

Отзывы рецензентов о книге

В основе методологии бережливого производства лежат научные методы: разработка гипотез, проведение экспериментов, сбор данных, получение информации и подтверждение или модификация гипотез. А/В-тестирование – это золотой стандарт разработки поддающихся проверке и повторяемых экспериментов, и эта книга подводит окончательный итог.

– Стив Бланк, адъюнкт-профессор Стэнфордского университета, отец современного предпринимательства, автор книг *The Startup Owner’s Manual* и *The Four Steps to the Epiphany*

Эта книга – отличный ресурс для руководителей, лидеров, исследователей или инженеров, желающих использовать контролируемые онлайн-эксперименты для повышения качества продукта, эффективности проекта или величины дохода. Я лично знаю, какое влияние оказала работа Кохави на Bing и Microsoft, и рад, что эти знания теперь доступны более широкой аудитории.

– Гарри Шам, исполнительный вице-президент, Группа искусственного интеллекта и исследований Microsoft

Замечательная книга, одновременно строгая и доступная. Читатели узнают, как проводить в своих организациях доверительные контролируемые эксперименты, которые произвели революцию в разработке интернет-продуктов.

– Адам Д’Анджело, соучредитель и генеральный директор Quora и бывший технический директор Facebook

Эта книга – отличный рассказ о том, как несколько компаний используют онлайн-эксперименты и А/В-тестирование для улучшения своих продуктов. Кохави, Тан и Сюй обладают богатым опытом и навыками, поэтому в книге есть множество практических примеров из реального мира и уроков, извлеченных за многие годы применения этих методов в больших масштабах.

– Джефф Дин, старший научный сотрудник Google и старший вице-президент Google Research

Хотите ли вы, чтобы ваша организация постоянно принимала лучшие решения? Это новая библия для желающих перейти от данных к решениям

в эпоху цифровых технологий. Читать эту книгу – все равно что присутствовать на собраниях в Amazon, Google, LinkedIn, Microsoft. Авторы впервые раскрывают способы принятия решений самыми успешными компаниями мира. Вместо болтовни и анекдотов из обычных книг про бизнес эта книга показывает, что надо делать и как делать это хорошо. Это руководство по принятию решений в цифровом мире с отдельными разделами для руководителей предприятий, инженеров и аналитиков данных.

– Скотт Кук, соучредитель и председатель исполнительного комитета Intuit

Управляемые онлайн-эксперименты – это мощные инструменты. Понимание того, как они работают, в чем их сильные стороны и как их можно оптимизировать, может послужить источником вдохновения как специалистам, так и более широкой аудитории. Эта книга представляет собой редкое сочетание авторитетных технических инструкций и увлекательного чтения и посвящена очень важным вопросам.

– Джон П.А. Иоаннидис, профессор медицины, здравоохранения, биомедицинских данных и статистики Стэнфордского университета

Какая онлайн-версия будет лучше? Нам часто приходится делать такой выбор, и мы часто ошибаемся. Чтобы определить, что на самом деле будет работать лучше, нам нужны строгие контролируемые эксперименты, также известные как А/В-тестирование. Эта прекрасная и яркая книга экспертов из Microsoft, Google и LinkedIn делится с нами теорией и передовым опытом А/В-тестирования. Ее обязательно должны прочесть все, кто хоть чем-то занят в интернете!

– Грегори Пятецки-Шапиро, доктор философии, президент KDnuggets, соучредитель SIGKDD и LinkedIn Top Voice в области науки о данных и аналитики

Рон Кохави, Диана Тан и Я Суй – ведущие мировые эксперты по онлайн-экспериментам. Я опираюсь на их опыт в течение многих лет и рад, что теперь они объединились для написания итогового руководства. Я рекомендую эту книгу всем своим студентам и тем, чья работа связана с онлайн-продуктами и услугами.

– Эрик Бриньолфссон, профессор Массачусетского технологического института и соавтор книги *The Second Machine Age*

Современный бизнес, основанный на программном обеспечении, не может оставаться конкурентоспособным без контролируемых онлайн-экспериментов. Написанная тремя наиболее опытными лидерами в этой области, эта книга представляет основные принципы, иллюстрирует их

убедительными примерами и дает глубокие практические советы. Настоятельно рекомендую к прочтению!

– Фостер Провост, профессор Школы бизнеса Стерна Нью-Йоркского университета и соавтор бестселлера
Data Science for Business

За последние два десятилетия технологическая индустрия наконец достигла то, что ученые знали на протяжении веков: контролируемые эксперименты являются одним из лучших инструментов для понимания сложных явлений и решения очень сложных проблем. Способность разрабатывать контролируемые эксперименты, проводить их в большом масштабе и интерпретировать результаты лежит в основе работы современных высокотехнологичных предприятий. Авторы разработали и реализовали несколько самых мощных в мире платформ для экспериментов. Эта книга – прекрасная возможность узнать на их опыте, как использовать эти инструменты и методы.

