

Содержание

<i>Об авторе</i>	11
<i>Предисловие</i>	13
01 Мир данных	15
Данные: мир, в котором мы живем	16
Данные: нехватка навыков	21
Данные: в чем причина нехватки навыков?	24
Данные: что дальше?	30
Краткое содержание главы	31
02 Четыре уровня аналитических методов	33
Данные и их анализ — целых четыре уровня?	33
Четыре уровня аналитических методов	35
Уровень 1: дескриптивные (описательные) аналитические методы	37
Уровень 2: диагностические аналитические методы	40
Уровень 3: предиктивные (предсказательные) аналитические методы	43
Уровень 4: прескриптивные (предписывающие) аналитические методы	47
Примеры использования четырех уровней аналитических методов в реальной жизни	48
Краткое содержание главы	52

03	Определение дата-грамотности.....	53
	Элемент 1: чтение данных.....	55
	Элемент 2: работа с данными	58
	Элемент 3: анализ данных.....	63
	Элемент 4: общение на языке данных.....	66
	Краткое содержание главы.....	69
04	Зонтик дата-грамотности	71
	Стратегия в сфере данных и аналитики	73
	Дата-грамотность и обработка данных	74
	Дата-грамотность и визуализация данных.....	78
	Дата-грамотность и топ-менеджмент.....	85
	Дата-грамотность и культура	87
	Дата-грамотность и качество данных.....	89
	Дата-грамотность и управление данными	90
	Дата-грамотность, этика и законодательство	91
	Краткое содержание главы.....	95
05	Чтение и общение на языке данных	97
	Чтение данных.....	101
	Свободное владение данными.....	105
	Словарь данных	108
	Стратегия чтения данных и свободного владения данными	110
	Пример из жизни организации.....	111
	Краткое содержание главы.....	113
06	Связь дата-грамотности с четырьмя уровнями аналитики	115
	Дата-грамотность и дескриптивный анализ	116
	Дата-грамотность и диагностический анализ.....	121
	Дата-грамотность и предиктивный анализ	130

Дата-грамотность и прескриптивный анализ	134
Дата-грамотность и четыре уровня аналитики: готовая мозаика	139
Краткое содержание главы	139
07 Стадии обучения дата-грамотности	141
Роль руководства в обучении дата-грамотности . . .	142
Роль общей аналитической стратегии и обучения дата-грамотности	145
Схема и стратегия обучения дата-грамотности . . .	146
Обучение четырем элементам дата-грамотности . .	152
Обучение и культура дата-грамотности	156
Краткое содержание главы	161
08 Три «С» дата-грамотности	163
Curiosity — любопытство	164
Creativity — творческий подход	174
Critical thinking — критическое мышление	181
Общение на языке данных	184
Краткое содержание главы	185
09 Принятие решений, подкрепленных данными	187
Схема принятия решений, подкрепленных данными	189
Степень 1. Спросить	190
Степень 2. Получить	193
Степень 3. Проанализировать	196
Степень 4. Интегрировать	198
Степень 5. Решить	204
Степень 6. Выполнить итерацию	206
Краткое содержание главы и пример	208

10 Дата-грамотность и стратегия в сфере данных и аналитики.....	211
Культура, основанная на данных	212
Бизнес-аналитика.....	215
Искусственный интеллект	218
Машинное обучение и алгоритмы.....	221
Большие данные	223
Внутренняя аналитика.....	225
Облачные сервисы.....	228
Периферийная аналитика.....	229
Геоаналитика	230
Краткое содержание главы.....	231
 11 Ваше личное путешествие в мир данных: первые шаги	 233
COVID-19, данные и аналитика	234
Рецепты.....	238
Проактивная и реактивная аналитика	241
Начинайте с основ.....	243
Данные и аналитика: геймификация	245
Найдите то, что вам интересно, и займитесь этим.....	246
Найдите свои «почему».....	247
Краткое содержание главы.....	249
 <i>Примечания</i>	 <i>251</i>

Об авторе

Джордана Морроу называют крестным отцом дата-грамотности. Он один из пионеров движения за дата-грамотность и постоянно стремится помогать руководителям и сотрудникам различных организаций раскрывать потенциал данных и аналитики. Вне мира данных Джордан Морроу — счастливый семьянин, у него пятеро детей. Кроме того, он увлеченный любитель бега по пересеченной местности и ультрамарафона.

Предисловие

«Есть три типа лжи — ложь, наглая ложь и статистика».

В этой цитате, которая приписывается самым разным людям, в частности Марку Твену, отражена вся наша сегодняшняя жизнь. Мы живем в мире, где данные — статистика и многое другое — находятся в свободном доступе. Они что-то нам рассказывают, мы черпаем из них информацию. Легко ли нам разобраться во всех данных, которые регулярно к нам поступают? Увы, зачастую они используются и интерпретируются совершенно неправильно. Так можем ли мы что-то с этим сделать?

Мое путешествие в мир дата-грамотности официально началось в июне 2016 года, но первые идеи по этой теме появились у меня еще раньше. Меня называли крестным отцом дата-грамотности и Главным Ботаником (второе прозвище мне и правда нравится). Я долго развивал и совершенствовал свои представления. В этой книге я хочу поделиться своими наблюдениями и помочь вам совершить путешествие в мир данных и аналитики.

