 75% всего населения Земли будет иметь постоянный доступ в интернет. Во-вторых, многократно возрастет количество умных гаджетов и домашних роботов, которые будут производить так называемые метаданные — служебную информацию, которой машины будут обмениваться между собой для слаженной работы. По сравнению с сегодняшним днем каждый человек будет в 20 раз чаще взаимодействовать с интернетом или с устройствами с выходом в интернет. Если сейчас среднее количество взаимодействий чуть больше 600, то к 2025 году мы будем сталкиваться с сетью 4800 раз в день²⁰.

Работа тысяч корпораций и государственных структур по созданию различного рода сервисов, предоставляющих услуги посредством интернета, уже привела к тому, что в крупных городах люди зачастую получают существенную долю государственных сервисов и приобретают различные товары и услуги за пару минут, пользуясь приложениями в телефоне. Более того, зачастую мы сталкиваемся с тем, что компании знают о наших предпочтениях и предпочтениях членов наших семей гораздо больше, чем мы сами, напоминая нам о том, какие продукты мы предпочитаем, когда необходимо запланировать очередную покупку и доставку этих продуктов к нам домой, когда необходимо сделать очередное ТО автомобиля, когда заканчиваются страховки и какие страховые продукты мы предпочли в прошлом периоде.

Все это, с одной стороны, делает жизнь легче и позволяет сосредоточиться на самом важном — семье, работе, хобби, общении с друзьями и родственниками, с другой стороны, еще больше ускоряет темп нашей жизни. Если вдуматься, мы экономим огромное количество времени на дороге, очереди, ожидании консультантов и оформлении документов в страховых компаниях, посещении всевозможных центров, предоставляющих различные государственные услуги.

Отдельно стоит отметить возможности для ИТ-специалистов, разработчиков различных приложений.

²⁰ <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>

1.2. Данные для науки

Наше время породило уникальное явление, позволяющее различным научным течениям проверить и скорректировать основополагающие теории и результаты открытий прошлого. Это возможность обратиться к самому ценному и честному архиву знаний о человечестве — данным, формирующим описание человеческой природы в размерах и формах, поражающих даже самое развитое воображение. Еще никогда научному сообществу не был доступен такой объем знаний о жизни, экономике, потреблении, передвижении и любых других сферах человеческой деятельности. Вместе с тем накопленный объем информации в частных компаниях, государственных, медицинских, образовательных и социальных институтах позволяет проводить уникальные по своим масштабам и качеству результатов исследования.

Нравится нам это или нет, но информация играет все более важную роль в жизни каждого из нас, и эта роль будет становиться значительнее. Сейчас в газетах встречаются целые разделы, полностью посвященные данным. В компаниях есть группы, единственная задача которых — анализ собранных данных. Инвесторы дают десятки миллионов долларов стартапам, если те могут собрать и сохранить большие объемы данных²¹.

Знания и инструменты работы с данными позволяют формировать эффективные и точные прогностические модели даже в таких неожиданных областях, как избирательный процесс, что ранее невозможно было представить ни в социологии, ни в политологии.

Приведем несколько цитат из книги Сета Стивенса-Давидовица «Все лгут»: «...например, информация о том, кто на самом деле будет принимать участие в выборах. Больше половины граждан, которые не голосуют, говорят исследователям, проводящим опросы непосредственно перед выборами, что они намерены пойти голосовать, что искажает оценку явки, в то время как данные о поиске в Google по фразам “как голосовать” или “где голосовать” за неделю перед выборами могут более точно предсказать, где предполагается большая активность на избирательных участках.

Можно даже узнать, за кого они пойдут голосовать. Мы со Стюартом Гэбриэлом, профессором университета штата Калифорния (Лос-Анджелес), нашли удивительную подсказку. Большой процент поисков, связанных с выборами, содержат запросы с именами обоих кандидатов. Во время выборов 2016 года, когда

²¹ Stephens-Davidowitz, S. Everybody lies. Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are. 2017. P. 40.