– Кевин Скотт, исполнительный вице-президент и технический директор *Microsoft*

Онлайн-эксперименты способствовали успеху Amazon, Microsoft, LinkedIn и других ведущих цифровых компаний. Эта практическая книга открывает читателю доступ к многолетнему опыту экспериментов в этих компаниях и должна быть на книжной полке каждого специалиста по данным, разработчика программного обеспечения и продакт-менеджера.

– Штефан Томке, профессор, лауреат гранта Уильяма Баркляя Хардинга, Гарвардская школа бизнеса, автор книги *Experimentation Works: The Surprising Power of Business Experiments*

Секрет успеха онлайн-бизнеса – это эксперименты. Но теперь это уже не секрет. В этой книге три мастера описывают ключевые приемы A/B-тестирования, чтобы вы тоже могли постоянно улучшать качество своих онлайн-продуктов.

– Хэл Вариан, главный экономист Google и автор книги *Intermediate Microeconomics: A Modern Approach*

Эксперименты – лучший инструмент для онлайн-продуктов и услуг. Эта книга полна практических знаний, полученных в результате многолетнего успешного тестирования в Microsoft, Google и LinkedIn. Понимание и передовой опыт объяснены реальными примерами и ловушками, выявлены их признаки и найдены решения. Я настоятельно рекомендую эту книгу!

– Престон Макафи, бывший главный экономист и вице-президент *Microsoft*

Экспериментирование – это будущее цифровой стратегии, а эта книга станет ее библией. Кохави, Тан и Сьюй – трое из наиболее заслуживающих внимания современных экспертов по экспериментированию, и их книга представляет собой практическое руководство по проведению экспериментов, которое полезно с первых строк. Тематические исследования по разоблачению экспериментальных заблуждений, которые они проводили на протяжении многих десятилетий в Microsoft, Amazon, Google и LinkedIn, трансформированы в простые для понимания практические решения с огромной глубиной и ясностью. Это обязательное чтение для любого менеджера цифрового бизнеса.

– Синан Арал, профессор менеджмента Дэвида Остина,
Массачусетский технологический институт,
и автор книги *The Hype Machine*

Незаменимая для любого серьезного экспериментатора, эта книга очень практична и необычайно глубока. Она настолько полезна, что кажется, будто ты обретаешь суперсилу. В этой книге вы найдете все – от тонкостей статистики до оценки результатов и измерения долгосрочного воздействия. Рекомендую к прочтению.

– Пип Лаха, эксперт по достижению конверсии,
основатель и директор CXL

Онлайн-эксперименты сыграли решающую роль в изменении корпоративной культуры Microsoft. Когда Сатья Наделла говорит об «установке на рост», экспериментирование – лучший способ попробовать новые идеи и извлечь из них уроки. После того как мы научились быстро повторять контролируемые эксперименты, Bing стал прибыльным и экспериментирование быстро распространилось по Microsoft через Office, Windows и Azure.

– Эрик Бойд, корпоративный вице-президент по
платформе AI, Microsoft

Как предприниматель, ученый и руководитель, я понял (на собственном горьком опыте), что один грамм данных ценнее, чем килограммы интуиции. Но как получить хорошие данные? Эта книга объединяет многолетний опыт работы в Amazon, Google, LinkedIn и Microsoft в доступное, хорошо организованное руководство. Это библия онлайн-экспериментов.

– Орен Эциони, генеральный директор Института
искусственного интеллекта Аллена и профессор
информатики Вашингтонского университета

Интернет-компании развили эксперименты до беспрецедентного масштаба, скорости и сложности. Авторы книги сыграли ключевую роль

в этих разработках, и читателям повезло, что они могут извлечь уроки из их совокупного опыта.

– Дин Эклс, профессор развития карьеры KDD в области коммуникаций и технологий Массачусетского технологического института и бывший ученый в Facebook

Замечательный богатый ресурс для изучения критически важной, но недооцененной области. Реальные тематические исследования в каждой главе иллюстрируют внутреннюю работу и уроки успешного бизнеса. Концентрация внимания на разработке и оптимизации критерия общей оценки (ОЕС) является особенно важным уроком.

– Джереми Ховард, Университет сингулярности, основатель fast.ai, бывший президент и главный научный сотрудник Kaggle

Я видел много пособий по A/B-тестированию, но лишь некоторые из них посвящены доверительным контролируемым онлайн-экспериментам. Я слежу за Ронни Кохави в течение 18 лет и вижу, что его советы отточены практикой и закалены работой в реальных условиях. Когда к нему присоединяются Диана Тан и Я Сюй, широта охвата не имеет себе равных. Я призываю вас сравнить эту книгу с любой другой – конечно, контролируемым образом.

– Джим Стерн, основатель саммита по маркетинговой аналитике и почетный директор Ассоциации цифровой аналитики

Чрезвычайно полезное руководство по проведению онлайн-экспериментов, сочетающее в себе аналитическую сложность, ясное изложение и добытые напряженным трудом уроки практического опыта.