Мой подход, возможно, отличается от традиционного, и кое-что в книге покажется вам неожиданным. Для большинства из нас данные стоят далеко не на первом месте в списке самых интересных тем на свете, но лично я ими просто очарован. Мне бы очень хотелось донести до вас правильное представление о мире данных и показать вам, что вы можете сделать, чтобы стать частью этого удивитель-

ного мира. Спасибо за то, что решили уделить мне время. Надеюсь, что смогу зажечь в вас искру любопытства и любви к данным: возможно, они смогут изменить вашу жизнь к лучшему, как это произошло со мной.

Мир данных

Задумывались ли вы когда-нибудь о том, какое оно — будущее? Каким, например, станет рынок труда? Неужели роботы и техника действительно монополизируют его и отберут у нас работу? Какие нас ждут открытия и когда уже можно будет заказать летающий автомобиль? Будущее всегда туманно, и мы пока не знаем, какие именно изобретения будут определять наше завтра. Бесспорно, в будущем появятся такие профессии, о которых мы пока даже не догадываемся. Но при всей этой неопределенности можно назвать одну вещь, которая точно станет частью будущего и уже присутствует в нашей жизни, — это данные. Будущее обещает нам множество изобретений и удивительных новых профессий, но этого мы только ждем — а вот с властью данных уже столкнулись. И это навсегда.

Мир данных поразителен, огромен и дает нам всем безграничные возможности для развития. Слишком долго попытки людей и организаций добиться успеха с помощью данных были неубедительными. Это неправильно, и мы должны помочь каждому научиться пользоваться всеми возможностями, которые предоставляет этот неисчерпаемый источник.

Данные называют новой нефтью, говорят, что они необходимы как воздух, — избитых штампов и дешевых преувеличений хватает. Но на самом деле данные — это актив, который при правильном использовании может помочь

всему человечеству двигаться вперед. С ними легче получать и усваивать знания, они не только готовят нас к будущему, но и закладывают прочное (и в то же время гибкое) основание для него. Такое комплексное понимание мира данных в наше время просто необходимо.

Данные: мир, в котором мы живем

Не секрет, что мы живем в мире, поглощенном технологиями и данными. Вряд ли можно пройти по улице какого-нибудь крупного города вроде Лондона или Нью-Йорка, не увидев людей, уткнувшихся в телефоны, вместо того чтобы смотреть на удивительный мир вокруг и на его обитателей. Убедитесь сами: в следующий раз, выйдя на улицу, посчитайте, сколько людей глядят в экраны, а сколько — смотрят по сторонам. Можете даже прибавить к последним тех, кто болтает друг с другом или приветствует кого-то. Да и сами не забывайте смотреть по сторонам, чтобы избежать неприятных сюрпризов: можно споткнуться о бордюр или не увидеть машину.

В последние лет пятьдесят, а в особенности в последние лет тридцать — с приходом в нашу жизнь интернета, персональных компьютеров, смартфонов и т. д. — мы стали свидетелями невероятного прогресса в развитии технологий и работе с данными. Задумайтесь об этом прогрессе хотя бы на минуту. Вселенная существует примерно 13,8 миллиарда лет¹, Земля — примерно 4,5 миллиарда², а мы говорим всего лишь о тридцати–пятидесяти годах. Но за эти несколько десятилетий у нас на глазах персональные компьютеры и сотовые телефоны стали обычным делом. В масштабах вечности это миг. Но сейчас, в наше время, нам уже трудно представить жизнь без этих технологических новшеств. И каждое из таких устройств генерирует данные. А интернет? Он еще «моложе» — проник в нашу жизнь в начале

1990-х³ и повсеместно распространился лишь к началу 2000-х. Но теперь это неотъемлемая часть нашей жизни — и на работе, и в быту. Чем больше становилось персональных компьютеров, тем быстрее развивались технологии: темп роста не падает и по сей день. Мы постоянно наблюдаем инновации, видим, как развиваются самые разные аспекты цифрового мира. Все это напрямую влияет на нашу жизнь. И в первую очередь — учитывая тему нашей книги — это влияет на работу с данными и на их мощь.

Вернемся к тем временам, когда интернет стал массовым достоянием и начал использоваться более активно. Организации, учебные заведения и отдельные люди теперь живут и работают иначе. И наша жизнь благодаря ему не просто изменилась, но и значительно улучшилась: интернет позволяет совершенствоваться, учиться и развиваться с невиданной скоростью. Когда к интернету стало можно подключить ПК, а теперь и смартфоны, мы смогли принести всю компьютерную мощь к себе домой. Теперь мы можем почти мгновенно получить столько же информации, сколько во всей «Британской энциклопедии»: больше сведений на нас вывалит только торговый агент, постучавшийся к нам в дверь. Мы гораздо быстрее получаем ответы на вопросы. Из стремления найти ответы и вырос Google — более того, это слово (google) даже получило статус глагола в словарях!⁴

Прогресс компьютерных технологий привел к развитию электронной коммерции и появлению Amazon и других компаний, которые смогли изменить потребительские привычки и захватить рынок. У нас на глазах лопнул пузырь доткомов: сначала стоимость IT-компаний, не производящих никакой продукции, вдруг стала огромной — а затем они обанкротились. Примером может служить компания Pets.com. Она возникла в 1998 году и закрылась в 2000-м⁵. Это время благополучно миновало, появилось множество других разнообразных сайтов, а затем началась эпоха социальных

сетей. Соцсети открыли двери к профилям потребителей и публичным данным, так что самая разная информация о нас — от селфи и фото ужинов до товаров, которые нам нравятся, — стала доступна всем желающим, причем как отдельным людям, так и бизнесу. Ведь так приятно, когда контекстная реклама решает за нас, что мы хотим сегодня на ужин!