соперничали Дональд Трамп и Хиллари Клинтон, некоторые люди делали запрос: “выборы: Трамп — Клинтон”. Другие искали: “Клинтон — Трамп, дебаты”. По сути, 12% поисковых запросов со словом “Трамп” включали и слово “Клинтон”. Более четверти поисковых запросов с фамилией Клинтон также содержали и фамилию Трамп. Мы обнаружили, что эти, казалось бы, нейтральные поиски могут дать нам некоторые подсказки о том, какого кандидата человек поддерживает. Как? Все зависит от порядка, в котором кандидаты появляются в запросе. Наши исследования показывают, что человек со значительно большей вероятностью поставит имя кандидата, которого он поддерживает, первым в поисковом запросе, содержащем имена обоих кандидатов»²².

«Даже если вы обманываете самого себя, Google способен узнать правду. За пару дней до выборов вы и некоторые из ваших соседей можете считать, что обязательно пойдете на избирательный участок и проголосуете. Но если ни вы, ни они не искали информацию о том, как и где проголосовать, специалисты по поиску и обработке данных вроде меня могут сказать: явка в вашем районе будет низкой»²³.

1.3. Данные для бизнеса

Отдельно следует упомянуть так называемую цифровую трансформацию — скоростное и основательное изменение жизненных процессов посредством цифровизации. Эти изменения зачастую до неузнаваемости меняют процессы на предприятиях и в государственном управлении, которым не один десяток лет, но результаты подобных трансформаций поражают. Как 30 лет назад сказал генеральный директор FedEx Фред Смит (Fred Smith), «информация о посылке так же важна, как и сама посылка»²⁴.

Существует несколько ключевых факторов, играющих важнейшую роль в процессе цифровой трансформации.

- Основой решений для цифровой трансформации должны стать платформенные решения.

²² Stephens-Davidowitz, S. Everybody lies. Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are. 2017. Pp. 21–22.

²³ Stephens-Davidowitz, S. Everybody lies. Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are. 2017. P. 138.

²⁴ FedEx Marks 40th Anniversary with Community Service, Eye to the Future <https://newsroom.fedex.com/newsroom/fedex-marks-40th-anniversary-with-community-service-eye-to-the-future>

- Центральным местом процесса цифровой трансформации должны быть данные и процессы, позволяющие их использовать.
- Все сотрудники компании, работающие с данными, будь то юристы, менеджеры по продажам или специалисты по логистике, должны быть «владельцами данных», вовлеченными в процесс цифровой трансформации.
- Основные предметные области должны быть основой трансформации, например клиент и всевозможные данные, относящиеся к нему.
- Должен быть обеспечен цифровой профиль основных данных: чем их больше, тем лучше.
- Ключевые менеджеры должны иметь доступ к различным срезам данных для формирования нововведений и творчества.
- При достаточном уровне качества данных искусственный интеллект становится эффективным инструментарием.
- Необходимы метрики стоимости работы с данными (стоимость хранения, стоимость использования данных, эффекты от работы с данными).
- Цифровая трансформация — это процесс, увеличивающий эффективность деятельности по мере его развития.
- Цифровая трансформация без ее глубинного понимания может стать самым дорогим проектом компании и привести к ее банкротству.

Цифровые технологии можно без преувеличения назвать фактором революционного изменения всех устоявшихся бизнес-процессов и существенного повышения эффективности деятельности. Организации, которые проигнорируют необходимость цифровой трансформации или допустят ошибки на стадии проектирования этого процесса, будут неэффективными и, возможно, будут вынуждены прекратить деятельность.

Тут стоит отметить, что мультинациональные корпорации одними из первых оценили возможности использования данных как в повседневных, так и в глобальных процессах трансформаций, поняв, что они все более и более зависимы от данных организации. Вместе с тем большинство современных мультинациональных компаний не смогли бы функционировать без единой ИТ-архитектуры и единой среды обеспечения данными. Практически все процессы крупнейших компаний мира основаны на эффективных инструментах работы с данными.