– Джим Манзи, основатель Foundry.ai, основатель и бывший генеральный директор и председатель совета директоров Applied Predictive Technologies, автор книги *Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics, and Society*

Экспериментальный подход совершенствуется каждый раз, когда его применяют в новой области: сельское хозяйство, химия, медицина, а теперь уже и электронная торговля в интернете. Эта книга трех ведущих экспертов изобилует практическими советами и примерами, объясняющими, как и почему нужно экспериментировать в интернете, чтобы вас не обманула собственная интуиция. Эксперименты могут стоить дорого; однако работа вслепую может обойтись еще дороже.

– Арт Оуэн, профессор статистики Стэнфордского университета

Это обязательная книга для руководителей предприятий и операционных менеджеров. Подобно тому, как операции, финансы, бухгалтерский учет и стратегия представляют собой обязательные строительные блоки бизнеса, сегодня, в эпоху искусственного интеллекта, проведение обдуманных контролируемых онлайн-экспериментов становится обязательным навыком. Кохави, Тан и Сюй изложили основы знаний для этой новой и важной прикладной области.

– Карим Р. Лахани, профессор и директор лаборатории инновационных наук Гарварда, член совета директоров Mozilla Corp.

Серьезные организации, основанные на данных, понимают, что аналитики недостаточно; они должны сосредоточиться на экспериментах. Эта удивительно доступная и очень интересная книга представляет собой манифест и руководство по организации эффективных экспериментов. Я считаю, что это вдохновляющий прагматизм. Самое главное, книга поясняет, как сочетание культуры с технической компетенцией становится критическим фактором успеха.

– Майкл Шраге, научный сотрудник Инициативы Массачусетского технологического института по цифровой экономике и автор книги *The Innovator's Hypothesis: How Cheap Experiments Are Worth More than Good Ideas*

Эта важная книга об экспериментировании вобрала в себя мудрость трех выдающихся лидеров крупнейших мировых технологических компаний. Если вы инженер-программист, специалист по данным или менеджер по продукту и пытаетесь внедрить в своей организации культуру бизнеса, основанную на данных, это книга станет для вас отличным практическим руководством.

– Даниэль Тункеланг, главный научный сотрудник Endeca и бывший директор по науке о данных и инжинирингу в LinkedIn

По мере того как каждая отрасль все больше становится цифровой и управляемой данными, проведение контролируемых онлайн-экспериментов и получение выгоды от них становится необходимым навыком. Кохави, Тан и Сюй предоставляют полное и хорошо проработанное руководство, которое станет обязательным чтением как для практиков, так и для руководителей.

– Евангелос Симудис, соучредитель и управляющий директор Synapse Partners; автор книги *The Big Data Opportunity in Our Driverless Future*

Авторы представляют результат более 10 лет упорного экспериментирования в самой стратегически важной книге для данной дисциплины.

– Колин МакФарланд, директор экспериментальной платформы Netflix

Практическое руководство по А/В-тестированию превращает опыт трех ведущих умов экспериментальной практики в простые и легко усваиваемые блоки ценных прикладных идей. Каждая глава знакомит вас с некоторыми из наиболее важных аспектов при проведении экспериментов – от выбора правильной метрики до пользы институциональной памяти. Если вы ищете наставника по экспериментам, который сочетает в себе науку и практичность, эта книга для вас.

– Дилан Льюис, руководитель службы экспериментов, Intuit

Единственное, что хуже, чем отсутствие эксперимента, – это ошибочное толкование эксперимента, потому что оно дает вам ложную уверенность! В этой книге подробно описаны технические аспекты тестирования, основанные на опыте проведения экспериментов в некоторых крупнейших компаниях мира. Если вы участвуете в онлайн-экспериментах в каком угодно качестве, скорее прочтите эту книгу, чтобы избежать ошибок и получить уверенность в своих результатах.

– Крис Говард, автор книги *You Should Test That!*, основатель и генеральный директор Widerfunnel

Это феноменальная книга. Авторы опираются на богатый опыт и создали удобочитаемое справочное пособие, которое в некотором роде является всеобъемлющим и исчерпывающим. Настоятельно рекомендую прочитать эту книгу всем, кто хочет проводить серьезные цифровые эксперименты.

– Пит Кумен, соучредитель, Optimizely

Авторы – пионеры онлайн-экспериментов. Созданные ими платформы и проводимые ими эксперименты изменили некоторые из крупнейших интернет-брендов. Их исследования и беседы вдохновили другие команды в отрасли на экспериментирование. Эта книга – авторитетный, но практичный труд, которого так ждала отрасль.