Наряду с соцсетями, предоставляющими массу интересных данных, в 2000-х появилась и новая технология, которая быстро вышла на передний план в области связи и сбора данных, — интернет вещей, или IoT. Что это такое? Все очень просто — это связь всего со всем через интернет. Возьмем, к примеру, датчики в автомобиле или самолете, которые собирают данные обо всем, что происходит в моторе или других частях механизма. Но знаете ли вы, когда на самом деле появился интернет вещей? Думаете, в 2000-х? А вот и нет, хотя многие об этом не знают. Термин появился в 1999 году, но одним из первых примеров интернета вещей можно считать давно знакомый нам... торговый автомат Coca-Cola. Все мы не раз видели такие автоматы, но конкретный автомат, о котором идет речь, находился в Университете Карнеги-Меллона. Он позволял покупателю связаться с холодильником через интернет и выяснить, холодный ли лимонад⁶. Об использовании данных для принятия более взвешенных решений благодаря «связи всего со всем» (то есть интернету вещей) люди задумывались уже давно: как собранные данные могут помочь нам делать правильный выбор в личной и профессиональной жизни? Достаточно вспомнить такие компании, как Amazon или Netflix: насколько часто они собирают наши данные, чтобы «рекомендовать» то, что нам может быть нужно? Очень часто... И знаете что? Эти рекомендации нередко оказываются верными!

Хотя интернет вещей берет начало в 1980-х, в силу он входит только теперь. Представьте, к примеру, любителя

бегать ультрамарафоны. Еще несколько лет назад практически не было способов сбора данных, которые могли бы помочь бегуну развиваться... а сейчас они есть. Но нужны ли нам на самом деле часы, которые показывают буквально все аспекты бега — от уклона трассы до скорости на разных участках? Сегодня бегун может получить от своих часов больше информации, чем ему, вероятно, требуется, — за все время бега она накапливается тоннами! Конечно, данные с часов потом интересно рассматривать и обсуждать, но ключевой момент заключается в том, что благодаря «связи всего со всем» и технологическому прогрессу современного мира генерируется все больше и больше данных о нашей жизни. Вы наверняка можете привести и другие примеры того, как работает интернет вещей. Но давайте вспомним именно те случаи, когда прогресс действительно определяет нашу жизнь и помогает нам.

Одна из компаний, культивирующих сетевое взаимодействие, цифровизацию и производство данных, — знаменитый Rolls-Royce. Сейчас это не просто мощная инженерная компания, производящая замечательные двигатели. Rolls-Royce — эффективная организация, управляемая на основе данных: она использует технологию интернета вещей и возможности связи для предоставления и получения данных, которые стали ценным активом компании⁷. Один из примеров использования Rolls-Royce возможностей данных — прогностический метод мониторинга техобслуживания двигателей⁸. С помощью датчиков, собирающих данные, Rolls-Royce успешно предсказывает возможные проблемы с авиационными двигателями и гарантирует, что они не откажут прямо в воздухе. Другой пример того, как «связь всего со всем» и интернет вещей улучшают нашу жизнь, можно найти в сфере здравоохранения. Использование данных, генерируемых оборудованием для физиотерапии, позволяет разрабатывать более эффективные программы

для тех, кто нуждается в физиотерапевтическом лечении⁹. Учитывая растущую стоимость медицинских услуг, такие более персонализированные программы будут помогать людям следить за здоровьем и реже попадать в больницу.

Еще одна сфера, в которой анализ данных способствует росту и развитию, — это спорт. Помните фильм «Человек, который изменил все» с великолепным Брэдом Питтом в главной роли? В нем говорится, что данные и их анализ могут очень серьезно влиять на судьбу спортивных команд, помогая добиваться побед. Речь, конечно, о бейсболе, но примеры работы с данными можно найти и в баскетболе, включая НБА (Национальную баскетбольную ассоциацию) — а это уже совсем другой масштаб. У большинства команд НБА (возможно, даже у всех) есть свои аналитики и эксперты по данным. Их задача — обнаруживать тенденции и закономерности в данных, которые они собирают: например, поиск недооцененных игроков и повышение их стоимости для продажи или обмена. Кроме того, команды НБА используют данные и технологии для отслеживания уровня утомляемости и качества сна своих игроков, что позволяет корректировать режим тренировок, предотвращать травмы и т. д. НБА даже проводит свой собственный ежегодный «хакатон», чтобы найти и привлечь новых талантливых аналитиков. Знаете ли вы, что количество трехочковых бросков в лиге выросло не в последнюю очередь благодаря анализу данных?¹⁰