Не секрет, что крупные международные корпорации в процессе своего развития зачастую создавали и приобретали компании в разных уголках мира, имеющих свою специфику в области информационных технологий и, соответственно, довольно разрозненные информационные системы относительно друг друга. В результате компания приобретала абсолютно неструктурированные и неупорядоченные данные, что существенно осложняло работу всех ИТ-систем в рамках

единой организации. В этих условиях важный инструмент работы с данными в мультинациональных компаниях и организациях, повышающий эффективность деятельности компаний и учитывающий широкий портфель продуктов, — внедрение единой платформы управления данными. Вот как в общих чертах выглядит подобная единая платформа управления данными на примере крупной международной страховой группы (рис. 1.2).

Одной из самых заметных тенденций в области работы с данными в корпоративном мире в последнее время становится построение бизнес-процессов компаний на основании данных. Формируется «дата-центричный» подход, что во многом обусловлено пониманием эффекта использования данных для различных направлений деятельности компаний. Сейчас уже принимается за правило выделять и рассматривать различные информационные системы как источники данных, обогащать эти данные и формировать их ценность. Проектирование архитектуры данных становится отдельным направлением деятельности организаций. Самые важные бизнес-процессы перестраивают именно на основе этой архитектуры (рис. 1.3).

Только после формирования архитектуры данных компании формируют функциональную и ИТ-архитектуру, описывая логическую и физическую модель работы с данными в организации. Все это приводит к тому, что фактическая функция ИТ становится вспомогательной по отношению к функции управления данными (рис. 1.4).

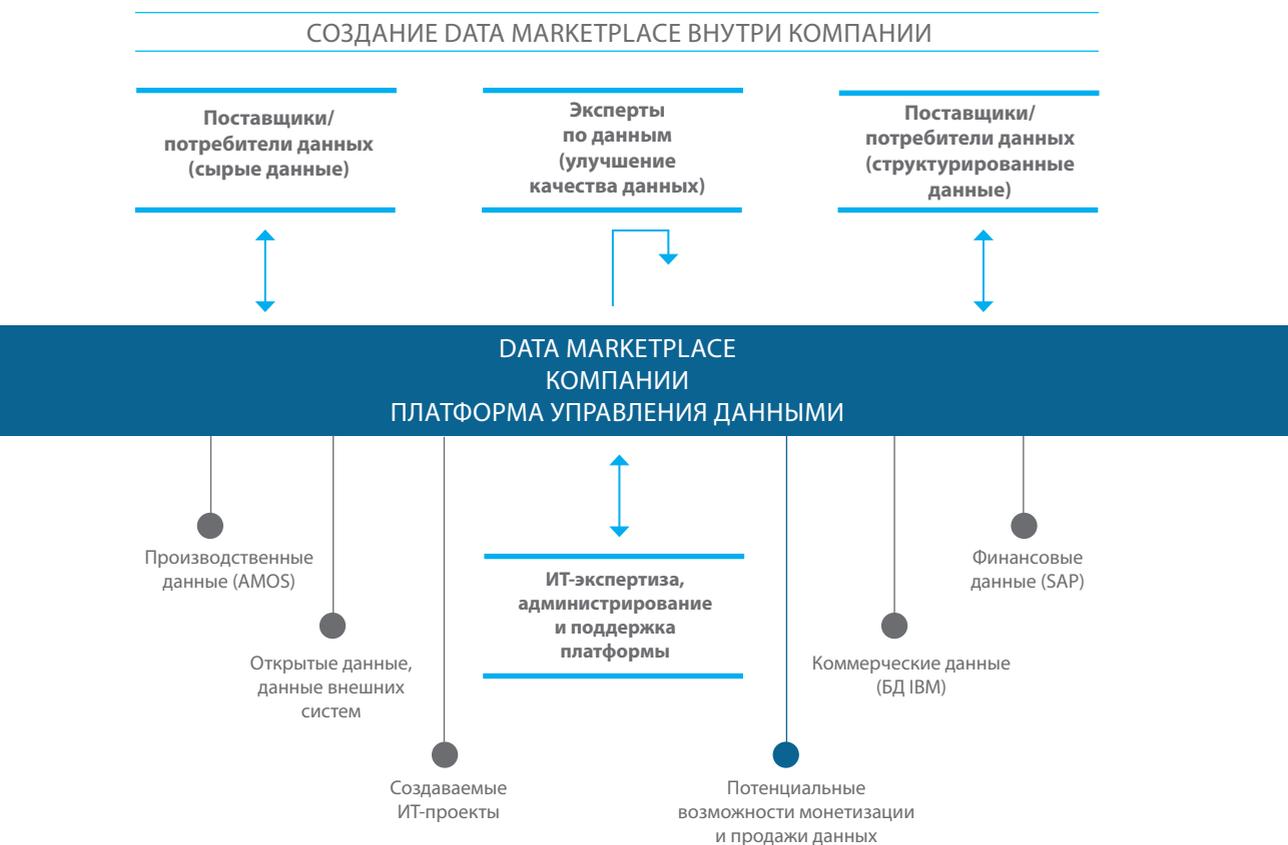
Чтобы оценить масштабы эффективности внедрения цифровых технологий, посмотрим на исследование, проведенное российской компанией «Юнидата». Оно показывает, что одним из самых востребованных направлений оптимизации бизнес-процессов с использованием решений для управления данными являются решения класса «Единый каталог товаров, работ и услуг» (ЕКТРУ), позволяющие повысить эффективность закупок для компаний и организаций. Те, кто пренебрегает подобными возможностями, с большой вероятностью станут неконкурентоспособными по отношению к коллегам, уделяющим должное внимание данному направлению деятельности.