– Адиль Айджаз, соучредитель и генеральный директор Split Software

Оглавление

| | |
|--|----|
| Отзывы рецензентов о книге | 5 |
| Предисловие от издательства | 18 |
| Вступление | 19 |
| Предисловие | 21 |
| Благодарности | 23 |
| ЧАСТЬ I. ВВЕДЕНИЕ ДЛЯ ВСЕХ | 25 |
| Глава 1. Введение и мотивация | 27 |
| 1.1. Терминология контролируемых онлайн-экспериментов..... | 29 |
| 1.2. Зачем нужны эксперименты? Корреляции, причинно-следственная связь и доверительность..... | 33 |
| 1.3. Необходимые ингредиенты для проведения эффективных контролируемых экспериментов..... | 35 |
| 1.4. Постулаты | 36 |
| 1.5. Постепенные улучшения | 39 |
| 1.6. Примеры интересных контролируемых онлайн-экспериментов | 41 |
| 1.7. Стратегия, тактика и их связь с экспериментами..... | 46 |
| 1.8. Дополнительное чтение | 50 |
| Глава 2. Проведение и анализ экспериментов. Пример полного цикла | 52 |
| 2.1. Условия демонстрационного эксперимента | 52 |
| 2.2. Проверка гипотез: установление статистической значимости..... | 56 |
| 2.3. Разработка эксперимента | 58 |
| 2.4. Проведение эксперимента и сбор данных | 61 |
| 2.5. Интерпретация результатов..... | 61 |
| 2.6. От результатов к решениям | 63 |
| Глава 3. Закон Тваймана и надежность экспериментов | 66 |
| 3.1. Неправильная интерпретация статистических результатов..... | 67 |
| 3.1.1. Нехватка статистической мощности | 67 |
| 3.1.2. Неправильная интерпретация p -значений..... | 67 |
| 3.1.3. Отслеживание p -значений | 69 |
| 3.1.4. Множественные проверки гипотез..... | 69 |
| 3.2. Доверительные интервалы..... | 70 |

| | |
|--|------------|
| 3.3. Угрозы внутренней достоверности..... | 70 |
| 3.3.1. Нарушения правила SUTVA | 70 |
| 3.2.2. Ошибка выжившего | 71 |
| 3.2.3. Вынужденное воздействие..... | 71 |
| 3.2.4. Несоответствие коэффициента выборки | 72 |
| 3.4. Угрозы внешней достоверности | 76 |
| 3.4.1. Эффекты первичности..... | 76 |
| 3.4.2. Эффекты новизны..... | 76 |
| 3.4.3. Выявление эффектов первичности и новизны..... | 78 |
| 3.5. Разделение по сегментам | 78 |
| 3.5.1. Сегментированное представление показателя | 79 |
| 3.5.2. Сегментированное представление эффекта (гетерогенность эффекта)..... | 80 |
| 3.5.3. Анализ эффекта по сегментам, вводящий в заблуждение | 81 |
| 3.6. Парадокс Симпсона | 82 |
| 3.7. Поощряйте здоровый скептицизм..... | 84 |
| Глава 4. Платформы и культура экспериментов | 85 |
| 4.1. Модели зрелости экспериментов..... | 85 |
| 4.1.1. Лидерство | 87 |
| 4.1.2. Процесс | 88 |
| 4.1.3. Разработать самим или купить готовый продукт? | 91 |
| 4.2. Инфраструктура и инструменты..... | 94 |
| 4.2.1. Разработка, настройка и управление экспериментом | 96 |
| 4.2.2. Развертывание эксперимента..... | 97 |
| 4.2.3. Инструменты для экспериментов..... | 100 |
| 4.2.4. Масштабирование экспериментов: тонкости назначения вариантов | 101 |
| 4.2.5. Параллельные эксперименты | 103 |
| 4.2.6. Анализ экспериментов | 105 |
| ЧАСТЬ II. ИЗБРАННЫЕ ТЕМЫ ДЛЯ ВСЕХ | 107 |
| Глава 5. Скорость имеет значение!..... | 111 |
| 5.1. Ключевое предположение: локальная линейная аппроксимация | 113 |
| 5.2. Как измерить быстродействие веб-сайта..... | 114 |
| 5.3. Схема эксперимента по замедлению | 116 |
| 5.4. Влияние различных элементов страницы | 118 |
| 5.5. Экстремальные результаты | 119 |
| Глава 6. Организационные показатели | 121 |
| 6.1. Таксономия показателей | 121 |
| 6.2. Выработка показателей: принципы и методы | 125 |
| 6.3. Оценка показателей..... | 128 |

| | |
|---|-----|
| 6.4. Развивающиеся показатели | 129 |
| 6.5. Дополнительное чтение | 130 |
| 6.6. Примечание: ограничительные показатели | 130 |
| 6.7. Примечание: преднамеренная манипуляция показателями..... | 132 |

Глава 7. Показатели экспериментов и общий критерий оценки..... 135

| | |
|---|-----|
| 7.1. От бизнес-показателей к показателям, подходящим для экспериментов..... | 136 |
| 7.2. Объединение ключевых показателей в ОЕС | 138 |
| 7.3. Пример: ОЕС для электронной почты на Amazon | 140 |
| 7.4. Пример: ОЕС для поисковой системы Bing. | 141 |
| 7.5. Закон Гудхарта, закон Кэмпбелла и замечание Лукаса | 143 |

Глава 8. Институциональная память и метаанализ..... 145

| | |
|--|-----|
| 8.1. Что такое институциональная память?..... | 145 |
| 8.2. Почему полезна институциональная память?..... | 146 |