Но довольна не только НБА. Всю мощь данных заключают и вещи, которыми мы пользуемся в быту: смарт-часы, смартфоны, посудомоечные машины, холодильники, системы обогрева и кондиционирования, автомобили и прочие транспортные средства... и многое, многое другое. В прочих сферах нашей жизни данные тоже производятся в невероятном количестве. Только представьте себе, сколько информации проходит через соцсети, торговые сайты вроде Amazon и eBay, платежные системы и т. д. Невероятные цифры. Давайте

взглянем на статистику, предоставленную Всемирным экономическим форумом. В 2019 году¹¹:

- ежедневно публиковалось 500 миллионов твитов;
- было отправлено 294 миллиарда электронных писем;
- каждый подключенный к интернету автомобиль произвел 4 терабайта данных;
- к 2025 году каждый день в мире будет производиться ориентировочно 463 эксабайта данных.

Если бы мы до сих пор смотрели DVD, а не Netflix, то такое количество данных было бы эквивалентно 212 765 957 дискам! Какое-то немыслимое число. Что это означает для нас? Это означает, что мы за всю жизнь не сможем просмотреть все DVD с данными, произведенными всего лишь за один день. И что нам делать со всей этой информацией?

Конечно же, организации и частные лица могут воспользоваться этим удивительным богатством. Согласитесь, любая организация ухватится за эту идею — как можно применить знания, находящиеся в ее распоряжении, для принятия взвешенных, подкрепленных данными решений. Однако в реальности все оказывается совершенно не так. Исследования демонстрируют нам истинное положение дел: в мире данных налицо чудовищная нехватка навыков, что и мешает организациям успешно применять этот ценнейший ресурс.

Данные: нехватка навыков

Чтобы понять причины такой нехватки, необходимо иметь представление о том, как обстоят дела с навыками дата-грамотности. Аналитическая компания Qlik помогает разобраться в общей картине и указывает на конкретные аспекты, где могут наблюдаться пробелы. Одно из исследований (август 2017-го — февраль 2018 года) обрисовало современное

состояние сферы дата-грамотности и необходимых навыков, а также предоставило ценнейшую информацию относительно уровня квалификации и уровня комфорта людей при взаимодействии с данными¹². Результаты исследования просто поражают. Выяснилось, что всего 24% людей, ответственных за принятие важных решений, чувствуют себя уверенно, работая с данными. Всего лишь 24%. Это крайне малое количество для тех, кто отвечает за будущее организации. В ряде случаев именно им приходится принимать решения именно на основании данных. Как же можно доверять таким решениям, принятым при нехватке навыков?

Когда организации определяют стратегию в области данных и их анализа (если это вообще происходит, потому что тенденции говорят об обратном), подразумевается, что команда топ-менеджеров должна сформулировать такую стратегию, задать направление и создать план использования данных для развития бизнеса. А знаете, какова доля руководителей высшего звена, уверенно пользующихся данными? Согласно результатам того же исследования — 32%! То есть приблизительно одна треть, причем, судя по всему, это субъективная оценка и она завышена: на самом деле еще меньше топ-менеджеров способны правильно использовать данные и обладают подлинной дата-грамотностью. Итак, если руководители высшего звена определяют представление организации о работе с данными и ее стратегию в этой области, то получается, что лишь 24% из них принимают правильные решения. Может ли стратегия быть эффективной, если за ее реализацию отвечают люди, не уверенные в своем умении обращаться с данными? Надеюсь, вы уже осознали масштабы проблемы нехватки навыков.

Интересно, что можно сказать о молодежи, которая только сейчас выходит на рынок труда? Все то же исследование Qlik выявило, что среди молодых людей — от 16 до 24 лет — уверены в своих способностях правильно обращаться

с данными лишь 21%! Возникает закономерный вопрос: почему молодые люди настолько не уверены в своих способностях? Разве они не даны им от природы или не формируются на более раннем этапе? Это очень интересный вопрос, требующий более пристального рассмотрения с демографической точки зрения. Люди, которым в 2017 году (когда проводилось исследование) было от 18 до 24 лет, родились в цифровом мире, когда интернет и персональные компьютеры уже стали частью повседневной жизни. Молодежь выросла в интернете, в соцсетях... но значит ли это, что ее учили правильно пользоваться данными и анализировать их? Да, она разбирается в компьютерах, но при этом не разбирается в данных.

Итак, в целом только один из пяти участников исследования уверен в своих навыках обращения с данными, так что для руководителей организаций должен быть очевиден огромный пробел, который необходимо заполнить. Здесь и кроется корень проблемы: если организации хотят извлечь выгоду из данных и аналитики, но при этом нужных специалистов критически не хватает, то как, собственно, извлечь эту выгоду? И какое влияние нехватка дата-грамотности и уверенности в умении обращаться с данными оказывает на организации? Не отражается ли этот недостаток на прибыли?