Внедрение ЕКТРУ позволяет компании обеспечить прозрачность процесса закупок, вести в одном месте закупаемую номенклатуру, иметь полную информацию о товарах/услугах, их технических характеристиках, поставщиках, стоимости, аналогах, условиях и сроках гарантийных обязательств, стоимости и условиях поставки, наличии изделий на складах и в производстве и многих других параметрах. Один из следующих шагов внедрения ЕКТРУ — создание различных интерпретаций личного кабинета, позволяющего поставщикам предоставлять всю необходимую информацию о поставляемых продуктах и услугах непосредственно



Рис. 1.2.

Единая платформа управления данными
на примере крупной международной страховой группы

**Рис. 1.3.**

Data Marketplace компании

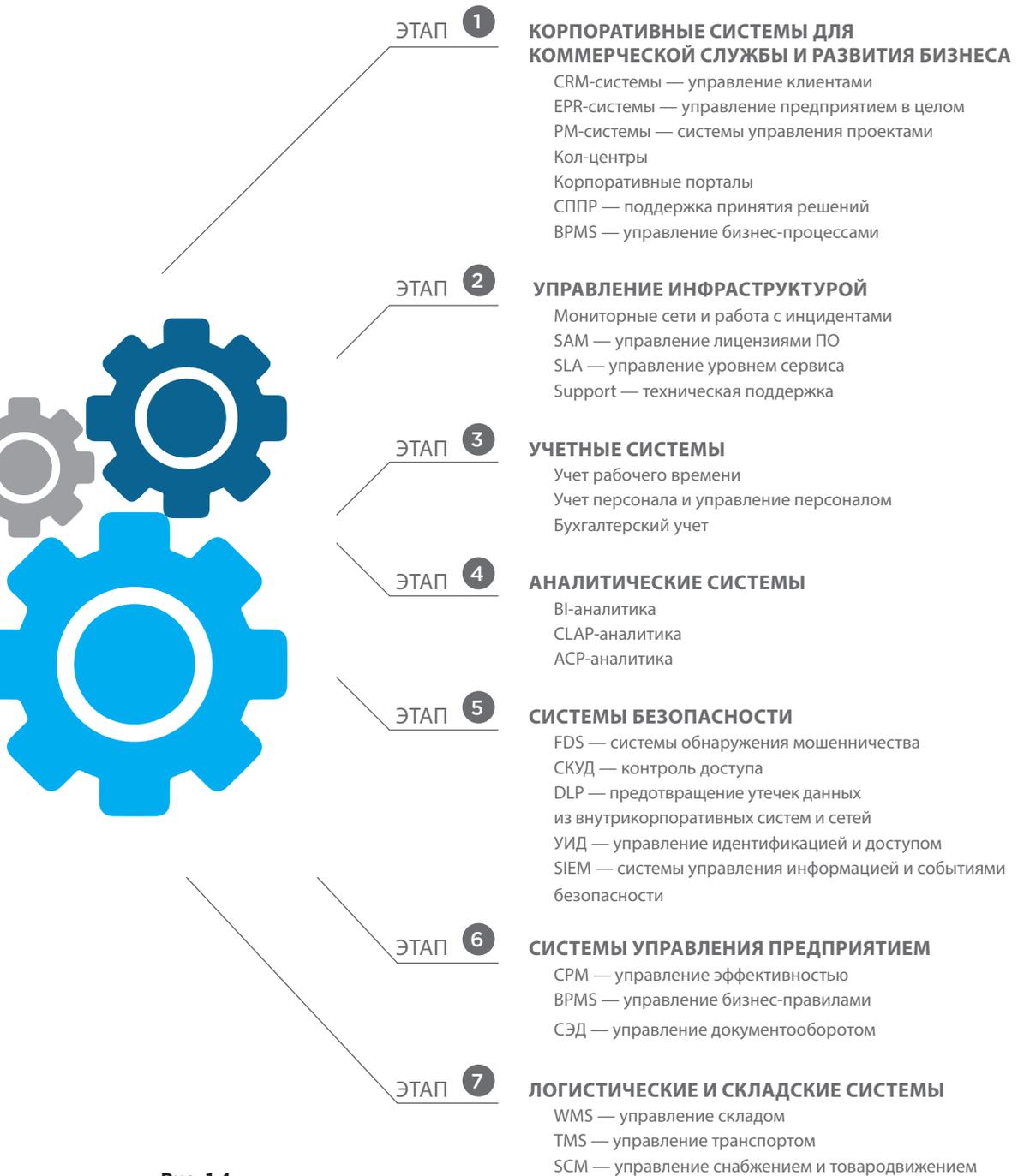


Рис. 1.4.
ИТ-ландшафт компании

заказчику, а тому в свою очередь иметь актуальную информацию для обеспечения своей деятельности и оптимизации бизнес-процессов.

Основные результаты внедрения

Единого каталога товаров, работ и услуг

- Создание и ведение параметрических классификаторов, обеспечивающих возможность формирования специализированного атрибутивного состава для каждого класса товаров, работ и услуг.
- Применение единого словаря терминов при ведении каталожных описаний.
- Возможность создания и ведения стандартных форматов описания (СФО) предметов потребления, на основании которых формируются каталожные описания конкретных номенклатурных позиций.