Глава 9. Этика контролируемых экспериментов 150

| | |
|--|-----|
| 9.1. Что лежит в основе этики | 150 |
| 9.1.1. Риски | 152 |
| 9.1.2. Преимущества и выгоды | 153 |
| 9.1.3. Возможность выбора | 155 |
| 9.2. Сбор данных | 155 |
| 9.3. Культура и процессы | 156 |
| 9.4. Примечание: идентификация пользователей | 157 |

ЧАСТЬ III. ДОПОЛНИТЕЛЬНЫЕ И АЛЬТЕРНАТИВНЫЕ МЕТОДЫ КОНТРОЛИРУЕМЫХ ЭКСПЕРИМЕНТОВ..... 159

Глава 10. Дополнительные методы..... 163

| | |
|---|-----|
| 10.1. Пространство дополнительных методов..... | 163 |
| 10.2. Анализ на основе журналов | 164 |
| 10.3. Экспертная оценка..... | 166 |
| 10.4. Исследование пользовательского опыта..... | 167 |
| 10.5. Фокус-группы | 168 |
| 10.6. Обзоры | 169 |
| 10.7. Внешние данные..... | 170 |
| 10.8. Подведем итог главы..... | 172 |

Глава 11. Наблюдательные исследования причинно-следственных связей 174

| | |
|--|-----|
| 11.1. Когда контролируемые эксперименты невозможны | 174 |
|--|-----|

| | |
|--|-----|
| 11.2. Планы для наблюдательных исследований причинно-следственных связей | 176 |
| 11.2.1. Прерывистый временной ряд | 176 |
| 11.2.2. Эксперименты с чередованием | 178 |
| 11.2.3. Метод разрывной регрессии | 178 |
| 11.2.4. Инструментальные переменные и естественные эксперименты..... | 180 |
| 11.2.5. Отбор подобного по склонности..... | 180 |
| 11.2.6. Дифференциальная разница..... | 181 |
| 11.3. Ловушки причинно-следственных связей | 182 |
| 11.4. Приложение: опровергнутые исследования причинно-следственных связей | 185 |

ЧАСТЬ IV. ПЛАТФОРМЫ ДЛЯ ЭКСПЕРИМЕНТОВ: УГЛУБЛЕННОЕ ИЗУЧЕНИЕ..... 189

| | |
|---|------------|
| Глава 12. Эксперименты на стороне клиента..... | 193 |
| 12.1. Различия между серверной и клиентской стороной..... | 193 |
| 12.1.1. Отличие №1: процесс выпуска..... | 194 |
| 12.1.2. Отличие №2: обмен данными между клиентом и сервером | 195 |
| 12.2. Следствия из компромиссов | 197 |
| 12.3. Выводы..... | 201 |

| | |
|--|------------|
| Глава 13. Инструментарий экспериментов..... | 202 |
| 13.1. Инструменты на стороне клиента и сервера | 202 |
| 13.2. Обработка журналов из нескольких источников..... | 204 |
| 13.3. Культура измерений..... | 205 |

| | |
|---|------------|
| Глава 14. Выбор единицы рандомизации..... | 206 |
| 14.1. Единица рандомизации и единица анализа..... | 208 |
| 14.1 Рандомизация на уровне пользователя | 209 |

| | |
|--|------------|
| Глава 15. Развитие эксперимента: компромисс между скоростью, качеством и риском. | 212 |
| 15.1. Что такое рампинг?..... | 212 |
| 15.2. Шаблон SQR для рампинга | 213 |
| 15.3. Четыре фазы рампинга..... | 214 |
| 15.3.1. Первая фаза рампинга: до MPR..... | 215 |
| 15.3.2. Вторая фаза рампинга: MPR..... | 216 |
| 15.3.3. Третья фаза рампинга: пост-MPR | 216 |
| 15.3.4. Четвертая фаза рампинга: длительное удержание или репликация..... | 216 |
| 15.4. Что после рампинга?..... | 218 |

| | |
|---|-----|
| Глава 16. Анализ масштабных экспериментов | 219 |
| 16.1. Подготовка данных | 219 |
| 16.2. Вычисление данных | 220 |
| 16.3. Формирование сводки и визуализация результатов..... | 222 |

ЧАСТЬ V. РАЗВЕРНУТОЕ ОПИСАНИЕ АНАЛИЗА ЭКСПЕРИМЕНТОВ

225

| | |
|---|-----|
| Глава 17. Статистика контролируемых онлайн-экспериментов | 229 |
| 17.1. Двухвыборочный t -тест | 229 |
| 17.2. p -значение и доверительный интервал | 230 |
| 17.3. Предположение о нормальности..... | 231 |
| 17.4. Ошибки типа I/II и статистическая мощность | 233 |
| 17.5. Смещение..... | 235 |
| 17.6. Множественное тестирование..... | 235 |
| 17.7. Метаанализ Фишера | 236 |