Влияние человеческого фактора на недостаток знаний о данных трудно переоценить. В исследовании 2019 года было выявлено, что лишь 32% опрошенных топ-менеджеров утверждают, что способны извлечь измеримую пользу из данных, и 27% — что их проекты в сфере данных и аналитики «дают им применимые на практике знания»¹³. Конечно же, это тоже следствие недостаточной дата-грамотности. Когда вспоминаешь, что миллионы, а то и миллиарды долларов вкладываются в аналитические проекты, ПО и технологии, связанные с данными, то поневоле задумываешься, сколько из них потрачено впустую. И если среди людей в целом лишь каждый пятый уверен в своих навыках обращения

с данными, а топ-менеджеры не осознают ценность этих навыков, то потери действительно огромны.

Какое же влияние нехватка дата-грамотности оказывает на нас, отдельно взятых людей? В упомянутом исследовании приведена лишь общая количественная оценка соответствующих навыков. Но если задуматься о сути исследования, то становится ясно: из-за отсутствия необходимых навыков люди совершенно не справляются с новыми технологиями и потоком данных. Более трети участников исследования признались: если бы перед ними стояла задача, которую можно решить с помощью данных, они предпочли бы искать иные способы ее решения. А целых 14% в таком случае постарались бы вообще уклониться от этой задачи. Картина всеобщей перегруженности данными представляется еще более убедительной, если обрисовать ее в виде рабочих часов, потерянных в результате неумения пользоваться данными и технологиями: более 5 рабочих дней (43 рабочих часа) на сотрудника за год. Сколько это в реальном денежном выражении? Очень немало! Согласно исследованию, для американской экономики эти потери составляют около 100 миллиардов долларов в год. По-вашему, это огромная цифра? Тогда давайте задумаемся, почему люди не справляются с данными и каковы причины такой нехватки навыков.

Данные: в чем причина нехватки навыков?

У такого огромного пробела в знаниях и навыках наверняка есть какие-то причины. Каковы же они? Очевидно, что факторов множество: от недостатков в системе образования до проблем с технологиями и программным обеспечением и даже до производства данных как такового. Сейчас мы

подробно рассмотрим эти возможные причины, а вы задумайтесь, как они могли повлиять на вашу карьеру и умение извлекать пользу из данных (будь то на личном уровне или во взаимодействии с людьми, с которыми вы работаете).

Программное обеспечение и технологии

Возможно, вы задумаетесь: как ПО и технологии могут быть причиной нехватки навыков дата-грамотности? Разве все не должно быть наоборот? Разве они не уменьшают этот пробел? Что ж, ответ вроде бы очевиден: да, должно быть наоборот; да, уменьшают. Задача ПО и технологий — помогать нам решать задачи, связанные с данными и их анализом, и получать реальные бизнес-результаты. Они должны дополнять человека, если он достаточно образован и обучен, чтобы ими пользоваться.

Проблема заключается в том, как именно вложения в технологии и ПО осуществляются людьми и организациями. Представьте себе, что вы руководите компанией и пытаетесь выстроить стратегию работы с данными, которая должна помочь вам добиться успеха в эпоху цифровой революции. К вам приходит замечательный торговый агент и заявляет: «Наше новое программное обеспечение предназначено для того, чтобы расширить возможности для решения задач по данным и аналитике». Или обещает: «Наша программа решит все проблемы с данными и их анализом». Словом, чтобы убедить вас купить его ПО, он засыплет вас самыми разнообразными убедительными аргументами. Он покажет вам примеры из практики и данные исследований. Полюбовавшись на все это, вы решите вложить свои деньги в ПО и внедрить его в своей организации. Когда организация обеспечивает всем своим работникам равный доступ к аналитическим программам, это называется «демократизация данных». Открою небольшой секрет. С одной стороны, демократи-

зация данных — именно то, что должна проделать каждая компания. С другой — это проблема. Давайте рассмотрим ее немного подробнее.

Исторически данные принадлежали сфере информации и технологий (или какой-то другой отдельной части бизнеса), так что лишь немногие сотрудники работали с ними, составляли отчеты и делали анализ. Таким образом, организация доверяла получение достоверных, практически применимых результатов этой ограниченной группе сотрудников. По мере того, как компании, предоставляющие инструменты бизнес-аналитики, такие как Qlik или Tableau, прогрессировали и превращались в гиганты разработки ПО, организации стали искать пути широкого распространения и использования данных. Таким образом они демократизировали данные и информацию, рассчитывая получить полезные знания и результаты. Но налицо одна большая проблема: многих ли в школе и университете учили работе с данными и их анализу? В последние годы количество таких людей растет, но что делать с теми, кто не получил соответствующего образования?

Может показаться, будто мы хотим сказать, что демократизация данных — это не решение. Не хотим! Демократизация данных полезна — благодаря инвестициям в данные и аналитику организации раскрывают свой потенциал. Демократизация данных позволяет использовать уникальные таланты и способности сотрудников, получать больше отдачи от вложений в ПО, данные и технологии. Причина, по которой демократизация данных может увеличивать недостаток знаний в организации, — образовательный фундамент ее коллектива. Если от людей, которые не обучались использованию данных и технологий, требуют освоить новое ПО и новые технологии или извлекать пользу из данных и информации, они никак не смогут проделать это эффективно. Как вы думаете, многие ли из них с готовностью (и радостью)

бросятся использовать эти новые технологии? Часто ли вы сами с восторгом соглашаетесь учиться чему-то новому, тому, чем вас внезапно нагружают по работе?