- Управление структурой каталога, классификаторами, опорными справочниками и другими объектами ЕКТРУ через пользовательский интерфейс платформы без привлечения разработчиков.
- Гибкая настраиваемая ролевая модель доступа к данным и функциям каталога.
- Реализация настраиваемых процессов согласования внесения изменений в каталог на основе утвержденного регламента ведения ЕКТРУ. Обеспечение автоматизированного контроля за ходом и результатами соблюдения регламента.
- Возможность применения встроенных в платформу управления данными механизмов консолидации, дедубликации, обеспечения качества данных.

Эффекты внедрения Единого каталога товаров, работ и услуг

- Увеличение рентабельности предприятия за счет снижения складских запасов (по экспертным оценкам, от 5 до 25%) и оптимизации резервирования материально-технических ресурсов.
- Оптимизация расходов на закупку материально-технических ресурсов за счет снижения количества избыточных заказов при закупке однотипной продукции (по экспертным оценкам, 3–5%).
- Уменьшение ошибок в корпоративной отчетности, связанных с неактуальностью справочной информации, и, как следствие, повышение качества принимаемых управленческих решений.
- Предотвращение материальных потерь, связанных с закупкой неликвида (по экспертным оценкам, в среднем предприятия теряют из-за приобретения неликвида от 10 до 30% от общего бюджета закупок).

