Моей прекрасной жене и пятерым чудесным детям.
Спасибо за поддержку в моем путешествии
к дата-грамотности.

Содержание

Об авторе 11

Предисловие 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
Ступень 1. Спросить 190
Ступень 2. Получить 193
Ступень 3. Проанализировать 196
Ступень 4. Интегрировать 198
Ступень 5. Решить 204
Ступень 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

Примечания 251

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.

01

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 лет[1], Земля — примерно 4,5 миллиарда[2], 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-х[3] 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-м[5]. 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 занимательных 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 лимонад[6]. 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 компании[7]. Один из примеров использования Rolls-Royce возможностей данных — прогностический метод мониторинга техобслуживания двигателей[8]. 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 лечении[9]. 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 году[11]:

- 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 данными[12]. 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

Конец ознакомительного фрагмента.
Приобрести книгу можно
в интернет-магазине
«Электронный универс»
e-Univers.ru