Вернемся к исследованию 2019 года о влиянии человеческого фактора на дата-грамотность и вспомним, какую проблему оно выявило: 36% участников «предпочли бы искать иные методы решения задачи вместо использования данных», а 14% — «постарались бы полностью уклониться от задачи», если бы им нужно было использовать данные. Это прекрасная демонстрация нехватки навыков работы с данными. А те, кто чувствует себя в этой области комфортно и уверенно, с большей охотой станут осваивать технологии. То же исследование показало, что почти три четверти участников (74%) чувствуют себя расстроенными или несчастными при работе с данными. Этот показатель свидетельствует: внедрение новых технологий и демократизация данных утомляет сотрудников. Корень проблемы именно в этой всеобщей усталости и нехватке навыков использования данных.

Производство данных

Как связаны производство данных и отсутствие умения с ними обращаться? Выше уже говорилось, сколько данных ежегодно производится в мире. Технологический прогресс ускоряется, количество данных растет, а в итоге организации и их сотрудники оказываются не готовы к обработке всего этого объема. Организации, которые возникли уже в цифровую эпоху, подготовлены к лавине данных несколько лучше, поскольку именно данные во многом определяют их рабочие процессы. Но что делать более старым компаниям? Их руководители отчаянно пытаются создать возможности для эффективного использования данных — и обнаруживают, что это не так-то просто. Речь не о том, что у них не выходит

получать и использовать данные. Они обнаруживают, что нельзя просто инвестировать в ПО и технологии, собрать данные... И вот ты уже царь горы — горы данных, доступных для анализа. В то же время, как показывают исследования, сотрудники просто не в состоянии достаточно быстро приспосабливаться к среде, которая становится все динамичнее. Отсюда и нехватка навыков работы с данными.

Отсутствие стратегии в сфере данных и их анализа

Еще одна причина нехватки навыков работы с данными — отсутствие у организаций соответствующей стратегии. Каков механизм этого? Для начала сядьте и подумайте: есть ли у вашей организации четкая и конкретная стратегия обращения с данными? Увы, многие из вас ответят, что нет. Отсутствие стратегии — это лишняя нагрузка на сотрудников, которые выбиваются из сил, пытаясь понять, как использовать ПО и технологии, в которые компания вложила деньги. (Помните исследование 2019 года? Люди действительно расстраиваются и устают!)

В чем важность для предприятия четкой и конкретной стратегии работы с данными и их анализом? Давайте вернемся к примеру с бегуном, которому также необходима определенная стратегия. Представьте себе начинающего (или не совсем начинающего, но в любом случае непрофессионального) спортсмена, который участвует в забеге.

Допустим, этот спортсмен — вы. Вы давно не бегали, но кто-то из ваших друзей, родных или коллег участвует в забегах. Вы видите, как это их воодушевляет, постоянно слушаете их разговоры (я сам бегун и прекрасно понимаю, что меня сложно заставить замолчать, если уж я оседлал любимого конька). И вы решаете тоже записаться на забег. До него остается несколько месяцев. Вы записались, но у вас

нет стратегии тренировок, вы не изучили трассу, вы не представляете, как правильно питаться и сколько нужно пить жидкости, — однако тренируетесь и едите, зная, что это необходимо. Кроме того, не имея стратегии, вы не знаете, какие вложения требуются для осуществления вашей мечты.

И вот наступил день забега. Вы стоите у стартовой линии, чувствуя себя более или менее подготовленным. У вас с собой, возможно, запас пищи и воды, и вы надеетесь, что тренировки помогут вам дойти до финиша. Но на самом деле вы не готовились должным образом — хотя, конечно, кое-как тренировались и даже купили спортивную форму. Если вы бежите полумарафон, то, возможно, справитесь с забегом, пусть и с трудом. Но и в полном, и в ультрамарафоне вам, скорее всего, придется сойти с дистанции.

А теперь представьте другой подход. Вы записываетесь на забег уже после того, как выработали стратегию, которая поможет вам добиться результата. Вы точно знаете, какое снаряжение вам необходимо, сколько питательных веществ и жидкости нужно вашему организму. Вы обратились к хорошему тренеру, который проследит за процессом вашей подготовки, давая необходимые указания. Итак, тренировки завершены, и вы выходите на старт. Теперь вы дойдете до финиша на любой дистанции — через боль, через усталость, но дойдете. Вы знали, что нужно делать, и достигли успеха именно благодаря стратегии и правильной подготовке.

В целом эти примеры наглядно показывают, что нужно организации от стратегии в сфере данных и их анализа. Слишком долго руководители полагались на более бессистемный подход. Они знали, что нужно инвестировать в данные и аналитику, но делали недостаточно, не разобравшись, какое снаряжение и зачем они приобретают (и подходит ли оно им вообще). Они не обращались к «тренерам», которые могли бы помочь им разработать и воплотить стратегию.