- Предотвращение материальных потерь за счет экономии человеческих ресурсов предприятия благодаря грамотному и качественному обмену данными между отделами, цехами и службами (по экспертным оценкам, от 5 до 25%).
- Достоверная информация о фактическом объеме складских запасов, статистике движения (поступления/расхода) материалов, увеличения необоснованных запасов, замораживания оборотных средств.
- Однозначная идентификация клиентов, повышение лояльности клиентов за счет наличия полной и достоверной информации о профиле клиентов, наличие информации о неблагонадежных партнерах/клиентах, повышение эффективности привлечения новых клиентов и групп клиентов за счет корректных фокус-групп.
- Повышение качества формирования закупок за счет накопленной статистики по товарам и товарным группам, поставщикам, параметрическим показателям товаров и услуг.
- Снижение затрат за счет отсутствия ошибок в адресах доставки товаров и корреспонденции.
- Сокращение поставок более дорогостоящих изделий с избыточными свойствами. Увеличение числа поставок более дешевых аналогов и заместителей, обладающих аналогичными свойствами оригинала, в том числе и при осуществлении импортозамещения.
- Снижение затрат на закупки при улучшенной консолидации закупочных партий.
- Повышение контроля в сфере закупок.
- Автоматизация всего закупочного процесса (подготовки документации для проведения закупки, определения начальной максимальной цены контракта, закупочной процедуры и последующих процессов оформления документов в строгом соответствии с закупочной процедурой).
- Снижение стоимости владения изделиями за счет подробного описания состава изделия и применимости ремонта и использования комплектующих при ремонте с учетом сроков гарантии каждой детали.

1.4. Данные для торговых сетей

Дальнейшее использование данных и технологий работы с ними изменит не только бизнес-процессы, но и сами отрасли.

Работа с данными позволит существенно повысить эффективность бизнес-процессов, снизить стоимость продукции при использовании унифицированных данных по продукции, создать единообразие продукции, увидеть наличие данных по аналогам и параметрической информации о товарах, статистику о закупаемых изделиях и наличии продукции на складах, объединять данные из информационных систем, оперирующих информацией о гарантийных обязательствах, сроках ремонта и обслуживания изделий.

Вместе с тем в будущем большинство организаций будут вынуждены кардинально перестроить свою деятельность.

Уже сегодня большой процент заказчиков покупает товары повседневной необходимости через интернет-магазины или приложения, что приводит торговые компании к необходимости оптимизации (зачастую закрытия) торговых сетей или отдельных торговых точек. Это также связано с тем, что покупатели довольно консервативны в своих предпочтениях и приобретают одни и те же товары, потребительские свойства которых им известны. Можно с уверенностью сказать, что спектр товаров, приобретаемых онлайн, будет расширяться. Производителям и торговым сетям эти изменения позволят иметь более точную информацию о своих клиентах, их предпочтениях и пожеланиях.

Тенденция распространится на товары более длительного использования, например одежду, электронику, автомобили. Это приведет к изменениям работы торговых сетей и их взаимоотношениям с производителями товаров, появлению новых экономических схем в сфере торговли. Модель работы изменится: в торговых сетях появятся нескольких флагманских магазинов, основной целью которых будет не офлайн-продажа, а знакомство покупателей с товарами и их потребительскими свойствами. Указанная модель обеспечит огромное преимущество торговым сетям, которые первыми придут к такому формату и фактически заставят остальные сети последовать своему примеру, поскольку экономия на капитальных и операционных затратах по сравнению с классическими торговыми сетями будет гигантской благодаря отсутствию затрат на аренду помещений, оборудования, коммунальные платежи, логистику, хранение товаров, зарплаты персонала в большом количестве торговых точек и других издержек. Эта модель уже получила широкое распространение: покупатели посещают магазины/салоны исключительно чтобы увидеть искомое воочию и получить консультации о потребительских свойствах и функционале товаров офлайн, но приобретают понравившийся товар в интернет-магазине, где им предлагают более выгодную стоимость и удобство доставки.

«В середине января 2020 года производитель аудиотехники Bose объявил о закрытии 119 магазинов по всему миру. Причиной такого решения Bose назвал