| | |
|---|-----|
| Глава 18. Оценка дисперсии и повышение чувствительности: подводные камни и решения | 238 |
| 18.1. Распространенные ошибки | 239 |
| 18.1.1. Дельта или процентная дельта? | 239 |
| 18.1.2. Показатели отношения: когда уровень анализа отличается от уровня эксперимента | 239 |
| 18.1.3. Выбросы..... | 241 |
| 18.2. Повышение чувствительности..... | 242 |
| 18.3. Дисперсия других статистических данных | 244 |

| | |
|--|-----|
| Глава 19. А/А-тестирование | 246 |
| 19.1. Почему нужны А/А-тесты? | 246 |
| 19.1.1. Пример 1: уровень анализа отличается от уровня рандомизации | 247 |
| 19.1.2. Пример 2: поощрение остановки эксперимента при достижении статистической значимости | 249 |
| 19.1.3. Пример 3: переадресация браузера | 249 |
| 19.1.4. Пример 4: неравное распределение по группам | 250 |
| 19.1.5. Пример 5: различия в оборудовании..... | 251 |
| 19.2. Как проводить А/А тесты | 251 |
| 19.3. Когда А/А-тест не подходит | 252 |

| | |
|--|-----|
| Глава 20. Включение по условию для повышения чувствительности | 254 |
| 20.1. Примеры включения по условию..... | 254 |

| | |
|--|-----|
| 20.1.1. Пример 1: преднамеренно частичное воздействие | 255 |
| 20.1.2. Пример 2: условное воздействие | 255 |
| 20.1.3. Пример 3: Увеличение охвата | 256 |
| 20.1.4. Пример 4: изменение покрытия | 256 |
| 20.1.5. Пример 5: контрфактическое включение для моделей машинного обучения | 257 |
| 20.2. Числовой пример | 258 |
| 20.3. Оптимальное и консервативное включение | 258 |
| 20.4. Общий эффект воздействия | 259 |
| 20.5. Достоверность включения | 261 |
| 20.6. Распространенные ошибки | 261 |
| 20.7. Открытые вопросы | 263 |

Глава 21. Несоответствие коэффициента выборки и другие ограничительные показатели

| | |
|---|-----|
| 21.1. Несоответствие коэффициента выборки (SRM) | 264 |
| 21.2. Причины возникновения SRM | 266 |
| 21.3. Устранение SRM | 268 |
| 21.4. Другие ограничительные показатели, связанные с доверием | 269 |

Глава 22. Утечка и интерференция между вариантами

| | |
|--|-----|
| 22.1. Примеры | 272 |
| 22.2. Некоторые практические решения | 275 |
| 22.2.1. Полезное правило: ценность действия в экосистеме | 276 |
| 22.2.2. Изоляция | 277 |
| 22.2.3. Анализ на уровне ребер графа | 279 |
| 22.2.4. Обнаружение и мониторинг взаимовлияния | 280 |

Глава 23. Измерение долгосрочных эффектов

| | |
|---|-----|
| 23.1. Что такое долгосрочные эффекты? | 281 |
| 23.2. Причины, по которым могут различаться краткосрочные и долгосрочные эффекты | 282 |
| 23.4. Зачем измерять долгосрочные эффекты? | 284 |
| 23.5. Длительные эксперименты | 285 |
| 23.6. Альтернативные методы для длительных экспериментов | 288 |
| 23.6.1. Метод №1: когортный анализ | 288 |
| 23.6.2. Метод № 2: постпериодный анализ | 288 |
| 23.6.3. Метод №3: воздействие с интервалом во времени | 290 |
| 23.6.4. Метод №4: сдерживание и обратный эксперимент | 292 |

Предметный указатель

293

Предисловие от издательства

Отзывы и пожелания

Мы всегда рады отзывам наших читателей. Расскажите нам, что вы думаете об этой книге, – что понравилось или, может быть, не понравилось. Отзывы важны для нас, чтобы выпускать книги, которые будут для вас максимально полезны.

Вы можете написать отзыв на нашем сайте www.dmkpress.com, зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу dmkpress@gmail.com; при этом укажите название книги в теме письма.

Если вы являетесь экспертом в какой-либо области и заинтересованы в написании новой книги, заполните форму на нашем сайте по адресу http://dmkpress.com/authors/publish_book/ или напишите в издательство по адресу dmkpress@gmail.com.

Список опечаток

Хотя мы приняли все возможные меры для того, чтобы обеспечить высокое качество наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг – возможно, ошибку в основном тексте или программном коде, – мы будем очень благодарны, если вы сообщите нам о ней. Сделав это, вы избавите других читателей от недопонимания и можете нам улучшить последующие издания этой книги.

Если вы найдете какие-либо ошибки в коде, пожалуйста, сообщите о них главному редактору по адресу dmkpress@gmail.com, и мы исправим это в следующих тиражах.

Нарушение авторских прав

Пиратство в интернете по-прежнему остается насущной проблемой. Издательство «ДМК Пресс» очень серьезно относится к вопросам защиты авторских прав и лицензирования. Если вы столкнетесь в интернете с незаконной публикацией какой-либо из наших книг, пожалуйста, пришлите нам ссылку на интернет-ресурс, чтобы мы могли применить санкции.