А потом обнаруживалось, что инвестиции — порой миллионы долларов — не дают ожидаемых результатов. К несчастью, таких руководителей и компаний очень много.

Организация должна построить стратегию получения и использования данных, чтобы расширять видение, ставить цели и решать задачи. На сегодня большинству организаций во всем мире не хватает такой стратегии.

При отсутствии стратегии навыки обращения с данными у сотрудников только ухудшаются. Вместо того чтобы вырабатывать четкую стратегию, определяющую и обеспечивающую поток инвестиций в ПО, технологии и обучение персонала, компании просто покупают ПО и технологии, рассчитывая, что все как-то само собой заработает. Налицо подмена понятий: руководители думают, что технологии — это и есть стратегия, и навязывают их сотрудникам. Но в таком случае стратегия не определяет, какую технологию использовать. В результате сотрудники отвергают новшества и возвращаются к старым способам решения задач. А новая технология как ненужная игрушка, которая лежит на полке и пылится.

Это приводит к двойной проблеме. Во-первых, ПО, купленное компанией, не внедряется или не используется сколько-нибудь эффективно. Во-вторых, вместо совершенствования навыков работы с данными сотрудники компании лишь еще больше отстают, потому что не желают осваивать программы и технологии, которые были приобретены, казалось бы, для их удобства.

Данные: что дальше?

Итак, мы выяснили, что при нехватке у сотрудников соответствующих навыков данные не используются достаточно эффективно. И что же делать? Действительно ли необходимо бороться с этим недостатком и ликвидировать пробелы —

или все-таки можно продолжать работать как раньше? Ответ очевиден: ликвидировать пробелы необходимо!

Как уже было сказано, к 2025 году будет производиться примерно 463 эксабайта данных ежедневно: это наше будущее. Давайте еще нагляднее, без примеров с DVD: один эксабайт — это единица с 18 нулями. Так что представьте себе число 463 и припишите к нему 18 нулей. По другим прогнозам, к 2025 году количество производимых ежедневно данных будет равняться 175 секстибайтам, а секстибайт — это единица с 21 нулем: 1 секстибайт равен триллиону гигабайт¹⁴. Так какой же прогноз верен? Или лучше задать другой вопрос: а так ли это важно? Ведь это просто *очень много данных*, и нет никаких сомнений, что где-то среди них скрывается много ценных знаний. И мы снова возвращаемся к той же проблеме: если нам так не хватает навыков в обращении с данными, смогут ли отдельные люди и организации воспользоваться этим огромным объемом данных с выгодой для себя? Или большинству придется бессильно наблюдать, как организации, умеющие использовать данные, легко обходят конкурентов?

Краткое содержание главы

В целом можно сказать, что мир данных, в котором мы живем, — удивительный, пугающий и неизведанный. В будущем появятся новые профессии, возможности и изобретения, и мы пока не можем сказать, что повлекут за собой все эти новшества. Наверняка возникнет необходимость в новых навыках, но мы не знаем, каких именно. Одно ясно уже сейчас: данные были и будут всегда! Сегодня налицо тенденция быстрого и масштабного накопления данных, которые необходимо обрабатывать. Есть и другая тенденция: все очевиднее нехватка навыков у сотрудников

организаций. Это мешает организациям добиваться успеха и эффективно использовать инвестиции в данные и их анализ. Что же мы можем сделать? Есть ли решение этой проблемы? Конечно же, да! Перед людьми и организациями, которые смогут освоить мир дата-грамотности, откроются огромные возможности.

Четыре уровня аналитических методов

Данные и их анализ —
целых четыре уровня?

Теперь, когда мы получили представление о мире данных, нам нужно целостное понимание методов анализа данных. Только в этом случае люди и организации смогут эффективно использовать данные и извлекать выгоду из аналитики. Понимание — ключ к реализации работающей стратегии в области работы с данными. Если понимания нет, руководители могут закупать сколь угодно дорогое и мощное ПО, получать данные и демократизировать инструменты работы с ними, но им будет непонятно, действительно ли их организации нужно именно это. Основа мира аналитики — четыре уровня: **дескриптивный** (описательный), **диагностический**, **предиктивный** (предсказательный) и **прескриптивный** (предписывающий). Чтобы осознать необходимость изучения этих уровней, давайте заглянем в прошлое: как организации исторически воспринимали данные и методы их анализа? Эта картина поможет нам понять, как относиться к ним теперь и как можно использовать эти четыре уровня для достижения успеха.

