

Spis treści

Przedmowa	15
Podziękowania	17

CZĘŚĆ I. Zagadnienia wstępne dla wszystkich

Rozdział 1. Wstęp i motywacja	21
Żargon kontrolowanych eksperymentów online	23
Po co eksperymentować? Korelacja, przyczynowość i wiarygodność	26
Niezbędne elementy przeprowadzania użytecznych eksperymentów kontrolowanych	28
Zasady	29
Poprawa w czasie	32
Przykłady ciekawych kontrolowanych eksperymentów online	33
Strategia, taktyka i ich związek z eksperymentami	39
Dodatkowe źródła	43
Rozdział 2. Przeprowadzanie i analizowanie eksperymentów — kompleksowy przykład	45
Kontekst przykładu	45
Testowanie hipotezy — Ustalenie istotności statystycznej	48
Projektowanie eksperymentu	51
Przeprowadzanie eksperymentu i pozyskanie danych	53
Interpretowanie wyników	54
Od wyników do decyzji	55
Rozdział 3. Prawo Twymana i wiarygodność eksperymentów	59
Błędna interpretacja wyników statystycznych	60
Przedziały ufności	63
Zagrożenia dla wewnętrznej trafności	63
Problemy z zewnętrzną trafnością	69

Różnice pomiędzy segmentami	72
Paradoks Simpsona	75
Wspieraj zdrowy sceptycyzm	77
Rozdział 4. Platforma i kultura eksperymentowania	79
Modele dojrzałości badawczej	79
Infrastruktura i narzędzia	88
 CZĘŚĆ II. Wybrane zagadnienia dla wszystkich	
Rozdział 5. Szybkość ma znaczenie. Całościowe studium przypadku	103
Kluczowe założenie: lokalna aproksymacja liniowa	105
Jak mierzyć wydajność strony?	106
Projekt eksperymentu spowalniającego	108
Różne elementy strony mają różny wpływ	110
Ekstremalne wyniki	111
Rozdział 6. Wskaźniki organizacyjne	113
Taksonomia wskaźników	113
Formułowanie wskaźników — zasady i techniki	117
Ocena wskaźników	119
Ewolucja wskaźników	120
Dodatkowe źródła	121
Rozdział 7. Wskaźniki dla eksperymentów i ogólne kryterium ewaluacji	126
Od wskaźników biznesowych	
do wskaźników właściwych dla eksperymentów	126
Łączenie kluczowych wskaźników w OKE	128
Przykład: OKE dla wiadomości e-mail w Amazonie	131
Przykład: OKE dla silnika wyszukiwarki Bing	132
Prawo Goodharta, prawo Campbella i krytyka Lucasa	134
Rozdział 8. Pamięć instytucjonalna i metaanaliza	135
Czym jest pamięć instytucjonalna?	135
Dlaczego pamięć instytucjonalna jest pożyteczna?	136
Rozdział 9. Etyka w eksperymentach kontrolowanych	140
Tło	140
Zbieranie danych	145
Kultura i procesy	146

CZĘŚĆ III. Techniki uzupełniające i alternatywne w stosunku do eksperymentów kontrolowanych

Rozdział 10. Techniki uzupełniające	151
Przestrzeń technik uzupełniających	151
Analiza oparta na logach	152
Ludzka ocena	154
Badania wrażeń użytkownika	155
Grupy fokusowe	156
Ankiety	156
Zewnętrzne dane	158
Wszystko razem	159
Rozdział 11. Obserwacyjne badania przyczynowe	161
Gdy kontrolowane eksperymenty są niemożliwe	161
Projekty w przyczynowych badaniach obserwacyjnych	163
Pułapki	169

CZĘŚĆ IV. Zaawansowane zagadnienia związane z budowaniem platformy do eksperymentowania

Rozdział 12. Eksperymenty po stronie klienta	177
Różnice między serwerem a klientem	177
Implikacje w zakresie eksperymentów	180
Podsumowanie	185
Rozdział 13. Instrumentacja	186
Instrumentacja po stronie klienta a instrumentacja po stronie serwera	186
Przetwarzanie logów z różnych źródeł	188
Kultura instrumentacji	189
Rozdział 14. Wybór jednostki randomizacji	190
Jednostka randomizacji i jednostka analizy	192
Randomizacja na poziomie użytkownika	193
Rozdział 15. Zwiększanie ekspozycji eksperymentu — kompromis między szybkością, jakością i ryzykiem	196
Czym jest faza rozbiegowa?	196
Model rozbiegowy SQR	197

Cztery etapy rozbiegu	198
Faza porozbiegowa	202
Rozdział 16. Skalowanie analizy eksperymentów	203
Przetwarzanie danych	203
Obliczenia na danych	204
Podsumowanie wyników i wizualizacje	206
 CZĘŚĆ V. Zaawansowane zagadnienia dotyczące analizy eksperymentów	
Rozdział 17. Statystyka kontrolowanych eksperymentów online	211
Test t Studenta	211
Wartość p i przedział ufności	212
Założenie normalności rozkładu	213
Błędy pierwszego i drugiego rodzaju oraz moc	215
Tendencyjność	216
Wielokrotne testowanie	217
Metaanaliza Fishera	217
Rozdział 18. Szacowanie wariancji i poprawa czułości	
— problemy i rozwiązania	219
Częste problemy	220
Zwiększanie czułości	223
Wariancja innych statystyk	225
Rozdział 19. Testy A/A	226
Po co robić testy A/A?	226
Jak przeprowadzać testy A/A?	231
Gdy test A/A daje negatywny wynik	233
Rozdział 20. Objęcie eksperymentem i poprawa czułości	234
Przykłady objęcia eksperymentem	234
Przykład liczbowy (Kohavi, Longbotham i inni, 2009)	238
Optymalne i konserwatywne objęcie eksperymentem	238
Ogólny efekt eksperymentalny	239
Wiarygodne objęcie eksperymentem	241
Częste problemy	241
Otwarte kwestie	243

Rozdział 21. Błąd niedopasowania proporcji próby	
oraz inne wskaźniki ochronne	244
Błąd niedopasowania proporcji próby	244
Analizowanie problemów niedopasowania proporcji próby	248
Rozdział 22. Wyciek informacji i zakłócenia między wariantami	251
Przykłady	252
Propozycje praktycznych rozwiązań	256
Wykrywanie i monitorowanie zakłóceń	260
Rozdział 23. Mierzenie długofalowego efektu eksperymentalnego	261
Czym są efekty długofalowe?	261
Powody, dla których efekt eksperymentalny może się różnić krótko- i długofalowo	262
Powody, dla których efekt eksperymentalny może się różnić	262
Po co mierzyć efekt długofalowy?	264
Długotrwałe eksperymenty	265
Alternatywne metody długoterminowych eksperymentów	268
Bibliografia	273
Skorowidz	293