Ссылку на подозрительные материалы можно прислать по адресу dmkpress@gmail.com.

Мы высоко ценим любую помощь по защите наших авторов, благодаря которой мы можем предоставлять вам качественные материалы.

Вступление

Получить числа легко; получить числа, которым можно *доверять*, сложно. Это практическое руководство от ведущих специалистов по экспериментам в Google, LinkedIn и Microsoft научит вас, как ускорить внедрение инноваций, используя *доверительные контролируемые онлайн-эксперименты* (trustworthy online controlled experiments), или, как их чаще называют, *A/B-тесты*. Основываясь на практическом опыте компаний, каждая из которых проводит более 20 000 контролируемых экспериментов в год, авторы делятся примерами, советами и предостережениями со студентами и профессионалами отрасли, приступающими к экспериментам, а также детально раскрывают сложные темы для опытных практиков, которые хотят усовершенствовать процесс принятия решений на основе экспериментальных данных в своей организации.

Прочитав эту книгу, вы научитесь:

- использовать научные методы для оценки гипотез с помощью контролируемых экспериментов;
- определять ключевые показатели и в идеале общий критерий оценки;
- проверять достоверность результатов и предупреждать экспериментаторов о нарушении предположений;
- выполнять быструю интерпретацию и итерации на основе полученных результатов;
- устанавливать ограничения для защиты ключевых бизнес-целей;
- создавать масштабируемые платформы, снижающие предельную стоимость экспериментов почти до нуля;
- избегать ловушек, таких как эффекты переноса, закон Тваймана, парадокс Симпсона и сетевые взаимодействия;
- решать на практике проблемы статистики, в том числе общие нарушения предположений.

Рон Кохави (Ron Kohavi) – вице-президент и технический сотрудник Airbnb. Он трудился над этой книгой, когда был членом технической команды и корпоративным вице-президентом в Microsoft. До этого он был директором по интеллектуальному анализу данных и персонализации в Amazon. Рон получил степень доктора информатики в Стэнфордском университете. Его статьи имеют более 40 000 цитирований, и три из них входят в 1000 наиболее цитируемых статей в области информатики.

Диана Тан (Diane Tang) – научный сотрудник Google, обладающий опытом в области крупномасштабного анализа данных и инфраструктуры, контролируемых онлайн-экспериментов и рекламных систем. Она получила

степень бакалавра гуманитарных наук в Гарварде и степень магистра/доктора в Стэнфорде, имеет патенты и публикации в области мобильных сетей, визуализации информации, методологии экспериментов, инфраструктуры данных, интеллектуального анализа данных и больших данных.

Я Сюй (Ya Xu) возглавляет отдел экспериментов и сбора данных в LinkedIn. Она опубликовала несколько работ по методике экспериментов и часто выступает на авторитетных конференциях и с лекциями в университетах. Ранее она работала в Microsoft и получила докторскую степень по статистике в Стэнфордском университете.

Предисловие

Если у нас есть данные, давайте рассмотрим их.

Если все, что у нас есть, – это мнения, делайте что я скажу.

– Джим Баркстейл,
бывший генеральный директор Netscape

Наша цель при написании этой книги – поделиться практическими уроками, извлеченными из многолетнего опыта проведения масштабных онлайн-экспериментов в Amazon и Microsoft (Рон Кохави), Google (Диана Тан), Microsoft и LinkedIn (Я Сюй). Хотя мы пишем эту книгу как частные лица, а не как представители Google, LinkedIn или Microsoft, мы опираемся на уроки и ошибки, с которыми столкнулись за годы нашей работы, и даем рекомендации как по программным платформам, так и по корпоративным аспектам использования контролируемых экспериментов для формирования культуры, основанной на данных, когда решения принимают исходя из достоверной информации, а не полагаясь на HiPPO (highest paid person's opinion, «мнение самой высокооплачиваемой персоны»). Мы считаем, что многие из этих уроков применимы в онлайн-среде, в больших или малых компаниях, или даже в командах и отделах внутри компании. Мы разделяем озабоченность по поводу необходимости оценки достоверности результатов экспериментов. Мы верим в скептицизм, выраженный законом Тваймана: любое значение, которое выглядит интересным или странным, обычно ошибочно; мы призываем читателей перепроверить результаты и провести тесты на достоверность, особенно если получены исключительно положительные результаты. Получить числа легко; получить числа, которым можно доверять, сложно!

Часть I предназначена для чтения всеми, независимо от уровня подготовки, и состоит из четырех глав:

- глава 1 представляет обзор преимуществ проведения контролируемых онлайн-экспериментов и вводит отраслевые термины;
- в главе 2 рассмотрен пример полного цикла организации и проведения эксперимента;
- в главе 3 описаны распространенные ошибки и способы повышения надежности экспериментов;
- в главе 4 рассказано, что нужно для создания экспериментальной платформы и масштабирования онлайн-экспериментов.