Итак, организации, данные и дата-аналитика: исторический аспект. Организации всегда вкладывали много средств в технологии, программное обеспечение и инструменты работы

с данными. ПО и технологии считались манной небесной, все были уверены, что с их помощью можно найти любые решения в области данных и аналитики. Руководители инвестировали в технологии, которые могли помочь им добиться целей и воплотить в жизнь самые смелые мечты из области данных и дата-аналитики (а такие мечты у руководителей есть всегда). Со времен появления самой первой версии Microsoft Excel в 1985 году¹⁵ (кстати, если вы не в курсе, то сначала эту программу выпустили для Apple Macintosh — странно, правда?) и даже более раннего ПО для электронных таблиц организации постоянно покупали либо продавали ПО, считая, что оно способно все решить, творя некую «магию». Итак, давным-давно существует ПО для хранения, использования и анализа данных и информации. По мере развития и совершенствования технологий также росли и инвестиции в них. По прогнозам, в 2019 году доходы, полученные во всем мире от анализа больших объемов данных и бизнес-информации, должны были составить 187 миллиардов долларов¹⁶. Этот рынок продолжает развиваться столь же бурно... однако существует и неприятная тенденция. Учитывая нехватку знаний у 24% людей, ответственных за принятие бизнес-решений, и у 32% топ-менеджеров, инвестиции в технологии сбора и анализа данных не приносят ожидаемых результатов. Эта нехватка знаний, о которой мы говорили в первой главе, мешает инвестициям приносить свои плоды.

Итак, каким же образом понимание четырех уровней аналитических методов влияет на грамотное внедрение технологий и возврат инвестиций в данные и аналитику? Когда организации и сотрудники разберутся, как работают все четыре уровня, это позволит понять, как коллектив и отдельные сотрудники со своими индивидуальными навыками при посредстве технических возможностей могут совместно построить правильную стратегию работы с данными и дата-аналитикой. А затем эту стратегию можно будет применить.

Четыре уровня аналитических методов

А сейчас, после краткого экскурса в историю, давайте попробуем разобраться в самих четырех уровнях аналитики. Затем мы покажем вам, как сотрудники на самых разных уровнях, от рядовых работников до топ-менеджеров, могут применять эти четыре уровня: 1) для продвижения стратегии работы с данными и аналитикой; 2) для принятия более разумных решений на основе данных; 3) для формирования правильного видения данных и аналитики. Понимание четырех уровней аналитики позволяет организации не стрелять по мишени вслепую, а осуществлять реальную работу по формированию стратегии.

Как уже говорилось выше, четыре уровня аналитических методов — это дескриптивный, диагностический, предиктивный и прескриптивный (см. рис. 2.1). Чтобы получить прочные базовые представления о каждом уровне, давайте для начала рассмотрим их определения и примеры. Кроме



Рис. 2.1. Четыре уровня аналитической мозаики

того, мы также познакомимся с конкретными программами и технологиями, соответствующими каждому уровню. Затем, разобравшись с каждым уровнем, мы увидим, как они работают в комплексе: формируют правильную аналитическую картину и помогают организации успешно освоить инвестиции в данные.

Прежде чем перейти к первому уровню, давайте разберемся, что на самом деле в нашем контексте означает слово «аналитика». Мы слышим его постоянно, но известно ли нам его истинное значение?

Если взглянуть на определение нужного нам значения слова «аналитика», то окажется, что аналитика — это «систематический вычислительный анализ данных или статистики»¹⁷.

И что же это такое? Взглянем на определение слова «анализ»: это подробное изучение элементов или структуры чего-либо. В мире данных и аналитики анализ — это способность глубоко «закапываться» в данные, *понимая*, что они нам говорят. Данные — это не всегда цифры, они могут представлять собой слова, символы и прочие элементы. Если мы понимаем, что говорят нам данные, то затем можем рассмотреть, каковы их элементы и структурные особенности, а также сделать с ними многое другое. Аналитика позволяет нам принимать более удачные решения, задавать более правильные вопросы и эффективно использовать имеющиеся у нас данные.

Четыре уровня аналитических методов — дескриптивный, диагностический, предиктивный и прескриптивный — помогают организациям по-настоящему разбираться в данных и информации, которые они получают, накапливают и используют для принятия бизнес-решений, совершенствования рабочих процессов и достижения успеха. Аналитические методы критически необходимы для успеха организации в современном цифровом мире, управляемом данными.

Уровень 1: дескриптивные (описательные) аналитические методы

Первый уровень анализа — это уровень дескриптивных (или описательных) методов.

Согласно одному из определений, «описание» означает «словесное изображение чего-либо или кого-либо, включающее все важные характеристики, качества или свойства».

Ну и что все это значит? В данном случае описательный метод — это метод, описывающий то, что имело или имеет место. Иными словами, дескриптивный анализ позволяет нам оглянуться на то, что уже произошло в бизнесе, и изучить это при помощи данных.

Однако это не всегда дает нам наиболее четкую картину. Чтобы ее уточнить, существуют другие методы — например, отчеты, работа со сводками или информационными панелями, наблюдения. Наверное, все это знакомо каждому из вас. Как часто на совещаниях или в рабочей электронной переписке упоминается слово «отчет»? Как часто мы видим сводки, KPI (ключевые показатели эффективности) и т.д.? Все эти понятия стали для нас такими обыденными, что сами слетают с языка, но они действительно необходимы для полного понимания описательного анализа. Описательный анализ строится на отчетах, сводках и наблюдениях, которые помогают выяснить, что происходило с организацией до этого момента или происходит прямо сейчас.

Научившись правильному применению методов описательного анализа, мы сможем понять и изучить роль, которую он играет в этой мозаике — в стратегии работы с данными. Но поможет ли это собрать всю мозаику? Здесь нужно учесть один ключевой момент: описательный анализ часто представляет для организаций серьезную проблему. Казалось бы,