Части со II по V могут быть прочитаны по мере необходимости, но они написаны с упором на конкретную аудиторию. Часть II содержит пять глав,

посвященных основам, таким как показатели организации. Главы части II рекомендуются всем, особенно руководителям и менеджерам среднего звена. Часть III состоит из двух глав, в которых представлены методы, дополняющие контролируемые онлайн-эксперименты. Эти методы могут оказаться полезными для руководителей, специалистов по обработке данных, инженеров, аналитиков, менеджеров по продуктам и других специалистов, стремящихся эффективно расходовать время и усилия. Часть IV посвящена созданию платформы для экспериментов и предназначена для инженеров и разработчиков. Наконец, в части V рассматриваются вопросы расширенного анализа, и она предназначена для специалистов по данным.

Наш веб-сайт <https://experimentguide.com> дополняет эту книгу. Он содержит дополнительные материалы, исправления и предоставляет форум для открытого обсуждения. Все доходы от этой книги авторы намерены пожертвовать на благотворительность.

Благодарности

Мы хотим поблагодарить наших коллег, которые работали с нами на протяжении многих лет. Хотя их было слишком много, чтобы перечислить каждое имя, эта книга основана на нашей совместной работе, а также на других исследованиях в отрасли, помимо исследований и проведения контролируемых онлайн-экспериментов. Мы многому у вас научились, спасибо.

Во время работы над книгой мы не переставали благодарить нашего редактора Лорен Коулз за сотрудничество с нами на протяжении всего этого процесса. Чери Вудворд дала нам отличные советы по редактированию и стилю и помогла объединить наши три голоса. Стефани Грей работала с нами над всеми графиками и рисунками, улучшая их раз за разом. Ким Вернон выполнила окончательное редактирование текста и проверку библиографии.

Что наиболее важно, мы глубоко признательны нашим семьям за то, что они позволили нам пожертвовать временем, проведенным с ними, чтобы поработать над этой книгой. Спасибо семье Ронни: Яэль, Орен, Иттаи и Нога, семье Дианы: Бену, Эмме и Лии, а также семье Я Сюй: Томасу, Лере и Тавису. Мы не смогли бы написать эту книгу без вашей поддержки и энтузиазма!

Google: Хэл Вариан, Дэн Рассел, Кэрри Граймс, Найл Кардин, Дейрдри О'Брайен, Хеннинг Хонхольд, Мукунд Сундарараджан, Амир Наджми, Патрик Райли, Эрик Тассон, Джен Дженнаи, Шеннон Валлор, Эрик Миралья, Дэвид Прайс, Кристал Дален, Тэмми Джи Мюррей, Лана Доннелли и все, кто работает над экспериментами в Google, спасибо.

LinkedIn: Стивен Линч, Яв Божинов, Джаида Лю, Вейтао Дуан, Нанью Чен, Гийом Сен-Жак, Элейн Колл, Мин Лю, Арун Свами, Киран Прасад, Игорь Перишич и вся команда экспериментов, спасибо.

Microsoft: Омар Алонсо, Бенджамин Араи, Джордан Атлас, Рича Баяни, Эрик Бойд, Джонни Чан, Алекс Денг, Энди Дрейк, Александр Фабиджан, Брайан Фраска, Скотт Гуд, Сомит Гупта, Адам Густафсон, Томми Гай, Рэнди Хенн, Эдвард Джезерски, Цзин Джин, Дону Ким, Уолдо Кейперс, Джонатан Литц, София Лю, Цзяннан Лу, Ци Лу, Даниэль Миллер, Карл Митчелл, Нильс Полманн, Вен Цинь, Томас Шрайтер, Гарри Шум, Дэн Соммерфилд, Гранат Ваз, Тоби Уокер, Мишель Зункер, а также команда анализа и экспериментов, спасибо.

Особая благодарность Марии Стоун и Маркусу Перссон за отзывы о книге, а также Мишель Н. Мейер за экспертный отзыв по главе, посвященной этике.

Среди тех, кто дал свой отзыв: Адиль Айджаз, Джонас Алвес, Алон Амит, Кевин Андерсон, Джоэл Барахас, Хуман Бедаят, Бо Бендер, Бахадор Биглари, Стюарт Бак, Джике Чонг, Джек Чоу, Павел Дмитриев, Юронг Фан, Георгий Георгиев, Илиас Геростатопулос. Мэтт Гершофф, Уильям Гроссо, Адитья Гупта, Раджеш Гупта, Шилпа Гупта, Крис Джек, Джейкоб Ярнвалл, Дэйв Кароу, Славек Кирнер, Пит Кумен, Дилан Льюис, Брайан Лю, Дэвид Манхейм, Колин МакФарланд, Танаполиндраунчрон, Адитья Рамдас, Андре Рихтер, Цзяньхонг Шен, Ганг Су, Энтони Тан, Лукас Вермеер, Роуэл Виллемс, Ю Ян и Юфэн Ван.

Спасибо всем, кто помогал нам, но здесь не упомянут.

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

e-Univers.ru